D. Jude Hemanth · *Editor*

Artificial Intelligence Techniques for Satellite Image Analysis

EARSeL

Springer

Remote Sensing and Digital Image Processing

Volume 24

Series Editor
Freek D. van der Meer, Faculty of Geo-Information Science and Earth Observation (ITC), Department of Earth Systems Analysis, University of Twente, Enschede, The Netherlands
Anna Jarocińska, Department of Geoinformatics, Cartography and Remote Sensing, University of Warsaw, Warsaw, Poland

Remote Sensing and Digital Image Processing book series. Remote sensing is the acquisition of Physical data of an object without touch or contact. Earth observation satellites have been used for many decades in a wide field of applications. With the advancements in sensor technology, earth imaging is now possible at an unprecedented level of detail. Imaging spectrometers and thermal multispectral systems acquire detailed spectroscopic information of physical properties of the earth's surface. Dynamic processes can now be studied with interferometric systems. SAR interferometry, laser altimetry and high-resolution imaging allow a very detailed, three-dimensional reconstruction of the earth surface. With the advent of multi-sensor mission comes a new era of imaging, opening the possibility of integrating date from various sensor systems. The books published in the series explore these topics in remote sensing and provide a framework for related advanced digital image processing approaches.

Accepted for inclusion in Scopus.

More information about this series at http://www.springer.com/series/6477

D. Jude Hemanth
Editor

Artificial Intelligence Techniques for Satellite Image Analysis

Springer

Editor
D. Jude Hemanth
Electronics and Communication Engineering
Karunya University
Coimbatore, Tamil Nadu, India

Preface

The necessity for the application of artificial intelligence (AI) techniques in the field of remote sensing is exponentially increasing in today's scenario. However, the literatures covering these two broad areas simultaneously are relatively scarce. This is one of the significant motivations behind the origin of this book. This book is highly interdisciplinary which covers readership from researchers/scientists in the field of AI, remote sensing, and image processing. Different practical applications are covered in this book which will create an interest among the budding engineers in these areas. On the other hand, in-depth analysis also has been carried out to attract the experts to further explore in these areas of research. This book, indeed, is a wholesome product which will cater the needs of the wide variety of academicians, scientists, and researchers. A brief introduction about each chapter is as follows.

Chapter 1 covers the application of augmented reality (AR) for satellite image analysis. The main objective of this work is to display the satellite image in a better way which will enhance the success of the subsequent image analysis methods. Chapter 2 deals with an intelligent method for clustering the satellite images into different groups. This method helps to identify the different objects in an efficient way. Chapter 3 reports about the application of deep learning approach for identifying the different vegetation regions in satellite images. This application is very significant in the context of enhancing the crop production in the agriculture field.

Chapter 4 illustrates the applications of artificial neural networks for detection/tracking of ships via satellite images. These methods are highly useful for controlling the maritime traffic apart from the surveillance application. Chapter 5 covers the intelligent image enhancement techniques which are one of the focal areas of research in remote sensing. Artificial Bee Colony (ABC) approach is used for enhancing the quality of the images. The analyses of synthetic aperture radar (SAR) images are carried out in Chap. 6. Noise removal is performed in this work with the help of mathematical transform-based approaches.

Chapter 7 deals with the application of fuzzy logic concepts for SAR image analysis. Fuzzy logic methods are used to detect the nutrition level of crops and subsequent planning for improving the crop production. Natural images and features

are detected from satellite images in Chap. 8. Convolutional neural networks are used in this chapter for object detection application. Chapter 9 covers the concepts of change detection of tropical mangrove ecosystem using mathematical approaches. This method is highly useful to detect the level of deforestation in a specific region.

The application of different AI-based classifiers for measuring the quality of vegetation land is explored in Chap. 10. The merits and demerits of different AI-based classifiers are also reported in this chapter. A survey on the various machine learning (ML) approaches for satellite image analysis is carried out in Chap. 11. The different applications of satellite image analysis are also dealt in this chapter. Water body extraction from satellite images is the emphasis of Chap. 12. Wavelet transform is used in this research work to carry out the extraction process.

We are grateful to the authors and reviewers for their excellent contributions for making this book possible.

Our special thanks go to Anna Jaroscinska, Van Der Meer, and Freek D (Series Editors of Remote Sensing and Digital Image Processing) for the opportunity to organize this edited volume.

We are grateful to Springer, especially to Petra van Steenbergen (Executive Editor), for the excellent collaboration.

This edited book covers the fundamental concepts and application areas in detail which is one of the main advantages of this book. Being an interdisciplinary book, we hope it will be useful to a wide variety of readers and will provide useful information to professors, researchers, and students.

Coimbatore, Tamil Nadu, India   D. Jude Hemanth
January, 2019

Contents

Chapter 1
Heightening Satellite Image Display via Mobile Augmented Reality – A Cutting-Edge Planning Model

Sagaya Aurelia

Abstract This paper summarizes on object detection, classification, analysis, and display for optical satellite image. Initially, all the existing object detection and object viewing system based on AI techniques are introduced. Various optical imaginary methods and the possibility of immersing optical and 3D data with other data sources are also explained. The surveyed literatures show that in most of the case, the detected objects are taken as resource for planning. We also observed that the image viewing and displaying model was ignored by many authors which is one of the key concepts for next phase. Satellite AR plays a vital role in displaying the images. Overall, it can be seen that optical image view along with AR display can be used for better planning, which is one of the popular research topics and has an excessive operational potential which is the need of the hour dealing with analyzing, predicting, and viewing large amount of data.

Keywords MAR · Object detection · Satellite AR · Image analysis · AI · Satellite view · Image display · Optical satellite data · Object recognition · Tracking · Rendering · 3D data

1.1 Introduction

Artificial intelligence is a field of academic study with its own vocabulary and specialist terms. It is a process where a computer solves a task in a way that mimics human behavior. Today, narrow AR—when a machine is trained to do one particular task—is becoming more widely used from virtual assistance to self-driving cars to automatic tagging your friends in your photos on Facebook. It makes to think like human, act humanly, and think and act rationally.

S. Aurelia (✉)
Christ University, Bengaluru, Karnataka, India
e-mail: sagaya.aurelia@christuniversity.in

Machine learning is an algorithm that allows computers to learn from examples without being explicitly programmed. Deep learning is a subset of ML which uses deep artificial neural networks as models and does not require feature engineering. Artificial intelligence involves various cybernetics such as expert systems, computational intelligence, distributed artificial intelligence, and intelligence interface. The evolution of fuzzy systems, neural networks, and evolutionary computation are some of the major parts of computational intelligence. Services and tools in AI are shown in Fig. 1.1. Intelligent agents are the major impact of distributed artificial intelligence (DAI). Back propagation, case-based reasoning, Bayesian nets, support vector machine, boosting, and genetic programming are the major subcomponents of artificial intelligence. It can also be extended in addition along with computer vision and natural language processing.

Cognitive science makes the system act as humans, whereas rational agents make the system work rationally. When the system thinks rationally, law of thought has become applicable, and finally when the system thinks like humans, an agent perceives its environment through sensors and acts on the environment through actuators. Human sensors are eyes and ears, and actuator effectors are hands, legs,
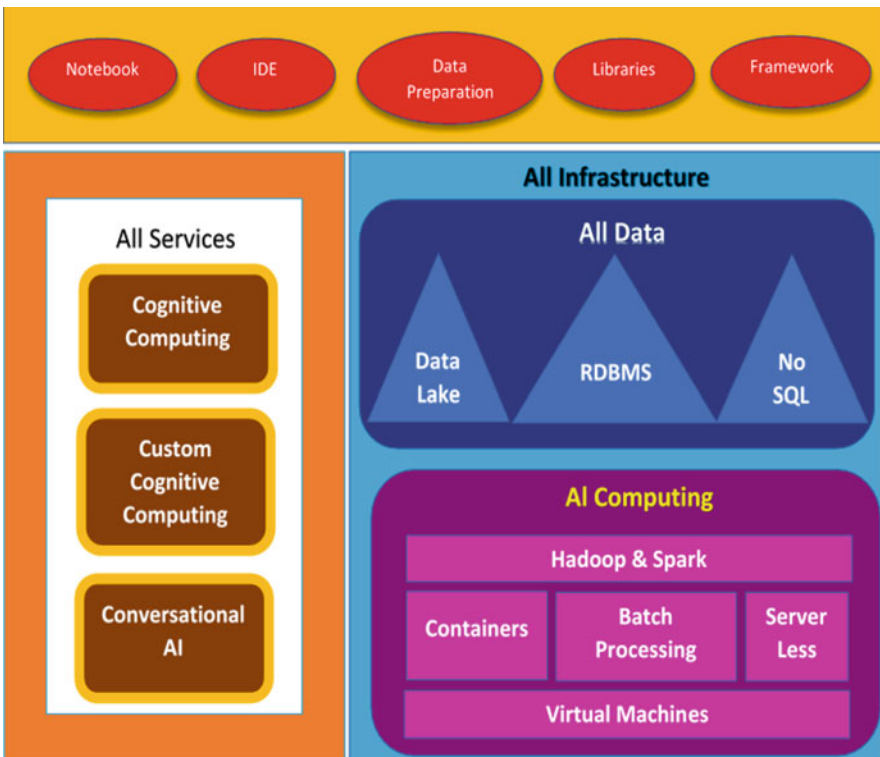
Fig. 1.1 Services and tools in AI
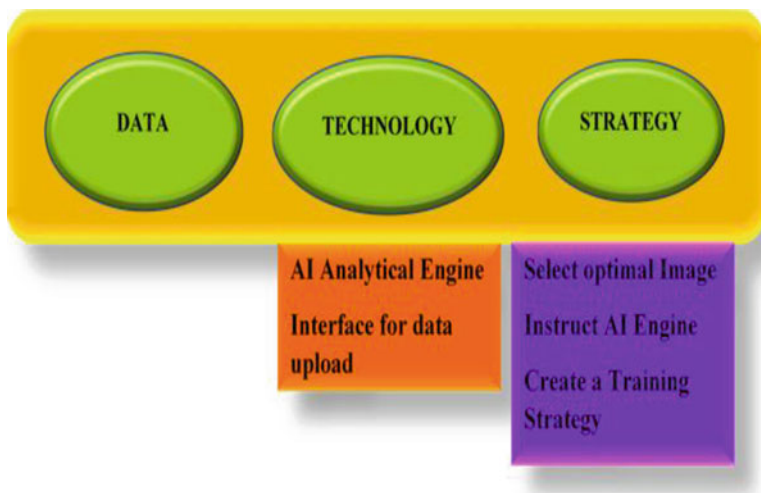
Fig. 1.2 Convergences of AI

and mouth. Sensors are cameras, sonar, lasers, ladder, and bump, and effectors are grippers, manipulators, and motors.

The artificial intelligence is shown in Fig. 1.2. The input layer receives input from Twitter, Facebook, and other social media site from which the static counting layer receives input and the knowledge representation layer gets the feedback from the models and decision generated from machine learning layer. The same process happens in analytics but without the existence of knowledge representation.

1.1.1 History of AI

The birth of AI took place in 1950s based on the assertion every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it.

The Mid-1950s and 1970s are called as golden years. Many new insights and successful programs took birth. AI faced financial setback and lot of critiques. The optimism raised, and expectations and impossibility level were high. As a result of failure in promised result, material funding for AI vanished. Expert system gave rebirth to AI during late 1980s. AI became an industry, and neural networks returned to popularity. Japan came forwarded and did an aggressive funding with projects for fifth-generation computers. The business community's eye fell on AI in between 1987 and 1993. Intelligent agents emerged during 1995. Evolutionary computation has a great impact even today. Some of the notable AI-based applications were IBM's Deep Blue which beats the world champion Garry Kasparov at chess. An

Fig. 1.3 History of artificial intelligence

automated system was introduced by Amazon in 2002 which replaced human editors. In 2011, Apple released Siri, a personal voice agent. Leading Go player was defeated by Lee Sedol, by Google's AlphaGo in 2016. Currently, AI-based deep learning is the trend. The ages of AI can be divided into three such as age of handcrafted knowledge, age of satellite learning, and age of contextual adaption. The history of AI is shown in Fig. 1.3.

1.2 Stages and Processes of Artificial Intelligence

In machine learning, data from various sources are gathered and are cleansed to have homogeneity. Based on selected right machine learning algorithm, model is built. The results give insight, and the data are visualized. Initially, data are designed by experiments and thenceforth leading to formulation of business problem in which the data are prepared. Analyzed data are explored and visualized. After visualizing, the data are classified and clustered which leads to predictive analysis, and finally, AI becomes achievable. AI attains its maturity after crossing seven different stages.

Perception gives an idea of what is happening now, and notification deals with information needed to know. Suggestion comprises of the user recommendation, and the procedures to be done is informed in automation. The expected task is described in prediction, and unwanted data are avoided in prevention, and finally, the current required is given in situational awareness.

1.3   Practical Application and Future of AI

Virtual personal assistant, climate, finance, agriculture, education, logistics and transportation, commercial, and health are some of the application of artificial intelligences as shown in Fig. 1.4. Virtual personal assistants provide assistance such as chatbots who can suggest products, restaurants, hotels, and services. Smart technologies play an importance in market prediction and customer's advice on transactions. In educational artificial intelligence, AI is used to create personalized offerings to optimize learning. Sales forecasting and recommending right product to the client are some of the major applications in commercial sector. Combat deforestation and in submarine vehicle to detect leaks in oil pipelines are the major area where AI plays an important role in reducing the energy consumption. In case of improving agricultural yields and warning of any adverse impact to the environment,
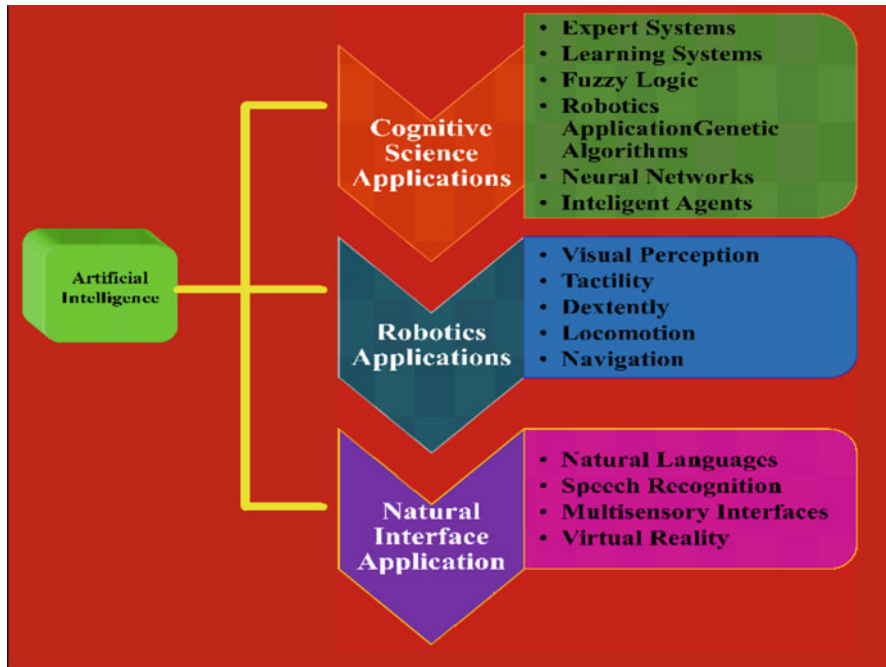
Fig. 1.4 Applications of artificial intelligence
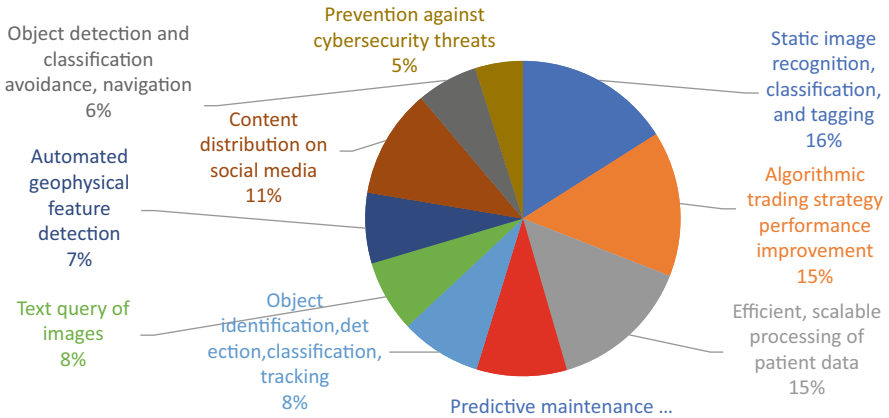
Fig. 1.5 The future of AI

AI is the most suitable format. Logistics and transportations are considered one of the major areas of AI application as far as avoiding collisions or traffic jams is considered and in optimizing traffic flow. Data collection generates patterns that help with identifying genetic factors which are susceptible to developing a disease.

When artificial intelligence is used to perform a single task better than a human, then it is termed as narrow AI, whereas when an assembly of narrow AI systems are used to achieve greater capabilities, then they fall under the group of narrow AI cluster. When more than one task can be performed, it falls under the group of enhanced narrow AI. The growth of AI and its milestones is remarkable.

The future of AI as shown in Fig. 1.5 is based on various areas in specific on static image recognition, classification, and tagging which will top with the maximum revenue, and therefore followed by algorithmic trading strategy performance improvement, efficient, scalable processing of patient data and finally object detection and classification, avoidance navigation, and prevention against cybersecurity threats will be the least revenue resource. Year-wise revenue as per Statista states that by 2025, AI market will expand worldwide with maximum limit.

Algorithmic trading strategy performance improvement will be the maximum revenue-generating resource and with contract analysis being the less source of revenue. Figure 1.6 shows the graphical representation of artificial intelligence revenue in world market.

1.4 Augmented Reality

Augmented reality (AR) is a new area of research which can be defined as augmenting existing reality. AR is a technology that homogenizes reality with the computer graphics. The superimposing of AR objects over real objects is viewed in
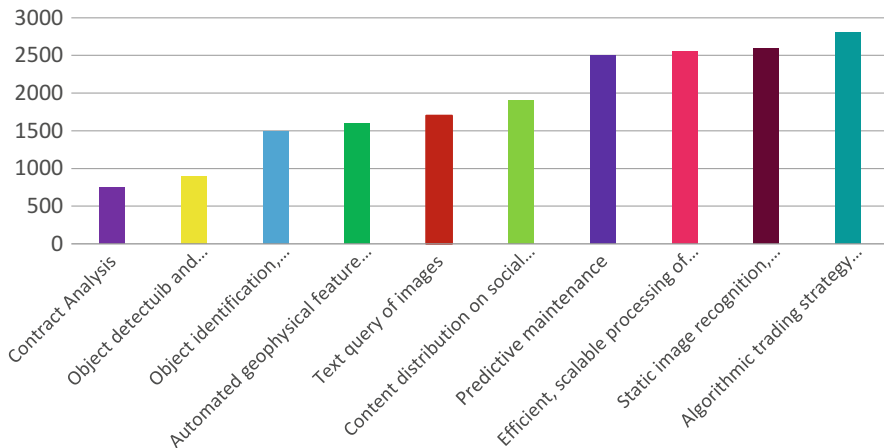
Fig. 1.6 Artificial intelligence revenue in world market
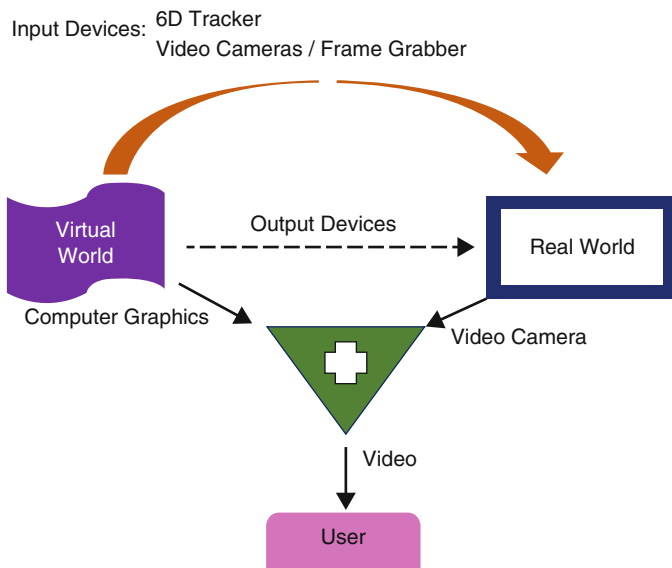
Fig. 1.7 Augmented reality contour

an AR-enabled user interface. AR is a field of computer science that concatenates device-generated digital data with the real-world environment as shown in Fig. 1.7. AR technologies enables and enriches the user's perspicacity about real world on the same way it makes the information about reality more interactive. Real-time object detection and recognition algorithms are used to recognize the surrounding objects in the real environment. These algorithms are used to position the superimposed content with real-time objects in the AR view.

AR is a combination of different techniques. The high level of immersion in AR is incredible. The main feature of augmented reality is real-time constructiveness. The main prosperity for AR users is the context of their surrounding environment is maintained; in addition to that, the preeminence of sensory information and input are also enjoyed. Augmented reality takes many forms in which visual and audio overlays on the world are ubiquitous.

1.4.1 History of Augmented Reality

The user views generated images and animations on dedicated display devices in computer graphics. Some of them are not just video projectors and conventional screens but also movies and television which convert digital images to classical media. In modernistic years, a lot of advanced display technologies for computer graphics are developed.

These advanced technologies provide large perceived display area and stereoscopic rendering. To name some, head-mounted displays [25], table-like display setups [26], domed projection system, and rooms made up of back-projected screens (CAVE) [27] are some of them. Virtual reality research is the genesis for most of these systems. They have a wide range of higher-end applications such as industrial design, entertainment settings, and scientific and medical visualization. Unfortunately, these devices focus only on computer-generated images, not providing any information about the real surroundings. There are multitudinous reasons to specify where, along with the artificial graphics, even the real environment is also necessary. This is where the actual ambiance supposed to be merged with the virtual information in an interactive display. The desideratum of such amalgamation of real environs and virtual image has led to the reinforcement of augmented reality (AR). Augmented reality is not utterly an unprecedented concept. The history of early computer graphics research substantiates it.

During 1968, Ivan Sutherland (named the "father of computer graphics") and his coadjutors developed a mechanically tracked 3D see-through head-worn display. Computer-effectuated information mongrel with physical object were able to be viewed. In between 1970 and 1980, it became a research topic in some of the major institutions, US Air Force Armstrong Laboratory being one of them. "Super Cockpit" project was developed by Tom Furness. A high-resolution heads-up overlay display for fighter pilots, supported by 3D sound, was achieved through this project. In 1989, the idiom virtual reality was first coined by Jaron Lanier, and the rudimentary commercial livelihood around virtual worlds was created. In 1990, it is believed that the term "augmented reality" was attributed by Tom Caudell, a former Boeing researcher. In 1993, Loomis and his colleagues from University of California and Santa Barbara developed a GPS-based outdoor system which dispensed navigational assistance for visually impaired with spatial audio overlays. Julie Martin in 1994 created "Dancing in Cyberspace," the first augmented reality theater production which constituents virtual objects on the physical stage and

acrobats who danced within. In 2000, Hirokazu Kato created ARToolkit, an open-source software library. It uses video tracking to overlay the artificial images on a video camera. It is used even today to accommodate AR view.

Later in 2009, ARToolkit was extended to web browser. Esquire magazine introduced AR. It asked the readers to scan the cover page and get the live view of Robert Downey. In 2013, Volkswagen MARTA app provided virtual step-by-step repair assistance, allowing service technicians to foresee how a repair process will look on the vehicle in front of them. Google introduced Google Glass in 2014; thus, a trend of wearable AR came into effect. As per 2016, virtual reality (VR) and augmented reality (AR) have detonated for $1.1 billion in investment, and it's still weighed the fountainhead of the forthcoming of computing. In the coming days, we can play a real-time strategy game on our computer or mobile phones and put on AR glasses. It is very much evitable that augmented reality is going to be next-generation ruling technology. Paul Milgram [27] introduced the brain wave of virtuality continuum, a continuous scale gallivanting between entire virtuality and entire reality. This virtuality continuum as shown in Fig. 1.15 girdles the difference and combination of virtual elements and real ambiance. The middle area between both the extremes is known as augmented reality where the reality and virtuality are merged.

AR varies from virtual reality (VR) in the output view. In VR, the user experiences are totally a computer-generated environment. No information about the real, original ambiance is given. Whereas in AR the environment is original, only an elongated incorporeal information is provided as virtual object. AR is a viaduct used to fill the gap between the virtual and real work in impeccable way [28]. To regiment how the computer-generated AR element appeared on the screen, we urge to use a personal computer or a handheld mobile device equipped with a camera. By showing a pre-recognized graphical code (called as marker) or graphics before the camera, a preassigned AR object, in form of 3D graphics, a video, or a flash movie, will be shown on the screen, replacing the marker/graphics. At present, both hardware and software components are required for an AR system to provide an augmented experience.

1.5   Substantive AR Hardware Components

There is a rapid refinement in the capabilities of AR hardware. A greater progress is seen in terms of less expensive and faster hardware used in AR application. The three basic hardware components used for all AR systems include:

• Processors
• Sensors
• Displays

1.5.1   Processors

A number of roles are consummated using processors which are one of the essential components of augmented reality system. Being the brain of the technological system, it procures the signal input from the sensors, executes the instructions from the application program, and then concocts a signal that activates the display system. Central processing unit (CPU) and special purpose graphics processing units (GPUs) are used in an AR system. GPUs are special three-dimensional optimized graphics computation processors used for good effects. Some of the preeminent specifications of either CPU or GPUs are:

- Processor speed
- Available memory
- Number of processor
- Available storage
- Network latency
- Graphic accelerator
- Network bandwidth

1.5.2   Sensors

Sensors play a vital role in getting the real-world information and passing that to the AR application. Other than providing the information regarding the pose of the user in the real world, some additional information such as lightness/darkness, temperature, and pH are also given. The primitive function of tracking is performed with the help of the sensors.

1.5.3   Displays

It is a device which responses to our sense-perceived signals. Signals are received from devices and sent to our eyes, our ears, our sense of touch, and our nose, and also a sense of taste. There are also other sensations provided by designed stimuli such as vestibular system [23, 24]. Many types of displays are available. To name, some are:

- Visual display

  - Stationary visual display
  - Visual display that moves with the participant's head
  - Visual display that moves with the participant's hands and other parts

- Audio display

  – Stationary audio display
  – Audio display that moves with the participant's head
  – Headphones
  – Earbuds
  – Audio display that moves with the participant's hands and other

- Haptic display

  – Skin sensations (traction)
  – Forces (kinesthetics)
  – Stereo display
  – Other sensory displays
  – Olfaction (smell)senses

1.5.4   Substantive AR Software Components

It is the software that makes the hardware to do exactly what the user wants to do. AR software makes the hardware capable of making our idea, in our own area of interest [24, 29]. Fig. 1.8 illustrates AR in a continuum of interface for accessing digital data. Based on creating and experiencing an AR application, the AR software are used for
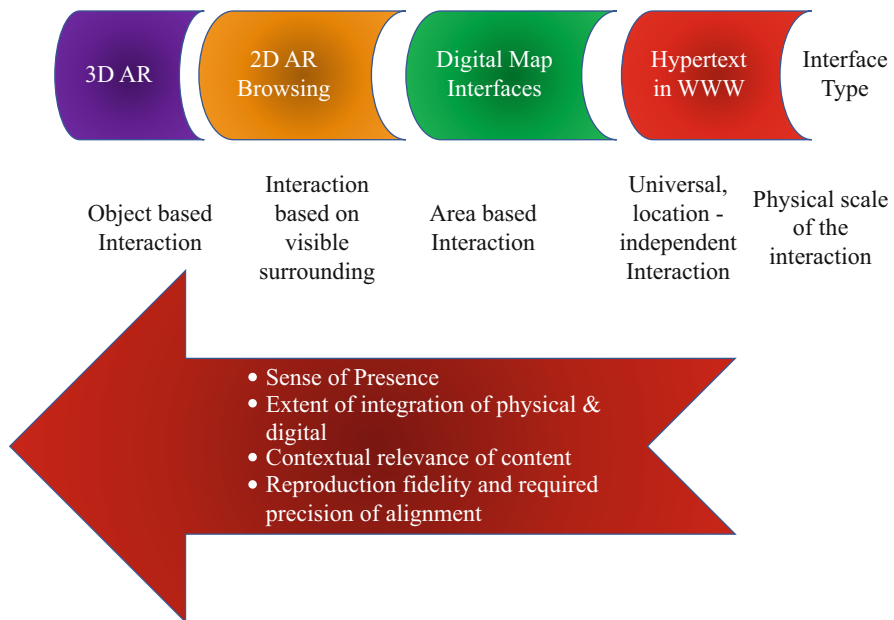
Fig. 1.8 AR in a continuum of interface for accessing digital data

- Creating an AR application
- Directly and indirectly used AR application
- Creating AR content for an AR application
- In specific for augmented reality system, software can also be conceptualized as
- Rendering libraries
- Application building libraries
- Programming libraries (tracking software)
- Stand-alone applications (AR authoring)
- Plug-in software (for existing application)

Besides AR hardware and software components, there are some additional elements required for attaining an AR experience, namely,

- AR application
- Technology
- Physical world
- Content
- Interaction
- Users

1.5.5 Mobile Augmented Reality

ARS (augmented reality system) merges virtual information along with physical environment. Mobile augmented reality systems explore the synergy of two promising fields of user interface research: geolocation which has the GPS tracking capabilities to pinpoint the user's location and its compass to detect device orientation and augmented reality.

Mobility is the most implicit way for AR due to it increased capacity to use computer vision and user position and computing. MAR is a momentous part of the AR and mobile computing fields. Mobile augmented reality systems (MARS) perform the task of integrating without considering the user's whereabouts. Figure 1.8 illustrates AR view on smartphone. MARS enhances user's perception by merging 3D virtual object into 3D ambiance in mobile context. MAR focus on multiple objects in the ambiance, and all MAR applications are highly decentralized. Due to dynamic nature, the most commonly faced challenges are unique design including real-time information retrieval, object tracking and recognition, information visualization, and user interaction [30]. Likewise, in MAR, a new interaction allegory requires designers to consider ways to meliorate the appeal and aesthetics. Smaller display size is one of the challenges which increases the complexity; therefore, a combination of tangible user interface and graphics might be required. On the other hand, smartphones have a very limited manipulation. Therefore, MAR design should promote user support, namely, application responsiveness and very subjacent physical effort.

Extend Azuma's [26] definition of AR to MAR in a more general way as follows:

- Combine real and virtual objects in a real environment.
- Interactive in real time.
- Register (align) real and virtual objects with each other.
- Run and/or display on a mobile device.

AR technologies are becoming credulous for navigation, for gaming, and mostly for browsing location-based content through widely adopted MAR applications such as Layar1, Junaio2, Wikitude3, and Acrossair4. Great revolution is how we understand, access, and enjoy the digital information in the future mobile world just by pointing a mobile camera toward an object and acquire context-sensitive information. This is truly a challenge for designing sector.

The architecture and fundamental components of mobile augmented reality are shown in Fig. 1.9. The user experience of MAR system is also great challenge to predict or envision even though MAR is a mind-blowing nascent technology which provides a novel interface to contextual information.

Fig. 1.9 Fundamental components of mobile augmented reality

MAR-based applications use different architectures. As a review, mobile augmented reality applications run on:

- Handheld devices (e.g., smartphones, tablets, etc.)
- Web (as web application)
- A cloud with a thin client
- Handheld system connected to remote server(s)
- Desktop/laptop computer
- Desktop/laptop computer connected to remote server(s)

1.5.6 Tracking in MAR

Calculating the relative pose of a camera in real time is known as tracking. Sensors perform the task of tracking. For a mobile augmented reality system, tracking is one of the most paramount components. Based on devices used, visual tracking, hybrid tracking, optical tracking, and sensor tracking are some of the atypical tracking method. Sensors provide information about different things in the real world (location and orientation). For instance, an MAR application can find the whereabouts of the user and his or her position in the real world. Visual tracking methods are of special interest in AR as the camera is already a part of the system. There is multifarious genre of tracking such as

- Optical tracking
- Acoustical tracking
- Electromagnetic tracking
- Mechanical tracking
- Depth sensors-based tracking
- Multiple sensors-based tracking

1.5.6.1 Optical Tracking

Being the advantageous and most commonly used tracking system especially for indoor applications, based on computer vision which is one of the reasons for opting optical tracking, camera is a specific sensor used for optical tracking. Not only cameras with visible light range are considered, but also cameras adaptable with ultraviolet and infrared are also desired. Motion capturing system is used in optical tracking where performance of complex motions of multiple entities is main criteria. The realistic motions for virtual content and tracking multiple entities in an AR application are made possible using optical tracking. Typically, in optical tracking, many cameras are present in an ambiance, and the virtual elements are tracked with marker or any other objects. They are identified and made easy through computer vision algorithm. Some of the advantages and disadvantages of optical tracking are as follows:

Advantages
- No attachments (like wires) needed to be attached with the object to be tracked
- Physical object and real ambiance don't need any sort of physical connection
- Many entities can be tracked simultaneously
- Chintzy.
- High-quality camera relatively has become ubiquitous

Disadvantages
- In real world, enough line of sight and light is a must
- Latency is introduced in a system by optical tracking

1.5.6.2 Types of Optical Tracking Methods

1. ID Marker-Based Tracking

Fiduciary markers or ID markers use the camera on smart devices to scan the visual markers. They can be easily detected since the structure is fixed with a black border. Many information can be configured using the pattern inside many markers. If the ID markers are in circular form, then they are called as circular markers, and if they are in form of templates, they are called as template markers.

2. 2D Barcode-Based Tracking

It is currently proving to be the more popular option largely because it's so simple to implement. Because of its simple nature to implement, it is the most widely and popularly used tracking system. They are patterns with black and white squares. Better results are achieved because the pattern influences the ID of the 3D virtual object it reflects, and therefore, no image matching is required.

3. Quick Response (QR) Code-Based Tracking

The 2D representation of data items is used in this flexible machine-readable optical tracking system. They have a very large storage capacity. A single QR code symbol can contain up to 4296 alphanumeric characters, 7089 numeric character, 1817 Kanji character, and 2953 bytes of binary data. The most broadly used applications of QR code are hospitals and journals and newspapers. Some of the superimposed information over the QR code can be data of any type such as video, Wi-Fi connection, SMS, vCard, image, and text.

4. Picture Marker-Based Tracking

It has a rectangular border which is strong and distinctive. Due to its distinctive border not like borderless markers, they can be detected faster. Any arbitrary image can be inside the boundary. This type of picture marker-based tracking is in between markerless and ID markers.

5. Markerless Tracking

When any part of real environment is used as target which is tracked to view the virtual objects, it is known as markerless tracking. The environment characteristics

and information can also be retrieved later. It is based on robust and specialized trackers. Complexity, increased computational cost, and device dependency are the most common challenges faced in this type of tracking. Most of the markerless applications are based on cloud computing due to the low computational power of the smartphone.

6. Location-Based Tracking

Accuracy of an indoor location is not acceptable to that level even though GPS in AR applications are very successful. Therefore, wireless LAN networks and infrared networks or RFID are used as alternative approaches. But they are not used very commonly because it is not cost-effective and also no guarantee of accuracy is provided.

7. Natural Feature Tracking

As long as the image is complex enough, natural feature tracking can be used. A mobile application that can recognize a movie poster is a good example of a natural feature tracking application. The application can analyze the poster and identify it by comparing the poster image to similar image in natural feature tracking [300]. In many of the natural tracking method-based applications, eye, face, and sound are also used.

1.5.7   Registration in MAR

When the real-world image acquired by the camera blends with the virtual object generated either mobile or computer, then the process of registration occurs. This process is one of the major key issues in MAR. Initially, the position between the user and the virtual elements is confirmed, and then the projection of virtual objects in visual boundary of the user happens. This is called as projection transformation. Viewing, modeling, creating and viewing frustum, and projection are the steps involved in the process of registration. The position of the camera is set as the external parameter, and projection matrix is set as internal parameter. Common classifications of registration in MAR are computer vision-based registration, knowledge-based registration, and tracker-based registration. Based on loop, registration is further divided into two types such as open loop-based registration and closed loop-based registration [31, 32].

1.5.8   Interaction

Interaction plays a vital role in the overall user proficiency. MAR itself is an interaction medium. It is the impact of one stage or thing on another. The response or reaction of one entity in MAR to the other one is coined as interaction. Many

interaction entities are indentified when the user expertise MAR. These interactions result into manipulation, communication, and navigation. Some of the interaction happens between:

- User and MAR application
- User and real ambiance
- User and virtual ambiance
- Real ambiance and virtual ambiance
- One user with the other one through MAR application

1.5.8.1 Object Recognition and Detection

The ability to identify specific object either in a video or in an image is called as object recognition. A large memory with higher graphic ability is expected in computer vision for object recognition. Model-based approach and appearance-based approach are the main two techniques of object recognition. Multiple algorithms with high memory storage requirements are used in model-based approach, whereas in appearance-based object, recognition uses only small database for colors and shapes.

The technique of detecting the type and subtype of an object is known as object detection in computer vision.

1.5.8.2 Satellite Image

A huge volume of data related to Earth are collected through satellite technologies. Urban areas expansion, agricultural land usage, and changes in climate are remotely monitored based on artificial and natural phenomena using satellite image [18, 19, 20, 21, 22]. Remote sensing analysis faces another challenge such as the granularity of the analysis. Two- or three-dimensional images of objects used in a radar is coined as synthetic-aperture radar (SAR) [13, 14, 15, 16, 17]. The motion of the radar antenna is used over a target region in SAR to provide finer spatial resolution [7].

1.5.8.3 Convergence of Satellite Image, Augmented Reality, and Artificial Intelligence

The ground equipment transmits signal to the satellite. The satellite amplifies the incoming signal and changes the frequency. Signal is transmitted back to Earth as raw data.

Geospatial data technology as shown in Fig. 1.10 uses various steps such as satellite image processing and digitization once it acquires the data followed by the spatial data analysis planning using AR which gets input and gives output to remote sensing provider. The satellite database is developed which maintains geo data and

Fig. 1.10 Geospatial data technology

all the 3D model. Special algorithms are used for training the data, for classifying them, and for segmentation and feature extraction. Better results equaling or even better than human are generated once the data are analyzed, viewed, and planned.

The raw data are acquired from the satellite image and sent for preprocessing. The reliability check is performed, and if it is found as error-free, then the data storage decision is taken (either real time, off-line), then filtration happens, and after that, the aggregation, complication, and result storage of data are done. Analysis based on AR view is done, and decision is executed and forwarded to the application. The raw data is compressed and then the radiance is found along with the geometric values, all the metrics for correction are determined, and finally, the reflectance is obtained [11, 12]. These are the steps followed in satellite data analysis process as shown in Fig. 1.11.

The geo data are derived from the satellite, and after preprocessing the acquired data, the changes are detected. AR module performs tracking, rendering, and AR display. 3D models and geo data are stored in the database which can be later retrieved as input to analysis support tool as shown in Fig. 1.11. In order to retain homogeneous region in the image, segmentation and clustering are performed. This results in a cluster map where pixels with similar reflectance values fall into the same category [8, 9, 10]. The stages of images are shown in Fig. 1.12.
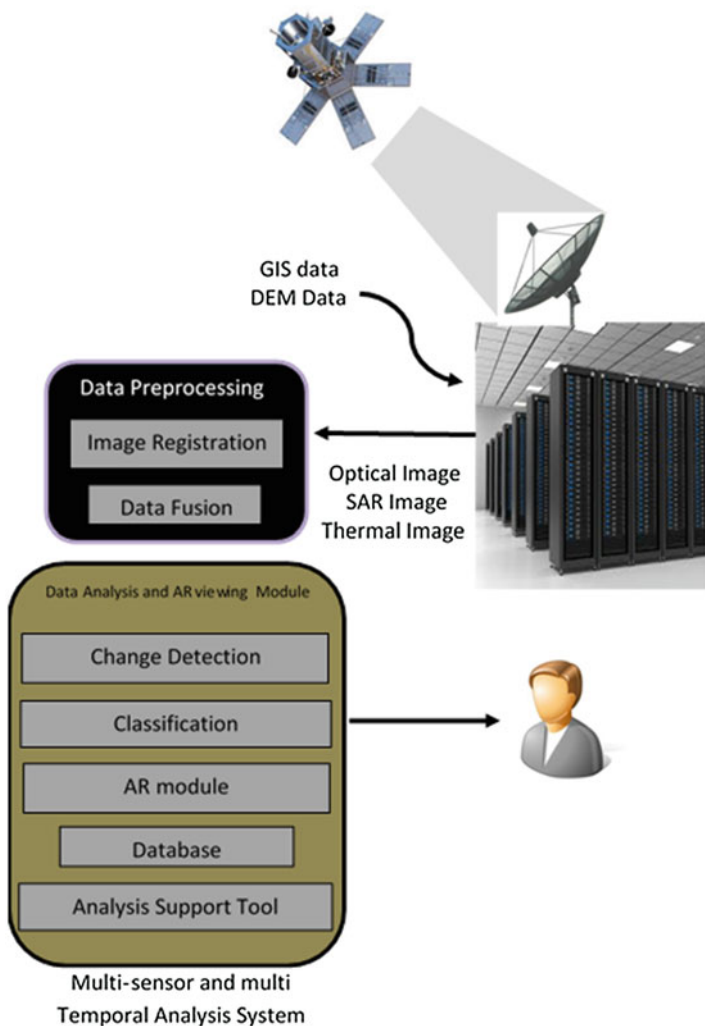
Fig. 1.11 Data analysis model for SAR image

1.6   Applications of Convergence of SAR, AI, and MAR

1.6.1   Damage Detection

Disaster response and relief require quick reaction and up-to-date information to enable decision maker taking the proper actions. Aerial and satellite imagery provide detailed information from remote sensing service provider. The data are preprocessed and the changes are detected [1, 2, 3, 4]. The damages are predicted
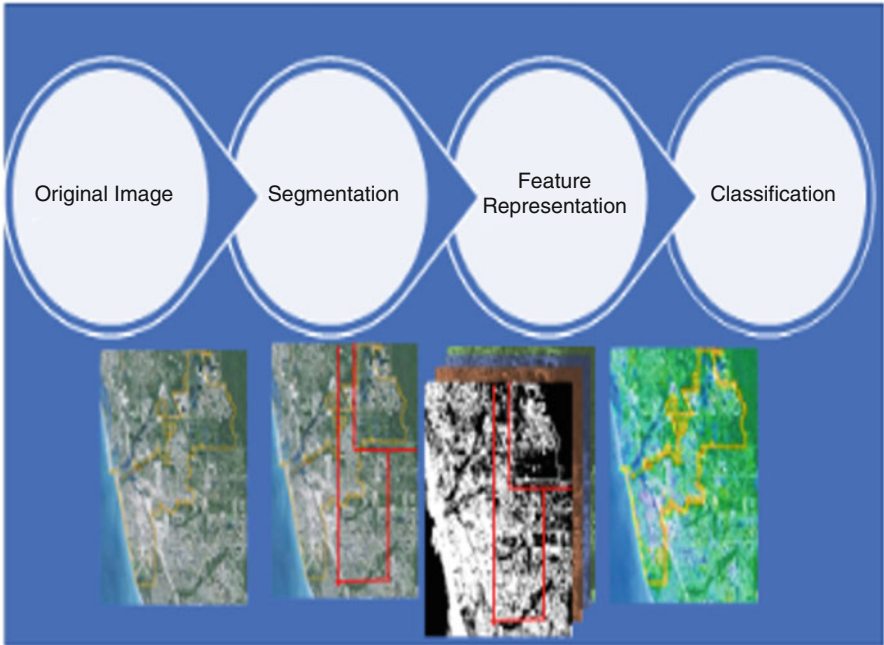
Fig. 1.12 Feature identification and tracking for satellite image
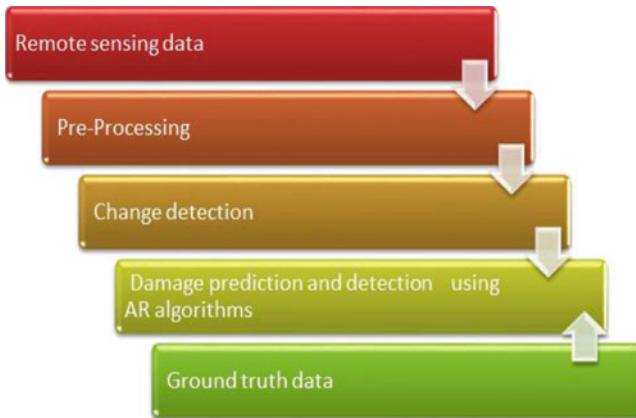
Fig. 1.13 Damage detection system for buildings

and predicted using AR and AI algorithms. The whole process as shown in Fig. 1.13 works based on the ground data input.
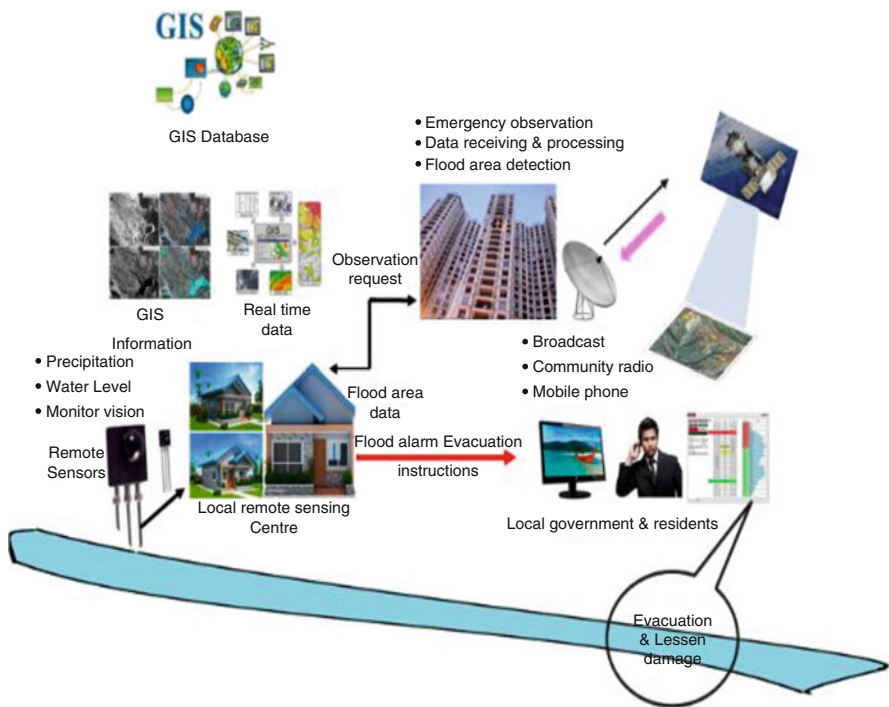
Fig. 1.14 Alarm system and flood monitor using mobile augmented reality

1.6.2 Flood Monitoring System

The ground water is detected through the GIS information obtained from the real-time data. The local remote sensing center is given an emergency flood alarm evacuation instruction to the local body. In turn, the local government alerts the resident regarding emergency observation and latest update based on the observation request as shown in Fig. 1.14.

1.6.3 Land Area Usage System

From the early land cover data, some samples are taken and spatial data are analyzed. AR module in providing the necessary 3D view is followed by classification of pixels based on regulation. After classification, the final AR display is viewed as shown in Fig. 1.15. Formation of collections of satellite pictures is done initially, implementation of results of images is computed, and the machines are trained. Data is classified based on correctness of geometric data. The ground truth data is modified for detection of the buildup area that is clipped [5, 6].
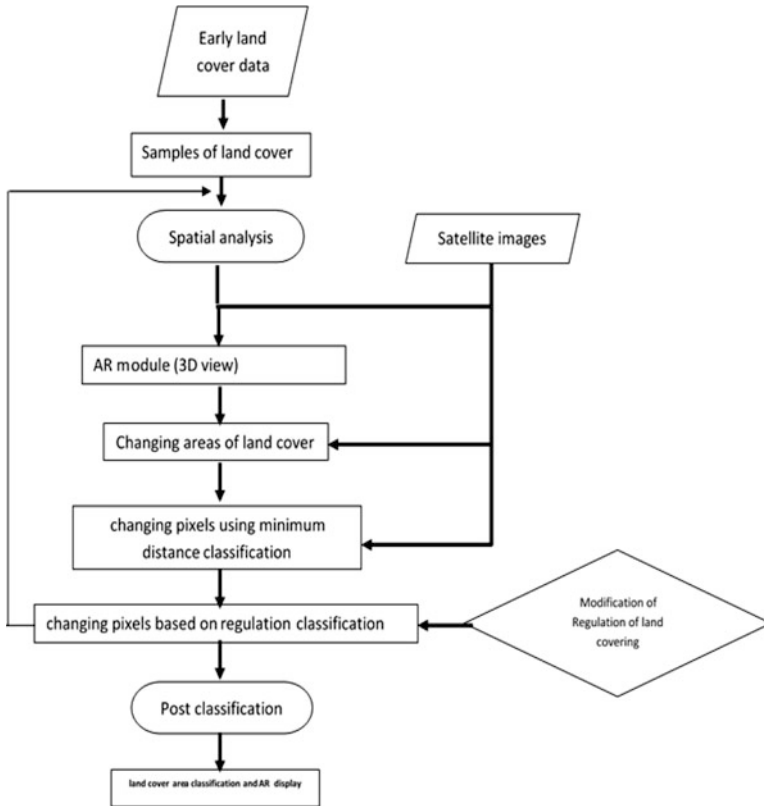
Fig. 1.15 Land coverage system

1.6.4 Accuracy Assessment System

Accuracy assessment measures many factors such as mean accuracy, overall accuracy, error inclusion, and error of exclusion.

The GIS data are classified based on decision tree classification, and the level of accuracy is measured. The satellite data are processed and trained based on the AR 3D data model.

1.7 Conclusion and Future Scope

Artificial intelligence by 2025 is going to be the highest revenue-generating resource in IT market. Augmented reality (AR) has become one of the most extensively researched fields in recent years. Learnability apperception and geo data manage-

ment system are undeniably the need of the hour. AR content creation (authoring) and AI with satellite image are both sides of a coin. They are becoming indispensable part of reality. Many of the integrated system works only for some limited transform domains which can be enhanced for other transforms also. Only IRS-ID images are mostly tested which can be enhanced to use other type of satellite images. Also, medical images can be detected and analyzed. Satellite image object domain and ontology work well for various objects and data properties. Segmentation of satellite images can also be done based on ancillary data. Ontological analysis can be done for better knowledge sharing among experts. Digital cartography can be enhanced in various aspects. More works and focus can be done on content-based image indexation concept.

References

1. Ding Y, Yao G, Zhao Y (2017) Remote Sensing Image Content Retrieval Based on Frequency Spectral Energy, International Congress of Information and Communication Technology. Proc Comput Sci 107:448–453
2. Thomas B, Geoffrey J, Maggi K, Stefan L, Peter H, Elisabeth A (2014) Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS J Photogramm Remote Sens 87:180–191
3. Raul Q, Freek M, Harald W, Frieke C, Dirk T (2017) Evaluating fuzzy operators of an object-based image analysis for detecting landslides and their changes. ISPRS J Photogramm Remote Sens Geomorphol 293:240–254
4. Andrianov DE, Eremeev SV, Kuptsov KV (2015) The review of spatial objects recognition models and algorithms, Procedia Engineering 129: 374–379; Arun PV, Katiyar SK (2013) An evolutionary computing frame work toward object extraction from satellite images, The Egyptian Journal of Remote Sensing and Space Sciences 16: 163–169
5. Jawaka SD, Rautb DA, Luisb AJ (2015) Iterative spectral index ratio exploration for object-based image analysis of Antarctic coastal oasis using high resolution satellite remote sensing data, International Conference On Water Resources, Coastal and Ocean Engineering (Icwrcoe 2015)
6. Lei M, Manchun L, Xiaoxue M, Liang C, Peijun D, Yongxue L (2017) A review of supervised object-based land-cover image classification. ISPRS J Photogramm Remote Sens 130:277–293
7. Tiana W, Biana X, Shaoa Y, Zhanga Z (2015) On the detection of oil spill with China's HJ-1C SAR image. Aquatic Proc 3:144–150
8. Mazur, Wåhlina, Krężel (2017) An object-based SAR image iceberg detection algorithm applied to the Amundsen Sea. Remote Sens Environ 189:67–83
9. Vahid S, Hamid E, Farshid F (2013) A new model for automatic normalization of multitemporal satellite images using Artificial Neural Network and mathematical methods. Appl Math Model 37:6437–6445
10. Khiali L, Ienco D, Teisseire M (2017) Object-Oriented Satellite Image Time Series Analysis using a Graph-Based representation. Ecol Inf 43:52–64
11. Dimple K, Minakshi K, Pradeep (2014) Automated object recognition from high resolution satellite images: a review. Int J Adv Res Comput Eng Technol 3:4277
12. Gerhardinger, E, Pesar S, Rottensteiner F, Hinz S (2005) Vehicles detection from very high resolution satellite imagery, 34,
13. Jenice R, Kumudha R (2016) An overview of technological revolution in satellite image analysis. Jf Eng Sci Technol Rev 9(4):1–5

14. Lakhwinder K, Vinod S, Guru (2014) Object Detection from the Satellite Images using Divide and Conquer Model, SSRG. Int J Comput Sci Eng 6

15. Sirmacek B, Angelo P, Reinartz P (2015) Detectıng complex buıldıng shapes ın panchromatıc satellıte Images for dıgıtal elevatıon model enhancement, ISPRS Istanbul Workshop 2010 on Modeling of optical airborne and spaceborne Sensors. Proc Eng 129:374–379

16. Brook RK, Kenkel NC (2002) A multivariate approach to vegetation mapping of Manitoba's Hudson Bay Lowlands, Taylor and Francis. Int J Remote Sens 23(21):4761–4776

17. Himanshu R, Nirvair N (2017) Water detection using satellite images obtained through remote sensing. Adv Comput Sci Technol 10(6):1923–1940

18. Samik B, Nitin G, Sukhendu D, Pinaki R, Sinha LK (2016) Detecting aircrafts from satellite images using saliency and conical pyramid based template representation. Indian Acad Sci:1155–1171

19. Aldighieri B, Testa B, Bertini A (2016) 3D exploration of the San Lucano Valley: virtual geo-routes for everyone who would like to understand the landscape of the Dolomites. Geoheritage 8(1):77_90

20. Brilha J (2018) Geoheritage and geoparks. In: Reynard E, Brilha J (eds) Geoheritage: assessment, protection, and management. Elsevier, Amsterdam, p 323_336

21. Vallejo M, Gallego E (eds) Towards the balanced management and conservation of the geological heritage in the new millenium. Sociedad Geoló´gica de Espan˜a, Madrid, pp 315–318

22. http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus

23. Rolland J, Hua H (2005) Head mounted display systems. In: Encyclopedia of optical engineering. Dekker, New York, pp 1–14

24. Park Y, Woo W (2006) The ARTable: an AR-based tangible user interface system. In: Pan Z, Aylett R, Diener H, Jin X, Göbel S, Li L (eds) Technologies for E-learning and digital entertainment. Edutainment 2006, lecture notes in computer science, vol 3942. Springer, Berlin, Heidelberg, pp 1198–1207

25. DeFanti TA, Dawe G, Sandin DJ, Schulze JP, Otto P, Giradoc J, Kuester F, Smarr L, Rao R (2009) The Star CAVE, a third-generation CAVE and virtual reality OptIPortal. Futur Gener Comput Syst 25:169–178

26. Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst, Vol. E77-D (12):1–15

27. Chang G, Morreale P, Medicherla P (2010) Applications of augmented reality systems in education. In: Gibson D, Dodge B (Eds.), In Proceedings of SITE 2010 society for information technology & teacher education international conference, San Diego, USA, pp 1380–1385.

28. Newman J, Schall G, Schmalstieg D (2006) Modelling and handling seams in wide-area sensor networks. In Proceedings of the 10th IEEE International Symposium on Wearable Computers, pp. 51–54

29. Hirano Y, Garcia C, Sukthankar R, Hoogs A (2006) Industry and object recognition: applications, applied research and challenges. Toward category-level object recognition, LNCS 4170, Springer-Verlag, pp 49–64

30. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001) Recent advances in augmented reality. IEEE Comput Graph Appl 21:34–47

31. Kumar Singh V, Ali A, Nair PS (2014) A report on registration problems in Augmented Reality. Int J Eng Res Technol 3(4):819–822

32. Adams B, Ghosh S, Wabnitz C, Alder J (2009) Post-tsunami urban damage assessment in Thailand, using optical satellite imagery and the VIEWSTM field reconnaissance system. In: The 1755 Lisbon Earthquake: Revisited, Geot Geol Earthquake, Springer Netherlands, pp 523–539

Chapter 2
Multithreading Approach for Clustering of Multiplane Satellite Images

Check for updates

C. Rashmi and G. Hemantha Kumar

Abstract This paper presents the clustering of multiplane high-resolution orthoimagery and multispectral satellite images. Two well-known clustering techniques k-means and ISODATA are usually used for classification. K-means clustering is used in this paper for the classification. Since the clustering of satellite images of pixel dimension greater than 1000×1000 has increased execution time, hence it is considered for the parallelism. This paper depicts the data parallelism exhibited by different threads in cores of a processor in the legacy system using GPU by assigning the tasks among different threads independently. A framework of parallel computation is exhibited for clustering multiplane high-resolution orthoimagery satellite images and Landsat MSS datasets. A parallel block processing implementation for clustering has been exploited and tested specifically on CPU achieving an efficient speedup on multicore processor by varying with 2, 4, 8, and 12 threads with variation in number of clusters 2, 4, 8, and 12. Around 10 samples of MSS sensor and high-resolution multiplane orthoimagery satellite images are considered for clustering with the usage of MATLAB 2017a environment. Hardware resources are efficiently used from the results obtained in parallel approach resulting in time depletion compared to serial k-means clustering. This approach can be applied for processing remote sensing images as results are acceptable.

Keywords Classification · K-means clustering · Multiplane · Parallel · Orthoimagery · Threads

C. Rashmi (✉) · G. Hemantha Kumar
High Performance Computing Project, Department of Studies in Computer Science, University of Mysore, Mysore, India

2.1 Introduction

Digital image processing with an aid of computer plays a major role for remote sensing applications involving the interpretation and manipulation of digital images. Nowadays, access to low-cost, efficient computer hardware and software is commonplace, and the digital image data sources are many and varied. These sources range from earth resource satellite systems to the meteorological satellites, airborne scanner to digital camera data, to image data generated by scanning microdensitometers and other high-resolution digitizing systems. Image classification and its analysis operations are used to digitally classify and identify pixels in an image data. It is necessary to classify features of an image in remote sensing applications. Feature identification of remote image data is automated by the replacement of visual analysis with the quantitative techniques which involve analysis of multispectral data for land cover detection. The decision rules are based on geometric shapes, sizes, and patterns in spatial pattern, but decision rules based on spectral radiances observed in the data are referred as spectral pattern recognition categorizing all pixels in a digital image into several themes are an intent of classification.

Supervised and unsupervised are the types of classification [1]. Unknown pixels are examined in an image and aggregated into cluster or number of classes present in image values in unsupervised image classification. Training data is not utilized in unsupervised classification. Classes that are spectral are the results from unsupervised classification. The natural groupings of spectral present in a dataset are determined by various clustering algorithm. K-means approach [2] is one common form of clustering, accepting number of clusters from analyst to be located in the data. K-means is computationally intensive because it is iterative in nature. Unsupervised training areas are chosen to contain numerous cover types at various locations throughout the scene. An iterative approach k-means considers only graphical distances between points using a chosen distance metric; hence, it is a lightweight algorithm for spectral clustering. Since k-means is widely used and hence is considered for parallel processing exhibiting task parallelism, SPMD parallel programming [3] data model is exhibited by parallel block processing for data parallelism. An image is processed in blocks rather than all at once. Due to high computation time for larger images, block processing is performed parallely in MATLAB programming. The main focus is to tackle this problem in time reduction. In this paper, parallel processing approach for clustering the multispectral satellite images using k-means is proposed.

This paper is organized as Sect. 2.2 explains about the survey and Sect. 2.3 illustrates parallel processing approach for clustering the multispectral satellite images with results.

2.2   Literature Survey

Three approaches utilized are distributed computing known as Message Passing Interface (MPI), shared memory programming known as Open specification for Multiprogramming (OpenMP), and heterogeneous computing for graphics processing unit known as CUDA programming techniques [4] explored for the performance of k-means clustering algorithm. Since GPUs have more cores which are suited for parallel calculations, GPUs are not well suited to process smaller datasets as overhead in communication. Their experimental results show around 35x speedup. For smaller images, OpenMP are used, while a CUDA outperforms larger images. Their experimental results show around 35x speedup [5]. describes the floating point divide unit is implemented for multispectral satellite images by applying k-means clustering algorithm. The usage of fp_dix, float2fix, and fix2float is exhibited for k-means clustering. FPGA hardware base is implemented for the computation of k-means. Implementation dominates the run time in data transfer time which includes the time for reading an image data from external input image [6]. illustrates about an automated clustering algorithm of nonparametric, and application to unsupervised image segmentation is described in this paper. Peak climbing approach is used for the analysis of the multidimensional histogram. The algorithm developed is parallelized, and a SEQUENT parallel is used for simulation. Parallel computing methods are exploited in [7] for retrieval of geophysical parameters for the case study of AOD, Aerosol Optical Data, retrieval. Parallel execution on multiple cores of processors and GPUs is implemented. Dynamic and static workload distributions are performed on a heterogeneous CPU and GPU configurations [8]. depicts the parallelization of high-resolution multispectral satellite images and its fragments using distributed programming model known as Message Passing Interface in MPP cluster environment. This has been implemented in GRASS GIS software. In their work, neural network-based radial basis function is utilized for processing of agricultural regions and forest area classification of high-resolution orthoimagery multispectral satellite images. Texture and fractal based approaches for processing forest and agricultural regions are retested in project framework [9]. It mainly focuses on fusion of data of several monospectral images on the segmentation of multispectral images. During segmentation, fusion of data is performed on different images. Parallely, connection machine 5 is implemented for the multispectral image segmentation. Mean shift algorithm based on CUDA is presented [10] for the segmentation of multispectral images. Despite the high level of noise, taking a long processing time in extraction of features is made visible clearly. Authors presented a method of analyzing multispectral images of highly noised comprising of different objects. Fourteen different image fusion techniques for the fusion of spectral information from multispectral images and spatial resolution of panchromatic images are presented in their work. The work depicts about the increase in time consumption for serial compared to parallel. The work [11] mainly focuses on parallelization of different monospectral images with an explicit exchange of information and synchronization, by treating image data fusion as an extension of segmentation

methods [12]. mainly focuses on classification of cluster area by consideration of different land cover types of satellite images. Three clusters are considered for Landsat 7 satellite images. The bands 2, 3, and 4 from Landsat satellite is considered for the classification in their work. Reinforcement programming [13] is proposed for optimization of k-means clustering for the segmentation of multispectral images. Synchronous parallel programming Parallaxis III on low-level image processing is focused [14]. Single-instruction multiple data (SIMD)-based data parallelism is implemented on the low-level routines of image processing operations. Parallaxis has been used in their work for an implementation of image processing algorithms.

2.3   Parallel Computation

Acceleration of processing speed in various algorithm is exploited by parallel computation. Task-level and data-level parallelism are the two kinds of parallelism for processors involving multiple cores. Task-level parallelism is considered as MIMD. Multiple tasks that are independent are processed simultaneously in order to increase the processing speed. Data-level parallelism is glanced as SIMD. It indicates that a large amount of data is split into smaller chunks and processed simultaneously by several processors using same algorithm. Although task-level parallelism exist, data-level parallelism can be exploited efficiently in computer vision and image processing algorithms for the following reasons. Firstly, task-level parallelism has a limited processing speed improvement. The computational requirements of most image processing significantly increases with an increase of image pixel resolution which can only be compensated by the exploitation of data-level parallelism. Secondly, data-level parallelism is more often present in image processing algorithms.

Data-level parallelism mainly exists in image processing algorithms of lower level, analyzing the pixel-level parallelism information. This may attribute the significant part of the computation requirement of the whole computer vision system. Data-level parallelism is exploited in most computer vision system that takes the advantage of SIMD computational mode, significantly increasing their processing speed. In parallel computation, thread is the most commonly used in this paper. Thread is a sequence of instruction that are treated independently. Thread can be used in both data-level and task-level parallel programs.

Multicores are used efficiently for processing the satellite images parallely. Time for processing is high for large images of pixel dimension greater than 1000×1000. However, time is reduced in parallel processing where the number of processors can be used efficiently, especially for the case of lower levels of image processing when processing is done at pixel level. Some operations require an image to be processed in blocks rather than performing operations on an entire image. An image is processed in blocks called sections rather than processing an entire image. Distinct blocks are processed simultaneously by data distribution among the collection of MATLAB sessions called workers in parallel block processing [15]. MATLAB pool

has a collection of workers. A specified function is called by different blocks at a time and consolidated to form an entire clustered output image.

Distinct block operations and sliding neighborhood are the basic operations for performing parallel block processing. A pixel-wise manner of processing an image is performed in sliding neighborhood operation. Neighborhood is a rectangular block sliding over an entire image with the center pixel. Fig. 2.1 shows the blocks of neighborhood with 2×3 sliding blocks in 6×5 matrix, where center pixel is represented as dot, i.e., some operation is performed on each pixel in order to determine the value of corresponding pixel in an output image.

Rather than a pixel at a time, an image is operated in subblocks in block processing. Based on the pixel dimension, image is divided into tiles of two, four, twelve, and so forth as shown in Fig. 2.2. Distinct threads process each block parallely at a time, performing clustering operations. Function to be performed on blocks, image size, and block size used for processing are the three factors affecting the performance of parallel block processing. Faster performance is exhibited, while
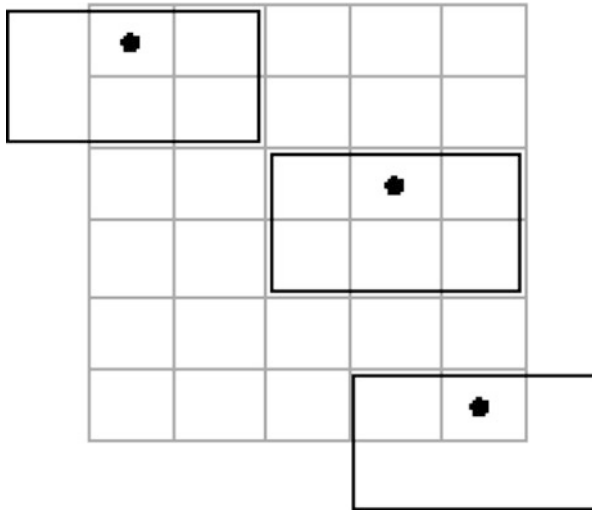
Fig. 2.1 Neighborhood Blocks
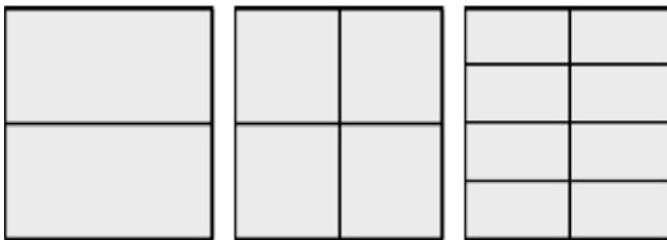
Fig. 2.2 Two, four, and eight subblocks division of an entire image

Fig. 2.3 Parallel block processing approaches row-shaped block, column-shaped block, and square block

processing larger subblocks rather than computing the same tasks using smaller blocks, sometimes algorithm applied to an entire image requires a certain block size and smaller blocks need to be used.

Column, row, and square shaped are the approaches resulting the influence of performance measurement as illustrated in Fig. 2.3. In first approach, height of an image is spanned in column-shaped block. Width of an image is traversed in row-shaped block in second approach, and a square block of equal size dimension is traversed in third approach. In our work, square block division over an entire image is exhibited for the optimization of clustering. Sequential execution by a single core and parallel execution of parallel block processing by multiple threads/cores is illustrated for clustering of multispectral images of 4 bands using k-means in Sect. 2.4 in detail.

2.3.1 K-Means Clustering of Multispectral Images

Multispectral images are multiband images of an electromagnetic spectrum that are captured from different sensors. It is collection of several monochrome images captured using different sensors. Multispectral images are widely used for remote sensing applications. Multispectral images [16] are captured from the sensor MSS—multispectral scanner of Landsat 4 and Landsat 5 satellite of MSS bands 1, 2, 3, and 4 of 0.5–0.6 μm, 0.6–07 μm, 0.7–0.8 μm, and 0.8–1.1 μm sensitivity. Each Landsat MSS scene is framed from the continuous MSS data swath; it covers 185×185 km area. A nominal scene consists of some 2340 scan lines with 3240 pixels per line or about 7,581,600 pixels per channel with four spectral observations/pixels each image dataset containing over 30 million observations. Landsat images have distinct area covered per frame and image scale compared to conventional aerial photographs, for example, more than 1400 aerial photographs at 1:20,000 scale covering an of single MSS image with no overlap are required. Images of Landsat is considered as a complementary interpretative tool compared to low-altitude aerial photographs. Most MSS images can be studied in two dimensions, and aerial photographs are acquired in stereo for the study of real-time applications of soil discrimination in vegetation, cultural feature identification, forest-type mapping, and so forth.

Clustering multispectral satellite images into distinct regions is a complex problem. In this paper, Landsat 4–5 MSS satellite images consisting of 4 spectral bands are considered for clustering. Dataset containing n objects are partitioned into k groups by an unsupervised k-means algorithm where k indicates the number of groups. Random centroid points are represented for each group. The data points are grouped together by the calculation of Euclidean distance between the centroid and the data points, aiming at minimization of square error function given as

$$J(C) = \sum_{i=1}^{n} \sum_{j=1}^{n_j} \left(x_i - c_j \right)^2 \tag{2.1}$$

where:

$x_i - c_j$ is euclidean distance between x_i and c_j.
x is set of points.
C is centroid.
c_j is the centroid at random points.
n_j is the number of data points in ith cluster.
n is the number of cluster centers.

New clusters are formed, and again the new centroids are calculated from the newly formed clusters by taking the mean of all corresponding data points. This process is repeated until they have grouped accurately. Serial k-means algorithm has been explained in algorithmic steps as follows.

Algorithmic steps for serial k-means algorithm.

1. Input: K, set of points $x_1, x_2, \ldots .x_n$.
2. Place the centroid $c_1, c_2 \ldots c_k$ at random points.
3. Repeat the process until convergence is met.

   – For each xi.

      Find the cj centroid.
      Assign point xi to cluster l.

   – For each cluster $l = 1, 2 \ldots k$.

      Mean of all points xi assigned to cluster j is the new centroid in previous step.

4. Stop when none cluster assignments change.

Memory and computational time for processing multispectral image data increase significantly; hence, it is considered for parallel processing. Figure 2.4 illustrates the block diagram of proposed parallel processing of k-means clustering for multispectral satellite images. It is a case of unsupervised method of classification where no prior knowledge is required for classification. In our proposed parallel approach, multispectral image data of 4 bands are considered for k-means clustering.

Fig. 2.4 Proposed system of parallel k-means
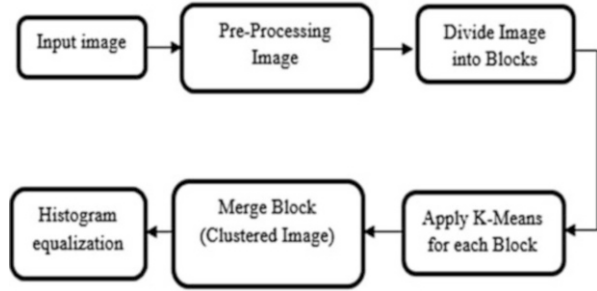
Initially, multispectral image containing bands 4, 3, 2, and 1 is read. Input image is converted to class double for further preprocessing. Contrast enhancement of linear stretch is applied on this input image using imadjust function in MATLAB. The function imadjust enhances contrast by mapping intensity values to new values. After the contrast enhancement is performed, large image is divided into subblocks of equal dimension (square-shaped) using blockproc function. Data-level parallelism is exhibited on the larger pixel dimension and processed by specified number of threads simultaneously with the usage of same algorithm. As in our work, k-means clustering is applied on these subblocks. Considering an image example of pixel dimension 4000×4000 of 32 bit depth with 2 resolution units to be processed by k-means algorithm, 100 processors running the same k-means algorithm can process these subblocks parallely leading to the decline of computation time by processing 1/100 of the image pixel (i.e., 400×400). Since same instruction, i.e., k-means algorithm, is operated on each nonoverlapping subblocks, this kind of structure is known as single instruction multiple data, hence the name data parallelism. Parallel block processing is more useful for larger images. Clustered sub-images of respective subblocks are aggregated into single large clustered image. Finally, post processing of histogram equalization is applied on the aggregated clustered image as stretchlim is not easily adapted with blockproc as it relies on the full-image histogram.

2.4 Experimental Results and Discussions

The experimental results obtained from the parallel processing of k-means for multispectral images are presented in this section followed by evaluation of performance.

The parallel image processing has been experimented on Intel Core i7-7500U CPU 2.90GHz x64 based processor 16GB RAM. Sequential and parallel block processing approach is implemented in MATLAB 2017b programming paradigm with the variation in number of clusters as 2, 4, 8, and 12 in sequential and parallel approach of k-means clustering. Number of threads is varied in parallel block

processing of k-means algorithm along with the variation in number of clusters. Three types of datasets NAIP [17], orthoimagery aerial satellite images [18], and Landsat MSS 1–4 spectral images of 4 spectral bands with 60 m spatial resolution of electromagnetic spectrum of red, green, blue, and near infrared are considered for testing. Four spectral bands are being used in our testing for parallel processing. For Landsat 4 of sensor multispectral scanner (MSS) data of 32 bit pixels cloud cover of less than 10% and less than 100%, all day and night indicator and different features such as airport, dam, forest, building, glacier, valley, and so forth with pixel dimension of ranging from 1200×800 to 9052×4965 of different size ranging from 1.11 MB to 176 MB are considered for testing. The performance is affected by the usage of blockproc function in MATLAB with number of times of read or write image file. Selecting the larger block size reduces the number of times blockproc has to access the disk, at the cost of using more memory to process each block.

