RESEARCH ARTICLE

Comparative Analysis and Implication of Hyperion Hyperspectral and Landsat-8 Multispectral Dataset in Land Classification

Neelam Dahiya[1] · Sartajvir Singh[2] · Sheifali Gupta[1]

Abstract

Remote sensing via hyperspectral imaging delivers the crucial earth's surface information in narrow spectral bands, which may not be possible with multispectral imaging. The classification algorithms play a vital role in highlighting or categorizing the essential features of the earth's surface with respect to spectral information and generate thematic maps for further processing in different applications. Therefore, it is essential to explore the impact of well-defined or emerging classifiers on hyperspectral and multispectral datasets. In the present work, the performance of various classifiers, i.e., support vector machine (SVM), feedforward neural networks (FF-NN) and maximum likelihood classifier (MLC), has been evaluated using Earth Observation-1 (EO-1) Hyperion and Landsat-8 Operational Land Imager and Thermal Infrared Sensor over a part of the North Indian states. The experimental outcomes have confirmed that the FF-NN classifier achieved higher accuracy (91.20% with Hyperion; 82% with Landsat-8) as compared to other classification methods, i.e., SVM (87.60% with Hyperion and 80% with Landsat-8) and MLC (84.40% with Hyperion and 72.40% with Landsat-8). This study is important in terms of exploring the potential of hyperspectral imaging with different classification algorithms in various emerging applications.

Keywords Hyperion · Landsat-8 · Feedforward neural network (FF-NN) · Support vector machine (SVM) · Maximum likelihood classification (MLC)

Introduction

Remote sensing is the art of collecting the earth's surface information with the help of aircraft or space-borne satellites to be utilized in applications (Sharma et al., 2013), e.g., oceanography, cryosphere, hydrology, agriculture, and weather monitoring service but not limited (Sood et al., 2021a, 2021b). In the land-use and land-cover (LULC) applications, it plays a vital role in the estimation of soil moisture and erosion, forest cover mapping, urban planning, crop yield monitoring and prediction, and management of natural resources (Bhosle et al., 2019; Taloor et al., 2020; Singh et al., 2022). With the continuous improvements in satellite imaging technology, the high-resolution earth's surface imagery is available at a huge range of spectral bands that can be utilized in numerous applications. However, there are still many challenges yet to be resolved for detecting land-cover changes in big cities, swath width problems, and resolution issues (Vivekananda et al., 2020). Multispectral imaging allows the acquisition of earth's surface information in different spectral bands, i.e., the red, green, blue, near-infrared, thermal infrared, and short-wave infrared. But, due to a wider bandwidth, some of the critical information may be lost, which may be retrieved from the advanced algorithms but not up to the extent as in hyperspectral imaging (Wang et al., 2019).

To overcome the limitations of the multispectral dataset, hyperspectral imaging can be proven significant in terms of extracting critical information about the different natural resources. Hyperspectral imaging allows the collection of earth's surface information in much narrower bands (10–20 nm). Observing information at such narrow spectral resolution has numerous advantages, such as

✉ Sartajvir Singh
    sartajvir.dhillon@gmail.com

    Neelam Dahiya
    neelam.rani@chitkara.edu.in

    Sheifali Gupta
    sheifali.gupta@chitkara.edu.in

1   Chitkara Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140 401, India

2   University Institute of Engineering, Chandigarh University, Mohali, Punjab 140 413, India

quantifying surface materials, identifying and quantifying molecular absorption, and discriminating in different crops via different classification algorithms (Mahesh et al., 2015; Caballero et al., 2020). The classification is a process of extracting information classes from a multi-band or hyperband raster image to form a thematic map. As compared to multispectral remote sensing, the classification of hyperspectral remote sensing data is one of the challenging tasks due to the availability of an enormous amount of information (Dahiya et al., 2023). On the other hand, the hyperspectral image classification offered better discrimination among the different class categories as compared to multispectral image classification (Dahiya et al., 2023; Jarocińska et al., 2023). Some of the previous studies have proven the potential of the hyperspectral dataset (HyspIRI) as compared to the multispectral (i.e., Landsat 8 and Sentinel-2) for the mapping forest alliances in Northern California (Clark et al., 2020). In another study, Clarke et al. (2009) explored the summer and multi-seasonal variable groups via the hyperspectral and multispectral datasets and concluded the better performance of hyperspectral as compared to multispectral. It is also suggested that target-specific absorption features could be considered in the classifiers to improve the outcomes.

Generally, classification algorithms are categorized into supervised and unsupervised or hard and soft classifiers. Some of the well-defined or commonly used classifiers are summarized in Table 1. In the past few years, various classification methods have been used to classify multispectral data, such as neural network (NN) (Zhong et al., 2020), support vector machine (SVM) (Negri et al., 2016), principal component analysis (PCA) (Licciardi et al., 2012), k-nearest neighbor (Huang et al., 2016), maximum likelihood classifier (MLC) (Sood et al., 2018), and linear mixer model (LMM) (Singh et al., 2021a, 2021b). Detailed information on classifiers can be found in the literature (Lu et al. 2007). Pu et al. (2008) performed the comparative analysis of multispectral, i.e., Advanced Land Imager (ALI) onboard Earth Observation (EO-1) satellite and Landsat-7 Enhanced Thematic Mapper Plus (ETM+) and hyperspectral, i.e., Hyperspectral Imager (Hyperion) using vegetation indices (VIs), spectral texture information and maximum noise fractions (MNCs), and multivariate prediction models. They concluded the effectiveness of the hyperspectral dataset in forest mapping and leaf area index (LAI) as compared to ALI and Landsat-7. However, it has been analyzed that the accuracy can also be improved with the help of machine learning or deep learning classification models. But very rare studies were conducted to analyze the performance of hyperspectral on the different classification methods. Moreover, it is also required to perform the comparative analysis with multispectral datasets such as Landsat-8 OLI/TIRS. Therefore, there is a need to perform a comparative analysis of different classifiers for both multispectral and hyperspectral datasets.

The focus of the present study is to evaluate the performance of various classifiers in land-use monitoring using hyperspectral and multispectral datasets. The objectives are divided as (a) to implement the different classifiers, i.e., SVM, MLC and feedforward neural network (FF-NN) using hyperspectral and multispectral datasets; (b) to compute the accuracy assessment of each classifier with different datasets; (c) to compare the performance of hyperspectral and multispectral dataset on each classifier; (d) to extract the discriminate the crops using hyperspectral imagery with the best classifier and compared with the multispectral dataset. This study has been conducted over a part of the Indian States, i.e., Haryana and Uttar Pradesh. This study has numerous applications in forestry, vegetation monitoring, and soil detection. It can also be used for monitoring crop stress, detecting various plant diseases, weather forecasting, and many more (Taloor et al., 2021).

Study Area and Satellite Dataset

Study Area

The study area is part of the North Indian States, i.e., Haryana, and Uttar Pradesh, having geographical coordinates between 30°8″ N and 29°16″ N in latitude and 77°16″ E to 77°6″ E in longitude, as shown in Fig. 1. Under these regions, the major class categories include vegetation/cropland, built-up area, barren land, and water. However, the study area covers the major portion of agricultural land. Moreover, these states are the biggest contributors to agriculture in India and agriculture in these states is one of the primary sources of income and employment, which plays a significant role in the improvement of the gross domestic product (GDP) of India. Therefore, the continuous monitoring and mapping of agricultural land are crucial for the effective management of agricultural land and accomplishing future requirements. Remote sensing offers a cost-effective solution for monitoring and mapping agricultural land and various types of crops.

