



# Extraction of Buildings in Urban Area for Surface Area Assessment from Satellite Imagery based on Morphological Building Index using SVM Classifier

R. Avudaiammal<sup>1</sup>  · P. Elaveni<sup>1</sup> · Shirley Selvan<sup>1</sup> · Vijayarajan Rajangam<sup>2</sup>

Received: 16 August 2019 / Accepted: 26 August 2020 / Published online: 10 September 2020  
© Indian Society of Remote Sensing 2020

## Abstract

Rooftops are essential features, extracted from satellite images for their significance in applications such as update of urban geodatabase, risk assessment and rescue map. In this work, a methodology (MBION-SVM) which integrates morphological, spectral, shape and geometrical features with SVM classifier to classify the objects within the satellite image into building rooftops and non-rooftops has been proposed. The probable buildings are detected using Morphological Building Index (MBI). The mislabeled rooftops are eliminated by combining Otsu thresholding and Normalized Difference Vegetation Index (NDVI). Geometrical features computed from identified building rooftops are used to train a support vector machine (SVM), and self-correction is performed for removing any mislabeled rooftops and to provide the data on surface area of the perfect rooftops. Here, we have used Very High Resolution (VHR) images of Worldview-2 and Sentinel-2. We have analyzed the performance of the proposed building extraction approach with classification algorithms such as linear discriminant analysis, logistic regression and SVM. Since the proposed method gives an accuracy around 99%, precision of 89%, a perfect recall of 1 and a F-score of 88%, it can be effectively utilized to extract the buildings from VHR images for any appropriate application.

**Keywords** Building extraction · MBI · Edge detection · Otsu · NDVI · SVM

## Introduction

The application domains of remote sensing have been significantly augmented with accessibility of sub-meter resolution data from high-resolution earth satellites, such as IKONOS, WorldView-2 and QuickBird (Liu et al. 2008, Gavankar et al. 2018). In order to distinguish natural

objects (vegetation, bare land, mountains, water bodies) and man-made objects (buildings, track, road) with more accuracy, researchers started working on Very High Resolution (VHR) images captured by these satellites. Rooftops are essential features to be extracted from satellite images due to their significance in applications such as urban cadastral management, 3D map reconstruction, updation of urban geodatabase, risk assessment and rescue map (You et al. 2018). Nowadays, the urban sector is rapidly growing and consequently more energy is being consumed. Due to depletion of the ozone layer, the weather becomes unpredictable, days become longer, and summer becomes hotter requiring more ventilation and air-conditioning units. The enormous energy required to drive these units in buildings utilizes high amount of electricity, one of our crucial resources. In order to effectively use electricity, insulation of buildings becomes a prime factor and installation of solar panels is required for generation of electricity. This requires knowledge of the building area.

---

✉ R. Avudaiammal  
dravudaiece@gmail.com

P. Elaveni  
elavenipalanivel@gmail.com

Shirley Selvan  
shirleycharlethenry@gmail.com

Vijayarajan Rajangam  
viraj2k@gmail.com

<sup>1</sup> Department of Electronics and Communication Engineering,  
St. Joseph's College of Engineering, Chennai, India

<sup>2</sup> SENSE, VIT, Chennai, India

Building extraction is carried out based on pixel and object-based methods that include both spectral and spatial information obtained from VHR images. Building extraction from urban areas faces challenges in the realm of accuracy and efficiency due to the diversity and complexity of building. Buildings comprising of multiple types may result in inappropriate extraction results (Hu et al. 2008). In an urban area, density of buildings is more and hence there is a possibility of detecting multiple buildings as a single building. The presence of shadows in an image causes errors in analysis. When shadow of a building falls on another building, conventional methods fail to detect the desired rooftop (Hu et al. 2011). Furthermore, nearby trees, road segments and bridges may be erroneously spotted as buildings. All these issues are due to complexity of spatial information and heterogeneity of spectral information present within the same class of objects (Shi et al. 2018).

Nowadays, Very High Resolution remote sensing images with multispectral bands are easily available due to the advancement of low-orbit earth imaging technology and provide exhaustive spatial information (You et al. 2018). Even though the resolution of VHR images provides exhaustive spatial information, it does not logically match up in providing higher image interpretation accuracies, especially in urban areas (Myint et al. 2004). Therefore, it is computationally intense for experts to label the buildings manually and is also not possible to identify complex information from VHR images (Gavankar et al. 2018). To ease the method of extracting buildings from VHR images, many researchers have introduced automated building extraction methods to reduce the burden of human to a greater extent (Zha et al. 2003). The advancements in machine learning (ML) led to the development of automatic methods with trained models to provide the desired results for various remote sensing applications. These ML methods build relationships between remotely sensed data samples and the outcome of interest. These methods predict output classes by estimating correlations among input variables and quantify uncertainty related to these predictions and estimates. The most common ML algorithms used in remote sensing data are classification, clustering, regression and dimension reduction (Holloway et al. 2018). In order to formulate an automatic building identification that includes both spectral and spatial information obtained from VHR images, huge number of studies are focused toward classification techniques. Supervised learning-based methods that combine both spatial and spectral information are commonly used (Benediktsson et al. 2003; Senaras et al. 2013). An automated building extraction method based on region segmentation and feature extraction is used to extract simple rectangular rooftops as well as complicated buildings (Xiaoying et al. 2005) and (Liu et al. 2008).

SVM is a type of nonparametric supervised algorithm that maps the support vectors present in N-dimensional feature space into two output classes. SVM creates the hyperplane in N-dimension feature space to perfectly classify objects into two classes by taking into consideration the maximum distance between data points of both classes. SVM is employed for analysis of remote sensing data for land use and land cover classification (Szuster et al. 2011), for crop classification (Mathur et al. 2008), for automatic recognition of man-made objects (Inglada 2007) and for automatic building extraction (Dayana et al. 2018).

Like classifier, the regression model is also trained based on a set of input features extracted from the associated classes (output). Logistic regression on satellite imagery data is employed to estimate the location and degree of deforestation (Bavaghar et al. 2015). Linear discriminant analysis (LDA) is one of the dimension reduction method-based ML algorithm.

A number of works have been conducted in the field of building extraction from high-resolution satellite imagery such as QUCIKBIRD images (Wei et al. 2004; Dutta et al. 2005; Lefèvre et al. 2007; Hu et al. 2008; Attarzadeh et al. 2018), IKONOS images (Lee et al. 2003; Xiaoying et al. 2005) and Worldview-2 images (Xu 2008; Huang et al. 2012; Kumar et al. 2012; Huang et al. 2017). Since the literature is vast in this area of research, some related papers are discussed in “[Literature Review](#)” section.

To defeat the weakness of the existing related research in order to reduce false positives (mislabeled buildings and missing buildings) and to improve the accuracy of building extraction for meeting up the requirements of application, Morphological Building Index-based MBION-SVM methodology has been proposed to automatically classify objects into rooftops and non-rooftops.

The outline of the paper is as follows: “[Literature Review](#)” section discusses the related work in the field of building extraction. “[Methodology](#)” section describes the proposed MBION-SVM methodology to automatically classify objects into rooftops and non-rooftops. “[Results and Discussion](#)” section presents performance analysis of classification using linear discriminant analysis (LDA), logistic regression and SVM. “[Conclusion](#)” section deals with conclusion and discussion on the performance of the proposed work and future scopes related to it.

## Literature Review

Even though high-resolution satellite imagery offers detailed information for building identification and extraction, it does not reasonably leave higher image interpretation accuracies because of the complexity of spatial information and heterogeneity of spectral

information, present within the same class of objects (Myint et al. 2004). This makes building extraction from satellite imagery a challenging task. Therefore, many researchers started formulating building extraction methods to utilize the complete spectral information or spatial information or by the combination of these information obtained from VHR images (You et al. 2018). In this section, various types of building extraction methods and their fundamental ideology and evolution have been discussed.

The traditional methods for building identification from satellite imagery were based on pixels that were located in the same region of space exhibiting very close spectral characteristics. The pixel-based techniques that exploit spatial information has the ability to differentiate spectrally similar targets. Spatial information includes texture, structure and contextual information. The well-known spatial features used in building extraction are the morphological profiles (Pesaresi et al. 2001; Benediktsson et al. 2003; Fauvel et al. 2008), gray-level co-occurrence matrix (GLCM) (Myint et al. 2004), wavelet texture (Ouma et al. 2006) and pixel shape index (PSI) (Huang et al. 2007). The pixel-based techniques were limited in building extraction because these approaches did not convey the contextual information, shape and spatial information present in the images (Blaschke 2003). Therefore, the research is focused toward object-based building extraction methods in which group of pixels characterizes the building and also their mutual relationship in the image is considered.

In object-based method, spatial information of the built-up areas obtained from segmentation along with spectral information such as the minimum, average and maximum values of each band and their variance are used. The spectral information-based methods are insufficient for differentiating spectrally alike classes such as buildings, roads, parking lots and open areas from high-resolution imagery. This is because these classes are constructed using the same type of material, and as a result, these have high variance among intraclass objects and low variance among interclass objects (Bruzzone et al. 2006). Therefore, building extraction methods are evolved by developing several transform-based methodologies (Hu et al. 2008; Sirmacek et al. 2009; Shi et al. 2018) and indices-based methodologies. The indices based have different types such as Normalized Difference Built-up Index (NDBI) (Zha et al. 2003), Index-based Built-up Index (IBI) (Xu 2008), PCA-based Build up Index (Kumar et al. 2012), Morphological Building Index (MBI) (Huang et al. 2011; Huang 2016; You et al. 2018) and Morphological Shadow Index (MSI) (Huang et al. 2012).

## Transform-Based Approaches

It is tough to differentiate buildings of different shapes and color from other man-made structures of similar color by simply using the local features. Lee et al. 2003 used Hough transformation along with Iterative Self-Organizing Data Analysis technique algorithm (ISODATA) for returning the locations and shapes of buildings automatically from IKONOS images. In order to extract small buildings from IKONOS satellite imagery spectral, structural and contextual information has been used (Xiaoying et al. 2005). It has the ability to extract 72.7% of buildings from other segments which have alike spectral characteristics and with a quality percentage of 58.8%. Conversely, the integration of features leads to the complexity.

Dutta et al. 2005 proposed an object-oriented classification method for developing buildings catalog of an urban area in aiding disaster risk analysis. In the study, image segmentation is carried out by setting a threshold value using a combination of both spectral and shape heterogeneity. They validated their study using QuickBird image and achieved 90% accuracy but failed in extracting the built-up features precisely covered with trees or vegetal cover. In another study, Hu et al. 2008 extracted buildings covered with trees or vegetal cover and roofs painted with green and with vegetal cover using Vegetation and Achromatic Objects Indifferent Transformation (VAOIT).

Then, Scale Invariant Feature Transform (SIFT) and graph theoretical tools-based methods were developed for detecting buildings captured in different lighting conditions by utilizing the grayscale information. But this method fails to detect buildings if the contrast between the rooftop and the background is low (Sirmacek et al. 2009).