2.4.1 Implementation

Standard data sample of multispectral image comprising of 4 spectral bands, red, green, blue, and near infrared, shown in Fig. 2.5 is considered for parallel block
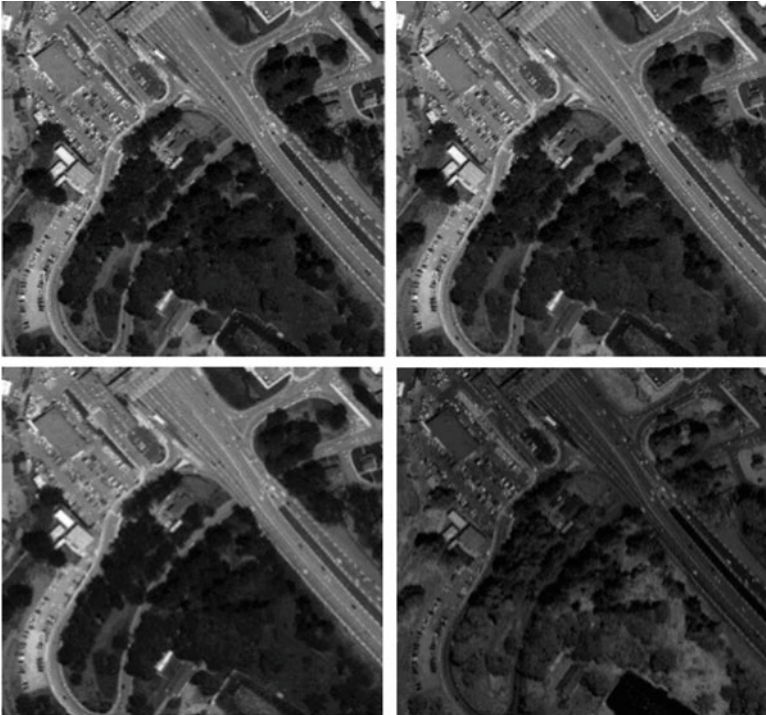
Fig. 2.5 Four bands of multispectral data

processing of k-means clustering with 10 iterations. Initially, an image obtained from MSS level 1 data products is used as an input. In our implementation, band 1, band 2, band 3, and band 4 are converted onto color space $L_*a_*b_*$ as this color space is easier to distinguish. Four clusters are segmented from the converted $L_*a_*b_*$ color space and enhanced as in Fig. 2.6. With distinct color using k-means clustering algorithm, clusters are distinguished with distinct feature space defined by its center as shown in Fig. 2.7. Image pixels are allocated the nearest cluster by measuring the euclidean distance. New clusters are formed from the new centers. Based on color-based segmentation, new clusters are formed. After cluster formation, images

Fig. 2.6 Contrast enhancement
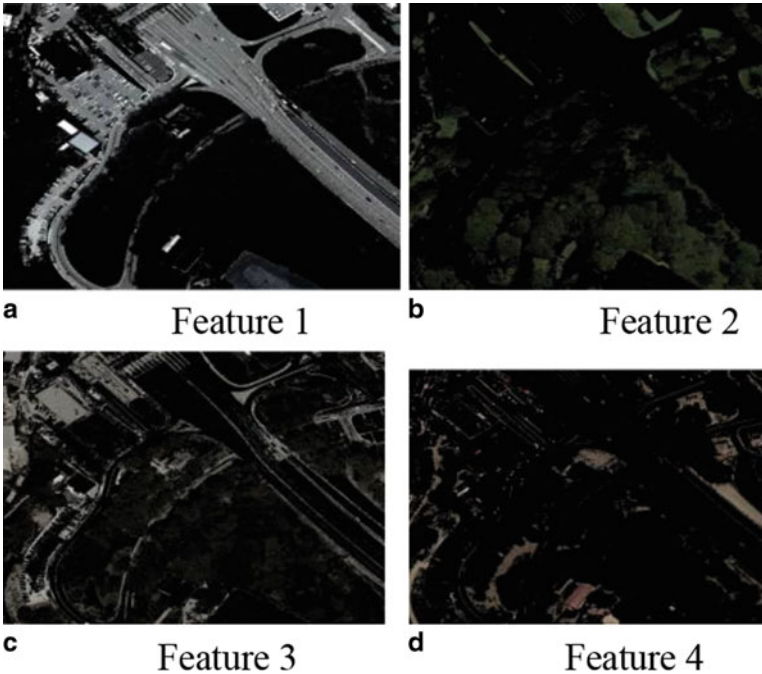
Fig. 2.7 Objects in different clusters

are classified based on pixels calculation of k-means clustering. Figure 2.8a depicts the sequential pixel classification and parallel approach for pixel classification as in Fig. 2.8b.

Figure 2.9 illustrates overall diagrammatic representation of proposed approach of k-means clustering by parallel block processing with the consideration of an image of red, green, and blue channels. An input image of pixel dimension 695 × 663 is divided into 4 subblocks of equal dimension 174 × 165. Each subblocks are processed using clustering algorithm simultaneously. Then the processed clustered subblocks are aggregated to get the final clustered output.

Following Fig. 2.10 shows sample datasets of sensor MSS, NAIP orthoimagery, and high resolution of orthoimagery aerial satellite images of four bands that are considered for testing. Sequential and parallel clustering of k-means algorithm are shown in following figures.

Following Fig. 2.11a illustrates the clustering results of serial k-means algorithm, and Fig. 2.11b depicts the clustering results of parallel block processing for cluster 2 and for 4 threads.

Figure 2.12 exhibits the output clustered images after applying the serial k-means algorithm and parallel block processing on k-means algorithm for cluster 8 with the consideration of 4 threads.
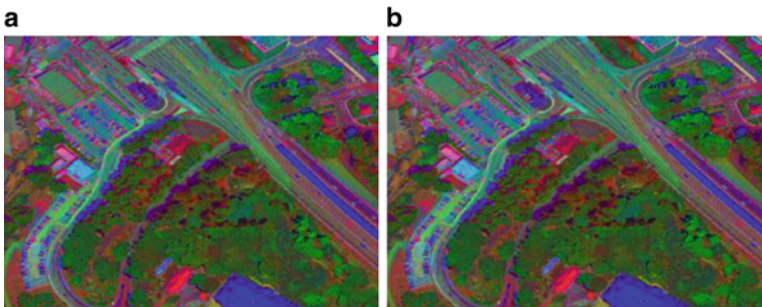
Fig. 2.8 Pixel classification (**a**) serial (**b**) parallel

Input   Sub-images   Processed sub-images   Output
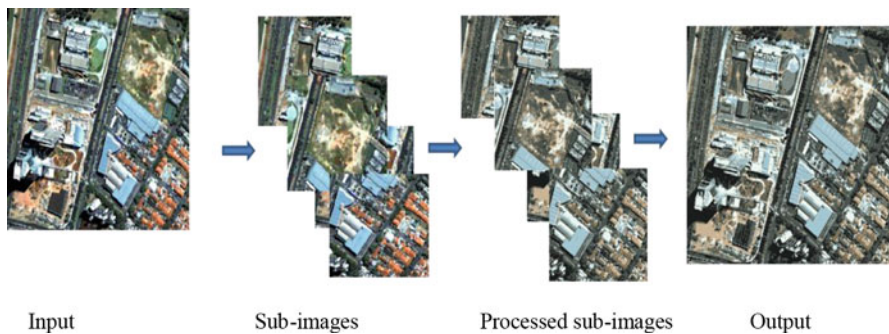
Fig. 2.9 Illustration of proposed system
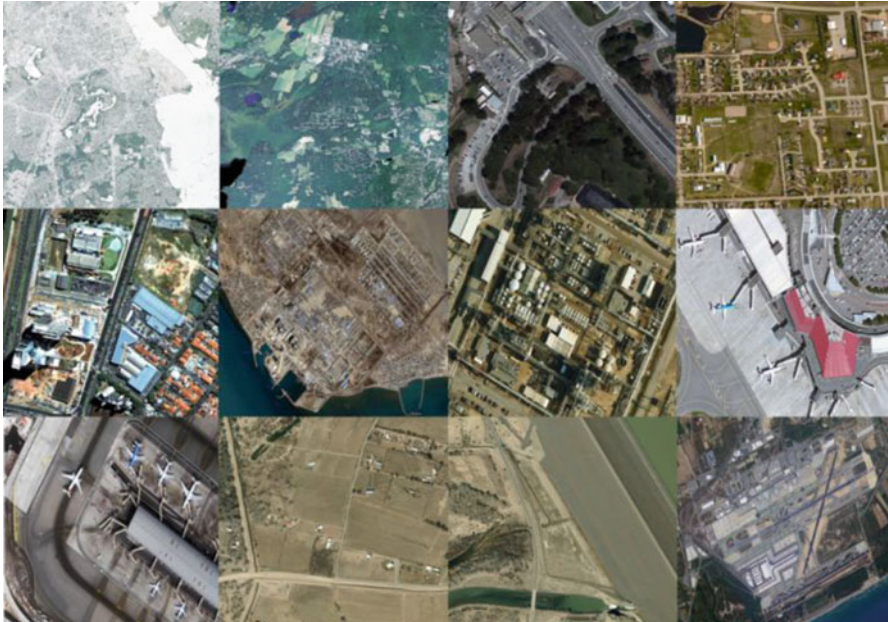
Fig. 2.10 Sample images of multispectral data
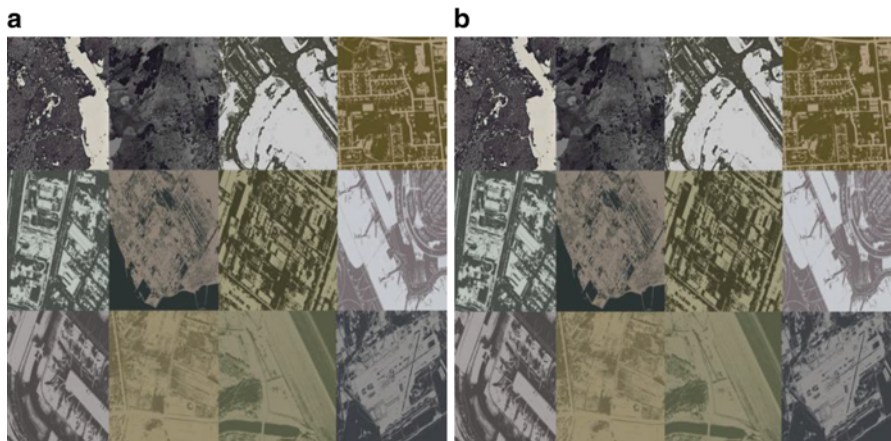
Fig. 2.11 K-means algorithm of cluster 2 (**a**) Sequential (**b**) Parallel Block Processing

Figure 2.13 demonstrates the output clustered images after applying the serial k-means algorithm and parallel block processing on k-means algorithm for cluster 12, 4 number of threads.

Tables 2.1, 2.2, 2.3 and 2.4 and highlight the comparison results of sequential and parallel computation time by variation in number of workers/cores/threads in
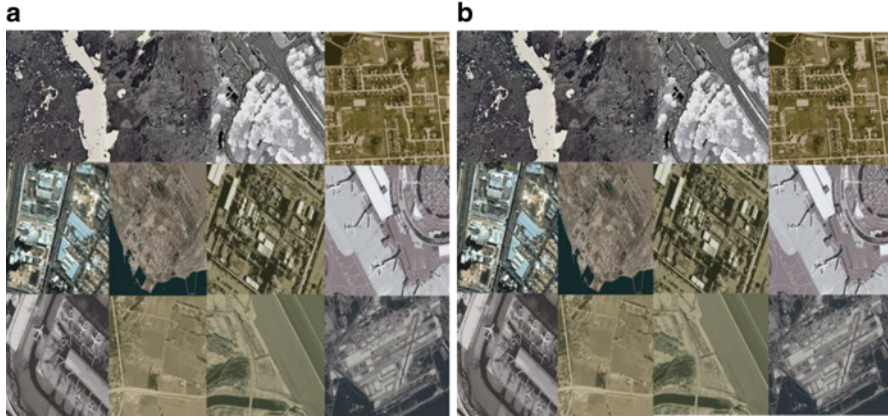
Fig. 2.12 K-means algorithm of cluster 8 (**a**) Sequential (**b**) Parallel Block Processing
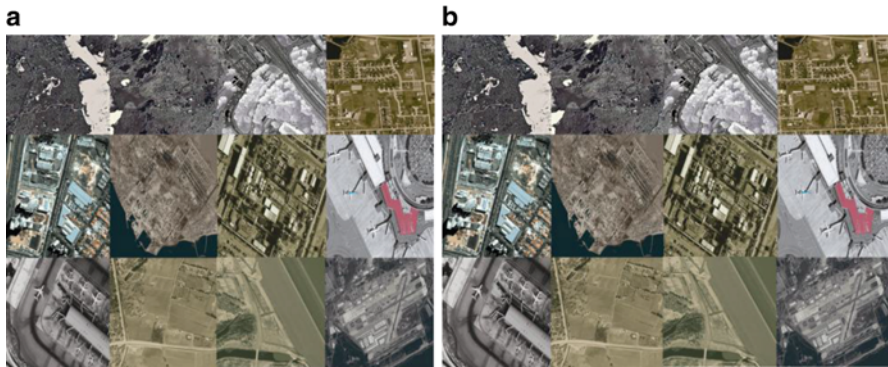
Fig. 2.13 K-means algorithm of cluster 12 (**a**) Sequential (**b**) Parallel Block Processing

MATLAB using Parallel Computing Tool (PCT) box as 2, 4, 8, and 12 for different number of clusters.

2.4.2   Performance Evaluation

The performance of parallel processing for an iterative approach of k-means algorithm has been evaluated with variation in number of workers for different pixel dimension in MATLAB programming environment.

Speedup calculation with two number of workers for cluster 2 is shown in Table 2.5 with its graphical representation in Fig. 2.14.

Table 2.1 Comparison results of serial and parallel time for cluster 2 for different threads 2, 4, 8, and 12

Cluster 2, time (ms)		Serial	Parallel			
Size	Resolution	1	2	4	8	12
650 KB	1024 × 768	0.05	0.036	0.019	0.0154	0.01
0.98 MB	1200 × 800	0.053	0.036	0.024	0.016	0.01
1.11 MB	1200 × 800	0.054	0.037	0.025	0.018	0.011
1.17 MB	1024 × 768	0.04	0.028	0.017	0.012	0.008
1.25 MB	695 × 663	0.03	0.026	0.019	0.016	0.011
2.51 MB	3729 × 2875	0.591	0.378	0.201	0.126	0.078
3.62 MB	1355 × 1255	0.091	0.065	0.032	0.022	0.017
17.7 MB	5528 × 5350	1.895	1.212	0.662	0.626	0.591
19.9 MB	2640 × 2640	0.436	0.273	0.114	0.128	0.119
50.7 MB	5490 × 5442	1.973	1.487	0.857	0.642	0.407
77.2 MB	4656 × 5793	1.717	1.2804	0.943	0.732	0.489
55.9 MB	9052 × 4965	2.4424	1.555	1.038	0.888	0.611
64.0 MB	4000 × 4000	0.416	0.288	0.278	0.288	0.272
175 MB	5990 × 7686	1.625	0.999	0.85	0.828	0.813
176 MB	6023 × 7693	2.212	1.126	0.877	0.835	0.807

Table 2.2 Comparison results of serial and parallel time for cluster 4 for different cores 2, 4, 8, and 12

4 cores, cluster 4 time(ms)		Serial	Parallel			
Size	Resolution	1	2	4	8	12
650 KB	1024 × 768	0.077	0.048	0.037	0.024	0.014
0.98 MB	1200 × 800	0.068	0.049	0.027	0.093	0.018
1.11 MB	1200 × 800	0.061	0.042	0.026	0.026	0.012
1.17 MB	1024 × 768	0.057	0.036	0.027	0.013	0.014
1.25 MB	695 × 663	0.047	0.032	0.026	0.011	0.011
2.51 MB	3729 × 2875	0.712	0.494	0.276	0.146	0.086
3.62 MB	1355 × 1255	0.118	0.0895	0.044	0.035	0.012
17.7 MB	5528 × 5350	2.406	1.579	0.812	0.693	0.437
19.9 MB	2640 × 2640	0.585	0.338	0.202	0.187	0.116
50.7 MB	5490 × 5442	2.568	1.929	1.257	0.813	0.547
77.2 MB	4656 × 5793	2.285	1.588	1.215	0.856	0.658
55.9 MB	9052 × 4965	3.095	2.136	1.397	1.211	1.054
64.0 MB	4000 × 4000	0.787	0.452	0.382	0.373	0.371
175 MB	5990 × 7686	1.689	1.237	1.114	1.082	1.021
176 MB	6023 × 7693	1.663	1.384	1.137	1.125	1.054

Table 2.3 Comparison results of serial and parallel time for cluster 8 for different cores 2, 4, 8, and 12

Cluster 8, time in ms

Size	Resolution	Serial	Parallel			
		1	2	4	8	12
650 KB	1024 × 768	0.092	0.071	0.045	0.032	0.028
0.98 MB	1200 × 800	0.094	0.029	0.038	0.029	0.013
1.11 MB	1200 × 800	0.092	0.065	0.037	0.023	0.019
1.17 MB	1024 × 768	0.063	0.056	0.023	0.016	0.018
1.25 MB	695 × 663	0.054	0.041	0.021	0.026	0.016
2.51 MB	3729 × 2875	0.935	0.686	0.359	0.256	0.134
3.62 MB	1355 × 1255	0.152	0.112	0.063	0.045	0.024
17.7 MB	5528 × 5350	2.951	1.958	1.165	0.865	0.651
19.9 MB	2640 × 2640	0.767	0.427	0.208	0.186	0.118
50.7 MB	5490 × 5442	3.261	2.382	1.539	1.049	0.782
77.2 MB	4656 × 5793	4.466	2.516	1.246	0.906	0.747
55.9 MB	9052 × 4965	5.152	3.775	2.518	1.558	1.164
64.0 MB	4000 × 4000	0.797	0.616	0.546	0.542	0.527
175 MB	5990 × 7686	2.303	1.799	1.592	1.58	1.536
176 MB	6023 × 7693	2.521	2.172	1.66	1.656	1.638

Table 2.4 Comparison results of serial and parallel time for cluster 12 for different cores 2, 4, 8, and 12

Cluster 12, time(ms)

Size	Resolution	Serial	Parallel			
		1	2	4	8	12
650 KB	1024 × 768	0.118	0.096	0.063	0.042	0.028
0.98 MB	1200 × 800	0.113	0.084	0.047	0.039	0.026
1.11 MB	1200 × 800	0.118	0.082	0.045	0.034	0.024
1.17 MB	1024 × 768	0.086	0.065	0.032	0.027	0.016
1.25 MB	695 × 663	0.062	0.056	0.033	0.028	0.018
2.51 MB	3729 × 2875	1.283	0.874	0.456	0.272	0.134
3.62 MB	1355 × 1255	0.199	0.145	0.087	0.041	0.033
17.7 MB	5528 × 5350	4.538	2.618	1.387	0.995	0.767
19.9 MB	2640 × 2640	0.861	0.518	0.379	0.296	0.146
50.7 MB	5490 × 5442	4.938	3.366	2.149	1.368	0.921
77.2 MB	4656 × 5793	5.415	3.734	1.772	1.133	0.848
55.9 MB	9052 × 4965	6.186	4.773	3.431	2.613	1.797
64.0 MB	4000 × 4000	1.042	0.782	0.72	0.717	0.705
175 MB	5990 × 7686	3.262	2.501	2.045	2.031	2.016
176 MB	6023 × 7693	3.292	2.522	2.177	2.155	2.062

Table 2.5 Calculation of speedup for cluster 2, 2 cores

Cluster 2, 2 cores, time (ms)			
Data size	Serial	Parallel	Speedup
695 × 663	0.03	0.026	1.153
1200 × 800	0.054	0.037	1.459
1355 × 1255	0.091	0.065	1.4
2640 × 2640	0.436	0.273	1.597
3729 × 2875	0.591	0.378	1.563
4656 × 5793	1.717	1.2804	1.34
5490 × 5442	1.973	1.487	1.326
6023 × 7693	2.212	1.126	1.964
9052 × 4965	2.4424	1.555	1.57
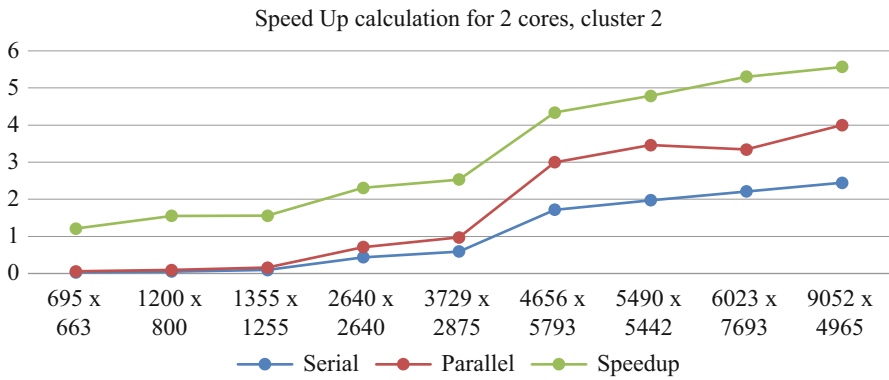
Fig. 2.14 Speedup for cluster 2, 2 cores

Table 2.6 Calculation of Efficiency for cluster 2 with varying number of cores

Efficiency for cluster 2			
Data Size	Cores	Speedup	Efficiency
5990 × 7686	2	1.626	81.3
5990 × 7687	4	1.911	47.79
5990 × 7688	8	1.962	24.53
5990 × 7689	12	1.998	16.656

Speed and efficiency calculation for cluster 2 with different number of cores is shown in Table 2.6, considering the pixel dimension of 5990 × 7686.

Speedup calculation with 2 cores for cluster 4 with different pixel dimension is shown in Table 2.7 with its graphical representation shown in Fig. 2.15.

Speed and efficiency calculation for cluster 4 with different number of cores is shown in Table 2.8 with the same pixel dimension of 5990 × 7686 and is graphically represented as in Fig. 2.16.

Table 2.7 Calculation of Speedup for cluster 4, 2 cores

Cluster 4, 2 cores, time(ms)			
Data size	Serial	Parallel	Speedup
695 × 663	0.047	0.032	1.468
1200 × 800	0.061	0.042	1.452
1355 × 1255	0.118	0.0895	1.318
2640 × 2640	0.585	0.338	1.73
3729 × 2875	0.712	0.494	1.441
4656 × 5793	2.285	1.588	1.439
5490 × 5442	2.568	1.929	1.331
6023 × 7693	1.663	1.384	1.201
9052 × 4965	3.095	2.136	1.448

Fig. 2.15 Speedup for cluster 4, 2 cores

Table 2.8 Calculation of Efficiency for cluster 4 with varying number of cores

Efficiency for cluster 4			
Data Size	Cores	Speedup	Efficiency
5990 × 7686	2	1.365	68.27
5990 × 7687	4	1.516	37.9
5990 × 7688	8	1.56	19.51
5990 × 7689	12	1.654	13.78

Speedup calculation with 2 number of workers for cluster 8 is shown in Table 2.9 with its graphical representation shown in Fig. 2.17 with different data size.

Cluster 4 Efficiency for cluster 4
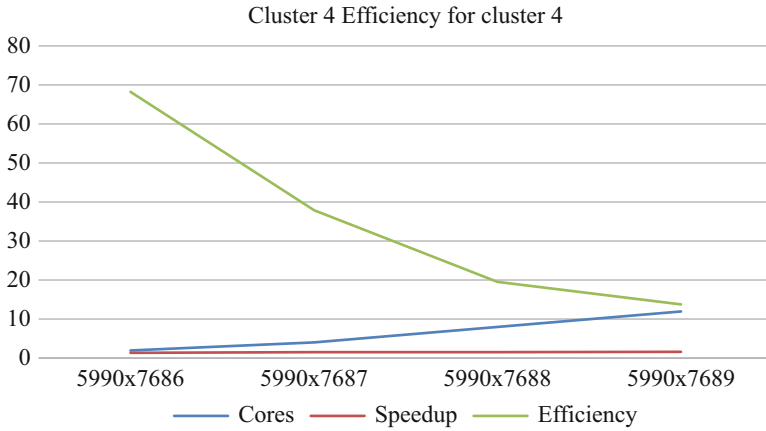
Fig. 2.16 Speedup for cluster 4, 2 cores

Table 2.9 Calculation of speedup for cluster 8, 2 cores

Cluster 8, 2 cores, time(ms)			
Data size	Serial	Parallel	Speedup
695×663	0.054	0.041	1.317
1200×800	0.092	0.065	1.415
1355×1255	0.152	0.112	1.357
2640×2640	0.767	0.427	1.796
3729×2875	0.935	0.686	1.362
4656×5793	4.466	2.516	1.775
5490×5442	3.261	2.382	1.369
6023×7693	2.521	2.172	1.16
9052×4965	5.152	3.775	1.364

Speed Up calculation for cluster 8
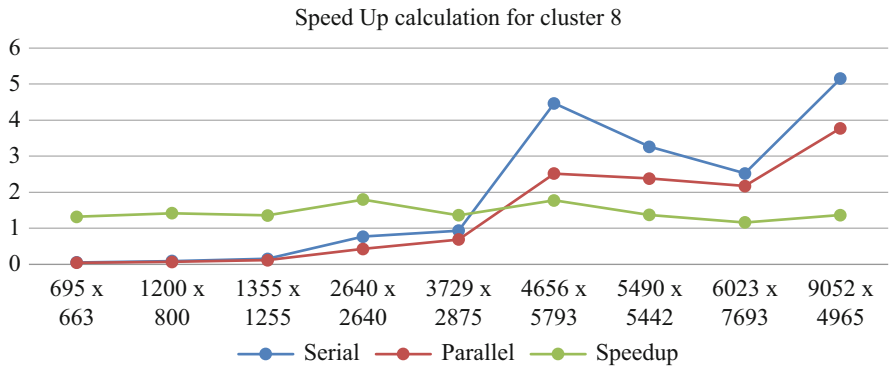
Fig. 2.17 Speedup for cluster 8, 2 cores

Fig. 2.18 Efficiency for cluster 8 with distinct cores

Table 2.10 Calculation of efficiency for cluster 8 with varying number of cores

Efficiency for cluster 8			
Data size	Cores	Speedup	Efficiency
5990 × 7686	2	1.28	64
5990 × 7687	4	1.446	36.16
5990 × 7688	8	1.457	18.21
5990 × 7689	12	1.499	12.49

Table 2.11 Calculation of speedup for cluster 12, 4 cores

Cluster 12, 4 cores, time(ms)			
Data size	Serial	Parallel	Speedup
695 × 663	0.062	0.033	1.878
1200 × 800	0.118	0.045	2.622
1355 × 1255	0.199	0.087	2.287
2640 × 2640	0.861	0.379	2.271
3729 × 2875	1.283	0.456	2.813
4656 × 5793	5.415	1.772	3.055
5490 × 5442	4.938	2.149	2.297
6023 × 7693	3.292	2.177	1.512
9052 × 4965	6.186	3.431	1.802

From the above Fig. 2.17, it is observed that there is reduction in parallel computation for 2 cores, cluster 8 compared to serial execution.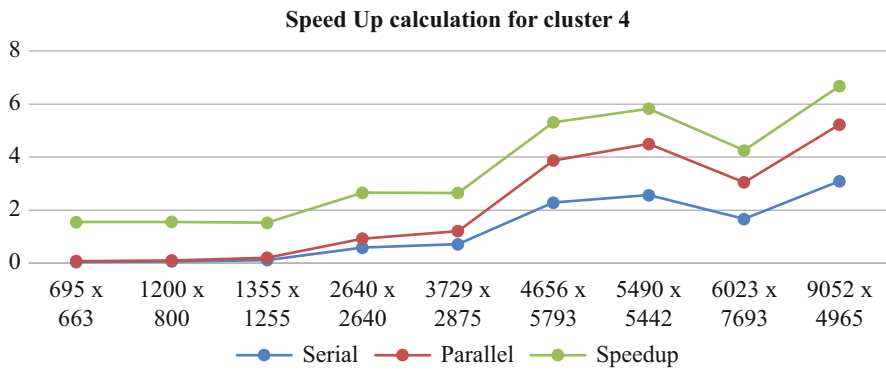 Speed and efficiency calculation for cluster 4 with different number of cores is shown in Table 2.8 with the same pixel dimension of 5990 × 7686 and is graphically represented as in Fig. 2.18 (Table 2.10).

Speedup calculation with 4 cores for cluster 12 is shown in Table 2.11 with its graphical representation shown in Fig. 2.19.

It can be observed from Fig. 2.19 that parallel time is reduced compared to serial execution for cluster 12 with the consideration of 4 cores.
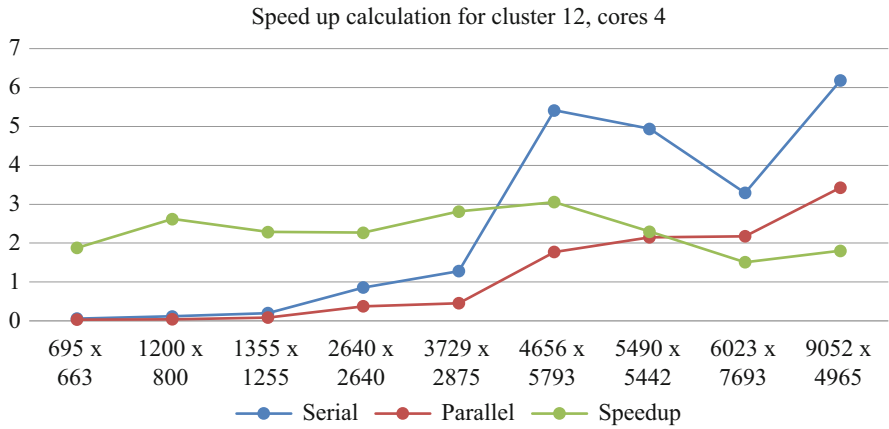
Fig. 2.19 Speedup for cluster 12, 4 cores

Table 2.12 Calculation of Speedup for cluster 12, 8 cores

Cluster 12, 8 cores, time(ms)			
Data size	Serial	Parallel	Speedup
695 × 663	0.062	0.028	2.214
1200 × 800	0.118	0.034	3.47
1355 × 1255	0.199	0.041	4.853
2640 × 2640	0.861	0.296	2.908
3729 × 2875	1.283	0.272	4.716
4656 × 5793	5.415	1.133	4.779
5490 × 5442	4.938	1.368	3.609
6023 × 7693	3.292	2.155	1.527
9052 × 4965	6.186	2.613	2.367

Table 2.13 Calculation of speedup for cluster 12, 12 cores

Cluster 12, 12 cores, time(ms)			
Data size	Serial	Parallel	Speedup
695 × 663	0.062	0.018	3.444
1200 × 800	0.118	0.024	4.916
1355 × 1255	0.199	0.033	6.03
2640 × 2640	0.861	0.146	2.069
3729 × 2875	1.283	0.134	9.574
4656 × 5793	5.415	0.848	6.385
5490 × 5442	4.938	0.921	5.361
6023 × 7693	3.292	2.062	1.596
9052 × 4965	6.186	1.797	3.442

Tables 2.12 and 2.13 illustrate the speedup calculation with 8 and 12 cores for cluster 12.

Following Fig. 2.20 is the graphical representation of speedup calculation for cluster 12 with 12 cores.

Speed Up calculation for cluster 12, cores 12
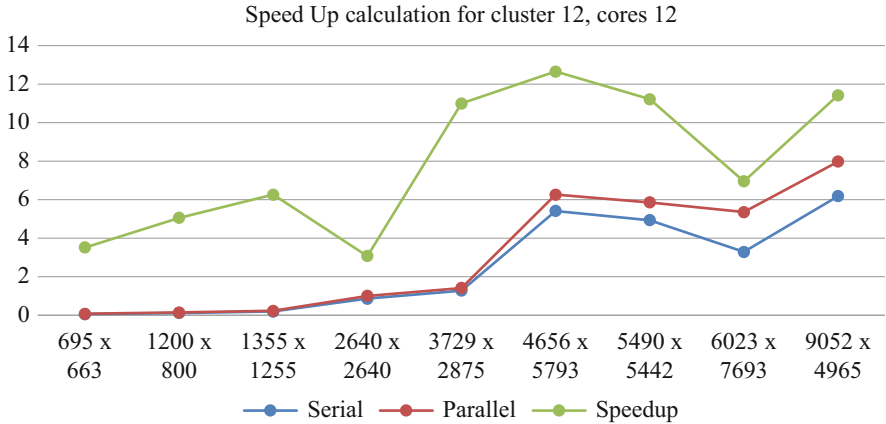
Fig. 2.20 Speedup for cluster 12, 12 cores

Table 2.14 Calculation of efficiency for cluster 12 with varying number of cores

Efficiency for cluster 12			
Data size	Cores	Speedup	Efficiency
5990 × 7686	2	1.304	65.21
5990 × 7687	4	1.595	39.87
5990 × 7688	8	1.606	20.07
5990 × 7689	12	1.618	13.48

Cluster 12 Efficiency
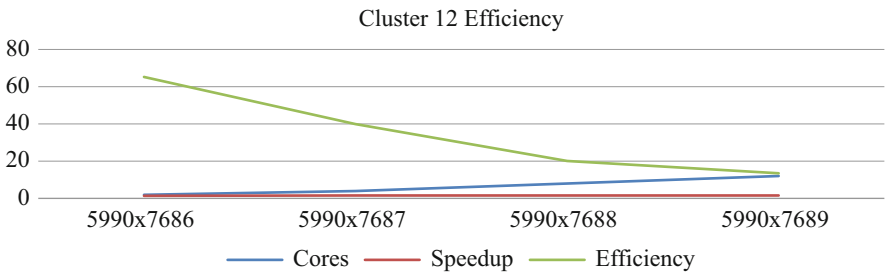
Fig. 2.21 Cluster 12 efficiency with distinct number of cores

Speed and efficiency calculation for cluster 12 with different number of cores is shown in Table 2.14, considering the same pixel dimension of 5990 × 7686, and is graphically represented as in Fig. 2.21.

It is observed from the above results of testing that speedup increases and efficiency decreases with an increase in number of workers for distinct block processing. There is reduction in execution time as the divided image of distinct blocks is processed independently by different workers in MATLAB at a time. Speedup is serially divided by parallel time. Efficiency is 100*speedup/number of workers. Efficiency is inversely proportional to the number of workers. As the

number of workers for processing increases, efficiency decreases as the work is divided among the cores/workers. Hence, the proposed work is more efficient in implementing parallel block processing rather than sequential.

2.5 Conclusion and Future Scope

Parallel processing for multispectral image clustering by implementing an unsupervised k-means algorithms using distributed programming model is focused in this paper. SPMD-based data parallelism is exhibited for parallel block processing. Multispectral images are segmented into four clusters of four features showing distinct content in satellite images. More clusters have more computational time. It is evident that an implementation of proposed system seems to be more efficient, leading to reduction in execution time which is obtained from the results of experimentation. Performance of parallel block processing has been experimentally and theoretically studied. Speedup increases with an increase in size of an image. From the experimentation, it is examined that processing time decreases with an increase in number of workers for clustering an image. This presented proposed approach can be applied for image processing applications.

Further proposed approach can be applied for clustering and classification of hyperspectral satellite images using coprocessor and graphical processing unit.

References

1. Anji Reddy M (2015) Remote Sensing and Geographical Information Systems, 3rd Edn. BS Publications
2. Lillerand K (1994) Remote Sensing and image interpretation, 2nd Edn
3. Parallel Matlab Guide (2015)
4. Bhimani J, Leeser M, Mi N (2015) Accelerating K-Means clustering with Parallel implementations and GPU computing. High Performance Extreme Computing Conference IEEE
5. Wang X (2007) Mirriam Leeser: K-means Clustering for Multispectral Images using Floating-Point Divide. 15th Annual IEEE Symposium on Field-Programmable Custom Computing Machines FCCM, pp 151–162
6. Srivastava A, Bouarfa A (1990) Image segmentation by a parallel, non-parametric histogram based clustering. J Pattern Recognit 23(9):961–973
7. Liu J, Feld D, Xue Y, Garcke J, Soddemann T, Pan P (2015) An efficient geosciences workflow on multi-core processors and GPUs: a case study for aerosol optical depth retrieval from MODIS satellite data. Int J Digit Earth:787–765
8. Sadykhov RK, Dorogush A (2007) Multispectral satellite images processing for forests and wetland regions monitoring using parallel MPI Implementation, Proceedings of Envisat Symposium

9. Bonnin P, Maurette C, Hoelzener-Douarin B, Pissaloux E (1995) A parallel implementation on CM5 of multispectral cooperative segmentation. Proceedings of 1st International Conference on Algorithms and Architectures for Parallel Processing

10. Zghidi H, Walczak M, Świtoński A (2014) Multispectral image segmentation using parallel mean shift and CUDA technology. AIP Conference proceedings

11. Bonnin PJ, Maurette C, Hoeltzener-Douarin B, Pissaloux EE (1995) Parallel cooperative segmentation method for multispectral images. Proceedings of SPIE Symposium on OE/Aerospace Sensing and Dual use Photonics

12. Phyo TZ, Khaing AS, Tun HM (2015) Classification of cluster area for satellite images. Int J Sci Technol Res 4

13. Rachmawan IEW, Barakbah AR, Harsono T (2015) Multiband Satellite Image Clustering using K-means optimization with Reinforcement Programming. 4th Indonesian-Japanese Conference on knowledge creation and Intelligent computing

14. Bräunl T (2001) Tutorial in data parallel image processing. Aust J Intell Inf Process Syst 6:164–174

15. Braunl T, Feyrer S, Wolfgang R, Reinchardt M (2001) Parallel image processing. Springer International Edition

16. USGS, Landsat Collection 1 Level 1, Version 1.0(2017)

17. USDA, National Agriculture Imagery Program NAIP (2015)

18. Earth Explorer http://earthexplorer.usgs.gov/

Chapter 3
Classification of Field-Level Crop Types with a Time Series Satellite Data Using Deep Neural Network

J. Jayanth, V. S. Shalini, T. Ashok Kumar, and Shivaprakash Koliwad

Abstract Crop-type classification has been relied upon on only spectral/spatial features. It does not provide the in-season information for researchers and decision makers for both practical and scientific purposes. While satellite images have desirable spectral and spatial information for classification, the ability to extract temporal information in satellite data remains a challenge due to revisiting frequency and gaps in the time period of capturing the data. To circumvent this challenge and generate more accurate results for an in-season crop-type classification, we have used Rectified Linear Unit (RLU) approach based on the concept of deep neural networks for intelligent and scalable computation of the classification process. The work was carried out on Nanjangud Taluk located in Mysuru District, Karnataka state on a Landsat data (multi-temporal scene) from 2010 to 2015. The results indicate that RLU shows an improvement of 5% to 15% for overall classification accuracy at 3 classes over the traditional against support vector machine. In comparison with KSRSC data set, this study reveals an accuracy of 85% for classifying rice and banana with an improvement of 10% over KSRCS crop-filed data.

Keywords Spectral · Temporal · Landsat · RLU · Rice · Banana

J. Jayanth (✉)
Department of Electronics and Communication Engineering, GSSS Institute of Engineering & Technology for Women, Mysore, Karnataka, India

V. S. Shalini
Department of Electronics and Communication Engineering, ATME College of Engineering, Mysore, Karnataka, India

T. Ashok Kumar
Sri Dharmasthala Manjunatheshwara Institute of Technology, Ujire, Karnataka, India

S. Koliwad
Department of Electronics & Communication Engineering, Malnad College of Engineering, Hassan, Karnataka, India

3.1 Introduction

Classifying different crop types in a high-resolution remote sensing data remains a practical and scientific challenge. Currently, the KSRSA unit has no in-season crop-type data set for accurate and timely data set for crop classification to provide a clear estimation, for monitoring the crops and also for decision-making applications in both public and private structure such as insurance for crop, commodity markets, etc. [1, 4].

In the current scenario, crop-type classification using remote-sensed data is done on the basis of extracting the spectral features on a single data during the crop-growing season which is based on the distinct spectral features in land cover data. Drawbacks of spectral feature crop-type classification are as follows [5–8]:

- Land use/land cover data may have a similar/distinct land cover features where these spectral features are used for classification.
- During the crop-growing season, the spectral information of the crop can have similarity with some other land cover features, which leads to misclassification of data.
- Misclassification of the data occurs when there are a similarities between the crop and natural vegetation like grass, trees, etc.

To overcome the above said drawbacks, researchers are utilising both temporal and spectral information to improve the classification accuracy of the crop-type data [5–7]. For example, in Nanjangud area, banana (Rasabale) is famous. In this area, grass usually starts during the spring season. These types of temporal features and time series information can improve the classification accuracy.

Karnataka state remote sensing agency (KSRSA) unit generates information of the field on the time series data which contains the dividing line for all the agricultural fields which is about $16_*16+/-20_*20$ satellite data in pixels [8]. Landsat data contains a weakness of lower temporal resolution, but it has also the advantage for in-season crop-type classification where Landsat data provides data from multiple growing season from temporal data for training and testing purposes which can increase the accuracy of an algorithm when applied with the other set of data of same year.

Artificial intelligence techniques have been a subject of study for different-type applications in image processing and natural language processing [2]. Currently, deep learning techniques such as deep neural network, deep belief network, convolutional neural network, recurrent neural network and restricted Boltzmann machine have been used for remote-sensed data classification. Traditionally, architecture of deep learning system is able to extract the abstract-level features from the spectral level from lower-level primitives to upper-level features for classification/regression. For remote sensing data classification, Chen et al. used auto-encoder set for feature selection and fed into support vector machine (SVM) for the classification of remote-sensed data and achieved an accuracy of 6% over SVM classifier [3]. Recently, LeCun et al. applied a multiscale sparse feature for the

classification of Spain data and show better classification accuracy over a back propagation optimisation technique [3].

An introduction to DNNs is briefly given in Sect. 3.2. The network setup, experimental results and comparison with spectral and combination of spectral and temporal are shown in Sect. 3.3. Finally, Sect. 3.4 concludes this paper.

3.2   Data and Methodology

3.2.1   Study Area

Nanjangud Taluk, Mysuru District, Karnataka state has a lat lon of 15.12°N 76.68°E 657 metres and is located in the south-western part as shown in Fig. 3.1. Rice and banana (Rasbale) are the famous crops in this study area and grown in farmland within 62% of this area. As per the field identified by the KSRSA unit, farmers grow only one crop per season. In fact, only banana is grown with a mixture of another plant. Thus, we need the field-level information for crop-type classification.

3.2.2   Data

This data set is acquired by the Landsat. Image dimension of the study area is 674×631 pixels in MS data covering the Nanjangud Taluk campus as shown in Fig. 3.1. Landsat 7 comprises of 6 spectral channels for 2010 to 2015 as shown
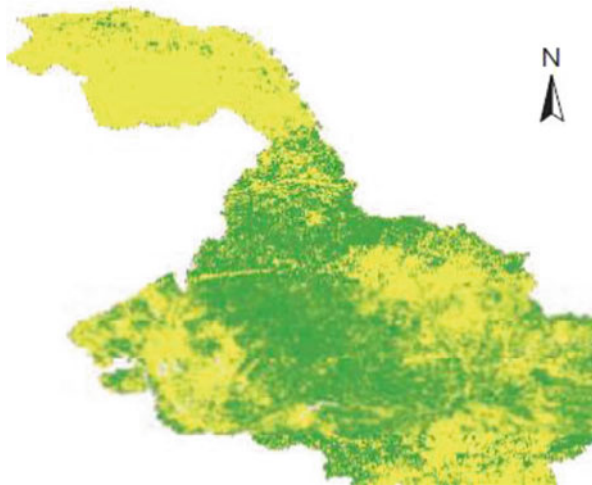
Fig. 3.1 Study area of Nanjangud Taluk

Table 3.1 Landsat data channel used

Year	Landsat 7 channel
2010	43
2011	45
2012	44
2013	44
2014	45
2015	44

Table 3.2 Detailed information of the six spectral bands from Landsat 7

Band	Name	Wavelength (micrometre)
1	Blue	0.44–0.54
2	Green	0.51–0.60
3	Red	0.63–0.69
4	NIR	0.77–0.89
5	SWIR-1	1.5–1.7
7	SWIR-2	2.1–2.3

in Table 3.1. In this data set, one noisy channel has been removed, and there are seven spectral channels like green, red, blue, near infrared, short wave infrared 1 and short wave infrared 2 investigated in this work. Table 3.2 provides information about different classes and their corresponding training and testing samples, which was used in the 2016 IRS contest conducted in NITK Surathkal Campus.

3.2.3 Support Vector Machine

Support vector machines (SVMs), relatively a new type of classifiers having their roots in the statistical learning theory, are introduced to address the problems related to classification of multispectral and hyperspectral remote-sensed (RS) data with growing popularity. Although it is a linear machine, it can solve non-linear problems by mapping of non-linear decision boundaries in the original data space into linear ones in a high-dimensional space using kernel function. This technique is said to be independent of the dimensionality of feature space, and the main idea behind this classification technique is separating surface through an optimisation procedure that finds the exemplars which form the boundaries of the classes. These exemplars are called the *support vectors*. Unlike statistical estimations, SVM performs class separation and finds the support vectors even with a small number of training samples having mean values very close to each other. Hence, the SVM does not suffer from the Hughes phenomenon (Hughes phenomenon states that for a limited number of training samples, the classification rate decreases as the data dimensionality increases) and made the SVM a very interesting classifier for multispectral image processing.

3.2.4 DNN-Based Classification Model

Deep learning usually involves neural layers learning on huge data sets, in the order of millions. However as data set is constrained, to obtain the good result, we adopted a dropout technique which plays an important role in high epoch, which measures the number of times all of the training vectors are used to update the weights. For the training of batches, all the training samples are passed through a learning algorithm in one epoch before the weights are updated. In this work, usage of ReLU activation function is justified because it is sector which is biologically plausible and practical.

DNN model uses 4 connected neural layers where the output of one layer acts an input to the next layer. The output dimension of neural layer is 4000 with a feature count of 300 for the 3030 readings from the input layer. ReLU activation functions are used for the layers with a dropout probability of 0.15. After the dropout probability, next layer assumes 4000 values for the input and 300 for output dimension. Next consecutive layers use 1200 inputs, and final output is reduced to 3 classes. Deep neural network uses activation function to transform the activation level of a unit (neuron) into an output signal; usage of smaller nodes computes the nontrivial problems of a non-linear function.

Rectified Linear Unit has become a common activation function approach for DNN; it has been used to vanish the gradient problem and speed up the learning process and leads to a non-bounded output by making zeros at negative values and identity at positive values.

- ReLU's computation does not contain any complicated math, so the model takes less time to train, learn and run.
- ReLU is zero for the negative values is to be determined with a concept called as "dying rule," because the neurons involved in the negative side do not play any role in discriminating the input; once a neuron gets negative, it's unlikely for it to recover.

As mentioned in the above points, samples can be approximated depending on the data composed of hidden nodes with an activation function; use of this activation function finds the optimal weight where the neural network uses activation function that changes the combined synthesis function of the input variables and the input information.

Average of weight has been used by the synthesis function to generate the input information, and the functions can be transferred through a composite value obtained from output stage with the value of a hidden stage, and it converts the values which are combined to the input information within a certain range.

In our work, ReLU is used as an activation function for the network output where the gradients often are needed to be truncated with a traditional bounded activation functions, representing the values which results in a dense representation in terms of learning. Thus, to train a valid recurrent network for the remote-sensed image classification, we designed the new activation function PRetanh, which has two major advantages: (1) producing a bounded output and (2) promoting sparsity adaptively.

For classification of spectral and temporal time series data for the year 2010 to 2015, reflectance sheet from KSRSA unit and Landsat data set have been taken as a reference for identifying the in-season crop-level information. For the classification of the spectral and temporal data, we have selected the 10% data from field-based information and selected 200 pixels for training for each class (rice and banana). In our work, we have considered the day of the year from 101 to 260 (from May to October) for the interpolation of different spectral bands with the temporal information in the study area, and 300 testing points were selected for rice from 2010 to 2015 data set and 380 points were selected for banana from 2010 to 2015 data set.

For in-season crop-type classification, we have built deep neural network (DNN) model using Rectified Linear Unit (RLU) activation function as f(z) = max (0, z) shown in Fig. 3.2.
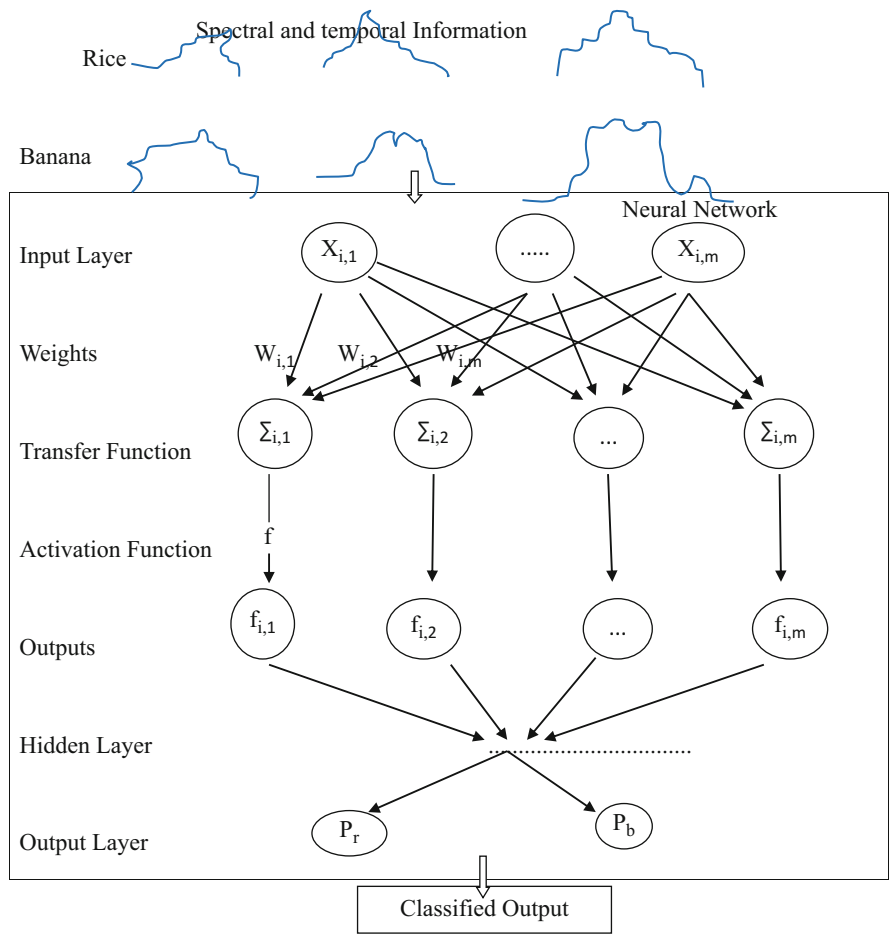
Fig. 3.2 Framework of the DNN-based crop-type classification

- Proposed schematic approach has four layers, where three are hidden layers and one is the output layer.
- For the input data (raw data has been provided with input values of Ja,b(b $\varepsilon[1,2,,\ldots n]$) where "b" consists of different spectral bands of Landsat 7 at a specific day of the year over a growing season and "a" represents the field data, n represents the number of channels in the data, and it changes according to the selected design.
- Transfer function provides the features from the radiometric data for classification.
- Activation function provides the stable performance of the selected features for classification of data.

3.2.5 Implementation

In this work, we have used multi-level filters to extract spatio-spectral and temporal information for the classification of crops like rice and banana. For processing the temporal and spectral information, each unit has been selected with three different types of regions, R1, R2 and R3, and each regions are treated with different classes like R1 represents rice crop in the mapping unit which has been overlapped by R3 units, R2 represents the regions which are covered by banana as a one-processing unit and R3 represents other classes which are ignored due to some overlap with other region R1.

In this work, ReLU is characterised with three regions, R1, R2 and R3 with a spectral correlation for band-band variability when there is a mapping between input pixel sequences and output labels.

These experiments were conducted on Windows 2008 OS with Intel I7 @ 2.40 GHz CPU, with 8GB memory and NVIDIA GeForce GTX 960 M 4GB DDR5 GPU using MATLAB 2018b software. Table 3.3 shows architecture used in the classification; the last layer dense_6 used the ReLU classifier in the experiments. Adam optimisation models were used for training with a learning rate $\alpha = 1 \times 10 - 3$, $\beta1 = 0.9$, $\beta2 = 0.999$, $\epsilon = 1 \times 10 - 8$ and no decay.

Table 3.3 Architecture of the ReLU DNN

Layer type	Output shape	Param#
Dense_3 (Dense)	(none, 512)	131,584
Dropout_4 (Dropout)	(none, 512)	0
Dense_4 (Dense)	(none, 512)	262,656
Dropout_5 (Dropout)	(none, 512)	0
Dense_5 (Dense)	(none, 512)	262,656
Dropout_6 (Dropout)	(none, 512)	0
Dense_5 (Dense)	(none, 10)	5130

3.3   Result Analysis

For achieving the overall accuracy, firstly, the data have been trained with its tuning parameters through the selected features based on the data collected from the training set, and then this data set has been applied for testing data set. The confusion matrix, overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and Kappa coefficients (Kappa) are used to evaluate the performance. The results are shown in Fig. 3.3 and Fig. 3.4. The results are tabulated in Tables 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14 and 3.15 with validation points of 2245 for 66.82Ha and training of 42.04Ha. To check the effectiveness of the proposed DNN using RLU, is compared with the vector-based classification approach known as support vector machine. SVM kernel with an RBF kernel has been used using libsvm package 2.

- For the year 2010, RLU shows an overall accuracy of 91.22% and SVM of 80.12% for the month of March.
- For the year 2010, RLU shows an overall accuracy of 89.22% and SVM of 79.12% for the month of June.
- For the year 2011, RLU shows an overall accuracy of 92.22% and SVM of 82.12% for the month of March.
- For the year 2011, RLU shows an overall accuracy of 90.44% and SVM of 79.12% for the month of June.
- For the year 2012, RLU shows an overall accuracy of 92.16% and SVM of 81.96% for the month of March.
- For the year 2012, RLU shows an overall accuracy of 89.22% and SVM of 79.12% for the month of June.
- For the year 2013, RLU shows an overall accuracy of 93.16% and SVM of 83.96% for the month of March.
- For the year 2013, RLU shows an overall accuracy of 89.92% and SVM of 78.46% for the month of June.
- For the year 2014, RLU shows an overall accuracy of 90.16% and SVM of 81.96% for the month of March.
- For the year 2014, RLU shows an overall accuracy of 93.12% and SVM of 83.12% for the month of June.
- For the year 2015, RLU shows an overall accuracy of 92.16% and SVM of 81.96% for the month of March.
- For the year 2015, RLU shows an overall accuracy of 90.22% and SVM of 80.12% for the month of June.
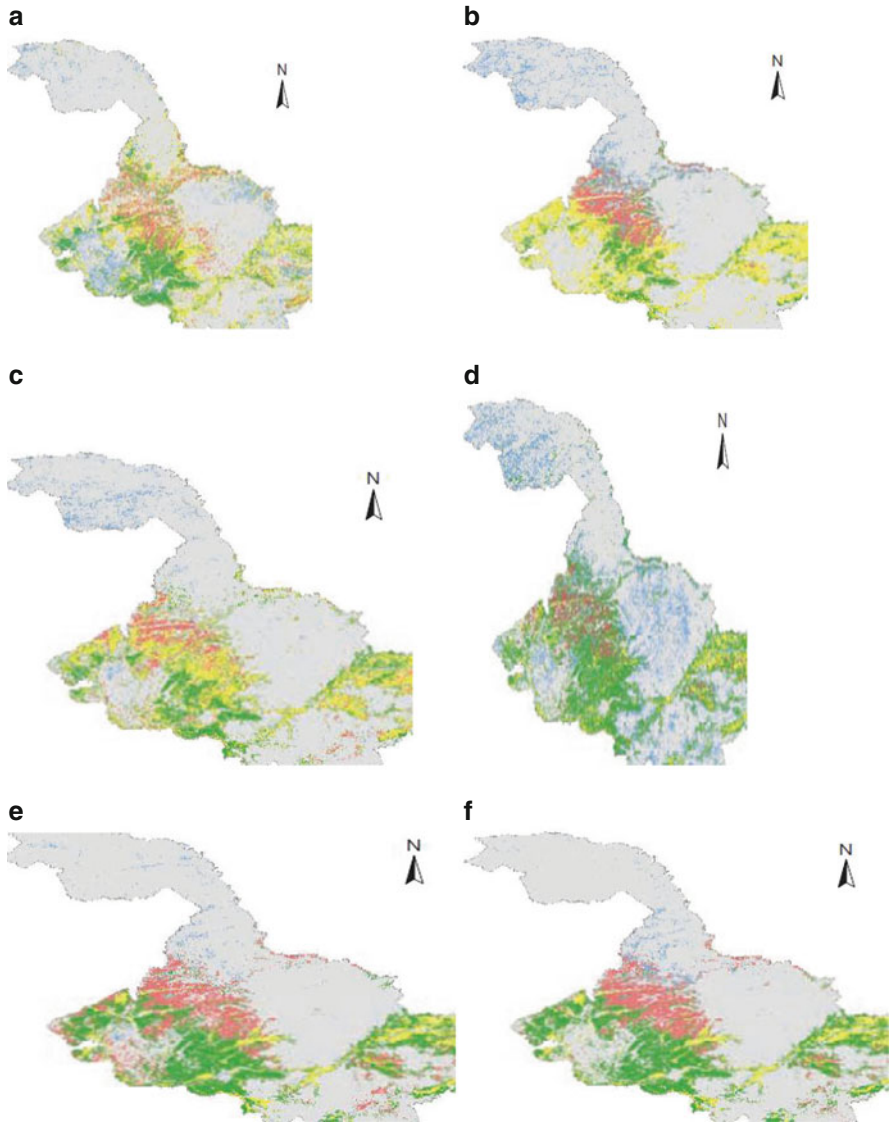
Fig. 3.3 Impacts of using different spectral-temporal for the final classification results using Recurrent Linear Unit. (**a**) 2010 data for March. (**b**) 2010 data for June. (**c**) 2011 data for March. (**d**) 2011 data for June. (**e**) 2012 data for March. (**f**) 2012 data for June. (**g**) 2013 data for March. (**h**) 2013 data for June. (**i**) 2014 data for March. (**j**) 2014 data for June. (**k**) 2015 data for March. (**l**) 2015 data for June

g

h

i

j

k

l
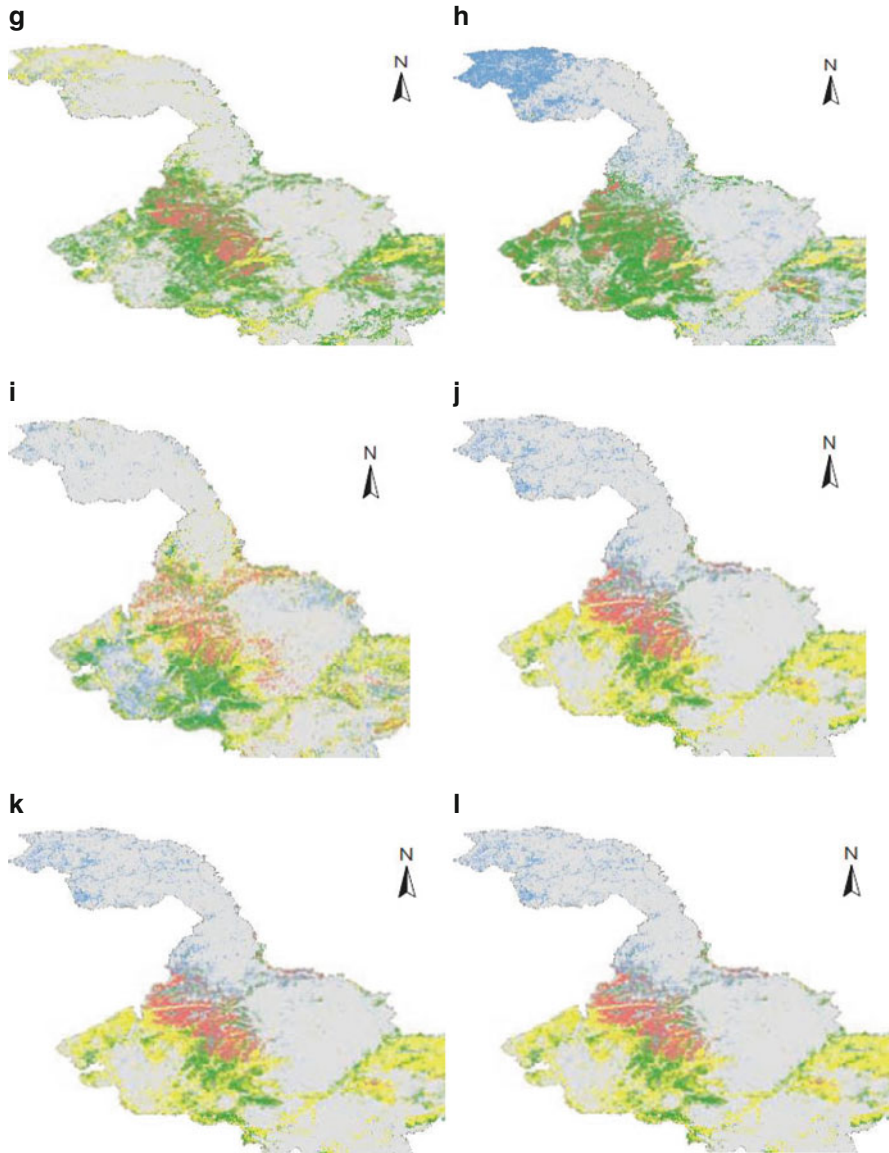
Fig. 3.3 (continued)
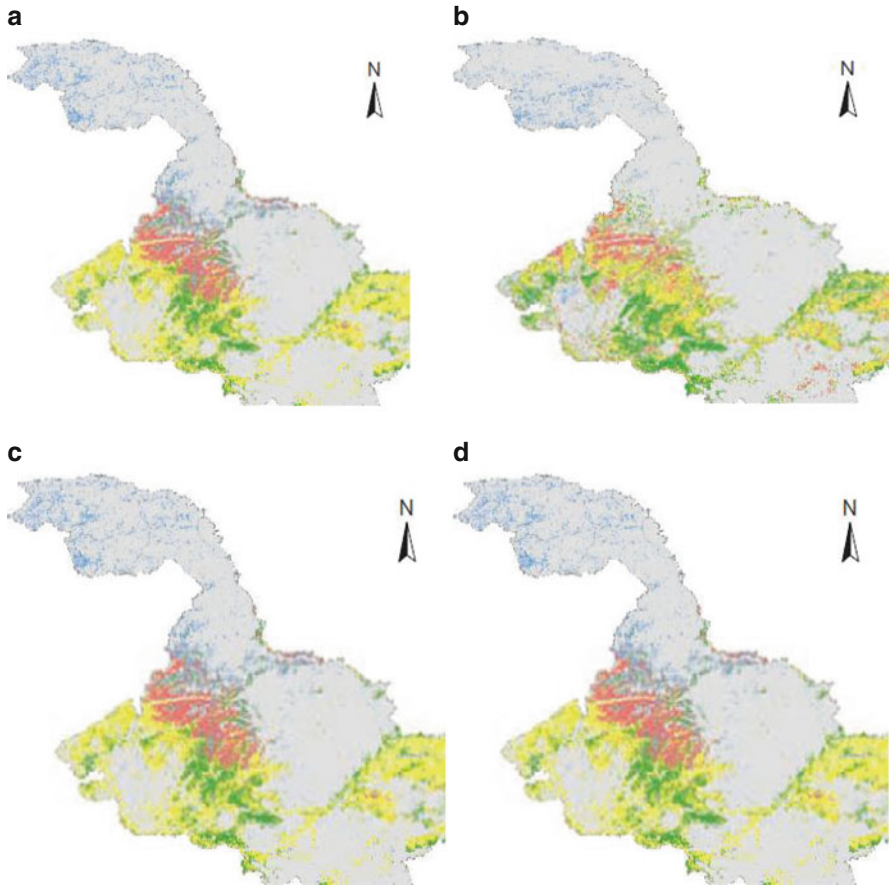
Fig. 3.4 Impacts of using different spectral-temporal for the final classification results using support vector machine. (**a**) 2010 data for March. (**b**) 2010 data for June. (**c**) 2011 data for March. (**d**) 2011 data for June. (**e**) 2012 data for March. (**f**) 2012 data for June. (**g**) 2013 data for March. (**h**) 2013 data for June.(**i**) 2014 data for March. (**j**) 2014 data for June. (**k**) 2015 data for March. (**l**) 2015 data for June
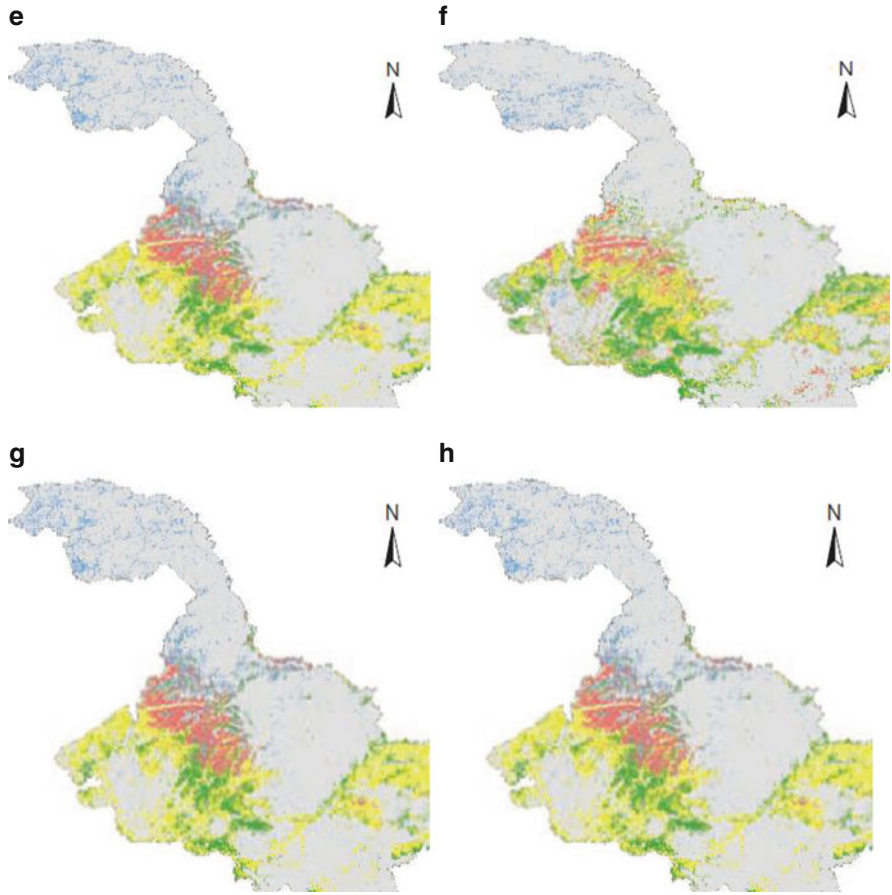
Fig. 3.4 (continued)

Fig. 3.4 (continued)

Table 3.4 Confusion matrix for 2010 March using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	95.7	91.0	Rice	78.8	25.6
Banana	97.5	89.7	Banana	50.8	61.0
Others	36.9	63.2	Others	14.7	61.3
OA (%) (RLU)	91.22		OA (%) (SVM)	80.12	

Table 3.5 Confusion matrix for 2010 June using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	95.89	90.92	Rice	63.8	42.36
Banana	97.01	92.48	Banana	30.8	61.0
Others	40.05	63.25	Others	24.7	61.3
OA (%) (RLU)	91.22		OA (%) (SVM)	79.12	

Table 3.6 Confusion matrix for 2011 March Using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	98.89	91.02	Rice	63.8	42.36
Banana	96.81	93.65	Banana	30.8	70.2
Others	39.05	64.52	Others	24.7	61.3
OA (%) (RLU)	92.22		OA (%) (SVM)	79.12	

Table 3.7 Confusion Matrix for 2011 June Using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	94.22	94.7	Rice	73.8	48.36
Banana	96.21	93.45	Banana	45.23	55.23
Others	40.25	63.25	Others	25.89	58.36
OA (%) (RLU)	90.44		OA (%) (SVM)	79.12	

Table 3.8 Confusion matrix for 2012 March using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	97.01	92.48	Rice	73.8	48.36
Banana	94.22	94.7	Banana	30.8	61.0
Others	25.89	58.36	Others	40.25	63.25
OA (%) (RLU)	92.16		OA (%) (SVM)	81.96	

Table 3.9 Confusion matrix for 2012 June using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	90.69	89.72	Rice	62.98	43.12
Banana	95.21	93.65	Banana	31.8	60.0
Others	38.05	67.3	Others	24.72	61.3
OA (%) (RLU)	89.22		OA (%) (SVM)	79.17	

Table 3.10 Confusion matrix for 2013 March using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	94.23	94.7	Rice	72.36	45.23
Banana	96.23	95.1	Banana	48.36	65.2
Others	47.2	67.1	Others	46.7	47.8
OA (%) (RLU)	93.16		OA (%) (SVM)	83.90	

Table 3.11 Confusion matrix for 2013 June using RLU and SVM

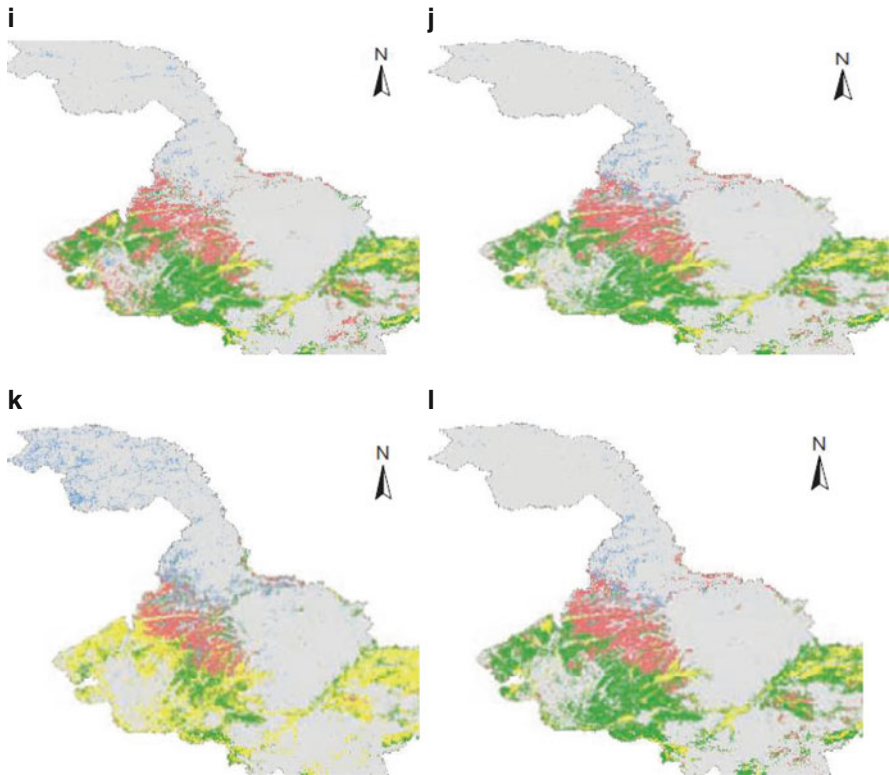| Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	91.65	88.23	Rice	63.58	42.36
Banana	92.13	95.32	Banana	87.21	22.36
Others	38.6	65.32	Others	45.23	66.23
OA (RLU)	89.92		OA (%) (SVM)	78.41	

Table 3.12 Confusion matrix for 2014 March using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	94.2	94.7	Rice	65.23	65.36
Banana	96.2	95.1	Banana	78.23	36.25
Others	48.65	47.89	Others	47.2	67.1
OA (%)(RLU)	90.16		OA (%) (SVM)	81.96	

Table 3.13 Confusion matrix for 2014 June using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	93.23	95.7	Rice	72.36	45.23
Banana	95.53	95.1	Banana	48.36	62.2
Others	47.2	67.1	Others	46.7	47.8
OA (%) (RLU)	93.12		OA (%) (SVM)	83.12	

Table 3.14 Confusion matrix for 2015 March using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	96.36	95.23	Rice	78.32	85.23
Banana	94.25	92.32	Banana	82.32	45.63
s	48.65	47.89	Others	47.2	67.1
OA (%)(RLU)	92.16		OA (%) (SVM)	81.96	

Table 3.15 Confusion matrix for 2015 June using RLU and SVM

Class	Producer's accuracy (%)	User's accuracy (%)	Class	Producer's accuracy (%)	User's accuracy (%)
Rice	95.63	95.7	Rice	78.25	56.32
Banana	92.36	98.23	Banana	48.36	62.2
Others	47.2	67.1	Others	46.7	47.8
OA (%) (RLU)	90.22		OA (%) (SVM)	80.12	

3.3.1 Time Series Profile for the Classified Data

Time series analysis has been done for the rice and banana fields in Nanjangud Taluk as shown in Fig. 3.5. As shown in Fig. 3.5, x axis represents the DOY, and y axis represents the index value which has been analysed from the selected fields for 5 years. Red color line represents rice crop; green color line represents the banana crop. Visible spectral bands like green, blue and green shown in Fig. 3.5 (a)–(c) show an overlap in the seasonal trajectories between rice and other and also with other and banana during the growing season (after DOY 150). For the NIR band,
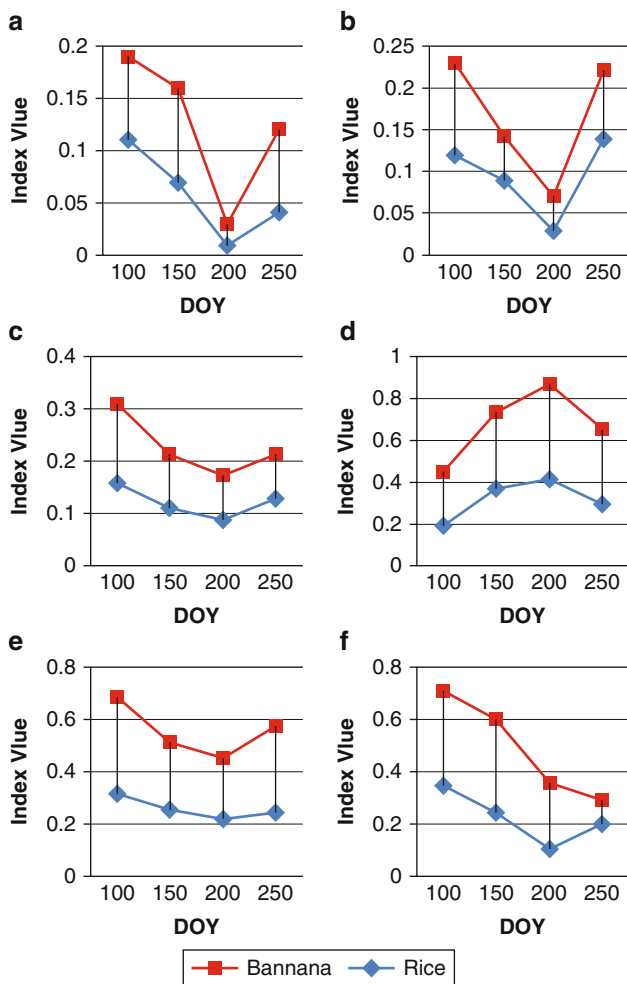
Fig. 3.5 Time series spectral band information and vegetation indices are aggregated for all the corn and soybean fields. (**a**) Blue. (**b**) Green. (**c**) Red. (**d**) NIR. (**e**) SWIR1. (**f**) SWIR2

there was a difference between the spectral bands of the rice and banana crop in all the stages. SWIR band shows a noticeable difference between rice and banana during the growing season which indicates that there is no overlap between rice and other and also with other and banana (DOY 180–210).

Overall classification accuracy (OCA) for the green band, blue band and red band shows a lesser spectral frequency of 0.44–0.62 between DOY 80 and 130 and increase in accuracy of 0.45–0.72 between DOY 130 and 180. Spectral frequency of near-infrared red band matches with the visible bands during the early stages of the crop, and it reaches peak performance of 250 during the longer period of time (DOY 210–240). Spectral frequency reaches DOY of 250 for the band SWIR for the crop rice and banana for the year 2010–2015. In this section, temporal and spectral information are merged for the entire season of crop to know the information of spectral bands for classifications of crops.

3.3.2   Discussion

For classification of crops using previous year data used in the training, the model and current year data is used for testing due to variations in spectral and temporal information. In this work, 2010 year data has been set as the initial set for classification for the month of March and June, and 2015 data set is used to predict the 2016 crop types. The results indicate that the variations in the spectral and temporal information for the classification of crop are coupled as an input for providing the necessary information for the identification of crops. Figure 11 shows that there is a general increase in the data in 2010, 2013, 2014 and 2015 due to combination of data to train and validate the crop types in 2016 and clearly shows the increase in the performance as shown in Figs. 3.3 and 3.4 for the data in 2011 and 2012 was considered as drought year for the crop information.

In literature, usage of SWIR bands has not been used in practice for identifying the crop-type classification, but SWIR bands relate the water content in the crop and provide notable difference between rice and other crops. Obtained results reveal that the inclusion of temporal information during classification has improved the accuracy by 10% when compared with KSRSA data set (Fig. 3.6). In this work, only Landsat spectral bands have been used with SWIR bands that provided the highest classification accuracy compared to other combination of data.

3.4   Conclusion

In this work, we have used Landsat 7 data, utilising a machine learning approach for in-season crop-type classification for rice and banana crop with detailed case study in Nanjangud Taluk, Mysuru, Karnataka. The classification model based on RELU of deep neural network was applied to distinguish rice and banana. Systematic
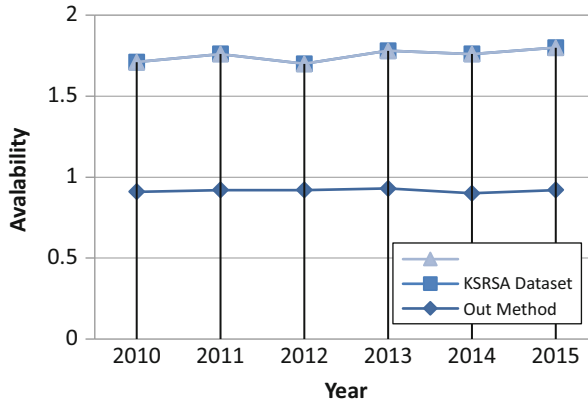
Fig. 3.6 Data comparison in improvement with KSRSA

experiments were conducted from the year 2010 to 2015 and showed an overall accuracy of 80%–88%. In addition to the crop-type classification, accuracy was verified with the time series profile for the temporal and spectral information and verified with the early stage and bands reached with a peak of DOY 250. Further improvements can be shown by using texture features from high temporal-spatial resolution data.

