Building Footprint Extraction from Very-High-Resolution Satellite Image Using Object-Based Image Analysis (OBIA) Technique

A. P. Prathiba, Kriti Rastogi, Gaurav V. Jain and V. V. Govind Kumar

Abstract Building footprints are an important input in several urban applications such as master plan preparation, development of 3D city building models, rooftop solar energy potential estimation, tax compliance evaluation, or study of population distribution in cities. The very high spatial resolution (VHR) image is invariably required for the extraction of building footprints. However, the conventional pixel-based approaches have limited success in building footprint extraction owing to inherent heterogeneity of the urban environment. In this research, we have therefore applied Object-Based Image Analysis (OBIA) for building footprints extraction from Cartosat-2 series data. The image was segmented into several objects on the basis of spectral and spatial homogeneity of pixels. The objects were thereafter classified using nearest-neighbor classification approach. Finally, these classified objects were further subjected to segmentation into smaller objects and classified using decision-rules. The combination of supervised nearest-neighbor classification with decision-rules resulted in an accuracy of over 82.5% in the extraction of building footprints. The results of the study will be used to develop a 3D city model of Ahmedabad city and assess the changes in the built-up volume in the city.

Keywords Building footprint extraction · OBIA · Nearest-neighbor classification

A. P. Prathiba (✉) · V. V. Govind Kumar
Indian Institute of Technology (Indian School of Mines) Dhanbad, Dhanbad, Jharkand, India
e-mail: prathi.gem@gmail.com

V. V. Govind Kumar
e-mail: vgkvilluri@iitism.ac.in

K. Rastogi · G. V. Jain
Space Application Centre (ISRO), Ahmedabad, Gujarat, India
e-mail: kritirastogi@sac.isro.gov.in

G. V. Jain
e-mail: gvj@sac.isro.gov.in

1 Introduction

The urban settlements cover about 0.5% of the earth's land surface area, but it holds more than half of the world's population [1]. The transition from rural to urban indicates the shift from an agriculture-based economy to the manufacturing and service-oriented economy. As per the study by the United Nations, by 2050 85.9 and 64.1% (source: World Bank record) of the developed and developing world will be urbanized respectively. Figure 1 shows the proportion of urban population in the total population in India from the year 1961–2011.

Urbanization leads to numerous problems like overcrowding, traffic congestion, growth of slums, insufficient resources, etc. To satisfy the basic needs of people and for dynamic growth, advanced technology is needed to be associated with urbanization. Building footprint is one of the recent technology, which can be used as a source for several urban applications like the development of 3D city models, tax compliance evaluation, rooftop solar energy potential estimation, master plan preparation.

Traditionally the aerial photography is used for urban studies owing to its high spatial resolution. With the advancement of technology, commercial satellites capable of imaging at sub-meter spatial resolution are becoming popular and easily accessible. Satellite images give detailed visibility of terrestrial features by reducing the heterogeneity of each pixel, which leads to the shift from aerial to satellite images for urban applications [2]. For urban features classification and information extraction, spatial information is more important than the spectral information. High-resolution satellite images (high spatial information) provide detailed information on the objects.

The main challenge with the very-high-resolution satellite images is a higher level of details. If the resolution of satellite image increases, the information of the finer objects comes to the picture, but it may reduce the accuracy of the digital classification in urban areas. For the effective identification of the object, the pixel size should be
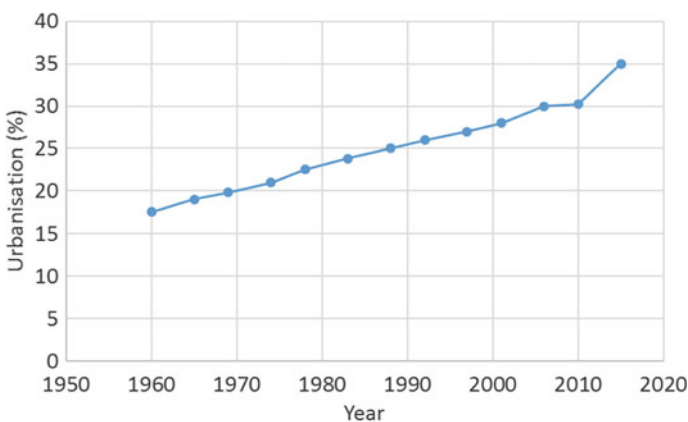
Fig. 1 Urbanization in India

half of the diameter of the smallest object of interest and a minimum four pixels are required for object identification [3]. The precision of the classification can be increased by adding the spectral information to the high-resolution image of the same spatial resolution. Visual interpretation of the urban objects using the multispectral image is advantageous. The advantages of both spectral and spatial resolution can be utilized effectively by fusion of multispectral image with the high-resolution panchromatic image [4].