Another building extraction method to detect various complex-shaped buildings based on SIFT is carried out by incorporating translation and prior template's rotation (Yang et al. 2012). In another SIFT-graph based method, Shi et al. 2018 collected building templates manually, and with the use of description vectors of the SIFT points and Euclidean distances among the SIFT points, the vertex set and edge set of the graph are built. The authors validated their study using GF-2 panchromatic HRS imagery and reported that the method is not suitable for automatic building extraction in an urban area that has wide range of buildings of various sizes and shapes as it involves the manual collection of the building templates.

The buildings appeared differently due to the presence of shadows, illumination difference and reflectance. The presence of shadows in an image causes errors in analysis. When shadow of a building falls on another building, conventional methods fail to detect the desired rooftop (Hu et al. 2011). Therefore, Shade Line and Shadow Line Angle

(SLSLA)-based shadow analysis is employed in identifying buildings. The limitation of this method is knowledge of information regarding building height, radiation from the sun, shadows by trees and buildings (Hu et al. 2011). Extraction of buildings through shadow-based methodologies is limited due to the poor detection performance of shadows (Ok 2013; Ok et al. 2013; Senaras et al. 2016). Buildings are detected by using the shadow mask obtained through graph cuts and solar angle information (Kovacs et al. 2015).

Transform-based methods are not sufficient to detect buildings in urban areas, if the contrast between the rooftop and the background is low and matching of regular features does not work correctly.

### Indices-Based Approaches

In the study, (Zha et al. 2003), Normalized Difference Built-up Index (NDBI), an exclusive spectral index of built-up areas, is employed for automatic classification of built-up areas. The use of NDBI helped in achieving an accuracy of 92.6% compared to the method of maximum likelihood classification. The authors concluded that NDBI is not suitable due to the following factors. As the reflectance of urban areas varies due to the influence of seasonal variation in the reflectance of vegetation area, it can map only broad urban land covers and cannot separate industrial, commercial and residential areas. As spectral response of urban areas and barren land is alike in all TM bands, it cannot separate urban areas from barren. These are difficult to be satisfactorily mapped even using the conventional supervised classification method anyway.

Xu 2008 derived a new index called Index-based Built-up Index (IBI), for preserving features of built-up area. The computation of IBI involves three thematic index bands of new image obtained using Soil Adjusted Vegetation Index (SAVI), the Modified Normalized Difference Water Index (MNDWI) and the Normalized Difference Built-up Index (NDBI) computed from original multispectral image. The authors validated their results in WorldView-2 satellite data. Nevertheless, the SWIR band required to extract NDBI is not offered in WorldView-2 PCA 1 and NIR 2-based normalized built-up index (PCABI) utilized for extracting built-up land exhibited higher accuracy (Kumar et al. 2012).

Another index, Morphological Building Index (MBI), is defined for automatic building extraction from high-resolution imagery (Huang et al. 2011). The limitation of MBI is its commission and omission errors. The commission errors are associated with bright soil, open areas and roads since they are also brighter than their adjacent region and their spectral characteristics are similar to buildings. The omission errors comprise of heterogeneous and dark roofs.

In another study, commission and omission errors are eliminated with the use of Morphological Shadow Index (MSI), Vegetation Index and dual thresholding (Huang et al. 2011). A post-processing framework is developed by integrating the morphological, spectral, contextual, geometrical information in a chronological practice to reduce all errors created in conventional MBI (Huang et al. 2017).

Another framework developed by combining MBI, object-oriented information and integration of spectral, geometrical and contextual information of buildings is also presented (Ding et al. 2018). You et al. 2018 improved the original MBI algorithm by extracting and optimizing local feature points using Gabor wavelet transform and saliency index, respectively. Then they used a spatial voting matrix for extracting built-up areas and MBI algorithm for detecting buildings from the extracted built-up areas. As the local feature points derived from small scattered buildings are few, these are mistakenly removed by the saliency index. Moreover, the method is not capable of extracting dark roof buildings because the MBI algorithm treats the building as a bright structure, whereas buildings with dark roofs have low MBI values and hence will be removed wrongly during binarization of the MBI feature image.

The integration of spectral, structural and contextual information features leads to the complexity (Xiaoying et al. 2005) and performs well (Ding et al. 2018; Attarzadeh et al. 2018). It is inferred from the review of different methodologies proposed by different researchers that in spite of complexity, a combination of spectral features, shape features and morphological operators works better for extracting buildings from satellite images.

The contribution of MBION-SVM method is to develop a simple object-based method to extract building rooftops. The principle behind this method lies in integration of MBI with Otsu thresholding and Normalized Difference Vegetation Index (NDVI) to minimize the number of missing and mislabeled rooftops. The standard geometrical features extracted from the obtained building rooftops are used to train the classifier and to perform self-correction for further minimization of mislabeled rooftops. The classification results using ML algorithms such as linear discriminant analysis (LDA), logistic regression and support vector machine (SVM) classifier are analyzed.

### Methodology

The objective of MBION-SVM is to develop an automatic building extraction method using a support vector machine. Our main contribution is the integration of structural features (MBI), spectral feature (NDVI), global thresholding, geometrical features-based self-correction in object-based

processing for extraction of building rooftops. The methodology works in four phases. In the first phase, structural information of the satellite imagery is extracted by deriving MBI using a set of differential morphological profiles. In the second phase, shadows are removed using Otsu's method and vegetation area is removed using spectral feature NDVI to refine MBI. In the third phase, buildings are extracted from the obtained built-up area using object-based image processing, as it is more suitable to provide vector output for creating and updating of database with appropriate relationship. In the fourth phase, shape features and geometrical features are extracted to form the feature set and to train the SVM. The workflow of the proposed methodology for building extraction is shown in Fig. 1.

## Study Area

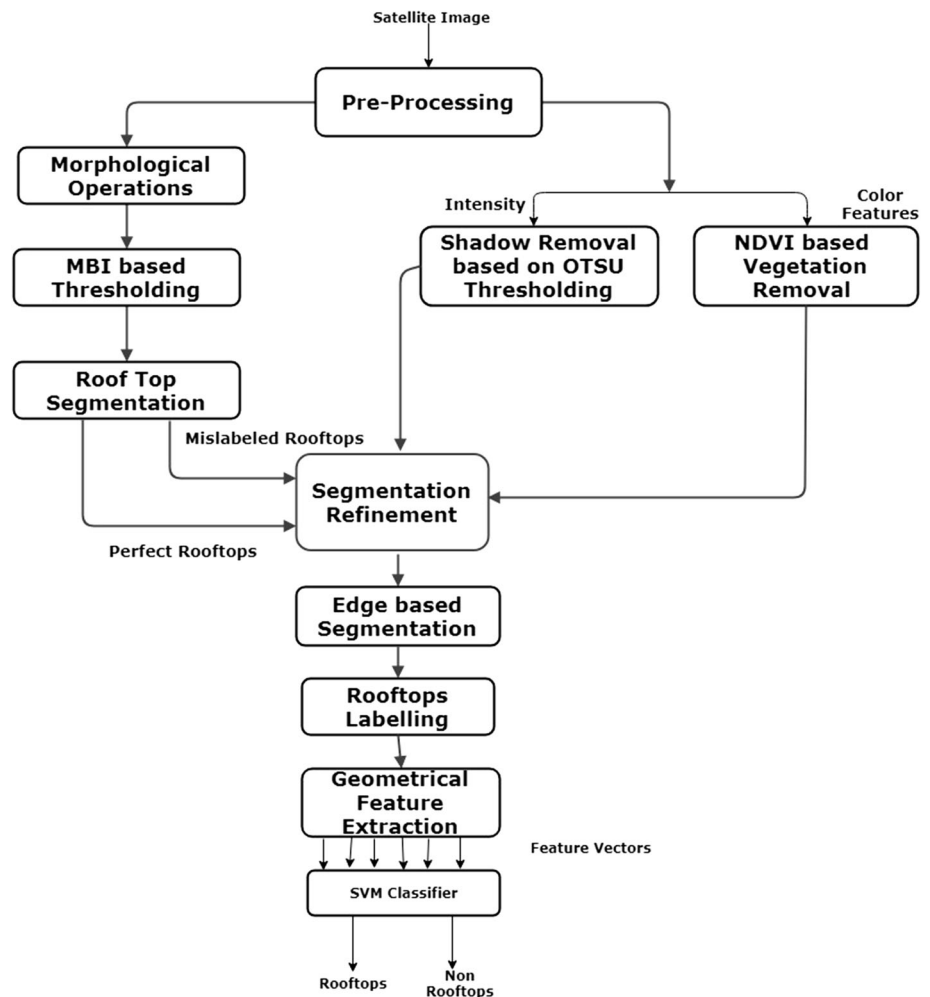
In this work, two images of Madurai city, Tamilnadu, India, from WorldView-2 satellite and two study images from Sentinel-2 satellite of which, one is a portion of

Mumbai, the Metropolitan city of Maharashtra state, India, and another is a portion of Chongqing, megacity in southwest China, are taken as study area for analysis. The images include various urban classes such as vegetation, barren land, bridges, roads, urban houses and commercial buildings.

## Pre-Processing

Pre-processing of satellite image data is done in order to extract detailed information that helps in extraction of features and classification of an image. Pre-processing of VHR images is usually carried out to prepare the images free from sensor-based distortions and platform-specific distortions of data. Geometrical distortions are caused by nonlinearities of a sensor's field of view and due to variations in the earth's curvature while capturing earth's surface. These distortions are corrected so as to represent the image in real-world geographic coordinates. The radiometric correction is done on the image to locate any mismatch of data between the sensor and noise in the

**Fig. 1** Block diagram of MBION-SVM method



atmosphere. The data are corrected for perfectly denoting the radiations emitted and reflected, for calculation by the sensor (Tarantino et al. 2012). The above-mentioned distortions are pre-corrected in the VHR image used in this work. The only preprocessing done in this work is converting the image under study into grayscale image by recording the maximum intensity value of R, G and B bands.

### Morphological Operations

In the first phase of building extraction, the structural information of the image is extracted by means of contour and shape-based processing on the image. This type of processing is performed by using morphological operations: dilation and erosion. These operations find each pixel value of the processed image by applying a structuring element (SE) on an input image and by comparing corresponding neighborhood pixels. Dilation makes objects more visible by inserting pixels to the boundaries of objects in an image, whereas erosion removes small objects by eliminating pixels on object boundaries.

### Morphological Profiles

The morphological profiles are derived by performing opening and closing operations sequentially. Morphological opening (erosion followed by dilation) is performed on an image to smooth the border of large objects, to preserve the shape and size of larger objects (buildings, roads and tracks) and also to remove small objects from an image. The opening of the image is carried out by using top-hat transform. Morphological closing (dilation followed by erosion) is performed on the image to fill small holes of an image and to maintain the shape and size of the objects in the image. The opening and closing of the image are carried out by using top-hat transform and bottom-hat transform, respectively. Differential morphological profiles (DMPs) are then derived by computing differences between morphological profiles with consecutive scales of structuring element (SE) to signify spatial pattern complexity of buildings.