References

1. Bolton DK, Friedl MA (2013) Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric For Meteorol 173:74–84. https://doi.org/10.1016/j.agrformet.2013.01.007
2. Gao F, Anderson MC, Zhang X, Yang Z, Alfieri JG, Kustas WP, Mueller R, Johnson DM, Prueger JH (2017) Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sens Environ 188:9–25. https://doi.org/10.1016/j.rse.2016.11.004
3. King L, Adusei B, Stehman SV, Potapov PV, Song X-P, Krylov A, Di Bella C, Loveland TR, Johnson DM, Hansen MC (2017) A multi-resolution approach to national-scale cultivated area estimation of soybean. Remote Sens Environ 195:13–29. https://doi.org/10.1016/j.rse.2017.03.047
4. Hinton G, Deng L, Yu D, Dahl G, Mohamed A, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T, Kingsbury B (2012) Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag 29(6):82–97. https://doi.org/10.1109/MSP.2012.2205597
5. Krizhevsky A, Sutskever I, Hinton GE (2012). ImageNet Classification With Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems

6. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539
7. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003
8. Schmidt G, Jenkerson C, Masek J, Vermote E, Gao F (2013) Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description. US Geological Survey

Chapter 4
Detection of Ship from Satellite Images Using Deep Convolutional Neural Networks with Improved Median Filter

S. Iwin Thanakumar Joseph, J. Sasikala, and D. Sujitha Juliet

Abstract Detection of maritime object is of greater attention in the field of satellite image processing applications in order to ensure the security and traffic control. Even though several approaches were built in the past few years, still it requires proper revamp in the architecture to focus toward the reduction of barriers to improve the performance of ship identification or appropriate vessel detection. The inference due to cluttered scenes, clouds, and islands in between the ocean is the greater challenge during the classification of ship or vessel. In this paper, we proposed a novel ship detection method called deep neural method which works very faster and based on the concept on deep learning methodology. Experimental results provide the better accuracy, and time complexity also reduces little further when compared to the traditional method.

Keywords Deep learning · Convolutional neural network · Artificial intelligence · Satellite images · Remote sensing · Ship detection · Classification

4.1 Introduction

Nowadays, remote sensing plays a vital role in the automatic detection of ship objects in the sea surface area that exhibit the capability of locating positions of ships from remote sensing images holding the wide range of applications like:

- Traffic surveillance.
- Illegal fishing surveillance.

S. I. T. Joseph (✉) · D. S. Juliet
Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India
e-mail: iwinjoseph@karunya.edu; sujitha@karunya.edu

J. Sasikala
Department of Information Technology, Annamalai University, Chidambaram, Tamil Nadu, India

- Smuggling activities monitoring.
- Vessels salvage.
- Naval warfare.
- Ship rescue.
- Oil discharge control.
- Sea pollution monitoring, etc.

Extracting the ship candidates merely needs the use of either synthetic aperture radar (SAR) images or mostly panchromatic images [33] for its high resolution. The exact detection of ships from the complex background [23, 48] is one of the hottest issues in the field of remote sensing. Periodic monitoring of vessels from satellite images exhibits the wide visual field and incorporating large sea area and thus targeted continuous monitoring of location and movements of vessels in their territory regions. In real-time application, the complexity of performance and computation should be considered while detecting and recognizing the satellite-based object [9, 10, 11]. Remote sensing plays a key role in monitoring ships due to its

- Long operating distance.
- Wide monitoring range.

In reverse, the sea surface also gives valid information well than the ship appearance. So proper sea surface analysis is also required to avoid

- Major losses.
- False alarm in detecting ship.

Due to lack of proper illumination and different sea surface conditions. The two important motivations behind this work to clarify the major challenges of ship detection in complex sea background are

1. Consistent and effective system [35] needed to work with huge data in a limited computational capability.
2. Difficulty in extracting ship object from a complex background scenarios.

The other difficulty is the variety of shapes and appearances which again increases the complexity of detecting the ship object.

Detecting ship object using synthetic aperture radar (SAR) images [27] has been studied widely in the earlier researches because SAR images [41] are having the following advantages:

- Less influenced by weather and time.
- Obtained both day and night images.
- Less impact of meteorological conditions.
- Utilized to estimate the moving target velocities.

However, SAR images face the following problems [34] like:

- High-level speckles and noisy response.
- Low resolution.
- Long revisit cycle [46] due to limitation in number of SAR satellites.

- Difficulty in the detection of nonmetallic small targets.

These problems restrict or degrade the usage of SAR images in the research field of ship detection. However, the potential of SAR data utilization [49] in the ship detection research field [39] is enhanced [45] by the successful launching of high-resolution SAR sensors which comprehensively overcomes the difficulties (Fig. 4.1).

Some of the recent research works have been done in the detection of ships using optical satellite images. When compared to SAR images, these images provide detailed characteristics of ships which can be further utilized for classification purposes. In addition to this, the signatures of nonmetallic ships are easily identified in optical images than that of in SAR images [20, 21]. Optical satellite images [30] provide better results for identifying ship targets due to high-resolution images that attract more and more research insights in recent years. The main challenges of detecting the ship objects using optical images are:

- Clouds, sea waves.
- Cluttered scenes [28].
- Variability of ship sizes.

Also facing the key problems in increase of processing time due to the complication of discriminating ship target from background and lots of false alarms, panchromatic optical satellite images are also used in the detection of ship targets. Here, the interpretation is much easier and also able to detect the very small object which is involved in fishing activities.

The two main issues of optical spaceborne images are:

1. Pseudo targets [31] for the detection of ships due to the difference in weather conditions [16, 17] like clouds, mists, ocean waves, etc.
2. Large data quantity due to high-resolution optical spaceborne images [18, 19] leads to complexity in real-time applications [24, 32].

Due to the advancement in optical sensor technology in both spatial and temporal resolution, the area of optical remotely sensed imagery leads to the consistent growth in enormous rate.

Fig. 4.1 Complex sea background

4.1.1 Outline

Section 2 focuses on the motivations and problem statements of ship detection. Section 3.1 deals with the related works done in the field of ship detection with the survey of the data collection and preprocessing. The list of data collected has been tabulated with the website, and the test outcomes are furthermore discussed in this part. Section 3.2 discusses in detail the literature about ship segmentation, and Sect. 3.3 describes about feature extraction and classification. Section 3.4 focuses on the main parameters undertaken for the effective ship detection under various complex environment. Section 4 describes about the proposed methodology and results and discussion, and Sect. 5 concludes the survey by highlighting the paramount components of each techniques used for ship detection applications.

4.2 Motivations and Problem Statement

The main motivation behind this research work on ship detection using various satellite images in recent years is due to the challenges or difficulties faced because of the interference of weather conditions such as clouds, mists, ocean waves, complex sea backgrounds, etc. [2]. An effective and computationally fast method is much needed [3].

1. To compute huge data in terms of satellite images.
2. To predict the ships accurately from the complicated backgrounds.
3. To detect the ships that differ in shapes and appearances effectively.

Most of the unsupervised algorithms [36] that are designed to detect the ships yield the good results but faced two difficult challenges [2]:

1. Detection of ships partially or fully occluded by clouds.
2. Detection of ships near harbor or land regions.

So after the stage of ship candidate extraction, some postprocessing methods are used to discard the background regions; however, the parameters or threshold values assigned lack the flexibility or robustness of the system. The performance of the system is easily affected by the illumination changes or sea surface conditions, since shape of the ship is not the only factor [25] to detect the ship region [12]. Sea surface also plays a key role in giving valid information about ships that captured from satellite images [4].

Most of the well-established algorithms [37] or methods [15] to detect the ships in complex or various sea surfaces follow two basic steps [5]:

1. Sea detection.
2. Ship detection in sea.

The satellite images captured in coastal areas [47] or nearby harbor region also include small fishing boats with or without nets. Ship candidate selection algorithms employed in major research papers face a challenge to define the shapes for these small targets. The intensity of the ships also varies with respect to time. Sometimes, ships may appear darker or brighter than the background regions. These are few basic problem statements which motivate various researchers to implement an effective system to detect the ships accurately using various parameters and metrics in satellite images [7].

4.3   Related Works

Fen Yang et al. [1] proposed a detection algorithm based on saliency segmentation and structure LBP features in order to improve the ship detection in satellite images. This algorithm works in two steps. First, an efficient saliency segmentation framework is formed along with integration of multiple visual cues to retrieve the expected candidate plot from the complex sea surface. Shape analysis is used next to avoid the false alarms. Finally, structure LBP feature is adapted to distinguish the true ship targets. The author experimentally proves that the algorithm works well when compared to the state-of-art method especially in detection time and accuracy in detecting the target [43]. The future work concerned about to improve the performance of detection in various environments especially to identify the ship near land regions.

Zikun et al. [2] proposed a ship-rotated bounding box space for extracting the ship region from high-resolution optical satellite images with complex backgrounds. Here, first [44], author checked the possibility of accurately covering all ships by rotated bounding boxes labels. In order to work it efficiently or reduce the searching, author used closed form ship-rotated bounding box spaces. Then by utilizing two cascaded linear model along with binary linear programming, number of highly potential candidates are able to be selected by applying scores for each latent candidate. Using the abovementioned algorithm, author is able to improve the efficiency and detection rate in a real time. The future work concerns about introducing some more visual cues in order to reduce the number of candidates to improve the detection accuracy.

Zhangxia Zou and Zhenwei Shi [3] proposed a novel ship detection algorithm called SVD network to ship detection approach. This algorithm is designed based on convolutional neural network and singular value decomposition algorithm. This algorithm is specifically designed to make the system fast, robust, and structurally compact. This algorithm positively overcomes the difficulty of interferences due to clouds and strong waves. Also, it overcomes the time complexity [29] of detecting both inshore and offshore ships. Author has done experiments in spaceborne optical images of GaoFen-1 and Venezuelan remote sensing images. The future work concerns about integrating the local and global environments to improve the detection accuracy [26] and reduce the memory space.

Guang Yang et al. [4] proposed ship detection algorithm based on sea surface analysis to detect the ships from optical satellite images. This algorithm first

analyzes whether the sea surface is similar or not by using two features. Then the ship candidate region is selected based on linear function, combining pixels [13] and region characteristics. On the next step, compactness and length width ratio is adapted to detect the false alarms. This algorithm efficiently blocks the noncandidate region and assigns weights for the selected candidate function to optimize the performance of detection. This method improves the accuracy of ship detection and efficiency.

Changren Zhu et al. [5] proposed the hierarchical method of detection of ship based on texture and shape features from optical images. Here, first, the simple shape analysis is adopted to eliminate the false candidates generated normally by segmentation of images with local and global features. Next, the semisupervised classification approach is used here based on various features to distinguish ship and nonship to eliminate the false alarms. Upon normally used [42] shape and texture features, local multiple pattern feature is used to upgrade the representation ability of feature set in extracting the features. This algorithm improves the ship detection performances satisfactorily. The future work concerns in better image preprocessing and also appropriate feature extraction, detailed selection, and also hierarchical classification.

Jiexong Tang et al. [6] proposed a method based on deep neural network and extreme learning machine to detect the ship on spaceborne optical images—the traditional methods facing difficulty in balancing complexity and performance. This method uses wavelet coefficients extracted from JPEG 2000 compressed domain combined with deep neural network and extreme learning machine. Compressed domain is applied for enduring fast extraction of ship candidate region followed by deep neural network for representation of high-level feature and classification. Extreme learning machine is used in this algorithm [14] for successful feature pooling and decision-making. This method reduces the detection time and improves the detection accuracy.

Nadia Proia and Vincent Page [7] proposed the ship detection method in satellite images based on characterization of Bayesian theory. This algorithm [22] is based on Bayesian decision theory and is not necessarily to undergo preprocessing steps. The tuning of this approach is based on two parameters: One is the size of the analysis window, and next one is the threshold used to take the decision. The two parameters are fixed from the receiver operating characteristics curves that used to get from various category [40] of experimental tests. The algorithm needs major improvement before integrating this method into an operating system. Several preprocessing and postprocessing steps are done to reduce the false alarms. That steps [38] have to be reduced in order to decrease the complexity of this approach.

Zhenwei Shi et al. [8] proposed the ship detection method based on anomaly detector and local shape feature. This method presents an approach to detect ship in a coarse to fine manner. This method first separates the ship and its background by rearranging the spatial adjacent pixels into a vector and then transforming panchromatic image into a pseudo hyperspectral format. Then hyperspectral algorithm is used to get the ship candidates and then validate the original ship region out of ship candidates to generate the hypothesis. This algorithm provides robust result even in

low-contrast regions. Some issues also occurred on the other side instead of positive results. İt gives highly poor performance while identifying ships near land regions.

4.3.1   Main Parameters

The main parameters undertaken by various authors for the research work of ship detection from complex background in both the cases of inshore and offshore are as follows:

1. Detection rate = number of correctly detected ships/number of all ships.
2. False alarm rate = number of detected false alarms/size of the image.
3. Recall rate = number of real detected ships/total number of real ships.
4. Precision rate = number of real detected ships/total number of detected ships.
5. Accuracy = number of correctly detected ships/number of real ships.
6. Missing ratio = 100% - accuracy.
7. False ratio = number of falsely detected candidates/number of detected ships.
8. Error ratio = missing ratio + false ratio.

4.4   Methodology

The overall methodology of convolutional neural network along with improved median filter used in the satellite images for detecting ships is given below (Fig. 4.2).

Input   Filtering process   Conv   ReLu   Pooling   Output
Image   Image
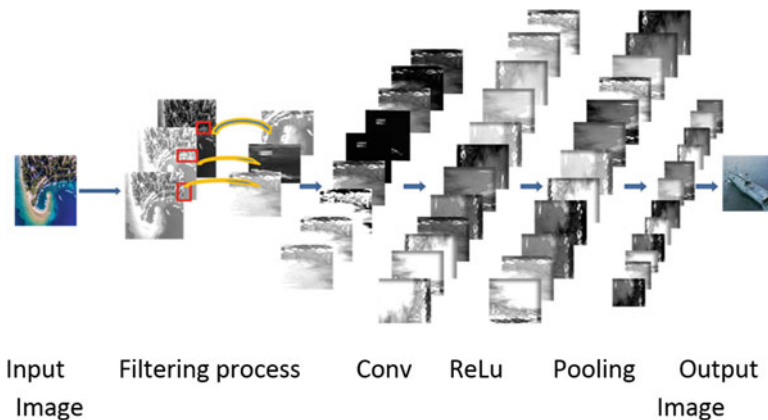
Fig. 4.2 Convolutional neural network with improved median filter

4.4.1 Improved Median Filter

The improved median filter is used in this architecture to enhance the accuracy performance of convolutional neural network in achieving the target of ship detection in optical satellite images.

Algorithm

Step 1: 3×3 size of two-dimension matrix is selected and centered to the noise-affected image of pixels $A(x,y)$.

Step 2: Arrange the pixels in the selected matrix in the ascending order. Trace the median pixel in the matrix by A_{median}, maximum pixel by A_{maxim}, and minimum pixel by A_{minim} of the sorted vector V_0. The first and the last component of the vector V_0 is A_{maxim} and A_{minim}, and the center component of the vector is A_{median}.

Step 3: If the processed one is within the range $A_{minim} < A(x,y) < A_{maxim}$, $A_{minim} > 0$ and $A_{maxim} < 255$, then it is categorized as unaffected one. Otherwise, it is considered as noise-affected one.

Step 4: If $A(x,y)$ is affected one, then the following two cases are reviewed.

Case 1: If $A_{minim} < A_{median} < A_{maxim}$ and $0 < A_{median} < 255$, then replace the affected one $A(x,y)$ with A_{median}.

Case 2: If the abovementioned condition is not satisfied, then A_{median} is a noisy one. In this category, find the variation between each pair of nearby pixels across the sorted vector V_0 and get the difference vector V_d. Analyze the maximum difference in the vector V_d and take its corresponding pixel in V_0.

Step 5: Repeat the process from step1 to step 4 until computation completed for an entire image (Fig. 4.3).

4.4.2 Convolutional Neural Network

A simple convolution neural network architecture is given in the following diagram (Fig. 4.4):

The three main processes done in this convolutional neural network architecture are convolution, max pooling, and neural network. The first layer of the figure
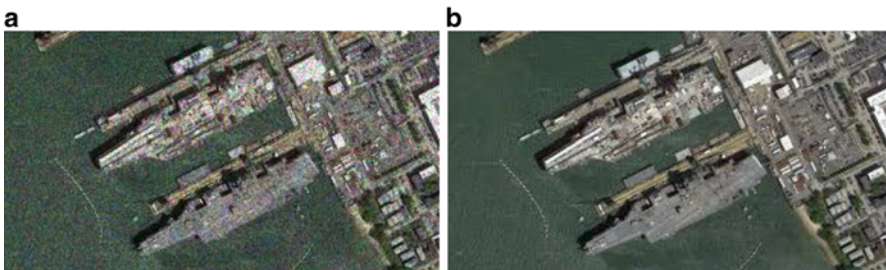
Fig. 4.3 (**a**) Input noisy image (**b**) Output of improved median filter
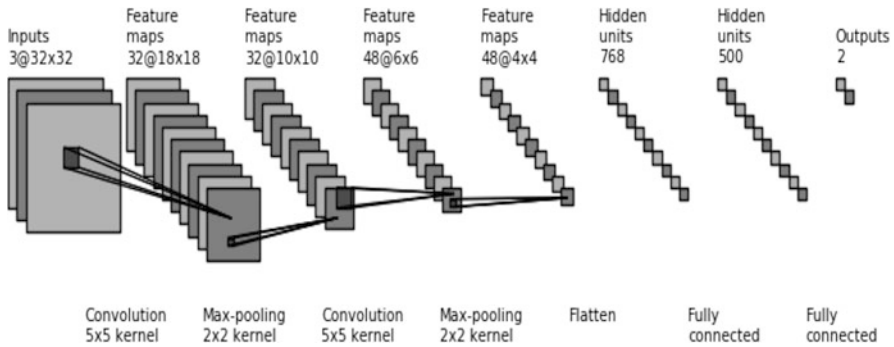
Fig. 4.4 Convolutional neural network (CNN) structure

represents the input layer where the input image is taken as it is, and the second layer represents the convolutional layer where it got several feature maps from the input image by the basic convolution step. Here, convolution does the process of feature extraction. The output of the first convolution layer acts as an input for the next convolution layer. The first layer output gives the low-level features, whereas the subsequent layer outputs give the high-level features. Pooling layer is the next layer; it gives the average or maximum of the input feature map. The following layer works in the same way as before; only the number and size of convolution varies. The output layer is fully connected, and it is the result of the classifier process in the end. The entire network is iterated or fine-tuned by two basic process: One is the feed forward, and the next is the back propagation through which the variation in the predicted output is altered in terms of assigned weight and further iteration occurs to tune the required result.

Algorithm

The convolution neural network follows the basic steps in classifying the output: convolution step, max pooling step, flatten step, and dense step. The algorithm for the deep convolution neural network after the convolution and max pooling process takes place in two basic steps:

1. Forward pass.
2. Reverse pass.

In forward pass, the first step is assign random weight to each node and then calculate the output of the first hidden layer using the following formula:

$$N = \sum W^T . x \tag{4.1}$$

where:

N – output of first hidden layer.
W – weight assigned to it.

The next step is to calculate the overall output in the feedforward step using the formula

$$O = f\left(W_n{}^T f\left(W_3{}^T f\left(W_2{}^T f\left(W_1{}^T x\right)\right)\right)\right) \qquad (4.2)$$

The cross entropy function is calculated in order to find out the difference in the output and the targeted output using the following formulae:

$$E = 1/N \sum \left[t_n \log\left(O_n\right) - \left(1 - t_n\right) \log\left(1 - O_n\right)\right] \qquad (4.3)$$

where:

N – total number of training data.
t – target value,
O – output value
n – number of neurons in the output layer.

After calculating the cross entropy function, the weights are adjusted in the reverse pass. There are several versions used to modify the weights in the reverse pass. In our methodology, we used adaptive moment estimation algorithm (ADAM) to adjust the random assigned weight in the reverse pass in order to tune the output toward the targeted one.

4.4.3 Results and Discussion

An exhaustive experimentation has been carried by applying the proposed algorithm on a vast set of varying and dynamic optical high-resolution images from the satellite imagery database. The images taken for experimentation have been investigated under three conditions: First category includes calm sea where 4200 subimages have been utilized. They are characterized by a smooth texture with no waves with a target object set at 150. The second category includes 3900 images under thick cloud cover where the sea is characterized with significant levels of swells with a target of 110 ships. Finally, the last category is the clutter condition with 1400 images and target of 40 vessels. These images are characterized by a thick cloud cover with many external disturbances. The implementation has been done in python 3.6.5 in anaconda environment running on Intel I7 2.8GHz CPU and 16GB RAM. The images are characterized by dimensions of 5000×5000 with 2.6 m resolution with ships of different categories such as coast guard ships, tankers, fishing vessels, and sailing vessels under different conditions of sea such as calm, rough, and turbulent conditions with small or medium to high-density cloud cover (Fig. 4.5).
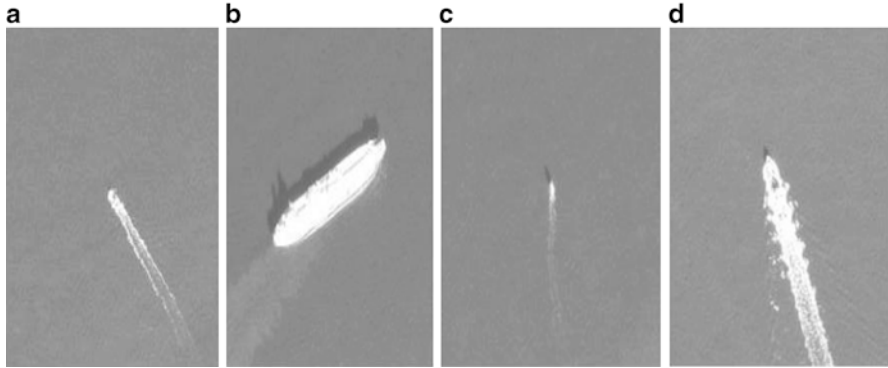
Fig. 4.5 Optical images of (**a**) fishing vessel, (**b**) tanker vessel, (**c**) sailing vessel, and (**d**) coast guard vessel

Table 4.1 Performance comparison with state-of-art method

	LBP	SVM	CNN	CNN with IMF
True positive rate	78.4	82.4	81.4	88.9
False positive rate	8.1	7.2	7.8	4.2
Accuracy	81.4	84.3	82.6	90.3

The proposed algorithm is compared with other state-of-art method such as local binary pattern (LBP), singular value decomposition (SVM), and salient mapping technique (SMT) for its performance in terms of accuracy, loss of function, and running time (Table 4.1).

$$\text{Accuracy} = (\text{correctly predicted class/Total testing class}) \times 100 \qquad (4.4)$$

$$\text{True positive rate} = TP/ (TP + FN) \qquad (4.5)$$

$$\text{False positive rate} = FP/ (FP + TN) \qquad (4.6)$$

From the abovementioned table, it is clear that the deep convolutional neural network method outperforms other ship detection algorithms in terms of its detection accuracy, false prediction, and computational time (Fig. 4.6).
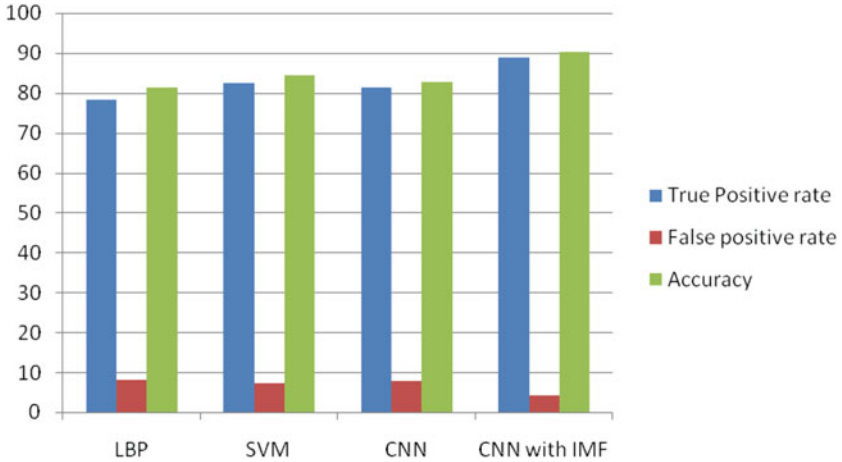
Fig. 4.6 Accuracy, true positive rate, false positive rate for LBP, SVM, CNN, CNN with IMF

4.5 Conclusion

In this work, the advantages and disadvantages of different classification techniques of ship detection are highlighted. The suitability of the techniques for various applications is also explained in the related work. Several hybrid approaches can be developed in order to increase the accuracy of ship detection system. The proposed methodology outperforms other state-of-art method in terms of accuracy and other key factors. The experimental results on real optical satellite images explain that the proposed method is not only efficient in terms of its computation but also consistent to different ocean backgrounds and with different targets in terms of size and shape when compared to the state of methods. Our future work will mainly focus on reducing the number of iterations in terms of hidden layers and finding the accuracy in that minimal number of hidden layers. This proposed work also aids in highlighting the significant contributions of satellite images for effective ship detection system.

References

1. Liu Z, Wang H, Weng H, Yang L (2016) Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds. IEEE Geosci Remote Sens Lett 13(8):1074–1078
2. Zou Z, Shi Z (Oct. 2016) Ship detection in spaceborne optical image with SVD networks. IEEE Trans Geosci Remote Sens 54(10):5832–5845
3. Yang G, Li B, Ji S, Gao F, Xu Q (Mar. 2014) Ship detection from optical satellite images based on sea surface analysis. IEEE Geosci Remote Sens Lett 11(3):641–645

4. Zhu C, Zhou H, Wang R, Guo J (2010) A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. IEEE Trans Geosci Remote Sens 48(9):3446–3456

5. Tang J, Deng C, Huang G-B, Zhao B (Mar. 2015) Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. IEEE Trans Geosci Remote Sens 53(3):1174–1185

6. Proia N, Pagé V (2010) Characterization of a Bayesian ship detection method in optical satellite images. IEEE Geosci Remote Sens Lett 7(2):226–230

7. Shi Z, Yu X, Jiang Z, Li B (Aug. 2014) Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature. IEEE Trans Geosci Remote Sens 52(8):4511–4523

8. Qi S, Ma J, Lin J, Li Y, Tian J (2015) Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images. IEEE Geosci Remote Sens Lett 12(7):1451–1455

9. Bi F, Zhu B, Gao L, Bian M (2012) A visual search inspired computational model for ship detection in optical satellite images. IEEE Geosci Remote Sens Lett 9(4):749–753

10. Shi Z, Yu X, Jiang Z, Li B (2014) Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature. IEEE Trans Geosci Remote Sens 52(8):4511–4523

11. Yang G, Li B, Ji S, Gao F, Xu Q (2014) Ship detection from optical satellite images based on sea surface analysis. IEEE Geosci Remote Sens Lett 11(3):641–645

12. Tang J, Deng C, Huang GB, Zhao B (2015) Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. IEEE Trans Geosci Remote Sens 53(3):1174–1185

13. Crisp DJ (2004) The state-of-the-art in ship detection in synthetic aperture radar imagery. Aust. Gov., Dept. Defence, Edinburgh, S. Aust., Australia, DSTO-RR-0272

14. Lure FYM, Rau Y-C (1994) Detection of ship tracks in AVHRR cloud imagery with neural networks. Proc IEEE IGARSS 3:1401–1403

15. Weiss JM, Luo R, Welch RM (1997) Automatic detection of ship tracks in satellite imagery. Proc IEEEIGARSS 1:160–162

16. Xia Y, Wan SH, Yue LH (2011) A novel algorithm for ship detection based on dynamic fusion model of multi-feature and support vector machine. In Proc IEEE 6th ICIG, pp. 521–526

17. Chen F, Yu W, Liu X, Wang K, Gong L, Lv W (2011) Graph-based ship extraction scheme for optical satellite image. In Proc IEEE IGARSS, pp. 491–494

18. Eldhuset K (1996) An automatic ship and ship wake detection system for space borne SAR images in coastal regions. IEEE Trans Geo Sci Remote Sens 34(4):1010–1019

19. Dragošvi'c MV, Vachon PW (2008) Estimation of ship radial speed by adaptive processing of RADARSAT-1 fine mode data. IEEE Geosci Remote Sens Lett 5(4):678–682, Oct

20. Li X, Chong J (2008) Processing of envis at alternating polarization data for vessel detection. IEEE Geosci Remote Sens Lett 5(2):271–275

21. Mirghasemi S, Yazdi HS, Lotfizad M (2012) A target-based color space for sea target detection. Appl Intell 36(4):960–978

22. Zhu C, Zhou H, Wang R, Guo J (Sep. 2010) A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. IEEE Trans Geosci Remote Sens 48(9):3446–3456

23. A. J. Morse and M. A. Protheroe, "Vessel classification as part of an automated vessel traffic monitoring system using SAR data," Int J Remote Sens, vol. 18, no. 13, pp. 2709–2712, Sep. 1997

24. Vachon PW, Campbell J, Bjerklund C, Dobson F, Rey M (1997) Ship detection by the RADARSAT SAR: validation of detection model predictions. Can J Remote Sens 23(1):48–59

25. Wang P, Chong J, Wang H (2000) Ship detection of the airborne SAR images. Proc IEEE IGARSS 1:348–350

26. Han ZY, Chong JS, Zhu MH (2005) Ship detection in SAR images using multi-polarimetric information. In Proc IEEEIGARSS, pp. 4729–4732

27. C. Liu, P. W. Vachon, and G. W. Geling, "Improved ship detection using polarimetric SAR data," in Proc. IEEE IGARSS, , 2004, vol. 3, pp. 1800–1803
28. Tello M, Martıacutenez CL, Mallorqui JJ (2005) A novel algorithm for ship detection in SAR imagery based on the wavelet transform. IEEE Geosci Remote Sens Lett 2(2):201–205
29. Margarit G, Mallorqui JJ, Fortuny-Guasch J, Lopez-Martinez C (2009) Exploitation of ship scattering in polarimetric SAR for an improved classification under high clutter conditions. IEEE Trans Geosci Remote Sens 47(4):1224–1235
30. Corbane C, Najman L, Pecoul E, Demagistri L, Petit M (2010) A complete processing chain for ship detection using optical satellite imagery. Int J Remote Sens 31(22):5837–5854
31. Weiss JM, Luo R, Welch RM (1997) Automatic detection of ship tracks in satellite imagery. Proc IEEE IGARSS 1:160–162
32. Tello M, Lopez-Martinez C, Mallorqui J (2005) A novel algorithm for ship detection in SAR imagery based on the wavelet transform. IEEE Geosci Remote Sens Lett 5(2):201–205
33. Lin I-I, Khoo V (1997) Computer-based algorithm for ship detection from ERS-SAR imagery. In Proc 3rd ERS Symp Space Service Environ, vol. 414, p. 1411
34. Crisp D (2004) Thestate-of-the-artinshipdetectioninsyntheticapertureradar imagery. Defence Sci Technol Org, Melbourne, Australia
35. Greidanus H, Clayton P, Indregard M, Staples G, Suzuki N, Vachoir P, Wackerman C, Tennvassas T, Mallorqui J, Kourti N, Ringrose R, Melief H (2004) Benchmarking operational SAR ship detection. Proc IEEE IGARSS, Anchorage, AK 6:4215–4218
36. Opelt A, Pinz A, Fussenegger M, Auer P (2006) Generic object recognition with boosting. IEEE Trans Pattern Anal Mach Intell 28(3):416–431
37. Oliver CJ, Blacknell D, White RG (1996) Optimum edge detection in SAR. Proc Inst Elect Eng—Radar, Sonar Navig 143(1):31–40
38. Liu C, Vachon P, Geling G (2005) Improved ship detection with airborne polarimetric SAR data. Can J Remote Sens 31(1):122–131
39. Ahonen T, Hadid A, Pietikainen M (2004) Face recognition with local binary patterns. M.S. thesis, Mach. Vis. Group, Infotech Oulu, University of Oulu, Oulu, Finland
40. Burgess DW (1993) Automatic ship detection in satellite multispectral imagery. Photogramm Eng Remote Sens 59(2):229–237
41. Corbane C, Marre F, Petit M (2008) Using SPOT-5HRG data in panchromatic mode for operational detection of small ships in tropical area. Sensors 8(5):2959–2973
42. Corbane C, Pecoul E, Demagistri L, Petit M (2008) Fully automated procedure for ship detection using optical satellite imagery. Proc SPIE-Remote Sens Inland, Coastal, Ocean Waters 7150:71500R-1–71500R-13
43. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In Proc IEEE Comput Soc Conf CVPR, pp 886–893
44. Wang Y, Liu A (2012) A hierarchical ship detection scheme for high resolution SAR images. IEEE Trans Geosci Remote Sens 50(10):4173–4184
45. Zhu CR, Wang RS (2012) Local multiple patterns based multiresolution gray-scale and rotation-invariant texture classification. Pattern Recogn 187:93–108
46. Qi S, Ma J, Tao C, Yang C, Tian J (2013) A robust directional saliencybased method for infrared small-target detection under various complex backgrounds. IEEE Geosci Remote Sens Lett 10(3):495–499
47. Wackerman CC, Friedman KS, Pichel WG, Clemente-Colon P, Li X (2001) Automatic detection of ships in RADARSAT-1 SAR imagery. Can J Remote Sens 27(5):568–577
48. Galleguillos C, Belongie S (2010) Context based object categorization: a critical survey. Comput Vis Image Understand 114(6):712–722
49. Hsu C, Chang C, Lin C (2004) A practical guide to support vector classification. [Online]. Available: http://www.csie.ntu.edu.tw/ ~cjlin/papers/guide/guide.pdf

Chapter 5
Artificial Bee Colony-Optimized Contrast Enhancement for Satellite Image Fusion

Anju Asokan and J. Anitha

Abstract Image fusion combines two or more images to a single image to extract all the necessary information from the source images. It minimizes the redundant information present in the source images. Fused images find wide applications in medical imaging, computer vision, remote sensing, change detection, and military applications. The success of the fusion technique is limited by the noise present in the source images. In order to overcome this limitations, an artificial bee colony (ABC)-optimized contrast enhancement for satellite image fusion is proposed to fuse two multitemporal satellite images. The ABC-optimized source images are given as input to the fusion stage. A hybrid contrast enhancement technique combining the histogram equalization and gamma correction techniques is used for the contrast enhancement of the source images. The contrast-enhanced images are fused using Discrete Wavelet Transform (DWT), Principle Component Analysis (PCA), and Intensity, Hue, Saturation Transform (IHS) individually. The proposed work further compares these conventional fusion techniques by computing performance measures for image fusion such as Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), entropy, Structural Similarity Index (SSIM), and Feature Similarity Index (FSIM). The experimental results show that the IHS-based image fusion technique outperforms the PCA- and DWT-based fusion techniques. Also, this method is computationally effective and simple in its implementation.

Keywords Image fusion · Remote sensing · Histogram equalization · Gamma correction · Multitemporal · PCA · IHS

A. Asokan · J. Anitha (✉)
Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India
e-mail: anithaj@karunya.edu

5.1 Introduction

Enhancement is a prominent image preprocessing technique which is applied in various fields such as medical imaging, satellite image analysis, computer vision, and so on. There can be some limitations as part of the image-capturing devices or some atmospheric interferences which degrade the quality of the image. Various local and global enhancement-based techniques are in use nowadays which are effective in enhancing the image quality so that the image can be perceived with no loss of information. Enhancement of an image can involve different aspects of the image correction such as tonal variations, sharpness, brightness, contrast correction, to name a few. Traditional method of histogram equalization is widely used for contrast enhancement. The main disadvantage of using this method is that it can result in loss of data and presents a very unnatural look by over-enhancing the image [1]. Various histogram modification techniques have been proposed to overcome the drawbacks of histogram equalization method. A histogram equalization method for enhancing the IR image is proposed [2]. A threshold value is used to divide the histogram into background and foreground parts. To prevent the over-enhancement of the image, the limits of the threshold are optimized using particle swarm optimization.

Gamma correction is another widely used method for enhancing the image contrast. Different modifications of gamma correction are being currently developed and used by the researchers such that the method can be made useful for different image types. Deciding the gamma value is a major challenge faced by the researchers today mainly in cases where the image has a combination of high and low contrast in them. An entropy-based gamma correction is proposed wherein the entropy is computed at each intensity level in the image so as to capture the local characteristics of the image. This is further used in the adaptive estimation of the gamma value [3]. The resulting image is more natural looking in comparison to that obtained using histogram equalization.

Although gamma correction and histogram equalization can provide an acceptable improvement in the image contrast, the extent of enhancement is uncontrollable and cannot be generalized for all types of images [4]. To overcome this limitation, an adaptive image enhancement using artificial bee colony optimization is proposed, which selects the optimal values for the control parameters in an incomplete beta function corresponding to the best curve in a grayscale transformation. Compared to genetic algorithm, it converges to its fitness value very fast. Single-value decomposition and DWT were used in combination for contrast and brightness enhancement of satellite image [5]. This method utilized ABC method for optimizing each of the decomposed bands on applying the DWT. Another method which utilizes gamma correction is based on the objective of improving the image contrast and keeping the mean brightness of the final image closer to the input image. Even though histogram equalization was one of the first choices to attain this objective, maintaining the mean brightness of the image was a huge challenge. Thus, the histogram equalization got replaced with nonlinear gamma correction

[6]. To generalize the technique for different image types, the control parameter optimization was also carried out. The method utilizes compression, expansion, and aggregation of image intensities for improved image contrast enhancement and brightness preservation. Another metaheuristic contrast enhancement is based on combination of genetic algorithm and ant colony optimization which generates the transfer function for contrast enhancement [7]. A multiobjective-based histogram equalization was further introduced which used an Otsu's thresholding method to segment the histogram [8]. This method could successfully maintain the mean brightness of the image. Modified differential algorithms and social spider-based contrast and brightness-preserving algorithms are widely in use today [9, 10].

A major conclusion drawn from all the previously discussed techniques is that a single method may not give acceptable results for different images which differ in image characteristics. This led to the use of hybrid method which involves combining the existing traditional methods to give improved results. A method which combines histogram equalization and adaptive gamma correction with weighted distribution is proposed [11]. This method overcomes the abovementioned limitations by analyzing the characteristics in the image and applying the enhancement technique based on its characteristics. The use of adaptive gamma correction helps in dynamically setting the control parameters which results in different transformations for different image types. The main advantage of this method is that because of its minimum time complexity, it can be used in a wide range of applications such as video processing, digital photography, and so on. Another method proposed was based on the adaptive gamma correction on the locally equalized histogram for improving the contrast of dark images [12]. There is no iterative approach involved as part of this technique. Hence, this technique is time efficient and can be applied easily on real-time images for effective image enhancement.

Besides the conventional techniques, a parametric transformation function is proposed for contrast enhancement which makes use of artificial bee colony (ABC) optimization [13]. This paper also introduces a new fitness function by introducing a new contrast measure. The main advantage of the method over the conventional techniques is its high-speed and efficient implementation along with the improved quality of the final image.

An image obtained from a single sensor may not provide enough information regarding the feature to be evaluated, which requires the images obtained from different sensors to improve the human vision and to overcome the limitations in each source image. The fused image should be such that it includes all the details in the individual source images and it should not add any new artifacts which can mislead the viewer. Multimodal image fusion combines information from different sensors in the fused image. Registration of these images needs to be done to transform them into the same coordinates. But multitemporal satellite images which are obtained from the same sensors do not require the registration stage [14]. Many efforts were made from the researchers in the area of satellite image fusion. Among the most prominent fusion techniques is the transform-based technique, mainly because of its simplicity. Remote sensing image fusion based

on Shift-Invariant Shearlet Transform (SIST) and regional selection is one such transform-based method [15]. SIST is one of the most widely used methods as it provides better directional representation along with spatial transformation so that high-detail information can be easily injected. Apart from the most commonly used wavelet transform, contourlet transform also finds wide use. But these transforms have the limitation that they cannot adequately represent images due to their subsampling. An important property of contourlet transform is that it is shift-variant and leads to loss of information on performing fusion. So, Nonsubsampled Contourlet Transform (NSCT) is being widely used in many areas like medical, remote sensing, etc. [16]. Contourlet-based image fusion offers good preservation of features when compared to shearlet- and wavelet-based image fusion. The major drawback of this technique is that it is applicable mainly for Synthetic Aperture Radar (SAR) and panchromatic (PAN) image fusion and needs to be extended to multispectral and hyperspectral image fusion [17].

In recent years, sparse representation (SR) is very useful for image classification, image feature extraction, image deblurring, and image fusion. Researchers have come up with fusion of multiple images from same or different sensors using sparse representation in combination with well-defined dictionaries. The similar patches from the source images are clustered, and a dictionary is constructed using only a few of the principal components that can describe each of joint patch clusters, and these are combined to form a dictionary [18]. But this technique is found to be computationally efficient when considering a small-sized dictionary. Constructing a good dictionary is the key to a successful image fusion technique in such sparsity-based models. This led to another sparse-based model which utilized a dictionary for its representation. Here, the image fusion was performed using cartoon-texture decomposition and sparse representation [19]. The input images were decomposed into their respective cartoon and texture components. The cartoon component used the energy-based fusion rule, while the texture component used the sparse representation-based fusion rule. The sparse representation used the dictionary learning method. The presence of a large number of matrix-based computations, however, posed a major limitation to this method in terms of the computation time.

The abovementioned schemes have their own limitations. The major limitation encountered is that the spectral information in the image is preserved effectively but lacks in the ability to express the spatial characteristics well. Hence, it can be observed that the abovementioned schemes cannot preserve the salient features in the source images and introduce unwanted artifacts in the fused image. In spite of all the available image fusion techniques, there still exists the unresolved issue of the inability to combine the individual contributions in the source images to form the fused image.

Fusions of images with different spatial and spectral resolutions play an important role in different applications such as satellite image change detection. Mostly, multispectral and panchromatic image fusion for improving efficiency of change detection is in use even today [20]. Different methods for image fusion are available. It can be classified as pixel-based, feature-based, and decision-level-based methods. The commonly used Principal Component Analysis (PCA) is mainly relevant for

spectral transformation. A novel method which combines the spatial PCA with spectral PCA to get a resolution similar to that obtained using multisensor resolution is proposed in [21]. This method also finds wide use in the area of medical image processing. Different medical images can be fused together via multimodal image fusion. Different methods for multimodal image fusion have been proposed recently. A method involving Convolutional Neural Network (CNN)-based image fusion in combination with shearlet transform is proposed [22]. But the complexity of this method limits its use in real-time applications relating to therapy and other diagnosis systems. The limitations of applying CNN to real-time applications can be overcome using an intuitionistic fuzzy set-based image fusion [23].

The main limitations in the currently used fusion techniques is the inefficiency in controlling the noise content in the source images which also limit the quality of the fused image. This paper explores an optimized satellite image contrast enhancement-based fusion technique which utilizes a hybrid contrast enhancement technique using histogram equalization and gamma correction where the control parameters are optimized using artificial bee colony optimization technique. The contrast-enhanced image is further fused using three traditional methods such as IHS-based fusion, PCA-based fusion, and DWT-based fusion, and the performance of the fusion technique is evaluated and compared in terms of PSNR, MSE, SSIM, FSIM, and entropy.

The remaining paper is organized as shown: Sect. 5.2 explains the software requirements and the proposed optimized contrast enhancement-based satellite image fusion method. The subdivisions under this section explain the histogram equalization, gamma correction, and the optimization technique in detail. Section 5.3 shows the experimental results of contrast-enhanced image after fusion using the traditional fusion algorithms in terms of various performance metrics such as PSNR, MSE, SSIM, FSIM, and entropy for 10 different samples. Finally, the conclusion of the paper along with the future works is presented in Sect. 5.4.

5.2 Materials and Methods

The efficiency of the proposed method is tested by applying it on different sets of satellite images obtained from the same sensor but at different times. This includes 10 sets of LANDSAT images obtained from earth explorer site (http://earthexplorer.usgs.gov/) and from global land survey (http://changematters.esri.com/). A set of 70 images are available of which we use 10 samples for analysis. All the experiments are programmed in MATLAB 2018a on an Intel(R) Core(TM) i3-4005U CPU @ 1.70GHz laptop with 8.00 GB RAM.

Table 5.1 shows the details of the satellite images used for study. The satellite images of the same location at different years are taken for analysis. Since the images are captured using the same sensors, image registration at the time of image fusion is not necessary.

From the table, a set of ten such images are taken for the analysis.

Table 5.1 Details of satellite imagery used for study

Satellite data	Spatial resolution	Source
LANDSAT ETM+	30 m	Earth explorer site (http://earthexplorer.usgs.gov/)
		Global land survey (http://changematters.esri.com/)
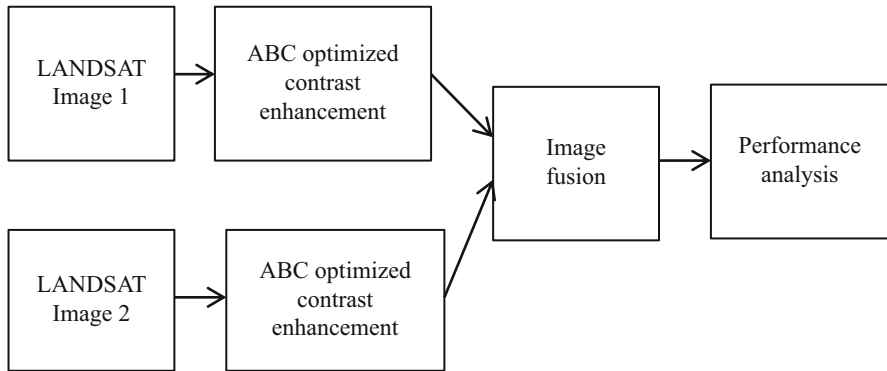
Fig. 5.1 Block diagram of the proposed method

5.2.1 Methodology

In the proposed method, the source images are subjected to contrast enhancement using an artificial bee colony optimization technique and then fused using various traditional fusion methods. The various fusion techniques applied are DWT-based fusion, PCA-based fusion, and IHS-based fusion. Figure 5.1 shows the block diagram of the proposed method.

The acquired satellite images of the same location at different times are preprocessed to enhance the image contrast using a gamma-corrected histogram equalization which is optimized using artificial bee colony optimization technique.

5.2.2 Contrast Enhancement of Satellite Image

Contrast enhancement forms one of the basic steps in image enhancement. It involves expanding the intensity levels in the image or modifying the image histogram so that it can be easily interpreted by humans. The most commonly used method in contrast enhancement is using global intensity transformation. This method involves a transformation function which maps the input intensity values to a new set of intensity values. Figure 5.2 represents the block diagram of the contrast enhancement technique used as part of this work.

The source image is subjected to global histogram equalization. A single stage of histogram equalization is not sufficient for the enhancement of such images
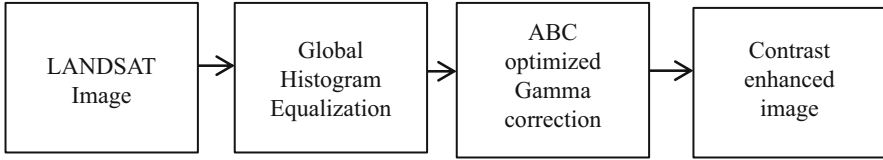
Fig. 5.2 Block diagram of the contrast enhancement technique

as it can limit the brightness levels in the image. A satellite image which is used in application such as image fusion should be such that the fused image preserves maximum information in the source images and should be useful for further analysis.

To attain this, an optimized gamma correction technique has been used for the further enhancement of the medium contrast images.

5.2.2.1 Global Histogram Equalization

Global histogram equalization (GHE) is a method used to increase the dynamic range of an image. It sets the intensity values in the input image such that there is a uniform intensity distribution on the output image.

Let $I = \{I(x,y)\}$ be the input image where $I(x,y)$ denotes the intensity values in the image I. n represents the total number of pixels in I, and the intensity values in I are divided into L levels $\{I_0, I_1, I_2, \ldots \ldots \ldots \ldots I_L\}$ [2].

The probability density function (PDF) of an intensity level I_k in the image can be formulated as given in Eq. (5.1):

$$p(I_k) = \frac{n_k}{n}, k = 0, 1, 2, \ldots \ldots L - 1 \tag{5.1}$$

Here, n_k represents the number of pixels with intensity level I_k.

Depending on the PDF, the cumulative distribution function (CDF) can be calculated from Eq. (5.2) as

$$c(I_k) = \sum_{s=0}^{k} p(I_s) = \sum_{s=0}^{k} \frac{n_s}{n}, k = 0, 1, 2, \ldots \ldots \ldots L - 1 \tag{5.2}$$

It is a known fact that from Eq. (5.2) $\sum_{k=0}^{L-1} c(I_k) = 1$.

Now, a transfer function $T(x)$ is formed based on the CDF which is given in Eq. (5.3) as

$$T(I_k) = I_0 + (I_{L-1} - I_0) \cdot c(I_k) \tag{5.3}$$

Now, the output image O which is the enhanced image is formulated as in Eq. (5.4):

$$O = T(I) = \left\{ T\big(I(x, y) \mid \quad \forall \quad I(x, y) \in \left\{ I_0, I_1, I_2, \ldots \ldots \ldots I_{L-1} \right\} \big) \right\}$$

(5.4)

The drawback of this method is the inability of GHE method to maintain the mean brightness of the output image. For this, some modifications can be done on the brightness so as to bring it to the desired levels. This is achieved using gamma correction.

5.2.2.2 Gamma Correction

The gamma correction is an image contrast enhancement method which is based on image-dependent and local exponential correction. It is a highly cost-effective method and is efficient when dealing with images which are both dim and bright. However, manually selecting the gamma value depending on the image can be time-consuming especially in the case of satellite images [11].

For a traditional gamma correction, the output image is obtained using Eq. (5.5) as

$$I_{\text{out}} = cI_{\text{in}}^{\gamma}$$

(5.5)

Here, I_{in} and I_{out} give the intensities of the input and output images respectively. The shape of the transformation curve is decided by the value of c and γ. Depending on the nature of the image, the c value changes.

An alternate method for contrast enhancement is to do compression of intensity levels toward darker level and simultaneously expanding the intensities toward brighter levels.

For images with low and medium contrast, the transformation function can be rewritten as given in Eq. (5.6):

$$I_{\text{out}} = \alpha I_{\text{cp}} + (1 - \alpha) I_{\text{ex}}$$

(5.6)

I_{cp} here is the source image compressed toward darker levels, and I_{ex} is the source image expanded to brighter levels.

I_{cp} is obtained using Eq. (5.7):

$$I_{\text{cp}} = I_{\text{in}}^{\gamma}$$

(5.7)

I_{ex} is obtained using Eq. (5.8):

$$I_{ex} = 1 - (1 - I_{in})^\gamma \tag{5.8}$$

I_{in} represents the histogram-equalized image.

From Eqs. (5.6), (5.7), and (5.8), it is clear that the contrast-enhanced image is controlled by two parameters, namely, α and γ. To ensure that the enhancement is within permissible levels without any color distortions, the parameters α and γ need to be modified depending on the image [6].

Here in the transformation function, α and γ are the parameters to be controlled. Hence, α and γ are the parameters to be optimized. Artificial bee colony optimization is used for optimizing the control parameters.

5.2.2.3 Artificial Bee Colony Optimization-Based Modeling

ABC algorithm is a metaheuristic algorithm used for solving various optimization problems. This algorithm is based on the foraging behavior of honeybees. The population in a bee colony can be divided as (i) employed bees, (ii) onlooker bees, and (iii) scout bees.

These bees move in the search space in search of food sources which are the solutions to the optimization problem under consideration. The ABC algorithm is illustrated using a flowchart as in Fig. 5.3.

The employed bees study the available food sources and convey the information to the onlooker bees. The food source with maximum score and maximum quality has a larger chance of being selected by the onlooker bees. The food source with minimum score has the least chance of being selected. It can also be rejected if the quality is low. In such a case, the scout bees will randomly search for a new food source. In each iteration involved during the search for food sources, three major steps are followed:

 (i) The employed bees search for food source and find their quality.
 (ii) On conveying the information about the food source, the food sources are selected by the onlooker bees.
(iii) On rejection of the food sources, the scout bees are sent in search of new food sources.

The main advantage of using ABC over other optimization is that since the algorithm carries out local and global search in each level of iteration, local optimum solution is not considered and the chances of finding the optimal solution is high.

For applying ABC, it requires a transformation function and a fitness function. In the proposed method, the image is enhanced using a gamma correction-based transformation function. Here, α and γ of the image are to be optimized. The bees move in the search space to find the optimal values of α and γ to enhance the contrast of the image by evaluating the fitness function.
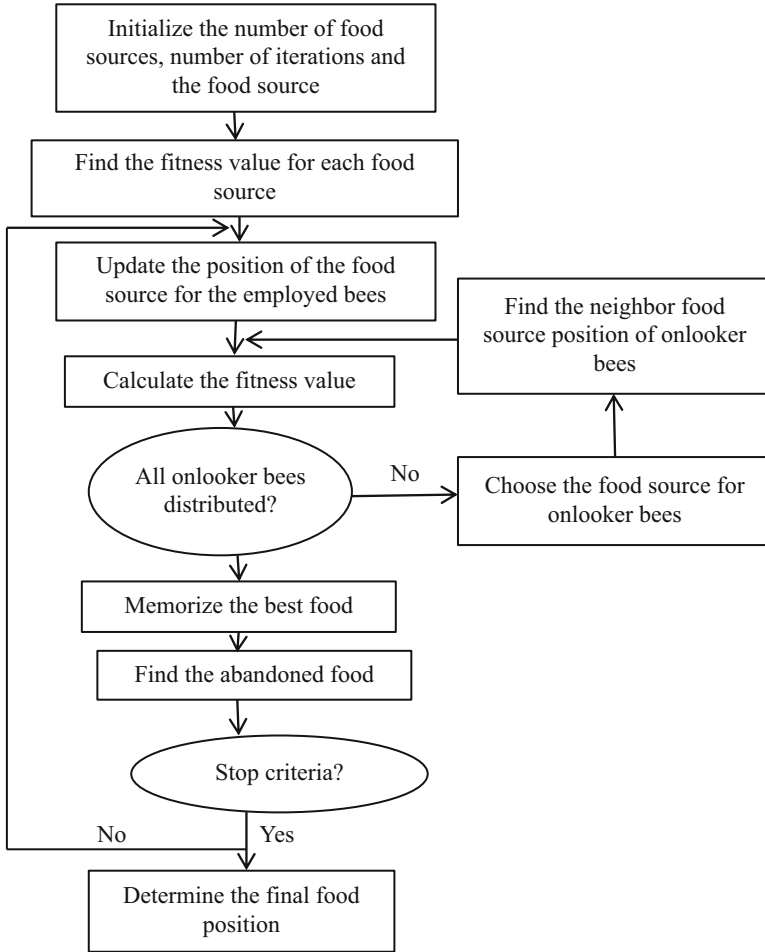
Fig. 5.3 Flowchart of ABC optimization technique

The fitness function is an objective function which evaluates the quality of the generated image. One such measure of the quality is the entropy. When compared to the original image, it is desired that the enhanced image has a higher value.

The ABC algorithm is implemented using the following steps:

- The population is initialized. The food sources are considered as the solution to the optimization problem. We use this algorithm to find the optimal values of α and γ. The parameters pertaining to ABC optimization are also set. The number of bees in the colony is set to 100, the number of iterations to 100. The nectar amount not updated is set to 5.
- Initialize by defining a 2D search space for the control parameters α and γ as $(\alpha, \gamma) = [[0,1], [1,5]]$.

$$X_i = \{\alpha_i, \gamma_i\}$$

- In the employed bee stage, each bee finds a new solution X_i^{n+1} within the neighborhood of the current solution X_i^n for every iteration using Eq. (5.9) as

$$X_i^{n+1} = X_i^n + \phi\left(X_i^n - X_j^n\right) \tag{5.9}$$

where is a random index and $\emptyset$ is a random value within $[-1,1]$.

The fitness value used for the technique is entropy. So the aim is to maximize the entropy value. It is calculated using Eq. (5.10) as

$$H = -\sum_{i=0}^{L} P(i)\log_2 P(i) \tag{5.10}$$

If the fitness value of the new solution X_i^{n+1} obtained using this equation is more than the current solution X_i^n, then employed bees proceed to this food source and abandon the old source. Otherwise, the old value is retained.

- İn the onlooker bee stage, once the employed bees complete their search, onlooker bee finds the solution based on a roulette wheel selection method which is stated using Eq. (5.11) as

$$P\left(X_i^n\right) = F\left(X_i^n\right) / \sum_{i=1}^{N} F\left(X_i^n\right) \tag{5.11}$$

where F represents the fitness value of X_i^n and N is the total solutions.

- In the scout stage, if there is no improvement in fitness value over many generations, a new solution is found.
- The iteration can be stopped if the maximum iterations are reached and fitness value has converged.

The RGB channels in the image are contrast enhanced initially using histogram equalization. The basic ABC parameters are initialized along with the control parameters for gamma correction. The value of I_{out} is calculated using Eq. (5.6). The fitness function is evaluated for all the possible solutions and the most suitable solution is chosen. With this, the optimized (α,γ) is used for the ultimate enhancement of the RGB image.

5.2.3 Satellite Image Fusion

Multiple images from different sensors or images of the same location from the same sensor at different times can be combined using image fusion. In this work, the images are obtained from same sensors, and hence no image registration is done.

The images obtained from satellites can be corrupted by atmospheric interferences such as noise. The success of the image fusion is decided by the quality of the input images. This necessitates the need for preprocessing techniques prior to image fusion. Image enhancement is an important part of preprocessing. It should be noted that over-enhancing the image can affect the image quality, and hence the purpose of image fusion to preserve the information content in the source image fails. To maintain the mean brightness of the source images, they are subjected to an image contrast enhancement using a combination of histogram equalization and gamma correction. The control parameters for gamma correction are optimized using the ABC optimization to obtain the enhanced image. The enhanced source images are the input to the fusion stage.

In this work, three main image fusion techniques are discussed. They are the IHS-based image fusion, PCA-based image fusion, and DWT-based image fusion. Each of the fusion techniques are discussed in detail in this section.

5.2.3.1 PCA-Based Image Fusion

Principle component analysis or PCA is useful in transforming a set of correlated variables into a set of uncorrelated variables. It is directed at reducing a large variable set to a smaller set that can still preserve the information available in the larger dataset. This technique helps in creating a reduced set of variables and using the reduced variable set for analysis. The reduced set of variables is called principle factor.

PCA method is mainly used in image compression, image enhancement, classification, object detection, and image fusion.

The PCA-based fusion method takes the pixel values at each location in the individual source image, adds a weight value to each of the pixels in the image, and finds the average of the weighted values at each location in the source image to produce the fused image.

PCA-based fusion method improves the resolution of the images such that the two images to be fused are divided into subimages having different frequencies. Now, the principle components are fused together, and the subimages are reconstructed such that the final image contains the informations in the individual source images. Also, the final image has enhanced information when compared to the individual images.

Thus, the PCA generates a fused image of a large number of input images in the form of a weighted superposition of individual source images. PCA is currently
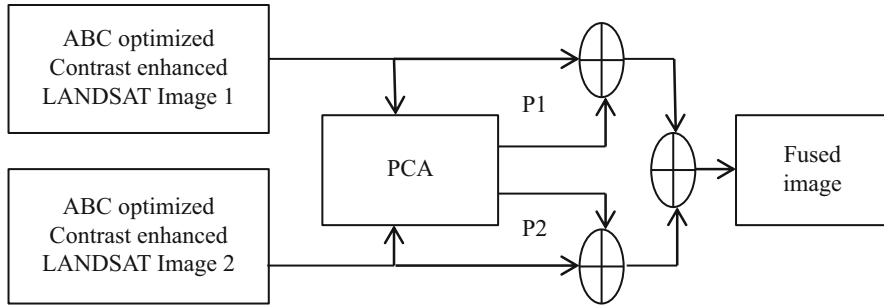
Fig. 5.4 Block diagram of PCA-based fusion technique

the most common method used for the fusion of multispectral and panchromatic images. Figure 5.4 shows the block diagram of PCA-based fusion technique.

The main idea of using PCA-based fusion method is that since it can remove the correlation between the different sets of multitemporal satellite images, the method requires only a subset of details to represent the information in the final fused image. But utilizing only a small set of data has the disadvantage that there is some loss of information when representing the fused image. So the method is considerably less accurate for image fusion.

5.2.3.2 DWT-Based Image Fusion

With the development of multiresolution images, wavelet-based fusion method has become one of the most promising methods in image processing. It is capable of generating images with very high spectral and spatial quality by minimizing color distortions in the image. DWT is the most fundamental and simplest of fusion methods among the available multiscale transforms. It is widely used in various applications over wide research areas such as in audio compression, video compression, pattern recognition, image fusion, and so on. It is most popularly used in comparison to Discrete Cosine Transform because of the high quality of fused image. DWT being a multiresolution technique can use images with different degree of resolutions.

Wavelet transform is an extension of the high-pass filtering technique. DWT uses filter banks for its implementation. The input image gets decomposed into high- and low-frequency components. The low-frequency components can be further decomposed until the desired resolution level is obtained. At each level of transformation, the image gets downsampled. At each level, the filtering followed by a decomposition stage is applied to the rows and then the columns in the image. This results in four images, namely, an approximation image and three wavelet coefficients. The approximation image is the result of applying low-pass filtering techniques to the rows and columns in the image. The next level of transformation is applied only to the approximation image.

After applying the wavelet transform, the result will be the approximation image and the wavelet coefficients for the individual images. Next step is to merge the approximation images and wavelet coefficients element by element from the two source images. Combining the individual images can be done on the basis of certain fusion rules. Once the coefficients are merged, the inverse transform is taken to obtain the fused image.

In this work, DWT-based image fusion is done by using the averaging rule to find the average of wavelet coefficients of the decomposed source images such that low-frequency information is not lost. Figure 5.5 represents the DWT-based image fusion method.

Different fusion rules can be applied on the coefficients based on combining the maximum of the elements in the individual source images, minimum of the elements in the individual source images, and average of the individual source images, randomly combining elements, or using either of the two source images.

The selection of the fusion rule in order to fuse the wavelet coefficients is an important factor to be considered. The fusion rule should be selected depending on the resolution and the quality of the individual source images. If the fusion rule is randomly chosen, the image quality of the fused image will be poor and the image cannot be used for further analysis.

The main disadvantage of using DWT is that the image quality can be improved with multiple decomposition levels. As the number of decomposition levels increases, the computational complexity increases; and with the overhead of huge image size also present, researchers needed to find alternative methods of fusion of satellite images.
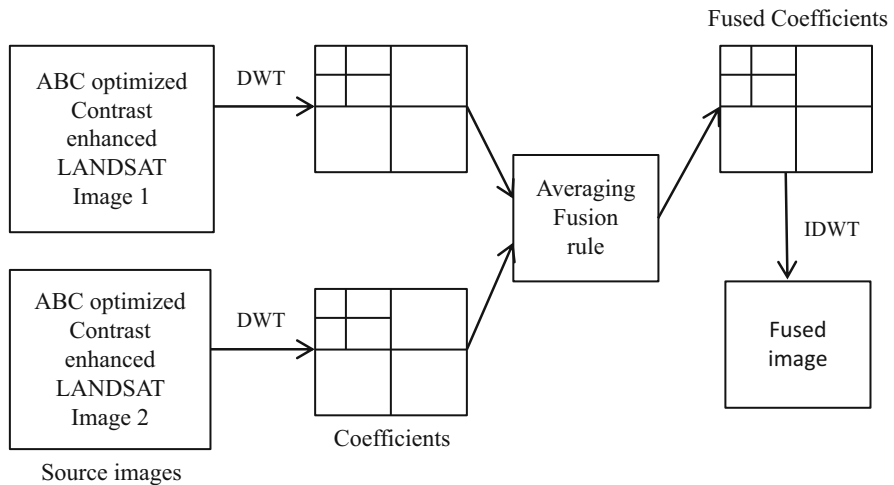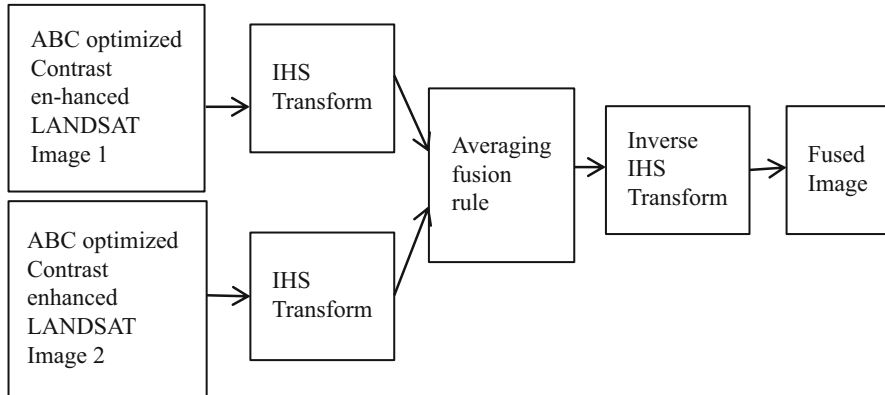
Fig. 5.5 DWT-based image fusion

Fig. 5.6 IHS-based image fusion

5.2.3.3 IHS Transform for Image Fusion

IHS or Intensity, Hue, Saturation Transform is a commonly used transform for image fusion. It finds application in enhancing color in an image, enhancing the features in an image, and improving the spatial resolution of the image. In the IHS domain, the spectral information is mostly contained in the hue and saturation components. The main use of IHS transform is to ensure that all the spatial and spectral information in the source images are reflected in the fused images. Figure 5.6 shows the block diagram of the IHS-based image fusion.

The RGB source images are divided into their respective intensity, hue, and saturation components. The hue, saturation, and intensity components in the individual source images are combined using the averaging rule in fusion to get the new hue, saturation, and intensity components. Inverse IHS transform on the new intensity, hue, and saturation components gives the fused image.

5.3 Results and Discussions

To prove the effectiveness of the proposed method, the results are obtained on a ten image sets by quantitative methods. The results are compared between three of the existing image fusion methods: IHS-based fusion, PCA-based image fusion, and DWT-based image fusion. The performance metrics such as Peak Signal-to-Noise Ratio (PSNR), Mean Square Error (MSE), Feature Similarity Index Metric (FSIM), Structural Similarity Index Metric (SSIM), and entropy are calculated for each of the techniques. Table 5.2 compares the contrast-enhanced image fusion results of the existing methods in terms of PSNR.

PSNR is a measure of the accuracy of the final image. It is computed using Eq. (5.12) as

Table 5.2 PSNR comparison
for the existing fusion
techniques on
contrast-enhanced images

Image dataset	PCA(dB)	DWT(dB)	IHS(dB)
Sample 1	30.6771	35.9709	40.8181
Sample 2	31.0766	36.1608	39.9114
Sample 3	33.0245	36.0774	38.7934
Sample 4	32.2615	37.2000	38.4679
Sample 5	33.7081	36.3310	38.1822
Sample 6	32.0224	35.3657	40.0575
Sample 7	30.0808	37.0617	40.3593
Sample 8	33.6824	36.1498	39.7005
Sample 9	32.5914	36.6965	38.4085
Sample 10	30.8632	34.0386	37.0513

$$\text{PSNR} = \log_{10}\left(\frac{255 * 255}{\text{MSE}}\right) \tag{5.12}$$

where MSE is the Mean Square Error and it is the sum of the square of error between the original image and the image which is affected by noise.

Each and every image has some level of uncertainty present in them. Fusing uncertain details in the images leads to an uncertain fused image. The proposed fusion method removes the uncertainties in the source images resulting in a fused image with better contrast in comparison to the traditional methods. In IHS transform, the intensity component depicts the brightness of the image. When the contrast-enhanced images are fused using IHS, the averaging of the individual components ensure that the overall brightness and contrast of the source images are preserved in the fused image. This is conveyed in Table 5.2 where it is clearly observed that the PSNR values of the IHS-based image fusion are superior over those obtained using PCA and DWT.

On an average, the PSNR values have increased from 36.1052 dB in the case of DWT and 31.9980 dB in the case of PCA-based fusion to 39.1750 dB in the case of IHS-based fusion technique. Therefore, we can say that there is a percentage increase of about 22.42 over the other techniques. Table 5.3 compares the contrast-enhanced image fusion results of the existing methods in terms of MSE.

The higher the value of PSNR, the lower will be the MSE. The MSE values are smaller for the proposed methods when compared to the existing methods.

On an average, the MSE values have decreased from 22.626 in the case of DWT and 26.988 in the case of PCA-based fusion to 17.057 in the case of IHS-based fusion technique. Therefore, we can say that there is a percentage decrease of about 36.79 over the other techniques.

SSIM is a measure of the structural similarity. It finds the similarity between the source images by performing comparisons between correlations of luminance, structure, and contrast on the individual images. It has a range between -1 and 1. Closer the value of SSIM to one indicates that the output images is structurally similar when compared to the source image. It is computed using Eq. (5.13) as follows:

Table 5.3 MSE comparison for the existing fusion techniques on contrast-enhanced images

Image dataset	PCA	DWT	IHS
Sample 1	24.112	20.430	15.761
Sample 2	25.241	22.077	16.550
Sample 3	26.213	25.447	17.063
Sample 4	27.618	23.516	18.616
Sample 5	25.981	20.804	14.722
Sample 6	26.682	19.057	16.412
Sample 7	28.956	20.463	16.080
Sample 8	27.943	22.329	16.912
Sample 9	28.068	25.918	18.329
Sample 10	29.067	26.204	20.133

Table 5.4 SSIM comparison for the existing fusion techniques on contrast-enhanced images

Image dataset	PCA	DWT	IHS
Sample 1	0.9224	0.9615	0.9660
Sample 2	0.8856	0.9392	0.9526
Sample 3	0.8775	0.9280	0.9648
Sample 4	0.8820	0.9042	0.9466
Sample 5	0.8779	0.9174	0.9528
Sample 6	0.8661	0.9262	0.9505
Sample 7	0.8729	0.9175	0.9639
Sample 8	0.8526	0.9225	0.9487
Sample 9	0.8821	0.9147	0.9668
Sample 10	0.8612	0.9094	0.9625

$$\text{SSIM} = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \qquad (5.13)$$

where μ_x and μ_y are the sample means of x and y, respectively; σ_x and σ_y give the sample variances of x and y, respectively; and σ_{xy} represents the sample correlation coefficient between x and y, and x and y are local windows in the two source images. Table 5.4 compares the contrast-enhanced image fusion results of the existing methods in terms of SSIM.

On an average, the SSIM values have increased from 0.9240 in the case of DWT and 0.8775 in the case of PCA-based fusion to 0.9575 in the case of IHS-based fusion technique. Therefore, we can say that there is a percentage increase of about 9.11 over the other techniques.

FSIM is a measure of the feature similarity between the source image and the processed image. It is calculated using Eq. (5.14):

$$\text{FSIM} = \frac{\sum\limits_{x \in X} S_L(x)\text{PC}_m(x)}{\sum\limits_{x \in X} \text{PC}_m(x)} \qquad (5.14)$$

where X represents the whole image, $S_L(x)$ denotes the similarity in the two images, and PC_m is the phase congruency map. Table 5.5 compares the contrast-enhanced image fusion results of the existing methods in terms of FSIM.

On an average, the FSIM values have increased from 0.8960 in the case of DWT and 0.8722 in the case of PCA-based fusion to 0.9327 in the case of IHS-based fusion technique. Therefore, we can say that there is a percentage increase of about 6.93 over the other techniques.

Table 5.6 compares the contrast-enhanced image fusion results of the existing methods in terms of entropy.

The entropy H is the amount of information contained in the image. It is calculated using Eq. (5.15) as

$$H = -\sum_{i=0}^{L} P(i)\log_2 P(i) \tag{5.15}$$

where P is the normalized histogram value of image i.

It is observed that the entropy values are higher in the proposed method when compared to the traditional methods. The higher values indicate that the fused image has extracted and preserved every detail in the source images effectively.

Table 5.5 FSIM comparison for the existing fusion techniques on contrast-enhanced images

Image dataset	PCA	DWT	IHS
Sample 1	0.8852	0.8903	0.9418
Sample 2	0.8803	0.8814	0.9243
Sample 3	0.8932	0.9046	0.9409
Sample 4	0.8732	0.9152	0.9348
Sample 5	0.8865	0.9137	0.9264
Sample 6	0.8754	0.9066	0.9414
Sample 7	0.8671	0.8993	0.9286
Sample 8	0.8457	0.8813	0.9492
Sample 9	0.8544	0.8907	0.9258
Sample 10	0.8612	0.8773	0.9145

Table 5.6 Entropy comparison for the existing fusion techniques on contrast-enhanced images

Image dataset	PCA	DWT	IHS
Sample 1	5.2704	5.9559	6.0266
Sample 2	5.0519	5.9056	6.2049
Sample 3	5.2055	6.3019	7.2550
Sample 4	4.6612	5.1758	6.1566
Sample 5	4.3464	5.4033	6.3258
Sample 6	4.6248	5.2021	5.5841
Sample 7	4.5564	5.2816	6.3183
Sample 8	4.8799	6.5108	7.1499
Sample 9	5.5918	6.0612	6.4390
Sample 10	4.9430	5.7270	6.2204

On an average, the entropy values have increased from 5.7525 in the case of DWT and 4.9130 in the case of PCA-based fusion to 6.3675 in the case of IHS-based image fusion. Therefore, we can say that there is a percentage increase of about 29.60 over the other techniques.

From the overall analysis of the PCA-, DWT- and IHS-based fusion methods, it can be observed that the IHS-based fusion method yields improved results in terms of PSNR, MSE, FSIM, SSIM, and entropy and gave acceptable levels of contrast enhancement in comparison to the other techniques. PCA-based fusion adds unwanted artifacts in the image and affects the image quality. DWT-based fusion is complex in terms of the multiple decomposition stages. IHS-based fusion is simpler and easier to implement and can preserve the color information in the source images in comparison to the other two methods.

Figure 5.7 shows the different levels of contrast enhancement for different values of gamma with and without ABC-optimized gamma correction.

From Fig. 5.7, it is clear that appropriate selection of gamma value is essential and the image contrast can drastically vary with gamma value. In order to improve the contrast and preserve the maximum information in the image, an optimal selection of gamma value is necessary. Otherwise, due to the poor contrast of the source images, the quality of the fused image can also get affected which limits its usage for further analysis. Only on applying ABC optimization for contrast enhancement we get a good contrast-enhanced image in which all the information
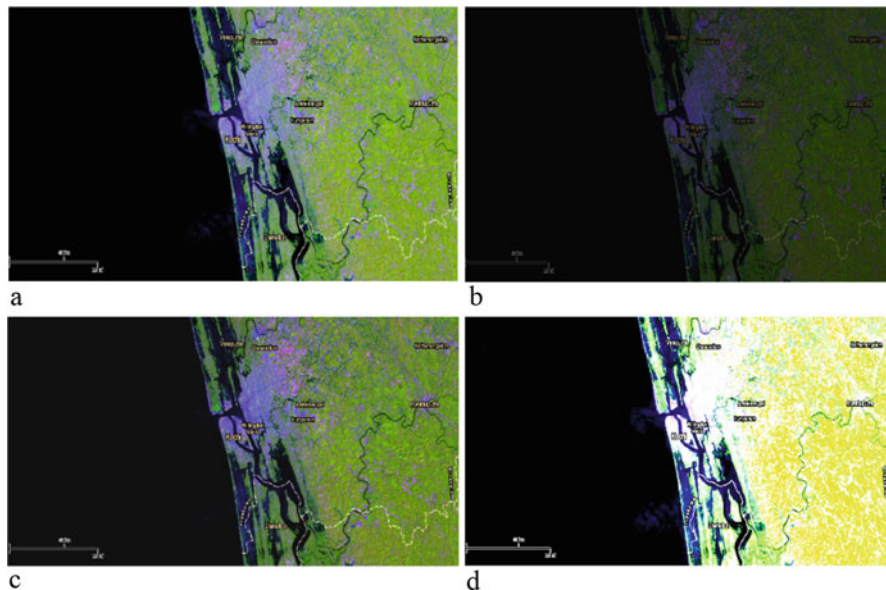
Fig. 5.7 (**a**) Input LANDSAT image. (**b**) Contrast-enhanced image for $\gamma = 1.2$, $\alpha = 0.2$ (without ABC optimization). (**c**) Contrast-enhanced image for $\gamma = 2.6$, $\alpha = 0.2$ (with ABC optimization). (**d**) Contrast-enhanced image for $\gamma = 4.7$, $\alpha = 0.2$ (without ABC optimization)

contents in the original image can be clearly distinguished. Randomly choosing γ and α values without ABC optimization can give poor or overly enhanced images which affect the image quality of the succeeding fusion stage.