Each pixel in the image carries the spectral and spatial information of the surface features such as color, texture, shape, etc., but by viewing individual pixels, it is difficult to identify. Object-based image analysis (OBIA) is the technique which is performed on the basis of an object, which human can interpret. OBIA groups the homogeneous pixels to form an object [5]. However, different objects have typical spectral reflectance, which helps the users to identify the object.

Image segmentation segments the image into several clusters by modifying the morphology of objects according to defined criteria, which serves as the framework for further analysis. eCognition provides several segmentation algorithms based on both top-down and bottom-up strategy. In which, multiresolution segmentation algorithm works on bottom-up strategy, based on relative homogeneity it merges the pixel with their neighbor. The main concept behind the multiresolution segmentation in eCognition is Fractal Net Evolution Approach (FNEA) [6]. Homogeneity criteria for multiresolution segmentation are the combination of scale, smoothness and compactness parameter [7]. The parameter for image segmentation is interpreted through the trial and error method. After several experiments, the most suitable parameter for multiresolution segmentation is selected.

The objects are classified into various classes based on hierarchy after completion of segmentation. The main type of classifiers is nearest-neighbor classifier or membership classifier. The standard nearest-neighbor classification is carried out by selecting the training samples of different location and varied reflectance for each classes using visual interpretation. The object-related features like layer value, texture, geometry, position, relation to neighbor objects and user-defined indices are used for the execution of fuzzy-based classification. The appropriate threshold value for each feature should be determined by the spectral reflectance curve of different classes, for accurate classification. After classification, the vector file for each class can be exported with object-related feature value for further analysis.

2 Study Area and Data Used

The study area used for the analysis is Ahmedabad, which is the sixth biggest city and seventh largest metropolitan city of India. The city in the first stage of the proposed smart city plan. As per census 2011, the population of the city was 6.3 million and it is estimated that the population will be 8.16 million in 2018. By 2021, there could be almost 9 million individuals living in the city [8]. The population density of the city is 9900 persons per square kilometer. For the study, three subsets area of 1 km^2 with
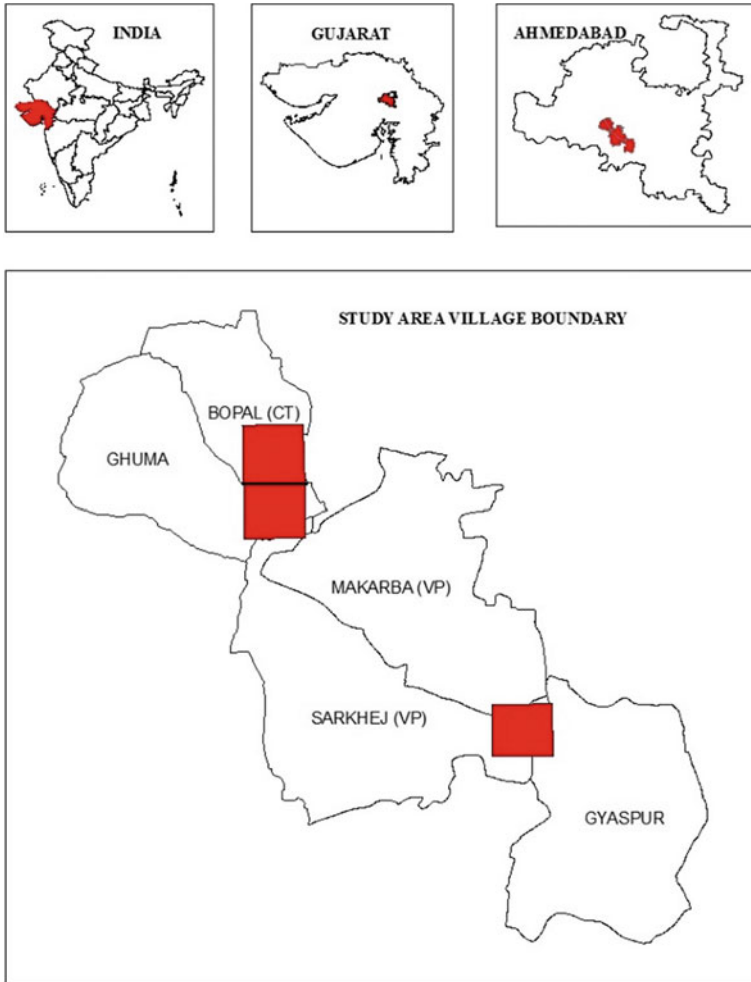
Fig. 2 Study area taken at few villages of Ahmedabad, Gujarat, India

different built-up type and density are selected to compare the accuracy of work. Figure 2 shows the location of the study area and areas selected for detailed accuracy assessment.