### Morphological Building Index (MBI)

MBI is used to identify building candidates from satellite images and is exclusively intended for detection of probable buildings in urban areas where huge number of buildings are present. In general, building regions appear more luminous than neighborhood locations. Since MBI provides a relationship between hidden characteristics of buildings such as brightness and contrast, it identifies probable building rooftops (Huang et al. 2011). The

probable building rooftops might include both mislabeled rooftop and perfect rooftop. The mislabeled rooftops are due to open areas, vegetation and roads being wrongly identified as buildings. Buildings have alike spectral reflectance as that of roads, but roads are generally extended in one or two directions, whereas buildings are more isotropic. High local contrast and isotropic features of buildings cause larger value of differential morphological profiles (DMPs) for buildings than roads. Therefore, Morphological Building Index (MBI) is designed on a series of linear structural elements (SE) to estimate the size and directionality of structures (Huang et al. 2012).

MBI is calculated by averaging DMPs as given in Eq. (1)

$$\text{MBI} = \sum \frac{\text{DMP}(D, S)}{D_N \times S_N} \quad (1)$$

where

$D$ —direction

$S$ —scale of the WTH transformation,

$D_N$ —total number of directions

$S_N$ —total number of scales

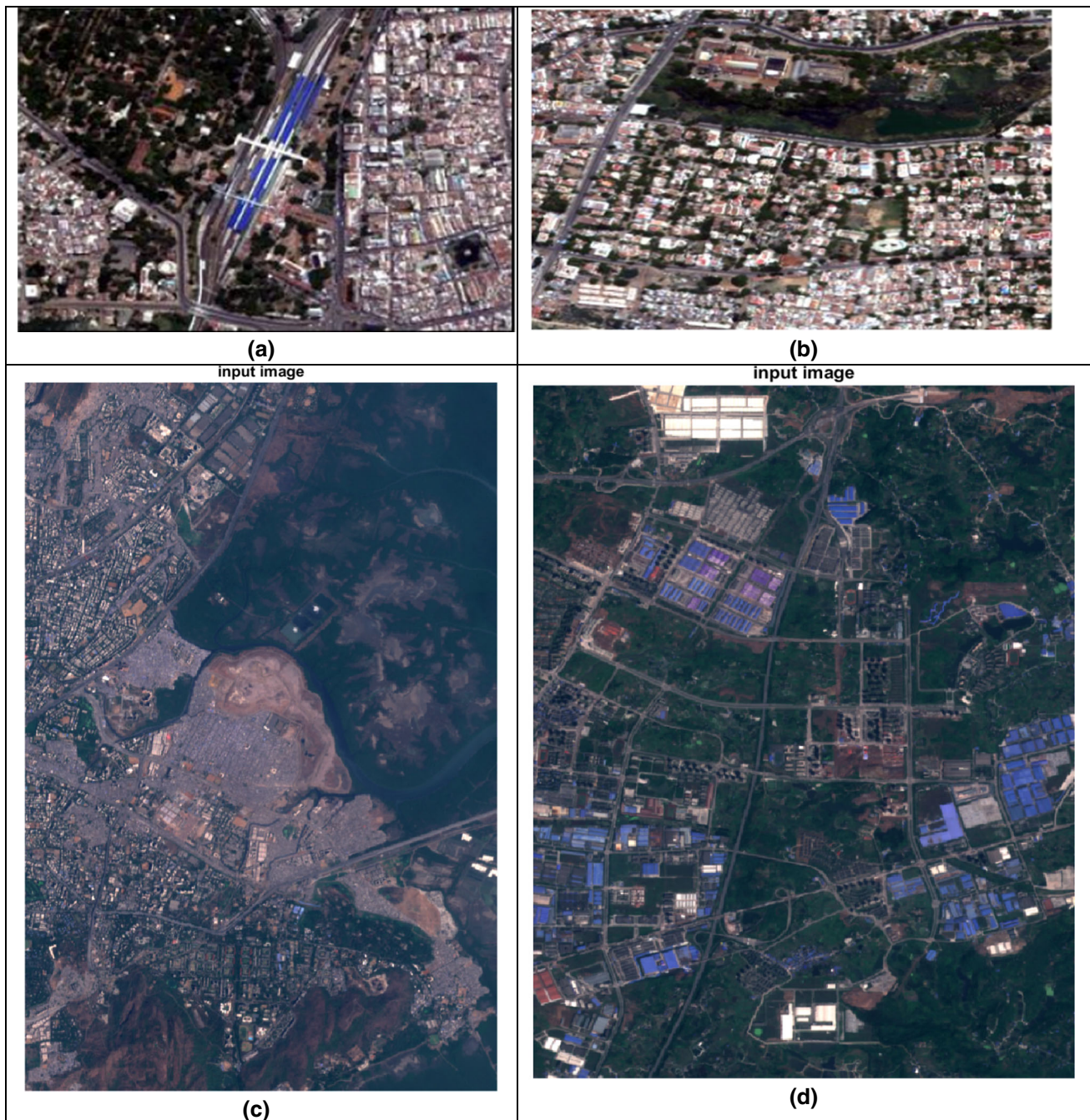
The computed MBI value will be high, if the density of buildings is high. In order to extract buildings, the MBI feature image is then binarized by using a preset threshold. The identified probable building rooftops might include both mislabeled rooftop and perfect rooftop. The mislabeled rooftops are due to open areas, vegetation and roads being wrongly identified as buildings.

The reason for mislabeling of buildings is the presence of shadows caused by trees and adjacent buildings. Hence, in an attempt to reduce the number of mislabeled rooftops, Otsu thresholding combined with a vegetation mask is incorporated. To isolate the shadow regions from the buildings, global thresholding technique using Otsu's threshold is employed as it is suitable for images with non-uniform lighting conditions. To isolate tree regions from buildings, Normalized Difference Vegetation Index (NDVI) is employed. In order to remove the vegetation in the test image (RGB), the green channel matrix value is compared with both red and blue channel matrices to highlight the green region. This image is inverted and then multiplied pixelwise with the input binary image to eradicate vegetation from the test image. NDVI is calculated as given in Eq. (2)

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} \quad (2)$$

where NIR refers near-infrared spectrum reflection and RED refers red spectrum reflection.

The refined MBI image after shadow removal and vegetation removal is given as an input to canny edge detector in order to extract the building rooftops (segments)

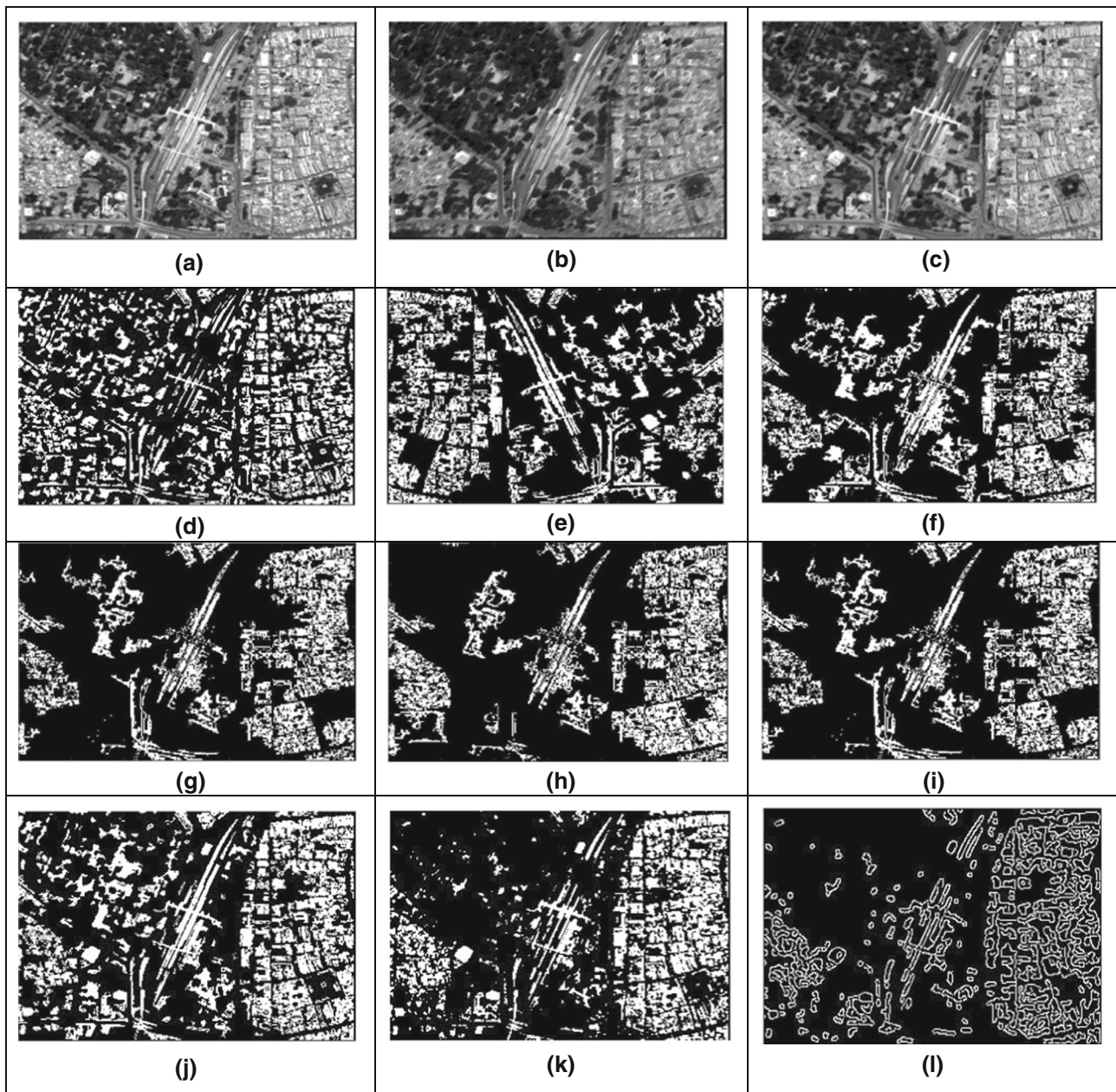


**Fig. 2** Study Sites **a** Study site 1. **b** Study site 2. **c** Study site 3. **d** Study site 4

from the image. Canny edge detection is implemented since it employs edge thinning process by means of non-maximum suppression method and double thresholding for efficient detection of edges. Building rooftops (segments) are then extracted and labeled using connected components regardless of the shape of buildings.