Dataset

In the present work, two cloud-free images from the Landsat-8 OLI/TIRS and EO-1 based Hyperion (hyperspectral) satellites were acquired on 12th March 2017 and 4 March 2017, respectively. The dataset was downloaded from the United States Geological Survey (USGS) earth explorer's online web platform (https://earthexplorer.usgs.gov/). The Landsat-8 consists of eleven spectral bands which include the wavelength of band 1 (0.43 µm—0.45

Table 1 A detailed comparison of various classification algorithms

Classifiers	Characteristics	Merits	Demerits	Example
K-nearest neighbor	It categorizes the data point into similar classes; It is based on Euclidean distance between test and training samples	It is useful for searching non-local pixels; It is helpful in image filtering and easy to implement	It is not suitable for large datasets; It does not work well with high dimensional data	ETM+[1] Meng et al. (2007); AVIRIS[2] Huang et al. (2016); Sentinel-2 Thanh et al. (2017)
Support vector machine (SVM)	The hyperplane is required to classify the different classes between data points; Used for binary classification problems	It works well with small datasets; It does not suffer from an overfitting problem	Not suitable for large datasets; It does not work well with noisy data	IKONOS Huang et al. (2016); Landsat-ETM+ Nijhawan et al. (2018); Landsat-8 Joshi et al. (2019)
Random forest (RF)	It contains a lot of trees; It is helpful in feature selection	It is a robust classifier; It shows low diversity	It is not interpretable; It tends to be overfitting	WorldView-2 Mutanga et al. (2012); MODIS[3] Ghasemian and Akhoondzadeh (2018); Landsat-7 Liu et al. (2018)
Decision tree (DT)	It is a graph-like structure; It includes root nodes, branches and leaf nodes	Domain knowledge is not required; It shows good results for high-dimensional data	Overfitting problem; It is complex	Landsat ETM+ Pal & Mather (2003); Landsat ETM+ Hua et al. (2017); Landsat-8 Acharya et al. (2016)
Maximum likelihood classification (MLC)	Pixels are assigned to the class which has the highest probability; Each band is normally distributed	Used for distributed data; It is a robust technique	Does not work well for non-distributed data; Computationally intense	Landsat ETM+ Otukei and Blaschke (2010); AWiFS[4] Sharma et al. (2013); MODIS Singh et al. (2018)
Neural network (NN)	It consists of neurons that are used to store information; Input is stored in its network	It consists of hidden layers; It has a fast learning capability	Overfitting problem due to more hidden layers; Its working depends on the hardware	AVIRIS (Dahiya et al. (2021); MODIS (Alexakis et al. (2016); Hyperion (Awad) (2014)
Multilayer perceptron (MP)	Connects multiple layers in the directed graph; Use as a supervised classification algorithm	Works well with large input data; Quick results after training	Time-consuming method; Computations are difficult	Landsat-5 Mishra et al. (2016); Landsat-8 Pham et al. (2017); Landsat 7 Yacoubi et al. (2019)

[1]ETM: Enhanced Thematic Mapper Plus;[2]AVIRIS: Airborne visible/infrared imaging spectrometer;[3]MODIS: Moderate-resolution imaging spectrometer;[4]AWiFS: Advanced Wide Field Sensor
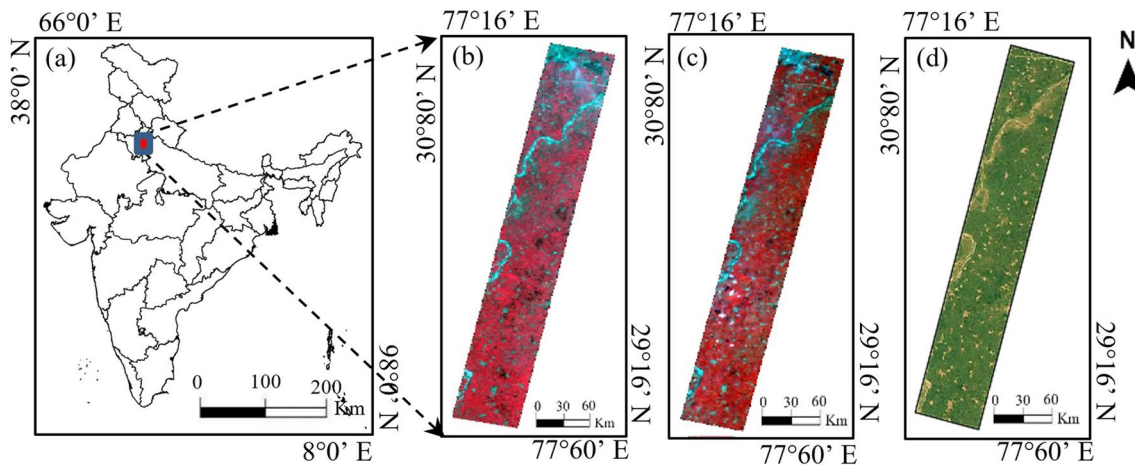
Fig. 1 Location of study site. **a** Image of India (highlighted area representing study site) **b** False color image (RGB: 40,30,20) of the study area (Hyperion) **c** False color image (RGB: 5, 4, 4) of the study area (Landsat-8) **d** Reference image

μ), band 2 (0.450 μm—0.51 μm), band 3 (0.53–0.59 μm), band 4 (0.64–0.67 μm), band 5 (0.85–0.88 μm), band 6 (1.57–1.65 μm), band 7 (2.11–2.29 μm), band 8 (0.50–0.68 μm), band 9 (1.36–1.38 μm), band 10 (10.6–11.19 μm) and band 11 (11.5–12.51 μm). The bands, i.e., 1–7 and 9, offered a spectral resolution of 30 m and band 8 had a spectral resolution of 15 m, whereas band 10 and 11 offered a spatial resolution of 100 m. On the other hand, the Hyperion EO-1 dataset includes 242 spectral bands with a separation of 10 nm with a wavelength coverage of 356–2577 nm at a spectral resolution of 30 m.

To validate the outcomes, the Pléiades constellation dataset was acquired from Google Earth history images at the spatial resolution of 0.5 m (panchromatic) and 2 m (multispectral). Airbus Defence and Space/Centre National d'Etudes Spatiales (CNES) oversees the operation of this satellite. It provides high-resolution imaging, which can give more specific information about the area. Google Earth viewer incorporated into the ERDAS Imagine version 2015 allows for image-based grounding on Google Earth. It allows users to connect with Google Earth, go to a predetermined area and analyze the satellite imagery with respect to Google Earth (Dahiya et al., 2023).

Methodology

The methodology of the proposed work is divided into three sections: (a) preprocessing of hyperspectral and multispectral datasets, (b) classification using MLC, SVM and FF-NN classifiers, and (c) accuracy assessment.

Preprocessing

Preprocessing is the first and most important step, which is to be taken care of after data collection. It is done for the successful removal of the numerous errors caused due to a variety of circumstances, including the location of the sun, varying air conditions, errors produced by satellite sensors, and errors resulting from rocky topography. If errors are not resolved timely, they may alter the result. The Hyperion EO-1 data was collected from the USGS website and consists of 242 spectral bands with a wavelength range of 356–2577 nm. The sensor's built-in visible near-infrared (VNIR) detector gathers information in bands 1 to 70, while the short-wave infrared (SWIR) detector gathers information in bands 71–242. During preprocessing, the bad bands (not informative) were removed from the datasets. The bands which were removed from the study include 1–7 (non-illuminated), 58–76 (overlap region), 221–224 (water vapor region), and 225–242 (not used). Out of 242, only 196 bands were used for classification purpose. Similarly, for Landsat-8 out of 11 spectral bands, only 9 informative bands were used for classification.