The contrast-enhanced images without optimization can give a dark image or an overly illuminated image, and such images if used for the succeeding fusion stage can give a fused image which does not convey the information present in the individual source images clearly and lead to poor-quality output image. Hence, by optimizing γ and α which are the parameters as part of the transformation function in gamma correction method, satellite images with good contrast enhancement can be obtained without manually tuning the control parameters which is extremely time-consuming for such large images.

The gamma values cannot be uniquely chosen for a set of images. It varies across images, and manual tuning of gamma values can be tedious especially when the image contains combined bright and dark contrast regions. Hence, optimization of the control parameters can solve this problem. This technique is faster and easier to implement when compared to other conventional techniques.

By using a hybrid contrast enhancement technique which combines histogram equalization and gamma correction helps to improve the image quality by enhancing the image contrast and preserving the brightness in the image. The disadvantage of the histogram equalization that it over-enhances the image is controlled by the gamma correction technique which utilizes an optimization technique to select the gamma value which gives acceptable level of contrast enhancement without loss of information. Also, the use of artificial bee colony optimization ensures that the fitness function of the solution converges fast. Figure 5.8 shows the result of fusion on the contrast-enhanced image.

It can be seen that the IHS-based image fusion outperforms DWT-based fusion and PCA-based fusion since it preserves the image brightness and the spatial information better in comparison to DWT- and PCA-based fusion since it treats the intensity, hue, and saturation components separately and fuse them. The quantitative analysis also further proves this by giving improved values of PSNR, SSIM, FSIM, and entropy for IHS-based fusion techniques over other fusion methods.

5.4   Conclusion

In this work, an ABC-optimized contrast enhancement-based fusion of satellite images has been presented. The main objective of the ABC-optimized contrast enhancement is that since the satellite images are captured under poor illumination, there is a need to preprocess the image to extract maximum information so that there will be maximum contents preserved in the subsequent fusion stage. In the proposed method, the contrast enhancement comprises of two levels: histogram equalization followed by gamma correction where the control parameters are optimized using artificial bee colony optimization. The enhanced images are input to an image fusion stage. Three traditional fusion techniques are used, namely, DWT-based fusion,
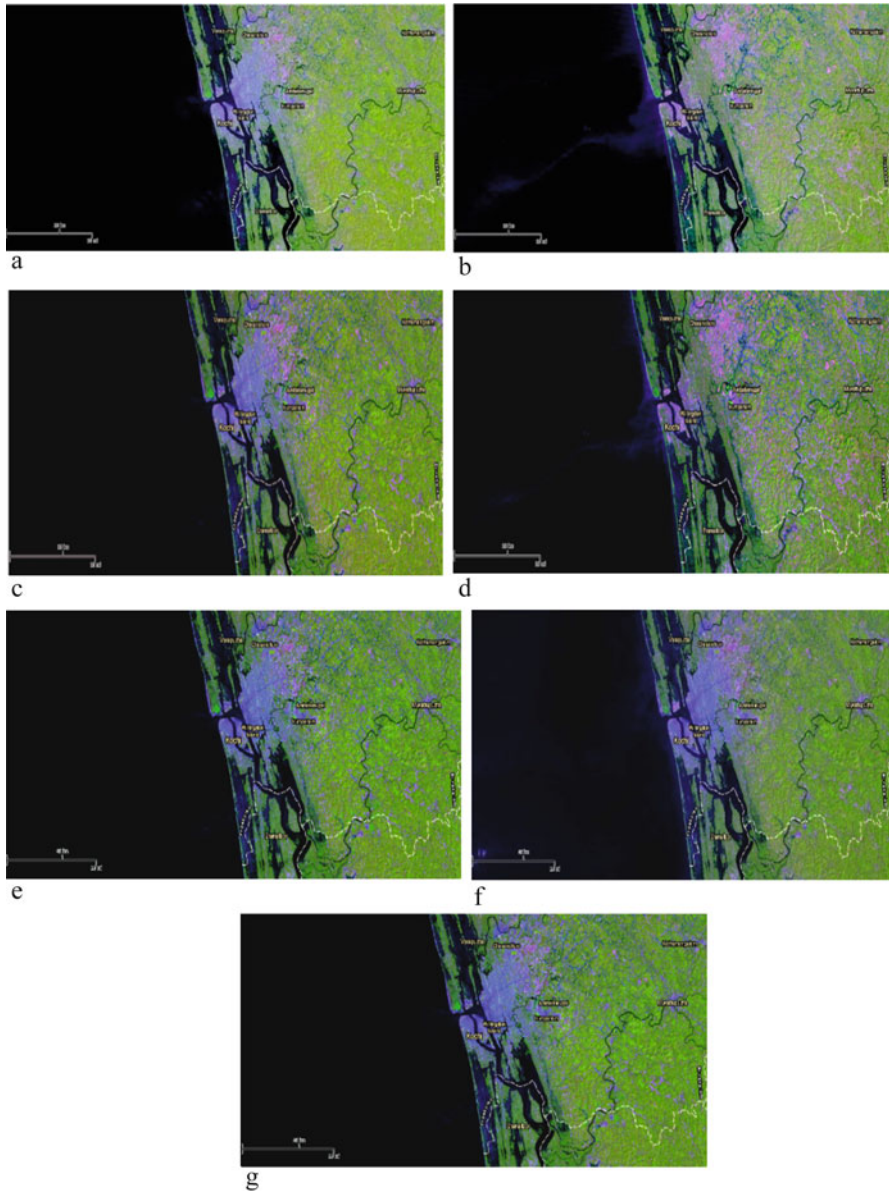
Fig. 5.8 (**a**) Input LANDSAT image 1. (**b**) Input LANDSAT image 2. (**c**) ABC-optimized contrast-enhanced LANDSAT image 1. (**d**) ABC-optimized contrast-enhanced LANDSAT image 2. (**e**) IHS-fused image. (**f**) DWT-fused image. (**g**) PCA-fused image

IHS-based fusion, and PCA-based fusion. The experimental results show that the use of contrast-enhanced image for fusion gives an improvement on the performance

of the fusion process. IHS-based image fusion has the advantage that it is easier to implement and faster than most of the conventional image fusion techniques. This method can be further extended to multi-focus image fusion and for fusing images from different sensors.

References

1. Wang X, Chen L (2017) An effective histogram modification scheme for image contrast enhancement. Signal Process Image Commun 58:187–198
2. Wan M, Gu G, Qian W, Ren K, Chen Q, Maldague X (2018) Particle swarm optimization-based local entropy weighted histogram equalization for infrared image enhancement. Infrared Phys Technol 91:164–181
3. Parihar AS (2017) Entropy-based adaptive gamma correction for content preserving contrast enhancement. Int J Pure Appl Math 117(20):887–893
4. Chen J, Li C-Y, Yu W-Y (2016) Adaptive image enhancement based on artificial bee colony algorithm. Int Conf Commun Electron Inf Eng 116:685–693
5. Bhandari AK, Soni V, Kumar A, Singh GK (2014) Artificial Bee Colony-based satellite image contrast and brightness enhancement technique using DWT-SVD. Int J Remote Sens 35(5):1601–1624
6. Jiang G, Wong CY, Lin SCF, Rahman MA, Ren TR, Kwok N, Shi H, Yu Y-H, Wu T (2015) Image contrast enhancement with brightness preservation using an optimal gamma correction and weighted sum approach. J Mod Opt 62(7):536–547
7. Hoseini P, Shayesteh MG (2013) Efficient contrast enhancement of images using hybrid ant colony optimisation, genetic algorithm, and simulated annealing. Dig Signal Process Rev J 23:879–893
8. Shanmugavadivu P, Balasubramanian K (2014) Particle swarm optimized multi-objective histogram equalization for image enhancement. Opt Laser Technol 57:243–251
9. Suresh S, Lal S (2017) Modified differential evolution algorithm for contrast and brightness enhancement of satellite images. Appl Soft Comput 61:622–641
10. Maurya L, Kumar Mahapatra P, Kumar A (2017) A social spider optimized image fusion approach for contrast enhancement and brightness preservation. Appl Soft Comput 52:572–592
11. Rahman S, Mostafijur Rahman Md, Abdullah-Al-Wadud M, Al-Quaderi GD, Shoyaib M (2016) An adaptive gamma correction for image enhancement. EURASIP J Image Video Process, Springer 35:1–13
12. Singh H, Agrawal N, Kumar A, Singh GK, Lee HN (2016) A novel gamma correction approach using optimally clipped sub-equalization for dark image enhancement. IEEE 16:497–501
13. Chen J, Yu W, Tian J, Chen L, Zhou Z (2018) Image contrast enhancement using an artificial bee colony algorithm. Swarm Evol Comput 38:287–294
14. Li Y, He Z, Zhu H, Zhang W, Wu Y (2016) Jointly registering and fusing images from multiple sensors. Inf Fusion 27:85–94
15. Luoa X, Zhang Z, Wua X (2016) A novel algorithm of remote sensing image fusion based on shift-invariant Shearlet transform and regional selection. Int J Electron Commun 70:186–197
16. Anandhi D, Valli S (2018) An algorithm for multi-sensor image fusion using maximum a posteriori and nonsubsampled contourlet transform. Comput Electr Eng 65:139–152
17. Li S, Kang X, Fang L, Hu J, Yin H (2017) Pixel-level image fusion: a survey of the state of the art. Inf Fusion 33:100–112
18. Kim M, Han DK, Ko H (2016) Joint patch clustering-based dictionary learning for multimodal image fusion. Inf Fusion 27:198–214

19. Zhu Z, Yin H, Chai Y, Li Y, Qi G (2018) A novel multi-modality image fusion method based on image decomposition and sparse representation. Inf Sci 432:516–529
20. Ghassemian H (2016) A review of remote sensing image fusion methods. Inf Fusion 32:75–89
21. Shahdoosti HR, Ghassemian H (2016) Combining the spectral PCA and spatial PCA fusion methods by an optimal filter. Inf Fusion 27:150–160
22. Hermessi H, Mouraliand O, Zagrouba E (2018) Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. Neural Comput Appl, Springer 30(7):2029–2045
23. Balasubramaniam P, Ananthi VP (2014) Image fusion using intuitionistic fuzzy sets. Inf Fusion 20:21–30

Chapter 6
Effective Transform Domain Denoising of Oceanographic SAR Images for Improved Target Characterization

S. Arivazhagan, W. Sylvia Lilly Jebarani, R. Newlin Shebiah, S. Vineth Ligi, P. V. Hareesh Kumar, and K. Anilkumar

Abstract Synthetic Aperture Radar (SAR) images are widely used for a variety of applications such as surveillance, agricultural assessment and classification, planetary and celestial investigations, geology and mining, etc., due to its remarkable characteristic of capturing it under all weather conditions. SAR images are highly prone to speckle noise due to the ingrained nature of radar backscatter. Speckle removal is highly essential to limit the difficulty encountered while processing the SAR images. An exhaustive work has been done by researchers to despeckle SAR images using spatial filters, wavelet transform, and hybrid approaches. This work aims at exploring the different despeckling techniques to identify the best and suitable methodology. On measuring the despeckling performance using Peak Signal-to-Noise Ratio, Edge Preservation Ratio, Speckle Suppression Index, Speckle Suppression and Mean Preservation Index, and Structural Similarity Index simultaneously for the various techniques experimented, ridgelet transform-based thresholding works well. It gives better results by applying ridgelet transform and processing the subbands with minimax thresholding. The type and characteristics of the scene imaged also influence the result.

Keywords Despeckling · Spatial filtering · Anisotropic diffusion · Multiresolution transform · Ridgelet transform · Minimax thresholding

S. Arivazhagan · W. Sylvia Lilly Jebarani · R. Newlin Shebiah (✉) · S. Vineth Ligi
Center for Image Processing and Pattern Recognition, Department of ECE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India
e-mail: sarivu@mepcoeng.ac.in; wsylvia@mepcoeng.ac.in; newlinshebiah@mepcoeng.ac.in; svinethligi@mepcoeng.ac.in

P. V. Hareesh Kumar · K. Anilkumar
Naval Physical and Oceanographic Laboratory, Kochi, Kerala, India

6.1   Introduction

Synthetic Aperture Radar (SAR) images, when acquired, suffer from patterns of constructive and destructive interference of backscattered signals from multiple distributed objects. This disturbance called as speckle noise, being the largest source of noise in SAR images, results in bright and dark spots in the image, provoking complications in image interpretation, like failure to extract bright spots which can be possible targets or giving false alarms by identifying other disturbances as targets. Speckle noise is a locally correlated noise in multiplicative form due to which the image processing techniques sense great difficulties when applied on SAR imagery. The heterogeneity of sea clutter, different weather conditions, and the unclear appearance of targets caused by different imaging angle in SAR imagery also adds to the difficulty in image interpretation [1]. In SAR oceanography, speckle noise is caused by the ripples and other objects on ocean that produce scattering. Oceanographic SAR images have been extensively used in diverse applications such as monitoring marine traffic, oil pollution, fishing activities, border surveillance, and crime control [2]. So the effect of speckle noise has to be reduced, in order to enhance the subsequent processes such as identification and characterization of the marine targets. A variety of methods have been attempted for despeckling by many researchers [3].

Speckle removal techniques in SAR images can be categorized into two, namely, the noncoherent or multi-look integration and adaptive image restoration or post-image formation methods. Multi-look integration involves averaging of the independent look images obtained from the azimuth spectra of the radar image [4]. In cases where multi-look capability is not possible, the second category of speckle removal technique will be followed, which has been discussed and experimented in this work. The post-image formation method performs filtering to remove the grainy appearance of speckle noise in image, which produces a smoothing appearance. Extensive smoothing of speckle effect leads to information loss. Therefore, a trade-off between noise reduction and information depth is required. Speckle noise reduction can be carried out either in the spatial domain or in the transform domain. Filtering in spatial domain consists of moving a window over the image and substituting the value of the center pixel of the subimage with the value derived mathematically. This process is continued until the entire image has been covered. In transform domain filtering, coefficients are thresholded for noise removal. Hybrid approaches combining both spatial domain and transform domain methods can also be followed to reduce speckle effect. An oceanographic speckle-affected SAR image with grainy appearance is shown in Fig. 6.1.

Different types of adaptive and nonadaptive filters can be used for speckle removal in spatial domain. Mean and median filters are the two simple filters used for despeckling SAR images, which involves just replacing the center pixel of the subimage with the mean and median values of the subimage, respectively. Mean filter produces image smoothing but smears the edges resulting in poor edge preservation. Median filter shows better edge preservation than mean filter but fails

Fig. 6.1 Speckled SAR image

to preserve the single pixel-wide features. These two filters are not adaptive in nature, and hence other adaptive filters such as Lee filter, Kuan filter, Frost filter, and adaptive forms of mean and median filters were used for despeckling. Lee filter [5, 6] performs despeckling by minimizing the mean square error based on the assumption that the mean and variance of the center pixel in the moving window are equal to the local mean and variance of the input image. Kuan filter [7, 8] is similar to Lee filter but with a different weighting function. Before despeckling, it transforms the multiplicative noise model into an additive model. Frost filter [9, 10] replaces the center pixel value in a moving window with the weighted sum of the values within that window by assuming that the noise is of multiplicative form with stationary statistics. Adaptive mean filter is a smoothing filter that operates based on the local statistics of the corrupted image. Adaptive median filter identifies the noisy pixels in the image and replaces those pixels with the median values of the neighbor pixels in the moving window. Anisotropic diffusion [11] is a Partial Derivative Equation (PDE)-based despeckling that uses a threshold function to preserve the edges by preventing diffusion across edges. In transform domain, wavelets [12], ridgelets [13], and curvelets [14] have also been used for denoising. Transform domain denoising methodology generally involves identifying the coefficients that satisfy a thresholding criterion and then reconstructing the thresholded coefficients to obtain the denoised image. Hybrid approach of speckle removal combining the effects of spatial filters and wavelets has been employed using Brute Force Thresholding [15]. In this technique, the best threshold is identified by thresholding the wavelet coefficients in an iterative process.

6.2 Proposed Approach

The proposed approach (Fig. 6.2) is designed to identify the best despeckling method among the various efforts, namely, filtering in spatial domain, processing in transform domain, and employing hybrid methodologies based on both the Peak
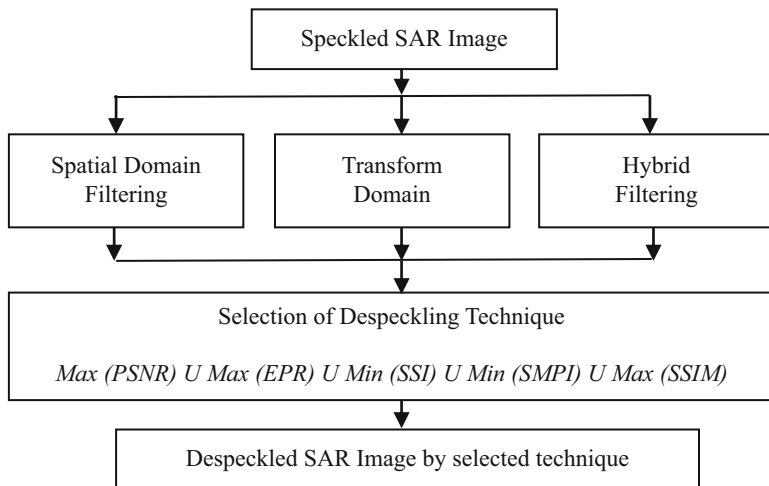
Fig. 6.2 Flow diagram for proposed approach

Signal-to-Noise Ratio (PSNR), Edge Preservation Ratio (EPR), Speckle Suppression Index (SSI), Speckle Suppression and Mean Preservation Index (SMPI), and Structural Similarity Index (SSIM). The trade-off that needs to be employed in speckle removal is preserving as much edges as possible with removal of the grainy appearance. Certain techniques give better PSNR while some others preserve the edges as well as their neighbors leaving a few pixels unattended. To address both the problems, the method which produces high PSNR along with high EPR, low SSI, low SMPI, as well as SSIM value close to one will be chosen as the best speckle removal method. The different despeckling techniques are summarized in the following sections.

6.3 Despeckling Usıng Spatial Domain Filters

In spatial domain, various adaptive and nonadaptive filters [3] are used for speckle removal. Some of the filters such as median filter, Frost filter, Kuan filter, Lee filter, adaptive mean and adaptive median filters, and anisotropic diffusion have been experimented to remove speckle noise in SAR images.

6.3.1 Median Filter

The median filter is similar to a mean filter. It replaces the center pixel value of the subimage window with the median of the neighboring pixel values [16].

It is given by

$$D(x, y) = \underset{(s,t) \in S_{xy}}{\text{median}} \{I(s, t)\} \tag{6.1}$$

where $D(x,y)$ represents the denoised image obtained by computing the median values in the subimage window of size $m \times n$ defined by S_{xy} in the noisy image $I(s,t)$. The median value is indeed one of the pixel values in the neighborhood, and therefore, it does not create any new improbable pixel values when it mounts on an edge. Hence, the median filter is much better than the mean filter in preserving sharp edges. Similar to low-pass filtering, the median filter smooths the image and thereby reduces noise.

6.3.2   Lee Filter

The Lee filter [5] is better at preserving edges by using the local statistics. It uses the variance as a basis to preserve the details. If the subimage has low variance, then it performs smoothing, but not for the area with high variance. Hence, it is capable of preserving details in both low- and high-contrast regions in an adaptive manner. Mathematically, it is given by

$$D(x, y) = m + W * (\text{Cp} - m) \tag{6.2}$$

where $D(x, y)$ is the pixel value after filtering, m is the mean intensity of the filter window, Cp is the center pixel value, and W is the filter window given by

$$W = 1 - \frac{C_u^2}{C_i^2} \tag{6.3}$$

where $C_u = \sqrt{\frac{1}{\text{ENL}}} = \sqrt{\frac{1}{\left(M_i / S_i\right)}}$ and $C_i = \frac{s}{m}$ in which ENL is the Equivalent Number of Looks, M_i is the mean value of the image, S_i is the standard deviation of the image, m is the mean intensity of the filter window, and s is the standard deviation of intensity within the window. The Lee filter has a drawback that the speckle noise near edges cannot be removed effectively.

6.3.3   Kuan Filter

The Kuan filter [7] is a local linear minimum mean square error filter under multiplicative noise. It transforms the multiplicative speckle model into the additive

model. It is advanced than Lee filter with no approximation involved. This filter is similar to Lee filter with same mathematical formulation but uses a different weighting function defined as

$$W = \frac{1 - \left(\frac{C_u^2}{C_i^2}\right)}{1 + C_u^2} \qquad (6.4)$$

6.3.4 Frost Filter

Frost filter [9] is devised to smooth out noise while retaining the edges or shape features in the image, using an exponentially damped convolution kernel by assuming multiplicative noise and stationary noise statistics. It makes use of an adaptive kernel which is based on the local statistical features. The Frost filter replaces the center pixel of the subimage with the weighted sum of the values in the moving kernel of size $n \times n$. The Frost filter follows the formula given by

$$D = \sum_{n \times n} c\alpha e^{-\alpha|t|} \qquad (6.5)$$

where $\alpha = \left(\frac{4}{n\bar{\sigma}^2}\right)\left(\frac{\sigma^2}{m^2}\right)$, c is the normalized constant, m is the local mean, σ is the local variance, $\bar{\sigma}$ is the image coefficient of variation value, n is the moving kernel size, and $|t| = |X - X_0| + |Y - Y_0|$ gives the position. The amount of exponential damping is determined by the damping factor, α. Larger damping values give better edge preservation with less smoothing, and smaller values smooth more. A damping value of zero produces output similar to that of low-pass filtered one.

6.3.5 Adaptive Mean Filter

The adaptive mean filter [16] is a linear filter that performs image smoothening in an adaptive manner. It depends upon the local statistics of the image.

It is given by

$$D(x, y) = I(x, y) - \frac{\sigma_n^2}{\sigma_L^2}[I(x, y) - m] \qquad (6.6)$$

where $D(x, y)$ is the pixel value after filtering, $I(x,y)$ is the pixel value of the noisy image, m is the local mean, σ_L^2 is the local variance, and σ_n^2 is the noise variance over the entire image. If $\sigma_n^2 = \sigma_L^2$, the filter returns the local mean, thus averaging out the noise. If $\sigma_n^2 \lll \sigma_L^2$, this is probably the location of an edge and the filter returns the

edge value, $g(x,y)$. When $\sigma_n^2 > \sigma_L^2$, then negative gray values are obtained which is likely a problem. When $\sigma_n^2 = 0$, there is no noise, so $g(x,y)$ is returned.

6.3.6 Adaptive Median Filter

The adaptive median filter [16] segregates the pixels as noisy or noiseless, by comparing each pixel in the image to its surrounding neighbor pixels. A pixel is labeled as impulse noise when it is different from a majority of its neighboring pixels and also when it is not structurally in range with its similar pixels. Those pixels that are labeled noisy are then replaced by the median value of the neighborhood pixels. The algorithm is given as follows:

```
Level A:   A1 = I_med - I_min
           A2 = I_med - I_max
           if A1 > 0 AND A2 < 0, go to level B
           else increase the window size
           if window size < S_max, repeat level A
           else output I_xy
Level B:   B1 = I_xy - I_min
           B2 = I_xy - I_max
           if B1 > 0 AND B2 < 0, output I_xy
           else output I_med
```

where I_{min} is the minimum gray-level value in S_{xy}, I_{max} is the maximum gray-level value in S_{xy}, I_{med} is the median of gray levels in S_{xy}, I_{xy} is the gray level at coordinates (x, y), and S_{max} is the maximum allowed size of S_{xy}. It adaptively changes the size of the neighborhood during operation, and the threshold for comparison is also adjustable.

6.3.7 Anisotropic Diffusion

Anisotropic diffusion or Perona–Malik diffusion [11] is a technique that performs denoising without discarding the valid image content, such as the edges, lines, or other details that are important for interpreting the image, by using Partial Differentiable Equations (PDE). In this method, a family of consecutively more and more blurred images is generated from an image by diffusion process based on a criterion. Each resulting image is a combination of the original image and a filter that depends on the local content of the original image. Therefore, anisotropic diffusion is a nonlinear and space-variant transformation of the original image. Anisotropic diffusion is defined as

$$\frac{\partial I}{\partial t} = \text{div}\left(c\left(x, y, t\right) \nabla I\right) = \nabla c . \nabla I + c\left(x, y, t\right) \Delta I \qquad (6.7)$$

where Δ denotes the Laplacian, ∇ denotes the gradient, div(...) is the divergence operator, and $c(x,y,t)$ is the diffusion coefficient that controls the rate of diffusion. It is chosen as a function of image gradient and is given by

$$c\left(\|\nabla I\|\right) = e^{-(\|\nabla\|/k)^2} \tag{6.8}$$

and

$$c\left(\|\nabla I\|\right) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2} \tag{6.9}$$

The diffusion coefficient depends on the constant K that controls the sensitivity to edges and is usually chosen experimentally or as a function of noise in the image. When the image gradient is higher than K, it is considered to represent an edge; and when it is lower than K, density diffusion takes place uniformly. For our implementation, K value is chosen to be 10 as Perona and Malik [10] suggested that K value can be chosen based on visual experiments.

6.4   Despeckling Using Multiresolution Transforms

In transform domain, denoising of SAR images is performed based on reconstructing the image by using only the transformed coefficients that satisfy a threshold condition. Wavelets, ridgelets, and curvelets are used for denoising.

6.4.1   Despeckling Using Wavelet Transform

Wavelet denoising aims at removing the noise content of the signal, without destroying the vital characteristics of the signal, mindless of its frequency content. As wavelet transform is localized in both time and scale, it enhances the edge preservation while denoising. In this work, denoising has been done on SAR images using two techniques, namely, wavelet-based thresholding and using Wavelet-Based Absolute Moments.

Despeckling Using Wavelet-Based Thresholding Wavelet thresholding [17] is a simple technique that compares each wavelet coefficient against a threshold value. If the coefficient is smaller than threshold, it is considered to be noisy and set to zero, else it is kept or modified. Applying inverse wavelet transform on the thresholded coefficients leads to reconstruction with the indispensable signal characteristics and less noise.
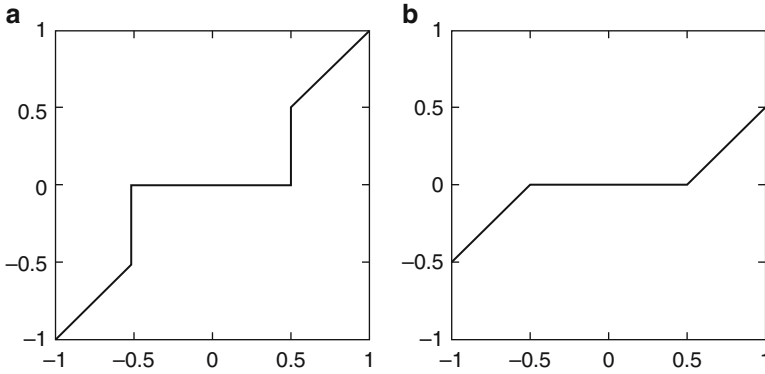
Wavelet thresholding involves the following three steps: a linear discrete wavelet transform, nonlinear shrinking of wavelet coefficients (thresholding step), and a linear inverse wavelet transform. Wavelet transform is applied with Haar basis [18]. Before thresholding, the threshold value has to be determined, which can be done using the wavelet shrinkage rules such as SURE Shrink which is an adaptive threshold selection using principle of Stein's Unbiased Risk Estimate (SURE), Visu Shrink also called as universal threshold, Risk Shrink or minimax thresholding, and Heuristic SURE Shrink which is a mixture of SURE and universal thresholding.

The universal threshold is a better estimate for soft thresholding when the number of samples is large. In SURE Shrink, the threshold is determined by decreasing the Stein's Unbiased Risk Estimate that depends on shrinkage function and multiresolution level assuming a unit noise variance. The SURE rule is weak when there is intense sparsity of the wavelet coefficients. In such cases, the noise contributed to the SURE by the multiple coordinates where there is zero signal swamps the information contributed to the SURE by the few coordinates where there is nonzero signal. Consequently, SURE Shrink uses a hybrid scheme, the Heuristic SURE. The concept of this hybrid scheme is that the universal threshold generates more loss than SURE for dense cases but smaller loss for sparse situations. So the threshold is set to universal threshold in dense situations and to SURE Shrink-based threshold in sparse situations. In simple words, if the signal-to-noise ratio is very small, the SURE estimate leads to noisy results. If such a situation is detected, then universal threshold is used. The minimax thresholding makes use of a fixed threshold to produce minimax performance that performs best in the worst possible case of a problem. In denoising using minimax thresholding, the minimax estimator recognizes the minimum of the maximum mean square error attained for the worst function in a given set.

A small threshold may yield an output close to that of the input, but it may still be noisy. On the other hand, a large threshold produces an output with a large number of zero coefficients. This results in a smooth signal. Concentrating too much on smoothness may destroy details and cause blur and artifacts. If we prefer to use a threshold such as the minimax threshold or the universal threshold that depends only on size of the input, n, then the threshold, λ, can be calculated initially, and then perform noise removal using the three-step procedure. Whereas if we prefer to use a data-adaptive threshold such as the threshold selected by SURE that depends not just on n but also on the data, U, then the threshold value must be determined first before thresholding.

Hard and soft thresholding [19] are employed to the wavelet decompositions of the input image. The shrinkage functions for the hard and soft thresholdings are used as given in [20] and pictorially represented in Fig. 6.3a, b, respectively.

Denoising Using Wavelet Absolute Moments Denoising of SAR images was also done using Wavelet-Based Absolute Moments [21]. The higher-order absolute moments of the noise residual are calculated as features for classification. In this work, WAM is used for calculation of denoised wavelet coefficients of image which are then reconstructed to obtain the denoised image.

Fig. 6.3 Shrinkage function for (**a**) hard thresholding and (**b**) soft thresholding. (Courtesy: http://what-when-how.com/embedded-image-processing-on-the-tms320c6000-dsp/wavelet-demising-image-processing-part-1)

6.4.2 Despeckling Using Ridgelet Transform

The ridgelet transform can effectively handle the line singularities. It utilizes the radon transform to convert the line singularities into point singularities. Wavelet transform is then applied to deal with the point singularities. Ridgelets are more powerful in handling directions and are highly anisotropic. Discrete ridgelet transform can be a near-optimal method for denoising [13], since it provides near-ideal sparsity representation for objects with both edges and smooth regions. The ridgelet image denoising involves partitioning the image into overlapping blocks and applying ridgelet transform on each block. The ridgelet coefficients obtained are thresholded with thresholding techniques similar to that applied for wavelet-based thresholding. The denoised image is then obtained by taking inverse ridgelet transform of the thresholded coefficients.

6.4.3 Despeckling Using Curvelet Transform

Curvelets provide multi-scale object representation. Wavelet transforms require large number of coefficients for image representation when compared to curvelets which has the advantage of representing the image with sparse coefficients. The edges are generally curved rather than straight; hence, efficient representation cannot be given by ridgelets alone. Curvelets can decompose an image at different scales and angles to represent more curvilinear objects and has better edge preservation. Denoising images using curvelet-based thresholding [14] involves computing of curvelet transform of the image, thresholding the obtained coefficients, and taking inverse curvelet transform of the thresholded coefficients. In this method also, the aforementioned thresholding techniques are used.

6.5 Despeckling Using Hybrid Approaches

The speckle noise in SAR images is removed using a combination of spatial filters and wavelet transform to provide a hybrid approach for despeckling. Here, the method called Brute Force Thresholding is used. This is an iterative process, followed to identify the best threshold value. Median filter [16] performs better at preserving the edge features. Savitzky–Golay filter [15] is a smoothing filter that performs much better than the standard averaging filters.

It tends to filter the significant part of the signal's high-frequency content along with noise. It is mathematically given as

$$
F_j = \sum_{i=-\frac{p-1}{2}}^{\frac{p-1}{2}} C_i \ f_{j+i}, \qquad \frac{p-1}{2} \le j \le n - \frac{p-1}{2} \tag{6.10}
$$

where C_i is the convolution coefficient, p is the set of convolution coefficients, n is the number of data points, f_{j+i} is the observed value, and F_j is the filtered value. The convolution coefficient varies based on the window size. For our implementation, since window size is taken as 5×5, the convolution coefficient used is $[-3\ 12\ 17\ 12\text{–}3]^T$. This filter efficiently preserves the appropriate high-frequency contents of the signal. The process involves three images, namely, the input image (A), Savitzky–Golay filtered image (B), and the median filtered image (C). Our approach has experimented with different options for B and C, such as Laplacian filter, adaptive mean, adaptive median filters, and approximation component of image.

6.6 Performance Measure

The performance evaluation of the work is done by calculating statistical measures like Peak Signal-to-Noise Ratio (PSNR), Edge Preservation Ratio (EPR), Speckle Suppression Index (SSI), Speckle Suppression and Mean Preservation Index (SMPI), and Structural Similarity Index (SSIM) [22, 23]. PSNR is used to give a quantitative evaluation of the denoising process. It is the measure of peak error in the data. It is fast and easy to implement. The PSNR value must be high for a better denoising algorithm. A higher PSNR indicates that the denoising is of good quality. PSNR is usually calculated as

$$
\text{PSNR} = 10 \log_{10} \left(255^2 / \text{MSE} \right) \tag{6.11}
$$

where

$$
\text{MSE} = \frac{1}{MN} \sum_{M} \sum_{N} \left[s(x, y) - d(x, y) \right]^2 \tag{6.12}
$$

where $s(x,y)$ is the speckled image and $d(x,y)$ is the despeckled image of size $M \times N$. EPR measures the ability to maintain details of the image. It is given by

$$\text{EPR} = \frac{\text{num}\,(M_r \cap M_d)}{\text{num}\,(M_r)} \tag{6.13}$$

where M_r is the edge map from reference image and M_d is the edge map from noisy image. If the detail structures and low-level texture features are retained after denoising, the EPR value will be high showing good edge preservation. SSI and SMPI are used to analyze the image quality and structural preservation. SSI is the coefficient of variation of the denoised image normalized by that of the speckled image.

It is given by

$$\text{SSI} = \frac{\text{SD}(d)}{M(d)} \cdot \frac{M(s)}{\text{SD}(s)} \tag{6.14}$$

where $\text{SD}(d)$ and $\text{SD}(s)$ are the standard deviation of the despeckled image and speckled image, respectively. $M(d)$ and $M(s)$ are the mean of the despeckled image and speckled image, respectively. SSI value is typically less than 1. Greater speckle suppression is obtained when the SSI value is smaller. SMPI is used as a measure of despeckling when the filter overestimates the mean value. The despeckling is better when the SMPI value is low. SMPI is given as

$$\text{SMPI} = K \ \times \ \frac{\text{SD}(d)}{\text{SD}(s)} \tag{6.15}$$

where K is obtained by

$$K = 1 + |M(s) - M(d)| \tag{6.16}$$

where $\text{SD}(d)$ and $\text{SD}(s)$ are the standard deviation of the despeckled image and speckled image, respectively. $M(d)$ and $M(s)$ are the mean of the despeckled image and speckled image, respectively. SSIM is used to quantify the quality of the despeckled image, given by

$$\text{SSIM} = \frac{(2M(s)M(d) + a)\,(2C\,(s,d) + b)}{\left(M(s)^2 + M(d)^2 + a\right)\left(\text{SD}(s)^2 + \text{SD}(d)^2 + b\right)} \tag{6.17}$$

where $\text{SD}(d)$ and $\text{SD}(s)$ are the standard deviation of the despeckled image and speckled image, respectively. $M(d)$ and $M(s)$ are the mean of the despeckled image and speckled image, respectively, and $C(s,d)$ is the covariance of the speckled and despeckled image. The constants a and b are used to stabilize the division when the denominator is weak.

6.7   Experimental Results and Discussion

Denoising of SAR images have been implemented using spatial filters, wavelet transform, ridgelet transform, curvelet transform, and hybrid approaches.

6.7.1   Denoising Results Using Spatial Domain Filters

In spatial domain, despeckling has been done using seven different spatial filters, namely, median filter, Lee filter, Kuan filter, Frost filter, adaptive versions of mean and median filter, and anisotropic diffusion.

Two European Remote-Sensing Satellite (ERS) SAR images collected from Google are used for experimentation. The ERS SAR images have a spatial resolution of 25 m with area coverage of 10,000 km^2. These images are acquired at 5.3 GHz frequency and linear vertical (VV) polarization [24]. The despeckled results of these two SAR images, using the abovementioned filters, are shown in Fig. 6.4, and the performance of these filters measured using PSNR, EPR, SSI, SMPI, and SSIM has been tabulated in Table 6.1.
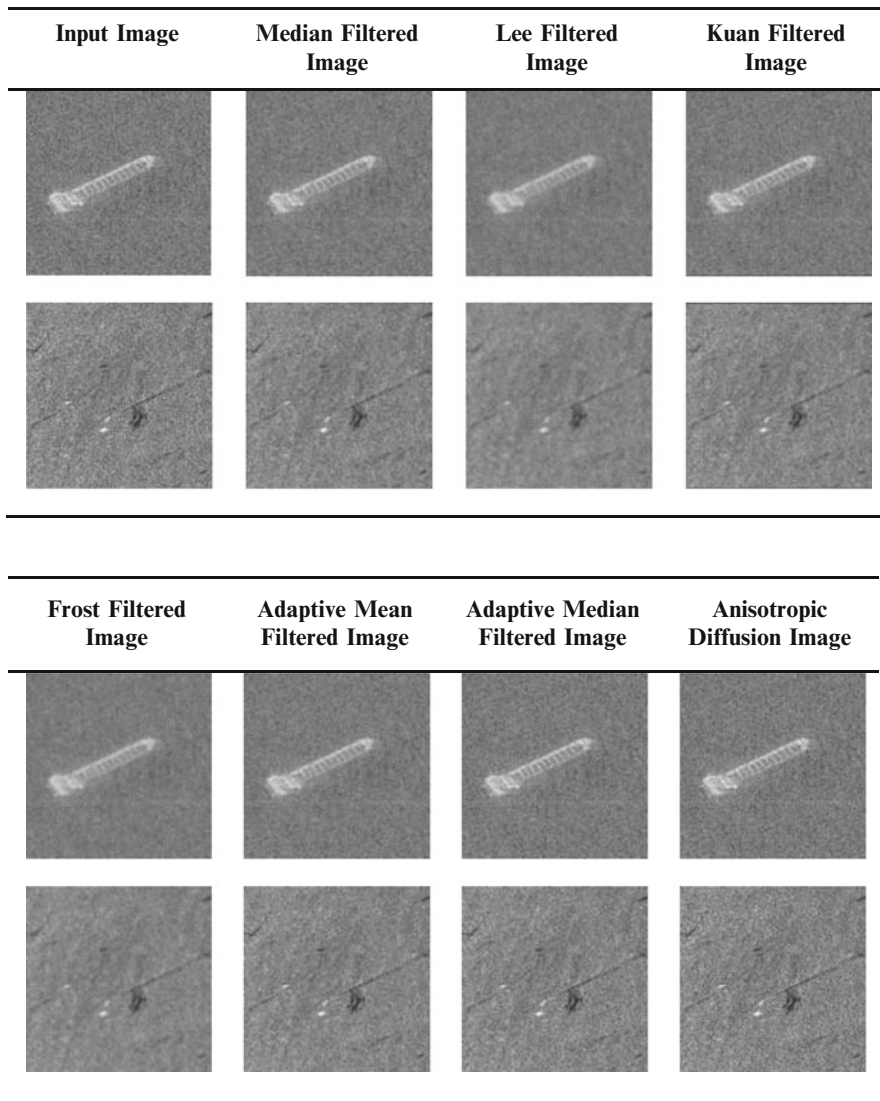
From Table 6.1, it is evident that anisotropic diffusion produces better despeckling than the other six filters. This is supported by the PSNR, EPR, and SSIM values. Even though the SSI and SMPI values of anisotropic diffusion are not better than those of Lee, it is also within the acceptable range. It is capable of removing the noise without smoothing out the edges. So anisotropic is concluded to be the best among the other filters that are experimented.

Unlike the other filters, it applies the diffusion law on the intensity values of the pixels to smooth the uniform regions in the image and retain the edge structures. A threshold function is used to prevent diffusion across edges. It is capable of producing better despeckling, irrespective of the level of clutter in the SAR image.

6.7.2   Denoising Results Using Multiresolution Transforms

In transform domain, denoising using wavelet transform, ridgelet transform, and curvelet transform was experimented.

In wavelet domain, the despeckling of SAR images was performed using wavelet-based thresholding and Wavelet-Based Absolute Moments (WAM). Both hard and soft thresholding were performed using shrinkage rules such as SURE Shrink, Visu Shrink, Heuristic SURE Shrink, and minimax thresholding. This work was done using Discrete Wavelet Transform and Stationary Wavelet Transform using Haar function.

Input Image	Median Filtered Image	Lee Filtered Image	Kuan Filtered Image

Frost Filtered Image	Adaptive Mean Filtered Image	Adaptive Median Filtered Image	Anisotropic Diffusion Image

Fig. 6.4 Despeckling results using spatial filters

The results for speckle removal using DWT-based soft thresholding and hard thresholding are displayed in Figs. 6.5 and 6.6, respectively. The despeckled images of SWT-based soft and hard thresholding are displayed in Figs. 6.7 and 6.8, respectively. The performance measures for DWT- and SWT-based thresholding are tabulated in Tables 6.2 and 6.3, respectively. Table 6.2 shows that hard thresholding with minimax shrinkage produces good performance measures such as high PSNR, high EPR, low SSI, low SMPI, and SSIM close to unity, when compared to the other

Table 6.1 Performance measures for despeckling using spatial filters

| Performance measures	PSNR (in dB)		EPR		SSI		SMPI		SSIM	
	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
Median	29.4323	39.5869	0.5838	0.4968	0.8137	0.6207	0.8872	1.1959	0.732	0.5252
Lee	26.6693	34.3338	0.2268	0.3962	0.707	0.4404	0.7073	0.4416	0.4313	0.2436
Kuan	26.896	34.4566	0.3658	0.5109	0.7099	0.4474	0.714	0.4497	0.4584	0.2839
Frost	28.211	27.9317	0.64	0.4828	0.8443	0.6277	1.2044	0.9774	0.709	0.5118
Adaptive mean	33.467	46.8912	0.9787	0.9953	0.8277	0.7511	0.8907	0.7914	0.8874	0.9129
Adaptive median	33.4263	45.928	0.8236	0.8177	0.9215	0.8274	0.9785	1.0296	0.9092	0.8157
Anisotropic diffusion	47.1745	51.5063	0.9855	0.9947	0.9785	0.9895	0.9797	0.99	0.9957	0.999
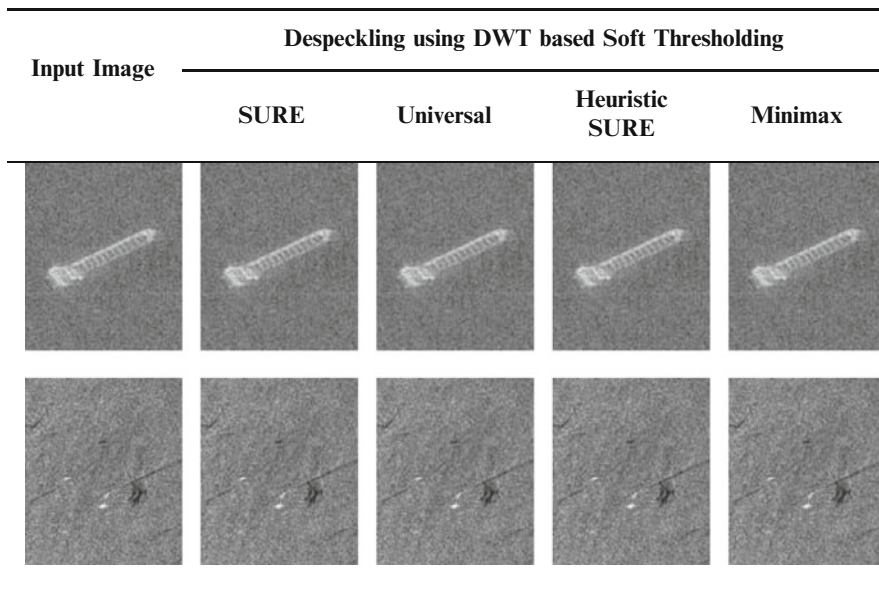
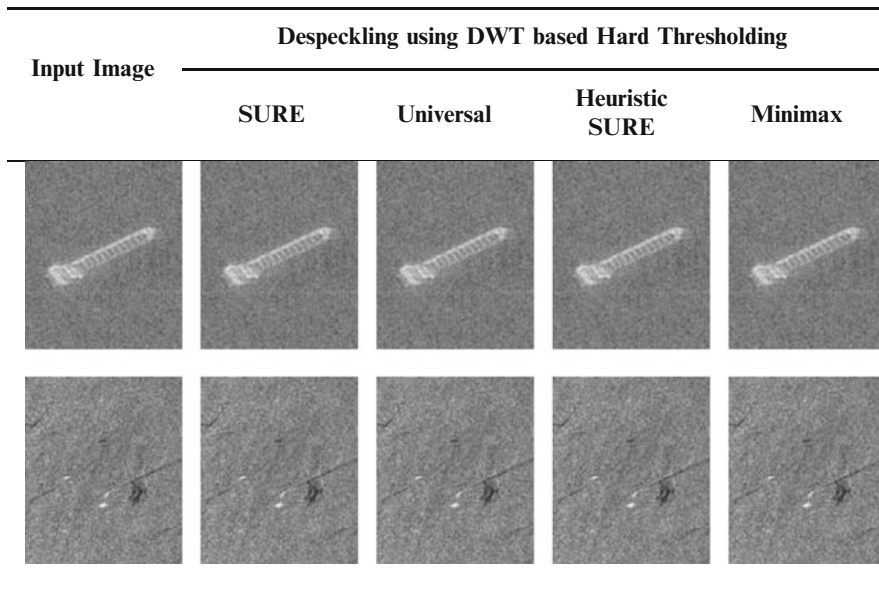**Fig. 6.5** Despeckling using DWT-based soft thresholding

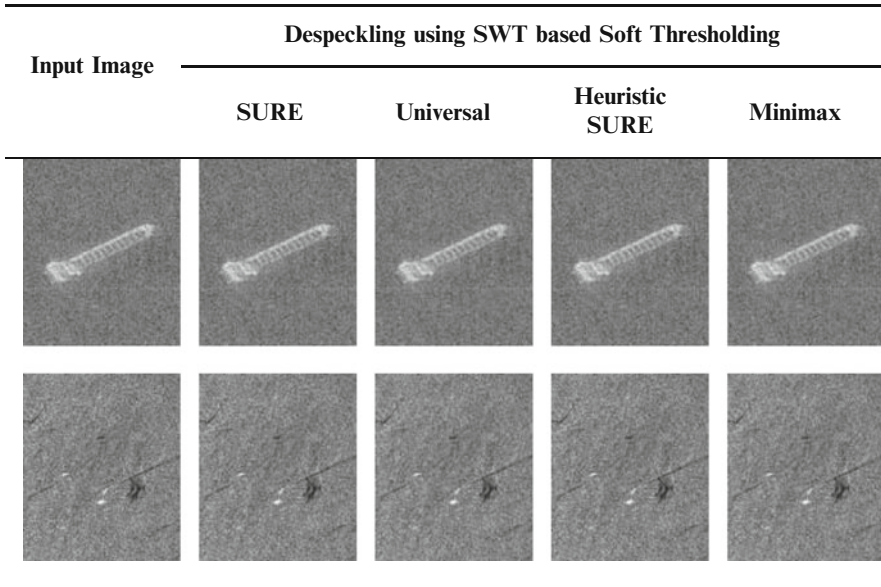
Fig. 6.6 Despeckling using DWT-based hard thresholding

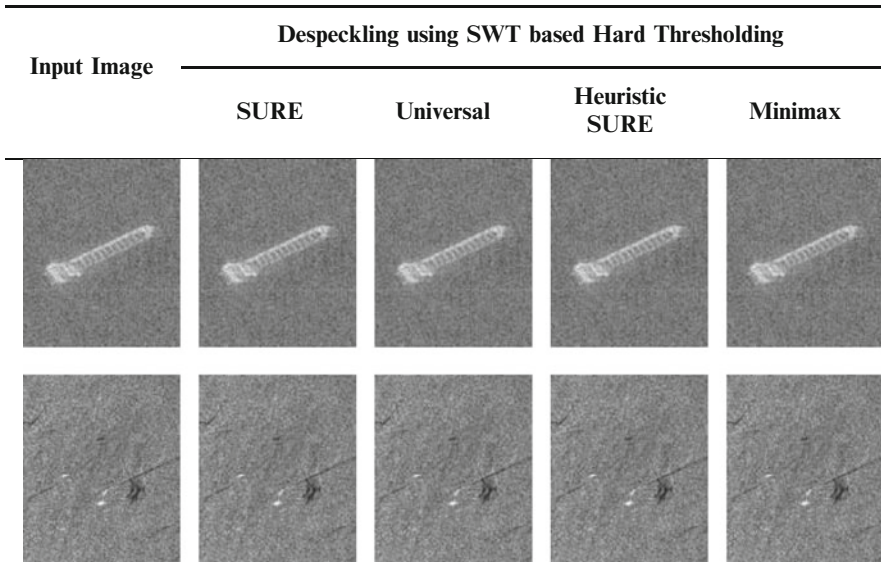| Input Image | Despeckling using SWT based Soft Thresholding | | | |
| | SURE | Universal | Heuristic SURE | Minimax |

Fig. 6.7 Despeckling using SWT-based soft thresholding

| Input Image | Despeckling using SWT based Hard Thresholding | | | |
| | SURE | Universal | Heuristic SURE | Minimax |

Fig. 6.8 Despeckling using SWT-based hard thresholding

Table 6.2 Performance measures for despeckling using Discrete Wavelet-Based Thresholding

		Performance measures									
		PSNR (in dB)		EPR		SSI		SMPI		SSIM	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	85.8817	NA	NA	NA	NA	NA	NA	NA	NA	NA
Universal	Hard	45.0293	47.9664	0.9702	0.9906	1.0037	1.0017	1.1354	1.0771	0.9943	0.9985
	Soft	38.5024	37.4834	0.9287	0.9671	1.0619	1.0879	1.1053	1.0991	0.9653	0.9816
Heuristic SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	85.8817	NA	NA	NA	NA	NA	NA	NA	NA	NA
Minimax	Hard	49.1573	52.1075	0.9864	0.9968	1.0037	1.0017	1.1354	1.0771	0.9943	0.9985
	Soft	41.3001	41.5536	0.9606	0.9829	1.0619	1.0879	1.1053	1.0991	0.9653	0.9816

Table 6.3 Performance measures for despeckling using Stationary Wavelet-Based Thresholding

		Performance measures									
		PSNR (in dB)		EPR		SSI		SMPI		SSIM	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Universal	Hard	47.2221	49.3798	0.9816	0.9945	1.0058	1.0026	1.0767	1.0561	0.9961	0.9989
	Soft	40.0065	38.5861	0.9661	0.9809	1.0753	1.0994	1.0941	1.1076	0.9786	0.9867
Heuristic SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Minimax	Hard	50.62	53.3366	0.9899	0.9973	1.0024	1.0011	1.0665	1.0504	0.9984	0.9996
	Soft	42.501	41.4161	0.9796	0.9894	1.0545	1.0695	1.0688	1.0791	0.9885	0.9933

Input Image	WAM based Despeckled Images	
	DWT	SWT
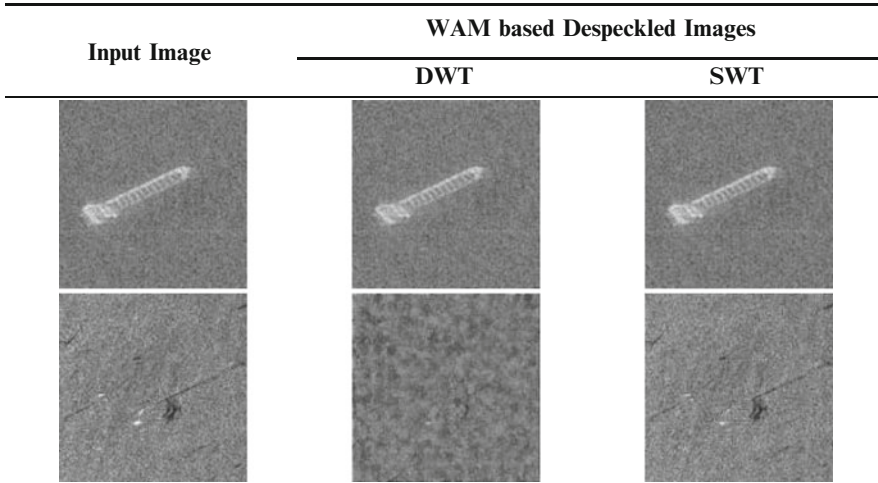
Fig. 6.9 WAM-Based Despeckling

thresholding methods. A similar situation prevails in SWT-based thresholding also as seen in Table 6.3. Therefore, hard thresholding with minimax shrinkage performs better for DWT- and SWT-based thresholding.

For some images, the threshold value obtained is very small in the order of 10^{-14}, and hence, no coefficient was identified as a noisy coefficient. Subsequently, no denoising operation was also carried out. For those images, the performance measures are not computed and the cases are indicated as not applicable (NA).

Despeckling using Wavelet-Based Absolute Moments (WAM) was implemented using both Discrete Wavelet Transform and Stationary Wavelet Transform. The despeckled images are displayed in Fig. 6.9.
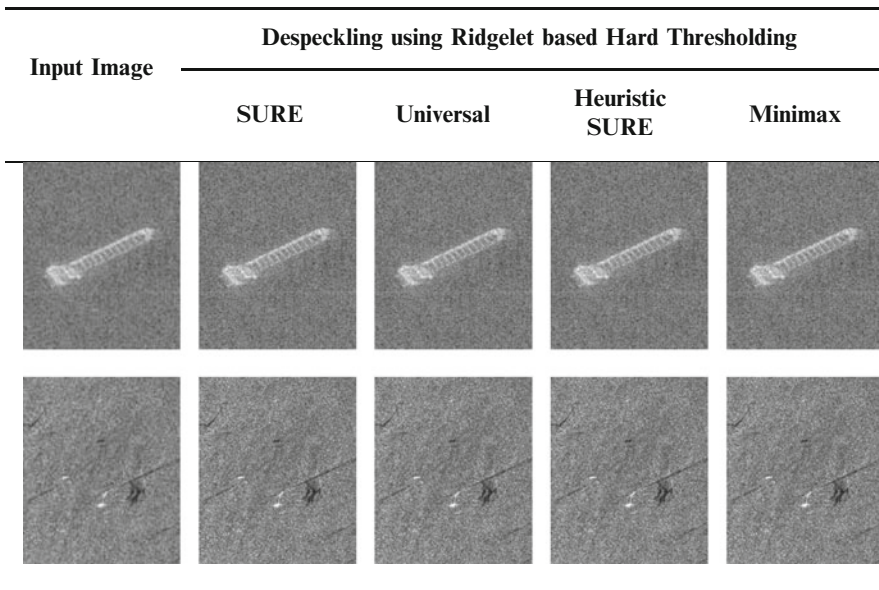
The obtained performance measures for WAM-based despeckling using DWT and SWT are tabulated in Table 6.4. The denoising results obtained using "db8" wavelet transform for WAM despeckling technique was better than other wavelet transforms but was not appreciable when compared to the previous techniques. Discrete Wavelet Transform is not suitable for WAM-based despeckling since most of the information content is degraded as evident in Fig. 6.10, leading to distortions in reconstructed image.

Therefore, comparing the despeckling techniques followed in wavelet transform domain, Stationary Wavelet Transform-based hard thresholding with minimax shrinkage function provides better speckle removal than Discrete Wavelet Transform Shrinkage.

Despeckling using ridgelet transform-based thresholding and curvelet transform-based thresholding was also implemented. Based on the despeckling results so far obtained, ridgelet transform-based thresholding provided better denoising performance than any other technique explained in the previous sections. This is

Table 6.4 Performance measures for WAM-Based Despeckling

Performance measures		DWT	SWT
PSNR (in dB)	Image 1	12.8728	26.2397
	Image 2	7.4564	16.6737
EPR	Image 1	0.4463	0.4552
	Image 2	0.4623	0.3936
SSI	Image 1	3.2843	1.0633
	Image 2	5.2806	2.9214
SMPI	Image 1	4.4239	1.0637
	Image 2	6.5949	4.556
SSIM	Image 1	0.0983	0.6496
	Image 2	0.018	0.1313

Fig. 6.10 Despeckling using ridgelet-based hard thresholding

supported by the obtained performance measure values tabulated in Table 6.5, and the corresponding denoised images are displayed in Fig. 6.11. Also in this case, soft thresholding performs better, whereas for wavelet-based despeckling methods, hard thresholding gives good results.

The speckle removal using curvelet transform-based thresholding produces good results for minimax thresholding function, but it was not as appreciable as ridgelet-based denoising. Also, curvelet transform-based hard thresholding produced better denoising than soft thresholding, which is inferred from the performance measures tabulated in Table 6.6 (Figs. 6.12 and 6.13).

Table 6.5 Performance measures for ridgelet-based despeckling

		Performance measures									
		PSNR (in dB)		EPR		SSI		SMPI		SSIM	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	59.6577	59.3531	0.9985	0.9998	1.0053	1.0063	1.0116	1.0302	0.9998	0.9999
Universal	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Heuristic SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Minimax	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	66.4145	65.3126	0.9999	1	1.0016	1.0022	1.002	1.0129	1	1

Input Image	Despeckling using Ridgelet based Soft Thresholding			
	SURE	**Universal**	**Heuristic SURE**	**Minimax**
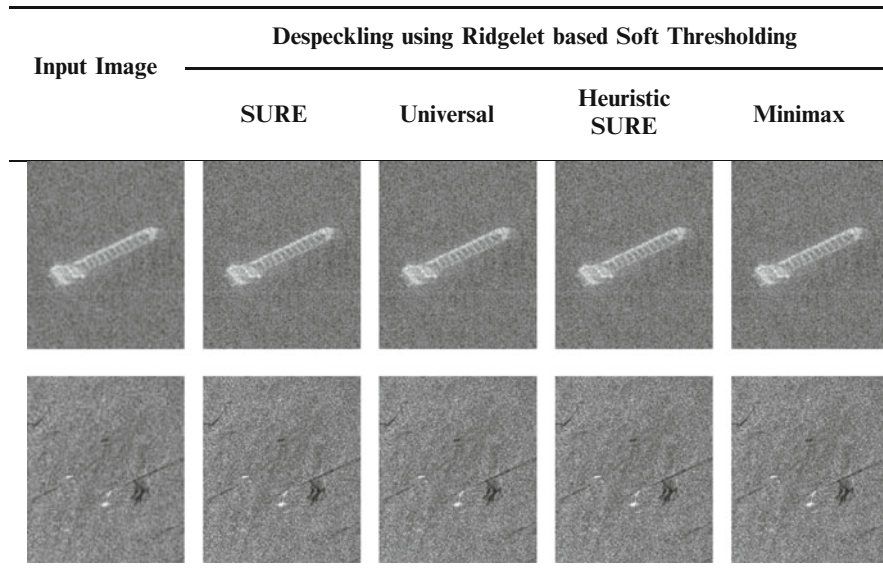
Fig. 6.11   Despeckling using ridgelet-based soft thresholding

6.7.3   Denoising Results Using Hybrid Approach

The despeckling of SAR images using hybrid approach is based on Brute Force Thresholding. This technique was experimented for various combinations of filtered images obtained using Laplacian filter, adaptive mean, adaptive median filters, and approximation component of the speckled image instead of the Savitzky–Golay filtered image and the median filtered image in the specified algorithm. The despeckled images obtained for the combination of Savitzky–Golay filter with other filters like Laplacian filter, adaptive mean filter, adaptive median filter, and median filter in Brute Force Thresholding are shown in Fig. 6.14.

The despeckled images obtained for the combination of approximation component of image with others are shown in Fig. 6.16. The corresponding performance measures PSNR, EPR, SSI, SMPI, and SSIM values of the abovementioned despeckling technique for various combinations are tabulated in Table 6.7, which shows that the performance measures for despeckling using Brute Force Thresholding with approximation subband of image and adaptive median filtered image are high when compared to other combinations. It is also evident that the combination of the approximation component of image with any one of the filtered images among median filtered image, Laplacian filtered image, and adaptive median filtered image also produces good results similar to the combination of approximation subimage and adaptive mean filtered image.

Table 6.6 Performance measures for curvelet-based despeckling

| | | Performance measures | | | | | | | | | |
		PSNR (in dB)		EPR		SSI		SMPI		SSIM	
		Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Universal	Hard	43.7246	45.1772	0.9547	0.9813	1.0616	1.0025	1.0626	1.1018	0.9595	0.9977
	Soft	37.2636	36.3641	0.9007	0.9428	1.0068	1.1017	1.0070	1.0028	0.9912	0.9763
Heuristic SURE	Hard	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Soft	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Minimax	Hard	49.7873	49.5051	0.9803	0.9931	1.0024	1.0008	1.0029	1.0012	0.9965	0.9992
	Soft	39.6946	39.0373	0.9428	0.968	1.0473	1.0729	1.0474	1.0731	0.9772	0.9878
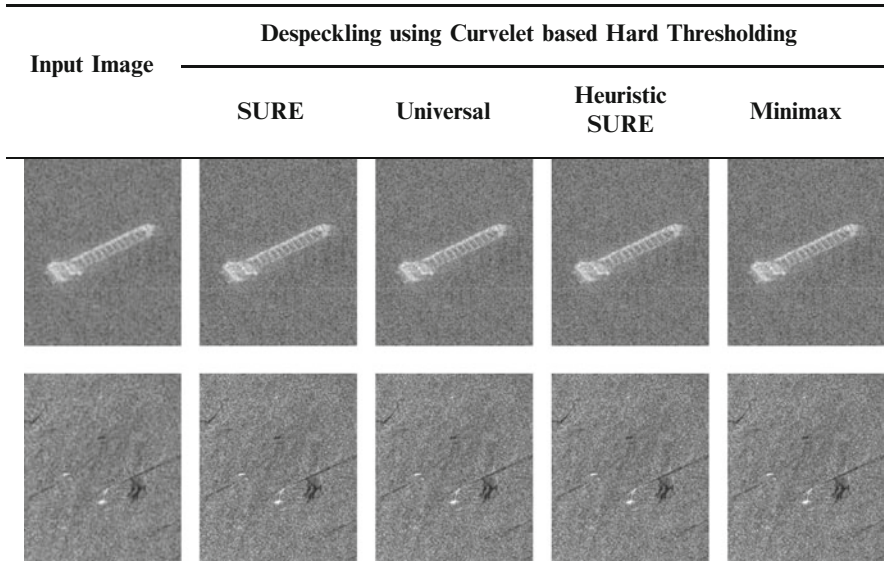
Fig. 6.12 Despeckling using curvelet-based hard thresholding
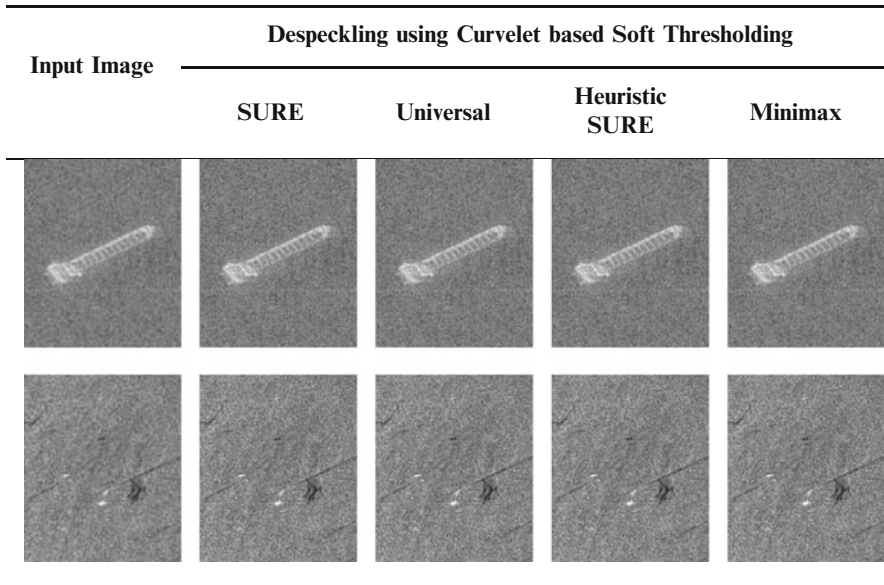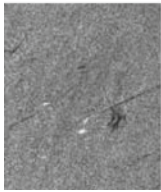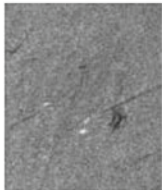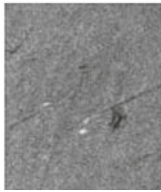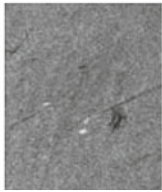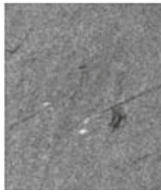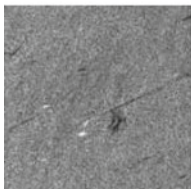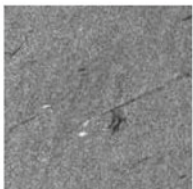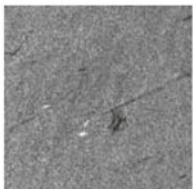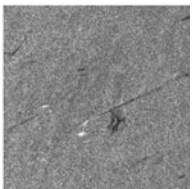
Fig. 6.13 Despeckling using curvelet-based soft thresholding

Input Image	Despeckling using Brute Force Thresholding			
	Savitzky Golay Filter & Laplacian Filter	Savitzky Golay Filter & Adaptive Median Filter	Savitzky Golay Filter & Adaptive Mean Filter	Savitzky Golay Filter & Median Filter

Fig. 6.14 Despeckling using Brute Force Thresholding with combination of Savitzky–Golay Filter

Input Image	Anisotropic Diffusion Image	SWT based Hard Thresholding - Minimax thresholding	Ridgelet Transform based Soft Thresholding - Minimax Thresholding

Fig. 6.15 Comparison of despeckling methods performed in different domains

Input Image	Target Detection in Speckled Image	Target Detection in Despeckled Image
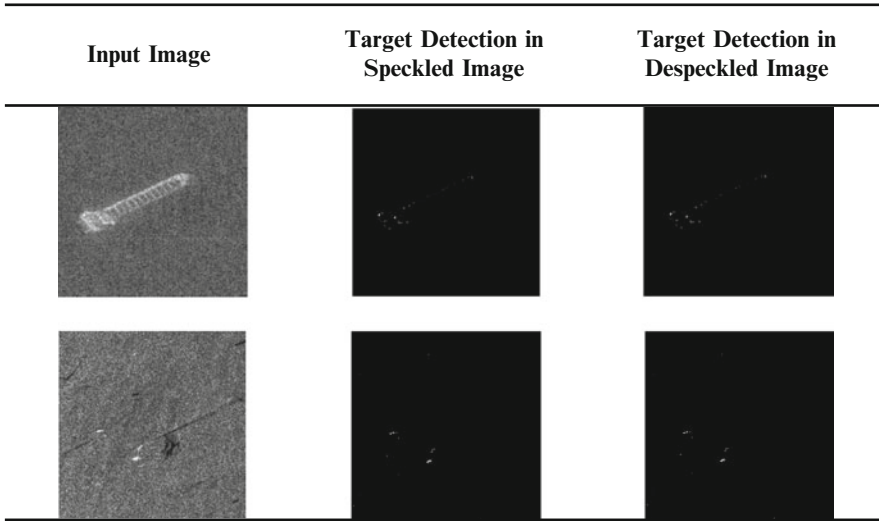
Fig. 6.16 Target identification

Input Image	Despeckling using Brute Force Thresholding			
	Approximation Component & Laplacian Filter	Approximation Component & Adaptive Median Filter	Approximation Component & Adaptive Mean Filter	Approximation Component & Median Filter
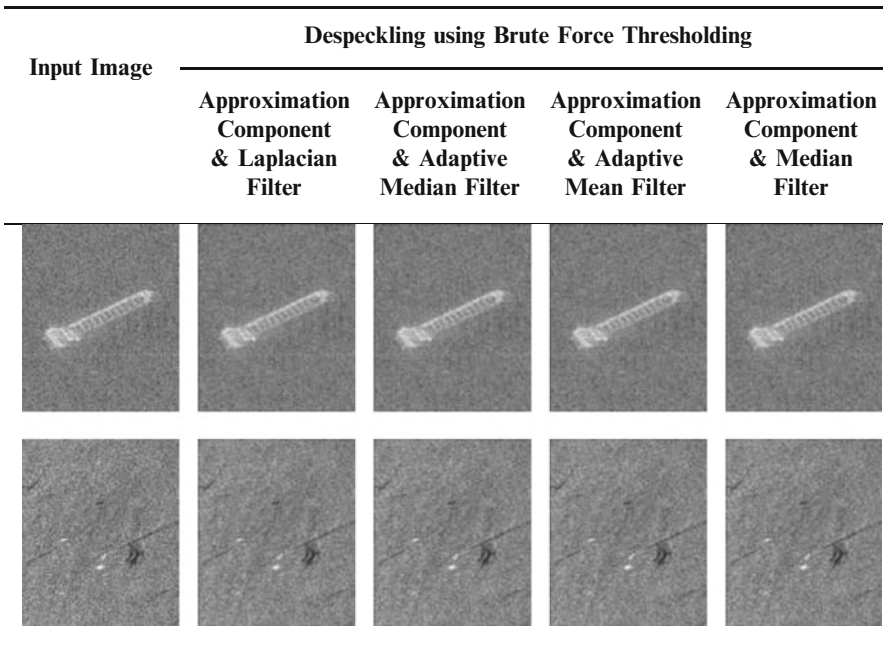
Table 6.7 Performance measures for despeckling using Brute Force Thresholding

	PSNR (in dB)		EPR		SSI		SMPI		SSIM	
	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2	Image 1	Image 2
Savitzky–Golay filter and Laplacian filter	32.2899	27.8128	0.8861	0.8652	1.2436	1.2077	0.8552	0.9699	0.8893	0.8891
Savitzky–Golay filter and adaptive median filter	32.7996	27.863	0.9	0.869	1.2419	1.2083	0.856	0.9698	0.9005	0.8402
Savitzky–Golay filter and adaptive mean filter	32.8017	27.8639	0.9002	0.869	1.2423	1.2083	0.8559	0.9698	0.9005	0.8403
Savitzky–Golay filter and median filter	32.7613	27.8128	0.8661	0.8652	1.2436	1.2077	0.8552	0.9699	0.8893	0.8391
Approximation component and Laplacian filter	32.2899	27.8128	0.8661	0.8652	1.2436	1.2077	0.8552	0.9699	0.8893	0.8391
Approximation component and adaptive median filter	32.8026	27.8644	0.8998	0.869	1.2428	1.2083	0.8556	0.9698	0.9005	0.8402
Approximation component and adaptive mean filter	32.8003	27.8637	0.8998	0.869	1.2424	1.2083	0.8558	0.9698	0.9005	0.8403
Approximation component and median filter	32.7619	27.8584	0.8986	0.8688	1.2433	1.2082	0.8553	0.9698	0.8996	0.8401

6.8 Comparison of Despeckling Methods Performed in Different Domains

On comparing the despeckling results of all the techniques experimented, despeckling using ridgelet transform-based soft thresholding with minimax shrinkage function provides good denoising with high PSNR as well as high EPR. The despeckled images of anisotropic diffusion, SWT-based hard thresholding with minimax shrinkage, and ridgelet transform-based soft thresholding with minimax shrinkage function, i.e., the best from every domain that yields an effective result among all the experimented techniques, are displayed in Fig. 6.15. The proposed approach acts as a foolproof approach, since the best despeckling technique is identified not only by maximum PSNR but also with other vital performance measures such as maximum EPR, minimum SSI, minimum SMPI, and SSIM close to unity.

Target identification with reduced false alarms is enhanced by appropriate denoising of the SAR images. Figure 6.16 shows the target detected images with speckled and despeckled images as input, respectively, using the CFAR detection algorithm [25].

6.9 Conclusion

Target detection in SAR images has spread its arms into a variety of vital applications. Due to inherent backscattering of the incoming radar waves from man-made objects, it is possible to discriminate the targets from background clutters. However, the disturbances such as the backscattered signal, sea clutter, and unfavorable weather conditions affect the target detection process. So the SAR images must be denoised efficiently, before subjecting to any subsequent processes such as Region of Interest extraction for target detection or segmentation, for effective image interpretation. Though ridgelet transform yielded the best result in our experimentation, the proposed approach suggests to perform denoising in different domains and select the denoising technique which gives better PSNR and simultaneously preserves the edges.

References

1. Hu C, Ferro Famil L, Kuang G (2013) Ship discrimination using polarimetric SAR data and coherent time-frequency analysis. Remote Sens 5:6899–6920
2. Margarit G, Barba Milanés JA, Tabasco A (2009) Operational ship monitoring system based on synthetic aperture radar processing. Remote Sens 1:375–392
3. Jaybhay J, Shastri R (2015) A study of speckle noise reduction filters. Signal Image Process Int J 6(3):71–80
4. Lee J-S (1981) Speckle analysis and smoothing of synthetic aperture radar images. Comput Graph Image Process 17(1):24–32
5. Lee J-S (1980) Digital image enhancement and noise filtering by use of local statistics. IEEE Trans Pattern Anal Mach Intell 2(2):165–168
6. Lee J-S (1983) A simple speckle smoothing algorithm for synthetic aperture radar images. IEEE Trans Syst Man Cybern 13(1):85–89
7. Kuan DT, Sawchuk AA, Strand TC, Chavel P (1985) Adaptive noise smoothing filter for image with signal-dependent noise. IEEE Trans Pattern Anal Mach Intell 7(2):165–177
8. Kuan DT, Sawchuk AA, Strand TC, Chavel P (1987) Adaptive restoration of images with speckle. IEEE Trans Acoust Speech Signal Process 35(3):373–383
9. Frost VS, Stiles JA, Shanmugan KS, Holtzman JC, Smith SA (1981) An adaptive filter for smoothing noisy radar images. In: Proceedings of the IEEE, pp 133–155
10. Frost VS, Stiles JA, Shanmugan KS, Holtzman JC (1982) A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Trans Pattern Anal Mach Intell 4(2):157–166
11. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell 12(7):629–639
12. Fukuda S, Hirosawa H (1998) Suppression of speckle in synthetic aperture radar images using wavelets. Int J Remote Sens 19(3):507–519
13. Chen GY, Kègl B (2007) Image denoising with complex ridgelets. Pattern Recogn 40:578–585
14. Starck J-L, Candès EJ, Donoho DL (2002) The curvelet transform for image denoising. IEEE Trans Image Process 11(6):670–684
15. Gir R, Jain L, Rai R (2015) Despeckling of SAR images using wavelet based iterative brute force thresholding with median filter and Savitzky-Golay filter. Int J Comput Appl 120(20):0975–8887
16. Rafael C, Gonzalez RE (2008) Woods: digital image processing, 2nd edn. Prentice Hall, Upper Saddle River
17. Xiao F, Zhang Y (2011) A comparative study on thresholding methods in wavelet based image denoising. Procedia Eng 15:3998–4003
18. Fukuda S, Hirosawa H (1999) Smoothing effect of wavelet-based speckle filtering: the Haar basis case. IEEE Trans Geosci Remote Sens 37:1168–1172
19. Donoho DL (1995) Denoising by soft-thresholding. IEEE Trans Inf Theory 41:613–627
20. Soman KP, Ramachandhran KI, Resmi NG (2004) Insight into wavelets: from theory to practice, 3rd edn. PHI Learning, New Delhi
21. Goljan M, Fridrich J, Holotyak T (2006) New blind steganalysis and its implications. In: Proceedings SPIE, Electronics imaging, security, steganography and watermarking of multimedia contents VIII, vol 6072, San Jose, CA, pp 1–13
22. Rajamani A, Krishnaveni V (2014) Performance analysis survey of various SAR image despeckling techniques. Int J Comput Appl 90(7):0975–8887
23. Biradar N, Dewal ML, Rohit M, Gowre S, Gundge Y (2016) Blind source parameters for performance evaluation of despeckling filters. Int J Biomed Imaging 2016:3636017
24. https://crisp.nus.edu.sg/ers/ers.html
25. Gao G, Liu L, Zhao L, Shi G, Kuang G (2009) An adaptive and fast CFAR algorithm based on automatic sensoring for target detection in high–resolution SAR images. IEEE Trans Geosci Remote Sens 47(6):1685–1697

Chapter 7
Fused Segmentation Algorithm for the Detection of Nutrient Deficiency in Crops Using SAR Images

V. P. Ananthi

Abstract The main aim of this chapter is to segment nutrient deficiency in crop images using fuzzy sets (FSs) theory. Fuzziness exists in images as the quantized level of brightness in each pixels. Processing of such uncertain images can be efficiently handled by using fuzzy sets, particularly IFSs. Before initiation of segmentation, crop images taken by satellite are fused to reduce uncertainty in the captured images. Finally, the fused image is processed for segmentation of deficiency in crop images using clustering method based on interval valued intuitionistic fuzzy sets (IVIFSs) with new distance function. Quantitatively, the segmented images are evaluated using precision-recall, ROC curves, and measure for structural similarity index, and their results are compared with results of existing methods. Performance measures reveal that the proposed method seems to segment deficiency better than other comparable methods. Segmentation of nutrient deficiency using the proposed method helps the agriculturist in differentiating various types of disease thereby estimating the rate of fertilization for the crop concerned, which improves economy.