### Feature Extraction

The standard geometrical features such as area, perimeter, convex area, solidity, centroid, major axis length and minor axis length of each labeled segment are extracted. Feature vector of each identified building rooftop is created based on these extracted features. A set of these vectors termed as feature set is then built to train the classifier to classify the identified building rooftops (segments) into perfect



**Fig. 3** Illustration keysteps of MBION-SVM on study site 1. (a) Pre-processed image. **b** Eroded image. **c** Open by reconstruction. Granulometry output with **d** SE (5,10). **e** SE (10,15). **f** SE (15,20).

**g** SE (20,25). **h** SE (25,30). **i** Building map after MBI. Image after **j** vegetation removal. **k** Shadow removal. **l** Output of image after refinement of MBI

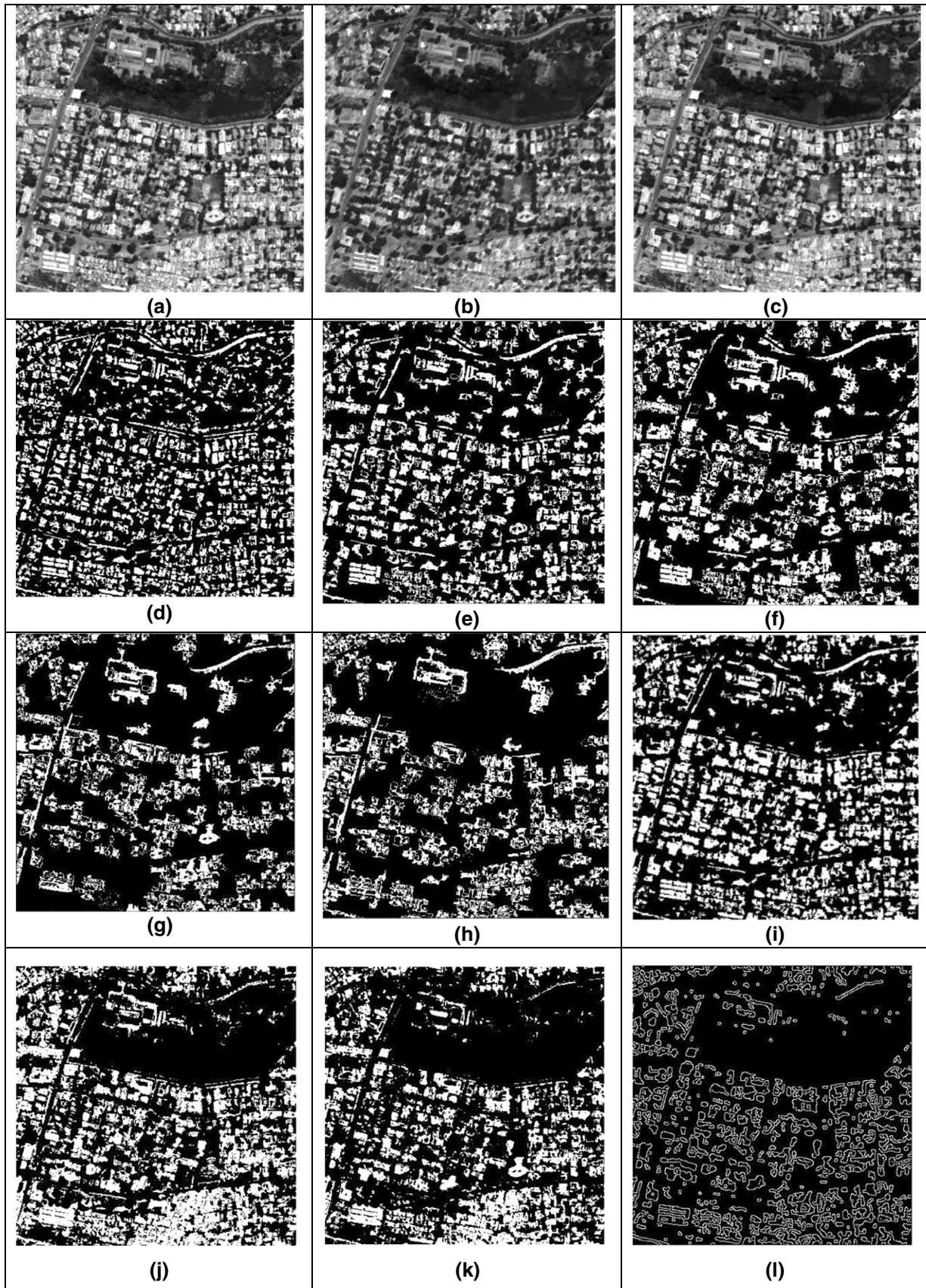
rooftops and mislabeled rooftops. The geometrical features extracted are described as follows:

- Area—specifies the number of pixels in the segmented area of the reference image
- Perimeter—specifies the number of boundary pixels in the segmented area of the reference image
- Centroid—specifies center of mass of the segmented region. The first and second elements of centroid are the horizontal and vertical coordinates.
- Solidity—specifies the (area/convex area).

- Major axis length—specifies the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region.
- Minor axis length—specifies the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region

Once the features are extracted, it is essential to classify the identified building rooftops into buildings and non-buildings. The building rooftops are classified using linear discriminant analysis, logistic regression and SVM.





◀**Fig. 4** Illustration of keysteps of MBION-SVM on study site 2 **a** Pre-processed image. **b** Eroded image. **c** Open by reconstruction. Granulometry output with **d** SE (5,10). **e** SE (10,15). **f** SE (15,20). **g** SE (20,25). **h** SE (25,30). **i** Building map after MBI. Image after **j** vegetation removal. **k** Shadow removal. **l** Output of image after refinement of MBI

## Results and Discussion

The proposed system is simulated in MATLAB1 and tested on a PC (Intel core 2 2.4 GHz with 2 GB RAM). The proposed MBION-SVM methodology is applied to the images taken from two datasets.

### Data Set-I

Worldview-2 is the first high-resolution 8-band multi-spectral commercial satellite launched by Digital Globe on October 8, 2009, to acquire remote-sensed imagery. Its data collection capacity is 975 K km<sup>2</sup>/day with an average revisiting time of 1.1 days around the earth. The sensors for this satellite image have eight spectral bands (coastal blue (400–450 nm), blue (450–510 nm), green (510–580 nm), yellow (585–625 nm), red (630–690 nm), red-edge (705–745 nm), NIR1 (770–895 nm), NIR2 (860–1040 nm)). The size of the image is 407 × 659. Two study sites of Madurai city taken from WorldView-2 satellite as shown in Fig. 2a, b are chosen for analysis of the methodology. The spatial resolution of the image is 1.84 m. The image of Madurai city was acquired on March 17, 2011.

### Data Set-II

The Copernicus Sentinel-2 satellite was launched by Vega VV05 rocket from Europe's spaceport near Kourou in French Guiana on June 23, 2015. These provide 13 band multispectral images with spatial resolutions ranging from 10 and 60 m in the visible, near-infrared (VNIR) and short-wave infrared (SWIR) spectral zones. Its revisiting time is 5 days at the equator and 2–3 days at mid-latitudes. The central wavelengths of 13 spectral bands ultra blue (coastal and aerosol)-443 nm, blue-490 nm, green-560 nm, red-665 nm, visible, near-infrared (VNIR)-(705–865 nm) and short-wave infrared (SWIR) (940–2190 nm) (Daudt et al. 2018). The portion of Mumbai image shown in Fig. 2c and Chongqing image shown in Fig. 2d taken from the Onera Satellite Change detection dataset compiled from Sentinel-2 are taken as the second image set for the analysis of the methodology. The sentinel sat API was used to ensure minimum cloud cover. The Mumbai image was captured on March 19, 2018. The size of Mumbai image is

858 × 557 × 13. The Chongqing image was captured on April 2, 2018, and its size is 730 × 544 × 13. Medusa toolbox2 was used to download and crop the images.

The pre-processing done in this work is converting the image under study into grayscale image by recording the maximum intensity value of R, G and B bands. The pre-processed image is then eroded using linear structuring element, after which it is dilated using the same structuring element. White top-hat (WTH) transformation is used to perform opening operation whose output image shows up high local diversity features of building candidates.

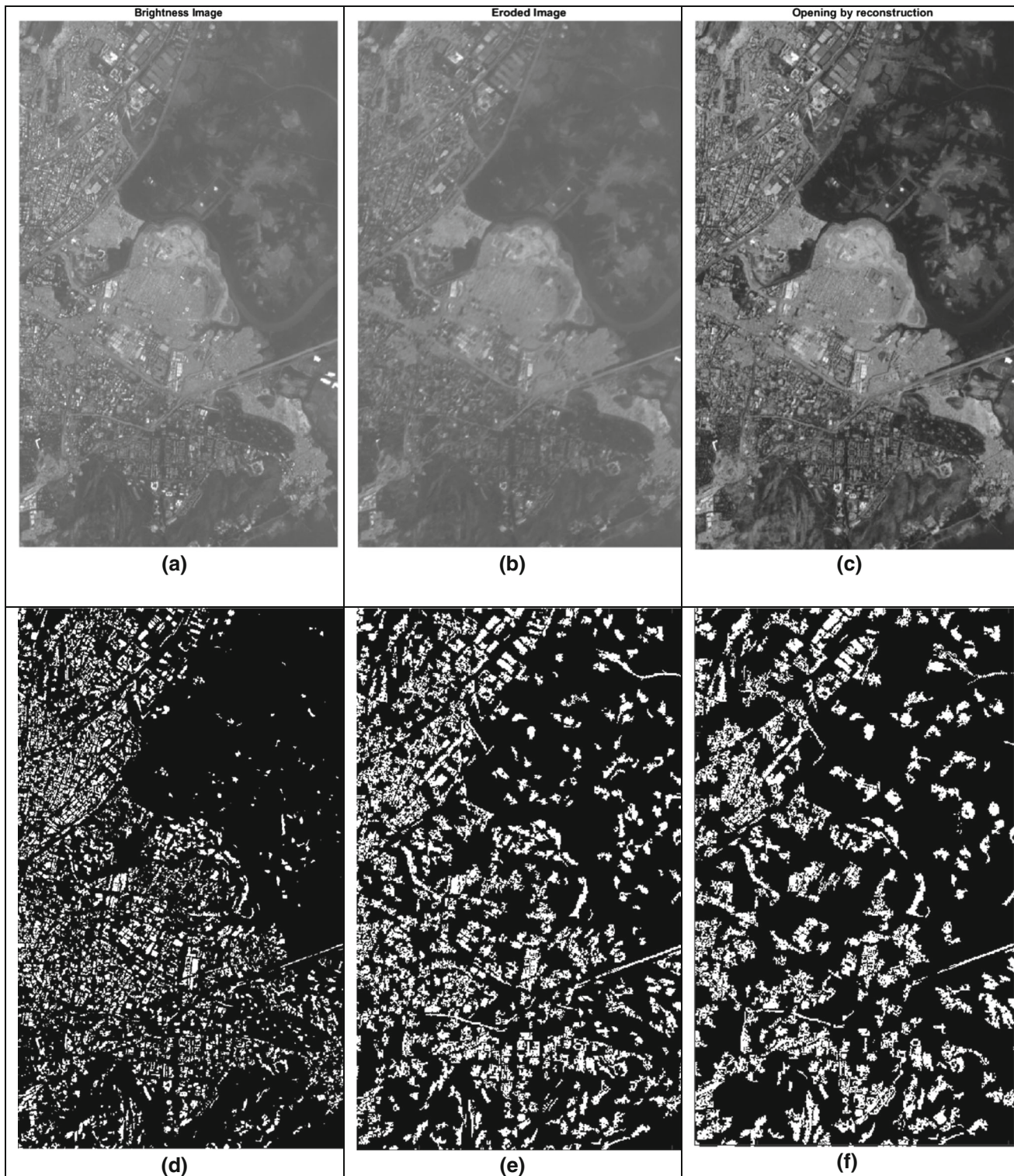
The structural information of the image is extracted by deriving morphological profiles. Based on the WTH transformation with various scale and directionalities, the differential morphological contours are calculated to indicate the complicated structural arrangement of buildings in different scales and directions. In this work, granulometry outputs are computed for different structuring element (SE) values such as (5, 10), (10, 15), (15, 20), (20, 25) and (25, 30). The directional threshold is finally calculated, by taking the average of thresholds obtained using different SE values. A large MBI value signifies the presence of a building structure. This value is considered to be the preset threshold for probable identification of building candidates. The result of identified building candidates includes mis-labeled building candidates, missing building candidates and correct building candidates. The reason for mislabeling of buildings is due to the presence of shadows, caused by trees and adjacent buildings. To isolate the shadow regions from the buildings, global thresholding technique using Otsu's threshold is employed as it is suitable for images with non-uniform lighting conditions. To isolate tree regions from buildings, NDVI is employed. In order to extract the building rooftops (segments) from the image, Canny edge detection is implemented since it employs edge thinning process by means of non-maximum suppression method and double thresholding for efficient detection of edges. The building rooftops (segments) extracted from the input image are then labeled using connected component labeling (CCL). Thus, the outlines of the buildings are extracted regardless of their shape. The key steps of MBION-SVM performed on study site 1, study site 2, study site 3 and study site 4 are illustrated in Figs. 3, 4, 5 and 6, respectively.