Moreover, the radiometric correction has also been performed over Hyperion EO-1 and Landsat-8 datasets using Fast Line-of-Sight Atmospheric Analysis of the Spectral Hypercubes (FLAASH) tool available in Environment for Visualizing Images (ENVI) v5.3 software. Here, digital numbers (DN) are used by the sensor to store the electromagnetic radiations (EMR) intensity. These DN need to be transformed into useful units like reflectance, radiance etc. Converting DN readings into radiance values is part of the estimation of reflectance. By considering the highest and minimum radiance values of each band, the DN imagery can be translated into a radiance value (Mishra

et al., 2009). According to Singh et al. (2018), the radiance Xi is computed as follows:

$$x_{i\lambda} = \left[\frac{(DM_i\lambda)}{(MGray)} \times \left(Bmax_\lambda - Bmin_\lambda \right) \right] + Bmin_\lambda \qquad (1)$$

where $Bmax_\lambda$ is the value of maximum radiance provided in the metadata, $Bmin_\lambda$ is the value of minimum radiance provided in the information, and DM_i is band pixel digital number. *MGray* represents the maximum DN value for a certain band. Calculating the various angles, including the solar zenith angle, azimuth angle, and elevation angle, is necessary. The distance between the Sun and its direct overhead position, or solar zenith angle, is measured in degrees.

Classification Algorithms

In the present paper, three popular supervised classification algorithms, namely (a) MLC, (b) SVM, and (c) FF-NN, have been implemented to classify hyperspectral and multispectral imagery as explained in subsequent sections. These classifiers are chosen for research purposes due to their numerous advantages as described in the coming section. The testing is done through other classifiers, also but for the present work MLC, SVM, and ANN show the best results. As shown in Table 1, KNN is not suitable for high-dimensional data, and RF and DT suffer from an overfitting issue which is resolved by SVM and showed better results for the current work. MP is a time-consuming classifier so not suitable for complex work. MLC is a robust classifier and NN is a fast learning classifier and has the capability to extract hidden features and improve the accuracy, so both are selected for the current work.

Maximum Likelihood Classifier (MLC)

The MLC is one of the most used supervised classifiers that computes the posterior probability of a pixel belonging to a specific class category. In other words, the pixel with maximum likelihood will be allocated to the corresponding class category and beneficial for more complex models of evolution (Lillesand et al., 2015). On the other hand, if the pixel has a smaller likelihood than the threshold value, it remains unclassified. It is also known as the parametric method as it is based on assumptions for the distribution of frequency for each class category. This approach is used to train the model for the classification of different classes into specific categories. The flow diagram of MLC is shown in Fig. 2a.

Steps for the execution of the MLC algorithm:

Step 1: In the prime stage, various training samples (*n*) are picked out based on different observations and spectral signatures.

Step 2: Select the number of classes.

Step 3: Afterward, files of spectral signatures of chosen class categories are generated for algorithm training.

Step 4: Compute the covariance matrices and mean vector as follows

$$M = \ln\ln(cj) - \frac{1}{2lnln(|\text{Cov}j|)} - \left[\frac{1}{2(P - mj)T\left(\text{Cov}j^{-1}\right)(P - mj)} \right] \qquad (1)$$

In Eq. (2), c_j is the probability of a class, Cov_j is the covariance matrix, P is used as the measurement matrix of the pixel, and m is used as the sample mean vector of class *j*.

Step 5: To train the model, 1000 samples were used which is subdivided into training (~80%), validation (~20%). After that, region of interest is selected (ROI) from the input image and the required class is selected from the ROI tool using Envi v5.3 software.

Step 6: Choose the probability threshold value as single and set the data scale factor value as 1. While converting integer-scaled reflectance or radiance data into floating-point values, the scale factor is employed as a division factor. After this, classification is performed.

Step 7: Then, the classified maps are visually interpreted with the reference data. If the desired result is not found, then go to step 1 and repeat the whole process.

Step 8: If the desired result is found, then a false color code is allocated to the class category of thematic images.

Support Vector Machine (SVM)

The SVM is a supervised algorithm which is used for both regression and classification purposes. It is further categorized into linear SVM for separable data and into nonlinear for inseparable data. In SVM, the hyperplane with maximum margin is selected which aims to separate the datasets into a distinct number of classes. It is implemented with the help of kernels which are used to convert the low dimensional input into the high-dimensional which helps to solve the algorithm problem (Maulik and Chakraborty, 2017). Various types of kernels like linear, polynomial, and radial basis function kernel can be used according to the requirement. The flow diagram of the SVM algorithm is shown in Fig. 2b.

Steps for the execution of the SVM algorithm:

Step 1: Select 'n' number of samples for algorithm training.

Step 2: A radial basis function kernel is selected for classification as it maps the input space in indefinite-dimensional space using Eq. (2) as given below.
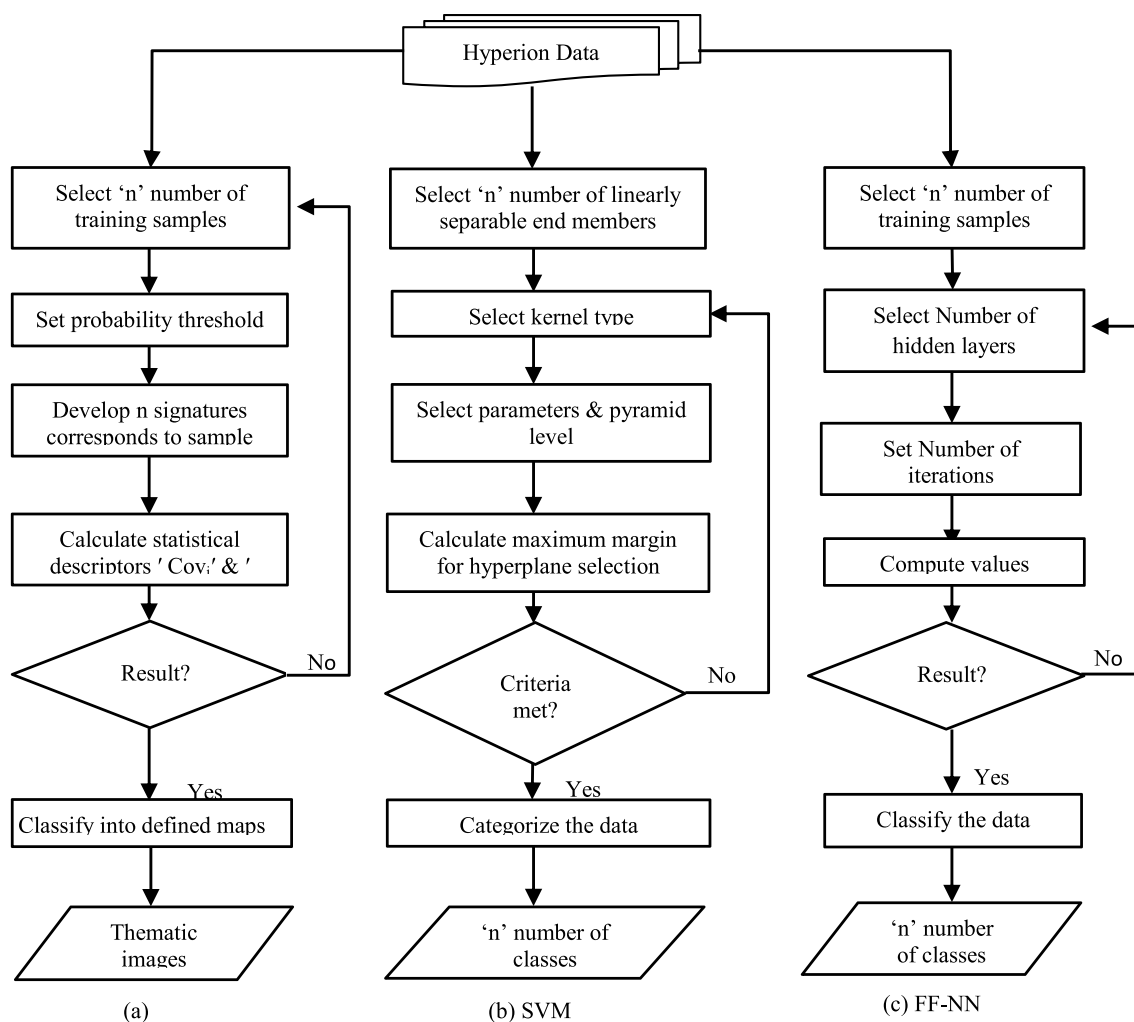
Fig. 2 Comparison of the methodology of different supervised classification algorithms: **a** MLC **b** SVM **c** FF-NN