7.1 Introduction

A central cohesive source for Indian economy is agriculture. Practicing traditional techniques in agriculture fields are expensive and doesn't contain surveillance system for instant change detection in crops/plants. Hence monitoring of plants/crops from their initial stage of growth till its maturity for the detection of pest attack/abnormal growth or disease due to deficiency of nutrients is necessary to

V. P. Ananthi (✉)
Department of Mathematics, Gobi Arts and Science College, Erode, Tamil Nadu, India

increase productivity. Monitoring crops by manual perception and by using conventional monitoring system through hand cameras is a tedious job. In order to overcome such difficulty, SAR monitoring system will be very useful [1]. SAR imaging can be used even in unrevealed weather. Processed SAR images can be exploited to check growth condition and pest/disease attack for the concerned field. In case growth retardation is identified as due to pest attack, the agriculturist can take immediate action for controlling it, thereby improving the yield. One of the important art of science in sensing the area of interest is related to Earth's surface of revolution. Such sensed information is acquired in various formats, namely, spectral, spatial, and temporal images. SAR provides information even in areas of cloud cover [2].

Radar sensors are sensitive to various features of imaging object and which works with the different wavelengths. For example, it is sensitive to the size of the region of interest (ROI). As in the case of radar, SAR imaging depends upon the ratio of the emitted and reflected signal for detecting ROI [3]. The signal/data acquired from SAR should be sent for preprocessing which accounts geometric distortions and distortions due to irregular illumination from satellite/aircraft. After the correction of the abovementioned distortion, noise due to reflection of signal from object features is reduced using speckle filtering. Multi-looking, radiometric calibration, and de-speckling are the other preprocessing tasks need to be done [4]. Edge features are enhanced using adaptive filtering. Similarly, radiometric resolutions can be improved by multi-looking.

India's indigenous RISAT had been launched to work even at unsupportable weather on April 26, 2012 which uses SAR technique to monitor agriculture and disaster. SAR images acquired from airborne system antenna/aperture electronically. It can be fixed in aircraft, which emits single beam to ROI, and its resolution is proportional to the ratio of bandwidth of the pulse used for sensing. SAR had been introduced by a mathematician Wiley in 1951 that are helpful in monitoring environmental applications. One of the important applications of SAR images is characterization or identification of crop images [5]. Crops imaged under complex environment may acquire uncertainty due to poor illumination or climatic condition.

Generally, optical imaging system renders images from high spatial to spectral resolution. But remote sensing system generates images with higher quality in the form of spectral or spatial resolution. Hence, sensors to be devised for rendering the images with the resolution of their fused form. One needs to verify the query whether the properties of both the form of resolution are the same; the answer is no. For instance, SPOT PAN images produce panchromatic of high resolution, and LANSAT TM renders multispectral images of low resolution [6, 7]. Image fusion is the technique that renders information from both the images by merging them with high clarity. Before initiating the process, resampled multispectral images are feed as an input for fusion to cope up with the properties of high-resolution images. Image fusion is a technique to merge these images and provide a high-resolution image. In such cases the multispectral images are resampled to have same resolution as that of panchromatic images before performing fusion process. Image fusion techniques are therefore useful in integrating a higher spectral with higher spatial resolution image [7].

In order to get images without uncertainty, multiple shots can be captured for a single projected field or single shot by different sensors and examined by fusing them along with the improvement of edge contrast of the image. Detection of changes can be identified by segmenting affected regions from the healthy regions. In order to get images without uncertainty, multiple shots can be captured for a single projected field and examined by fusing them along with the improvement of contrast of the image. Detection of changes can be identified by segmenting abnormal regions from the normal regions. Handling of uncertain data can be managed using fuzzy set theory. Fuzzy set (FS) has been initiated by Zadeh in 1965 [8]. In recent decades, clustering of uncertain data have been done perfectly by employing fuzzy set theory. In 1986 the concept of intuitionistic fuzzy set (IFS) has been introduced by Atanassov [9]. IFS is the generalized FS; its hesitation degree plays a leading role in uncertainty analysis. In 2012, Chaira studied the detection method for retrieving edges from the images by reducing uncertainty in defining intensity of image pixels [10]. Many researches have paved the way for analyzing uncertainty in images through type-2 fuzzy sets [11], whose membership grades are themselves fuzzy.

Interval-valued intuitionistic fuzzy set (IVIFS) has been utilized in various digital fields such as decision-making in hospitals regarding state of patient, time series forecasting in unsupportive weather, ROI extraction, and so on. Uncertainty is designed as the interval in type-2 FS and IVIFS; the interval is large, and then uncertainty should be modeled in such a way to minimize it. Atannassov and Gargov have been introduced IVIFS in 1989 [12].

Segmentation is a generic process which separates an image into meaningful regions or objects, and the segmented objects are called as the foreground, and the rest of the image is the background [13]. That is, an image can be divided into regions (set of pixels) which are related in some way. For example, they may have similar brightness or color that indicates that they belong to the same object. Typically, it is difficult to design a computer vision system that is entirely autonomous. Numerous segmentation techniques have been proposed in the past decades, and some basic approaches of segmentation are thresholding [14], edge detection [15], region-based segmentation, clustering, and matching. The major goal of these techniques was to segment an image perfectly. Perfect image segmentation is assigning each pixel in an image to the correct object. But this may be impossible because a pixel may span the "real" boundary of objects such that it partially belongs to two (or even more) objects, which lead to oversegmentation (pixels belonging to the same object are grouped to different classes) or undersegmentation (pixels belonging to different objects are grouped to the same class). Most of the recent segmentation methods attempt to assign a pixel to a single segment that is adequate for most of the applications. The present chapter concentrates on clustering-based segmentation technique for classifying the affected area of the image.

Possibilistic partition matrices contain entries that are constructed using the possibility of each pattern to a cluster and the column, respectively; sum constraint does not necessarily sum to 1 over any column [16]. Fuzzy partition matrix contains a membership value of each pattern based on the relative distance between the

pattern and a fuzzy prototype. FCM is comparatively vulnerable to outliers but hard to initialize [17]. Even though PCM is robust to outliers, it is sensitive to the selection of parameters and initialization, and this might consequently produce reduced clusters than the pre-intended ones. If the clusters of the dataset are relatively close to each other, then PCM may generate overlapping clusters or miss some clusters. Distance measure between data points is another important element of clustering algorithm. Euclidean distance measure is sufficient in case of samples in the dataset have same physical units, but it misleads even in simple cases. In case of data vectors that are not immediately comparable, measures should be introduced based on the domain knowledge. So, a new distance measure is utilized to get rid of such problem during clustering. Mostly clustering technique has a great importance in image processing, data mining, and in the field of recognizing patterns. But these methods may be affected with fuzziness due to intensity level variation in the considered image. Main focus of the present work is in removing uncertainty during segmentation of nutrient deficient fields.

Fuzzy c-means (FCM) algorithm is the traditional fuzzy clustering technique and is robust in the absence of noise. It also depends upon the fuzzier utilized for the construction of membership matrix. Fuzzifier is also uncertain in many circumstances as like that of distance function [18]. In this chapter, the segmentation of SAR images using interval-valued intuitionistic fuzzy sets for detection of nutrient deficiency has been analyzed.

Images of crops are obtained using sensors of satellite or by artificial sensor and are stored in a data for processing. After all the preprocessing work, the images of the same ROI acquired are fused to get an uncertain fused image with information from both the fusing images by using IFS fusion technique. After merging the data, color and texture features are extracted, and they are utilized in distance measure which in turn used in detection of membership value of the membership matrix. The newly created membership matrix are the resulting clusters, and from them one can detect deficient region.

Section 7.2 seeds some preliminary ideas about images in digital form and how it is represented in fuzzy and their extended forms. Proposed method of clustering is explained in Sect. 7.3. Experimental results and analysis on the results are presented in Sect. 7.4. Finally conclusion has been drawn in Sect. 7.5.

7.2 Preliminary Ideas

7.2.1 Image

A two-dimensional function $I(i, j)$ with i and j representing plane (spatial) coordinates is called an image [13]. The intensity of the image I at any pair of coordinates (i, j) is known as amplitude of I at that level.

7.2.1.1 Digital Image

An image is called a digital image, if i, j and the amplitude values of I are finite and discrete values. It is framed by countably finite number of components named as image elements, pixels, picture elements, and pels; among them, pixel is the most frequently used term. Pixels are ordered as a rectangular array with number of columns and rows denoting the width and height of the image. Hence, an image matrix I having $P \times Q$ array of pixels is symbolized as

$$I = \begin{bmatrix} I(1,1) & I(1,2) & \ldots & I(1,Q) \\ I(2,1) & I(2,2) & \ldots & I(2,Q) \\ \vdots & \vdots & \vdots & \vdots \\ I(P,1) & I(P,2) & \ldots & I(P,Q) \end{bmatrix},$$

where $0 \leq I(i,j) \leq L-1$, $1 \leq i \leq P$, $1 \leq j \leq Q$, and $L = 2^k$, $k = 1, 2, \ldots$ denote number of bits.

Resolution is the spatial scale of the image pixels. For instance, an image of 2416×1356 pixels represented with resolution of 300 pixels per inch (ppi) would be a real-world image of size $11'' \times 8.5''$. ppi is related to pixel arrays and dpi is related to dots per inch concerned with printer resolution are the terms utilized to clarify resolution.

Type of image is concerned with type of intensity that is utilized for each pixel. For example, in an image, if there is an intensity variation from black (darkest gray) to white (lightest gray), then it is termed as black and white image. Similarly, intensity variation of red, green, and blue colors from the darkest to the lightest shade with several compositions produces a color image. Black-and-white color images are the most primary types of digital images and are well known as grayscale and RGB images, respectively. Bits are utilized for defining intensity value in digital images. One single bit contains a single binary value, either 0 or 1. There are 256 possible values with intensity, ranges from 0 to 255 for an 8-bit image and is represented mathematically as 2^k; $k = 8$ denotes the number of bits. That is, there are $2^1 = 2$ and $2^8 = 256$ possible values for 1-bit image and 8-bit image, respectively. Most commonly standard digital imaging equipments utilize 8-bit intensity range. Grayscale image has a single 8-bit intensity range. In color images, each color has 8-bit intensity range, that is, totally they have $3 \times 8 (= 24)$ bit intensity. Experimental results in present study are executed on 8-bit images, therefore, throughout the chapter $L = 256$.

7.2.2 Emergence of Fuzzy Sets in Digital Image

Zadeh [8] proposed FS theory; since then it is utilized in numerous fields. Even though FSs are applied in various fields with single valued membership function, which cannot have the ability to convey the evidence of the confirmation, resistance, and hesitation of a particular element or object under consideration. IFS is an extended FS, introduced by Atanassov [9], which are the evolution of the traditional FSs. IFS considers three aspects of knowledge into account, namely, membership, nonmembership, and hesitation degree. Fuzzy concept of "non-this non-that" can be efficiently described by nonmembership function, which exquisitely demonstrates the vague nature of the objective world. Hence, the IFSs have more flexibility and applicability in treating of fuzzy information and uncertainty than the conventional FSs. Digital images contain fuzziness; for instance, medical images are vague due to poor illumination during imaging, and such uncertainties can be eliminated using IFS. The third (hesitation) degree makes IFSs more flexible than FSs. Chaira [10] showed the capability of IFS in removing uncertainties while detecting image edges. Even though it eliminates uncertainties in digital image with brightness level as fuzzy using the third degree, there arises a query whether the assigned value is appropriate for the considered pixel. So for this case, one can allot a range of values to a pixel instead of a single value with width of the range representing the level of hesitation. Thus for all the cases, it is not desirable to employ IFS having a constant membership and nonmembership degree; instead, a range of values can be utilized [15].

Atanassov and Gargov [12] generalized IFS by defining membership and non-membership as an interval instead of an exact number. Interval-valued intuitionistic fuzzy set (IVIFS) is the only set which perfectly models the abovementioned uncertainty. Nowadays, the extended FSs such as interval-valued fuzzy set (IVFS), IFS, and IVIFS are being utilized in a great extent to handle the problem of uncertain data in numbers of various domains such as medicine, image processing, data mining, and so on. Bustince and Burillo [19] have initiated the way for building IVIFS from IFS in theoretical point of view and fail to explain about the type of uncertainty and how it can be modeled for applications. In 2010, Xu and Wu [18] have extended c-means clustering algorithm for clustering IVIFSs. Virtually real-world problems own certain amount of ambiguity. Among such various problems, present study concentrates on the problems of fusion, segmentation, and noise removal in uncertain images using IFSs and their extended sets. Digital images, which are mappings of natural scenes, are always accompanied by some degree of uncertainty (fuzziness) mainly due to:

 (i) Imprecision of gray values of the pixels;
 (ii) Ambiguity resulting from the image acquisition and mapping mechanism;
(iii) Vague information in the region boundaries.

This concept of fuzziness justifies the development of algorithms based on IFSs and their extended sets for several tasks of image analysis. The following sections

briefly describe how fuzziness occurs in images and how they are suppressed using FSs, IFSs, and IVIFSs.

7.2.2.1 Fuzzy Sets of an Image

In images, naturally there may arise a doubt about fuzziness as what makes an image fuzzy? Numerous image properties like edges, levels of intensity, and so on are fuzzy due to inherent defects in imaging equipment, poor illumination, or acquired image vagueness. Since images can acquire uncertainty within the pixel due to possible multivalued levels of intensity, which is taken as fuzzy throughout this chapter. Since uncertainty arises in the intensity of the image, images are transformed to fuzzy domain, which helps to stretch out the membership function over its whole range of intensity level. Fuzzification of images is primarily done by quantizing and normalizing intensity values.

Based on FSs, $P \times Q$ dimensioned image with L levels of grayness is considered as a $P \times Q$ array of fuzzy singletons concerning the intensity values of the pixels. For more details, one can refer [20]. Thus, a fuzzy image (I_F) of an image (I) is defined as a mapping described below

$$I_F : P \times Q \xrightarrow{I(i,j)=g} G \xrightarrow{\mu_F(g)} [0, 1]$$

where $I(i, j)$ is (i, j)th pixel value of the image I, $I(i, j) = \{g \in G | G = \{0, 1, 2, \ldots, L-1\}\}$ with G holding positive integers representing the gray values of the image. $\mu_F(g)$ denotes the membership of the element g in the image set I defined as

$$\mu_F(I(i, j)) = \frac{g - g_{\min}}{g_{\max} - g_{\min}} \tag{7.1}$$

where $g_{\min}$ and $g_{\max}$ are, respectively, the lowest and highest values of the gray levels of the image I. Hence, the image I in the fuzzy domain is defined by

$$\begin{cases} I_F = \left\{ \langle I(i, j), \mu_F(I(i, j)) \rangle \right\}, \\ 0 \leq I(i, j) \leq L - 1, 0 \leq \mu_F(I(i, j)) \leq 1, \ 1 \leq i \leq P, \ 1 \leq j \leq Q. \end{cases} \tag{7.2}$$

7.2.2.2 Intuitionistic Fuzzy Set of an Image

Image processing based on FS theory is introduced to address the problem of vagueness in image properties such as brightness and edges by patterning membership function. Though FS removes such uncertainty, there arises hesitation during the allotment of quantitative value of brightness to the considered pixel. The principal aim of the extended fuzzy image processing is to reduce gray-level vagueness along

with the elimination of ambiguity in the allotment membership values to those uncertain image pixels. These reasons make one to transfer an image in a FS domain to an IFS domain. Usually, experts may opt or define membership functions in an intuitive way. Therefore, vagueness in images ought to be reduced by considering a membership function without deciding the best among the choice of membership functions. The above discussion motivates the introduction of IFS in images to reduce the hesitation in assigning values to the brightness levels due to lack of knowledge/personal error. Such hesitation makes the membership values of IFS to lie in a range [10].

The membership degree of the image I in IFS domain is calculated as

$$\mu_A(I(i, j)) = 1 - (1 - \mu_F(I(i, j)))^\gamma, \ \gamma \geq 0, \tag{7.3}$$

where the value of μ_F is obtained using Eq. (7.1).

The nonmembership degree is computed as

$$\nu_A(I(i, j)) = \frac{1 - \mu_A(I(i, j))}{1 + \gamma.\mu_A(I(i, j))}. \tag{7.4}$$

The hesitation degree is defined as

$$\pi_A(I(i, j)) = 1 - \mu_A(I(i, j)) - \nu_A(I(i, j)). \tag{7.5}$$

The image I in IFS domain is described as

$$\begin{cases} I_A = \left\{ \langle I(i, j), \mu_A(I(i, j)), \nu_A(I(i, j)), \pi_A(I(i, j)) \rangle \right\}, \ 0 \leq I(i, j) \leq L - 1, \\ 0 \leq \mu_A(I(i, j)) \leq 1, 0 \leq \nu_A(I(i, j)) \leq 1, \\ 0 \leq \pi_A(I(i, j)) \leq 1, \ 1 \leq i \leq P, \ 1 \leq j \leq Q. \end{cases} \tag{7.6}$$

7.2.2.3 Interval-Valued Intuitionistic Fuzzy Set of an Image

Though hesitation in choosing membership function is reduced using IFS, the transformation from IFS domain to IVIFS domain makes one to ask query stated as why IVIFS is essential and what are uncertainty left over in images even after they are processed by IFSs? Even though the allotment of membership values using IFS to the chosen pixel are appropriate, still there emerges doubt about the exactness of the values allotted. These reasons make one to utilize IVIFS to remove the vagueness that are left over by IFSs, which contains interval of values instead of a constant membership value. IVIFS is generated from IFS as in [19] by a mapping ϕ as

$$\phi : IFS(I) \rightarrow IVIFS(I) \tag{7.7}$$

defined as

$$\phi(I_A) = \left\{ \langle I(i, j), M_{\phi(A)}(I(i, j)), N_{\phi(A)}(I(i, j)) \rangle | I(i, j) \in \{0, 1, \ldots, L - 1\} \right\} = I_{\widetilde{F}},$$

such that

1. $M_{\phi(A)L}(I(i, j)) = M_{\widetilde{F}L}(I(i, j)) = \mu_A(I(i, j)) - s \cdot \pi_A(I(i, j)), \ 0 \le s \le \frac{\mu_A(I(i,j))}{\pi_A(I(i,j))}$.
2. $M_{\phi(A)U}(I(i, j)) = M_{\widetilde{F}U}(I(i, j)) = \mu_A(I(i, j)) + \alpha \cdot \pi_A(I(i, j)), \ 0 \le \alpha \le 1$.
3. $N_{\widetilde{F}L}(I(i, j)) = \nu_A(I(i, j)) - t \cdot \pi_A(I(i, j)), \quad 0 \le t \le \frac{\nu_A(I(i,j))}{\pi_A(I(i,j))}$.
4. $N_{\widetilde{F}U}(I(i, j)) = \nu_A(I(i, j)) + \beta \cdot \pi_A(I(i, j)), \ 0 \le \beta \le 1$, with $0 \le \alpha + \beta \le 1$, $0 < \alpha + s \le 1$ and $0 < \beta + t \le 1$.
5. $W_{M\widetilde{F}}(I(i, j)) = M_{\widetilde{F}U}(I(i, j)) - M_{\widetilde{F}L}(I(i, j)) = (\alpha + s) \cdot \pi_A(I(i, j))$.
6. $W_{N\widetilde{F}}(I(i, j)) = N_{\widetilde{F}U}(I(i, j)) - N_{\widetilde{F}L}(I(i, j)) = (\beta + t) \cdot \pi_A(I(i, j))$.

It is clearly seen from the definitions of $W_{M\widetilde{F}}(I(i, j))$ and $W_{N\widetilde{F}}(I(i, j))$ that width of the membership and nonmembership intervals does not surmount intuitionistic fuzzy index (π_A). If $I_A \in FS(I)$, then $\phi(I_A) = I_F$. Hence, the image I in IVIFS domain is constructed as

$$\begin{cases} I_{\widetilde{F}} = \left\{ \langle I(i, j), M_{\widetilde{F}}(I(i, j)), N_{\widetilde{F}}(I(i, j)) \rangle \right\}, \\ M_{\widetilde{F}}(I(i, j)) = [M_{\widetilde{F}L}(I(i, j)), M_{\widetilde{F}U}(I(i, j))], \\ \\ N_{\widetilde{F}}(I(i, j)) = [N_{\widetilde{F}L}(I(i, j)), N_{\widetilde{F}U}(I(i, j))], \\ 0 \le I(i, j) \le L - 1, \ 1 \le i \le P, \ \ 1 \le j \le Q. \end{cases} \quad (7.8)$$

7.2.3 Fuzzy Image Processing

There are three main phases in fuzzy image processing, namely: image fuzzification, modification of membership values, and image defuzzification as described in Fig. 7.1. Fuzzification of image data and defuzzification of the results are the potential steps to work on images with fuzzy techniques. Fuzzy image processing majorly depends upon the middle step, in which membership values are modified. The image data are transformed from gray-level plane to the membership plane by using appropriate fuzzification technique to modify the membership values. Fuzzy integration approach, fuzzy rule-based approach, fuzzy clustering, and so on can be implemented for such modification/patterning membership values [21].

7.2.3.1 Image Fusion

Image fusion is the process of aggregating two or more same or various images into a single compound image that holds significant features from each image. For instance, tracing of ecological situations demands an image that renders atmospheric
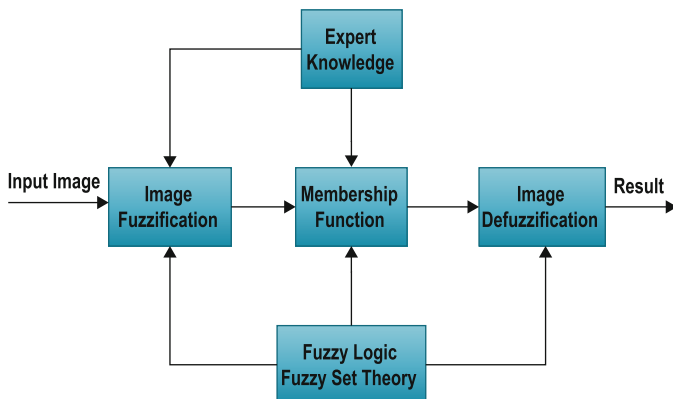
Fig. 7.1 Schematic diagram of fuzzy image processing

conditions for calculating cloud cover in the region along with the level of water vapor at each point. Such demands motivate one to study the field of image fusion.

Fusion situations are versatile depending upon the nature of the field to which it is applied [22], and the four common fusion situations are concisely explained below:

1. Fusion of various images from the same sensor: Fusion of several channels on the same satellite or of multi-echo images in magnetic resonance imaging (MRI).
2. Fusion of various images from different sensors: Fusion of positron emission tomography (PET) and MRI images or earth remote sensing (ERS) satellite and Satellite Pour l'Observation de la Terre (SPOT) images.
3. Condensing information of various elements from the same image: Fusion by several operators, classifiers, sensors, and so on; each depends upon unrelated image characteristics.
4. Fusion of images and another source of information: Fusion of digital atlas with the data provided by experts.

The first and second type of fusion situation which are helpful in medical, military, and forensic purposes has been analyzed in present study. Images acquired by different modality or different illumination are fused to get an image that renders information from both the images, which helps the analyst to classify them quickly in short span of time. For example, instead of rendering two or more images of brain image to a medical practitioner to check the existence of tumor, it is better to provide a single image that has enough information from all the images.

Fusion technique used in this chapter under fuzzy domain is shown below with image A and image B being SAR image and gray-scaled near-infrared (NIR) images, whose fused image is shown in image C. For more details refer [23, 24]. Fusion mechanics is represented in the following Fig. 7.2.
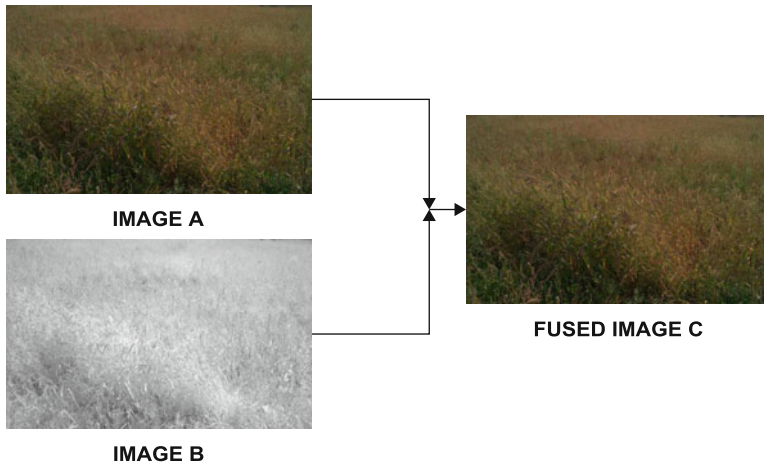
Fig. 7.2 Fused output by the proposed technique

### *7.2.4   Identification of Disease Infected Crop SAR Images*

If sunlight struck the healthy plant parts, near-infrared (NIR) and green band are reflected high compared to blue and red bands that are sensed by satellites like Landsat and Sentinel [5, 25]. Due to high reflectance of green light, the healthy vegetation looks green through our vision; see Fig. 7.3.

Therefore segregation of healthy and diseased crop fields can be identified using normalized difference vegetation index (NDVI) [26].

$$NDVI = (NIR-Red)/(NIR+Red)$$

Normalized values vary from -1 to 1, in which the values near -1 represent water, 0 represents soil, and near 1 represents healthier plant regions. From the values one can identify the fields that are having less than 1 and near to zero are unhealthy. After detection of such unhealthy region, the need for ROI is imaged multiple times for further processing. Unhealthiness of the crop may be due to improper irrigation, imbalance fertilization, and environmental condition. For example, Fig. 7.5 renders the ROI of a fields that has both healthy and non-healthy regions.

The color changed region in Fig. 7.4 may be a unhealthy region or the crop at the stage of harvesting. In this figure, it is found that they are disease-affected crop field. Crop production and yield detection have been recently done by Francis et al. in 2018 [27]; in order to improve the productivity in between such detection, search for change detection in crop images needs to be done from the field.
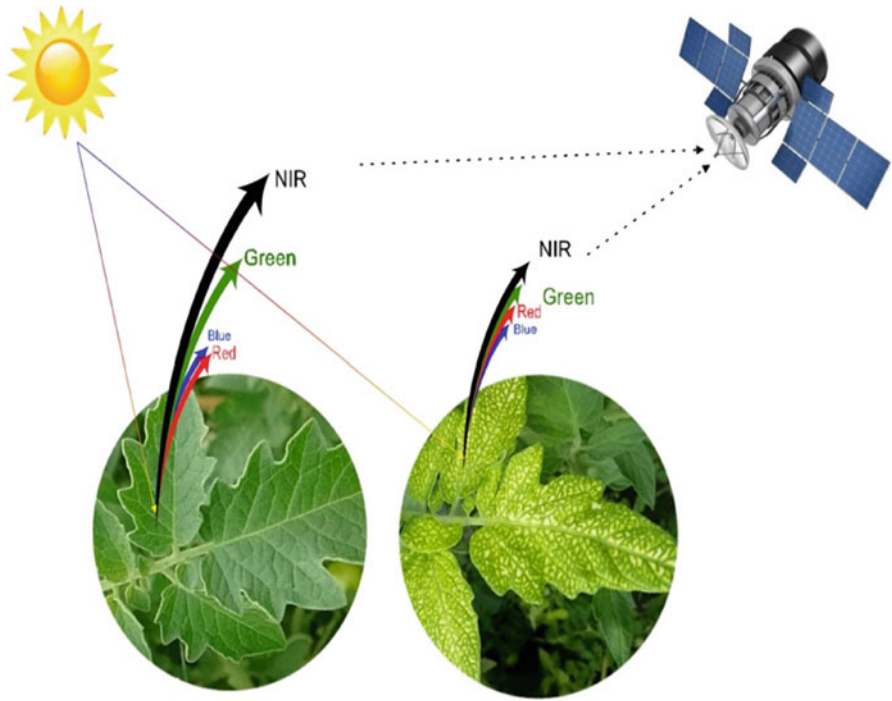
Fig. 7.3 Basic idea about sensing of plant images with different color and texture

Fig. 7.4 Yellow-boxed region shows healthy field and red-boxed region showed the unhealthy region of crop
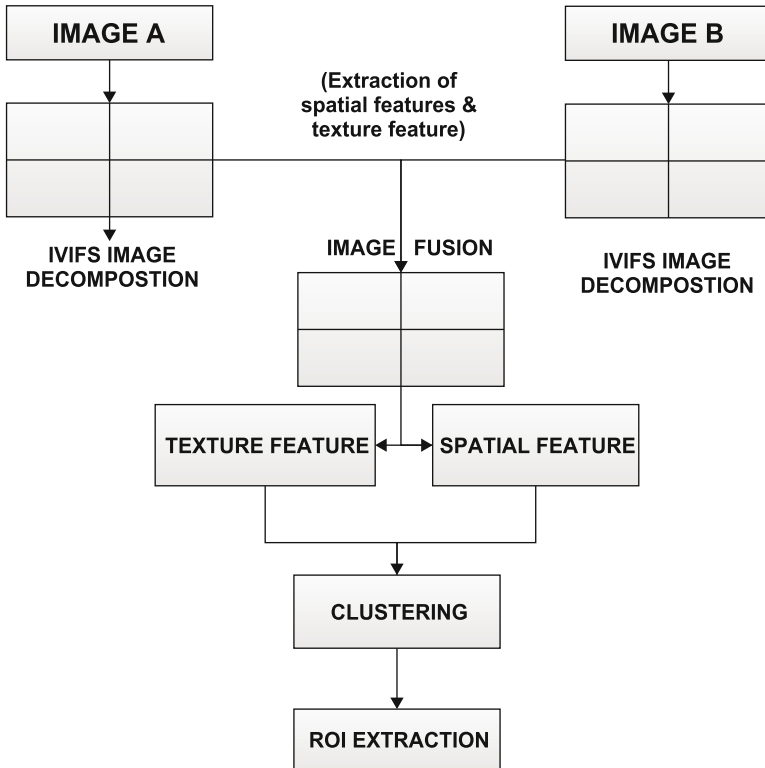
Fig. 7.5 Schematic diagram of segmentation

7.3 Proposed Classification of Affected Crop Region Using IVIFS

This section explains the process of the proposed fused clustering technique for the segmentation of nutrient deficient region; its flowchart has been shown in Fig. 7.5.

7.3.1 Fusion of SAR Images

After the preprocessing SAR images, ROI are imaged either under same sensor or different sensor are initially registered in database. The two images of same ROI are initially fuzzified as in Sect. 7.2.2.1 by using Eqs. (7.2) and (7.6). Image in IFS domain is stretched as an interval using Eq. (7.8), and then the images are then decomposed into images of smaller size. Then their spatial feature is extracted in terms of intensity variation, and their texture features are extracted using gray-level

co-occurrence matrix (GLCM). Fuzzy uniformity, fuzzy contrast, and homogeneity over the image neighbors are extracted using GLCM. Then each corresponding image features of the same ROI are compared and fused to get a composite image. For instance, contrast feature of the block of image A is high, and then A is retained; otherwise block of the image B is retained, which in turn produces a contrast-enhanced image.

7.3.2 Proposed Interval-Valued Intuitionistic Fuzzy C-Means Clustering Technique (IVIFCM) Segmentation Technique

This section explains the process of the proposed segmentation technique that is utilized in the extraction of ROI from the fused image. IVIFCM is a newly emerged clustering algorithm introduced to eliminate uncertainty which is left over by IFSs. Rather than fuzzy c-means (FCM), intuitionistic fuzzy c-means (IFCM) clusters the image by minimizing the following new objective function, defined as

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{*m} d_{ik}^* + \sum_{i=1}^{c} W_i^* e^{1-W_i^*}, \; m \geq 1,$$

where $d_{ik}^* = d(x_k, v_i)$ is Euclidean distance function between the cluster center v_i and the pixel x_k, the fuzzification parameter $m > 1$, and $u_{ik}^* = \underline{u}_{ik} + \overline{u}_{ik}$; here $\underline{u}_{ik}$ and $\overline{u}_{ik}$ denote the lower and upper membership value of kth point (x_k) in ith cluster (v_i), and the second term on the right hand describes entropy to accumulate the pixels exactly in each class. This term minimizes the entropy of the image and $W_i^* = \frac{1}{n} \sum_{k=1}^{n} W_{ik}$, where W_{ik} (degree of hesitation) describes uncertainty of kth point in ith cluster. Steps of IVIFCM clustering algorithm for segmentation of SAR image are given as follows:

Step 1: Let I be a fused SAR image of size n (= PQ). Fix the cluster class c, let $m_1, m_2 > 1$, and its average $m > 1$ be the three fuzzification indices, end limit $\epsilon > 0$, and iteration counter t.

Step 2: Initialize the cluster center $V^{(t)}$ at $t = 0$.

Step 3: For tth iteration, compute the membership matrix $U^{(t)}$ by adopting the following expression

$$\underline{u}_{ik}^{(t)} = \begin{cases} \dfrac{1}{\left(\dfrac{d_{ik}^{(t)}}{\sum_{j=1}^{c} d_{jk}^{(t)}} \right)^{\frac{2}{m_1-1}}}, & \text{if } d_{ik}^{(t)} > 0; \\[2em] \dfrac{1}{\left(\dfrac{d_{ik}^{(t)}}{\sum_{j=1}^{c} d_{jk}^{(t)}} \right)^{\frac{2}{m_2-1}}}, & \text{otherwise,} \end{cases} \qquad (7.9)$$

$$\overline{u}_{ik}^{(t)} = \begin{cases} \dfrac{1}{\left(\dfrac{d_{ik}^{(t)}}{\sum_{j=1}^{c} d_{jk}^{(t)}}\right)^{\frac{2}{m_1-1}}}, & \text{if } d_{ik}^{(t)} \leq 0; \\[4ex] \dfrac{1}{\left(\dfrac{d_{ik}^{(t)}}{\sum_{j=1}^{c} d_{jk}^{(t)}}\right)^{\frac{2}{m_2-1}}}, & \text{otherwise,} \end{cases} \tag{7.10}$$

where $d(x_k, v_i) = \frac{1}{2}\sqrt{\text{dist}}$ with

$$\text{dist} = ||x_k - v_i||_{RGB}^2 + ||x_k - v_i||_{\text{Texture}}^2.$$

Step 4: Update cluster center $V^{(t+1)}$ by utilizing the following expression

$$v_i^{(t+1)} = \frac{\sum_{k=1}^{n} \left(u_{ik}^{*(t+1)}\right)^m x_k}{\sum_{k=1}^{n} \left(u_{ik}^{*(t+1)}\right)^m}, \tag{7.11}$$

where $u_{ik} = \frac{u_{ik} + \overline{u}_{ik}}{2}$.

Step 5: If distance between the membership matrix got from the present $(t+1)$th and previous (t)th iteration is less than ϵ, then stop the process. Else, go to step 3 by fixing t as $t+1$.

7.4   Experimental Analysis

Experimentally algorithm is on large database; few of them are provided in this chapter to show the performance of IVIFCM algorithm. SAR images have been Uavsar, Copernicus data search, National Remote Sensing Centre, and Tamil Nadu Agricultural University, India. In this section, nutrient deficiency of the crop images has been clustered using IVIFCM algorithm. In order to prove the efficiency of the segmentation algorithm, the following measures are evaluated [28]. The proposed method is compared with the fuzzy clustering and intuitionistic fuzzy clustering method with Euclidean distance measure.

7.4.1   Accuracy

It is evaluated to test the overall classification rate of the classifier and is estimated as $Accuracy = (T_p + T_n)/(T_p + T_n + F_p + F_n)$, T_p – number of positive classes clustered as positive, T_n – number of negative classes clustered as negative, F_p – number of negative classes clustered as positive, F_n – number of positive classes clustered as negative.

7.4.2 Precision

It calculates the percent of predictions done by the method are positive and are correct.

$$\text{Precision} = T_p/(T_p + F_p)$$

7.4.3 Recall

It calculates the percent of positive patterns predicted by the method are correct.

$$\text{Recall} = T_p/(T_p + F_n)$$

7.4.4 Precision-Recall Curves

Precision-recall curves render the connection between precision and recall as segmentation cutoff limits vary [29].

7.4.5 ROC Curves

Receiver operating characteristic (ROC) is a graph that plots true-positive and false-positive rates along the variation of cutoff limits. True positive rate is nothing but recall, and false-positive rate is the ratio of false positive to the total negative [28].

7.4.6 SSIM

All natural images are well organized, and their image elements are vigorously dependent which inherit vital data regarding structure of an image. It works in three phases, namely, luminance, contrast, and structural comparison. The third phase correlates the internal patterns of intensities among luminance and contrast of a normalized image.

$$\text{SSIM} = \frac{(2\mu_{I_{\mathcal{R}}}\mu_{I_{\mathcal{F}}} + K_1)(2\sigma_{I_{\mathcal{R}}I_{\mathcal{F}}} + K_2)}{(\mu_{I_{\mathcal{R}}}^2 + \mu_{I_{\mathcal{F}}}^2 + K_1)(\sigma_{I_{\mathcal{R}}}^2 + \sigma_{I_{\mathcal{F}}}^2 + K_2)},$$

where K_1 and K_2 are constants which are appended to retain stability whenever $\mu^2_{I_\mathcal{R}} + \mu^2_{I_\mathcal{F}}$ and $\sigma^2_{I_\mathcal{R}} + \sigma^2_{I_\mathcal{F}}$ are approaching zero.

7.4.7   Results and Discussion

Figure 7.6 shows the segmented results of the images A to E clustered by the proposed algorithm, IFCM, and FCM algorithm. Second row of Fig. 7.6 shows the image A, and its three segmented image acquired using IVIFCM, IFCM and FCM techniques. The result acquired by IFCM clustering method has oversegmented regions. Similarly, segmented output got by FCM algorithm has over some segmented regions than that of IFCM and proposed method. That is, it has segmented other than the defected regions. The proposed segmentation technique has segmented the image approximately than IFCM and FCM methods.
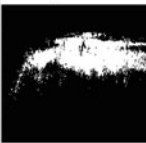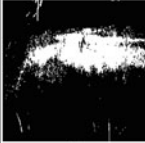
Fig. 7.6 Segmentation results of various nutrient deficient crop images from A to E

The third and fourth row of Fig. 7.6 shows the segmented outputs of the images B and C by the three methods, and from these outputs, the proposed method seems to be more effective than IFCM algorithm. From the segmented results of the images D and E given in the fifth and sixth row of Fig. 7.6, it is visually vivid that the proposed method segments the ROI more perfectly than IFCM and FCM methods.

Similarly, Fig. 7.7 shows the clustered results of the image F to J by the three methods. The results of the proposed method given in the third column of Fig. 7.7 seem to be more desirable segmentation than the result provided by IFCM and FCM algorithm in the fourth and fifth column of Fig. 7.7. The results rendered by IFCM and FCM algorithms show some oversegmented regions. From Figs. 7.6 and 7.7, the proposed method is qualitatively better when compared to the existing intuitionistic fuzzy and fuzzy algorithms.

Figure 7.8 renders the values of the quantitative metric accurately plotted as a bar graph for all ten images for the three methods. Proposed IVIFCM algorithm seems to have high accuracy rate than the other two algorithms. Figure 7.8 eloquently shows the efficiency of the IVIFCM fused segmentation technique.
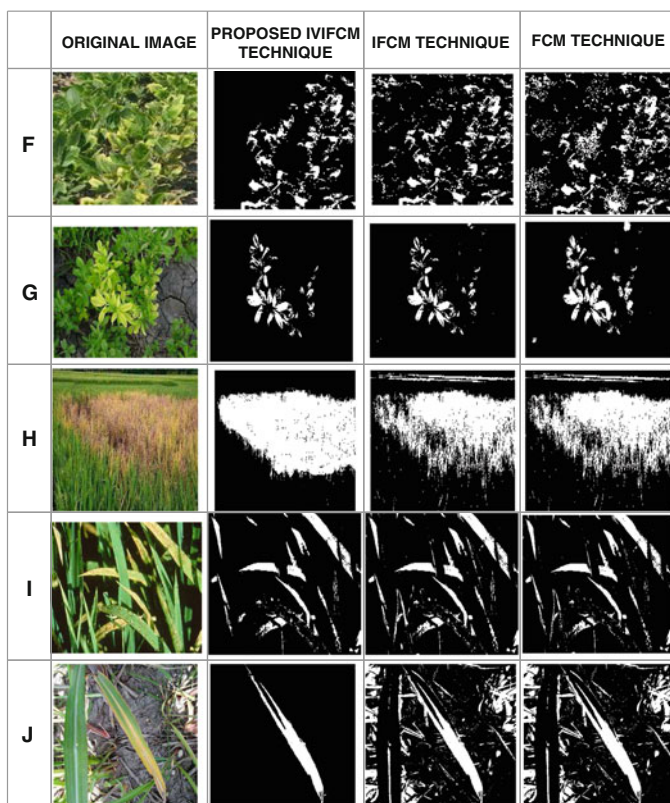
Fig. 7.7 Segmentation results of various nutrient deficient crop images from F to J

Fig. 7.8 Accuracy values of the ten segmentation images from A to J
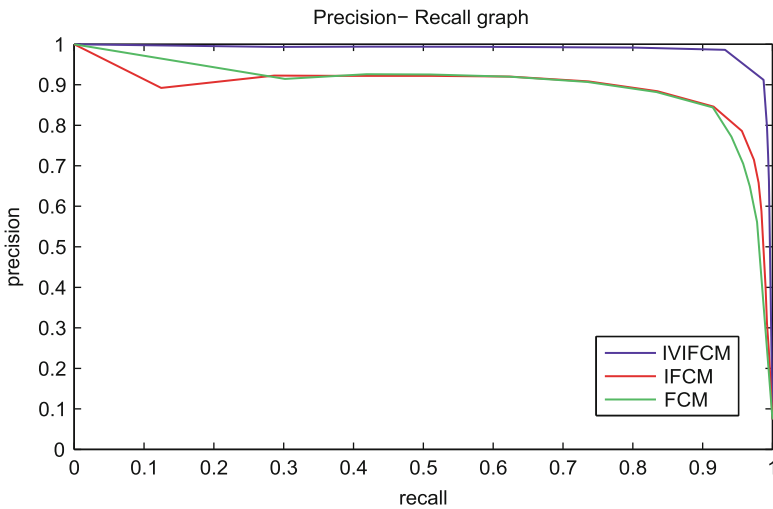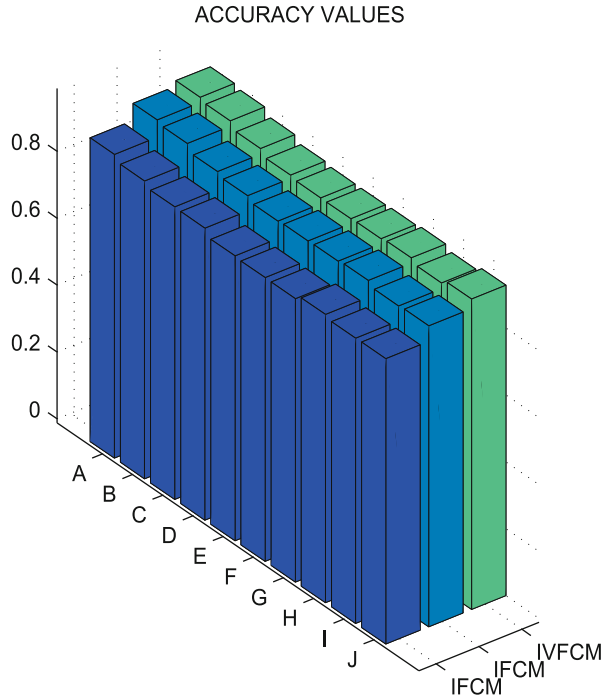
Fig. 7.9 Precision-Recall graph

Figures 7.9 and 7.10 picturizes the precision-recall graph and ROC curves of the segmented images of the input images A to J. Both the graphs vividly represent the effectiveness of the proposed method over IFCM and FCM algorithm. All the
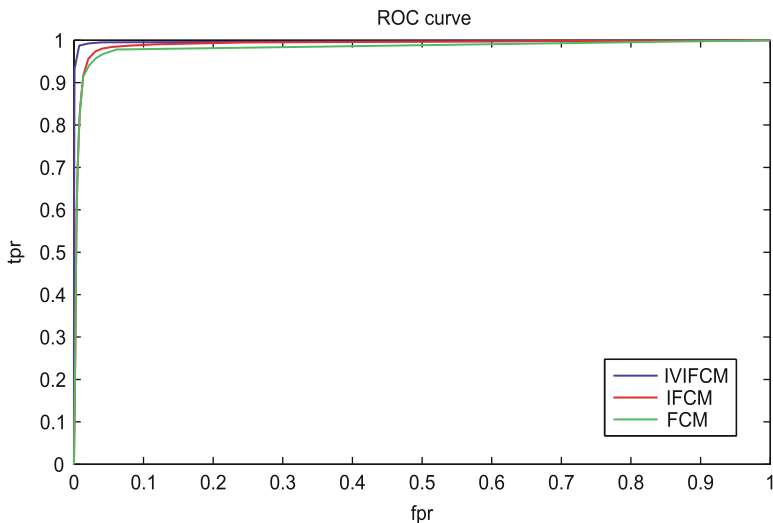
Fig. 7.10 ROC curve
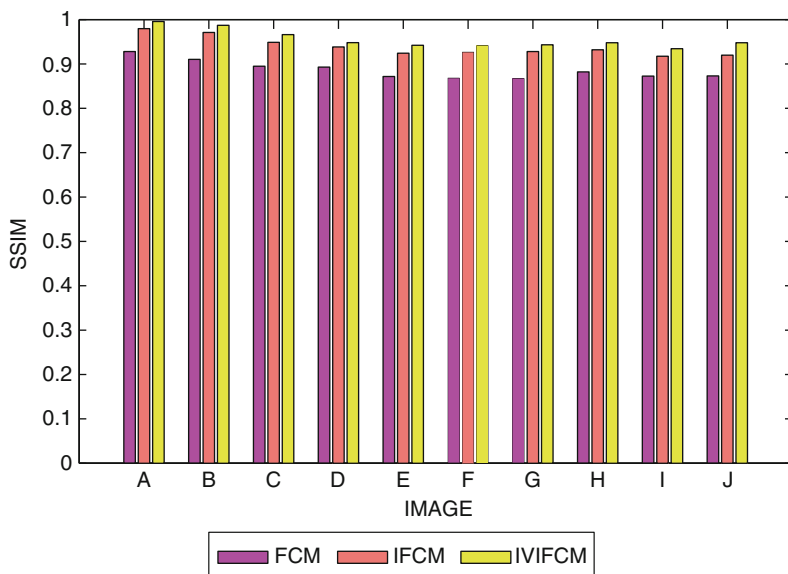
Fig. 7.11 SSIM values of the ten segmented images

figures quantitatively and qualitatively reveal that the proposed method clustered the ROI without over- or undersegmentation of other non-deficient region.

SSIM values of the segmented images are evaluated, and the values are plotted in a bar graph in Fig. 7.11. From the SSIM values of the proposed method are high

compared to the other two methods. All the figures quantitatively and qualitatively reveal that the proposed method extracts the ROI in a better way than that of the existing algorithm.

Though the proposed method seems to extract the ROI better than the existing method, let us discuss about importance and improvement of the proposed method over the existing method. Generally, clustering technique depends upon the Euclidean distance function with the spatial intensity of the image. Proposed method extends the same concept for both spatial and texture features of the image for improvement in the detection of ROI. Proposed method initiates by eliminating uncertainty in this membership grades. Uncertainty in IVIFS has been modeled as width of the interval to reduce the confusion about the acquired membership grade acquired by the considered membership function. Spatial features such as intensity sharpness, spectral resolution, and textural information such as homogeneity, color, entropy, and statistical measure are acquired using gray-level co-occurrence matrix. Textural and spatial information are both directly cannot be related to each other. One should make use of conversion of these details as an intensity value for correlation. Usually working in fuzzy domain seems to be working in the normalized pixel intensity range of value. Hence nature of intensity value will be preserved. Therefore spatial and textural information can be combined to form a fused image. For instance, vegetation region accurately or perfectly captured by NIR sensors and SAR imagining can be converted as RGB scale for easy perception. Fusing these images from these sensors will provide more information about vegetation/field in the image, which is ROI. Again clustering of the fused image is done according to the distance measure for spatial or color and texture feature, respectively. That is, distance measure is separately checked for color components R,G,B and texture feature like homogeneity or entropy; it is minimum then the considered pixels belong to the particular group. Hence this makes the proposed clustering technique to identify ROI perfectly than the existing technique.

7.5 Conclusion

In this chapter, segmentation of the fused images has been introduced to handle the problem of finding deficiency region using SAR images. Since images obtained from sensors at a particular time are not compatible for detection for all ROI due to impact of environmental condition. Images of particular field are imaged multiple times, and then they are fused by reducing the uncertainty about its intensity level. Then the fused image is clustered using IVIFS with distance function having both color and texture features. The deficiency region is finally extracted from the clustered output, from which one can estimate the area of field affected with disease. From the area detected, the agriculturist can do the concern need for the crop. If such investigation/monitoring is initiated from the beginning of the seeding/planting, the agriculturist can rectify the nutrient deficiency/affected region at the beginning (or spread) of such effects for improving productivity. The major reason for the

better result of the proposed IVIFCM is that it works on the fused SAR image and IVFCM considers both the spatial and textural information while clustering. If the disease/pest attack is found at earlier stage, then it simultaneously improves the productivity immediate and proper manuring and fertilization for that crop, thereby improving the country's economy. As a future work, SAR image's 1 pixel size represents $2.5 \times 2.5\,\text{m}^2$ ground area, from that one can find the area of deficient region which in turn we can find the rate of fertilization for the concerned field.

References

1. Kurosu T, Fujita M, Chiba K (1995) Monitoring of rice crop growth from space using the `ERS-1 C-band SAR`. IEEE Trans Geosci Remote Sens 33(4):1092–1096
2. Tan CP, Ewe HT, Chuah HT (2011) Agricultural crop-type classification of multi-polarization sar images using a hybrid entropy decomposition and support vector machine technique. Int J Remote Sens 32(22):7057–7071
3. Teimouri M, Mokhtarzade M, Valadan Zoej MJ (2016) Optimal fusion of optical and sar high-resolution images for semiautomatic building detection. GIsci Remote Sens 53(1):45–62
4. Massonnet D, Souyris J-C (2008) Imaging with synthetic aperture radar. EPFL, Lausanne
5. Zhou T, Pan J, Zhang P, Wei S, Han T (2017) Mapping winter wheat with multi-temporal sar and optical images in an urban agricultural region. Sensors 17(6):1210
6. Dong Z, Wang Z, Liu D, Zhang B, Zhao P, Tang X, Jia M (2013) Spot5 multi-spectral (ms) and panchromatic (pan) image fusion using an improved wavelet method based on local algorithm. Comput Geosci 60:134–141
7. Ananthi VP (2017) Studies on processing of images with Uncertainty using intuitionistic fuzzy sets. Unpublished doctoral dissertation, Gandhigram Rural Institute-Deemed University, India
8. Zadeh LA (1965) Information and control. Fuzzy sets 8(3):338–353
9. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96
10. Chaira T (2012) A rank ordered filter for medical image edge enhancement and detection using intuitionistic fuzzy set. Appl Soft Comput 12(4):1259–1266
11. Hwang C, Rhee FC-H (2007) Uncertain fuzzy clustering: interval type-2 fuzzy approach to c-means. IEEE Trans Fuzzy Syst 15(1):107–120
12. Atanassov K, Gargov G (1989) Interval valued intuitionistic fuzzy sets. Fuzzy Sets Syst 31(3):343–349
13. Gonzalez RC, Woods RE (2002) Digital image processing. Prentice Hall of India, New Delhi
14. Gonzalez RC, Woods RE (2002) Thresholding. In: Digital image processing, pp. 595–611
15. Bustince H, Barrenechea E, Pagola M, Fernández J (2009) Interval-valued fuzzy sets constructed from matrices: application to edge detection. Fuzzy Sets Syst 160(13):1819–1840
16. Krishnapuram R, Keller JM (1993) A possibilistic approach to clustering. IEEE Trans Fuzzy Syst 1(2):98–110
17. Bezdek JC (2013) Pattern recognition with fuzzy objective function algorithms. Springer, New York
18. Xu Z, Wu J (2010) Intuitionistic fuzzy c-means clustering algorithms. J Syst Eng Electron 21(4):580–590
19. Bustince H, Burillo P (1995) A theorem for constructing interval valued intuitionistic fuzzy sets from intuitionistic fuzzy sets. Notes on Intuitionistic Fuzzy Sets 1(1):5–16
20. Pal SK, King R (1981) Image enhancement using smoothing with fuzzy sets. IEEE Trans Syst Man Cybern 11(7):494–500
21. Chaira T, Ray AK (2009) Fuzzy image processing and applications with MATLAB. CRC Press, Boca Raton

22. Raol JR (2015) Data fusion mathematics: theory and practice. CRC, New York
23. Ananthi VP, Balasubramaniam P (2015) Image fusion using interval-valued intuitionistic fuzzy sets. Int J Image Data Fusion 6(3):249–269
24. Meera Gandhi G, Parthiban S, Thummalu N, Christy A (2015) Ndvi: vegetation change detection using remote sensing and gis–a case study of vellore district. Proc Comput Sci 57:1199–1210
25. Ortiz B, Shaw J, Fulton J (2011) Basics of crop sensing. Alabama cooperative extension system, 1–3
26. McNairn H, Wiseman G, Powers J, Merzouki A, Shang J (2014) Assessment of disease risk in canola using multi-frequency sar: preliminary results. In: EUSAR 2014; Proceedings of 10th European conference on synthetic aperture radar. VDE, Berlin, pp 1–4
27. Canisius F, Shang J, Liu J, Huang X, Ma B, Jiao X, Geng X, Kovacs JM, Walters D (2018) Tracking crop phenological development using multi-temporal polarimetric radarsat-2 data. Remote Sens Environ 210:508–518
28. Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manag 45(4):427–437
29. Balasubramaniam P, Ananthi VP (2016) Segmentation of nutrient deficiency in incomplete crop images using intuitionistic fuzzy c-means clustering algorithm. Nonlinear Dyn 83(1–2): 849–866

Chapter 8
Detection of Natural Features and Objects in Satellite Images by Semantic Segmentation Using Neural Networks

Vihar Kurama, Samhita Alla, and Sridevi Tumula

Abstract In recent years, Neural Networks have become one of the most research focused areas of Artificial Intelligence. From detecting objects in real time to the classification of images, these Neural Networks are efficient and are achieving maximum possible accuracies based on the given inputs. In this work, we use Neural Networks for detecting features in satellite images. Using image segmentation and object detection techniques, we find objects, like roads, buildings, trees, and other resources, in the satellite images. In this work, Neural Network architecture used for segmentation of the images is ConvNet also called Convolutional Neural Network. U-Net which has a convolutional autoencoder architecture maps the layers to find the features and resources in the given satellite images. U-Nets do per-pixel semantic alignment for finding objects and features which result in segregation of resources. By using these, each feature or resource in the satellite image is segmented in different colors with regards to the distinct features allowing us to estimate the resources.

Keywords CNN · Image segmentation · U-Nets · Object detection · Neural networks · Artificial intelligence

8.1   Introduction

These days, satellite images have improved our understanding of the planet in all respects. They are being experimented using AI and Deep Learning techniques to formulate decisions. Satellite images provide us with a wide variety of applications. However, it poses a challenging problem due to the complexity of the pictures and also the wide variations of objects available. Therefore, most of the current object detection techniques are not suitable when dealing with satellite images [12].

V. Kurama · S. Alla · S. Tumula (✉)
Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: sridevi_t@cbit.ac.in

Object recognition in aerial images is becoming the current research topic due to the recent advancements in Deep Learning and Computer Vision. Detecting buildings, vegetation, and roads in satellite images have a wide variety of applications, like for creating maps, to mitigate disasters, and to perform environmental surveys [7]. Image segmentation is to divide the whole image into a patchwork of regions, each of which represents an image again. Image classification is to assign a predefined class label to every pixel in the image. But, the above two concepts are interlinked; a classifier segments an image, and segmentation leads to classification.

8.1.1 Rise in Deep Learning

Artificial Neural Network can be thought of as a composite function which mimics the Neural Network present in the human brain. A Neural Network is made up of neurons, weights, and biases. Neurons are the powerful computational units that take in an input signal and produce the output signal by applying an activation function to the input. These neurons are present across several layers of the Neural Network. Weights refer to the strength of the connections between the neurons. Bias is used to shift the output, i.e., either increase or decrease it. There can be several layers present in a Neural Network, an input layer, an output layer, and either several or no hidden layers. An input layer accepts the input and passes it on to the next layers; hidden layers apply transformations on the received input, and the output layer gives the predicted features.

8.1.2 Image Semantic Segmentation

Semantic segmentation refers to classifying the images pixel-wise into one of the predefined classes. The current advancements in the Deep Learning have driven the researchers to explore semantic segmentation, a pixel-level classification task. Computer Vision techniques, one of the promising applications of AI, can be applied for satellite images to extract the objects embedded in it. It helps in gaining high-level understanding of the complex satellite images. And with Neural Networks and Deep Learning, technology is taken to a different level. Neural Networks form the crux of Deep Learning applications in Computer Vision.

Segmentation in satellite images can be carried out using various techniques, say thresholding, clustering, region-based, and Artificial Neural Networks. Among all those, Artificial Neural Networks are the most successful concerning accuracy. It has also laid down the ground for Convolutional Neural Networks. It is a supervised approach used for segmentation purposes. The output should define the localization of labels for every pixel, for which we need considerable datasets to attain a higher accuracy. To overcome the above, Ciresan et al. [3] defined an architecture wherein a sliding window is used to take in the pixel regions and then it predicts the class labels. The disadvantage with this is, it consumes a lot of time.

8.1.3   Deep Learning Over Machine Learning

Machine Learning is the parsing of data and then learning the patterns present in the data to make informed decisions. It has become one primary utility for every researcher and is progressing day by day by solving various complicated problems. It explains the outcomes that can be drawn from the data based on the associations, similarities, and dependencies present among multiple attributes. It involves a good many complex mathematical concepts in eliciting the patterns existing in the dataset.

Deep Learning is a unique way of representing the Machine Learning algorithms. It is the subset of Machine Learning and its functions, but its capabilities are different. Deep Learning structures, Artificial Neural Networks, can learn and make decisions intelligently on their own.

Deep Learning needs massive amounts of data to operate and perform well, but Machine Learning algorithms perform well irrespective of the data size. The hardware required by Deep Learning is also of high-end, like GPUs (graphics processing units) and TPUs (tensor processing units) to run the algorithms smoothly, whereas Machine Learning algorithms do not need such high-end hardware. Features in a Machine Learning model need to be identified by an expert and then hard-coded, but a Deep Learning model automates the task of identification of elements. Hence there is no requirement to develop a new feature extractor. A Deep Learning model consumes more amount of time for training unlike a Machine Learning model, but during the testing phase, Deep Learning takes much less time. In the case of images, a Machine Learning model divides the problem into modules and then executes it; a Deep Learning model does all that in one go.

Deep Learning is a breakthrough in the fields of Computer Vision, information retrieval, natural language processing (NLP), and medical diagnosis. Deep Learning models include several parameters which need to be trained to attain higher accuracy.

With time, Deep Learning algorithms are getting better regarding performance attained on the given data. Feature engineering can be skipped whenever a problem is solved using Deep Learning algorithms. This feature engineering involves exploring and performing analysis on the data before it is fed into any classical Machine Learning algorithm. In Deep Learning, the data can be directly sent into the network for predictions. This data can be of any format which includes text, audio, images, and videos. One other advantage of using Deep Learning algorithms over Machine Learning is that they are adaptable and transferable. We need not train the Neural Network over time; once the model is trained, we can save the model and transfer it to any other domain which can be implemented several times unless there is a change in the network architecture or the dimensions of the input data. We always need to train a Machine Learning algorithm whenever we need a new prediction which again consumes a lot of computational resources and time.

8.2 Artificial Neural Networks

A feed-forward network is the one where the inputs pass from one layer to the next, finally reaching the output layer. In a back-propagation network, the errors are given starting from the output layer and then the hidden layers, eventually reaching the input layer. The weights are then updated to reduce the errors in the subsequent feed-forward propagation of inputs (Fig. 8.1).

8.2.1 Back Propagation

Back-propagation algorithm [16] can update the weights efficiently and is used in conjunction with gradient descent optimization method. It updates the weights repeatedly to minimize the errors between the actual output vector and the desired output vector. As a part of this weight adjustment, internal nodes which are neither inputs nor outputs play a significant role in a task domain [16]. It involves finding the derivative of the cost function and then back propagating the errors to update the weights. Cost function in most cases is calculated using mean squared error (MSE) to find the deviations of predicted output over the actual production in case of supervised Machine Learning algorithms. Back-propagation tries to minimize the following error function by calculating for each weight w_{ij}^k, in other words, for node j in layer k for an incoming node i:

$$E(X, \theta) = \frac{1}{2N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \tag{8.1}$$
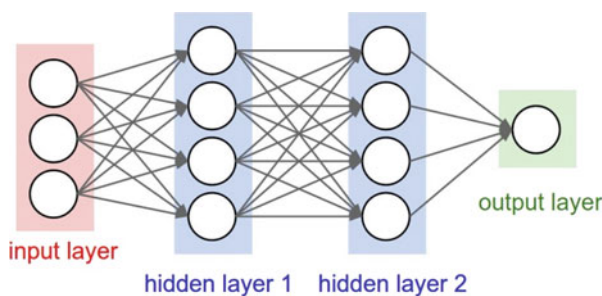
Fig. 8.1 Simple neural network

where E is the error function, X is the input, and θ is the weight vector used. N is the total number of training examples used, $\hat{y}_i$ is the computed output, and y_i is the actual output.

The derivative can be calculated for each weight individually and then combined finally.

$$\frac{\partial E(X, \theta)}{\partial w_{ij}^k} = \frac{1}{N} \sum_{d=1}^{N} \frac{\partial}{\partial w_{ij}^k} \left(\frac{1}{2} \left(\hat{y}_d - y_d \right)^2 \right) = \frac{1}{N} \sum_{d=1}^{N} \frac{\partial E_d}{\partial w_{ij}^k} \tag{8.2}$$

Back-propagation proceeds by applying chain rule to the error function.

$$\frac{\partial E}{\partial w_{ij}^k} = \frac{\partial E}{\partial a_j^k} \frac{\partial a_j^k}{\partial w_{ij}^k} \tag{8.3}$$

where a_j^k is the activation function of node j in layer k. Starting from the output layer, back-propagation proceeds backward to the input layer calculating error at every step using the chain rule, and then the weights are updated (Fig. 8.2).

Activation functions are applied to the outputs produced at every hidden layer and the output layer to make the predictions more accurate by inducing complexity into the Neural Network. Sigmoid or logistic activation function produces the output between 0 and 1; a tanh or tangent hyperbolic activation function gives output lying between -1 and 1. ReLU (which is the most used activation function in Convolution Neural Networks) lies between 0 and infinity (Fig. 8.3).
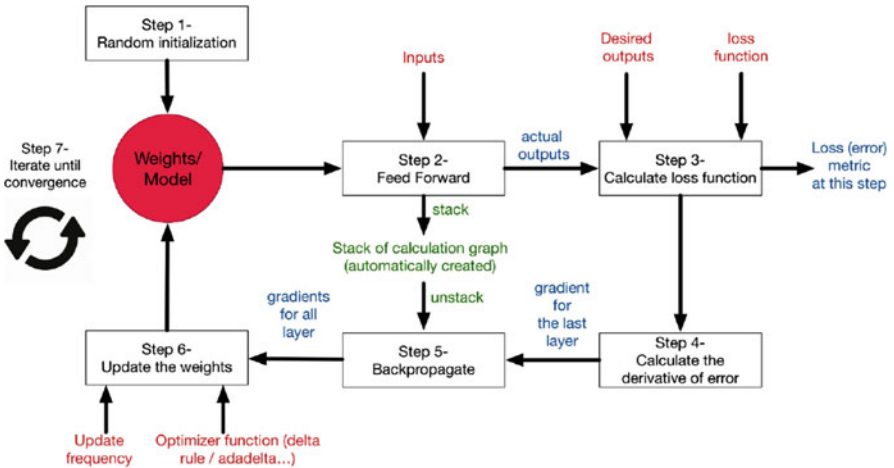
Fig. 8.2 Step-by-step approach to implement a Neural Network [13]
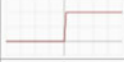
Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a. k. a Soft step)		$f(x) = \dfrac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
TanH		$f(x) = \tanh(x) = \dfrac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \dfrac{1}{x^2 + 1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \dfrac{1}{1 + e^{-x}}$

Fig. 8.3 Different activation functions used in Neural Networks [1]

8.3 Hyperparameters for Neural Network Training

All Machine Learning models include a set of variables for better predictions and optimization of the algorithms which are known as hyperparameters. These are defined before optimizing the algorithm based on the hyperparameters that are assigned the model's accuracy, training time, and optimization changes. These are often chosen randomly and corrected based on how the algorithm works. Deep Neural Networks depend more on these hyperparameters, including those which specify the architecture of the neural system itself and those which determine how the Neural Network model is trained. A few significant hyperparameters are learning rate, loss function, and momentum.

8.3.1 Mini-Batch Gradient Descent Hyperparameters

The outputs that are expected by the Neural Network will not only depend on the network architecture; it also primarily depends on the network parameters. Previously we've encountered mini-batch gradients, these are the parameters that have to be considered whenever the training is done in batches. As the optimization and cost

function of the network are updated for every new training data, mathematically, we define the mini-batch gradient descent update of network parameters θ as

$$\theta^{(t)} \leftarrow \theta^{(t-1)} - \epsilon_t \frac{1}{B} \sum_{t'=Bt+1}^{B(t+1)} \frac{\partial L(z_{t'}, \theta)}{\partial \theta} \tag{8.4}$$

where

B – Mini-batch size
T – Number of iterations
L – Loss function
ϵ t – Learning rate at t-th iteration.

8.3.2   Learning Rate

Learning rate defines how fast does the gradient update with the iterations of the input data. Basically, for most of the Neural Network architectures, the learning rate is considered in the range of 0.01 to 0.00001. Learning is directly proportional to the optimization of network training time and inversely proportional to the optimization. However, the learning rate of the model can be fine-tuned over a period of training, here, with the vast amount of data present, the ideal learning rate considered is 0.001.

8.3.3   Loss Function

The loss function is used for comparison of the original values or the ground labels that are presented in the training data to the predicted values by the Neural Network. The most commonly considered loss function is the squared Euclidean distance which will be the squared difference between the expected and the actual cost.

$$L = \frac{1}{2} \sum_i (y_i - z_i)^2 \tag{8.5}$$

Whenever the systems have a softmax function and have a probability-based reduction, then cross entropy loss is used which is mathematically defined by

$$L = - \sum_i y_i \log(z_i) \tag{8.6}$$

The loss function can be changed based on the training data and the type of network architecture that is implemented.

8.3.4 Mini-Batch Size

These are used when the model needs to be trained on large datasets; the data is split into mini batches and computed in parallel in the same architecture. Value of mini quantity is directly proportional to the efficiency of the network. One more advantage of using the mini-batch size as hyperparameter is that the training will be finished in less time as they are computed in parallel. This can be set to the accurate value irrespective of the other hyperparameters.

8.3.5 Momentum

Momentum (g) is one of the hyperparameters which is mostly used for optimization in converging the rates on Deep Neural Networks. This makes the gradient descent smoother by a leaky integrator filter with parameter β which is mathematically defined as

$$\bar{g} \leftarrow (1 - \beta)\bar{g} + \beta \frac{\partial L(z_t, \theta)}{\partial \theta} \tag{8.7}$$

8.4 Convolutional Neural Networks (CNN)

LeCun et al. introduced the Convolutional Neural Networks [9]. They operate over volumes and always assume that the input is a multi-channelled image. A ConvNet arranges its neurons in three dimensions (width, height, depth) to avoid building a Neural Network comprising a significant number of neurons as seen in the case when the input is an image.

CNN architecture: It has a hierarchical architecture consisting of convolutional layers, pooling layers, and fully connected layers. Every layer transforms one volume of activations to the other using a differentiable function.

8.4.1 Convolutional Layer

Convolutional layer – The dot product of the input image tensor and the filter is computed which is one by-product, and it is then applied to an activation function, say ReLU. A filter is a group of neurons which recognize patterns at different locations of an image. The above process is repeated until we arrive at an optimal

solution. Also, with regards to the depth, we don't mention hyperparameters, as the convolution by default is performed along the whole thickness of the image. The number of convolutions being conducted can be defined by a hyperparameter called stride which determines the gap between two scanned regions [2]. Padding can also be used to make the size of the output image the same as the input image size, and it helps in retaining the border values too. Therefore, the output size of the convolutional operation can be calculated as shown in the below formula. Below is the first convolutional architecture that is proposed by LeCun et al. [10] (Fig. 8.4).

Around the entire input when the convolution kernel is applied, the image is shrinked. To retain the dimensions of the original image, we add the padding data. You add padding data with a width equal to the kernel width minus one (or height equal to kernel height minus one if it's above and beneath), so that the kernel can look at the extreme edge. Mathematically, the output size output size can be written with padding as (Fig. 8.5)

$$\text{outputsize} = \frac{\text{inputsize} - \text{filter} + (2 * \text{padding})}{\text{stride}} + 1 \tag{8.8}$$
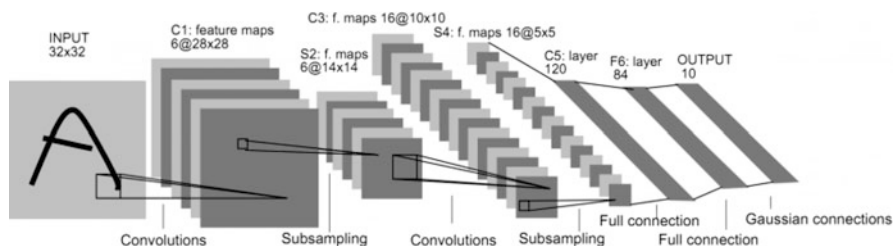
Fig. 8.4 Architecture of convolutional neural network, LeCun et al. [10]

Fig. 8.5 Receptive field (filter), padding and stride [2]
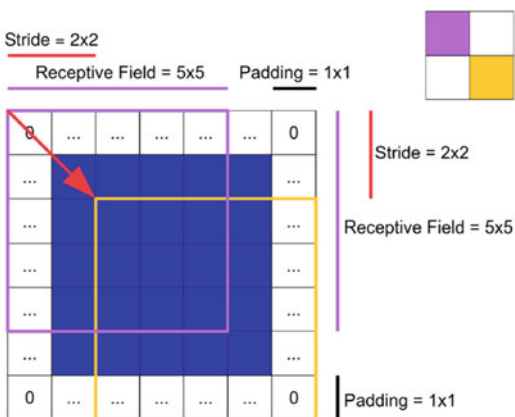
8.4.2 Pooling Layer

The main difference between a pooling layer and a convolutional layer is that in a pooling layer there are no parameters to learn, i.e., it is not parameterized [17]. There are no weights and biases to learn. It performs some static function on the input given. The most common type of pooling operation is max pooling wherein the size of the receptive field is fixed and then moved along the whole image. The maximum value in each region is returned. Average pooling is another way of implementing pooling wherein the average of all the values is taken (Figs. 8.6 and 8.7).

8.4.3 Fully Connected Layer

Each neuron in a layer is connected to every other neuron in the next layer. It classifies the image into various classes of the training dataset. Below is the output of a sample input once the features are recognized using a fully connected layer. The results obtained using Convolution Neural Networks on various Machine Learning models have been remarkable.
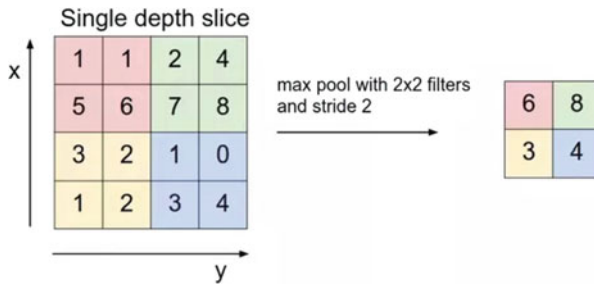
Fig. 8.6 Pooling operation on kernels [18]
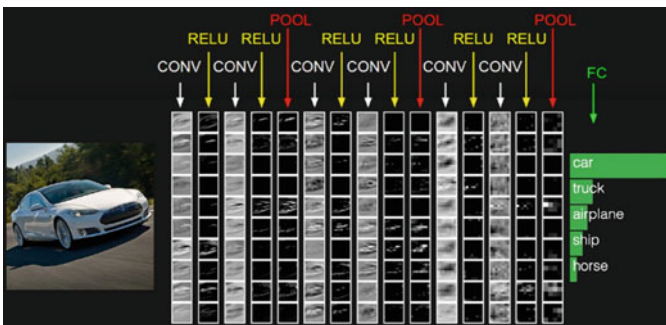
Fig. 8.7 Sample output of a fully connected layer [18]

Assuming that the input signal is x, then each subsequent layer x_j would be

$$x_j = \rho W_j x_{j-1} \tag{8.9}$$

where W_j is a linear operator and ρ indicates the nonlinearity. ρ is the rectifier $\max(x, 0)$ or sigmoid $\frac{1}{1+\exp(-x)}$, and W_j is a stack of convolutional filters. Each layer can be written as the sum of convolutions of the previous layers [8]:

$$x_j(u, k_j) = \rho\left(\sum_k (x_{j-1}(., k) * W_{j,k_j}(., k))(u)\right) \tag{8.10}$$

* is the convolution operator

$$(f * g)(x) = \sum_{u=-\infty}^{\infty} f(u)g(x - u) \tag{8.11}$$

8.5 Image Semantic Segmentation Using ConvNets

We begin the proposal background of image segmentation in this section, to make readers have a better understanding of all the image segmentation and object detection research progress and application fields that can be applied on satellite images.

8.5.1 FCN (Fully Convlutional Neural Network)

They take in arbitrarily sized images as input and then produce the output with efficient learning and inferences. Previously, Convolutional Neural Networks have been used to label every class of its enclosing region or object [3, 4, 14]. In an FCN, the fully connected layers are replaced by convolutional layers. This has a massive set of advantages over the conventional CNN since the entirely connected layers in CNN focus on classifying the outputs generated into their respective classes, whereas, in an FCN, the feature maps are made until the last, involving convolutions and deconvolutions.