## Classification

Supervised classification algorithm is used to classify every extracted segment as either a rooftop (building) or non-rooftop. Classification algorithms such as logistic regression, linear discriminant analysis and SVM are implemented. The performance of the classifiers is analyzed and

compared using correlation plot, parallel coordinates plot, ROC curve and confusion matrix.

Correlation plot provides a two-dimensional data visualization of any two variables and impact of variation of one variable on another variable. Correlation plot shown in



**Fig. 5** Illustration of keysteps of MBION-SVM on study site 3 **a** Pre-processed image. **b** Eroded image. **c** Open by reconstruction. Granulometry output with **d** SE (5,10). **e** SE (10,15). **f** SE (15,20).

**g** SE (20,25). **h** SE (25,30). **i** Building map after MBI. Image after **j** vegetation removal. **k** Shadow removal. **l** Output of image after refinement of MBI

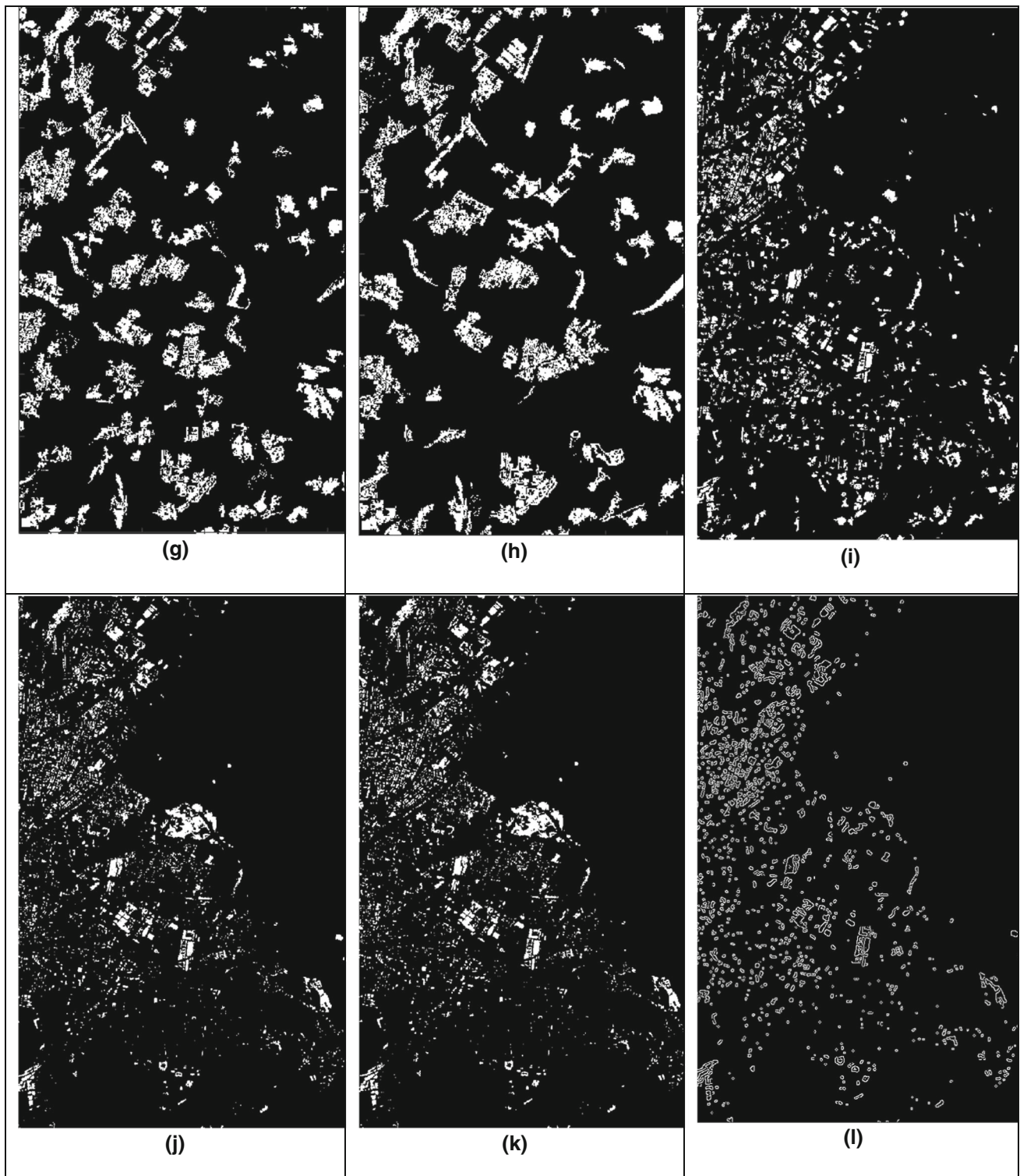
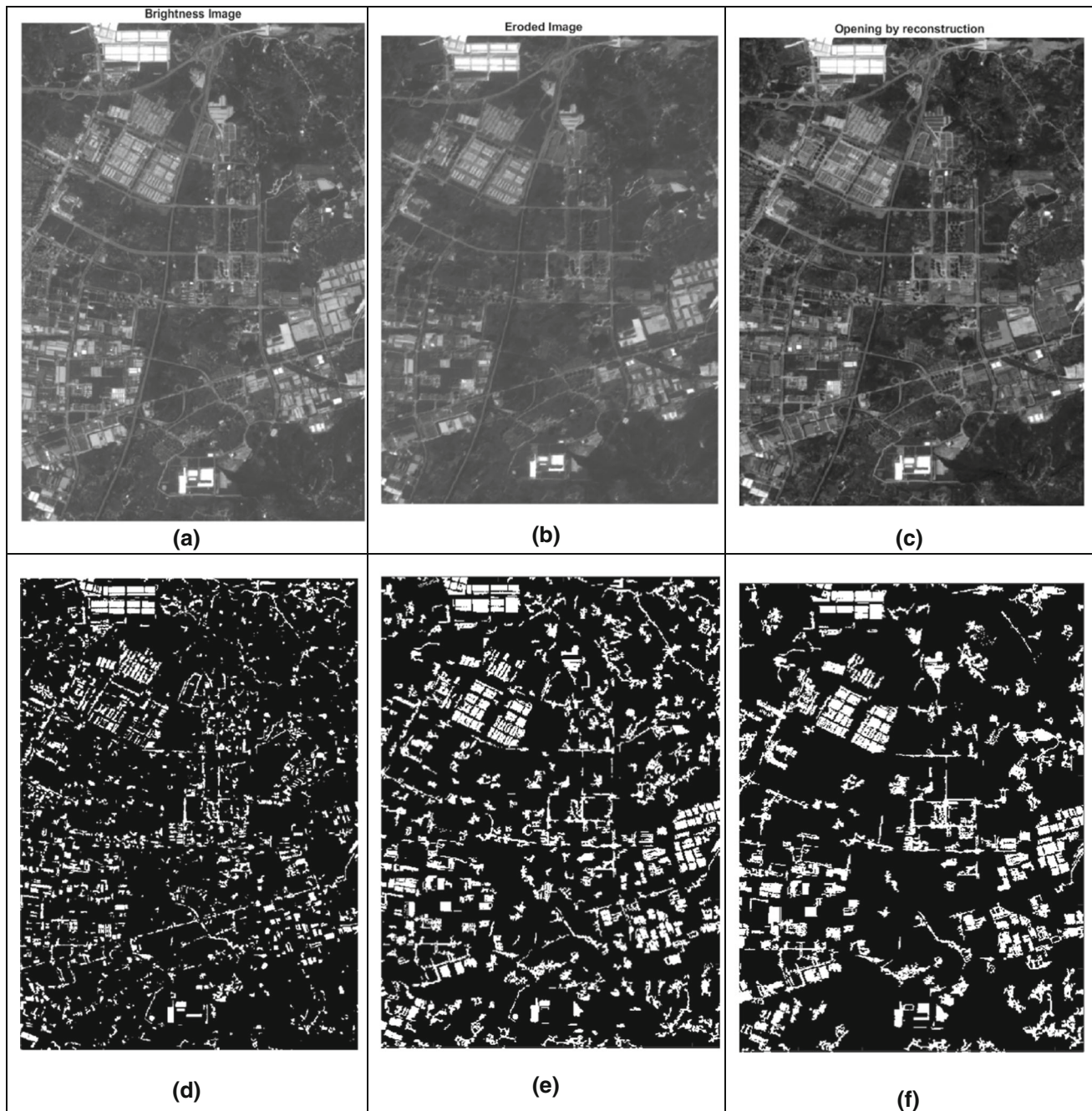


Fig. 5 continued

Fig. 7 depicts the correlation between the variables (area and major axis length) of a set of buildings. Blue color indicates non-rooftop, and orange color indicates rooftop.