$$k(x, xi) = \exp(-\text{gamma} * \text{sum}(x - xi)\text{pow2}) \qquad (2)$$

In Eq. (3) *k* stands for the kernel, the gamma value ranges from 0 to 1, and its value is assigned manually during the model training. Here *x* and x_i are the data points used for margin selection.

Step 3: Set all the parameters to find a hyperplane.

Step 4: Compute the hyperplane as given below.

$$w.x + b = 0 \qquad (4)$$

In Eq. (4) w is a vector which is normal to the hyperplane and *b* is an offset.

Step 5: To train the model, 1000 samples were used which is subdivided into training (~80%) and validation (~20%) After that, the region of interest is selected (ROI) from the input image and the required class is selected from ROI tool using Envi v5.3 software.

Step 6: Choose the gamma value in the kernel function as 0.005. After kernel selection, a penalty parameter is chosen that regulates the compromise between allowing for training mistakes and enforcing strict margins. The cost of incorrectly categorizing points rises as the penalty parameter's value is increased and its default value is 100. After this, classification is done.

Step 7: Then, the classified maps are visually interpreted with the reference data. If the desired result is not found, then go to step 2 and repeat the whole process.

Step 8: If the desired result is found, then a false color code is allocated to the class category.

Feedforward Neural Network (FF-NN)

The FF-NN is a subpart of an artificial neural network (ANN) and is also known as the multi-layered network of neurons (MLN). It consists of many layers, i.e., the input

Table 2 Accuracy assessment of different classifiers for (Hyperion EO-1) hyperspectral dataset

Classifier	Classified data	Accuracy parameters						Overall	
		RT	CT	CN	PA (%)	UA (%)	Kc	OA (%)	Kc
Maximum likelihood classification (MLC)	Dense vegetation	11.2	11.2	9.00	67.86	67.86	0.6380	84.40	0.7701
	Deciduous vegetation	18.4	17.6	17.06	78.26	81.82	0.7772		
	Built-up	10	17.6	9.00	76.00	76.00	0.7333		
	Water	50	51.2	56.87	96.00	93.75	0.8750		
	Barren	10.4	10	8.05	65.38	68.00	0.6429		
Support vector machine (SVM)	Dense vegetation	11.2	12.8	11.41	89.29	78.13	0.7537	87.60	0.8149
	Deciduous vegetation	16.4	15.6	14.61	78.05	82.05	0.7853		
	Built-up	10.4	10	8.67	73.08	76.00	0.7321		
	Water	52.8	51.2	56.62	93.94	96.88	0.9338		
	Barren	9.2	10.4	8.67	82.61	73.08	0.7035		
Feedforward neural network (FF-NN)	Dense vegetation	11.2	12.4	10.96	89.29	80.65	0.7820	**91.20**	**0.8711**
	Deciduous vegetation	16	15.6	15.35	87.50	89.74	0.8779		
	Built-up	9.6	10	8.33	79.17	76.00	0.7345		
	Water	50.8	50	53.50	96.06	97.60	0.9512		
	Barren	12.4	12	11.84	87.10	90.00	0.8858		

RT: Reference Total, *CT*: Classified Total, *CN*: Correct Number, *PA*: Producer Accuracy, *UA*: User Accuracy, *Kc*: Kappa value, *OA*: Overall Accuracy
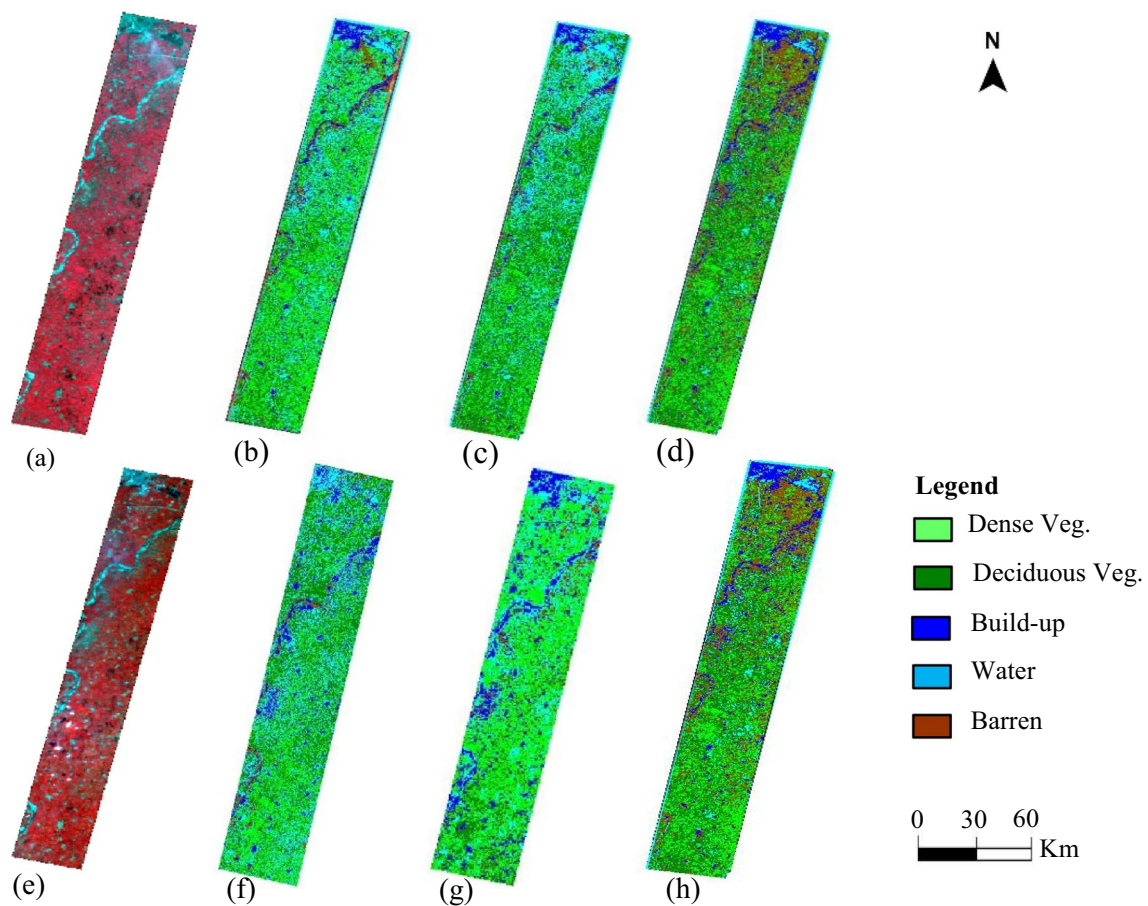
Fig. 3 Input and classified outcomes from Hyperion EO-1 and Landsat-8 using different classifiers **a** Input Hyperion EO-1 Image (RGB: 40-30-20) **b** MLC **c** SVM **d** FF-NN **e** Input Landsat-8 image **f** MLC **g** SVM (h) FF-NN