CNN identifies the objects in an image irrespective of where they are placed. The convolution, pooling, and activation operations are applied on the local regions. If x_{ij} is the data vector at location (i, j) for one layer and y_{ij} for the following layer [11]

$$y_{ij} = f_{ks}(\{x_{si+i, sj+j}\}_{0 \leq \delta i, \delta j \leq k}) \tag{8.12}$$

where k is the size of the kernel, s is the stride, and $f(k, s)$ determines the layer type, max, or average pooling. The function is maintained under the composition

$$f_{ks} \circ g_{k's'} = (f \circ g)_{k'+(k-1)s', ss'} \tag{8.13}$$

An FCN takes in the input of any size and then produces the resampled dimensions. It gives heat map as the output and hence is used in semantic segmentation. Since a heat map is necessary for satellite images, FCN can be used to segment the image into various objects semantically.

8.5.2 R-CNN (Regional Convolutional Neural Network)

This model revolutionized the way objects are detected in an image. It initially involves breaking the image into various regions using selective search technique [19]. This forms regions based on the similarity existing among the adjacent pixels and is used as inputs to the Convolutional Neural Networks. The final layer has an SVM (support vector machine) to classify the objects. Finally, we optimize the purpose to fit the exact size using linear regression.

8.5.3 Fast R-CNN

It takes in an input image and object classes. The network processes the image with convolutional and max-pooling layers to produce the feature maps. Then, the RoI pooling layer extracts a feature vector from the feature map. RoI uses max pooling to convert features inside a RoI into a feature map. These are passed onto the fully connected layers which give two output layers, one is the softmax probability over all the object classes considered, and the other provides four numbers for each object class [5]. The four values define the bounding box for the objects being detected. Fast R-CNN uses a softmax classifier and regressor, unlike R-CNN which uses SVM.

8.5.4 Mask R-CNN

Mask R-CNN is an extension to Fast R-CNN which in addition to drawing bounding boxes around the objects also predicts the object mask. The primary essential element of a Mask R-CNN is pixel-to-pixel [6]. The mask branch is an FCN applied to each RoI predicting the cover. It is simple to implement and adds only a small overhead over Fast R-CNN.

It makes pixel-level identification of objects in an image. The first stage uses regional proposal network (RPN) to find the region proposals, the second stage predicts the classes, and a binary mask is used to mark 0s where the object is not present and 1s to represent the presence of an object. A technique called RoIAlign is proposed to remove the misalignment due to pixel-level sensitivity.

8.6 U-Net

The main idea in a U-Net is replacing the pooling operations with upsampling operations [11]. Hence, these layers will increase the output image resolution. To attain localization, high-resolution features are combined with the upsampled features [15]. There are no fully connected layers, and the segmentation or the heat map only has the pixels for which the full image is available in the input. To take in the border region, the input image is extrapolated using the mirror of it, without which the GPU memory would decrease the resolution. When passing in the features through the U-Net model, it more or less seems like taking in a U shape, hence the name U-Net. The main advantage with U-Net is it reduces the distortions of the neighboring elements.

U-Net Architecture It is made of two parts, the left one is called the contracting path, and the right one is named as the expansive path. The contracting path has a convolutional architecture. It comprises two 3×3 convolutions, followed by a rectified linear unit (ReLU) [15]. Then, 2×2 max-pooling operation with a stride of 2 is used for downsampling.

We double the feature maps while downsampling. The expansive path reduces the number of feature channels by half and concatenates the respective feature map from the contracting path. Two 3×3 convolutions are applied followed by ReLU activation layers. At the final segment, a 1×1 convolution operation is used to map 64-component vector to classes. In total, it has 23 convolutional layers (Fig. 8.8).

8.7 Convolution Neural Networks for Satellite Images

Detecting features and objects in the satellite images using Deep Convolutional Neural Networks needs a consistent and quality dataset; in our work we use an open-source database by Defence Science and Technology Laboratory (DSTL).

8.7.1 Satellite Images Dataset

This dataset includes 2 sets of images which include 16 band and 13 band images. These images are loaded with 20 channels of TIFF format; out of 20, 3
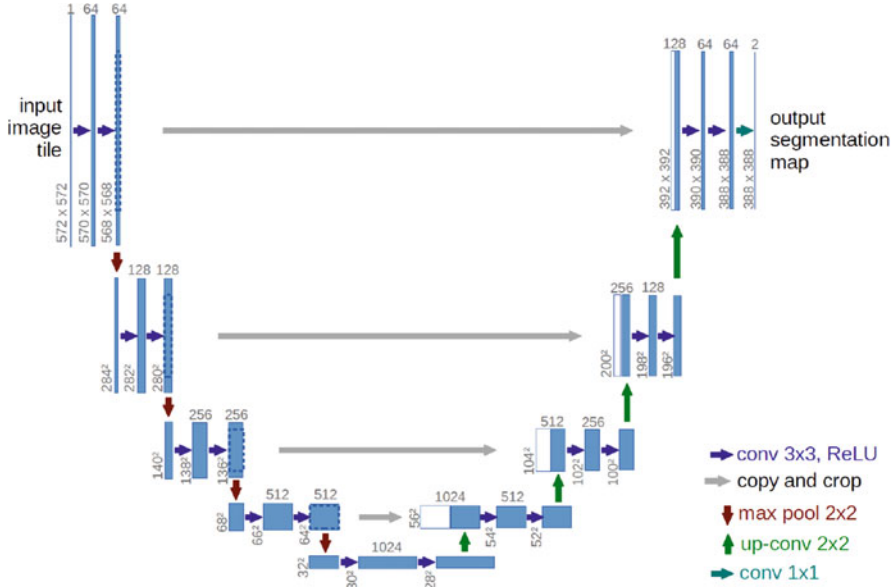
Fig. 8.8 U-net architecture (e.g., 32 × 32 pixels). The blue box represents the feature map. x, y, size is provided at the lower left edge of the box. The number of channels is denoted on top of the box. White boxes represent the copied feature maps

comprise RGB band and all the remaining channels go into A(1195–2365 nm), M(400–1040 nm), and P(450–690 nm) bands. The labels in the training dataset contain roads, tracks, trees, waterways, crops, vehicles, and miscellaneous data. These images have the wavelength of 400–1040 nm with multispectral range. These images are captured at nadir with the sensor of following specifications: panchromatic (the sensitivity of lens) ranging from 1.24 m, and SWIR (shortwave infrared): 14-bits per pixel. One square kilometer is captured in each satellite which is present in the dataset. Below are few test samples of the training dataset (Fig. 8.9).

8.7.2 Preprocessing and Mask Generation

The images are of TIFF format; to make them ready for network training, we need to preprocess them. Initially, the work was carried out in finding the building in the satellite images; to label them we create masks using Computer Vision for final outputs of the satellite buildings. We used the grid sizes which were given in the training datasets and scalers for finding the ground labels of the polygon masks of the structures. From the original images, the building data is stored in the forms of polygons and is stored in numpy arrays. Numpy is a Python framework which is used for n-dimensional representation and manipulation of data. The masked images
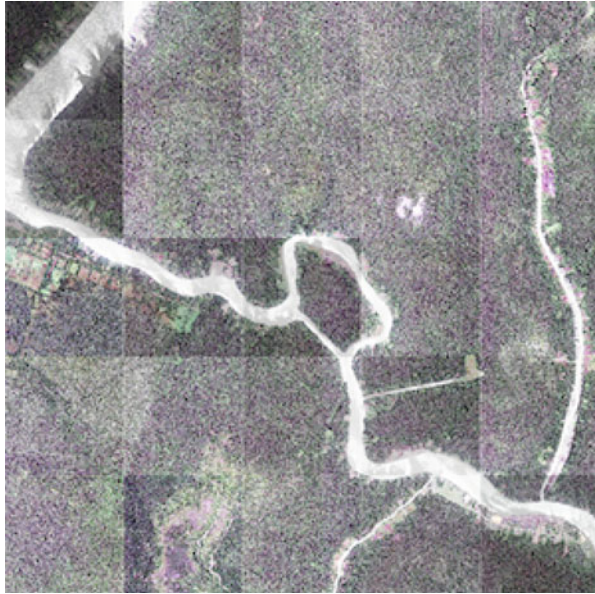
Fig. 8.9 Peek over dataset
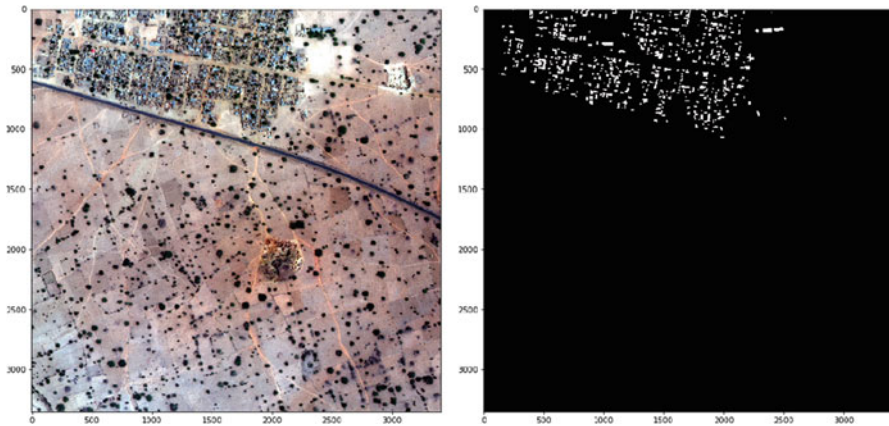
Fig. 8.10 Generated mask for sample satellite image 1

are generated using these polygon arrays which are used as labelled data for the training. All the masked pictures are of resolution 512×512 since the dimensions of the Neural Network should always be the same for every training image. These masks are generated for satellite images that are present in the three band training set (Figs. 8.10, 8.11, 8.12, and 8.13).
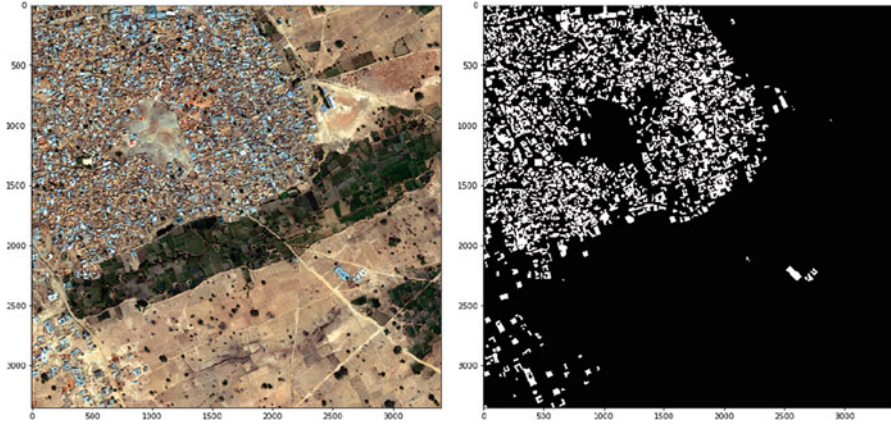
Fig. 8.11 Generated mask for sample satellite image 2
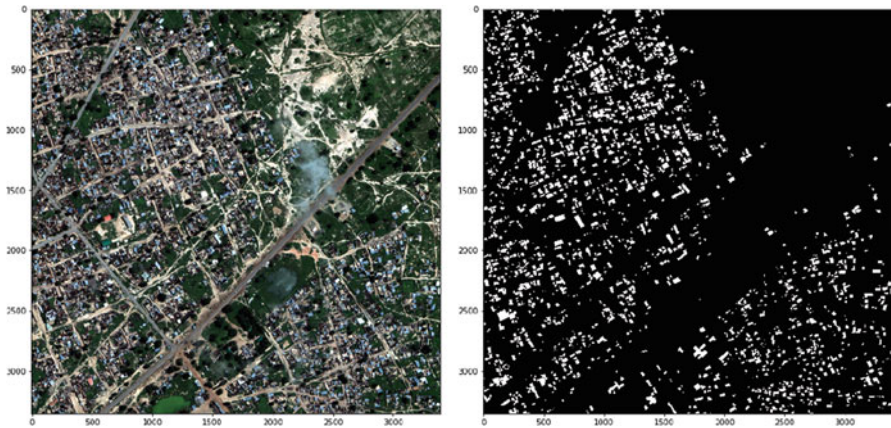
Fig. 8.12 Generated mask for sample satellite image 3
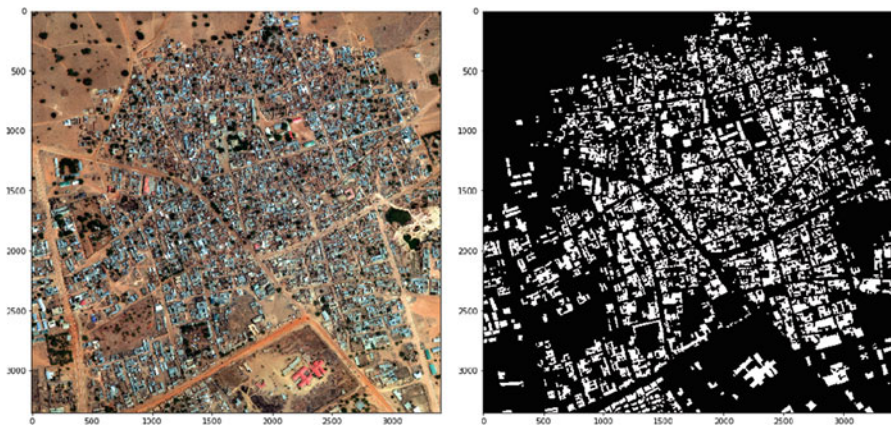
Fig. 8.13 Generated mask for sample satellite image 4

8.8   Convolutional Neural Network for Finding the Buildings in the Satellite Images

We developed a nine-layer convolutional architecture for detecting the buildings in the given satellite images; the same architecture can be used for finding the other classes in the given inputs like water, trees, and crops, etc. The proposed architecture takes the input information as the dimensions of the input image with the labelled masks.

8.8.1   Proposed Model

In the proposed nine-layered convolutional architecture, the first four convolutional layers are followed by max-pooling operation; these help the network to learn the highly concentrated features of the images here, in this case, the buildings. Once the data is passed through the first four layers, the fifth layer is immediately followed by one more convolutional layer without any max-pooling layer as the images might lose its features and deform; in this case we used a new sampling layer which is used to increase the image resolution and also to highlight the main elements in the image. From the sixth layer, the convolutional layer is preceded by an upsampling concentration layer for bringing back the original resolution of the image, and also for training the model with proper stride and padding values.

8.8.2   Training the Model

Neural Networks take time to train on a standard CPU as the data is vast and it needs to be updated for every input. Hence, we prepared the following U-Net architecture on a GPU over 2000 images which took around 8 hours of training time. We implemented several Neural Network architectures but were finally convinced with the U-Net architecture as it has shown outstanding results for the previously mentioned problems. The parameters that are mainly considered for the network training are loss function, optimizer, learning rate, number of epochs, and the accuracy metrics. For this U-Net architecture, the implemented loss function was Adam optimizer which was used for all the considered 25 epochs of training. Binary cross entropy was used for optimizing the loss throughout the training. These hyperparameters gave an accuracy of 98% in detecting the buildings in the satellite images. 1e-4 was the learning rate that the networks were trained with; it is the rate at which the loss gets updated each time until the network reaches the minimum gradient descent. This training was done on an NVIDIA GPU using a Python framework named Keras. The training and validation set were divided in the 30–70 ratio. The validation set is used for finding the metrics of the trained model. Below is the image of a sample predicted image (Fig. 8.14).
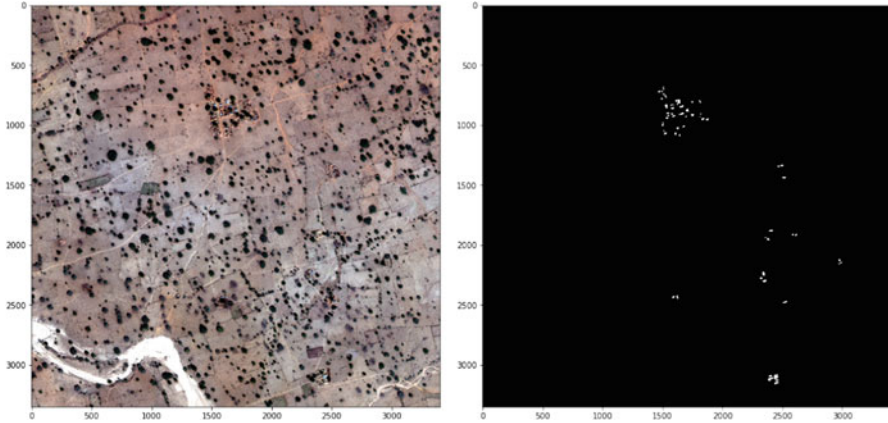
Fig. 8.14 Predicted output and generated mask

8.9 Generic U-Net for Finding Several Classes in the Satellite Images

The primary motivation for making the architecture generic is by using the Jaccard index J(A,B). This Jaccard index is used for finding the intersections over union and intersection coefficient values of the given inputs. To see the Jaccard index, we need to carry two images as the input, one the original image, the other will be a masked image. This returns the percentage of the coincidence and the intersection of the images. For the model, the prediction for each class was evaluated independently using average Jaccard index (also known in the literature as Intersection over Union), and the class-wise scores were averaged over all ten labels with equal weights. Mathematically, Jaccard index is given by

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{8.14}$$

8.9.1 Preprocessing Images for Network Training

Initially, before the inputs are passed into the Neural Network architecture, few preprocessing steps are applied to the training dataset. As we described the dataset in Sect. 7.1, first we change the image resolution of every TIFF file to a unique 3600×3600 pixel image. Since the pictures are of a different resolution in the dataset, we need to make a constant proportion of the input dimensions which is equal to the U-Net architecture, for which we consider the above-used decision as the ideal one. We give the image padding of one as we know that the resolution of the image changes wherever a convolutional kernel is applied to the image pixel

matrix. The filling is used for all the three bands which are R, G, and B. Once the masks are prepared for all the classes, these are divided into grids which are used for counting and segregation of the labelled classes. Once the input of these preprocessed images is sent to the U-Net, we find the predictions for each category, and then using Jaccard index, we calculate the accuracy, and then in the output image, we merge the forecasts of all the classes. Basically, combine predictions back into the original size. The outputs that are generated by the convolutional network include padding and are projected upside down; the projections should be changed by rotating the axis into 90 degrees horizontal, and the extra padding is removed.

We implement a similar procedure for detecting 11 other classes, the way we have done for building detection in an image previously. The classes include roads, tracks, trees, waterways, crops, vehicles, and miscellaneous data. Since the classes are different and the whole data attaining a higher accuracy would be difficult, we train separate models for each and every class. The training dataset comprises the original satellite images and the mask images which are sent into the U-Net model. We use Nadam optimizer, i.e., Adam with Nesterov momentum for the optimization purposes while training the U-Net model. The network is trained for 25 epochs with a learning rate of $1e^{-3}$ and an extra 25 epochs using a learning rate of $1e^{-4}$. The number of batches is 64, and each batch has 100 images. The images are cropped into batches and fed into the network.

8.9.2 Model Implementation

We developed a ten-layer convolutional architecture for detecting the objects in the given satellite images. The first four convolutional layers have max-pooling operations and batch normalizations which help the network to learn the most focused features, and the latter one helps in normalizing the pixel values in the image. The fifth layer has no max pooling and is immediately followed by upsampling to generate the pictures again, i.e., deconvolution happens to create the images with a high resolution. The four layers, starting from fifth and ranging until ninth, will include upsampling, convolution, batch normalization, concatenation, and activation (ELU (Exponential Linear Unit)) applied to the concatenated layers. The standard filter size used across the whole net is 3×3. The final tenth convolution layer is where the output is generated using the sigmoid activation function.

8.9.3 Loss Function

The metric used is average Jaccard index. The most common loss function used is categorical cross entropy, but the classes in our dataset are mutually exclusive, so binary cross entropy is most used. However, Jaccard index helps us in understanding the intricate details in a complex image.

$$\text{loss} = \text{binary\_cross\_entropy} - \log{(\text{jaccard\_approximation})} \qquad (8.15)$$

$$\text{binary\_cross\_entropy} = -\sum{(y_{\text{true}} \log{y_{\text{pred}}} + (1 - y_{\text{true}}) \log{(1 - y_{\text{pred}})})} \qquad (8.16)$$

$$\text{jaccard\_approximation} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum{y_{\text{true}} * y_{\text{pred}}}}{\sum{y_{\text{true}}} + \sum{y_{\text{pred}}} - \sum{y_{\text{pred}} * y_{\text{true}}}} \qquad (8.17)$$

Below is the diagram of the implemented model with all the input and output dimension connections and a total number of params used (Fig. 8.15).

8.10 Experimental Results and Discussion

The below-predicted diagram is concerning the detection of buildings in satellite images, and the same generic U-Net architecture, described above, is used for the detection of other natural resources as well. The buildings have been highlighted in blue to separate them from the rest and for the ease of identification. The same procedure can be applied to trees, roads, waterways, etc. which wholly contribute to detecting various objects in the satellite images (Fig. 8.16).

We have identified the natural features like buildings, roads, crops, waterways, trees, and tracks using Convolutional Neural Networks. The model is generic for all the features and is developed to identify the natural resources persisting in the satellite images. This is trained as per the parameters mentioned in Sect. 8.8.2. The model is trained over 2000 iterations to achieve maximum accuracy. We define epoch as a single pass through the whole training dataset. We measure the model's loss and accuracy with respect to epoch. As the number of epochs increases, the loss decreases indicating the indirect proportionality existing between loss and the number of epochs. Similarly, the model's accuracy increases as the number of epochs increases indicating the direct proportionality existing between the accuracy and the number of epochs.

Below are the generated graphs for the trained model which gives the visualization between the loss, accuracy, and epochs (Figs. 8.17 and 8.18).

8.10.1 Software and Processor for Model Training

We used four test-time augmentation for generating masks and preprocessing the images since satellite images are entirely captured in raw. It was complicated on a standard CPU to manipulate and preprocess them in a conventional RAM. We trained our model on a P5 GPU of four CPUs and a single NVIDIA GPU GTX 1080 of 12 GB memory also known as Titan on a cloud infrastructure; the CPU is configured with Keras, a Python framework. Each epoch took over 169 seconds on

Fig. 8.15 Proposed U-Net model architecture for identifying natural resources in satellite images
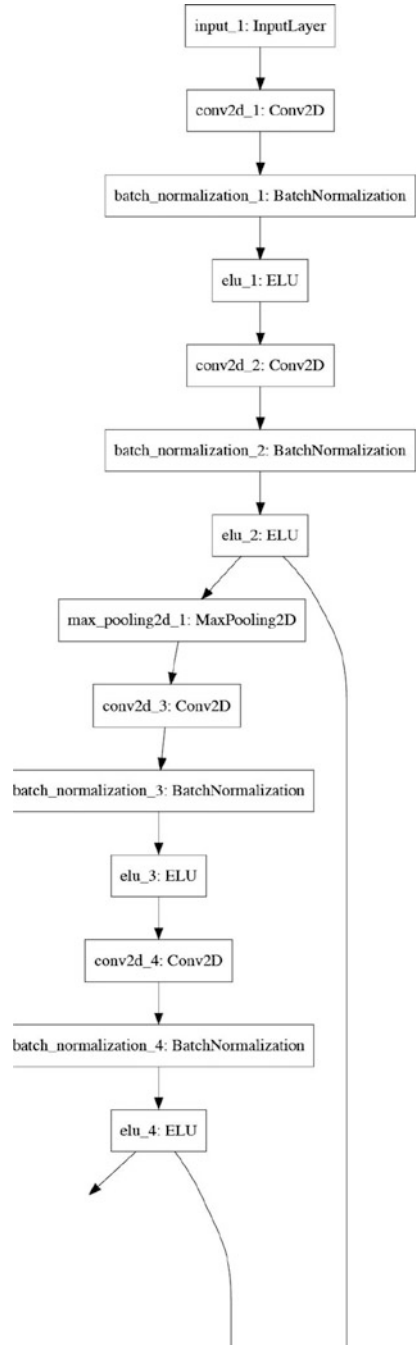
Fig. 8.15 (continued)
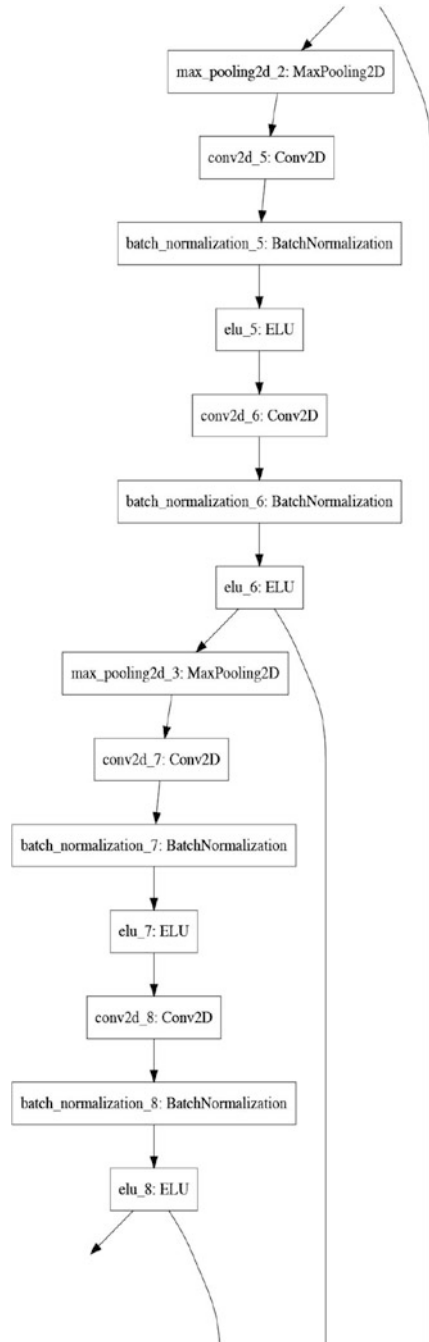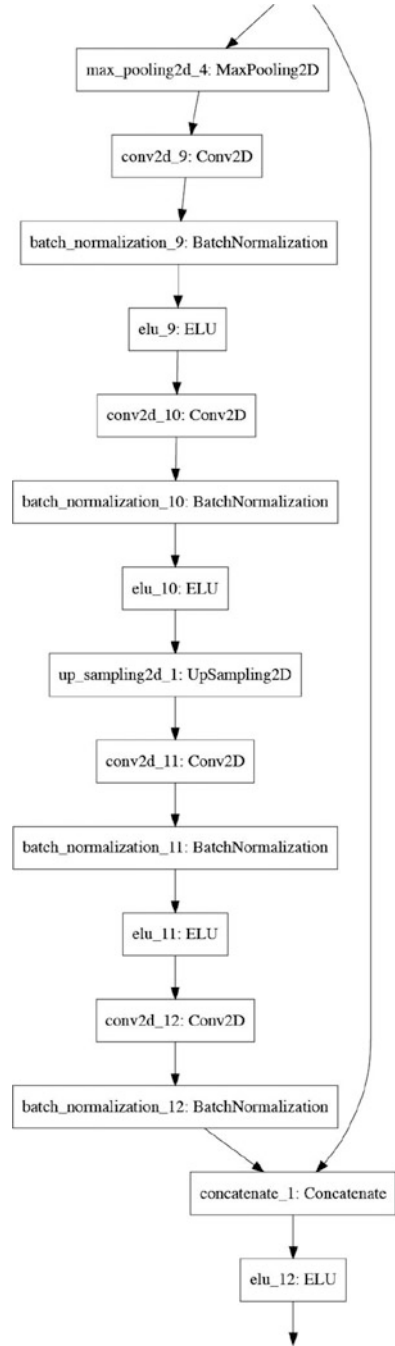
Fig. 8.15 (continued)
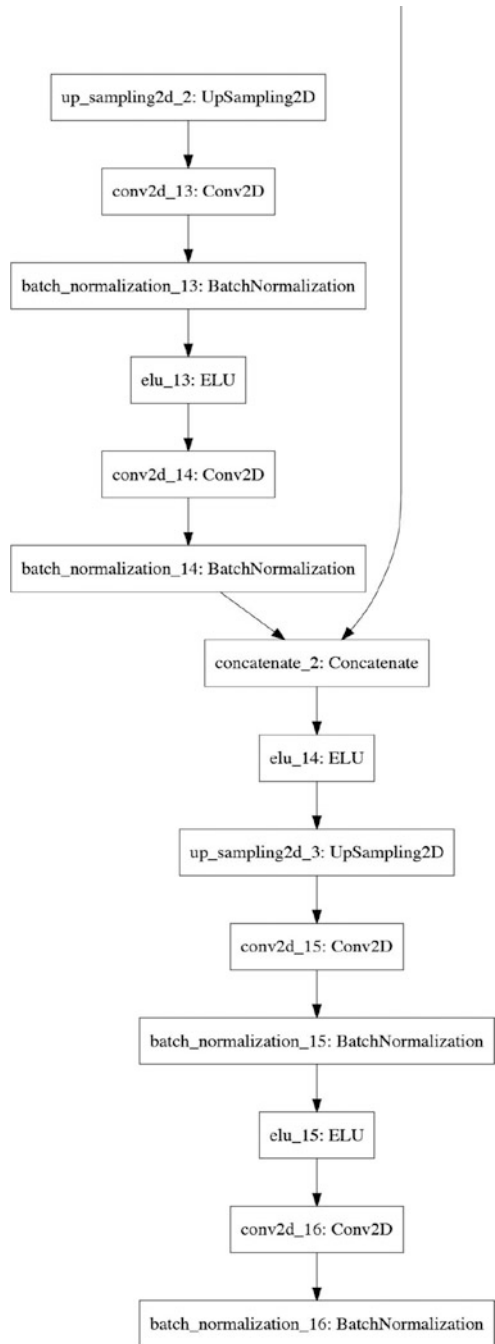
Fig. 8.15 (continued)

Fig. 8.15 (continued)
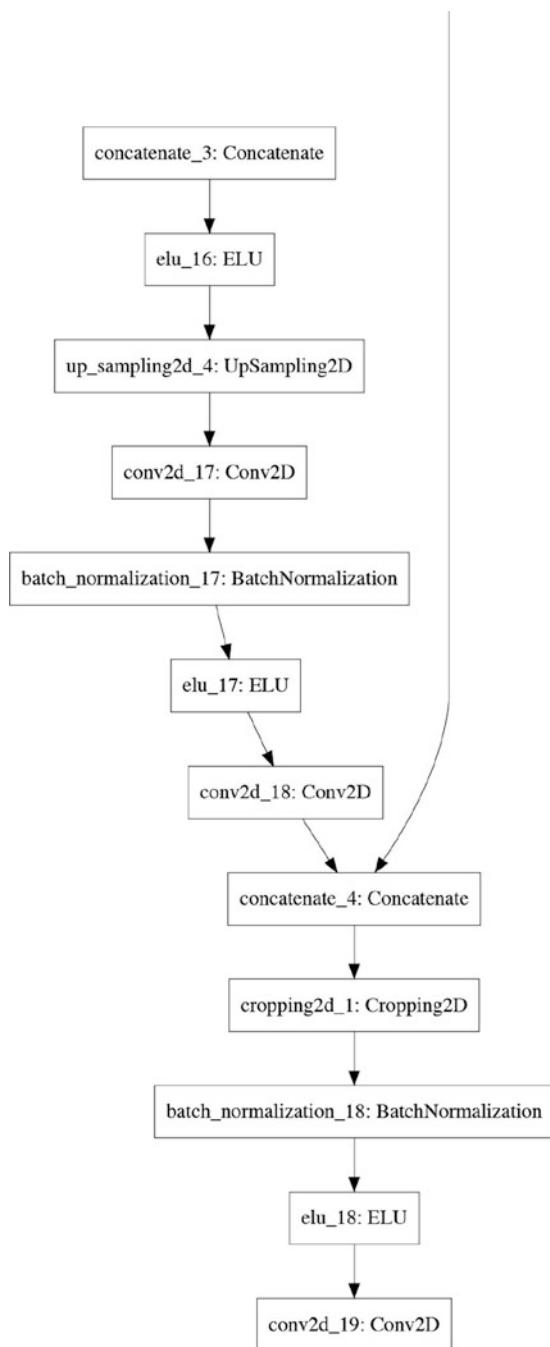
Fig. 8.16 Prediction from trained model
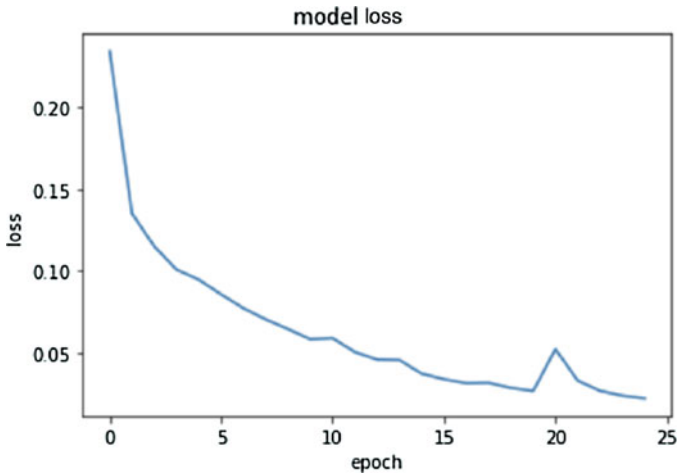
Fig. 8.17 Graph between model's loss and epochs

a GPU and over 35 min on a standard CPU which is almost 97% more time that is consumed on the CPU. With this, we could successfully achieve identifying the features and objects that are present in the satellite images using deep Convolutional Neural Networks through semantic segmentation.
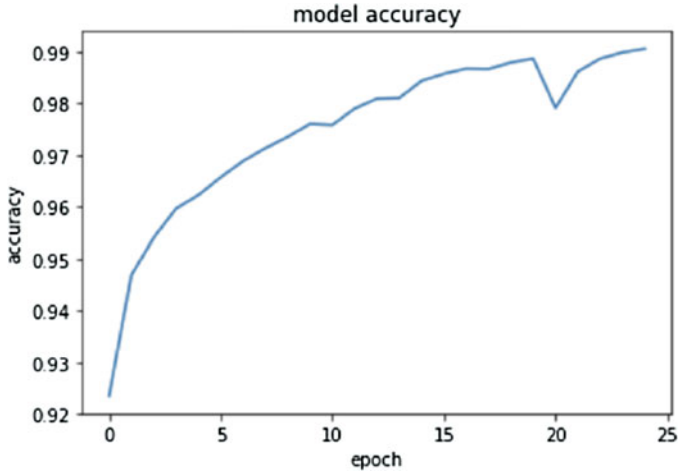
Fig. 8.18   Graph between model's accuracy and epochs

8.11   Conclusion

In this research, we have used Convolutional Neural Networks for detecting natural features in the satellite images and also reviewed all the available image segmentation techniques that are used widely employed for feature detection. The images are preprocessed and then trained on a Neural Network for segmenting the natural features with one another. We proposed an upgraded U-Net convolutional architecture and used the satellite images. The loss, metrics, and other statistical attributes are explained based on the network parameters used and mentioned in Sect. 8.8.2. The efficiency of the image segmentation is calculated on the basis of IoU (Intersection over Union) values and mask overlays for identifying each natural resource in the satellite images. We could achieve finding features like roads, building, and waterways in the given satellite images using this proposed generic model. This model is also widely dependent on the hardware requirements as the training is processed in the TIFF images. Further works include overlay masking of segmented images and high resolution of the images using deconvolutional techniques.

References

1. Activation functions (2015) https://ml-cheatsheet.readthedocs.io/en/latest/activation-functions.html
2. Altenberger F, Lenz C. A non-technical survey on deep convolutional neural network architectures. arXiv:1803.0219

3. Ciresan DC, Gambardella LM, Giusti A, Schmidhuber J (2012) Deep neural networks segment neuronal membranes in electron microscopy images. In: Advances in neural information processing systems, pp 2852–2860
4. Farabet C, Couprie C, Najman L, LeCun Y (2013) Learning hierarchical features for scene labelling. IEEE Trans Pattern Anal Mach Intell 35:1915–1929
5. Girshick R, Fast R-CNN, Microsoft Research. arXiv:1504.08083
6. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. Facebook AI Research (FAIR). arXiv:1703.06870
7. Iglovikov V, Mushinskiy S, Osin V. Satellite imagery feature detection using deep convolutional neural network: a kaggle competition. True Accord, Open Data Science, AeroState. arXiv:1706.06169
8. Koushik J (2016) Understanding convolutional neural networks. Language Technologies Institute, Carnegie Mellon University, Pittsburgh. arXiv:1605.09081
9. LeCun BB, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1990) Handwritten digit recognition with a back-propagation network. In: Advances in neural information processing systems
10. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
11. Long J, Shelhamer E, Darrell T (2014) Fully convolutional networks for semantic segmentation. UC, Berkeley. arXiv:1411.4038 [cs.CV]
12. Ma Z, Wang Z, Liu C, Liu X. Satellite imagery classification based on deep convolution network
13. Neural networks and back propagation explained in a simple way (2017) https://datathings.com/blog/post/neuralnet/
14. Ning F, Delhomme D, LeCun Y, Piano F, Bottou L, Barbano PE (2005) Toward automatic phenotyping of developing embryos from videos. IEEE Trans Image Process 14(9):1360–1371
15. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation
16. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536
17. Schaeffer J, Burch N, Björnsson J, Kishimoto A, Müller M, Lake R, Lu P, Sutphen S (2007) Checkers is solved. Science 317(5844):1518–1522
18. Stanford course material CNN (2018) http://cs231n.github.io/convolutional-networks/
19. Uijlings J, van de Sande K, Gevers T, Smeulders A (2013) Selective search for object recognition. IJCV 59:167–181

Chapter 9
Change Detection of Tropical Mangrove Ecosystem with Subpixel Classification of Time Series Hyperspectral Imagery

Dipanwita Ghosh and Somdatta Chakravortty

Abstract This chapter aims to use hyperspectral imagery to categorize saline blank classes amidst mangrove mixtures and analyze its changing patterns in the Sunderban Mangrove Forests of West Bengal, India. This research derives fractional abundance of mangrove endmembers at subpixel level with Fully Constrained Linear Spectral Unmixing (FCLSU) based on Least Square Error optimization criteria. NFINDR algorithm has been applied on time series hyperspectral image data of 2011 and 2014 to recognize pure saline blank and mangrove endmembers in the thickly forested study area followed by FCLSU to estimate mangrove species distribution maps of 2 years. The estimates in location 21° 34′ 24.81″N and 88° 17′ 36.89″E indicate a pure saline blank patch showing 74.47% occurrence with *Phoenix paludosa, Avicennia alba,* and *Ceriops decandra* showing 9.87%, 12.67%, and 2.99% presence in 2011. In 2014, the coordinate shows an increase in occurrence of saline blanks and *Ceriops decandra* but reduction in *Phoenix paludosa* and *Avicennia alba*. *Ceriops decandra* are salt-tolerant mangrove species that show an increase in abundance with increase in saline blanks. *Phoenix paludosa* which is salt intolerant shows a decrease in abundance with increase in saline blank areas. It is observed that mangroves, namely, *Excoecaria agallocha* and *Ceriops decandra,* are common and dominant around the saline blank areas. Salt-tolerant mangroves such as *Avicennia marina* and *Avicennia alba* are also observed to survive in certain locations of saline blanks.

Keywords Change detection · Mangroves · Saline blanks · Time series data · Spectral unmixing · Least square error

D. Ghosh · S. Chakravortty (✉)
Maulana Abul Kalam Azad University of Technology, West Bengal, Kolkata, India

9.1 Introduction

This chapter tries to utilize hyperspectral imagery to recognize the changing pattern of saline blank and mixed mangrove forest with time series data. In this study, remote sensing can be used as an effective tool for research and management of tropical mangrove forests of large areal extent and identification of the saline blanks that remain interspersed within it. It is a common observation that within deep mangrove forests, certain open patches exist that remain non-vegetated in terms of tree or shrub or mangrove cover but marked by presence of grasses and other saline aquatic species [1]. These are known as saline blanks. These are formed as a result of geological action by coastal processes that operate at the mouth of pro-grading delta by way of continuous sedimentation process. With a steady increase in salinity, these areas gradually get devoid of mangrove cover. As a result of differential accumulation of sediments (brought down by rivers from terrestrial sources and further reworking by tidal influences), there occurs differential settling, compaction, and subsidence at some parts and swelling at the other (leading to formation of saline blanks). If the sediment deposition continues, these saline blanks get further elevated into slightly elevated areas with development of more such saline blanks in the adjoining areas. Finally, with less flushing, the saline blanks assume the marginal nature of "beyond-intertidal zone." It is observed that prevalence of mangroves with high pneumatophore density and better stilt-root system further facilitates aggravation of saline blank formation, as stilt-root system acts as excellent sediment trappers and helps in entrapping of sediments within the root-encased network spaces [2]. Saline blanks may be also caused by upswelling of tidal flats by tectonic causes, as observed in case of Digha-Contai region of southern Bengal.

The existence of saline blanks has got tremendous implications and significance for growth and distribution of mangrove ecosystem, and its impacts vary from one species to the other. For example, saline blanks represent disturbed state of ecological equilibrium for certain climax mangrove species such as Sundari (*Heritiera fomes*), Dhundul (*Xylocarpus granatum*), and Possur (*Xylocarpus mecongensis*) which have great economic importance in terms of timber value. For these species, saline blanks create a strong negative impact and deter their growth and proliferation [3, 4]. On the contrary, incidence of saline blanks is beneficial for certain other groups of mangroves like *Avicennia marina* and *Avicennia alba* which have stilt-root system with high density (therefore can trap more sediments in erosion-prone areas) and prefer more saline environment (salt marsh/saline blanks) for their survival. But *A. marina* and *A. alba* represent adverse habitat in terms of quality timber production and severely affect the prey success of tigers [5, 6]. The mangroves of Sunderban symbolize one of the most endangered and susceptible ecosystems that has gone through a disturbing reduction at spatial level and species variety over the past few decades. It is therefore essential for a complete species level

categorization and mapping of mangroves that would assist to give the idea of the amount of damaged one to this floral community. The application of image classification algorithm for identification of saline blanks and related mangrove species with time series hyperspectral data is a sole attempt in India for Indian mangroves. Additionally, change detection of saline blanks and identification of patterns of significant mangrove species that surround these fundamental locations are barely mentioned in national and international forum. Hyperspectral data have high spectral resolution of 10 nm bandwidth which allows accurate classification of mangroves at species level. When this data is combined with subpixel level classification algorithms, they are likely to exhibit the possibility for consistent and comprehensive categorization of mangrove forests together with leaf area, canopy elevation, and biomass. The large number of narrow spectral bands of hyperspectral data may be helpful in detailed-level mapping such as mangrove species identification, assessment of plant stress, and more. The results so far achieved indicate great potentialities for HSRS in future years. However, lack of proper Hyperspectral Image Processing tools and techniques pose an impediment to high-end hyperspectral research. Hyperspectral images captured from "airborne sensors" are relatively expensive but feature wide array of beneficial applications as well. With launching of more numbers of Spaceborne Satellite Hyperspectral Sensors, the disadvantage of extra cost would be lowered. The huge number of fine spectral bands in hyperspectral imagery also leads to computation exhaustive processing steps and strongly correlated data. The spectral bands that contain information depicting maximum spectral differences among mangroves species are considered most suitable for mangrove discrimination and mapping. In this research, signatures of mangrove endmembers have been extracted using NFINDR algorithm and taken as input for spectral unmixing of mixed pixels comprising saline blank and mangrove endmembers. The fractional abundance calculations on time series hyperspectral data indicate that, with time, as saline blanks become more profuse in the island, the existence of salt-tolerant mangroves increases; however, the abundance of salt-intolerant ones reduces. However, in areas where the mangrove species still rule and saline blanks are in its shaping phase, there is still a likelihood of regrowth of the salt-prone mangrove species.

A study has been completed for uncovering changes in the distribution of mangroves within Moreton Bay, southeast Queensland, Australia [7]. They have used two methods for assessment, such as spatial- and time-based pattern analysis method and analysis of change identification methods. In the first method pattern, metrics analysis provided the detail information about mangrove pattern changes in this region and change detection analysis gave more detail information about spatially explicit changes of mangrove pattern. Lastly, it was concluded from this study that the amount of mangroves are reduced day by day due to the effect of increasing urban development, agriculture, and aquaculture. Moreover, it was

also realized that the landward mangroves are sustained and seaward mangroves are mainly destroyed for the increasing salinity level. Another author studied that the mangrove forests change according to their dynamics [8]. These changes were estimated by using multi-temporal satellite data. It was also realized that these change has been influenced mainly by anthropogenic forces. The forest was changing due to erosion and degradation of forests. Connection of NDVI statistics with canopy has been used for detecting change in this study. This remote sensing technology is proved to be fruitful for this study. According to the result of this study, it is concluded anthropogenic forces should be reduced to sustain the mangrove dynamics. In [9], the authors highlighted on the land cover changes, especially in mangrove forest area, shrimp farms, urban areas, and agricultural land. They have used NDVI composite image as remote sensing tool to detect the change in the land cover. It was concluded that this technique was very useful for detecting the extent of mangrove forest and other land. In another study, authors have observed the change pattern of mangroves with the change of rainfall [10]. They have observed changing pattern of mangroves over the period of 32 years including wetter period and drier period, using remote sensing technology. From this study, a strong relationship has been established between rainfall variables and landward mangroves expansion. So remote sensing technology is very efficient for this study. In [11], change detection has been done using multi-temporal assessment of data that proved useful for mapping deforestation/regeneration and mangrove dynamics associated with varying patterns of sedimentation process. In this study, ALOS PALSAR data and L-band SAR images have been used for mangrove forest detection. So this type of satellite imagery was proved to be useful for the detection of pattern changes of mangrove forestry. In [12], object-based method has been used for mangrove mapping. This method was tested on satellite imagery and SPOT XS data. It identified the pattern change of mangrove species between the year 1986 and 2000. This also identified the growing pattern of mangroves with the change of rainfall. So in this study, this technology gave the best result to identify the change of mangrove patterns.

9.2   Study Area

The mangrove ecosystem of the well-known Sunderban Delta of West Bengal is identified for this research study. This study will concentrate on the Henry Island of Sunderbans, West Bengal (Fig. 9.1). The total area of Sunderbans is around 10110 sq. km. that spreads over the two neighboring countries of India (West Bengal) and Bangladesh. Out of this, the Indian portion of Sunderbans covers 4110 sq. km., while the rest forest patch lies in Bangladesh part.

Henry represents the westernmost border of the Indian part of Sunderban. It is among the 102 islands of Indian Sunderbans that face the impact of coastal hazards.

Fig. 9.1 A Glimpse of a typical mangrove habitat at Henry Island, Sunderban, India

Named after a British surveyor, the Henry Island is typically Criss-Crossed by numerous coastal streams and rivulets and offers a paradigm mix of mangrove forests interlaced with saline blanks and beachfronts. The area of the island is only 10 sq. km. and falls within the administrative jurisdiction of South 24-Parganas district of West Bengal. The historic Dumpier-Hodges line (the imaginary line that marks the northernmost jurisdiction of Sunderban) passes well above the northern edge of the Henry Island. Literature survey reveals that the Henry was once within the core Sunderban forest area. With increasing human aggression and consequent pollution, there has been a fast recession of mangrove forests from this area. Yet, the area still supports a rich biodiversity, including luxuriant growth of mangrove vegetation. The mangrove habitat of Henry Island extends from 21° 36′ 00″N to 21° 34′ 00″N latitude and 86° 16′ 30″E to 88° 18′ 30″E longitude (Fig. 9.2).

9.3 Methodology

A flowchart of the methodology followed is presented as under:
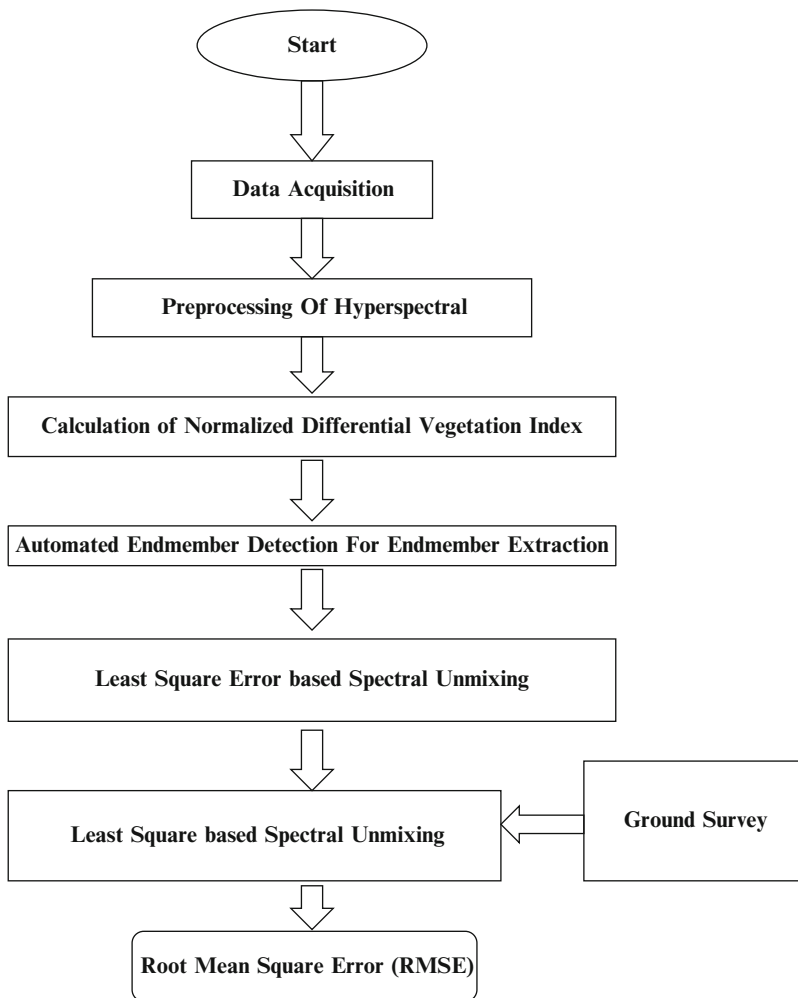
9.3.1 Data Acquisition

Hyperspectral data that was acquired on 27 May 2011 and 24 November 2014 by the Hyperion sensor on the Earth Observatory-1 (EO-1) satellite (Fig. 9.2) has been used in this research study. The study is based on hyperspectral imagery specially captured by the United States Geological Survey (USGS) under Data Acquisition Request made in proper format by the researcher. The system was essentially programmed to perform a trajectory over the study area for the sake of capturing the forest imagery for present research purpose. The wide spectral range of Hyperion data (ranging from 380 to 2500 nm of electromagnetic radiation) and high spectral

Fig. 9.2 Hyperspectral image of Henry Island located in Sunderban. (Courtesy: United States Geological Survey)

resolution (10 nm bandwidth and less) allow accurate detection and classification of surface canopies and ground features, through application of superior Hyperspectral Image Processing algorithms. The acquired hyperspectral image strips of 2011 and 2014 are shown in Figs. 9.3a and 9.3b.

9.3.2 Ground Survey

The study area offers an ideal backdrop for application of hyperspectral remote sensing for mapping of wide range of pure and mixed patches of mangrove species [2, 6].The ecological feebleness of the Sunderban rainforest is plainly visible by the declining density and diversity of the mangroves of Henry Island. The delicate nature of the mangrove ecosystem could be effectively interpreted through study of diverse endmembers (mangrove species) that cast a deep weight on the integrated whole.

The mangroves of Henry Island of Sunderban (the study area) are dominated by *Excoecaria agallocha, Ceriops decandra, Avicennia marina,* and *Avicennia alba* which constitute about 70% of the mangrove species in the Indian Sunderbans. Out of the 94 species of mangroves reported from the total Sunderban, only 7 pure (homogeneous) patches of species have been traced in this island. The habitats of a wide range of other mangrove species of Sunderban are practically inaccessible to

Fig. 9.3a Hyperion image
strip of Sunderban of 2011

Fig. 9.3b Hyperion image
strip of Sunderban of 2014
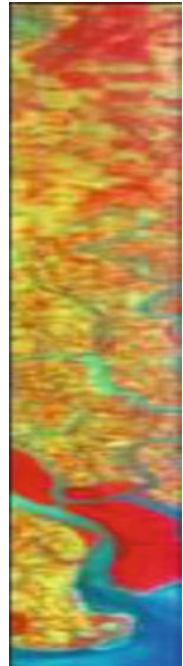
man. The mangroves of Henry are clustered in zones with dominance of species *Avicennia* and *Excoecaria* on the seaward edge. Away from the waterline lie the mixed patches of *Bruguiera cylindrica, Phoenix paludosa, Ceriops decandra,* and *Excoecaria agallocha.* Sundari is in decline in Henry owing to high salinity hazard. Pure patch of Sundari in a 30 m × 30 m matrix is virtually absent here. Although artificial afforestation programs of Sundari have been taken up by the Forest Department at select patches, the efforts are far from satisfactory. The dense mangroves forest of Henry is crisscrossed by numerous small canals with evidence of wildlife, namely, deer, wild pig, crocodile, and poisonous snakes, posing threats to man.

Ground survey was conducted to recognize and gather samples of mangrove species whose image categorization has been carried out. This was made in June 2011 on immediate acquisition of data. GPS (Global Positioning System of Make: GARMIN GPSMAP 78 s with high-performance marine handheld 3 axis compass and barometric altimeter and accuracy of 4 m) has been used to accurately position the geographical coordinates of the study track. Field Quadrat Survey-cum-Sample technique was adopted to evaluate the regularity of species dominance. The quadrat dimension was taken as 30 m × 30 m that is equal to the spatial resolution of Hyperion imagery. Thirty quadrat plots were chosen in "pure" and "mixed patch" forests that were plotted on imagery with GPS. Diverse species present within each quadrat were counted. The species with more than 50% abundance was considered as "dominant," and the corresponding pixel represents a "pure pixel." The quadrats in which none of the mangrove species achieve 50% dominance are chosen a "mixed patch." These ground survey areas are used as reference points for exactness evaluation of image-extracted findings (Table 9.3). Ground truthing exercise was again repeated in May 2012 and June 2013. Some changes in physical abundance of individual plants could be noted within this time span.

9.4 Calculation of Normalized Differential Vegetation Index (NDVI)

As the main objective of this study is to identify saline blanks and related mangrove species of the study scene, we have removed the cloud zones, pisciculture, and water bodies to focus only on saline blanks and mangrove endmembers for faster processing. The mangrove-forested region has been retained by evaluating the Normalized Differential Vegetation Index (NDVI) [13]. NDVI is based on the difference between reflectance between two bands, red and infrared. Studies have shown that upper ranges of NDVI values specify the occurrence of vegetation. The lesser index range between 0.1 and 0.2 indicates soil. Hence, the upper range of NDVI values (0.5–1.0) and the lower range between 0.1 and 0.2 have been retained and the other areas removed. Now, we have an image of only the forested portions and soil of the study scene.

9.5 Automated Endmember Detection Algorithm for Endmember Extraction

Endmember identification determines the set of pure spectra present in a hyperspectral image. N-FINDR algorithm aims to locate the simplex of maximum volume that can be inscribed within the hyperspectral data set by means of a simple nonlinear inversion. This study applies NFINDR algorithm for programmed endmember detection of saline blank and mangrove species endmembers.

In N-FINDR, the convex nature of hyperspectral data helps determine pure pixels in an image by locating points in the imagery that correspond to the vertices of the simplex containing the data [14, 15]. The N-FINDR algorithm assumes that the N-dimensional volume created by a simplex is at all times bigger than that created by other arrangement of pixels. It is understood that the quantity of existing endmembers, p, in the data is known a priori. After estimation of pure endmembers in the data set, the realization of the algorithm is initialized by a simplex whose vertices are formed by an arbitrary set of pixels used as opening endmembers chosen from the data set. The algorithm then estimates the volume of the simplex defined by vertices. For each image pixel vector, it recalculates the dimensions of p simplices, each of which is formed by replacing one pixel with the sample vector. If none of these recalculated volumes is larger than the previous volume, no endmember sample is replaced. The similar process is repeated for other vectors until all pixel vectors in the original data set have been utilized.

9.6 Least Square Error-Based Spectral Unmixing

We consider the measured spectrum of a pixel in a hyperspectral image that is known to contain a single endmember spectra or a combination of spectra of multiple endmembers and may be represented as

$$S = EA + e \tag{9.1}$$

where S represents the mixed spectrum, E represents the individual endmember spectra, and A represents the percentage of each endmember in each pixel.

The objective of the linear unmixing equation is to simultaneously calculate all the difference equations of pure endmembers for all values of their abundances. We solve to find the minimum value of these simultaneous equations that is where the abundance values make this output a minimum:

$$S{-}EA = \min \tag{9.2}$$

The goal is to find the smallest difference which is achieved by least square error equation where we create a function that is the square of this equation and then solve for the minima:

$$S^2 - (E_1 A_1 + E_2 A_2 + \ldots)^2 \qquad (9.3)$$

Linear algebra is used to solve multiple equations with multiple variables, in this case, abundances of endmembers. The best fit values of abundances are calculated, and the percentage of each endmember present within a mixed pixel is estimated. The estimated abundance values are multiplied by their respective pure endmember spectra to get the output pixel intensities. When this is repeated for each pixel in the image, we get fractional abundance images for each endmember, thus generating abundance images for the entire scene. When the abundance values are estimated, two constraints have been considered: Sum of all abundance values is one, and abundance values cannot be negative.

9.6.1   Root Mean Square Error (RMSE)

To authenticate the outcome obtained using linear unmixing model, the fractional abundance values of the endmembers recognized have been evaluated. The total inaccuracy in abundance value of each endmember has been defined as the total difference between the value obtained by the linear model and that by the real ground value of a specific location (coordinate). The RMSE has been estimated as follows:

$$\text{RMSE} = 1/NR \sum_{k=1}^{N} |f_{\text{ck}} - f_{\text{mk}}|$$

where f_{ck} is the real ground value fraction output and f_{mk} is the linear model output; N is the number of endmembers; and R is the number of coordinates (locations) whose RMSE is to be estimated.

9.7   Experimental Results and Discussions

Automated endmember extraction algorithm, NFINDR, has been executed on the hyperspectral images of 2011 and 2014 to extract pure pixels of dominant endmembers present in the study area. NFINDR has successfully identified seven endmembers of mangrove species and two endmembers of saline blanks from the image scene. The coordinates of extracted endmembers and their relevant spectral profiles for both image scenes are shown in Figs. 9.4, 9.5, 9.6, and 9.7. The derived endmembers have been validated through ground survey made in the study trail.
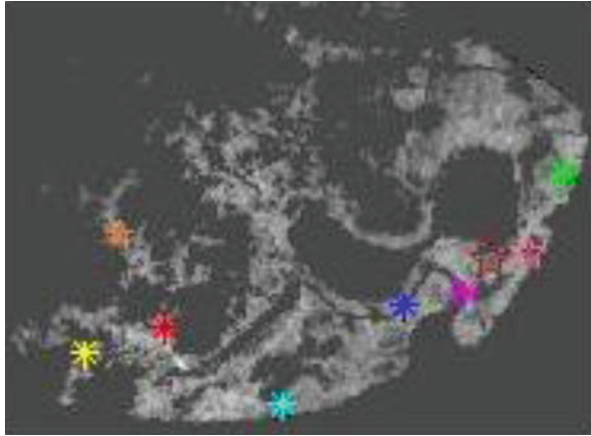
Fig. 9.4 Endmember extracted locations in image of 2011
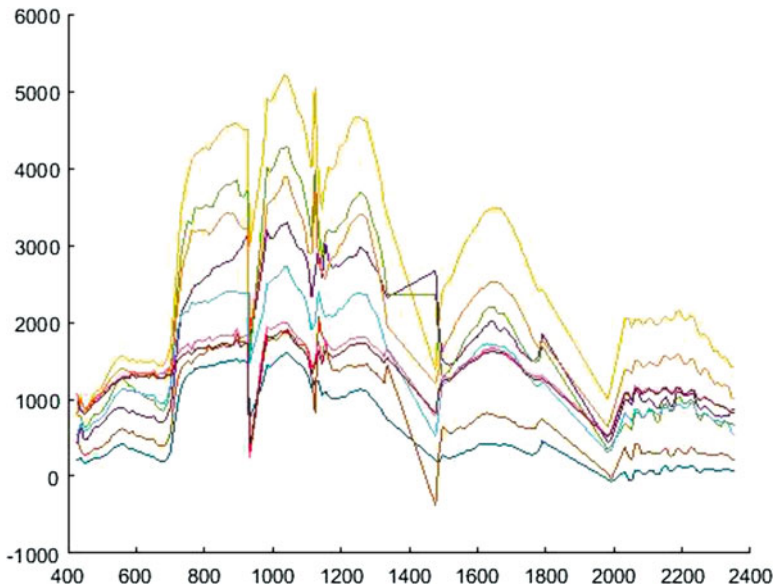
Fig. 9.5 Spectral profile of dominant mangrove species of hyperspectral image of 2011

The unique signatures of saline blank and mangrove endmembers derived from NFINDR have been put in for spectral unmixing of our mixed pixels. Least square error-based Fully Constrained Linear Spectral Unmixing (FCLSU) has been implemented to estimate abundance values of endmembers (subpixels) within a pixel. The abundance mapping of salt-tolerant mangroves and saline blanks in the image scene along with their RMSE values is shown in Tables 9.1 and 9.2.
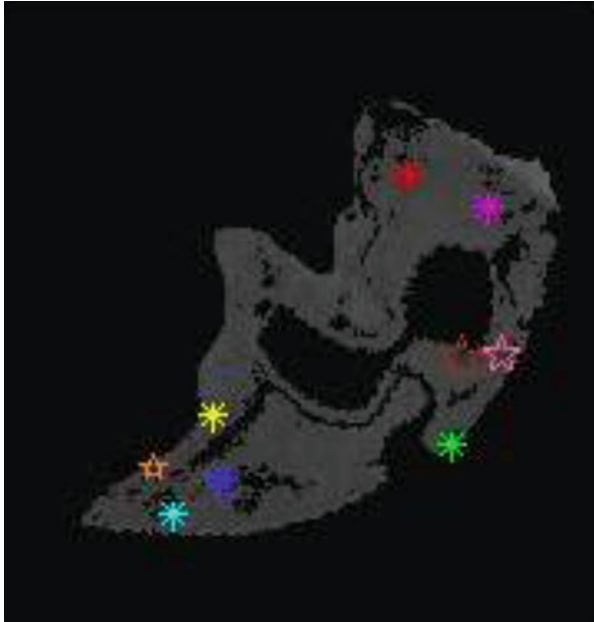
Fig. 9.6 Endmember extracted locations in image of 2014
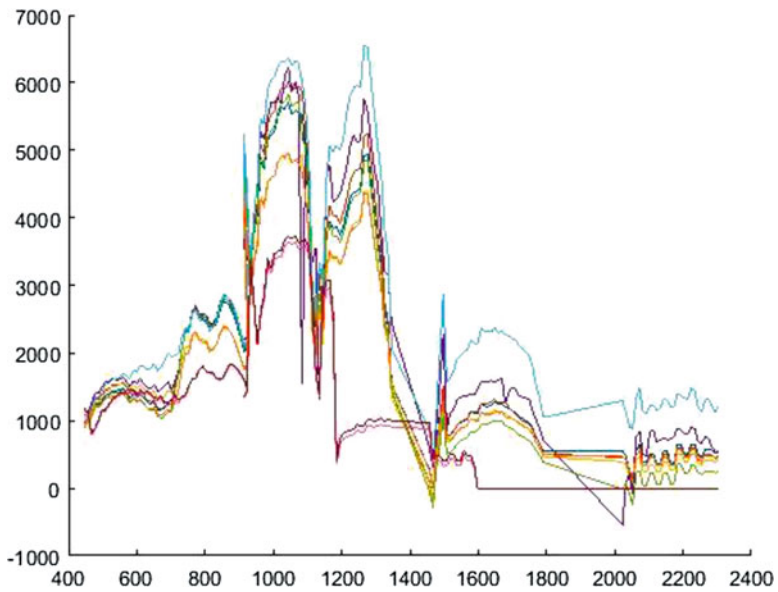
Fig. 9.7 Spectral profile of dominant mangrove species of hyperspectral image of 2014

Table 9.1 Mangrove species associated with saline blank areas (2011)

Species name	Latitude	Longitude	Abundance (2011)	Ground truth value (2011)	RMSE
Excoecaria agallocha	21° 34′ 24.81″	88° 17′ 36.89″	0	0	0.0606
Ceriops decandra			0.0299	0.05	
Phoenix paludosa			0.0987	0.10	
Avicennia alba			0.1267	0.15	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0.7447	0.75	
Excoecaria agallocha	21° 34′ 21.86″	88° 17′ 38.95″	0	0	0.0826
Ceriops decandra			0.0704	0.05	
Phoenix paludosa			0.0147	0.05	
Avicennia alba			0.021	0.05	
Avicennia marina			0.0239	0.05	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.0284	0.05	
Saline blank			0.8415	0.85	
Excoecaria agallocha	21° 34′ 24.70″	88° 17′ 50.45″	0	0	0.0598
Ceriops decandra			0.0425	0.05	
Phoenix paludosa			0.0353	0.05	
Avicennia alba			0.0532	0.05	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.0747	0.10	
Saline blank			0.7942	0.80	
Excoecaria agallocha	21° 34′ 25.69″	88° 17′ 49.41″	0	0	0.0494
Ceriops decandra			0.6437	0.65	
Phoenix paludosa			0.0642	0.05	
Avicennia alba			0	0	
Avicennia marina			0.1851	0.20	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.1032	0.10	
Saline blank			0.0037	0	
Excoecaria agallocha	21° 34′ 23.71″	88° 17′ 52.52″	0	0	0.0700
Ceriops decandra			0.0623	0.05	
Phoenix paludosa			0.0127	0.05	
Avicennia alba			0	0	
Avicennia marina			0.0491	0.05	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.0805	0.10	
Saline blank			0.7954	0.80	

(continued)

Table 9.1 (continued)

Species name	Latitude	Longitude	Abundance (2011)	Ground truth value (2011)	RMSE
Excoecaria agallocha	21° 34′ 21.77″	88° 17′ 51.46″	0	0	0.0754
Ceriops decandra			0.17	0.20	
Phoenix paludosa			0.0627	0.05	
Avicennia alba			0.0196	0.05	
Avicennia marina			0.0563	0.05	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.085	0.10	
Saline blank			0.6064	0.60	
Excoecaria agallocha	21° 34′ 26.63″	88° 17′ 53.59″	0	0	0.0679
Ceriops decandra			0.0668	0.05	
Phoenix paludosa			0.0529	0.05	
Avicennia alba			0.0225	0.05	
Avicennia marina			0.0812	0.10	
Bruguiera cylindrica			0.0335	0.05	
Avicennia officinalis			0	0	
Saline blank			0.7431	0.75	
Excoecaria agallocha	21° 34′ 22.84″	88° 17′ 38.96″	0	0	0
Ceriops decandra			0	0	
Phoenix paludosa			0	0	
Avicennia alba			0	0	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0	0	
			1	1	
Excoecaria agallocha	21° 34′ 23.81″	88° 17′ 40.01″	0.1311	0.15	0.0593
Ceriops decandra			0	0	
Phoenix paludosa			0	0	
Avicennia alba			0	0	
Avicennia marina			0.0382	0.05	
Bruguiera cylindrica			0.0865	0.10	
Avicennia officinalis			0	0	
Saline blank			0.7442	0.75	

In Tables 9.1 and 9.2, the abundance values in location 21° 34′ 24.81″N and 88° 17′ 36.89″E show an almost pure saline blank patch showing 74.47% occurrence along with presence of *Phoenix paludosa, Avicennia alba,* and *Ceriops decandra* with 9.87%, 12.67%, and 2.99% presence in 2011. The identical location shows an increase in occurrence of saline blanks and *Ceriops decandra* but a reduction in occurrence of *Phoenix paludosa* and *Avicennia alba* in the year 2014. Similarly, in another geographic location 21° 34′ 24.70″N and 88° 17′ 50.45″E of the

Table 9.2 Mangrove species associated with saline blank areas (2014)

Species name	Latitude	Longitude	Abundance (2014)	Ground truth value (2014)	RMSE
Excoecaria agallocha	21° 34′ 24.81″	88° 17′ 36.89″	0.1061	0.10	0.0668
Ceriops decandra			0	0	
Phoenix paludosa			0.1084	0.10	
Avicennia alba			0.0163	0.05	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0.7692	0.75	
Excoecaria agallocha	21° 34′ 21.86″	88° 17′ 38.95″	0	0	0.0614
Ceriops decandra			0.042	0.05	
Phoenix paludosa			0	0	
Avicennia alba			0	0	
Avicennia marina			0.0849	0.10	
Bruguiera cylindrica			0.0366	0.05	
Avicennia officinalis			0.1256	0.15	
Saline blank			0.711	0.70	
Excoecaria agallocha	21° 34′ 24.70″	88° 17′ 50.45″	0.0946	0.10	0.0717
Ceriops decandra			0	0	
Phoenix paludosa			0.0656	0.05	
Avicennia alba			0.0143	0.05	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0.8255	0.85	
Excoecaria agallocha	21° 34′ 25.69″	88° 17′ 49.41″	0.2028	0.20	0.0677
Ceriops decandra			0.4749	0.45	
Phoenix paludosa			0.0428	0.05	
Avicennia alba			0.0243	0.05	
Avicennia marina			0.1896	0.20	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0.0657	0.05	
Excoecaria agallocha	21° 34′ 23.71″	88° 17′ 52.52″	0.0974	0.10	0.0337
Ceriops decandra			0	0	
Phoenix paludosa			0.0011	0	
Avicennia alba			0.0077	0	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			0.8939	0.90	

Table 9.2 (continued)

Species name	Latitude	Longitude	Abundance (2014)	Ground truth value (2014)	RMSE
Excoecaria	21° 34′ 21.77″	88° 17′ 51.46″	0	0	0.0564
agallocha			0.0746	0.05	
Ceriops decandra			0	0	
Phoenix paludosa			0	0	
Avicennia alba			0.0994	0.10	
Avicennia marina			0.0412	0.05	
Bruguiera cylindrica			0.146	0.15	
Avicennia officinalis			0.6388	0.65	
Saline blank					
Excoecaria	21° 34′ 26.63″	88° 17′ 53.59″	0	0	0.0545
agallocha			0.0394	0.05	
Ceriops decandra				0	
Phoenix paludosa			0	0	
Avicennia alba			0.1108	0.10	
Avicennia marina			0.047	0.05	
Bruguiera cylindrica			0.1361	0.15	
Avicennia officinalis			0.6668	0.65	
Saline blank					
Excoecaria	21° 34′ 22.84″	88° 17′ 38.96″	0	0	0
agallocha			0	0	
Ceriops decandra			0	0	
Phoenix paludosa			0	0	
Avicennia alba			0	0	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0	0	
Saline blank			1	1	
Excoecaria	21° 34′ 23.81″	88° 17′ 40.01″	0.0081	0	0.0622
agallocha			0.0405	0.05	
Ceriops decandra			0.1382	0.15	
Phoenix paludosa			0.0207	0.05	
Avicennia alba			0	0	
Avicennia marina			0	0	
Bruguiera cylindrica			0	0	
Avicennia officinalis			0.7924	0.80	
Saline blank					

study area, a pure patch of saline blank shows 79.43% abundance in 2011 and 82.55% abundance in 2014. The location shows an increase in *Ceriops decandra* but decrease in *Phoenix paludosa* with abundance estimates of 4.25% and 3.53%, respectively. Geographic locations 21° 34′ 23.71″N and 88° 17′ 52.52″E and 21° 34′ 22.74″N and 88° 17′ 51.47″E of the study area also show a similar trend. Coordinates 21° 34′ 22.84″N and 88° 17′ 38.96″E show an extremely pure pixel of saline blank with no presence of any mangrove species on it.

The results signify that over the years, as saline blanks become more plentiful in Henry Island, the presence of mangrove species that are salt tolerant increases but the abundance of salt-intolerant ones decrease. *Ceriops decandra* are salt-tolerant mangrove species that show an increase in abundance with increase in saline blanks. *Phoenix paludosa* which is salt intolerant shows a decrease in abundance with increase in saline blank areas. It is also observed from field visits that in areas where the mangrove species are present in abundance and saline blanks are in its primitive state, there is a likelihood of regeneration of salt-tolerant mangrove species. Figure 9.8a–i displays the fractional abundance images of individual mangrove and saline blank endmembers generated by least square-based linear spectral unmixing algorithm.

The integrated fractional abundance image of mangrove species endmembers and saline blanks has been displayed in Fig. 9.9. The area covered by each leading mangrove species in the entire study area has been calculated and the change in area estimated. Table 9.3 displays the total pixel area covered by each species in 2011 and 2014. The number of pixels occupying a particular class has been estimated from the classified image and the area calculated by multiplying the spatial resolution (30 m × 30 m) of the image with the number of pixels.

It is observed from Table 9.3 that the total area covered by dominant mangrove species in and around the saline blank region has shown certain variations over the years. The area calculated shows that with increase in area of saline blanks from 2011 to 2014, there have been a significant increase in salt-tolerant mangrove species such as *Ceriops decandra*. The presence of salt-intolerant mangrove species such as *Phoenix paludosa* and *Bruguiera cylindrica* has decreased during the timeline.