A parallel plot shown in Fig. 8 is used to identify the significance of the extracted features of buildings even though the variables are completely different in terms of range and units. The variables area, perimeter, centroid,



**Fig. 6** Illustration of keysteps of MBION-SVM on study site 4 **a** Pre-processed image. **b** Eroded image. **c** Open by reconstruction. Granulometry output with **d** SE (5,10). **e** SE (10,15). **f** SE (15,20).

**g** SE (20,25). **h** SE (25,30). **i** Building map after MBI. Image after **j** vegetation removal. **k** Shadow removal. **l** Output of image after refinement of MBI

solidity, major axis length and minor axis length are shown as clusters, and each cluster is represented in a different color. A specified range of each variable is analyzed. Based on the analysis, area and major axis length are chosen for performing self-correction. In general, mislabeled rooftops, such as barren land, bridges and roads, generally have high major axis length or occupy large area, whereas very small objects like vehicles occupy very small space on the

ground. Though urban area buildings generally occupy sizeable area, they do not occupy very large area. Based on the features of common building candidates in this study area, a suitable threshold range for area (1–73 pixels) and major axis length (1.154701–29.84843 pixels) have been set, to eradicate the mislabeled rooftops. The resolution of image and type of the area (residential or commercial or industrial) present in the study area of the image vary the

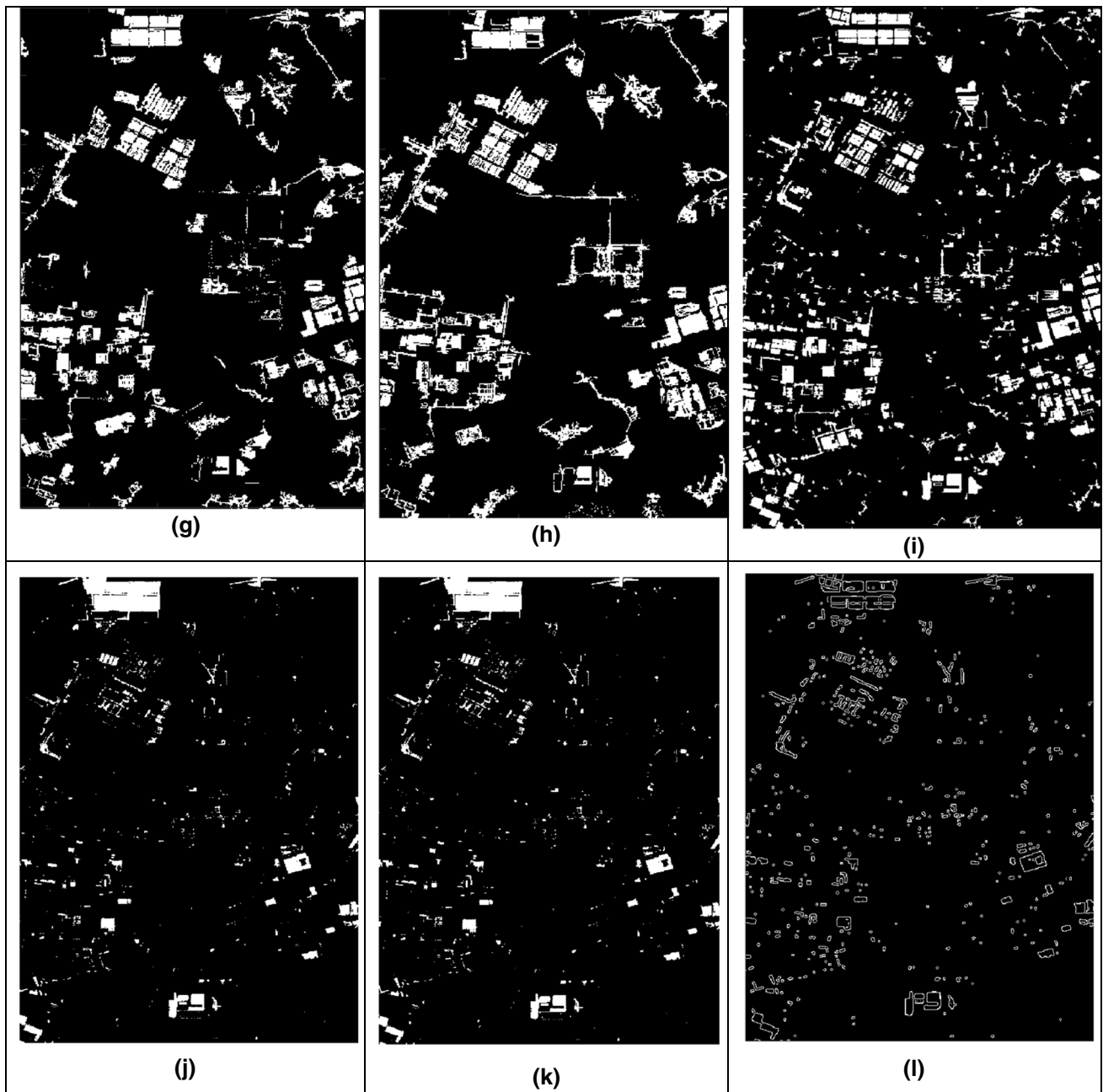
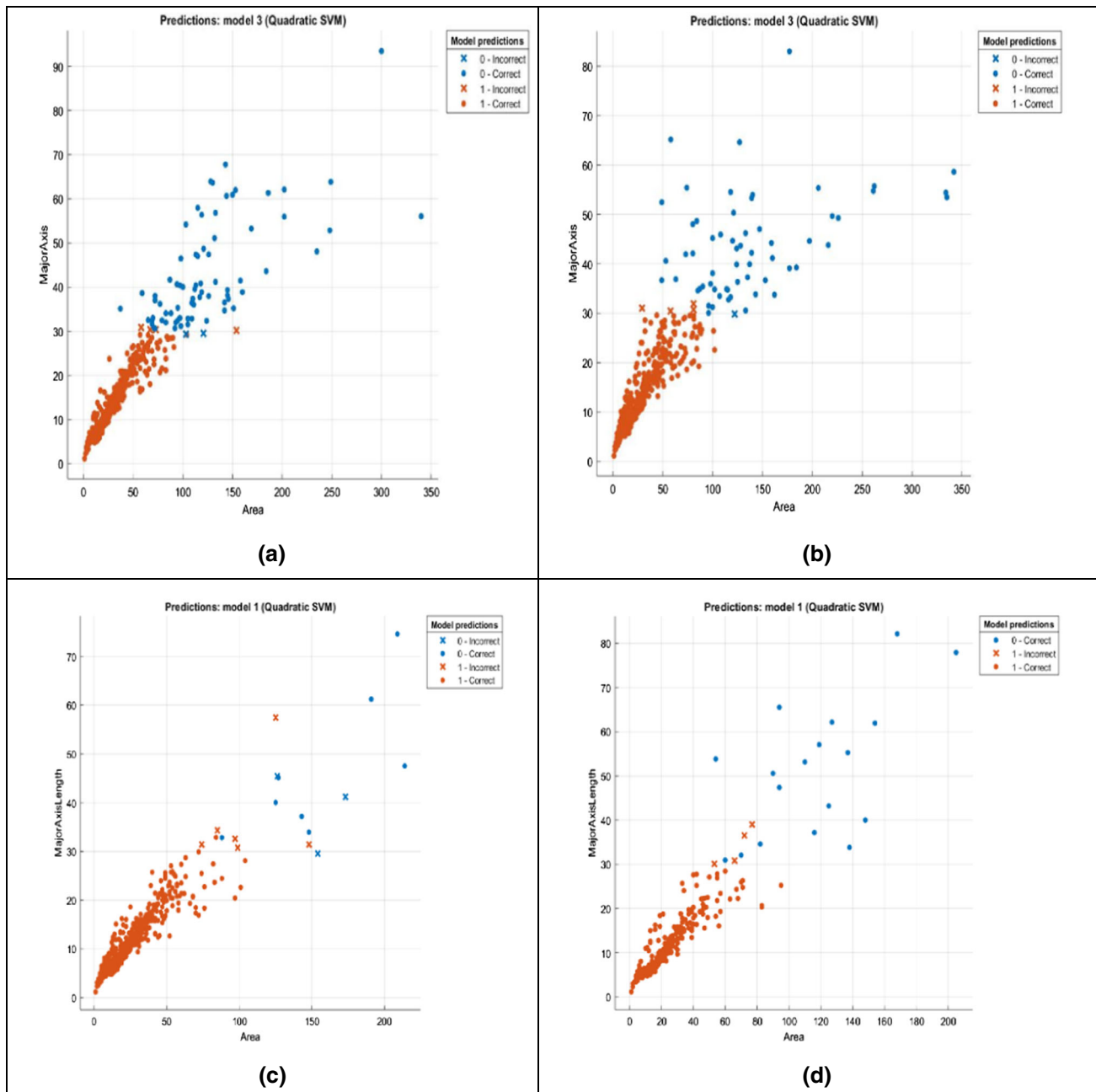


Fig. 6 continued

range of the threshold value. The process of self-correction is therefore carried out to remove mislabeled roof tops further. It is seen from receiver operating characteristic (ROC) plot of SVM shown in Fig. 9 proving that its performance is good with an area under curve (AUC) of 1.