Table 3 Accuracy assessment of different classifiers for (Landsat 8 OLI/TSRI) multispectral dataset

Classifier	Classified data	Accuracy parameters						Overall	
		RT	CT	CN	PA (%)	UA (%)	Kc	OA (%)	Kc
Maximum likelihood classification (MLC)	Dense vegetation	19.6	20	19.88	73.47	71.43	0.6517	72.40	0.6550
	Deciduous vegetation	22.4	20	22.09	71.43	72.92	0.7423		
	Built-up	21.2	20	20.99	71.70	76.00	0.5631		
	Water	19.2	20	19.37	72.92	70.00	0.6287		
	Barren	17.6	20	17.67	72.73	64.00	0.5631		
Support vector machine (SVM)	Dense vegetation	22.8	20	21.5	78.18	86.00	0.8205	80.00	0.7500
	Deciduous vegetation	22	20	22.5	78.95	90.00	0.8705		
	Built-up	20	20	18.5	86.05	74.00	0.686		
	Water	18	20	17.5	77.78	70.00	0.6341		
	Barren	17.2	20	20.00	80.00	80.00	0.75		
Feedforward neural network (NN)	Dense vegetation	19.2	20	20.97	89.58	86.00	0.8205	**82.00**	**0.7600**
	Deciduous vegetation	18.4	20	17.56	78.26	72.00	0.8705		
	Built-up	22.4	20	21.97	80.36	90.00	0.686		
	Water	21.6	20	20.48	77.78	84.00	0.6341		
	Barren	18.4	20	19.02	84.78	78.00	0.75		

RT: Reference Total, *CT*: Classified Total, *CN*: Correct Number, *PA*: Producer Accuracy, *UA*: User Accuracy, *Kc*: Kappa value, *OA*: Overall Accuracy

layer, the output layer, and the hidden layer. The hidden layer between the input and the output layer aims to perform the non-linear transformation of the input layer to produce the desired output (Paoletti et al., 2019). Each perceptron in one layer is connected to every perceptron on the next layer which allows the constantly transferring or "feed forward" from one layer to the next. It allows more generalized and accurate results as compared to other supervised algorithms. It is used to solve various problems such as the validation of data and helps to find the patterns in data. The flow diagram of the FF-NN algorithm is shown in Fig. 2c.

Steps for the execution of the FF-NN algorithm:

Step 1: Select the 'n' number of training samples.

Step 2: Afterward, select the number of hidden layers which is the size between the input and the output layer.

Step 3: Select the number of iterations to train the model.

Step 4: Compute the values as given below.

$$y = a(W1 * X1 + W2 * X2 \ldots \ldots \ldots .. Wn * Xn) \tag{5}$$

In Eq. (5) a is the activation function, $(W1, W2 \ldots Wn)$ are the weights, and $(X1, X2 \ldots Xn)$ are the input neurons.

Step 5: To train the model, 1000 samples were used which is subdivided into training (~ 80%) and validation (~ 20%) After that, the region of interest is selected (ROI) from the input image and the required class is selected from ROI tool using Envi v5.3 software.

Step 6: Choose the activation method and adjust the value of threshold training in between 0 and 1.0. The training algorithm dynamically modifies the weights between nodes and, if necessary, the node thresholds. The internal weights of the node are unaffected by setting the Training Threshold Contribution to 0. Better classifications could result by adjusting the internal weights of the nodes, while poor generalizations might result from using too many weights.

Step 7: Then, the classified maps are visually interpreted with the reference data. If the desired result is not found, then go to step 2 and repeat the whole process.

Step 8: If the desired result is found, then a false color code is allocated to the class category.
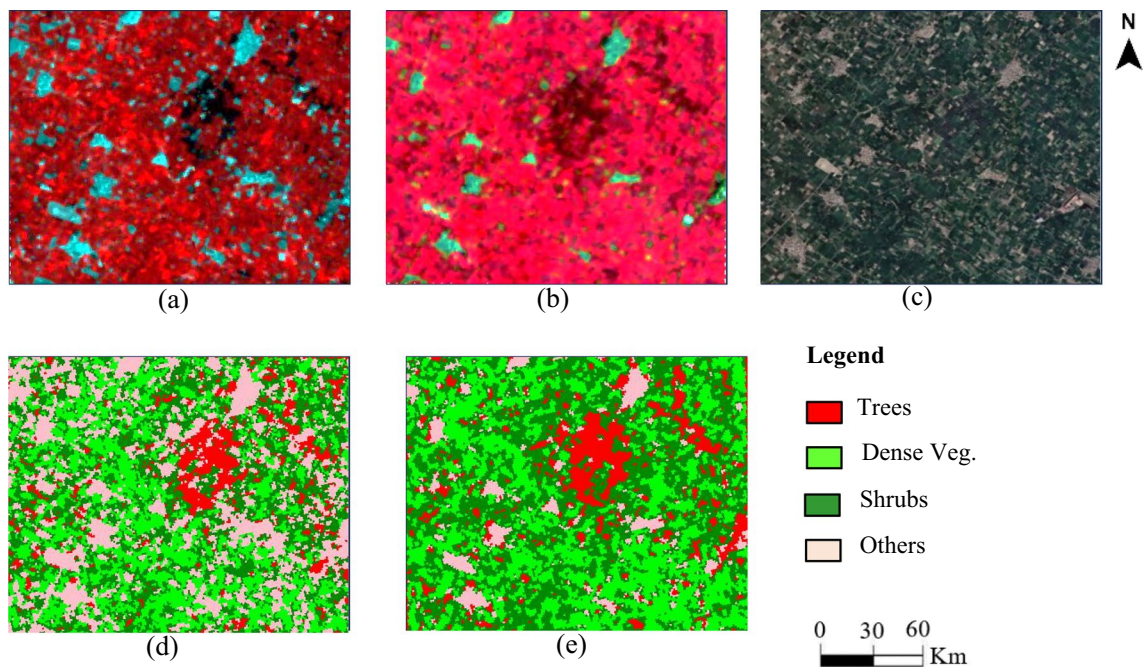
Accuracy Assessment