9.8 Conclusion

This research has recognized endmember signatures of pure mangrove endmembers and saline blanks present in Henry Island, Sunderbans, through implementation of unsupervised endmember extraction model, NFINDR. The exclusive signature spectra of saline soil endmembers and pure mangroves have been used to determine their fractional abundance values if present in a mixed pixel using least square error-based FCLSU. *Excoecaria agallocha* and *Ceriops decandra* are the mangrove species prevalent around saline blank areas. *Avicennia marina* and *Avicennia alba* are also observed to exist in abundance in and around highly saline soils. It may be concluded that, with time, as saline blanks increase in abundance in the island, the occurrence of mangrove species that are salt tolerant increases but the existence of salt-intolerant ones reduce. There has been a significant increase in salt-tolerant mangrove species in the study area such as *Ceriops decandra, Excoecaria agallocha, Avicennia marina, and Avicennia alba,* whereas the presence of salt-intolerant mangrove species such as *Phoenix paludosa* and *Bruguiera cylindrica* has reduced during the timeline.
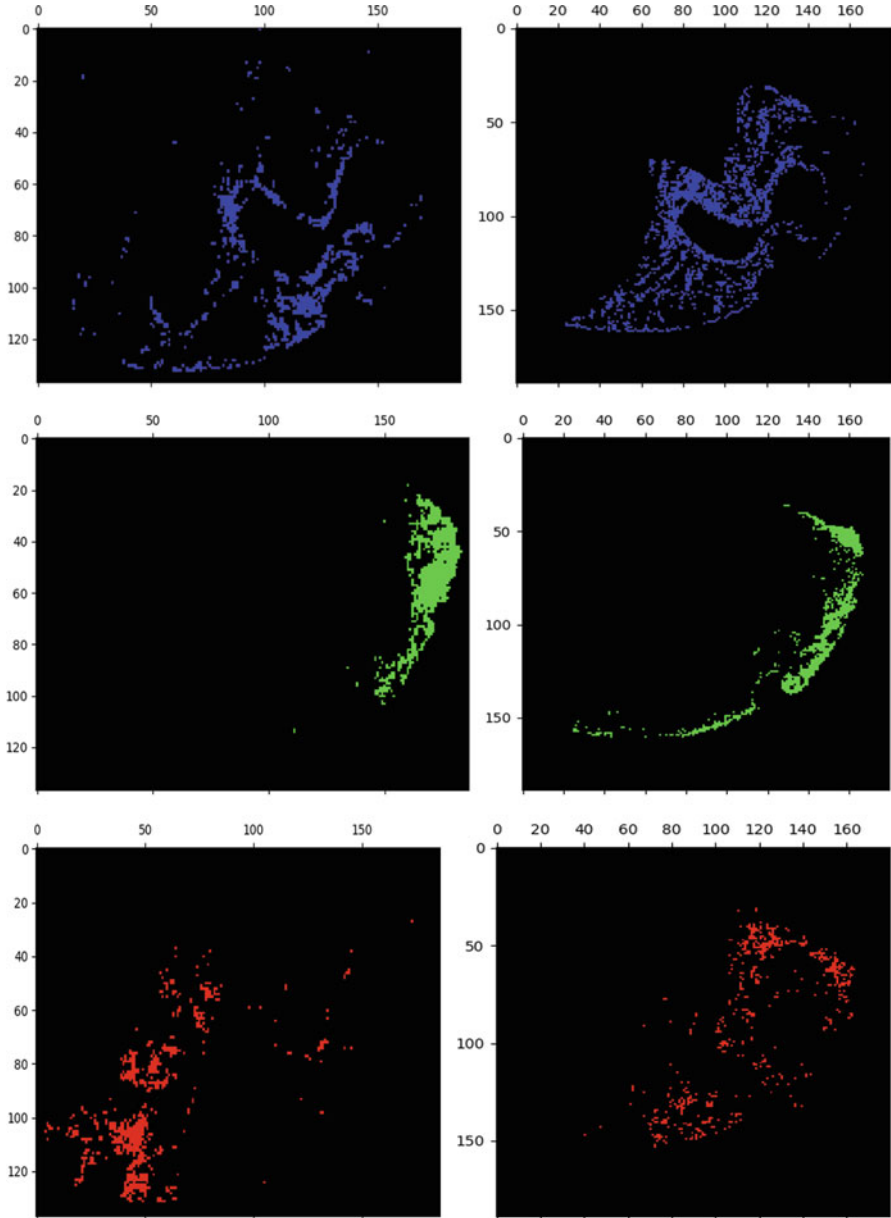
Fig. 9.8 Mangrove species distribution map obtained after least square error-based linear unmixing for the data sets in 2011 (left) and 2014 (right)
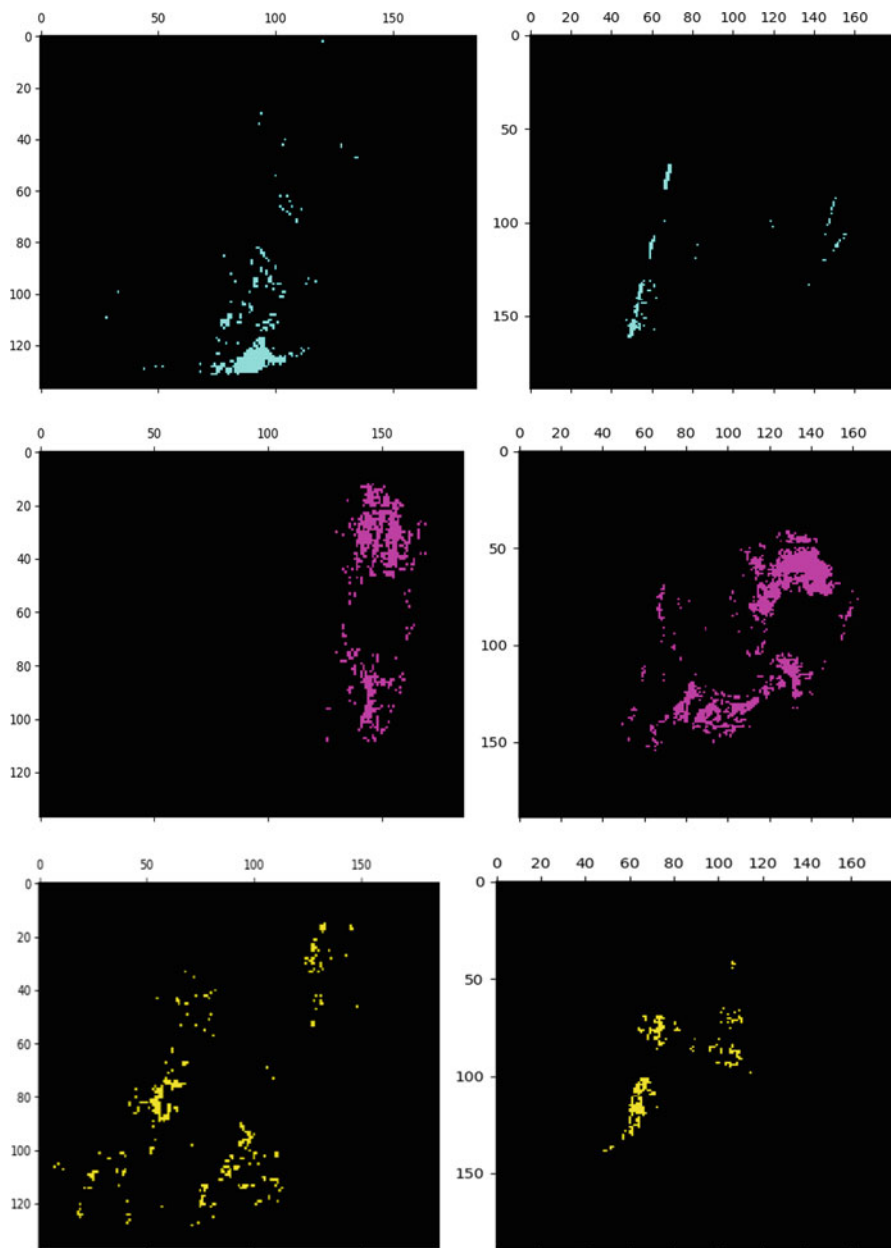
Fig. 9.8 (continued)
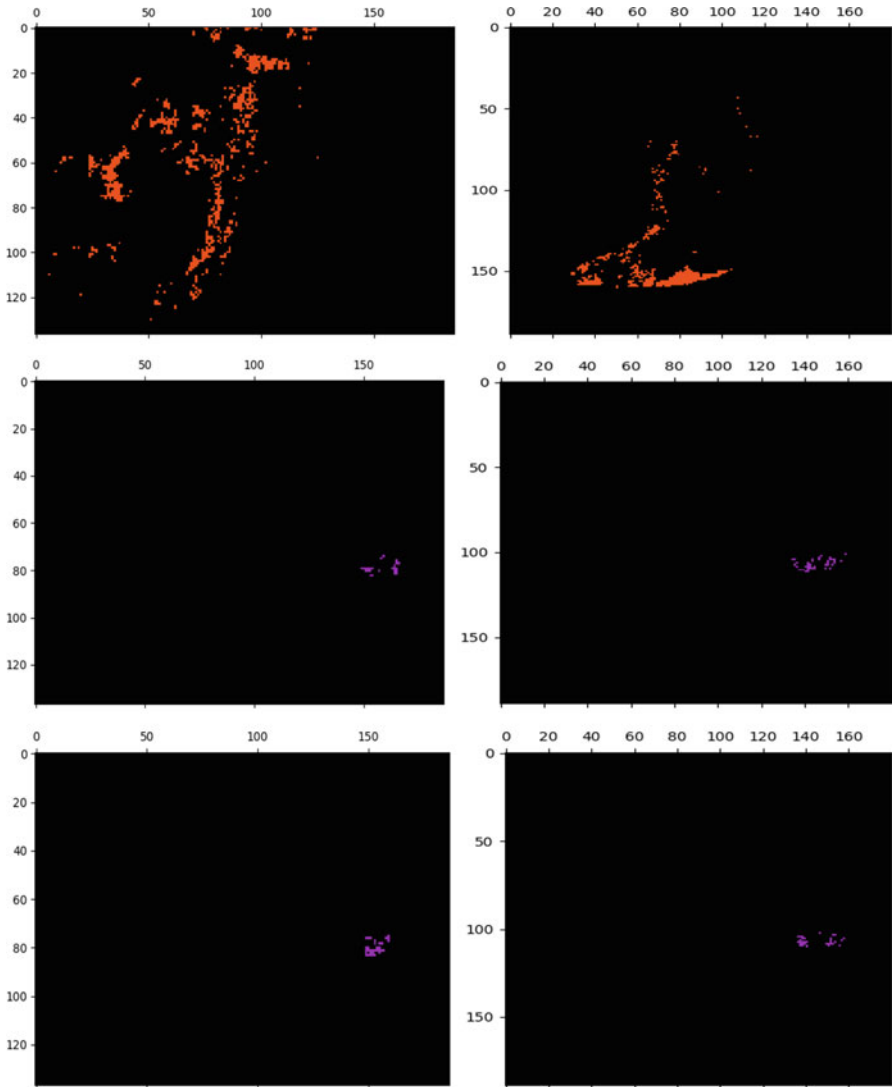
Fig. 9.8 (continued)
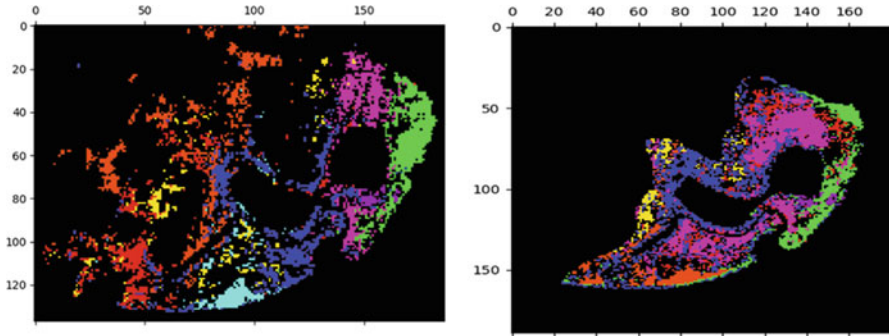
Fig. 9.9 Integrated fractional abundance images of 2011 and 2014

Table 9.3 Area covered by dominant mangrove species of Henry Island, Sunderbans

Mangrove species	Area in year (2011)	Area in year (2014)
Ceriops decandra	709*900 = 6,38,100	880*900 = 7,92,000
Phoenix paludosa	671*900 = 6,03,900	564*900 = 5,07,600
Bruguiera cylindrica	349*900 = 3,14,100	258*900 = 2,32,200
Avicennia officinalis	855*900 = 7,69,500	473*900 = 4,25,700
Saline blank	60*900 = 54,000	77*900 = 69,300

This study will enable creation of a proper land-use map for proper utilization of fallow saline wastelands and can be taken up at the right earnest. Adoption of proper preventive, restoration and utilization measures to combat saline banks in remotely located Sunderban islands can also be proposed on case-specific basis which would lead to sustainable management of this pristine mangrove ecosystem.

References

1. Mohanty S (2015) Mangrove denundation scar on sanctuary. Indian Express News, 24th May, 2015. http://www.newindianexpress.com/states/odisha/Mangrove-Denudation-Scar-on-Sanctuary/2015/05/24/article2830205.ece
2. Chowdhury A, Maiti SKA (2014) Comparative study of variations in mangrove biodiversity at central and eastern parts of the Sundarban Biosphere Reserve, India. Am Int J Res Formal Appl Nat Sci 5(1):27–31
3. GhoshThakur D (2009) Saline poison in Sunderban ecosystem. Anandabazar News, 4th December
4. Giri S, Mukhopadhyay A, Hazra S, Mukherjee S, Roy D, Ghosh S, Ghosh T, Mitra D (2014) A study on abundance and distribution of mangrove species in Indian Sundarban using remote sensing technique. J Coast Conserv 18(4):359–367
5. Nayak S, Bahuguna A (2001) Application of remote sensing data to monitor mangroves and other coastal vegetation of India. Int J Remote Sens 30(4):195–213
6. Samanta K, Hazra S (2012) Landuse/landcover change study of Jharkhali Island Sundarbans, West Bengal using remote sensing and GIS. Int J Geom Geosci 3(2):299

7. Manson FJ, Loneragan NR, Phinn SR (2003) Spatial and temporal variation in distribution of mangroves in Moreton Bay, subtropical Australia: a comparison of pattern metrics and change detection analyses based on aerial photographs. Estuar Coast Shelf Sci 57(4):653–666

8. Giri C, Pengra B, Zhu Z, Singh A, Tieszen LL (2007) Monitoring mangrove forest dynamics of the Sundarbans in Bangladesh and India using multi-temporal satellite data from 1973 to 2000. Estuar Coast Shelf Sci 73(1–2):91–100

9. Muttitanon W, Tripathi NK (2005) Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. Int J Remote Sens 26(11):2311–2323

10. Eslami-Andargoli L, Dale PER, Sipe N, Chaseling J (2009) Mangrove expansion and rainfall patterns in Moreton Bay, Southeast Queensland, Australia. Estuar Coast Shelf Sci 85(2):292–298

11. Lucas RM, Mitchell AL, Rosenqvist AKE, Proisy C, Melius A, Ticehurst C (2007) The potential of Lband SAR for quantifying mangrove characteristics and change: case studies from the tropics. Aquat Conserv Mar Freshwat Ecosyst 17(3):245–264

12. Conchedda G, Durieux L, Mayaux P (2008) An object-based method for mapping and change analysis in mangrove ecosystems. ISPRS J Photogramm Remote Sens 63(5):578–589

13. Haboudane D, Miller JR, Pattey E, Zarco-Tejada PJ, Strachan IB (2004) Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Remote Sens Environ 90(3):337–352

14. Winter ME (1999) N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. In: SPIE's international symposium on optical science, engineering, and instrumentation, pp 266–275

15. Du Q, Raksuntorn N, Younan NH, King RL (2008) Variants of N-FINDR algorithm for endmember extraction. SPIE Remote Sens 7109:71090G1-8

Chapter 10
Crop Classification and Mapping for Agricultural Land from Satellite Images

A. Kalaivani and Rashmita Khilar

Abstract Agriculture is the backbone of Indian production which is the vital sector for food production. It is very important for national government to know what type of crops are being grown in which region for budget planning to import and export food products. Traditional ground survey method is laborious, time-consuming, and expensive. Along with this, continuous monitoring of crops is highly difficult. Crop area estimation is a key element in crop production forecasting and estimation. Crop classification and mapping are the most challenging tasks among the land use/land cover classification problems.

In agriculture domain, the common approach used by the government (farmers) for crop monitoring is to go to the field and acquire the images using cameras for estimation of the crop yield. So in this context, a fast, reliable, and automated system is required which provides the exact crop mapping using satellite images. In recent years, crop identification and area monitoring from satellite images are given more and more attention.

The stages are image acquisition, image preprocessing, feature extraction, and image classification. Satellite images are preprocessed and features are extracted from input images. Based on the features extracted, images are classified based on the extracted features. The proposed automated system should provide better accuracy than the existing in the literature.

Keywords Crop mapping · Crop classification · Satellite images · Feature subset selection · Supervised classifier

A. Kalaivani
Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Kuthambakkam, Tamil Nadu, India
e-mail: kalaivania.sse@saveetha.com

R. Khilar (✉)
Department of IT, Panimalar Engineering College, Chennai, Tamil Nadu, India

10.1 Introduction

Agriculture is a boom of Indian economy, which accounted for 10% of the total export earnings and serves as an important raw material for many food production industries. Agriculture is the major source of food production and serves as a backbone for reduction of poverty. National growth across global market depends on the agricultural growth which depends on agricultural production. The agriculture production can be improved by introduction of new varieties of technology in cultivation and production of crops. Changing of crop pattern can better improve agricultural production and makes a position rise in global market.

Remote sensing refers to the detection, identification, and classification of the objects by observing and recording their signals which are collected through aerial- or satellite-based sensors. The RS technology started with an oblique photograph of a village near Paris from a balloon. It expanded from black and white photographs to color photographs based on sensors. The sensors used to record the signal are divided into two classes of active and passive remote sensing. Active remote sensors use a transmitter and receiver to record the reflected or scattered electromagnetic radiation. The remote sensor systems which measure energy emitted or scattered from the objects are called as passive sensors. A schematic diagram showing a solar radiation-based passive remote sensors system is shown in Fig. 10.1.

In India, remote sensors are used in crop area estimation, and a number of studies are also explored for crop inventory, acreage estimation, and yield prediction. Later, the Ministry of Agriculture and Department of Space collaborated for large-scale projects for demonstrating the applications of optical and microwave RS data for crop studies. The operational project FASAL supports India in preharvest estimation of crops and predicts multiple crop production on different scales. CAPE was the first large-scale project in India that used remotely sensed data for crop studies to identify crops and estimate production. The satellites launched by India from the source ISRO 2016 are listed in Table 10.1.

India is one of the major food grain-producing countries in the world similar to China, USA, Russia, Canada, etc. To withstand in global food grain market, India should be aware of the demand and production of other countries. The agricultural statistics at the state level are done by State Agricultural Statistics Authorities (SASAs). At the all-India level, the agricultural statistics are done by the directorate
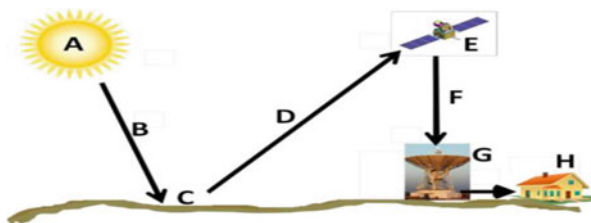
Fig. 10.1 Passive remote sensor system

Table 10.1 Satellites launched by India

Satellite	Launch date	Launch vehicle	Major payloads
CARTOSAT – 2C	Jun 22, 2016	PSLV-C34	PAN
SARAL	Feb 25, 2013	PSLV-C20	AltiKa, ARGOS, SCBT
RISAT-1	Apr 26, 2012	PSLV-C19	SAR (C-band)
Megha-Tropiques	Oct 12, 2011	PSLV-C18	MADRAS, SAPHIR, ScaRaB, ROSA
RESOURCESAT-2	Apr 20, 2011	PSLV-C16	LISS-3, LISS-4, AWiFS, AIS
CARTOSAT-2B	Jul 12, 2010	PSLV-C15	PAN
Oceansat-2	Sep 23, 2009	PSLV-C14	OCM, SCAT, ROSA
RISAT-2	Apr 20, 2009	PSLV-C12	SAR (X-band)
IMS-1	Apr 28, 2008	PSLV-C9	Mx, HySI
CARTOSAT-2A	Apr 28, 2008	PSLV-C9	PAN
CARTOSAT-2	Jan 10, 2007	PSLV-C7	PAN
CARTOSAT-1	May 05, 2005	PSLV-C6	PAN (Fore), PAN (Aft)
RESOURCESAT-1	Oct 17, 2003	PSLV-C5	LISS-3, LISS-4, AWiFS
TES	Oct 22, 2001	PSLV-C3	PAN
Oceansat (IRS-P4)	May 26,1999	PSLV-C2	OCM, MSMR
IRS-1D	Sep 29,1997	PSLV-C1	PAN, LISS-3, WiFS
IRS-P3	Mar 21,1996	PSLV-D3	WiFS, MOS, IXAE, CBT
1RS-1C	Dec 28,1995	Molniya	PAN, LISS-3, WiFS
IRS-P2	Oct 15,1994	PSLV-D2	LISS-2A, LISS-2B
IRS-1E	Sep 20,1993	PSLV-D1	LISS-1, MEOSS
IRS-1B	Aug 29,1991	Vostok	LISS-1, LISS-2A, LISS-2B
SROSS-2	Jul 13,1988	ASLV	GRB, MEOSS
IRS-1A	Mar 17,1988	Vostok	LISS-1, LISS-2A, LISS-2B
RS-D2	Apr 17,1983	SLV-3	Smart sensor, L-band beacon
Bhaskara-II	Nov 20,1981	C-l Intercosmos	TV cameras, SAMIR

Source: ISRO, 2016

of economics and statistics. Although this system is well recognized, the major shortcomings are subjectivity in the crop acreage estimation and delays in crop forecasts.

Moorthi et al. [15] developed FASALSoft, an ISRO software framework for crop production forecast using primarily remote sensor data analysis. In India, the FASALSoft is accepted as operational for making multiple in-season crop production forecasts and implemented by MNCFC, New Delhi. During the year 1880–2010, economic growth increased due to the significant land use and land cover (LULC). The interactions of human activities, climate systems, and ecosystem in the country are not specified by accurate database of LULC. Tiana et al. [12] used high-resolution RS datasets from Resourcesat-1 and historical archives at district and state levels to generate LULC datasets during 1880–2010 in India. The experimental results identified that there is a significant loss of forests, and cropland area increased significantly during 1880–2010. Due to this farm mechanization,

electrification and introduction of high-yielding crop varieties are introduced in government policies to achieve self-sufficiency in food production.

A methodology for monitoring progress of Rabi crop area at country scale was developed by Nigam et al. [16]. He used temporal vegetation index derived or captured from Indian geostationary satellite (INSAT 3A). The 10-day maximum NDVI composite products were generated and used over six crop dominant states of India. The estimates showed a deviation of 18.1% to 14.6% deviations, and the inter-seasonal variability in the estimate was consistent with the reported statistics for Rabi crop. The authors recommend NDVI product with finer spatial resolution satellite data for crop area monitoring for their country. Singla et al. [21] have justified the role of geo informatics to be used for discriminating different crops at various levels of classification, monitoring crop growth, and prediction of the crop yield. The author justified that in addition to RS technology usage of ground observations, reviews, GIS, and soil analysis.

10.2   Land Utilization and Crop Pattern in Tamil Nadu

In Tamil Nadu, agricultural crops are broadly classified into food crops and non-food crops. Human-consumed crops are food crops which are further classified as cereals and pulses. The cereals crops produced all over our state includes rice, wheat, jowar, bajra, maize, ragi, and other millets, and pluses include gram, arhar, moong, peas, and masoor. The non-food crops are divided into commercial crops which are cultivated for sale. The other category of non-food crops is a plantation crop which covers a large estate. The final category of horticultural crops are crops with rich source of carbohydrates, proteins, organic acids, vitamins, and minerals for human nutrition.

Land utilization depends on environmental factors such as soil characteristics, climate, topography, and vegetation. Land can be properly utilized by humans for agriculture, industry, forestry, energy production, settlement, recreation, water catchments, and storage. For better utilization of the land, identify the natural characteristics, extent and location, and its quality, productivity, suitability, and limitations. The high utilization of land can be achieved through application of bio fertilizers, hybrid seeds, double cropping, and modern methods (Fig. 10.2).

In Tamil Nadu, the population is very high and land resources are not utilized due to the mal-utilization of land. So, if proper planning of land utilization yields to higher agricultural production, our states can concentrate on the major crops such as rice, paddy, cholam, cumbu, ragi, sugarcane, cotton, groundnut, and million hectares.

In Tamil Nadu, the districts are classified into 13 districts which are (1) Chengalpattu (Chennai, Kancheepuram, Thiruvallur), (2) South Arcot (Cuddalore, Villupuram), (3) North Arcot (Thiruvannamallai, Vellore), (4) Salem (Namakkal), (5) Dharmapuri (Krishnagiri), (6) Coimbatore, (7) Erode, (8) Tiruchirapalli (Karur, Perambalur), (9) Pudukkottai, (10) Thanjavur (Thiruvarur, Nagapattinam), (11)
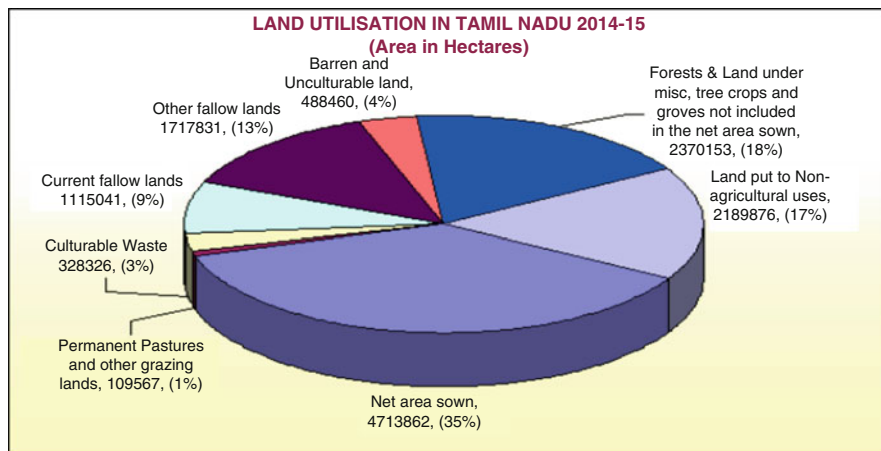
Fig. 10.2 Land utilization in Tamil Nadu

Madurai (Dindigul, Theni), (12) Ramanathapuram (Kamarajar, Pasumpon Muthu-ramalingam, Sivagangai), and (13) Tirunelveli (Thoothukudi).

The data for cropping pattern of each district are extracted from Season and Crop Reports of Tamil Nadu during the period from 1981–1982 to 2005–2006 specified crops average yield of 25 years of data for thirteen districts and state as a whole which are shown in Tables 10.2 and 10.3.

10.3   Literature Survey

Smith et al. [22] and Rachel Barrett et al. [1] reported how the current technology used to recognize and differentiate between horticultural crops in Tasmania, Australia. A methodology for the systematic recognition of individual crops was developed with a better utilization of the technology in a commercial environment. The focus is tailored individually to five commercially significant horticultural crops such as poppies, pyrethrum, potatoes, peas, and onions. These crops are selected due to the predictive accuracy higher in comparison to 17 different crops investigated. The logical conclusion from both the perspective of agronomic and computational level is that the probability of achieving a correct outcome is enhanced by incorporating horticultural and environmental conditions used in both the training and testing data. The predictive accuracy can be improved by specifically tailoring interpretative images to cropping sequences.

Computer-assisted photo interpretation (CAPI) tool funded by National Control and Paying Agencies (NCPA) checks the proper spending of funds by identifying crops via remotely sensed imagery. The major issue of the proposed method is time-consuming and requires skilled interpreters. The objective of May et al. [14]

Table 10.2 District-wise crop data in Tamil Nadu – sugarcane, cotton, groundnut, tapioca

Districts (1)	Sugarcane		Cotton		Groundnut		Tapioca	
	Extent of cultivation (percent) (10)	Average yield (kg ha) (11)	Extent of cultivation (percent) (12)	Average yield (kg ha) (13)	Extent of cultivation (percent) (14)	Average yield (kg ha) (15)	Extent of cultivation (percent) (16)	Average yield (kg ha) (17)
Chengalpattu	2.93	108	0.04	351	16.58	2190	0.05	32156
South Arcot	0.96	113	1.37	362	17.66	1601	2.03	32608
North Arcot	7.09	82	1.01	361	33.10	1333	0.08	32222
Salem	3.43	125	3.37	321	21.27	1417	6.01	37199
Dharmapuri	3.23	86	2.59	336	12.59	1392	2.24	32726
Coimbatore	3.66	105	4.41	383	10.87	1327	0.25	40915
Erode	6.50	119	2.29	385	17.09	1552	0.76	41670
Trichy	3.70	109	3.11	253	11.40	1398	1.01	38985
Pudukkottai	1.70	108	0.31	339	21.60	1232	0.01	31012
Thanjavur*	2.29	101	0.63	397	3.23	1639	0.05	34124
Madurai	3.18	111	5.70	322	9.25	1524	0.05	35721
Ramanathapuram	1.07	102	7.71	209	5.41	1029	0.01	27575
Tirunelveli	0.77	104	10.71	216	2.61	1546	0.04	31286
Tamil Nadu	3.77	105	3.22	269	13.38	1455	1.21	33627

Table 10.3 District-wise crop data in Tamil Nadu – paddy, cholam, cumbu, ragi

Crops in Tamil Nadu

Districts (1)	Paddy		Cholam		Cumbu		Ragi	
	Extent of cultivation (per cent) (2)	Average yield (kg ha) (3)	Extent of cultivation (percent) (4)	Average yield (kg ha) (5)	Extent of cultivation (percent) (6)	Average yield (kg ha) (7)	Extent of cultivation (percent) (8)	Average yield (kg ha) (9)
Chengalpattu	66.18	2896	0.09	1259	0.50	1882	1.37	1799
South Arcot	35.61	3052	1.75	985	10.00	1323	0.89	1961
North Arcot	27.09	2865	4.54	1143	2.51	1079	2.43	1976
Salem	10.01	2900	10.38	1025	3.43	1375	3.49	2131
Dharmapuri	9.83	2935	7.91	1221	1.18	1510	15.5	1610
Coimbatore	5.96	3382	27.35	508	0.53	1565	0.28	1998
Erode	16.93	3859	10.03	500	2.38	1022	4.03	1195
Trichy	22.62	3068	20.33	777	7.54	635	0.21	1827
Pudukkottai	51.92	2515	0.99	1073	0.39	1423	1.10	1473
Thanjavur*	67.36	2559	-	-	0.04	1597	0.03	2370
Madurai	21.36	3549	16.17	1254	2.69	1341	0.34	1970
Ramanathapuram	51.04	1940	2.21	945	2.71	1130	1.79	1728
Tirunelveli	26.96	3671	3.80	1510	7.26	1340	0.64	2416
Tamil Nadu	31.94	2855	7.72	910	3.31	1113	2.27	1750

developed a reliable control system to replace CAPI for crop identification. The proposed method reduced control costs and completion time. The authors proposed an automatic control system (ACS) which classifies crops based on a reliability requirement. The main concern of this method require proper decision-maker to set a high reliability level which restricts automatic crop identification which classify crops at high certainty.

The authors Jonas Schmedtmann and Manuel L. Campagnolo [20] quantified the accuracy of the proposed approach and analyze the trade-off between the reliability level and the proportion of parcels that can be automatically controlled. The automatic classification of agricultural land parcels is reliable, and it provides a good performance even when crops are difficult to discriminate.

Foster and Kesselman [9] used spectral mixture modeling to produce the land cover maps for Tumkur District, Karnataka. Spectral Mixture Analysis (SMA) was performed and evaluated on Landsat-8 ETM (Enhanced Thematic Mapper Plus) data, and the results are compared with the ground truth data. The specified techniques give more accurate results by taking the absolute difference between actual and modeled estimations with homogenous coconut land cover. SMA is easy and low computational method used successfully to classify different vegetation covers in intensive agricultural areas. The methodology is used in Czajkowski et al. [4] satellite data selection and suitable method for classification and checking the accuracy. The most important challenges by Beeresh et al. [2] are multiple crops identification and differentiation of crops of same. Good soft computing and analysis skills are required to classify and identify the class from multispectral and hyper-spectral images. Researchers focused on supervised and unsupervised classification methods along with hard classifiers and also with soft computing techniques like fuzzy C mean and support vector machine.

Remote sensing images are good source for decision-making related to crops monitoring and mapping in optical region, Rajesh K Dhumal et al. [7] uses a multispectral images which gives much detail for overall vegetation mapping in large area. Whereas it is having limitation due to broad wavelength and spatial resolution this paper lacks in differentiating crops of similar type, this problem overcomes by hyper-spectral images. Selection of spectral bands in hyper-spectral images is also a quite challenging task. Some limitation of optical remote sensing can be overcome by fusing optical remote sensing images with microwave remote sensing images.

Foster et al. [11] carried out a study on pixel-based cropland classification by fusion of data from satellite images with different resolutions. Methodology is based on various multispectral images acquired at different resolutions by different imaging instruments, Landsat-8 and Rapid Eye. The proposed method overcomes the shortcomings of different instruments in the particular cropland classification scenario situated on the plains of Vojvodina in northern Serbia. Researchers proposed a data fusion method with a robust random forest classifier. It improves the overall classification performance with a coarser spatial resolution in the given specific cropland classification task. The method developed by Predrag Lugonja

et al. [10] provides an improvement over the existing pixel-based classification approaches through the combination of different data sources. Another contribution of this paper is the employment of crowdsourcing by Bahram Salehi et al. [19] in the process of reference data collection via dedicated smartphone application. NIR band contributed a significant influence in classifiers that use the proposed data fusion method.

Nataliia Kussul et al. [13] used unsupervised neural network (NN) for optical imagery segmentation which helps in restoring missing data due to clouds and shadows and an ensemble of supervised NNs. A fully connected multilayer perceptron (MLP) and RS community random forest approaches are used and compared with convolution NNs (CNNs). The proposed method 2-D CNN provides better performance than 1-D CNNs, and small objects in the final classification were smoothed and misclassified which leads to further modifications. The methodology proposed by Daniel M. Howard et al. [5] is to classify major crops in the Greater Platte River Basin (GPRB) in addition to existing crop classification products. The input for the system includes Moderate Resolution Imaging Spectroradiometer-normalized differential vegetation index, average long-term growing season temperature, average long-term growing season precipitation, and yearly start of growing season. The overall accuracy of 78% is achieved for a test sample of roughly 215,000 independent points that were withheld from model training. Ten 250 m resolution annual crop classification maps were produced and evaluated for the GPRB region, one for each year from 2000 to 2009. The proposed method validates and focuses on spatial distribution and county-level crop area and later compared with NASS CDL and county statistics. Crops are classified based on the spatial distribution and exhibited a close linear agreement from USDA datasets.

The methodology developed by Yetkin Özüm Durgun et al. [8] is based on phenological characteristics of different crop types applied using 100-m Proba-V NDVI data for the season 2014–2015. The postclassification rules were applied to aggregate the crop type at the plot level. The methodology is applicable to 100-m Proba-V and used in crop area mapping across the world. The accuracy ranges from 65% to 86%, and the kappa coefficient varied between 0.43 and 0.84 depending on the site and the temporal window used. The main objectives of researchers Bahram Salehi et al. [19] is to evaluate how the number of multi-temporal images used in an OBIA framework affects the classification results for both RE and RS-2 data. The above technique combines the object- and pixel-based image classification method for classification of accurate crop type. The above methodology can be improved, and accurate classification at discrete locations (approximately 25×25 km frames) can be applied in a separate procedure to increase the accuracy of crop area estimation at the regional to provincial scale by linking these local, very accurate spatially discrete results to national wall-to-wall continuous crop classification maps.

A separate quantitative accuracy assessment for segmentation needs to be conducted. In spite of the considerable progress in segmentation algorithms, there is still lack of a global segmentation quality measure [6], and visual assessment is still

widely used and required. Moreover, since only spectral information of segments (objects) was utilized in the proposed framework and not shape (or geometry), the geometry of the segment does not have much impact on the classification results. In other words, as long as there is oversegmentation for crop fields (i.e., segments are smaller than crop fields), the segmentation results are valid. Therefore, it is important to, visually, ensure that there is no undersegmentation (one segment covering more than one crop field) in image segmentation. In multi-resolution segmentation, this can be controlled by selecting a scale parameter that results in segments that are smaller than actual crop fields. With several SAR polarimetric indicators and features extracted from the decomposition and scattering matrix in the context of OBIA, it is expected to achieve improved classification results compared to what was achieved solely based on the four intensity channels. This is recommended to be explored in a future research.

Cankut Ormeci et al. [17] are the first who use two pixel-based classification algorithm unsupervised classification and maximum likelihood supervised classification on the ISODATA and object-based classification algorithm. The accuracies of classifications achieved by the confusion are overall accuracy, user's accuracy, producer's accuracy, and kappa coefficient. An error matrix is obtained as a simple cross tabulation of mapped class label and data for a sample of cases at specified locations. The methodology searches the efficiency of satellite images in crop type and area is determined. The performance of crop areas identification was acceptable both by pixel-based classification and object-based classification techniques, but the object-based classification gives superior results compared to pixel-based classification techniques. Rose M. Rustowicz et al. [18] provides a comparison of mono- vs. multi-temporal results and justifies that temporal information is helpful for successful crop classification. The test accuracies are 92% for the multi-temporal case and 85% for the mono-temporal case. When input data features are compared, the current data points are normalized spectral irradiance values at satellite sensor. The system is adaptable for more classes and a more sporadic placement of crop types. The areas need to be extend and individual needs to get involved in order to improve the above methodology.

The approach explained by Gong Cheng et al. [3] are a comprehensive review of the recent progress in this field and proposed a large-scale, publicly available benchmark dataset by analyzing the limitations of existing datasets. The current state-of-the-art scene classification methods performed on NWPU-RESISC4 are investigated. Users performed knowledge discovery by crowdsourcing of information through these location-based social media data. The methodology maps "what-is-where" easily on the surface of the Earth using the "what" and "where" aspects of the information. A comparison is carried out that with remote sensing images, the ground photos uploaded by user hold higher resolution and are quite different from satellite remote sensing in the observation direction, which can well capture the detail and vertical characteristics of ground objects. The additional information are very useful for the classification and recognition of remote sensing images.

New methods need to be proposed which combine remote sensing data and information coming from social media and spatial technology which can be deployed to promote the state-of-the-art remote sensing image scene classification.

10.4   Remote Sensing System Data Classification

Remote sensing system classification techniques are classified into three categories, i.e., manual, automated, and hybrid classifications. Manual satellite image classification approaches are robust and effective methods. Most of the time, manual methods consume more time. In manual methods, the analyst must have high degree of familiarity with the area covered by the satellite image. Efficiency and accuracy of the classification largely depend on analyst subject knowledge and familiarity of the theme under study. Automated remote sensing classification methods use algorithms that are applied systematically to the entire image for grouping meaningful pixels into various categories. Majority of the classification methods falls under this category. Hybrid image classification methods combine the advantages of both the methods. These methods use automated satellite image classification methods to do initial classification; further manual methods are used to refine classification output by adding value to it based on the analyst's subject knowledge. Automated data classification methods are classified into supervised and unsupervised classification methods. Supervised classifiers can be parametric or nonparametric. Parametric classification is based on the observed measurement vectors obtained for each class in each spectral band during the training phase and is normally distributed Gaussian vectors. In nonparametric classification, no such assumption is made and the maximum likelihood classification is the popular and widely used supervised classifier. The most important data classification is discussed in detail.

10.4.1   *Minimum Distance from Mean (MDM) Classification*

MDM is the simple supervised classification technique which determines the spectral distance between the measurement vector for the pixel and the mean vector for each training class signature. It is used to classify unknown image pixels into classes which minimizes the distance between the image pixel and the class in multifeatured space. To compute minimum distance between the data and the class having separable class variances, Euclidean distance or normalized euclidean distance is used. It requires the least computational time among other supervised methods, and class variability is not taken into consideration by this method. The spectral feature space formed by the multisensory data needs to be normalized with respect to different radiometric ranges present in the datasets to produce the desired results.

10.4.2   Mahalanobis Distance Classification (MDC)

The Mahalanobis distance is calculated as group in m-dimensional space defined by m variables and their covariance. This method is used if there is a correlation among the axes in feature space. This method takes the variability of classes which is more useful than minimum distance which considers weighting factors. However, this method tends to overclassify signatures with relatively large values in the covariance matrix. The proposed method is slower in computation and depends on a normal distribution of the data in each input band.

10.4.3   Maximum Likelihood Classification

MLC is one of the most popular supervised classification techniques used in the analysis of RS data. The proposed methods define the posterior probability of a pixel belonging to a particular class. The MLC for each class in each band follows normal distribution and calculates the probability accordingly. The pixels are classified into the class with the highest probability, and the other pixel remains unclassified. The proposed method needs long time of computation and depends on a normal distribution of the data in each input band. Sufficient ground truth/training sites have to be selected for computing the variance-covariance matrices of population.

10.4.4   K-Means Classification

The K-means algorithm is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The K-means clustering algorithm is a partition-based cluster analysis method which is the most popular unsupervised clustering classification techniques. In clustering techniques, the pixels are grouped into different clusters and the clusters exhibit property of homogeneity among the same cluster and heterogeneity between different clusters. In K-means, K stands for number of clusters to be formed which may be defined by user or from any automatic techniques or initializes K randomly. The algorithm is iterative in nature and each pixel is assigned to an exclusive cluster. In the iteration process, it minimizes within cluster scatter and the process repeats until the scatter is less than a threshold value or reaches the maximum number of iterations. The image is divided into K clusters and the mean of each cluster is computed. In the next iteration, the pixels are assigned to the nearest classes and new class means are computed. The final partitioning of the clusters is based on the final K means obtained after the completion of iteration, hence the name K-means.

10.4.5   ISODATA

Iterative Self-Organizing Data Analysis (ISODATA) is the most popular unsupervised clustering method. It is used to partition the spectral image into number of spectral classes based on the statistical information inherent in the image. ISODATA creates predefined number of clusters in a satellite image. Cluster centers are randomly placed and pixels are assigned based on the shortest distance to center method. The standard deviation within each cluster and the distance between cluster centers are calculated. Clusters are further divided if standard deviation is greater than the user-defined threshold otherwise they are merged. Further iterations continue until either the average intercenter distance falls below the user-defined threshold, the average change in the intercenter distance between iterations is less than a threshold, or the maximum number of iterations is reached. Unsupervised classification can produce more accurate results than supervised classification which are ideal to large, complex, and heterogeneous areas with lack of intimate familiarization and field information.

10.4.6   Fuzzy Classification

In fuzzy classification, the image pixels are grouped into a fuzzy set whose membership function truth is the value of a fuzzy propositional function. A fuzzy set allows its members to have different grades of membership in the interval [0, 1]. Fuzzy classification is a soft classification which takes heterogeneous and imprecise nature (mix pixels) of the real-world scenario. Proportions of the multiple classes within a pixel (e.g., 30% bare soil, 30% forest, and 40% crop) can be obtained. In the case of a hard classification (ISODATA, K-means, MLC, etc.), each pixel belongs to the class, and in the case of soft classification, each pixel can belong to more than one class and has membership grades for each class. If needed, defuzzification can produce a crisp result (one pixel to only one class) from fuzzy membership grades. In fuzzy classification, the boundary between two neighboring classes is assumed as a continuous, overlapping area within which an object has partial membership in each class. This viewpoint reflects the reality of many applications in which categories have fuzzy boundaries. The fuzzy classifier is computationally complicated and is not recommendable, if conventional approach yields a satisfying result.

10.4.7   Artificial Neural Networks (ANN)

Artificial neural network is a parallel distributed processor made up of simple processing units called neurons. Artificial neural network or ANN is a computational

system that resembles the organizational principles present in biological nervous systems. It has a normal tendency for storing experiential knowledge like a human brain and uses this for pattern recognition. In ANN, the basic computational element is known as the neuron or node. The neuron processes data in stages. First, the messages received are aggregated by an internal activation function. Then, the information is sent to transfer functions, which determine whether the neurons will send the output message or not. Multiple neurons are connected together in layers. These layers called the input layers are set up to receive input information, process the data through one or more hidden layers, and produce a corresponding output pattern through the output layer.

ANN have a distinct advantage over statistical classification methods in achieving higher training accuracy—the capabilities of ANN for nonlinear function approximation, data classification, nonparametric regression, and nonlinear decision-making. ANN approach are ideally suitable to crucial applications such as land use/cover-related feature classification from a satellite imagery. They are nonparametric and require little or no a priori knowledge of the distribution model of input data. ANNs have high processing speed, robustness, and capability to deal with high-dimensional data spaces.

10.4.8 Support Vector Machines (SVM)

Support vector machine (SVM) is a supervised nonparametric learning technique independent of any assumptions for the underlying data distribution. It finds a hyperplane which separates the dataset into a discrete predefined number of classes. SVMs are linear binary classifiers and assume that the multispectral feature data are linearly separable in the input space. SVMs are not widely used by remote sensing community as SVM performs similar to other established method.

10.4.9 Decision Rule-Based Tree Classification

Decision tree classifier is one of the most popular supervised classification methods, which is nonparametric and does not require the data in normal distribution. Decision tree classifiers are easy to train and they learn quickly from examples. It follows a tree-structured graph or model of decision rules and their possible result. Following a tree architecture, it is composed of a root node and a series of internal nodes and leaf nodes. Every node can have only one father node and two or more child nodes. Nodes are connected with each other by branches. Each node is passed through certain test properties. Similarly, each leaf node corresponds to a class property. Not only decisions rules are designed following a tree architecture, but also a group of IF-THEN rules is also used.

Classification rules in turn are easy to interpret and can serve as a knowledge base for further classification of satellite image. It is easy to insert additional layers of ancillary data with decision trees due to its nonparametric nature. The basic scheme in decision tree classification is to mask every target as an image layer, so that the influence of one target on the other is minimum. In comparison to decision properties, rules are more popular, because of their simplicity, flexibility, and convenience to use to build up the base of an expert system. The decision tree algorithms are ID3, CD4.5, CART, etc. Decision tree classification is widely used for the classification of remote sensing images for the extraction of information and utilization of land use coverage.

10.5   System Methodology

The Landsat datasets are taken from UCI machine repository generated by the Australian Centre for Remote Sensing. A sample database includes 82 rows and 100 columns from the original data. The binary values were converted to their present ASCII and the classification for each pixel was performed. This data is of numerical and at a single resolution and is highly suitable for standard maximum likelihood classification (Fig. 10.3).

The Landsat MSS imagery consists of four digital images in which two of them are in the visible region mostly in green and red regions and two are in the infrared regions. Each pixel is an 8-bit binary word, with 0 corresponding to black and 255 to white, and the pixel spatial resolution is about 80m x 80m. Each image contains
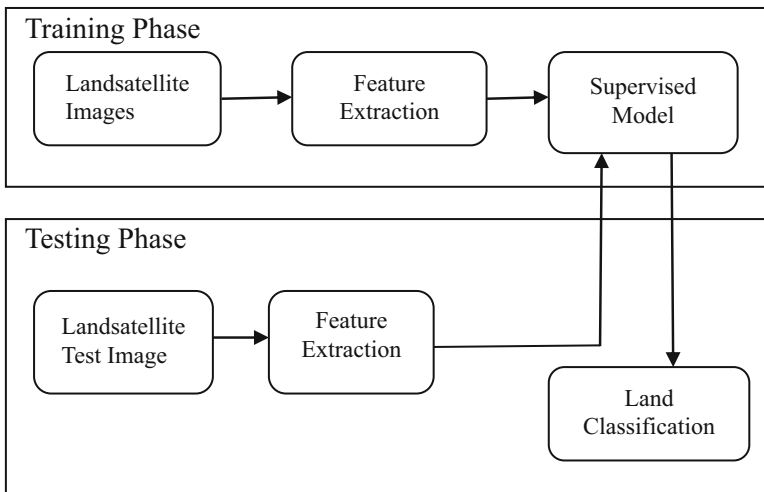
Fig. 10.3 Proposed system methodologies

2340 × 3380 such pixels. The database is a (tiny) subarea of a scene, consisting of 82 × 100 pixels. Each line of data corresponds to a 3 × 3 square neighborhood of pixels completely contained within the 82 × 100 subarea. Each line contains the pixel values in the four spectral bands (converted to ASCII) of each of the 9 pixels in the 3 × 3 neighborhood and a number indicating the classification label of the central pixel. The classes used in the datasets are red soil, cotton crop, gray soil, damp gray soil, soil with vegetation stubble, and very damp gray soil. The four spectral values for the top-left pixel are followed by the four spectral values for the top-middle pixel and then those for the top-right pixel and so on with the pixels read out in sequence from left to right and top to bottom. The spectral values for the central pixel are given by attributes 17, 18, 19, and 20.

10.6   Results and Discussion

For the experimentation, Weka toolkit is used for land coverage data and the performance and effectiveness of the data are identified through prediction models. The various Feature Selection Methods chosen are cfssubset evaluation, consistency subset evaluation, and filtered subset evaluation. The features selected and the features subjects are listed in Table 10.4.

The classifier model is built for the chosen feature subset such as J48 classifier, BF Tree, J48 Graft, NB Tree, Random Forest, and LAD Tree supervised classifier. The performance of an identified classifier is validated based on error rate and computation time. The classification accuracies are measured in terms of precision, recall, and F-score.

The evaluation parameters are the correctly classified instance (CCI) and incorrectly classified instance (ICCI). The error measures recorded for the classifier performance are mean absolute error (MAE) and root mean square error (RMSE). The classifier performance results are shown from Tables 10.5, 10.6, 10.7, 10.8, 10.9, and 10.10.

Table 10.4 Optimal feature subset

Feature selection methods	Search methods	Features selected	Total features subset
Cfssubseteval	Best first	1,2,4,5,6,9,10,12,13, 14,16,17,18,20,21, 22,24,25,26,28,29, 32,33,34,36,class	26
Consistency subseteval	Best first	4,5,9,11,13,16,17,18, 19,23,27,33,36,class	13
Filtered subseteval	Best first	1,2,4,5,6,9,10,12,13, 14,16,17,18,20,21, 22,24,25,28,29, 32,33,34,36, class	24

Table 10.5 J48 classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-Score	MAE	RMSE
AD features	81.67	18.32	81.11	81.17	81.20	5.6	21.47
Cfssubseteval	82.45	17.50	81.5	82.4	81.7	5.63	21.17
Consistency subeval	82.26	17.73	81.4	82.3	81.6	5.73	20.99
Filtered subseteval	82.44	17.55	81.5	82.4	81.7	5.63	21.17

Table 10.6 BF classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-score	MAE	RMSE
All features	85.18	16.81	82.1	83.6	82.2	6.8	19.85
Cfssubseteval	83.59	16.40	82	83.6	82.2	6.83	19.8
Consistency subeval	82.90	17.09	82.1	82.9	82.3	5.86	20.5
Filtered subseteval	83.59	16.40	82.0	83.6	82.2	6.83	19.8

Table 10.7 J48 graft classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-score	MAE	RMSE
All features	85.57	14.12	81.4	83.0	82.0	5.26	20.86
Cfssubseteval	83.45	16.54	81.7	83.5	82.3	5.37	20.68
Consistency subeval	82.63	17.36	81.4	82.6	81.8	5.61	20.8
Filtered subseteval	83.45	17.36	81.7	83.5	82.3	5.37	20.68

Table 10.8 NB tree classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-score	MAE	RMSE
All features	76.06	23.03	81.0	78.9	79.3	6.12	23.78
Cfssubseteval	78.56	21.43	80.5	78.6	78.9	6.11	23.27
Consistency subeval	77.60	22.39	78.8	77.6	77.4	6.52	23.4
Filtered subseteval	78.56	21.43	80.5	78.6	78.9	6.11	23.27

Table 10.9 Random forest classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-score	MAE	RMSE
All features	83.11	16.81	87.1	88.3	87.0	5.81	15.88
Cfssubseteval	87.75	12.24	87.0	87.8	86.5	5.82	15.99
Consistency subeval	79.15	20.81	74.8	79.2	76.6	8.3	20.3
Filtered subseteval	87.75	12.24	87.0	87.8	86.5	5.82	15.99

The performance metrics of CCI and ICCI for different supervised tree classifier chosen for efficient features subsets are listed in Table 10.11. The comparative performance measure graph is shown in Fig. 10.4.

Table 10.10 LAD classifier results

Feature selection methods	Performance		Accuracy			Error measure	
	CCI	ICCI	Prec	Recall	F-score	MAE	RMSE
All features	85.75	14.12	77.2	80.4	78.7	80.25	20.26
Cfssubseteval	87.75	12.24	87.0	87.8	86.5	5.82	15.99
Consistency subeval	79.15	20.81	74. 8	79.2	76.6	8.3	20.3
Filtered subseteval	80.89	19.10	77.8	80.9	99.2	8.05	19.98

Table 10.11 Comparative classifier performance analysis

Supervised classifiers	Performance	
	CCI	ICCI
All features	81.67	18.32
J48 classifier	82.45	17.5
B-Tree classifier	83.59	16.40
J48 graft	83.45	16.54
NB tree	78.56	21.43
Random forest	87.75	12.24
LAD tree	87.75	12.24
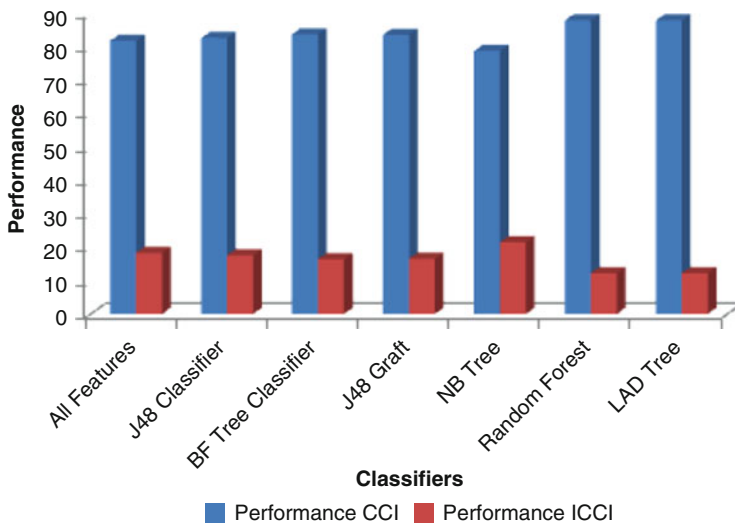
Fig. 10.4 Comparative classifier performance analysis

The classifier accuracy in terms of precision, recall, and F-score is listed in Table 10.12. The comparative performance measure graph is shown in Fig. 10.5.

The classifier error measures in MAE and RMSE are listed in Table 10.13. The comparative performance measure graph is shown in Fig. 10.6.

From the experimental results analyzed, cfs subset feature subset selection method produces the optimal features of 26 features from the total features of 37 features. The optimal subset features produces the better classifier accuracy results

Table 10.12 Comparative
classifier accuracy analysis

Supervised classifiers	Precision	Recall	F-score
All features	81.11	81.17	81.20
J48 classifier	81.11	81.17	81.20
B-tree classifier	82.1	83.6	82.2
J48 graft	81.4	83.0	82.2
NB tree	81.0	78.9	79.3
Random forest	87.1	88.3	87.0
LAD tree	77.2	80.4	78.7
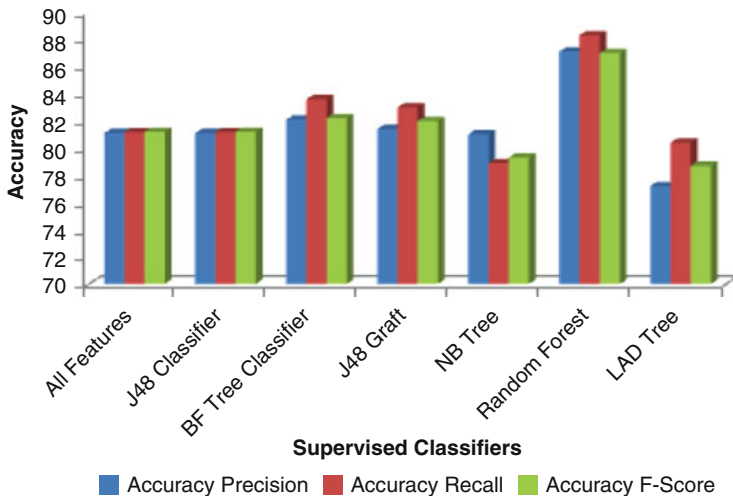
Fig. 10.5 Comparative classifier performance analysis

Table 10.13 Comparative
classifier error measure
analysis

Supervised classifiers	MAE	RMSE
All features	5.4	20.12
J48 classifier	5.6	21.47
B-tree classifier	6.8	19.85
J48 graft	5.26	20.86
NB tree	6.12	23.78
Random forest	5.81	15.28
LAD tree	5.82	15.99

with supervised classifiers. The supervised tree classifier such as J48 classifier, BF Tree, J48 Graft, NB Tree, Random Forest, and LAD Tree classifier are compared on the feature subsets. The comparative metrics chosen for assessment of better classifier are on the performance, accuracy, and error measures. Random Forest classifier produces outperforming results with other supervised classifier. So it is concluded that Random Forest classifier seems to be the better model for identification of crop category. The Random Forest classifier identifies the crop classification correctly for 88% and 12% only produces misclassification.
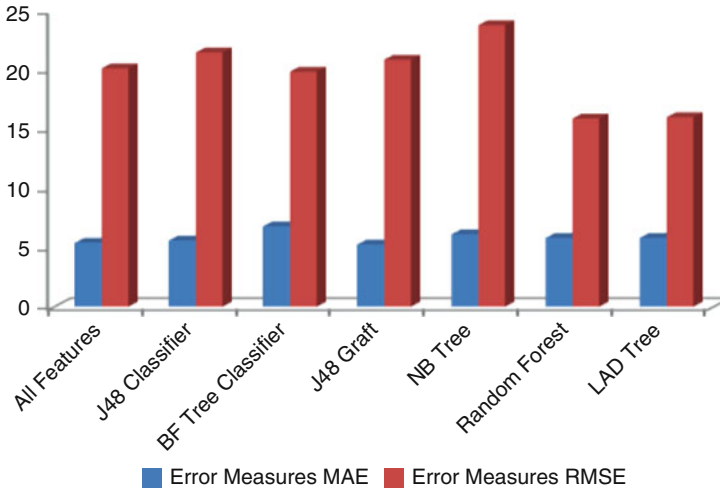
Fig. 10.6 Comparative classifier error measure analysis

10.7 Conclusion

In this chapter, crop classification and mapping technique have been devised for crop classification and mapping, and their performance is also measured and compared with other supervised classifier model. The produced classification results are very much promising with 88% accuracy of correct classification, and F-score is of 87% with reduced error measures of 12%. The proposed method may provide an adequate support to the farmers in differentiating crops suitable for the given land cover with high accuracy and of low error measures. The research can be further enhanced to develop better preprocessing techniques with an improved classifier performance to reduce both false positives and false negatives, employing high-resolution real-time satellite images of Indian agriculture system.

References

1. Barrett R, Crowther P, Laurence R, Lincolne R (2000) Agricultural crop identification using spot and landsat images in Tasmania. In: International archives of photogrammetry and remote sensing. vol XXXIII, Part B7. Amsterdam
2. Beeresh HV, Latha BM, Thimmaraja Yadava G, Dandur N (2014) An approach for identification and classification of crops using multispectral images. Int J Eng Res Technol (IJERT) 3(5). https://doi.org/10.1007/978-3-319-11933-5_65. ISSN: 2278-0181
3. Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: benchmark and state of the art. Accepted by proceedings of the IEEE
4. Czajkowski K, Fitzgerald S, Foster I, Kesselman C (2001) Grid information services for distributed resource sharing. In: 10th IEEE international symposium on high performance distributed computing. IEEE Press, New York, pp 181–184

5. Daniel M, Howard A, Bruce K, Wylie B, Larry L, Tieszen B (2012) Crop classification modeling using remote sensing and environmental data in The Greater Platte River Basin, USA. Int J Remote Sens 33(19):6094–6108
6. Dey A, Dwivedi BS, Bhattacharyya R, Datta SP, Meena MC, Das TK, Singh VK (2016) Conservation agriculture in a rice-wheat cropping system on an alluvial soil of north-western Indo-gangetic plains: effect on soil carbon and nitrogen pools. J Indian Soc Soil Sci 64(3):246–254. https://doi.org/10.5958/0974-0228.2016.00034.7
7. Dhumal RK, Rajendra Y, Kale KV, Mehrotra SC (2013) Classification of crops from remotely sensed images: an overview. Int J Eng Res Appl (IJERA) 3(3):758–761. ISSN: 2248-9622 www.Ijera.com
8. Durgun YÖ, Gobin A, Van De Kerchove R, Tychon B (2016) Crop area mapping using 100-M Proba-V time series. Remote Sens 8:585. https://doi.org/10.3390/Rs8070585
9. Foster I, Kesselman C (1999) The grid: blueprint for a new computing infrastructure. Morgan Kaufmann, San Francisco
10. Lugonja P, Brkljac B, Brunet B (2014) Classification of small agricultural fields using combined landsat-8 and rapideye imagery: case study of Northern Serbia Vladimir Crnojevic. J Appl Remote Sens 8(1):083512
11. Foster I, Kesselman C, Nick J, Tuecke S (2002) The physiology of the grid: an open grid services architecture for distributed systems integration. Technical report, Global Grid Forum
12. Tiana H, Bangera K, Bo T, Dadhwal VK (2014) History of land use in India during 1880–2010: large-scale land transformations reconstructed from satellite data and historical archives. Glob Planet Change 121:78–88. https://doi.org/10.1016/j.gloplacha.2014.07.005
13. Kussul N, Lavreniuk M, Skakun S, Shelestov A (2017) Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci Remote Sens Lett 14(5):1–60
14. May P, Ehrlich HC, Steinke T (2006) ZIB structure prediction pipeline: composing a complex biological workflow through web services. In: Nagel WE, Walter WV, Lehner W (eds) Euro-Par 2006. LNCS, vol 4128. Springer, Heidelberg, pp 1148–1158
15. Moorthi MS, Oza MP, Misra I, Gambhir RK, Darji NP, Sharma S, Jain DK, Dhar D, Padia K, Ramakrishnan R, Chowdhury S, Parihar JS (2014) FASALSoft – An ISRO software framework for crop production forecast using remote sensing data analysis. J Geomatics 8(1):27–33
16. Nigam R, Bhagia N, Vyas S, Manjunath R (2015) Estimation of rabi crop area progression from INSAT 3A CCD, SAC/EPSA/BPSG/CAD/FASAL-R&D/SR/03/2015
17. Ormeci C, Alganci U, Sertel E (2010) Turkey identification of crop areas using spot – 5 data, an article
18. Rustowicz RM (2017) Crop classification with multi-temporal satellite imagery, an article
19. Salehi B, Daneshfar B, Davidson AM (2017) Accurate crop-type classification using multi-temporal optical and multi-polarization SAR data in an objectbased image analysis framework. Int J Remote Sens 38(14):4130–4155. https://doi.org/10.1080/01431161.2017.1317933
20. Schmedtmann J, Campagnolo ML (2015) Reliable crop identification with satellite imagery in the context of common agriculture policy subsidy control. Campagnolo Remote Sens 7:9325–9346
21. Singla N, Babbar B (2015) Rodent damage and infestation in wheat and rice crop fields: district wise analysis in Punjab State. Indian J Ecol 37(2):184–188
22. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197

Chapter 11
Next-Generation Artificial Intelligence Techniques for Satellite Data Processing

Neha Sisodiya, Nitant Dube, and Priyank Thakkar

Abstract In this chapter, we have tried to cover majority of the artificial intelligence (AI) techniques that has contributed to the remote sensing community in the form of satellite data processing, right from the basics to advanced level. A wide variety of applications and enormous amount of satellite data growing exponentially has critical demands in speedup, cost cutting, and automation in its processing while maintaining the accuracy. We have started with the need of AI techniques and evolution made for revolutionary changes in remote sensing and other areas. Subsequently, the traditional ML techniques and its limitations, advancements, and need of introducing DL in various applications are reviewed with what is the present requisites and expectation from AI community to overcome the issues and meet the upraised demands by emerging applications. We concluded that ML and DL technology should integrate with big data technologies and cloud computing to meet the future needs.

Keywords Satellite images · AI · Machine learning · Deep learning · Hyperspectral · Multispectral

11.1 Introduction

Satellite data requirements and advancements in sensor technology have led to large number of operational satellites and a huge volume of satellite. There are number of sensors operating in optical, microwave, and infrared bands, with different spectral, spatial, and radiometric resolutions. This data is being used in large number of

N. Sisodiya (✉) · P. Thakkar
Institute of Technology, Department of Computer Science Engineering, Nirma University, Ahmedabad, India
e-mail: neha.sisodiya@nirmauni.ac.in

N. Dube
Space Application Center–Indian Space Research Organization, Ahmedabad, India

applications in the field of defense intelligence, policy decision-making, urban planning, vegetation monitoring, natural resource monitoring, climate change studies, geo-hazards monitoring, ocean monitoring, and many more. Current satellite data processing techniques are based on physical retrieval, probabilistic approaches, and statistical models and are associated with a wide range of challenges [1] which includes high dimensionality, uncertainty, nonlinearity, spatial and spectral redundancy, etc. Apart from the current automated techniques for satellite data processing, still a large number of applications require manual intervention and use of human intelligence for decision-making. Artificial intelligence and machine learning techniques are trying to address these issues machine (computer) process like humans. Artificial intelligence (AI) is a term coined by Alan Turing, who was a mathematician and has laid the foundations of the computers for modern age. In around 1950s, his work transpires into broad popular comprehension and has given birth to the idea of "General AI." But the characteristics of human intelligence, that include reasoning, interaction, understanding, and thinking, should also be possessed by the computer. The idea was to narrow down the extent of AI technologies that could be tasks specific such as various gaming software, recommendation systems, spam emails identification, etc. [2]. Machine learning in the last two decades has drawn the attention of most of the researchers as all of these tasks exhibit certain portion of human intelligence in the process.

In general, machine learning comprises of set of algorithms or models that have data (training data) specific to an application, that is, can be used for training and inference to be drawn from observed patterns (features trends) obtained through training the data. We get a generous amount of (cleaned or uncleansed) data, with the features defined manually (e.g., "weight," "color," "spam email," etc.). Prediction on new data (testing data) from the inference drawn from data trained for specific features and are relatively tuned all features. For example, if we have one image and we need to make out whether the object within the image is a human being or table, we need to take a large amount of image data which contains vegetation land and water body (this is termed as "labeled data"). The algorithm now will identify the features common in all of those images for both vegetation and water. The same algorithm has to be applied to the unlabeled data, the provided image, to predict whether the image contains human being or animal. In the field of satellite image analysis, machine learning plays a vital role, since AI-driven satellite data applications are in demand for a number of important reasons:

- The rapid advancements in machine vision in the last few years have made challenging tasks (such as identifying cars, buildings, or changes in scene over time) and can be done by machines.
- With the satellite proliferation, camera technology enhancement, improved data storage, and transfer competencies, there is an exponential increase in amount of data being produced from satellites.
- Satellite image analysis and interpretation performed by human imagery analysts are a costly affair

  .

Machine learning models are able to recognize man-made structures in a satellite image as well as airplanes parked at an airport. Beside of all these abilities, a lot of human intervention is required in ML, such as manually choosing the features (e.g., look for the shape, size, color, texture, etc.).

Limitations of traditional machine learning model in remote sensing are:

- It is lacking in representation of large and complex models for the systems not linear in nature.
- Unable to handle large volumes of data
- Difficult to learn hierarchical features and generalize
- Lack of training data and use of domain knowledge
- Compatibility to high-performance computing architectures
- Accuracy in prediction, forecasting, and classification
- Address optimization, scalability, and portability

At present, around 180 plus remote sensing satellites have already been launched in 2006–2018 and are operational. The expected number is going to rise by twice or thrice in next few decades [1]. Deep learning (DL) is a subset of ML which in turn is a subset of AI. It is based on neural networks, a conceptual model of the brain that has been around since the 1950s but largely ignored earlier. That's because they are very computationally expensive. In recent times, processing has become sufficiently cheap and powerful through graphics processing unit (GPUs) and field-programmable gate array (FPGAs), and there has been enough data to feed the DL algorithms, and hence DL is becoming more popular nowadays. DL handles almost every limitations imposed by traditional machine learning with its advanced algorithms and power of computation. Though deep learning is also facing the challenges imposed by satellite data, we are able to overcome most of the challenges by traditional machine learning algorithms. In this chapter, we have discussed both traditional and new trends in machine learning for satellite data processing with focus on present status and future directions in both.

11.2 Machine Learning in Satellite Data Processing

Several machine learning techniques have contributed to reveal the insights of the data provided by satellites orbiting Earth. Machine learning has to be integral part of every step involved in processing satellite data, from preprocessing to decision-making. Stages of processing satellite data are shown in Fig. 11.1. It shows expansion of information domain, i.e., revealing the facts hidden inside, as we process the data from acquisition to decision and policy making to analytics.

The three primary applications, namely, classification (grouping similar pixel together), segmentation (dividing the image into different regions to detect objects), and denoising (making an estimate of the obtained image), have been studied in [3]. Application areas utilizing satellite image data such as change detection, land use land cover, vegetation monitoring, etc. require classification of satellite image under
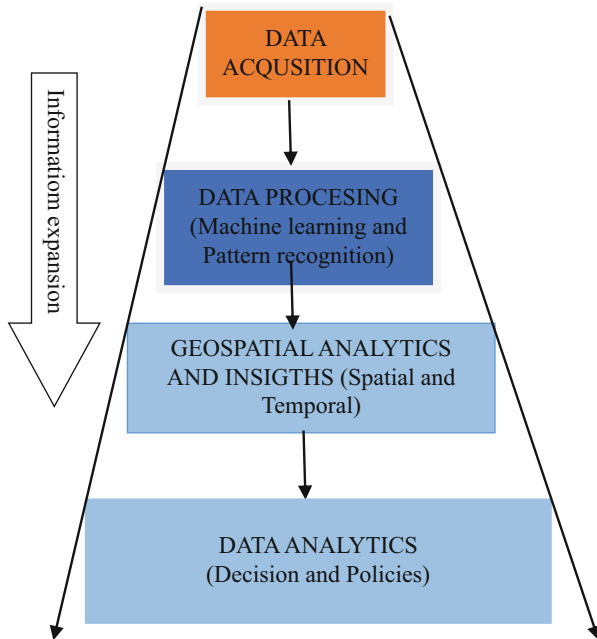
Fig. 11.1 Stages involved in processing satellite data

investigation, whereas segmentation is required for urban growth monitoring, road extraction, building extraction and detection, etc., and denoising is a preprocessing stage that shares the importance with both classification as well as segmentation. In order to accomplish this task, various algorithms have been introduced which require the number of feature to be selected by investigator. Traditional manual features are the Histogram of Ordered Gradients (HOG), the Scale-Invariant Feature Transform (SIFT) and its variants, Speeded-Up Robust Features (SURF) [4], color histograms, etc. Principal components analysis (PCA) is one of the most commonly used feature space (dimensionality) reduction techniques. Regularized linear discriminant analysis (LDA) is used for classification of hyperspectral image. K-means clustering is popular in clustering satellite data. Sparse coding is used for sparse representation of data. There is a lot of research that has already been registered in the area of hyperspectral image (HSI) data analysis, kernel-based methods, methods with statistical learning for HSI, spectral distance functions for classification, hidden Markov random field (HMRF), generalized least squares, multi-classifier systems, fuzzy-based, spectral-spatial classification, change detection, vertex component algorithm (VCA), orthogonal subspace projection, support vector domain description (SVDD), Gaussian processes (GP), genetic algorithm, manifold learning, graph-based methods, transfer learning, endmember extraction, and spectral unmixing [5]. These algorithms in combination with few others like selection of feature through genetic algorithm in combination with fuzzy logic-

Table 11.1 Various traditional machine learning algorithms with applications

Processing stage	ML method	Application
Preprocessing	GMM, ANN	Denoising
	PCA, LLE, ISOMAP	Dimensionality reduction
Processing/analysis	Sparse coding	Sparse representation
	HOG, SURF, SIFT, decision trees, random forest, genetic algorithm, HMRF, SVM, MRF	Feature selection
	Decision tree, multilayer perceptron, logistic regression, SVM, K-means, GP, NN, ARIMA	Segmentation and classification
	K-nn, SVM, SOM, GMM, K-means, fuzzy clustering, hierarchical clustering, hybrid clustering	Clustering
	Image transformations, correlation analysis	Change detection

based classification for accuracy improvement over high-dimensional HSI data are also found in literatures. We will discuss about few recent advancements in these fields with applications and benefits and elaborate them in subsequent sections (Table 11.1).

11.2.1 Satellite Image Classification

Machine learning in general is categorized broadly into two categories: supervised and unsupervised. Supervised learning deals with the training data which has labels associated with it. Neural network [6] and SVM have gained most successful models under this category for processing satellite data. In unsupervised learning, data itself is used for learning the patterns (features). Fuzzy -based clustering [7, 8] and multiobjective optimization [9] are used for developing spatial membership relations. A fusion of information from multiple sources has been introduced with graph cuts [10], projection based [11], hierarchical clustering [12], hidden Markov random fields, Markov random fields (MRF) [13] for contextual regularization, self-organizing maps (SOM), and hybrid genetic algorithms [20] and has been reviewed [1]. Segmentation, region extraction is applied to the image with multiple components using ANN and genetic algorithm based approach is being introduced in [14] is done in an image with multiple components. All methods mentioned above work well for spatial data, i.e., these are considered as pixel-based classification, but there are certain applications that appeal consideration of the pixel location over a period of time along with the captured scene termed as spatiotemporal analysis. People also have worked on dynamic clustering strategies for spatiotemporal reasoning [15] and visualization [16]. Sometimes in an application like climate change and prediction-based application, time series analysis techniques are necessary and critical part of

our analysis. Linear regression and autoregressive moving average (ARIMA) for time series analysis are used in [17] to study rainfall and temperature trends in Bangladesh. In [18], authors have used data from Moderate-Resolution Imaging Spectroradiometer (MODIS) with a spatial resolution of 500 m to exploit global urban extent map using supervised decision trees for classification of images taken over a period of 1 year.

Change detection is one of the applications that uses spatiotemporal data to see the change in trends that have been observed over a period of time at a particular location. A machine learning model using decision tree has been employed to observe change in multiscale imagery [19]. Object-based change classification with combining object correlation images (OCIs), object-based change classification integrating neighborhood correlation images (NCIs), object-based change classification without contextual features, per-pixel change classification with NCIs, and traditional per-pixel change classification using only bitemporal image data are done. Along with this, a machine learning decision tree and nearest neighbor were also investigated. Comparison between the OCI and the NCI variables was evaluated. Object-based change classifications with incorporation of OCIs or the NCIs resulted more accurate change detection classes [20].

Image transformation subtraction, rotation, analysis of change vector, or cross-correlation analysis are also utilized as change detection approaches [1]. In the past few decades, neural networks and kernel methods are also widely used approaches. Composite kernels have been specifically designed for the combination of multitemporal, multisensor, and multisource information [21, 22]. People are now focused on exclusion or reduced to a large extent the human intervention by utilizing completely unsupervised or semisupervised approaches, respectively [23, 24].

11.2.2 Kernel Base Extraction

Working with hyperspectral images is challenging due to its high dimensionality and nonlinear nature results in higher computational time, and the presence of high colinearity and noisy bands vitiate the model's quality. Spectral bands represent the characteristics or features of the elements under consideration to model. Various techniques have been proposed to study feature selection in satellite images classical discriminative criteria [25], and a bit advanced one uses machine learning, such as genetic algorithms [26] or SVM-based recursive feature elimination [27, 28]. Recently, more attention has been paid on feature extraction methods. PCA is one of the most widely used linear methods. Later on, multivariate kernel machines were proposed to deal with nonlinearities in the data [29]. In [30], Gaussian process-based classification is also done for hyperspectral images.

11.2.3 Pure Pixel-Based Extraction

A pixel in satellite images is a blend of features (spectral signatures) of various objects or materials found in the spatial extent in the scene. An automatic extraction mechanism right from the image, to get the pure spectral pixels, termed as endmembers is required to be developed: NFINDER algorithm, VCA (vertex component algorithm), orthogonal subspace projection, and SVDD (support vector domain description) [1]. Subpixels in the images can be identified using pure pixels as they serve as basis to represent all other pixels as a linear (or nonlinear) combination of them or mineral mapping. Support vector domain description (SVDD) has also been used for pure pixels selection [31].

11.2.4 Regression

Prediction models are based on establishing and learning the relationship between the evaluated (observed values) and the ground truth available to validate the relationship. Empirical models tune to learn the relationship between the obtained spectra and actual measurements done on ground. Due to certain drawbacks, parametric models have landed to inaccuracy in prediction over new data, which has led to nonparametric and nonlinear regression techniques to come into picture, such as neural networks, support vector regression (SVR) [33], semisupervised SVM for parameter estimation [34], relevance vector machines (RVM) [35], or Gaussian processes (GP) [30, 36]. Due to lack of interpretability and dependency on training data, few analytical models with better accuracy also limit its utilization.

11.3 Recent Trends in Machine Learning for Satellite Data Processing

11.3.1 Manifold Learning

Manifold learning deals with dimensionality reduction and nonlinear feature extraction. The field is span throughout computer science, machine learning, image processing, etc. Manifold learning is focused toward projection of high-dimensional data into a lower dimension and provides a better analysis by preserving main features of the original data. High-dimensional data visualization and understanding become viable. Intrinsic structure of satellite data cannot be described using traditional linear dimensionality reduction methods. Isomap [37, 38] and Laplacian methods, such as unnormalized graph Laplacian, are used as regularization technique for SVM [39]. A manifold regularization technique also includes Laplacian regularization (LapR and HLapR) in [40]. Graph-based Laplacian energy [41] is

used for hierarchical image analysis. Local linear embedding (LLE) transforms a very high-dimensional space-embedded images into two dimensional and makes visualization much simpler to analyze the whole data [42]. A nonlinear dimensionality reduction using LLE is done in [43]. The extension over LLE has been proposed in [44] to provide a supervised feature extraction technique. Image denoising is also done using LLE [45]. People also have made efforts to strengthen the discrimination capability and ability to generalize the representation of embedded data [46]. Some algorithms that analyze the intrinsic dimensionality of hyperspectral images have been mentioned [47].