Table 1 shows the confusion matrix of logistic regression, linear discriminant analysis and support vector machines for all the four sites under study. In case of study site 1, it is observed that in LDA, out of 611 rooftops, 3 rooftops have been missed and 10 non-rooftops have been detected falsely, whereas in logistic regression, 5 rooftops

have been missed and 1 non-rooftop has been detected falsely. In SVM classifier, 2 rooftops have been missed and 4 non-rooftops have been detected falsely. In case of study site 2, it is observed that in LDA, out of 534 rooftops, 5 rooftops have been missed and 3 non-rooftops have been detected falsely, whereas in logistic regression, 3 rooftops have been missed and 4 non-rooftops have been detected falsely. In SVM classifier, only 1 rooftop has been missed and only 4 non-rooftops have been detected falsely. In case of study site 3, it is observed that in LDA, out of 1012 rooftops, 4 rooftops have been missed and 12 non-rooftops



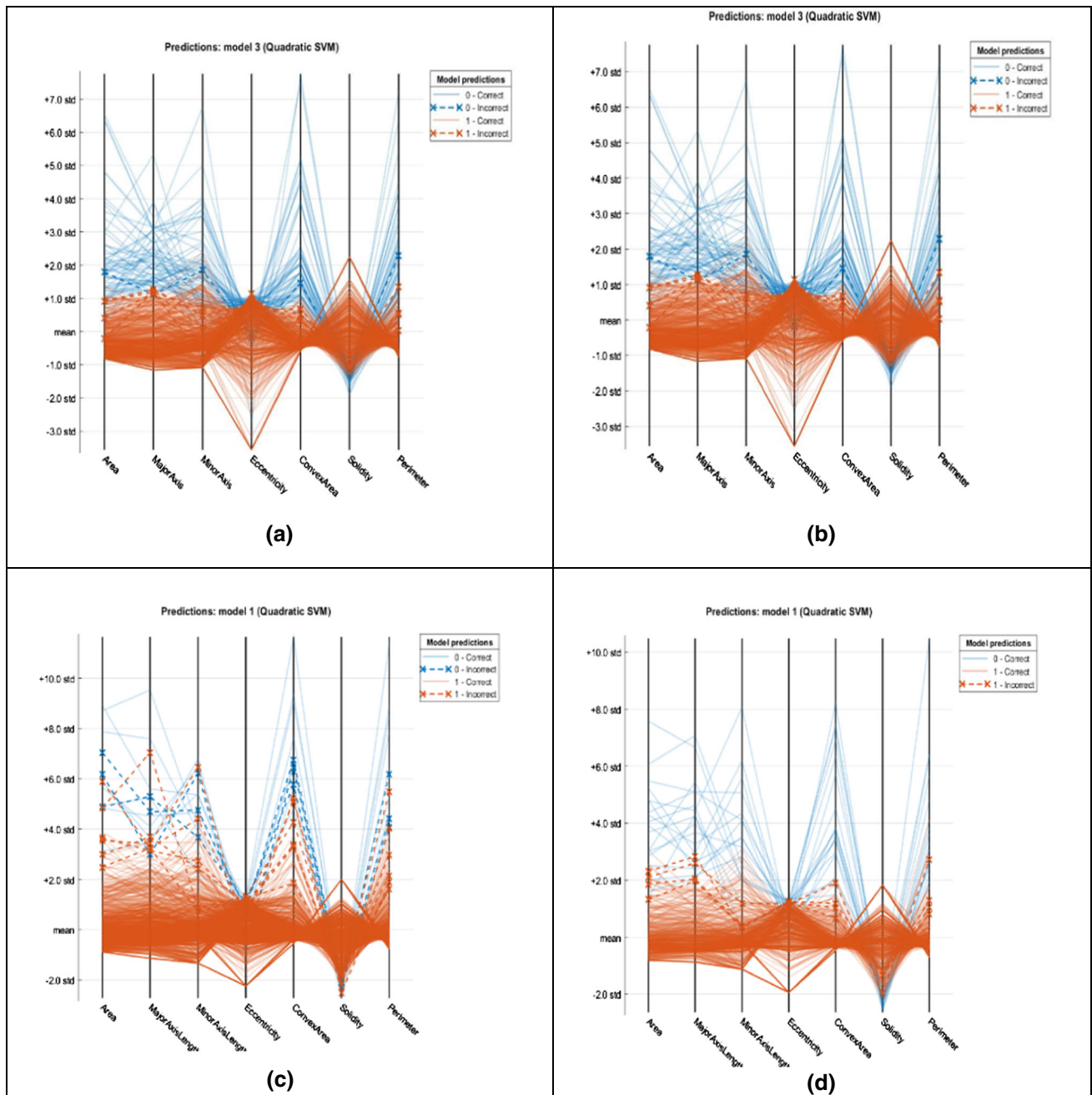
**Fig. 7** Correlation plot of SVM classifier for **a** Study site 1. **b** Study site 2. **c** Study site 3. **d** Study site 4

have been detected falsely, whereas in logistic regression, 12 rooftops have been missed and 5 non-rooftops have been detected falsely. In SVM classifier, 2 rooftops have been missed and 7 non-rooftops have been detected falsely. In case of study site 4, it is observed that in LDA, out of 491 rooftops, 20 non-rooftops have been detected falsely, whereas in logistic regression, 20 non-rooftops have been detected falsely. In SVM classifier, 16 non-rooftops have been detected falsely. Thus, the proposed method using SVM helps in reducing false detection to a great extent.

In order to quantitatively evaluate the proposed work, the number of extracted buildings is manually labeled and these labeled buildings are treated as ground truth. The standard four metrics such as Shape Accuracy, Precision P (correctness), Recall R (completeness) and F-score F (Overall quality) are evaluated to find the performance of the proposed method of building extraction.

$$\text{Shape Accuracy} = 1 - \frac{|A_{GT} - A_S|}{A_{GT}} \tag{3}$$

where



**Fig. 8** Parallel plot of SVM classifier for **a** Study site 1. **b** Study site 2. **c** Study site 3. **d** Study site 4

$A_{GT}$ —area of buildings from the ground truth,  $A_S$ —area of buildings from the extraction results

$$\text{Precision } (P) = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{Fscore } (F) = \frac{TP}{TP + FP + FN} \tag{6}$$

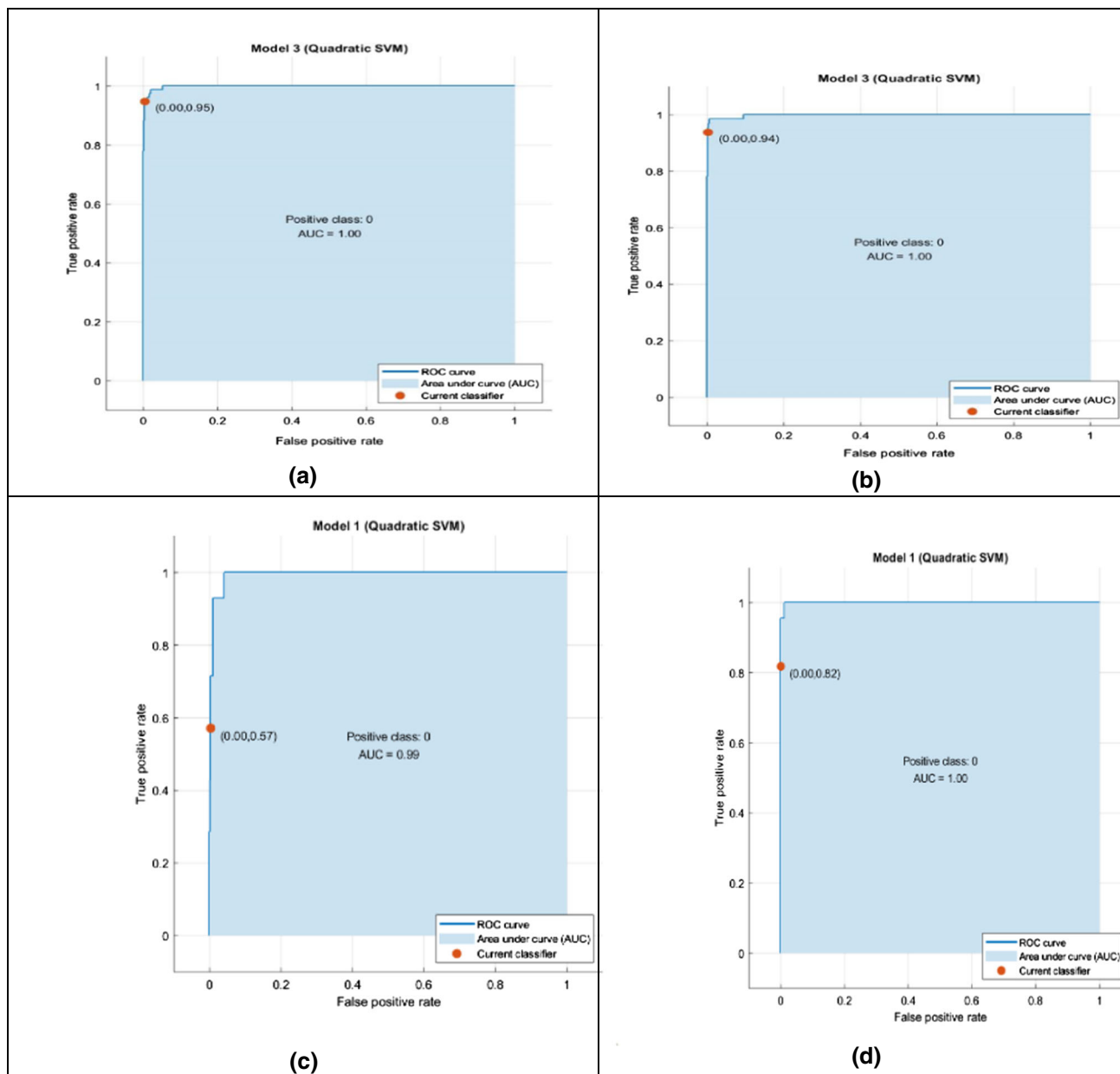
where

TP—true-positive indicates the number of pixels assigned as buildings in both segmentation and ground truth results.

FP—false positive indicates the number of pixels assigned as buildings in segmentation result but not in ground truth

FN—false negative indicates the number of pixels assigned as buildings in ground truth but not in segmentation result.





**Fig. 9** Receiver operating characteristic (ROC) plot of SVM for **a** Study site 1. **b** Study site 2. **c** Study site 3. **d** Study site 4

The performance of the classifiers is presented in Table 2.

It is well known that the optimal values for the four metrics are 1. F-score values cannot be higher than neither P nor R. The proposed MBION-SVM method gives an accuracy, precision, recall and F-score for study site 1 as 99%, 88%, 99% and 88%, respectively. These values for study site 2 are 99.1%, 89%, 100% and 88%, respectively. These values for study site 3 are 99.1%, 99%, 99% and 99%, respectively. These values for study site 4 are 99.6%, 99%, 100% and 99%, respectively. With the good results

achieved in classifying, the work can be adopted to suit any applications.