To validate the outcomes, the accuracy assessment has been computed for each classified map generated from MLC, SVM, and FF-NN. The reference dataset is also used for assessment to determine the accuracy of the classified results. Field surveys and the visual interpretation of high-resolution Google Earth pictures were used to gather the reference data. The research area was divided into five land-cover classes, all of which could be seen in the field and in pictures (from Google Earth), i.e., dense vegetation, deciduous vegetation, built-up, water, and barren. Envi
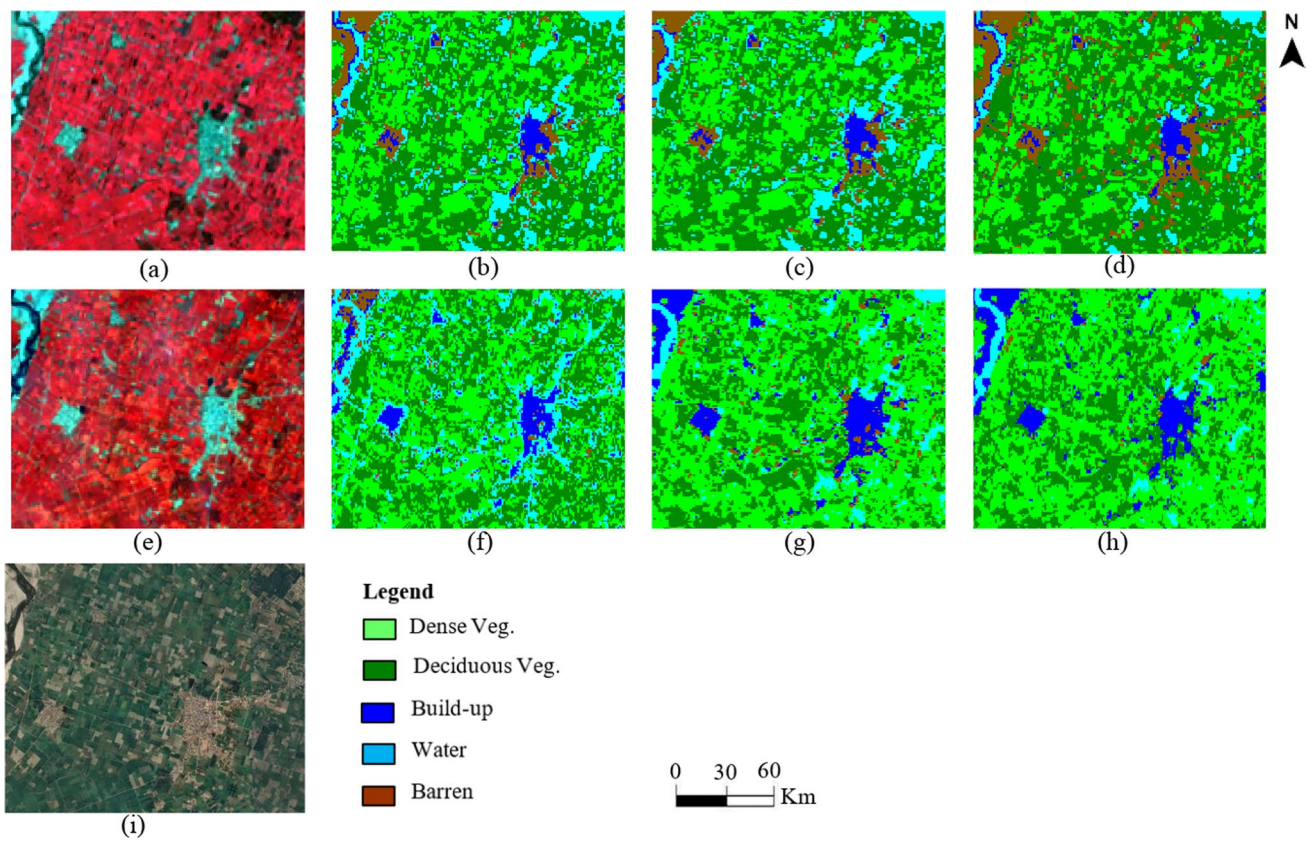
Fig. 4 Representation of subset **a** Hyperion EO-1 imagery classified by **b** MLC, **c** SVM, and **d** FF-NN; **e** Landsat-8 imagery classified by **f** MLC, **g** SVM, and **h** FF-NN; and **i** reference dataset

Fig. 5 Discrimination of ornamental crops **a** Landsat-8 Input Image **b** Hyperion EO-1 Input Image **c** Reference Image **d** FF-NN Landsat-8 **e** FF-NN Hyperion EO-1

Table 4 Accuracy assessment of discriminant ornamental crops using FF-NN classifier for (Hyperion EO-1) Hyperspectral dataset and (Landsat-8) Multispectral dataset

Classifier	Classified data	Reference total (%)	Classified total (%)	Correct number (%)	Producer accuracy (%)	User accuracy (%)	Kappa value	Overall accuracy (%)	Overall kappa
Feed Forward-Neural Network (FF-NN) Hyperion EO-1	Trees	35	33	28	79.41	100	1.000	**91.67**	0.9023
	Dense veg	44	43	40	88.88	100	1.0000		
	Shrubs	14	14	13	100	92.85	0.8059		
	Others	07	10	9	100	65	0.6745		
Feed-Forward-Neural Network (FF-NN) Landsat-8	Trees	32	27	27	84.38	100	1.000	87.88	0.8224
	Dense veg	47	43	40	85.11	93	0.8684		
	Shrubs	15	20	15	100	75	0.7059		
	Others	06	10	06	100	60	0.5745		

v5.3 tool was used to collect a total of 996 locations using stratified random sampling. According to its size, each land-cover type was given a certain number of points. The total number of samples is further divided into training (~80%) and validation set (~20%). In total, 1000 (approx.) samples were collected for both Hyperion EO-1 and Landsat-8 datasets. Multiple training samples (50 to 60 polygons) were chosen from each class category from the Hyperion and Landsat-8 datasets across the agricultural region in Haryana and Uttar Pradesh in order to train the model. The fivefold cross-validation method was used 10 times on the samples to determine the final accuracy. The dataset was shuffled before each repetition randomly and new folds were created to improvise the model performance. The essential components of the accuracy assessment included the producer's accuracy (PA), user's accuracy (UA), overall accuracy (OA), and kappa coefficient (Kc). The PA defines the probability of correct classification with respect to reference pixels and the probability of pixels fall under the correct class category, whereas OA and Kc represent the collective accuracy and distinction between actual and expected outcomes, respectively (Dahiya et al., 2023).

Results and Discussion

In the present study, two datasets, namely Hyperion EO-1 and Landsat 8 OLI/TIRS, are used as input. Three supervised classifiers, i.e., MLC, SVM, and FF-NN, have been implemented using hyperspectral as well as a multispectral dataset to classify the different categories and to narrate the impact on LULC over the part of Indian states, i.e., Haryana and Uttar Pradesh. During the classification process, various categories are explored such as dense vegetation, deciduous vegetation, built-up, water, and barren. The MLC algorithm

has been implemented according to Eq. (1) for assigning a class category to a pixel based on maximum likelihood. Afterward, SVM has been implemented using Eqs. (2) and (3) in which kernels are selected according to the desired result and the hyperplane is selected with maximum margin. According to Eq. (4), the FF-NN is implemented to select the hidden layers and a number of iterations to fetch the maximum features. The final classified outputs from each classifier using hyperspectral and multispectral imagery are shown in Fig. 3.