11.3.2 Semisupervised

An area similar to manifold learning is semisupervised learning, which makes use of both labeled data and the wealth of data (unlabeled) samples for development of model using manifold data structure. In perspective of remote sensing data, a variety of methods have been developed that are either generative or discriminative. Conditional density estimation in inclusion of generative models has been presented [48]. A number of graph-based methods have been developed for classification [49, 50], regression, as well as target detection. An adapted graph-based SVM classification technique for time series data is used for analysis and classification of satellite images and a new graph kernel is being designed for the same [51]. The design of cluster and bagged kernels has been successfully presented [52]. A technique applied for image classification and change detection [53] which has been using SVM known as transductive SVM (TSVM) has been developed [54, 55]. Also, a modified TSVM is proposed in [56] to address ill-posed problems in remote sensing. In [57], Fisher discriminant classifier was proposed which is also a semisupervised kernel. The problem faced by these methods is incapability to handle large-scale dataset, i.e., if the number of unlabeled samples is very large, these methods cannot be applied directly.

11.3.3 Transfer Learning

Transfer learning or domain adaptation in view of remote sensing is the problem that arises when training samples are available for a particular time, and we need to classify time series data to update land cover maps. Few methods, like NN and domain adaption SVM (DASVM), have been used [1]. Another problem to address is that for classification of an image with samples taken from different images, which results biasing in sample selection or covariance shift. This has been addressed by defining proper kernel machines [58]. Recently, in [59], authors have proposed maximum margin-based clustering that has used the common features from both the domains (target and source) avoiding the samples only from source domain.

11.3.4   Active Learning

Recently, selection of most relevant sample to train the model is introduced which is termed as active learning. In [60], authors have mentioned object-oriented classification using SVM for pixel-based classification using maximum likelihood classifiers. The extension of this approach that uses boosting for weighing of few selected pixels has been done iteratively. Also, an information-based active learning was introduced for target detection. For very high-resolution satellite images, a model-independent active learning method is also reviewed [1]. In order to achieve improved accuracy in classification, an improved training samples selection strategy has been proposed with the active learning approach utilizing two-staged spatial computation [61]. There is a lot of scope in the field of active leaning since selection of the most informative sample is difficult. To address this issue, authors have employed a method that considers both classification and localization of the detected object [62]. Two metrics have been proposed, one of which evaluates the overlapping of the object and bounding box (resultant of prediction) and the second shows stability of detection in presence of noise.

11.3.5   Structured Learning

Majority of the methods mentioned in previous sections assumes binary output with two classes. But, most of the application of satellite data deals with multiclass output which further increases the complexity for both classification and prediction. This is the part of structured learning where multiple labels can be predicted simultaneously; computer vision-based structured SVM (SSVM) [63] and very few applications have been presented [64] for structured learning.

Kernel based image classification using structured learning in [64] is introduced to overcome issues of noise in very high-resolution image. An efficient with enhanced performance-based multiclass classification-based hierarchical spectral clustering is introduced very recently for incorporating the scalability aspect of processing large satellite image data [65].

The issues with almost all of the methods mentioned in previous sections can be addressed by making use of many intermediate layers that can take care of signature extraction automatically and also capability of learning from a large dataset. Secondly, to deal with time series data, a memory-based architecture is required to be used. These issues have drawn attention of the researchers to move on to the area well known as deep learning, which eliminates the human intervention for feature extraction, making it robust and easy to use for complex applications in satellite data processing.

11.4   Deep Learning

Deep learning refers to a deeper network, with many strata (layers), typically at least four or five layers deep, of (usually) nonlinear transformations. Its unsupervised nature of learning from labeled data and ability to generalize over unlabeled data made it more popular to use in the field of satellite data processing. Deep neural network (DNN) is just not a solution to image classification or prediction but can also be used for feature-based image registration [66] in order to get a robust and accurate match in presence of noise in synthetic aperture (SAR) data. Change detection is also being done using deep learning architecture [67].

11.4.1   Convolutional Neural Network (CNN)

CNN are well-known deep learning network, a stratified architecture which is comprised of convolutional layer, nonlinear layer, and pooling layer. CNN transform input to an output class (prediction) and recognize hierarchically. PCANet proposed in [68] an unsupervised convolutional deep learning network architecture formed using cascading principal component analysis to learn filter banks having multiple stages followed by binary hashing used for indexing purpose and block-wise histograms serve as pooling layer. For estimation of crop yield [69], CNN is being utilized through Convolutional Architecture for Fast Feature Embedding (Caffe), a framework for deep learning, in two different ways: a model with two inner product layer and the other one using single inner product layer with rectified linear unit (ReLU) as an activation function.

In [70], authors have reviewed recent advancement in the use of CNN in classification of hyperspectral images. Supervised CNN with 1-D, 2-D, and 3-D CNNs along with their comparison on performance basis have been reviewed. Unsupervised CNN has also been used for hyperspectral image classification to learn spectral-spatial features. The videos obtained from cameras mounted on satellites don't have spatial resolution well enough to interpret the scene due to motion of objects on earth as well as motion of camera; a spatiotemporal analysis by fusion of multispectral images and space videos is done using CNN [71]. A pretrained model transferred to supervised CNN [72] is used to handle high-dimensional data with a simple and computationally efficient approach. In [73], summary of almost all types of CNN and modifications made over it are presented with application and data. Not only for earth, Convolutional Neural Networks have also been used to automatically detect geological landforms on Mars [74].

11.4.2   Recurrent Neural Network

The applications of standard neural networks (and also convolutional networks) are limited because they only accept a fixed-size vector as input (e.g., an image) and

produce a fixed-size vector as output (e.g., probabilities of different classes). Also, these models use a fixed amount of computational steps (e.g., the number of layers in the model). Recurrent neural networks are unique as they allow us to operate over sequences of vectors—sequences in the input, the output, or in the most general case both.

A novel approach for ocean and weather prediction is proposed as a doctoral thesis referred in [73]. Modification on RCNN is done to learn the equations in mesoscale meteorological model. A recurrent neural network (RNN) works for tracking of multiple objects in presence of occlusion in data and is mentioned in [73].

11.4.3 Recursive Neural Network

A recursive NN can be seen as a generalization of the recurrent NN, which is in fact a recursive neural network with the structure of a linear chain. Recursive NNs operate on hierarchical structure. Recurrent NN operates on progression of time. Long short-term memory (LSTM) which is a type of recursive NN used with convolution named as Convolutional LSTM is used for precipitation nowcasting [75]. The ConvLSTM is an extension of fully connected LSTM (FC-LSTM).

11.4.4 Deep Belief Network

A deep belief network (DBN) is a probabilistic, generative model made up of multiple layers of hidden units. It is a composition of simple learning modules that contributes to each layer. A DBN can be used generatively to pretrain a DNN by using the learned DBN weights as the initial DNN weights. Back-propagation or other discriminative algorithms can then be applied for fine-tuning of these weights. DBNs are particularly helpful when limited training data are available. These pretrained weights are closer to the optimal weights than randomly chosen initial weights. People have used DBN [76] for urban planning to effectively extract the features and improve the performance of classification. Detection of aircraft in high-resolution satellite imagery, object recognition, traffic flow prediction, urban land use and land cover (LULC), vehicle detection, nighttime vehicle sensing in far infrared (IR), and prediction of drought index utilize DBN. Classification of polarimetric SAR and HSI made use of Restricted Boltzmann Machine (RBM) and DBN to perform spectral information-based classification [73].

There are few other classes of DNN as well, one of which is stacked auto encoders (SAE) which can be used for HSI classification and dimensionality reduction. All of the methods mentioned above are summarized in Table 11.2 with their application and advantages.

Table 11.2 Comparison of various deep learning algorithms with applications and advantages

DL method	Data type	Application	Advantages
CNN	SAR, HIS, multispectral images, LiDAR data, RGB, pansharpening images	Image segmentation, urban area (building, road, etc.) classification and extraction, urban growth prediction, road extraction, oceanographic target classification, thematic classification, automatic target detection, super-resolution for sea surface temperature analysis, object detection and recognition, 3D object classification, human detection and activity classification, semantic labeling multisource earth observation data, anomaly detection, disease recognition in plants, denoising, poverty, multiscale classification	Automatic feature extraction for learning, from large labeled dataset provided as input. Easy to train. A pretrained model can be prepared using CNN and can be used for transfer learning. CNN can be combined with other types of DL architectures for performance enhancement of the those architectures, e.g., recurrent CNN, convolutional LSTM, etc.
Recurrent NN	Spatiotemporal and time series data	Ocean and weather forecasting, tracking for multiple objects	Used for sequential data
Recursive NN (LSTM)	Spatiotemporal data	Precipitation nowcasting	The presence of memory unit helps it to work well with sequential data.
Deep belief network (DBN)	Polarimetric SAR, HSI, high-resolution image, time series data, radar data, spatiotemporal data	Classification, drought index prediction, object detection, prediction of traffic flow, urban LULC	Can be used with limited available labeled dataset

11.5 Case Studies

We are presenting particular use cases of artificial intelligence in this section. Depending on the way of utilizing the information obtained from processing of satellite data, these can be categorized into two categories:

(a) *Direct application*: Satellite imagery is processed through machine learning techniques straight away that provides insights of the scene, e.g., object detection, vehicle tracking, urban boundaries, road segmentation, building detection, change detection, etc.

(b) *Derived application*: Certain complex and sophisticated model that uses the set of features which in turn used to derive some conclusion or making decision

policies and also utilizes the data other than the data obtained through satellite, e.g., making decision policies for farmers, profit prediction of retailers, etc.

(i) *Object detection in a high-resolution image*

Many challenging applications of satellite data processing has already been discussed in previous sections. Object detection in particular is a challenging task in satellite imagery due to the following reasons:

- Within hundreds of megapixels in satellite image, a tiny part (approximately 10–20 pixels to few hundred pixels) of it constitutes the object in search.
- Lack of availability of training data.
- Optimization of algorithm to detect small objects.
- Adaptation of algorithm to work well with different scales and objects.

Segmented satellite image with identification of several objects is shown in Fig. 11.2. Different colors used for different objects are marked in legend.

The major issues in the object detection, change detection, and time series analysis-based application have been addressed by few popular deep learning architectures such as FasterRCNN and You Only Look Once (YOLO) [77]. The presence of unsolved challenges made this research area still in an immature stage.

(ii) *Change detection* is another area of application which uses the direct output produced by processing satellite imagery. This represents the difference in the image taken for a particular geographical location at two different times (temporal changes), where significant changes in the area (region) are expected to occur as shown in Fig. 11.2. There can be a simple
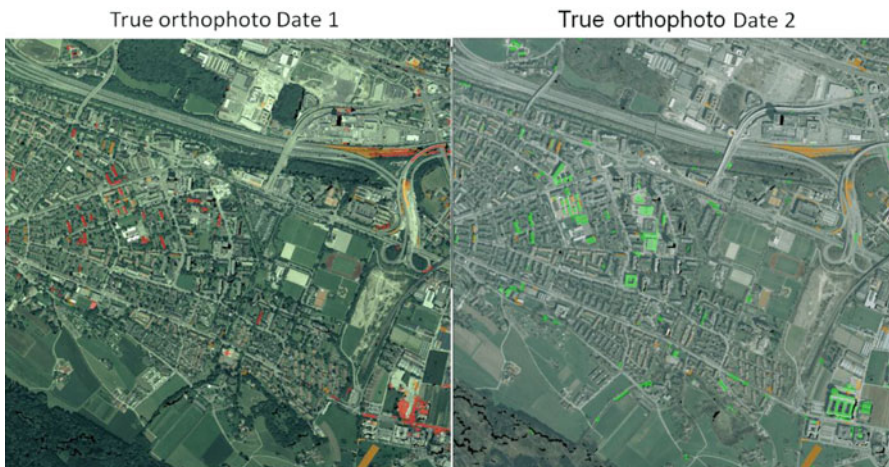
Fig. 11.2 Demonstration of time series analysis and change detection algorithm. The changes in buildings and roads are highlighted red (before) to green (after) [78]

binary change, where a pixel belongs to either of the two expected class or a multiclass change, where a pixel can be labeled from a set of class labels.

Multiclass classification requires supervised techniques to be used, which uses training data (ground truth observation) for classification. Developing a solution to multiclass classification is extremely difficult and complex due to following issues:

- Complex and expensive ground truth data collection.
- Data normalization.
- Effects of lenses, climate changes, and other natural changes.
- Thorough knowledge of remote sensing is required to choose a machine learning technique.

(iii) *Profit earned by retailers*: Unlike direct applications mentioned previously, derived ones extract information from satellite imagery which serve as base features to complex systems. In retail sector, number of cars parked at parking lot can give the estimate about the profit that can be made by the retailer. The predicted profit then can be reported in form of monthly, quarterly, or annual reports. The count for number of cars present in the parking area for the scene captured by satellite image is presented in Fig. 11.3.

(iv) *Crop yield estimation and price prediction*: Normalized vegetation index (NDVI) is derived product from various satellites like Landsat, MODIS, OCM, Sentinel, etc. NDVI provides a crucial information used for crop yield estimation and price prediction. The farmers, commodity traders, insurance policy makers, government policies in agribusiness, and many more are utilizing agriculture-related intelligence to accomplish their needs. AI techniques are utilized in identifying higher yield area to help farmers to choose the best time and places for farming for a particular crop. Figure 11.4 shows the domination of sugar plantation in the island.

(v) *Economic growth monitoring*: Satellite images can uncover economic activities of the countries which are hard to reach. The factors affecting or showcasing the economic status of a country are number of high-rise buildings and increase in rate of construction, electricity consumption (can be measured in night hours by measuring the luminosity), number of cars, roads, etc. The prediction of poverty can also be done with measuring the factors mentioned as well as making a per capita income map, which also contributes to policies for making sustainable development growth (Fig. 11.5).

Fig. 11.3 Car-counting algorithm [79]

11.6 Conclusion

Artificial intelligence that includes using neural networks, deep learning, or other algorithms that make the computer do tasks that would be hard to program with conventional methods can be more efficient and easier way. Data processing of satellite images and the evaluation of extremely large datasets are two common use cases of AI in space applications. The future of satellite processing lies in use of deep architectures as well as big data technology. Since there are a number of open challenges that need to be addressed in this area and big data is one of them. Storage and archival of that huge satellite data is another problem to address that can be fixed by utilizing cloud computing. We have reviewed many AI algorithms that have been utilized to address few problems faced by satellite image processing community, but these are restricted to the type of data, type of sensors, availability of limited labeled dataset, etc. Now, there is a need of a common platform wherein the data, irrespective of the source, type and data inadequacy, etc.,
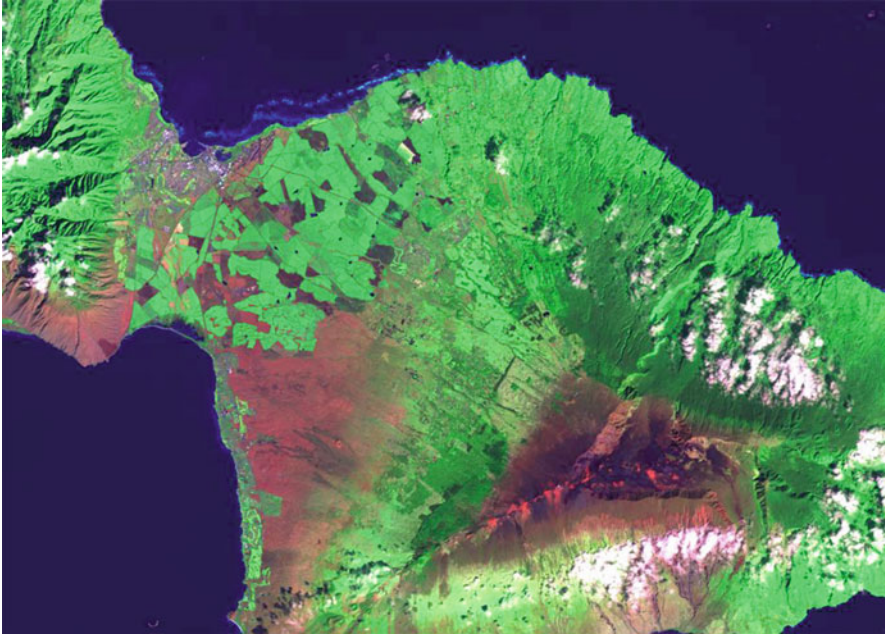
Fig. 11.4 In this infrared view of Maui presented in false color, the large sugar fields that dominate the agriculture of that island can be seen in bright green at the narrowest part of the island, wedged in the low lands between the mountains of Pu'u Kukui on the east and Haleakala in the west [80]

Fig. 11.5 Construction rates monitoring (urbanization) using shadow detection [81]

can be processed automatically with least or no human intervention. The extension of machine learning and deep learning architectures for big data technologies in composition with cloud computing architectures are still open issues.

References

1. Camps-Valls G (2009) Machine learning in remote sensing data processing. IEEE international workshop on machine learning for signal processing, pp 1–6, September 2009
2. https://towardsdatascience.com/the-10-deep-learning-methods-ai-practitioners-need-to-apply-885259f402c1
3. Dev S, Wen B, Lee YH, Winkler S (2016) Machine learning techniques and applications for ground-based image analysis. CoRR abs/1606.02811. http://arxiv.org/abs/1606.02811
4. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (SURF). Comput Vision Image Understand 110(3):346–359
5. Ball JE, Anderson DT, Chan CS (2017) A comprehensive survey of deep learning in remote sensing: theories, tools and challenges for the community. CoRR abs/1709.00308. http://arxiv.org/abs/1709.00308
6. Radhika K, Varadarajan S (2018) A neural network based classification of satellite images for change detection applications. Cogent Engineering 5(1):1–9
7. Bilgin G, Erturk S, Yildirim T (2008) Unsupervised classification of hyperspectral-image data using fuzzy approaches that spatially exploit membership relations. IEEE Geosci Rem Sens Lett 5(4):673–677
8. Cannon RL, Dave JV, Bezdek JC, Trivedi MM (1986) Segmentation of a thematic mapper image using the fuzzy c-means clustering algorithm. IEEE Trans Geosc Rem Sens GE-24(3):400–408
9. Bandyopadhyay S, Maulik U, Mukhopadhyay A (2007) Multiobjective genetic clustering for pixel classification in remote sensing imagery. IEEE Trans Geosc Rem Sens 45(5):1506–1511
10. Tyagi M, Bovolo F, Mehra AK, Chaudhuri S, Bruzzone L (2008) A context-sensitive clustering technique based on graph-cut initialization and expectation-maximization algorithm. IEEE Geosc Rem Sens Lett 5(1):21–25
11. Bachmann CM, Donato TF, Lamela GM, Rhea WJ, Bettenhausen MH, Fusina RA, Du Bois KR, Porter JH, Truitt BR (Oct 2002) Automatic classification of land cover on smith island, va, using hymap imagery. IEEE Trans Geosc Rem Sens 40(10):2313–2330
12. Marcal ARS, Castro L (2005) Hierarchical clustering of multispectral images using combined spectral and spatial criteria. IEEE Geosc Rem Sens Lett 2(1):59–63
13. Sarkar A, Biswas MK, Kartikeyan B, Kumar V, Majumder KL, Pal DK (2002) A MRF model-based segmentation approach to classification for multispectral imagery. IEEE Trans Geosc Rem Sens 40(5):1102–1113
14. Awad M, Chehdi K, Nasri A (2007) Multicomponent image segmentation using a genetic algorithm and artificial neural network. IEEE Geosc Rem Sens Lett 4(4):571–575
15. Heas P, Datcu M (2005) Modeling trajectory of dynamic clusters in image time-series for spatio-temporal reasoning. IEEE Trans Geosc Rem Sens 43(7):1635–1647
16. Jia X, Richards JA (2002) Cluster-space representation for hyperspectral data classification. IEEE Trans Geosc Rem Sens 40(3):593–598
17. Rahman MR, Lateh H (2017) Climate change in Bangladesh: a spatio-temporal analysis and simulation of recent temperature and rainfall data using GIS and time series analysis model. Theoret Appl Climatol 128(1-2):27–41
18. Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. Environ Res Lett 4(4):044003
19. Levien LM, Roffers P, Maurizi B, Suero J, Fischer C, Huang X 1999 A machine-learning approach to change detection using multi-scale imagery. American Society of Photogrammetry and Remote Sensing. Annual conference, Portland, Oregon, May 20, 1999
20. Tang Z, Tang H, He S, Mao T (2015) Object-based change detection model using correlation analysis and classification for VHR image. IEEE international geoscience and remote sensing symposium (IGARSS), pp 4840–4843, July 2015
21. Tuia D, Ratle F, Pozdnoukhov A, Camps-Valls G, (2008) Multi-source composite kernels for urban image classification. IEEE Geosc Rem Sens Lett (Accepted)

22. Camps-Valls G, Gómez-Chova L, Mu noz Mar´ı J, Rojo-Álvarez JL, Mart´ınez-Ram´on M (2008) Kernel-based framework for multi-temporal and multi-source remote sensing data classification and change detection. IEEE Trans Geosci Remote Sens 46(6):1822–1835

23. Bovolo F, Bruzzone L, Marconcini M (2008) A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. IEEE Trans Geosc Rem Sens 46(7):2070–2082

24. Ghosh S, Bruzzone L, Patra S, Bovolo F, Ghosh A (2007) A novel context-sensitive technique for unsupervised change detection based on hopfield type neural networks. IEEE Trans Geosc Rem Sens 45(3):778–789

25. Bruzzone L, Serpico SB (2000) A technique for features selection in multiclass problems. Int J Rem Sens 21(3):549–563

26. Pal M (2006) Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. Int J Rem Sens 27(14):2877–2894

27. Archibald R, Fann G (2007) Feature selection and classification of hyperspectral images with support vector machines. IEEE Geosc Rem Sens Lett 4(4):674–679

28. Shaw G, Manolakis D "Signal processing for hyperspectral image exploitation," in IEEE Signal Processing Magazine, vol. 19, no. 1, pp. 12–16, Jan. 2002. doi: 10.1109/79.974715

29. Camps-Valls G, Bruzzone L (2009) Kernel methods for remote sensing data analysis. Wiley, Hoboken

30. Bazi Y, Melgani F (2008) Classification of hyperspectral remote sensing images using gaussian processes. IEEE Intl Geosci Rem Sens Symp (IGRASS) 2:II-1013–II-1016

31. Keshava N, Mustard JF (2002) Spectral unmixing. IEEE Signal Process Mag 19(1):44–57

32. Camps-Valls G, Bruzzone L (2009). Kernel Methods for Remote Sensing Data Analysis. 10.1002/9780470748992.

33. Camps-Valls G, Bruzzone L, Rojo-Álvarez JL, Melgani F (2006) Robust support vector regression for biophysical parameter estimation from remotely sensed images. IEEE Geosc Rem Sens Lett 3(3):339–343

34. Camps-Valls G, Mu˜noz-Mar´ı J, Gómez-Chova L, Richter K, Calpe-Maravilla J (2009) Biophysical parameter estimation with a semi-supervised support vector machine. IEEE Geosc Rem Sens Lett

35. Camps-Valls G, Gomez-Chova L, Vila-Franc´es J, Amor´os- L´opez J, Mu˜noz-Mar´ı J, Calpe-Maravilla J (2006) Retrieval of oceanic chlorophyll concentration with relevance vector machines. Rem.Sens Environ 105(1):23–33

36. Pasolli L, Melgani F, Blanzieri E (2008) Estimating biophysical parameters from remotely sensed imagery with Gaussian processes. In: IEEE international geoscience and remote sensing symposium. IGARSS'08, Boston

37. Bachman CM, Ainsworth TL, Fusina RA (2005) Exploiting manifold geometry in hyperspectral imagery. IEEE Trans Geosci Rem Sens 43(3):441–454

38. Bachman CM, Ainsworth TL, Fusina RA (2006) Improved manifold coordinate representations of large-scale hyperspectral scenes. IEEE Trans Geosci Rem Sens 44(10):2786–2803

39. Gomez-Chova L, Camps-Valls G, Munoz-Mari J, Calpe J (2008) Semisupervised image classification with Laplacian support vector machines. IEEE Geosci Rem Sens Lett 5(3):336–340

40. Ma X, Liu W, Li S, Zhou Y Hypergraph-Laplacian regularization for remote sensing image recognition. Comput Vision Pattern Recog. Submitted, June 2018

41. Huigang Z, Xiao B, Huaxin Z, Huijie Z, Jun Z, Jian C, Hanqing L (2013) Hierarchical remote sensing image analysis via graph Laplacian energy. IEEE Geosci Rem Sens Lett 10(2):396–400

42. Boschetti F (2005) Dimensionality reduction and visualization of geoscientific images via locally linear embedding. Comput Geosci 31(6):689–697

43. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Sci Am Assoc Adv Sci 290(5500):2323–2326

44. de Ridder D, Kouropteva O, Okun O, Pietikainen M, Duin RPW (2003) Supervised locally linear embedding. Artificial neural networks and neural information processing, Springer, Berlin/Heidelberg, pp 333–341

45. Shi R, Shen I, Chen W (2005) Image denoising through locally linear embedding. Intl Conf Comput Graph Imag Visual:147–152

46. Zhang S-q (2009) Enhanced supervised locally linear embedding. Pattern Recog Lett 30(13):1208–1218

47. Chang C-I, Du Q (2004) Estimation of number of spectrally distinct signal sources in hyperspectral imagery. IEEE Trans Geosc Rem Sens 42(3):608–619

48. Jackson Q, Landgrebe DA (2001) An adaptive classifier design for high-dimensional data analysis with a limited training data set. IEEE Trans Geosc Rem Sens 39:2664–2679

49. Réjichi S, Chaabane F (2015) Satellite image time series classification and analysis using an adapted graph labeling. International workshop on the analysis of multitemporal remote sensing images (Multi-Temp), pp 1–4, July 2015

50. G´omez-Chova L, Camps-Valls G, Mu˜noz Mar´ı J, Calpe, J (2008) Semi-supervised image classification with Laplacian support vector machines. IEEE Geosc Rem Sens Lett 5:336–340

51. Camps-Valls G, Bandos T, Zhou D (2007) Semi-supervised graph-based hyperspectral image classification. IEEE Trans. Geosc. Rem. Sens. 45(10):3044–3054

52. Tuia D, Camps-Valls G (2009) Semi-supervised remote sensing image classification with cluster kernels. IEEE Geosci Remote Sens Lett 6(1):224–228

53. Bovolo F, Bruzzone L, Marconcini M (2008) A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. IEEE Trans. Geosc. Rem Sens 46(7):2070–2082

54. Bruzzone L, Chi M, Marconcini M (2006) A novel transductive SVM for semisupervised classification of remote sensing images. IEEE Trans Geosci Remote Sens 44(11):3363–3373

55. Chi M, Bruzzone L (2007) Semi-supervised classification of hyperspectral images by SVMs optimized in the primal. IEEE Trans Geosci Rem Sens 45(4):1870–1880

56. Bruzzone L, Chi M, Marconcini M (2006) A novel transductive SVM for semisupervised classification of remote-sensing images. IEEE Trans Geosci Remote Sens 44(11):3363–3373

57. Dundar M, Langrebe A (2004) A cost-effective semisupervised classifier approach with kernels. IEEE Trans Geosci Remote Sens 42(1):264–270

58. Gómez-Chova L, Camps-Valls G, Bruzzone L, Calpe-Maravilla J Semi-supervised remote sensing image classification based on clustering and the kernel mean map. In IEEE international geoscience and remote sensing symposium, IGARSS'08, Boston, USA, 2008

59. Saha S, Banerjee B, Merchant SN (2016) Unsupervised domain adaptation without source domain training samples: a maximum margin clustering based approach. In Proceedings of the Tenth Indian conference on computer vision, graphics and image processing, ICVGIP'16 ACM, pp 56:1–56:8, 2016

60. Bietti A (2012) Active learning for object detection on satellite images. Technical report, California Institute of Technology, Pasadena

61. Lu Q, Ma Y, Xia G-S (2017) Active learning for training sample selection in remote sensing image. Rem Sens Lett Taylor & Francis 8(12):1210–1219

62. Kao C-C, Lee T-Y, Sen P, Liu M-Y (2018) Localization-aware active learning for object detection. J CoRR, arXiv

63. Blaschko MB, Lampert CH (2008) Learning to localize objects with structured output regression. In: Forsyth D, Torr P, Zisserman A (eds) Computer vision: ECCV2008, pp 2–15, Springer

64. Tuia D, Kanevski M, Munoz Mari J, Camps-Valls G (2009) Structured SVM for remote sensing image classification. In IEEE workshop machine learning and signal processing, MLSP09, Grenoble, France

65. Qu Y et al (2017) Joint hierarchical category structure learning and large-scale image classification. IEEE Trans Image Process 26(9):4331–4346

66. Quan D, Wang S, Ning M, Xiong T, Jiao L (2016) Using deep neural networks for synthetic aperture radar image registration. Intl Geosci Rem Sens Symp 2016:2799–2802
67. Zhao J, Gong M, Liu J, Jiao L (2014) Deep learning to classify difference image for image change detection. International joint conference on neural networks (IJCNN), pp 411-417
68. Chan T, Jia K, Gao S, Lu J, Zeng Z, Ma Y (2015) PCANet: a simple deep learning baseline for image classification? IEEE Trans Image Process 24(12):5017–5032
69. Kuwata K, Shibasaki R (2015) Estimating crop yields with deep learning and remotely sensed data. IEEE Intl Geosci Rem Sens Symp (IGARSS):858–861
70. Zhu XX, Tuia D, Mou L, Xia G, Zhang L, Xu F, Fraundorfer F (2017) Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci Rem Sens Mag 5(4):8–36
71. Mou L, Zhu XX (2016) Spatiotemporal scene interpretation of space videos via deep neural network and tracklet analysis. IEEE Intl Geosci Rem Sens Symp (IGARSS) 2016:1823–1826
72. Marmanis D, Datcu M, Esch T, Stilla U (2016) Deep learning earth observation classification using ImageNet pre-trained networks. IEEE Geosci Rem Sens Lett 13(1):105–109
73. Ball A, Chan (2017) Online content provided with paper. A comprehensive survey of deep learning in remote sensing: theories, tools and challenges for the community.
74. Palafox LF, Hamilton, CW, Scheidt SP, Alvarez AM (2017) Automated detection of geological landforms on Mars using convolutional neural networks. Computers and Geosciences, Elsevier Limited, vol 101, pp 48–56, April 2017
75. Shi X, Chen Z, Wang H, Yeung D-Y, Wong W-K, Woo W-C (2015) Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In Cortes C, Lawrence ND, Lee DD, Sugiyama M, Garnett R (eds) Advances in neural information processing systems, vol 28. Curran Associates, Inc, pp 802–810. http://papers.nips.cc/paper/5955-convolutional-lstm-network-a-machine-learning-approach-for-precipitation-nowcasting.pdf
76. Lv Q, Dou Y, Niu X, Xu J, Li B (2014) Classification of land cover based on deep belief networks using polarimetric RADARSAT-2 data. IEEE Geosci Rem Sens Symp:4679–4682
77. https://deepsense.ai/deep-learning-for-satellite-imagery-via-image-segmentation/
78. https://www.techemergence.com/ai-applications-for-satellite-imagery-and-data/
79. https://rctom.hbs.org/submission/bean-counters-in-space-how-orbital-insights-sees-the-world/
80. http://spaceref.com/earth/space-imaging-helps-predict-crop-yields.html
81. https://www.gislounge.com/orbital-insight-exemplifies-how-artificial-intelligence-ai-can-be-used-to-study-images-from-space/

Chapter 12
A Wavelet Transform Applied Spectral Index for Effective Water Body Extraction from Moderate-Resolution Satellite Images

R. Jenice Aroma and Kumudha Raimond

Abstract The outbreak of pollution due to rapid urbanization has led to a huge threat unto the ecosystem where the water reserves are severely affected. In order to protect these coastal regions, drinking water resources, and artificial reservoirs, the satellite image-based monitoring models could be deployed. For developing such models using moderate-resolution data, the extraction of water body from the large land cover would be more complex. This limitation of effective water body extraction has been addressed in this paper through using the combination of wavelets and image transformation methods. A novel Wavelet-based Water Index (WaWI) has been proposed, and the results achieved are quantitatively assessed with both the spectral- and clustering-based water body extraction results such as Normalized Difference Moisture Index (NDMI), Modified Normalized Difference Moisture Index (MNDMI), and K-means clustering-based water body extraction using the Image Quality Assessment metrics like correlation coefficient, structural similarity index (SSIM), and Jaccard's similarity measures. The results achieved profoundly justifies the effectiveness of WaWI in water body extraction.

Keywords Water body · Spectral indices · NDMI · Wavelet · Segmentation

12.1 Introduction

The rapid growth of population density has led to increased urbanization that ruined the global ecosystem. The depletion factors like reduced rainfall, increasing air pollution, and degraded underground water resources have severely affected the life of human in earth. This increased population rate has also made an impact on the need for employment through industrialization of cultivable lands which

R. Jenice Aroma (✉) · K. Raimond
Karunya Institute of Technology & Sciences, Coimbatore, India

has polluted the water resources nearby through release of toxin substitutes from industries. This state of urbanization without any future perspective has led to tremendous loss of natural wealth gifted and raises the need for developing stand-alone models for monitoring the changes in our ecosystem.

In olden days, such monitoring applications are a distant dream but with the advent of remote sensing activities, which is nothing but a field of tracing any object or a natural phenomenon through sensing either the own or artificially emitted electromagnetic radiation reflected from those objects, without any direct physical contact. It can be carried out through either passive or active mode of remote sensing.

In case of performing active remote sensing, the artificial sensors generate those signals and its corresponding reflected radiations are observed. But on using passive mode of remote sensing, the naturally available radiations due to sun illumination are observed [1]. These observed signals that are received at the ground stations from those distant objects get digitized as satellite images. Generally, due to the lack of sun illumination, the passive sensors are limited on night sensing. But this limitation is not a par with active sensors which can acquire images throughout the day and also during cloudy atmospheric conditions. Thus, satellite images are grouped into four different types such as spectral, spatial, temporal, and radiometric resolution. The surface area covered within every image pixel is termed as spatial resolution of a satellite image. The temporal resolution is termed as the time period between the image acquisition of individual satellites. This temporal availability of satellite images based on their acquisition time period is a major differentiating factor in case of change mapping studies. The spectral resolution is termed for the region of electromagnetic spectrum observed in individual spectral bands of any satellite image. The total number of spectral bands and the narrowness of those bands determine the type of spectral resolution in a satellite image. Probably, the multispectral resolution images have 3–10 spectral bands, whereas the hyperspectral resolution data could have thousands of bands which could hold more specific details of the subjective phenomenon. And the radiometric resolution of a satellite image is nothing but the smallest energy differences observed from the electromagnetic reflectance [2].

The above said differential factors play a major role in satellite image acquisition for spatial-based applications like climate assessments, natural hazard post case studies, and environmental applications like weather forecast, rainfall prediction, and water body monitoring [3]. The technological outgrowth nowadays has led to an advent increase in revolutionized information technology that leads to developing unmanned aerial vehicles and spy robots. These stand-alone intelligent models could be a great assistive device in military and defense [4].

Other than this, the advanced spatial data-assisted models such as agricultural crop health tracking applications, terrain modeling, forest fire prediction models, and many surveillance-based applications such as traffic monitoring, vehicle navigation, catastrophic modeling, urban land usage studies, and other such smart development projects are also using satellite images [5, 6]. Even the simulation

environment for training the intelligent spy agent models virtually is also developed using satellite images.

Other than moderate spatial resolution images, the use of point cloud data such as RAdio Detection And Ranging (RADAR) and Synthetic Aperture Radar (SAR) data also could favor better mineral and underground water exploration activities [7, 8]. The terrain modeling applications are based on such point data for three-dimensional modeling.

These latest advancements in geospatial technologies favor deployment of effective surveillance and change monitoring applications [9, 10]. In a moderate-resolution satellite image, the land cover constitutes different spatial objects, and the machine learning techniques with high reasoning capability are used for better discrimination. However, these techniques suffer with rigorous training and longer computation time due to huge dimension of data; hence, land cover classification has been attempted with smaller subregions of images [11].

In order to locate or detect the water region from satellite images, different models based on contours and clusters such as Self-Organizing Maps (SOM), k-means, and Fuzzy C-Means (FCM) clustering are widely used for analyzing satellite images [12, 13]. The SOM is the widely used unsupervised method for classifying land cover features in satellite images. A combination of wavelets to acquire reduced levels of information using Discrete Wavelet Transform (DWT) and SOM for water region extraction has been performed [14]. Similarly, the bimodal distributions of sparseness histograms which are also a compact representation of data have been derived from the Airborne Laser Swath Mapping (ALSM) data to detect the presence of water features and followed by clustering to extract the water bodies [15]. The use of unsupervised methods than supervised methods is preferred due to the unavailability of ground truth data.

The above said complexities of satellite image analysis such as huge data dimensions and unavailability of ground truth labels are the major limitations in satellite image analysis. But in case of spectral information-based image transformation, a simple image ratioing approach over the surface reflectance of an atmospherically corrected image could extract the appropriate region from the images. It leads to the development of spectral indices for tracing the climatic change events [16, 17]. The specific band properties of a satellite image which may vary in different sensor instruments according to their respective wavelengths can be studied and combined for extracting the exact land cover features through image transformation [18].

The Normalized Difference Vegetation Index (NDVI) is the most popular spectral index which is used for tracing the vegetation health and to build drought prediction models [19]. The Normalized Difference Moisture Index (NDMI), Modified Normalized Difference Moisture Index (MNDMI), and Automated Water Extraction Index (AWEI) are the notable spectral indices that are widely used for water resources mapping where NDMI is represented as Normalized Difference Water Index (NDWI) in few research works [20]. These water indices apply different band combinations based on their spectral reflectance properties to achieve the accurate water body extraction [21]. However, these spectral indices may face difficulties in water region extraction due to the mixed water pixels in very low-resolution satellite

images [22]. Many fusion-based schemes are available to fuse high- and moderate-resolution image bands for increasing the specific resolution of image products. The panchromatic band fusion methods can also be availed for overcoming these limitations [23, 24].

In general, the human eyes are capable of detecting even very small changes in brightness and color information of a visual scene, so that even smaller fragments of detail also can be more apparent. Similarly, wavelet-based analysis which can decompose the image details into varied levels of information can be opted for attempting to acquire better vision capability in machines [25]. Wavelet transform has also been used for reducing the redundant details in an image efficiently to improve compression schemes in satellite images [26]. In coastal monitoring, the changes in the sea due to ships sailing such as oil spills and ship wakes have been traced using the wavelet-based analysis. The use of RADAR data plays a vital role in tracing these mesoscale features related to coastal zone monitoring [27].

This paper aims at bridging the gap of existing water extraction methods based on clustering by applying the proposed wavelet-based image transformation method. The following section details the workflow of the proposed wavelet-based model for water extraction results. Further, the quantitative evaluation done for justifying the correctness of the proposed model has been carried out through comparing the results of NDMI, MNDMI, and K-means clustering-based water body extraction using generic image similarity measures. Also, the outcome and benefits of the proposed model with possible extensions are portrayed in the next preceding sections.

12.2   Wavelet-Based Water Extraction Method

The proposed model for simple water body extraction from satellite images has been designed using the multi-temporal Landsat 8 images of two different lakes which have been acquired from USGS Earth explorer archive [28]. The Landsat series is a joint venture of NASA and USGS which is the largest provider of moderate-resolution images in the recent days. These datasets are free for the registered public users and academicians for carrying out research. The different versions of satellite images are Landsat 5, 7, and 8. These are the most widely used moderate-resolution images. Nowadays, the satellite image-based models are applied for numerous research activities right from land cover, forest, and agricultural applications to even surveillance tasks.

In general, the Landsat 7 satellite images are available with two variants as 30 m spatial resolution for multispectral data and 15 m spatial resolution for panchromatic mode (PAN) images. The spectral band details of Landsat 5 and the other Landsat 7 and 8 are different. Similarly, the spectral and spatial details of satellite images from different satellite vendors are also different and certain prominent data satellites are shown in the following Table 12.1.

Table 12.1 Multispectral satellite data [18]

Satellite	Sensor	Spatial resolution	Revisit capability
Landsat 7	Enhanced Thematic Mapper (ETM+)	30 m (multispectral) and 15 m (PAN)	Every 16 days
SPOT 5	Panchromatic multispectral	5 m (PAN) and 10 m in multispectral	11 times on every 26 days
Resourcesat	LISS IV	5.8 m (multispectral)	5 days
RADARSAT-2	Ultrafine/Quad-Pol	3 m (ultrafine) and 25 m (Quad-Pol)	Every few days
QUICKBIRD	Panchromatic and multispectral	0.6 m (PAN) and 2.4 m (multispectral)	1.5–3 days
IKONOS	Panchromatic and multispectral	1 m (PAN) and 4 m (multispectral)	1.5–3 days

The two notable Indian lakes such as Sambhar Lake in Rajasthan (26°58′N 75°05′E) and Govind Ballabh Pant Sagar (GBPS) in Uttar Pradesh (24°12′9″N 83°0′29″E) are the study areas chosen. A multi-temporal dataset of Landsat 8 images comprising four different dates has been chosen in both the study regions for evaluation. In order to check the prominence of the model, (a) Sambhar Lake dataset that includes data during four different months of two different years, 2014 and 2015, and (b) GBPS lake dataset that includes data considering four different months in 2015 are used. The Sambhar Lake is the largest inland salt lake which holds the highest salt production of India, next to Lake Chilika in the east coast [29]. The GBPS is an artificial lake which is surrounded by many thermal power stations due to the availability of huge coal deposits. Hence, Singrauli district, where GBPS Lake is nearby, got named as the energy capital of India with more than 10,000 MW of electricity production [30].

The Sambhar Lake is located in a highly drought-prone regions of India as it lies in the northwestern part of India, nearer to Thar Desert of Rajasthan which records the hot climate; hence, it is suitable for tracking the changes in water body on multiple dates. The GBPS Lake is otherwise termed as Rihand Reservoir which is more vulnerable to contamination due to the release of toxins from chemical industries and in specific for the cooling purpose of thermal power plants around.

The individual satellite image bands are generally not good for visualization; hence, a false color composite (FCC) which is a combination of near-infrared (NIR), green, and blue spectral bands of Landsat 8 image acquired from USGS has been created using QGIS for Sambhar Lake and GBPS Lake region as shown in Figs. 12.1a, 12.1b, and 12.1c, respectively. In FCC, the land cover regions are in red and water bodies are in green color. The absence of water depicted as brown color in the salt extraction zone during October 2014 is clearly seen by comparing both FCC composites of Sambhar Lake. Since it is not a seasonal activity, the evaporation of water for precipitation is carried out throughout the year. In Fig. 12.1c, the circled regions depict the presence of thermal stations around GBPS Lake which lead to heavy water pollution.

Fig. 12.1a FCC of Sambhar Lake on Oct 2014-Landsat 8

Fig. 12.1b FCC of Sambhar Lake on Oct 2015-Landsat 8

Every Landsat 8 image comprises 11 bands with different spectral reflectance properties, and the details have been shown in Table 12.2, which can be useful in different applications like cloud masking, healthy vegetation monitoring, and other environmental monitoring [31].

The focus of the proposed model is to apply wavelet-based image transformation to extract only the water region from the satellite images. Hence, the bands 3 and 5 which are green and NIR bands, respectively, have been selected for tracing the water region since these bands combination have been proved to be more efficient for water region extraction in many image transformation-based approaches [32].

The workflow of the proposed model has been shown in the following Figs. 12.2a and 12.2b.

Fig. 12.1c FCC of GBPS Lake from 2015-Landsat 8

Table 12.2 Landsat 8 band details

Bands	Wavelength(μm)	Spatial resolution(m)
Coastal – Band 1	0.43–0.45	30
Blue – Band 2	0.45–0.51	30
Green – Band 3	0.53–0.59	30
Red – Band 4	0.64–0.67	30
[a]NIR – Band 5	0.85–0.88	30
[a]SWIR 1 – Band 6	1.57–1.65	30
SWIR 2 – Band 7	2.11–2.29	30
[a]PAN – Band 8	0.50–0.68	15
Cirrus – Band 9	1.36–1.38	30
[a]TIRS 1 – Band 10	10.6–11.19	100
TIRS 2 – Band 11	11.5–12.51	100

[a]where *NIR* near infrared, *SWIR* shortwave infrared, *PAN* panchromatic, and *TIRS* thermal infrared sensor

The Landsat 8 images available in USGS archive are orthorectified. However, the atmospheric variations due to different illumination factors and climate changes can affect the quality of an image. The chosen multi-temporal dataset has been selected without huge cloud coverage through cloud screening. Then these images are atmospherically corrected and converted into surface reflectance (SR) values [33]. The Dark Object Subtraction method (DOS1) is the widely used atmospheric correction method for converting the digital numbers of raw satellite images into SR products [34]. Here, the chosen raw Landsat 8 images are atmospherically corrected using DOS1 method in Quantum GIS. Then, Stationary Wavelet Transform (SWT) which is an undecimated wavelet transform has been applied on the chosen SR bands to acquire the water-concentrated pixels. In human eyes, a lossless
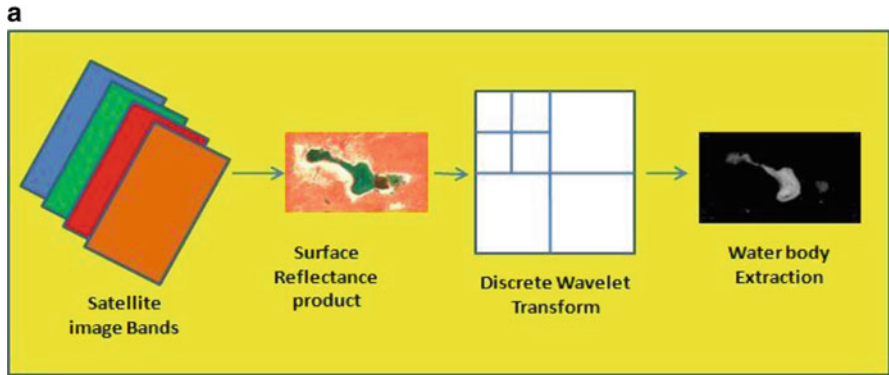
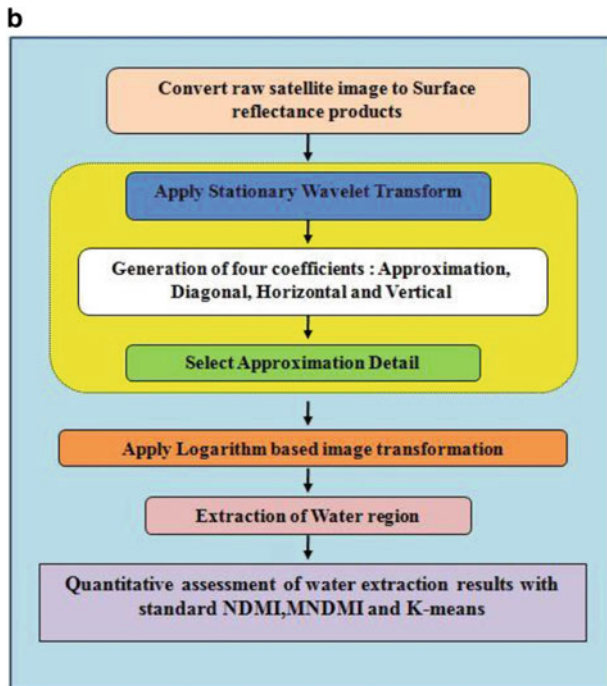**Fig. 12.2a** Wavelet-based water region extraction model

Fig. 12.2b Phases of wavelet-based water extraction method

compression of visual scenes will take place before being sent to the visual cortex for processing [35]. Similarly, to acquire lossless detail from the satellite images, SWT which does not include the down sampling as in DWT has been chosen [36]. The SWT holds shift invariance capability which could ensure better spatial localization capability that favor its usage even in structural damage detection models for composite structures [37].

On applying this SWT method using Daubechies wavelet (db1) which is similar to that of Haar wavelet, the satellite image will undergo upsampling to generate four components such as approximation, horizontal, vertical, and diagonal details. Among them, the approximation detail of an image alone has been chosen for further processing since it holds the abstract detail of the respective image than the other three components. Here, the choice of spectral bands and image ratioing approach for WaWI has been derived from the standard NDMI method for water body extraction as in Eq. (12.1) [38]:

$$\text{NDMI} = \frac{(b_3 - b_5)}{(b_3 + b_5)} \qquad (12.1)$$

where b_3 and b_5 represent the SR values of green and NIR bands, respectively.

The proposed WaWI index performs image ratioing as in Eq. (12.2) by applying natural logarithm on the acquired approximation coefficients of both the spectral bands from a chosen image. The logarithm is a monotonically increasing function that could gain the maximum values without any alteration of the underlying values [39]:

$$\text{WaWI} = \frac{(\log(A_{b5}) - \log(A_{b3}))}{(\log(A_{b5}) + \log(A_{b3}))} \qquad (12.2)$$

where A_{b5} and A_{b3} represent wavelet approximation coefficients of the SR values of green and NIR bands, respectively. In order to differentiate the spectral band detail and wavelet approximation detail of the spectral band, A_{b5} and A_{b3} have been used instead of b_5 and b_3. (Unlike the NDMI ratioing, here NIR and green bands have been relocated in ratioing positions for normalization.) As stated earlier, here the choice of SWT using Haar wavelet over DWT is due to improved shift invariance capability that could offer better spatial localization.

12.3 Experimental Results and Discussion

This spectrally segmented water body is then subjected to change differencing using absolute and mean differencing as shown in Figs. 12.3a and 12.3b. The Sambhar Lake region is a center for salt production throughout the year where the optimum level of water and extreme hot weather induces the salt production. The lake water is allowed to spread across the salt pans for being evaporated and then crystalline gets collected. Hence, there is a zone of suspended salts without water in these salt pan regions of the lake during crystal formation [40]. It is evident from the extracted water bodies of Sambhar Lake from Landsat 8 image on October 2014 using NDMI, MNDMI, WaWI, and K-means-based extraction methods which has been portrayed in the following Figs. 12.3a, 12.3b, 12.3c, and 12.3d that the waterless region of
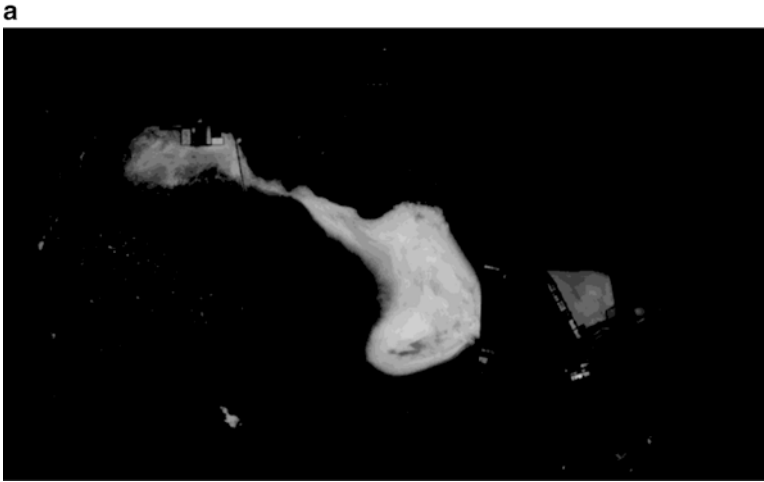
Fig. 12.3a WaWI extraction of water body (Landsat 8-Oct 2014)
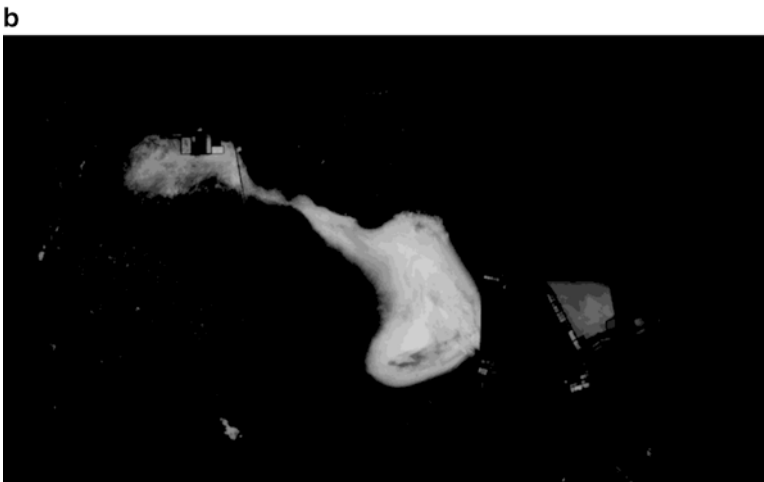
Fig. 12.3b NDMI extraction of water body (Landsat 8-Oct 2014)

salt pans is not visible. Similarly, water in salt pans during October 2015 has been extracted as shown in Figs. 12.3e, 12.3f, 12.3g, and 12.3h. The FCC of Sambhar Lake as shown in Figs. 12.1a and 12.1b in the above section can be referred for visualizing the water region in green color. The accurate extraction of water body in both of these years can be clearly seen in these figures which accounts for extreme drought pattern in the lake, where WaWI can be used for effective water body extraction.
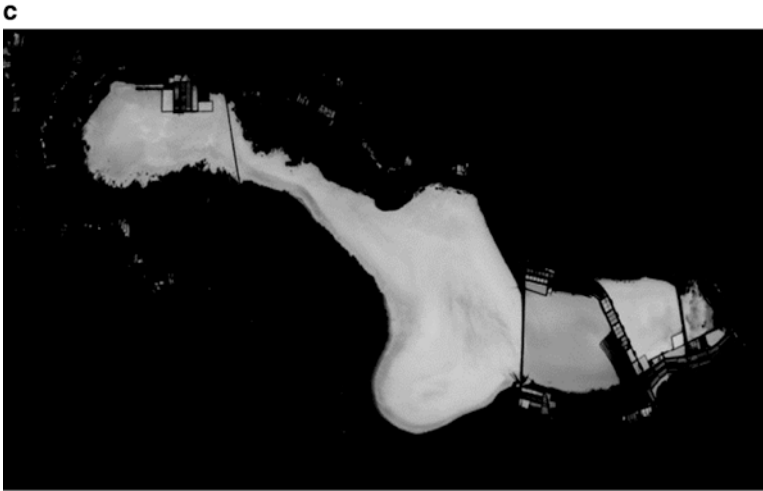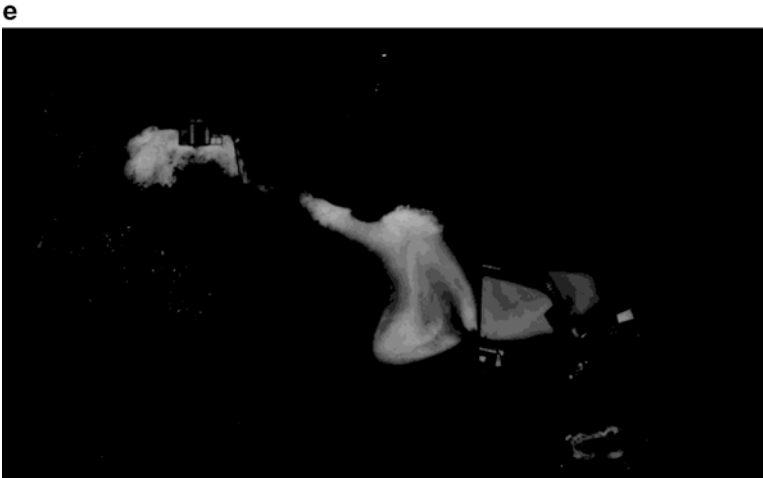
c

Fig. 12.3c MNDMI extraction of water body (Landsat 8-Oct 2014)

d

Fig. 12.3d K-means extraction of water body (Landsat 8-Oct 2014)

Similarly, the water body of GBPS Lake also can be clearly examined in the following Figs. 12.4a, 12.4b, 12.4c, and 12.4d where both the WaWI- and NDMI-based water extraction can be compared. Though many recent studies have led to an awareness on implementing proper check measures, more suitable steps must be taken to monitor the lake pollution [41].

Also, from the above water extraction results, the water contaminants are made visible near the coastlines where the thermal power stations are situated. These extracted water features can be examined in further along with the use of in situ measurements of water concentration samples to monitor the contamination level of

e

Fig. 12.3e WaWI extraction of water body (Landsat 8-Oct 2015)

f

Fig. 12.3f NDMI extraction of water body (Landsat 8-Oct 2015)

lake due to industrialization where the WaWI index can be used for effective water body extraction.

The proposed WaWI extracted the water region clearly from the water body as that of the exact water representation observed in FCC visualization. It also has effectively overcome the computational issues of traditional classifier-based water segmentation methods. In 2015, a comprehensive review has been done for ana-
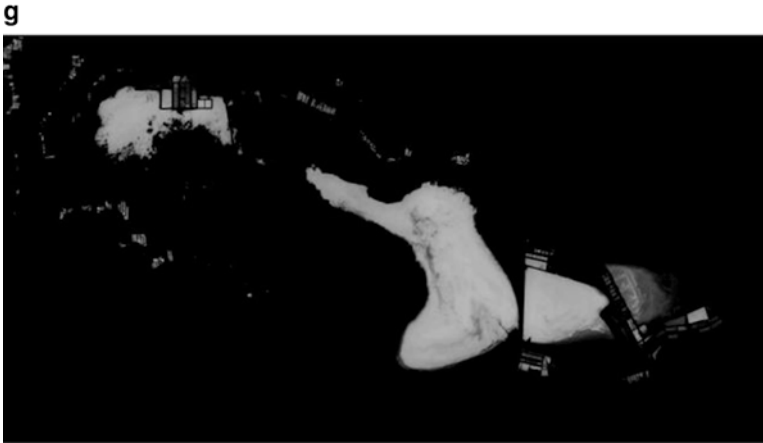
Fig. 12.3g   MNDMI extraction of water body (Landsat 8-Oct 2015)

Fig. 12.3h   K-means extraction of water body (Landsat 8-Oct 2015)

lyzing the various image transformation methods for delineating the water bodies from satellite images. The study includes different satellite image datasets like World View, Landsat, MODIS, and ASTER [42]. Thus, using image transformation methods for spectral-based spatial object detection along with the advanced soft classifiers for automated change detection can be a major boon to environmental monitoring applications. The achieved results have been clearly found to be similar in both the mentioned methods, through qualitative analysis on visual comparison with NDMI and MNDMI Water extraction results.
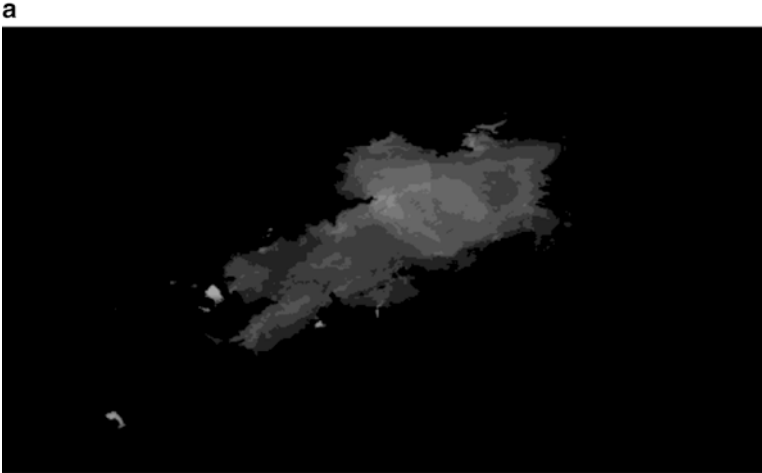
Fig. 12.4a WaWI extraction of GBPS Lake (MAY 07, 2015)
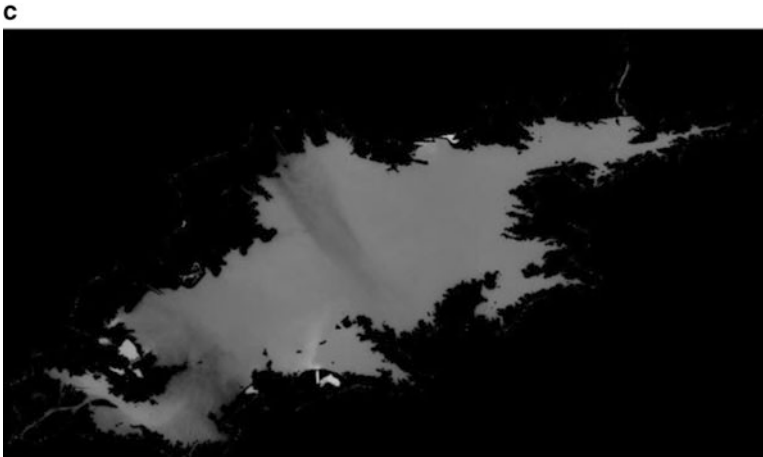

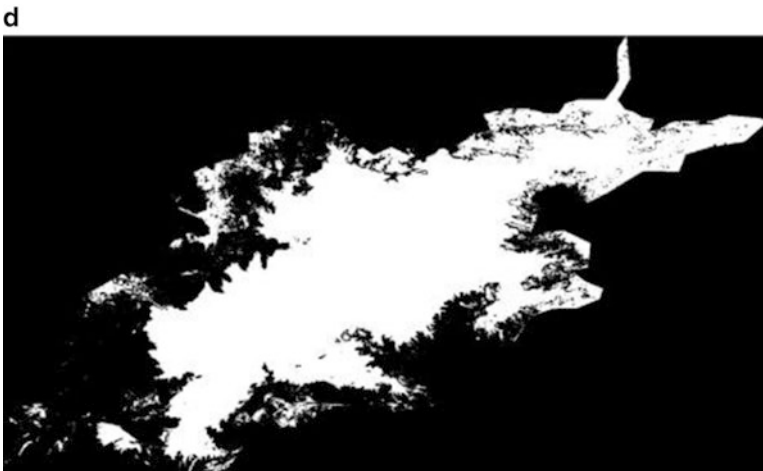**Fig. 12.4b** NDMI extraction of GBPS Lake (MAY 07, 2015)

The MNDMI is also a water extraction index similar as NDMI, but image ratioing has been applied with green and shortwave infrared (SWIR) band rather than NIR band as in Eq. (12.3)

$$\text{MNDMI} = \frac{(b_3 - b_6)}{(b_3 + b_6)} \tag{12.3}$$

where b_3 and b_6 represent the SR values of green and SWIR bands, respectively.

Fig. 12.4c MNDMI extraction of GBPS Lake (MAY 07, 2015)

Fig. 12.4d K-means extraction of GBPS Lake (MAY 07, 2015)

Hanqiu Xu in 2006 has compared both MNDMI and NDMI water extraction results and found that MNDMI stands good in water extraction for a land cover having numerous built-up areas. It is due to the fact that SWIR band is good at observing the SR values of entire coastal region of a water body and not only the pure water pixels [47]. Hence, even the region with little marsh can also be treated as the presence of water which could lead to misclassification of water available in a water body. But NDMI based on NIR spectral band has the spectral reflectance value of pure water pixels ranging from 0 to 1 where other spatial features range

between -1 and 0 which could offer better water extraction even in the case of saline or brackish water bodies and mixed vegetation [48].

For example, in Bihar, Chowdary et al. have performed a study to detect the water-logged areas using NDWI (otherwise termed for NDMI) values where the deep water bodies are traced with NDWI value of $+1$ and vegetation with value of -1 [49]. Similarly, the soft clustering-based water body extraction also admits misclassification of marshy region to be as water pixels. On applying simple K-means clustering for water body extraction in Sambhar Lake, the similar pattern of water body extraction as MNDMI with misclassification of brackish region as water pixels has been obtained. It is a widely used clustering method which is based on the clusters computed upon the closest centroid. The computation ends with user-given cluster number 'k' [50]. Here, on trying different cluster numbers, the water body extraction using cluster number 4 is found to be good.

Thus, on comparing the water extraction results of K-means clustering, MNDMI, NDMI, and proposed WaWI through qualitative assessment, the pure water pixels extraction is evident only in NDMI and WaWI. Further to ensure the performance of water body extraction, a quantitative assessment has been done using standard image similarity metrics which is shown in the following section.

12.3.1 Performance Evaluation

In order to evaluate the performance of the achieved water extraction results using NDMI and proposed WaWI, a quantitative analysis has been carried out using effective image similarity measuring metrics like correlation coefficient, SSIM, and Jaccard's similarity coefficient. The correlation coefficient (r) given in Eq. (12.4) [43] is the widely accepted metric for measuring the relative similarity even while comparing the results of different spectral indices. In the earlier days, when image transformation indices are being derived from Landsat 5 data, the emerging image ratioing indices are quantitatively assessed using correlation coefficient [44]. Here, the achieved result needs to be similar to that of the standardized NDMI method for water extraction.

$$r = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A} \right) \left(B_{mn} - \overline{B} \right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \overline{A} \right)^2 \right) \sum_m \sum_n \left(B_{mn} - \overline{B} \right)^2 \right)}} \tag{12.4}$$

where A and B be two different images with $\overline{A}$ and $\overline{B}$ as their respective mean values. Here, m and n denote the image pixel locations.

Jaccard's coefficient is defined to be the ratio of the intersection and union of the two entities chosen for similarity comparison [45]. Here, two different images to find the similarity of water extraction using WaWI and NDMI have been compared

and validated using Eq. (12.5). If the similarity ratio is 1, then these outcomes are more similar:

$$\text{Jaccard Coefficient} = \frac{|A \cap B|}{|A \cup B|} \tag{12.5}$$

The standard SSIM function as shown in Eq. (12.6) has also been applied for examination [46]:

$$\text{SSIM}\,(x,\,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{12.6}$$

where μ_x, μ_y, σ_x, σ_y, and σ_{xy} are the local means, standard deviations, and cross-covariance for the ground truth and segmentation images $x,\, y,$ and two values of c be the constant.

The results for similarity measures for both the study areas have been illustrated in the following Tables 12.3 and 12.4.

However, in order to prove the effectiveness of the achieved results, the water region that has been extracted using MNDMI and K-means is compared with WaWI results as in Table 12.5.

Table 12.3 NDMI vs. WaWI – water body extraction of Sambhar Lake

NDMI vs. WaWI	Correlation coefficient (r)	SSIM	Jaccard index
05/Oct/2014	0.9957	0.9846	0.9955
21/Oct/2014	0.9943	0.9866	0.9956
06/Nov/2014	0.9957	0.9907	0.9967
08/Oct/2015	0.9912	0.9856	0.9952

Table 12.4 NDMI vs. WaWI – water body extraction of GBPS Lake

NDMI vs. WaWI	Correlation coefficient (r)	SSIM	Jaccard index
07/May/2015	0.9843	0.9785	0.9959
23/May/2015	0.9773	0.9721	0.9953
15/Nov/2015	0.9929	0.9736	0.9922
01/Dec/2015	0.9922	0.9715	0.9925

Table 12.5 Quantitative assessment for Sambhar Lake data on Oct 2014 (Landsat 8) – WaWI with NDMI, MNDMI, and K-means water body extraction results

Water body extraction in Oct 2014 image	SSIM	Correlation coefficient (r)	Jaccard index
WaWI vs. NDMI	0.9846	0.9957	0.9955
WaWI vs. K-means	0.7576	0.6361	0.8859
WaWI vs. MNDMI	0.8568	0.8172	0.9246

Thus, from the above result of WaWI in both visual and quantitative assessment, it has been clearly inferred that effective water body extraction can be achieved using proposed WaWI method than other existing approaches for segmentation.

12.4   Conclusion

The importance of water body extraction is made evident with the development of successful change monitoring applications for effectively tracing the depletion of water resources. The achieved water extraction results can be useful in environmental monitoring applications that can assist the government bodies to apply measures for preventing drought and any pollution effects. The proposed wavelet-based water extraction model using WaWI has been successfully compared with the more standard NDMI, MNDMI, and K-means clustering results. It has been inferred from the above comparison that WaWI-based results outstand the rest. In order to extend this model, the choice of wavelets to prevent loss of detail on image ratioing and the improved shift invariance capability for change detection models from extracted regions for any water resources can be explored for better results.

References

1. Aggarwal S (2003) Principles of remote sensing, satellite remote sensing and GIS applications in agricultural meteorology. In: Proceedings of the training workshop, pp 23–38
2. Kumar M, Singh RK (2013) Digital image processing of remotely sensed satellite images for information extraction, conference on advances in communication and control systems (CAC2S 2013)
3. Sakamotoa T, Gitelsonb AA, Nguy-Robertsonb AL, Arkebauerc TJ, Wardlowb BD, Suykerb AE, Vermab SB, Shibayamaa M (2012) An alternative method using digital cameras for continuous monitoring of crop status. Agric For Meteorol 154-155:113–126
4. Fisher JRB, Acosta EA, Dennedy-Frank PJ, Kroeger T, Boucher TM (2018) Impact of satellite imagery spatial resolution on land use classification accuracy and modeled water quality. Remote Sens Ecol Conserv 4(2):137–149
5. Khan MA, Ectors W, Bellemans T, Janssens D, Wets G (2017) UAV-based traffic analysis: a universal guiding framework based on literature survey. Trans Res Procedia 22:541–550
6. Valasek J, Henrickson JV, Bowden E, Shi Y, Morgan CLS, Neely HL (2016) Multispectral and DSLR sensors for assessing crop stress in corn and cotton using fixed-wing unmanned air systems. Int Soc Opt Photon
7. Kodikara GRL, Woldai T, Ruitenbeek F JAv, Kuria Z, van der Meer F, Shepherd KD, Van Hummel GJ (2011) Hyperspectral remote sensing of evaporate minerals and associated sediments in Lake Magadi area, Kenya. Int J Appl Earth Obs Geoinf 14:22–32
8. Elbeih SF (2014) An overview of integrated remote sensing and GIS for ground water mapping in Egypt. Ain Shans Eng J 6(1):1–15
9. Baboo SS, Shereef IK (2010) An efficient weather forecasting system using artificial neural network. IJESD 1(4):321–326
10. Samadzadegan F, Haahn M, Saeedi S (2007) Position estimation of aerial vehicle based on a vision aided navigation system, In: Proceedings of visualization and exploration of geospatial data-Stuttgart

11. Behery GM (2013) An automatic neural networks system for classifying dust, clouds, water and vegetation from Red Sea area. IJAE 4(2):27–44
12. Praveena S (2016) Clustering and classification of images using ABCFCM and naive bayes classifier. IJSR 5(6):199
13. Prasad SVS, Satya Savithri T, Iyyanki V, Krishna M (2013) Object classification of satellite images using cluster repulsion based kernel Fcm and Svm classifier. IOSR-JECE 7:25–35
14. Janahiraman TV, Kong W (2011) SOM based segmentation method to identify water region in LANDSAT images. IJECCT 2(1):13
15. Yuan X, Sarma V (2011) Automatic urban water-body detection and segmentation from sparse ALSM data using spatially constrained model driven clustering. IEEE Geosci Remote Sens Lett 8
16. Siva Sankari T, Chandramouli AR, Gokul K, Mangala Surya SS, Saravanavel J (2015) Coastal vulnerability mapping using geospatial technologies in Cuddalore-Pichavaram Coastal Tract, Tamil Nadu, India. Aquat Procedia 4:412–418
17. Ko BC, Kim HH, Nam JY (2015) Classification of potential water bodies using Landsat 8 OLI and a combination of two boosted random forest classifiers. Sensors 15(6):13763–13777
18. Jenice Aroma R, Raimond K (2015) A review on availability of remote sensing data. In: IEEE international conference on technological innovations in ICT for Agriculture and Rural Development (TIAR), pp 150–155
19. Patel JH, Oza MP (2014) Deriving crop calendar using NDVI time series. Int Arch Photogram Rem Sens Spatial Inform Sci XL-8:869–873
20. Du Y, Zhang Y, Ling F, Wang Q, Li W, Li X (2016) Water Bodies' mapping from Sentinel-2 imagery with modified normalized water index at 10-m spatial resolution produced by sharpening the SWIR band. Remote Sens 354(8):1–19
21. Rokni K, Ahmad A, Selamat A, Hazini S (2014) Water feature extraction and change detection using multitemporal Landsat imagery. Remote Sens 6:4173–4189
22. Jiang H, Feng M, Zhu Y, Lu N, Huang J, Xiao T (2014) An automated method for extracting Rivers and lakes from landsat imagery. Remote Sens 6:5067–5089
23. Nikolakopoulos KG (2008) Comparison of nine fusion techniques for very high resolution data. Photogram Eng Remote Sens 74:647–659
24. Du P, Liu S, Xia J, Zhao Y Information fusion techniques for change detection from multitemporal remote sensing images. Inf Fus 14:19–27
25. Dalavi SJ, Patil MS, Patil SR (2014) Content based image retrieval using Daubechies wavelet transform. Int J Innov Res Sci Eng Technol 3
26. Meenakshi A, Hariharan M (2013) Wavelet bi-orthogonal filter coefficient based image compression. IJSET, p 2
27. Liu AK, Peng CY, Chang SY-S (1997) Wavelet analysis of satellite images for coastal watch. IEEE J Ocean Eng 22:9–17
28. USGS Earth explorer (2016). Available at: http://earthexplorer.usgs.gov/. Accessed September
29. Bhat AH, Sharma KC, Banday UJ (2015) Impact of climatic variability on salt production in sambar lake, a Ramsar wetland of Rajasthan, India. Middle East J Sci Res 23(9):2060–2065
30. Pankaj S, Prabhakar K, Ananad B, Sameer V, Kaushal M, Mahure NV, Ratnam M (2012) Monitoring the contamination in the Gobind Ballabh Pant Sagar, Rihand, U.P., India. IJES 3(1):487–497
31. Landsat 8-Fact sheet (2013) U.S. Geological Survey, August 2013
32. Landsat 8 Band details (2016). Available at: https://blogs.esri.com/esri/arcgis/2013/07/24/band-combinations-for-landsat-8/. Accessed Sept 2016
33. Mahiny ARS, Turner BJ (2007) A comparison of four common atmospheric correction methods. Photogram Eng Remote Sens 73:361–368
34. Makarau A, Richter R, Müller R, Reinartz P (2014) Haze detection and removal in remotely sensed multispectral imagery. IEEE Trans Geosci Remote Sens
35. Mittal A, Moorthy AK, Bovik AC (2012) Visually lossless H.264 compression of natural videos. Comput J 56(5):617–627

36. Kumar BS, Nagaraj S (2013) Discrete and stationary wavelet decomposition for image resolution enhancement. IJETT 4(7):16
37. Arivazhagan S, Newlin Shebiah R, Salome Magdalene J, Sushmitha G (2015) Railway track derailment inspection system using segmentation based fractal texture analysis. IJIVP 6:1060–1065
38. Sahu AS (2014) Identification and mapping of the water-logged areas in Purba Medinipur part of Keleghai river basin, India: RS and GIS methods. Int J Adv Geosci 2:59–65
39. Importance of Logarithm (2018). Available at: https://en.wikipedia.org/wiki/Likelihood_function. Accessed Sept 2018
40. Singh BP, Singh N, Singh SP (2013) Modern salt (halite) deposits of the Sambhar lake, Rajasthan and their formative conditions, scientific correspondence. Curr Sci 104(11):1482–1484
41. Prabhakar K, Anand B, Sharma SN, Sharma P, Gupta SL (2014) Influence of a chemical industry effluent on water quality of Gobind Ballabh pant Sagar – a long term study. IJESI 3:07–13
42. Jawak SD, Kulkarni K, Luis AJ (2015) A review on extraction of lakes from remotely sensed optical data with a special focus on Cryospheric Lakes. Adv Remote Sens 4:196–213
43. Correlation Coefficient (2018). Available at: http://www.stat.wmich.edu/s216/book/node122.html. Accessed Sept 2018
44. Lautenschlager LF, Perry Jr CR Comparison of vegetation indices based on satellite-acquired spectral data, technical document, U.S Department of Agriculture, Stat Rep Serv
45. Jaccard Index. http://en.wikipedia.org/wili/ard_index
46. Zhou W, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13:600–612
47. Xu H (2006) Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int J Remote Sens 27:3025–3033
48. Sahu AS (2014) A study on Moyna basin water-logged areas (India) using remote sensing and GIS methods and their contemporary economic significance. Geogr J 2014:9
49. Chowdary VM, Chandran RV, Neeti N et al (2008) Assessment of surface and sub-surface water logged areas in irrigation command areas of Bihar state using remote sensing and GIS. Agric Water Manag 95:754–766
50. Kanungo T, Mount DM, Netanyahu NS, Piatko CD, Silverman R, Wu AY (2002) An efficient k-means clustering algorithm: analysis and implementation. IEEE Trans Pattern Anal Mach Intell 24(7)