### Conclusion

In this paper, we propose MBION-SVM method that employs morphological, spectral, shape and geometrical features to identify building rooftops in Very High Resolution (VHR) images efficiently. The morphological features are used to extract structural information of the satellite imagery through the computation of

**Table 1** Confusion matrix

|                            | Study site 1 |              | Study site 2 |              | Study site 3 |              | Study site 4 |              |
|----------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|                            | Rooftops     | Non-rooftops | Rooftops     | Non-rooftops | Rooftops     | Non-rooftops | Rooftops     | Non-rooftops |
| <i>LDA</i>                 |              |              |              |              |              |              |              |              |
| Rooftops                   | 530          | 3            | 460          | 5            | 1008         | 4            | 491          | 0            |
| Non-rooftops               | 10           | 67           | 3            | 66           | 12           | 2            | 20           | 2            |
| <i>Logistic regression</i> |              |              |              |              |              |              |              |              |
| Rooftops                   | 528          | 5            | 467          | 3            | 1000         | 12           | 491          | 0            |
| Non-rooftops               | 1            | 76           | 4            | 60           | 5            | 9            | 20           | 2            |
| <i>Quadratic SVM</i>       |              |              |              |              |              |              |              |              |
| Rooftops                   | 531          | 2            | 469          | 1            | 1010         | 2            | 491          | 0            |
| Non-rooftops               | 4            | 73           | 4            | 60           | 7            | 7            | 16           | 6            |

**Table 2** Performance metrics of classifiers

| Images       | Classifier                   | Accuracy | Precision | Recall | F-score |
|--------------|------------------------------|----------|-----------|--------|---------|
| Study site 1 | Linear discriminant analysis | 97.9%    | 0.89      | 0.99   | 0.88    |
|              | Logistic regression          | 99%      | 0.87      | 0.99   | 0.87    |
|              | Quadratic SVM                | 99%      | 0.88      | 1      | 0.88    |
| Study site 2 | Linear discriminant analysis | 96.6%    | 0.87      | 1      | 0.87    |
|              | Logistic regression          | 98.7%    | 0.89      | 0.99   | 0.88    |
|              | Quadratic SVM                | 99.1%    | 0.89      | 1      | 0.88    |
| Study site 3 | Linear discriminant analysis | 98.4%    | 0.98      | 0.99   | 0.98    |
|              | Logistic regression          | 98.3%    | 0.99      | 0.98   | 0.98    |
|              | Quadratic SVM                | 99.1%    | 0.99      | 0.99   | 0.99    |
| Study site 4 | Linear discriminant analysis | 98.8%    | 0.98      | 1      | 0.98    |
|              | Logistic regression          | 98.8%    | 0.98      | 1      | 0.98    |
|              | Quadratic SVM                | 99.6%    | 0.99      | 1      | 0.99    |

Morphological Building Index (MBI). Following which, Otsu's thresholding and Normalized Differential Vegetation Index (NDVI) are used to refine the output of Morphological Building Index (MBI) by removing shadows and vegetations, respectively, and also to extract possible rooftops. Based on the analysis of geometrical features of the possible rooftops, threshold values are fixed for performing self-correction to minimize false detection of rooftops. In this work, we have proved the performance of our proposed methodology on two images of Madurai city, Tamilnadu, South India, acquired by WorldView-2 satellite and on two study images acquired by Sentinel-2 satellite of which, one is a portion of Mumbai, the Metropolitan city of Maharashtra state, India, and another is a portion of Chongqing, megacity in southwest China. In this work, we have analyzed the performance of building extraction approach with classification algorithms such as linear discriminant analysis, logistic regression and SVM. The proposed method gives an accuracy of 99%, precision of 89%, a perfect Recall of 1 and a F-score of 88%. The good

performance analysis achieved in classification shows that the work can be adopted to suit any application.

However, we have used four images here for evaluating our methodology of building extraction approach. The classifiers can be trained with more number of images to improve and generalize the performance of the proposed methodology. Future work can be focused on integrating some more spectral indices to further improve the performance of the building extraction approach. The work can be extended on multiclass classification to further classify the non-rooftops into vegetation, water bodies, barren land samples and the rooftops into single-storey and multi-storey buildings.

## References

- Andrea, M.-K., & Ali, O. (2015). Building detection from monocular vhr images by integrated urban area knowledge. *IEEE Geoscience and Remote Sensing Letters*, 12(10), 2140–2144.

- Attarzadeh, R., & Momeni, M. (2017). Object-based rule sets and its transferability for building extraction from high resolution satellite imagery. *Journal of the Indian Society of Remote Sensing*, 46, 169–178.
- Attarzadeh, R., & Momeni, M. (2018). Object-based rule sets and its transferability for building extraction from high resolution satellite imagery. *Journal of the Indian Society of Remote Sensing*, 46, 169–178.
- Bavaghar, P. M. (2015). Deforestation modelling using logistic regression and GIS. *Journal For Science*, 61, 193–199.
- Benediktsson, J. A., Pesaresi, M., & Amason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9), 1940–1949.
- Blaschke, T. (2003). Object-based contextual image classification built on image segmentation. In IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, October 27–28, 2003, Greenbelt, MD, USA, pp. 113–119.
- Bruzzone, L., & Carlin, L. (2006). A multilevel context-based system for classification of very high spatial resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9), 2587–2600.
- Daudt, R.C., Le Saux, B., Boulch, A. & Gousseau, Y., (2018). Urban change detection for multispectral earth observation using convolutional neural networks. In IEEE International Geoscience and Remote Sensing Symposium (IGARSS). July 22–27, 2018, Valencia, Spain. pp. 2115–2118
- Dayana, S., & Avudaiammal, R. (2018). Automatic building extraction from VHR satellite image and denoising with various filters. *International Journal of Pure and Applied Mathematics*, 118(11), 121–128.
- Dutta, D., & Sarkar, N. M. K. (2005). Urban building inventory development using very high resolution remote sensing data for urban risk analysis. *International Journal of Geoinformatics*, 1(1), 109–116.
- Fauvel, M., Benediktsson, J. A., Chanussot, J., & Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11), 3804–3814.
- Gavankar, N., & Ghosh, S. (2018). Automatic building footprint extraction from high-resolution satellite image using mathematical morphology. *European Journal of Remote Sensing*, 51, 182–193.
- Holloway, J., & Mengerser, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sensing MDPI*, 10, 1365.
- Hu, J., Chen, W., Li, X., & He, X. (2008). Roof confusion removal for accurate vegetation extraction in the urban environment. In International workshop on earth observation and remote sensing Applications. 30 June–2 July, 2008, Beijing, China, pp. 1–7.
- Hu, L., Zheng, J., Gao, F. (2011). A building extraction method using shadow in high resolution multispectral images. In IEEE international geoscience and remote sensing symposium, July 24–29, 2011, Vancouver, BC, Canada, pp. 1862–1865.
- Huang, X., Yuan, W., Li, J., & Zhang, L. (2017). A new building extraction postprocessing framework for high-spatial-resolution remote-sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(2), 654–668.
- Huang, X., & Zhang, L. (2011). A multidirectional and multiscale morphological index for automatic building extraction from multispectral GeoEye-1 imagery. *Photogramm. Engineering and Remote Sensing*, 77(7), 721–732.
- Huang, X., & Zhang, L. (2012a). Morphological building/shadow index for building extraction from high-resolution imagery over urban areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, 161–172.
- Huang, X., & Zhang, L. (2012b). Morphological building/shadow index for building extraction from high-resolution imagery over urban areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1), 161–172.
- Huang, X., Zhang, L., & Li, P. (2007). Classification and extraction of spatial features in urban areas using high resolution multispectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 4(2), 260–264.
- Inglada, J. (2007). Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. *ISPRS-J. Photogramm. Remote Sens.*, 62, 236–248.
- Kumar, A., Pandey, A., & Jeyaseelan, A. (2012). Built-up and vegetation extraction and density mapping using WorldView-II. *Geocarto International*, 27, 557–568.
- Lee, S. D., Shan, J., & Bethel, J. S. (2003). Class-guided building extraction from IKONOS imagery. *Photogrammetric Engineering and Remote Sensing*, 69(2), 143–150.
- Lefèvre, S., Weber, J., Sheeren, D. (2007). Automatic building extraction in VHR images using advanced morphological operators. In Proceedings of IEEE/ISPRS joint workshop on remote sensing and data fusion over urban areas (URBAN). April 11–13, 2007. Paris, France. pp.1 - 5.
- Liu, Z., Cui, S., Yan, Q. (2008). Building extraction from high resolution satellite imagery based on multi-scale image segmentation and model matching. international workshop on earth observation and remote sensing applications, pp. 1–7
- Mathur, A., & Foody, G. M. (2008). Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, 29, 2227–2240.
- Myint, S. W., Lam, N. S. N., & Tyler, J. (2004). Wavelets for urban spatial feature discrimination: Comparisons with fractal, spatial autocorrelation, and spatial co-occurrence approaches. *Photogramm. Eng. Remote Sens.*, 70(7), 803–812.
- Ok, A. O. (2013). Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts. *ISPRS Journal of Photogrammetry*, 86, 21–40.
- Ok, A. O., Senaras, C., & Yuksel, B. (2013). Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3), 701–1717.
- Ouma, Y. O., Ngigi, T. G., & Tateishi, R. (2006). On the optimization and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application toward suburban-tree delineation. *International Journal of Remote Sensing*, 27(1), 73–104.
- Pesaresi, M., & Benediktsson, J. A. (2001). A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2), 309–320.
- Senaras, C., Ozay, M., & Yarman Vural, F. T. (2013). Building detection with decision fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 295–1304.
- Senaras, C., & Yarman Vural, F. T. (2016). A self-supervised decision fusion framework for building detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1780–1791.
- Shi, W., Mao, Z., & Liu, J. (2018). Building extraction from high-resolution remotely sensed imagery based on multi-subgraph matching. *The Journal of the Indian Society of Remote Sensing*, 46, 2003–2013.

- Sirmacek, B., & Unsalan, C. (2009a). Urban-area and building detection using SIFT keypoints and graph theory. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4), 1156–1167.
- Sirmacek, B., & Unsalan, C. (2009b). Urban-area and building detection using SIFT keypoints and graph theory. *IEEE Transactions on Geoscience and Remote Sensing*, 47(4), 1156–1167.
- Szuster, B. W., Chen, Q., & Borger, M. A. (2011). Comparison of Classification techniques to support land cover and land use analysis in tropical coastal zones. *Applied Geography*, 31, 525–532.
- Tarantino, C., Adamo, M., Pasquariello, G., Lovergine, F., Blonda, P., & Tomaselli, V. (2012) 8-Band Image Data Processing of the Worldview-2 Satellite in a Wide Area of Applications. <https://doi.org/10.5772/27499>.
- Wei, Y., Zhao, Z., & Song, J. (2004). Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*. September 20–24, 2004, (Vol. 3, pp. 2008–2010). Anchorage, AK, USA.
- Xiaoying, J., & Davis, C. H. (2005). Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. *EURASIP Journal on Advances in Signal Processing*. <https://doi.org/10.1155/ASP.2005.2196>.
- Xin, H., Xiaopeng, H., Liangpei, Z., Jianya, G., Wenzhi, L., Jon, B. (2016). Generalized differential morphological profiles for remote sensing image classification. In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. pp. 1–34.
- Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing*, 29(14), 4269–4276.
- Yang, J., Wang, Y.-H. (2012) Towards automatic building extraction: Variational level set model using prior shape knowledge. In *International conference on image analysis and signal processing*, pp. 1–6.
- You, Y., Wang, S., Ma, Y., Chen, G., Wang, B., Shen, M., et al. (2018). Building Detection from VHR Remote Sensing Imagery Based on the Morphological Building Index. *Remote. Sens.*, 10, 1287.
- Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24, 583–594.
- Zhe, D., Wang, X. Q., Li, Y. L., & Zhang, S. S. (2018). Study on building extraction from high resolution images using MBI. *International Archives of the Photogrammetry, Remote Sensing and Spatial and Spatial Information Sciences.*, 42(3), 283–287.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.