To compute the effectiveness of thematic or classified images, the accuracy assessment is one of the important steps to evaluate errors and the efficiency of the model. Tables 2 and 3 represent the accuracy assessment parameters computed for hyperspectral and multispectral imagery, respectively. From the statistical analysis, the accuracy assessment table has confirmed the effectiveness of the FF-NN (91.20% with Hyperion, and 82% with Landsat-8) classifier as compared to other classification methods, i.e., SVM (87.60% with Hyperion and 80% with Landsat-8) and MLC (84.40% with Hyperion and 72.40% with Landsat-8). The accuracy assessment of various supervised classifiers is performed based on various accuracy assessment parameters such as PA, UA, KC, and OA. From the experimental outcomes, it is evident that the FF-NN algorithm not only improves the accuracy for a given LULC region but also obtained the highest accuracy among various supervised classifiers using the hyperspectral dataset as compared to the multispectral dataset.

Along with the above analysis, the potential of hyperspectral has also been testified on subset representation of input and processing datasets as shown in Fig. 4. From the visual interpretation, the difference between the outcomes

of the hyperspectral and multispectral can be easily analyzed. Along with the above analysis, the potential of hyperspectral has also been testified on discrimination of different vegetation types using FF-NN classifier and compared with the Landsat-8 dataset as shown in Fig. 5. Moreover, the statistical analysis has also been computed as shown in Table 4. These outcomes depict the effectiveness of hyperspectral-classified images as compared to the multispectral classified image. The main reason behind such results is due to the potential of hyperspectral to deliver the narrow band information and produces spectra of all pixels. On the other hand, the multispectral dataset is easy to process but provides only limited information only, which results in the loss of vital information. The major challenge associated the hyperspectral imagery is the impact on the computing speed while dealing with FF-NN. However, some of the alternate or advanced approaches of ANN can also be explored to increase the computing capacity in the classification process.

Conclusion

In the present work, the Hyperion EO-1 and Landsat 8 data are evaluated over a region of Indian states, namely Haryana and Uttar Pradesh. This article shows the potential of three well-defined supervised classifiers, i.e., MLC, SVM, and FF-NN using hyperspectral and multispectral datasets. From the experimental outcomes, it is apparent that the FF-NN classification method obtained the highest accuracy (91.20% with Hyperion and 82% with Landsat-8) as compared to other classification methods, i.e., SVM (87.60% with Hyperion and 80% with Landsat-8) and MLC (84.40% with Hyperion and 72.40% with Landsat-8). It is also apparent that the hyperspectral can generate accurate information from classified maps compared to the multispectral dataset. Moreover, the potential of hyperspectral in vegetation discrimination is also evident as compared to the multispectral. This study can be further used for different applications in different regions such as crop identification, disease detection, and crop growth for sustainable crop production.

Declarations

References

Acharya, T. D., Lee, D. H., Yang, I. T., & Lee, J. K. (2016). Identification of water bodies in a Landsat 8 OLI image using a J48 decision tree. *Sensors, 16*(7), 1075. https://doi.org/10.3390/s16071075.

Alexakis, D. D., & Tsanis, I. K. (2016). Comparison of multiple linear regression and artificial neural network models for downscaling TRMM precipitation products using MODIS data. *Environmental Earth Sciences, 75*(14), 1–13. https://doi.org/10.1007/s12665-016-5883-z.

Awad, M. (2014). Sea water chlorophyll-a estimation using hyperspectral images and supervised artificial neural network. *Ecological Informatics, 24*, 60–68. https://doi.org/10.1016/j.ecoinf.2014.07.004.

Bhosle, K., & Musande, V. (2019). Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images. *Journal of the Indian Society of Remote Sensing, 47*(11), 1949–1958. https://doi.org/10.1007/s12524-019-01041-2.

Caballero, D., Calvini, R., & Amigo, J. M. (2020). Hyperspectral imaging in crop fields: Precision agriculture. *Data handling in science and technology* (Vol. 32, pp. 453–473). Elsevier.

Clarke, G. K. C., Berthier, E., Schoof, C. G., & Jarosch, A. H. (2009). Neural networks applied to estimating subglacial topography and glacier volume. *Journal of Climate, 22*(8), 2146–2160. https://doi.org/10.1175/2008JCLI2572.1.

Clark, M. L. (2020). Comparison of multi-seasonal Landsat 8, Sentinel-2 and hyperspectral images for mapping forest alliances in Northern California. *ISPRS Journal of Photogrammetry and Remote Sensing, 159*, 26–40. https://doi.org/10.1016/j.isprsjprs.2019.11.007.

Dahiya, N., Gupta, S. and Singh, S. (2021). Quantitative analysis of different land-use and land-cover classifiers using hyperspectral dataset. In *2021 Sixth International Conference on Image Information Processing (ICIIP)* (Vol. 6, pp. 256–260). IEEE. https://doi.org/10.1109/ICIIP53038.2021.9702568.

Dahiya, N., Gupta, S., & Singh, S. (2023). Qualitative and quantitative analysis of artificial neural network-based post-classification comparison to detect the earth surface variations using hyperspectral and multispectral datasets. *Journal of Applied Remote Sensing, 17*(3), 032403–032403. https://doi.org/10.1117/1.JRS.17.032403.

El Yacoubi, S., Fargette, M., Faye, A., de Carvalho Junior, W., Libourel, T., & Loireau, M. (2019). A multilayer perceptron model for the correlation between satellite data and soil vulnerability in the F Senegal. *International Journal of Parallel, Emergent and Distributed Systems, 34*(1), 3–12. https://doi.org/10.1080/17445760.2018.1434175.

Ghasemian, N., & Akhoondzadeh, M. (2018). Introducing two Random Forest based methods for cloud detection in remote sensing images. *Advances in Space Research, 62*(2), 288–303.

Hua, L., Zhang, X., Chen, X., Yin, K., & Tang, L. (2017). A feature-based approach of decision tree classification to map time series urban land use and land cover with Landsat 5 TM and Landsat

8 OLI in a Coastal City. *China. ISPRS International Journal of Geo-Information, 6*(11), 331. https://doi.org/10.3390/ijgi6110331.

Huang, K., Li, S., Kang, X., & Fang, L. (2016). Spectral–spatial hyperspectral image classification based on KNN. *Sensing and Imaging, 17*(1), 1. https://doi.org/10.1007/s11220-015-0126.

Jarocińska, A., Kopeć, D., Niedzielko, J., Wylazłowska, J., Halladin-Dąbrowska, A., Charyton, J., Piernik, A., & Kamiński, D. (2023). The utility of airborne hyperspectral and satellite multispectral images in identifying Natura 2000 non-forest habitats for conservation purposes. *Scientific Reports, 13*(1), 4549. https://doi.org/10.1038/s41598-023-31705-6.

Joshi, P. P., Wynne, R. H., & Thomas, V. A. (2019). Cloud detection algorithm using SVM with SWIR2 and tasseled cap applied to Landsat 8. *International Journal of Applied Earth Observation and Geoinformation, 82*, 101898.

Licciardi, G., Marpu, P. R., Chanussot, J., & Benediktsson, J. A. (2012). Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles. *IEEE Geoscience and Remote Sensing Letters, 9*(3), 447–451. https://doi.org/10.1109/LGRS.2011.2172185.

Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. Wiley.

Liu, J., Feng, Q., Gong, J., Zhou, J., Liang, J., & Li, Y. (2018). Winter wheat mapping using a random forest classifier combined with multi-temporal and multi-sensor data. *International Journal of Digital Earth, 11*(8), 783–802. https://doi.org/10.1080/17538947.2017.1356388

Lu, D. & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing, 28*(5), 823–870. https://doi.org/10.1080/01431160600746456.

Lu, D., Mausel, P., Brondízio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing,, 25*(12), 2365–2407. https://doi.org/10.1080/0143116031000139863.

Mahesh, S., Jayas, D. S., Paliwal, J., & White, N. D. G. (2015). Hyperspectral imaging to classify and monitor quality of agricultural materials. *Journal of Stored Products Research, 61*, 17–26. https://doi.org/10.1016/j.jspr.2015.01.006.

Maulik, U., & Chakraborty, D. (2017). Remote sensing image classification: A survey of support-vector-machinebased advanced techniques. *IEEE Geoscience and Remote Sensing Magazine, 5*(1), 33–52. https://doi.org/10.1109/MGRS.2016.2641240.

Meng, Q., Cieszewski, C. J., Madden, M., & Borders, B. E. (2007). K nearest neighbor method for forest inventory using remote sensing data. *GIScience & Remote Sensing, 44*(2), 149–165. https://doi.org/10.2747/1548-1603.44.2.149.

Mishra, V. D., Negi, H. S., Rawat, A. K., Chaturvedi, A., & Singh, R. P. (2009). Retrieval of sub-pixel snow cover information in the Himalayan region using medium and coarse resolution remote sensing data. *International Journal of Remote Sensing, 30*(18), 4707–4731. https://doi.org/10.1080/01431160802651959.

Mishra, V. N., & Rai, P. K. (2016). A remote sensing aided multilayer perceptron-Markov chain analysis for land use and land cover change prediction in Patna district (Bihar), India. *Arabian Journal of Geosciences, 9*(4), 249. https://doi.org/10.1007/s12517-015-2138-3.

Negri, R. G., Dutra, L. V., & SantAnna, S. J. S. (2016). Comparing support vector machine contextual approaches for urban area classification. *Remote Sensing Letters, 7*(5), 485–494. https://doi.org/10.1080/2150704X.2016.1154218.

Nijhawan, R., Raman, B., & Das, J. (2018). Meta-classifier approach with ANN, SVM, rotation forest, and random forest for snow cover mapping. In: Chaudhuri, B., Kankanhalli, M., Raman, B. (eds) *Proceedings of 2nd International Conference on Computer Vision & Image Processing. Advances in Intelligent Systems and Computing*, vol 704. Springer, Singapore. https://doi.org/10.1007/978-981-10-7898-9_23.

Pal, M., & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment, 86*(4), 554–565. https://doi.org/10.1016/S0034-4257(03)00132-9.

Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2019). Deep learning classifiers for hyperspectral imaging: A review. ISPRS *Journal of Photogrammetry and Remote Sensing, 158*, 279–317. https://doi.org/10.1016/j.isprsjprs.2019.09.006.

Pham, B. T., Tien Bui, D., Pourghasemi, H. R., Indra, P., & Dholakia, M. B. (2017). Landslide susceptibility assesssment in the Uttarakhand area (India) using GIS: A comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theoretical and Applied Climatology, 128*(1–2), 255–273. https://doi.org/10.1007/s00704-015-1702-9.

Pu, R., Kelly, M., Anderson, G. L., & Gong, P. (2008). Using CASI hyperspectral imagery to detect mortality and vegetation stress associated with a new hardwood forest disease. *Photogrammetric Engineering & Remote Sensing, 74*(1), 65–75. https://doi.org/10.14358/PERS.74.1.65.

Sharma, J. K., Mishra, V. D., & Khanna, R. (2013). Impact of topography on accuracy of land cover spectral change vector analysis using AWiFS in Western Himalaya. *Journal of the Indian Society of Remote Sensing, 41*(2), 223–235. https://doi.org/10.1007/s12524-011-0180-5.

Singh, S., & Talwar, R. (2018). An intercomparison of different topography effects on discrimination performance of fuzzy change vector analysis algorithm. *Meteorology and Atmospheric Physics, 130*(1), 125–136. https://doi.org/10.1007/s00703-016-0494-5.

Singh, S., Tiwari, R. K., Gusain, H. S., & Sood, V. (2020). Potential applications of SCATSAT-1 satellite sensor: A systematic review. *IEEE Sensors Journal, 20*(21), 12459–12471. https://doi.org/10.1109/JSEN.2020.3002720.

Singh, S., Tiwari, R. K., Sood, V., & Gusain, H. S. (2021a). Detection and validation of spatiotemporal snow cover variability in the Himalayas using Ku-band (13.5 GHz) SCATSAT-1 data. *International Journal of Remote Sensing, 42*(3), 805–815. https://doi.org/10.1080/2150704X.2020.1825866.

Singh, S., Tiwari, R. K., Sood, V., Gusain, H. S., & Prashar, S. (2021b). Image fusion of Ku-band-based SCATSAT-1 and MODIS data for cloud-free change detection over western Himalayas. *IEEE Transactions on Geoscience and Remote Sensing, 60*, 1–14. https://doi.org/10.1109/TGRS.2021.3123392.

Singh, S., Tiwari, R. K., Sood, V., Kaur, R., & Prashar, S. (2022). The legacy of scatterometers: Review of applications and perspective. *IEEE Geoscience and Remote Sensing Magazine.* https://doi.org/10.1109/MGRS.2022.3145500.

Sood, V., Gupta, S., Gusain, H. S., & Singh, S. (2018). Spatial and quantitative comparison of topographically derived different classification algorithms using AWiFS data over Himalayas, India. *Journal of the Indian Society of Remote Sensing, 46*(12), 1991–2002. https://doi.org/10.1007/s12524-018-0861-4.

Sood, V., Gupta, S., Gusain, H. S., Singh, S., & Taloor, A. K. (2021a). Topographic controls on subpixel change detection in western Himalayas. *Remote Sensing Applications: Society and Environment, 21*, 100465. https://doi.org/10.1016/j.rsase.2021.100465.

Sood, V., Gusain, H. S., Gupta, S., Taloor, A. K., & Singh, S. (2021b). Detection of snow/ice cover changes using subpixel-based change detection approach over Chhota-Shigri glacier, Western Himalaya, India. *Quaternary International, 575*, 204–212. https://doi.org/10.1016/j.quaint.2020.05.016.

Taloor, A. K., Kothari, G. C., Manhas, D. S., Bisht, H., Mehta, P., Sharma, M., Mahajan, S., Roy, S., Singh, A. K., & Ali, S. (2021). Spatio-temporal changes in the Machoi glacier Zanskar Himalaya

India using geospatial technology. *Quaternary Science Advances, 4*, 100031.

Taloor, A. K., Kumar, V., Singh, V. K., Singh, A. K., Kale, R. V., Sharma, R., Khajuria, V., Raina, G., Kouser, B., & Chowdhary, N. H. (2020). *Land Use Land Cover Dynamics Using Remote Sensing and GIS Techniques* (p. 37). Springer.

Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors*. https://doi.org/10.3390/s18010018.

Vivekananda, G. N., Swathi, R., & Sujith, A. V. L. N. (2020). Multi-temporal image analysis for LULC classification and change detection. *European Journal of Remote Sensing*. https://doi.org/10.1080/22797254.2020.1771215.

Wang, Z., Stoffelen, A., Zhang, B., He, Y., Lin, W., & Li, X. (2019). Inconsistencies in scatterometer wind products based on ASCAT and OSCAT-2 collocations. *Remote Sensing of Environment*, *225*, 207–216. https://doi.org/10.1016/j.rse.2019.03.005.

Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, *221*, 430–443. https://doi.org/10.1016/j.rse.2018.11.032.