The very high spatial resolution image acquired by the Cartosat-2 series satellites were used in this study. The Cartosat-2 series satellites launched by Indian Space Research Organization (ISRO) offers sub-meter spatial resolution data in panchromatic mode. The captured images are useful for several cartographic, urban and land applications. The panchromatic sensor takes a single band image at less than 1 m spatial resolution in wider bandwidth but multiband sensor takes an image in the

Table 1 Specification of Cartosat-2 sensors

Satellite sensor	Period of capture	Spatial resolution	Spectral bands
Cartosat-2 series merged product	October 2016	<1 m	B1-Blue B2-Green B3-Red B4-NIR

accurate multiple bands at a moderate resolution [9]. The specifications of the data used for the study are given in Table 1.

3   Methodology

The overall flow of methodology is shown in Fig. 3. The details of each of the steps are discussed in the subsequent sections.
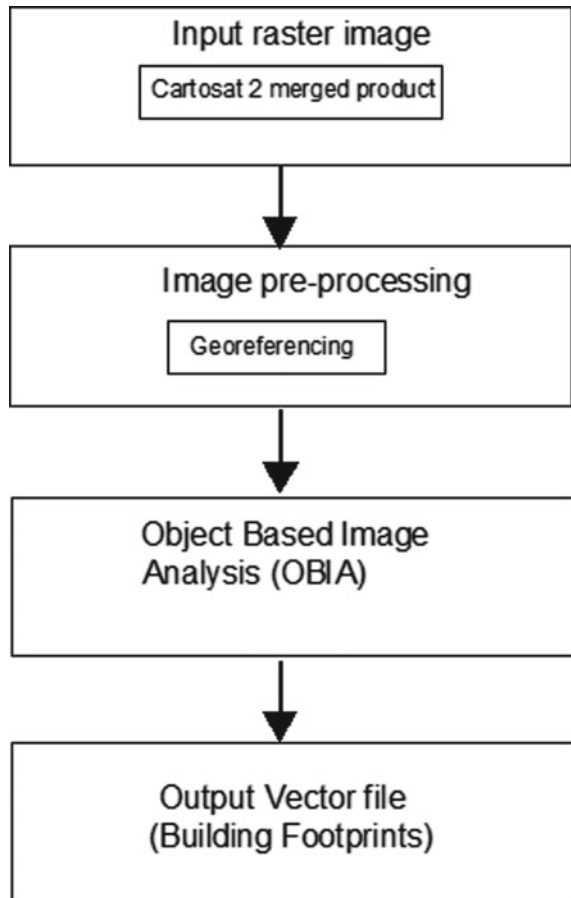
3.1   Image Preprocessing

The quality of the image taken from the sensor depends on its radiometric resolution, cloud cover, wavelength and bandwidth of the band, etc. Based on the application, either of these parameters specifications are considered. For urban studies, the image with high spatial resolution with multispectral bands is mostly preferred. In this study, the merged product of panchromatic high spatial resolution image with corresponding low-resolution multispectral image was used. The images were georeferenced with the corresponding geo-rectified reference images. The alignment of the image with respect to known coordinates is very important for the subsequent process of work.

3.2   Object-Based Image Analysis (OBIA)

Object-Based Image Analysis (OBIA) groups the pixels into meaningful clusters, on the basis of spectral, contextual and textural information. This approach helps to develop the ruleset based on statistics using spectral, spatial characteristics (such as shape, rectangularity fit, length, etc.), textural parameter, relation to neighbor objects for extraction of image objects [10]. It helps in discriminating the spectrally similar objects that improve the classification process, which cannot be obtained in pixel-based method. OBIA segments the image into several objects that are the building blocks for further classification. eCognition is the first commercial software for OBIA, previously it was known as "Definiens" [11].

Fig. 3 Methodology used for building footprint extraction from very-high-resolution satellite image

```
┌─────────────────────────────────────┐
│        Input raster image           │
│   ┌───────────────────────────────┐ │
│   │ Cartosat 2 merged product     │ │
│   └───────────────────────────────┘ │
└─────────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────────┐
│       Image pre-processing          │
│   ┌─────────────────────┐           │
│   │ Georeferencing      │           │
│   └─────────────────────┘           │
└─────────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────────┐
│       Object Based Image            │
│       Analysis (OBIA)               │
└─────────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────────┐
│       Output Vector file            │
│       (Building Footprints)         │
└─────────────────────────────────────┘
```

The process is completely based on programmable workflow and in-built with Fractional Net Evolution approach [12]. Compared with single pixels, the segmented objects have additional spectral information like minimum, maximum values, mean of band, standard deviation, etc. [13]. Figure 4 shows the complete process of OBIA. Process in OBIA is carried out in two phases, segmentation followed by classification.

3.2.1 Image Segmentation

Image segmentation is a fundamental framework for object-based image analysis. In the object-based classification prototype, the region with a homogenous group of pixels is considered as objects. Multiresolution segmentation [12] in eCognition software is working on the basis of the Fractal Net Evolution Approach (FNEA),
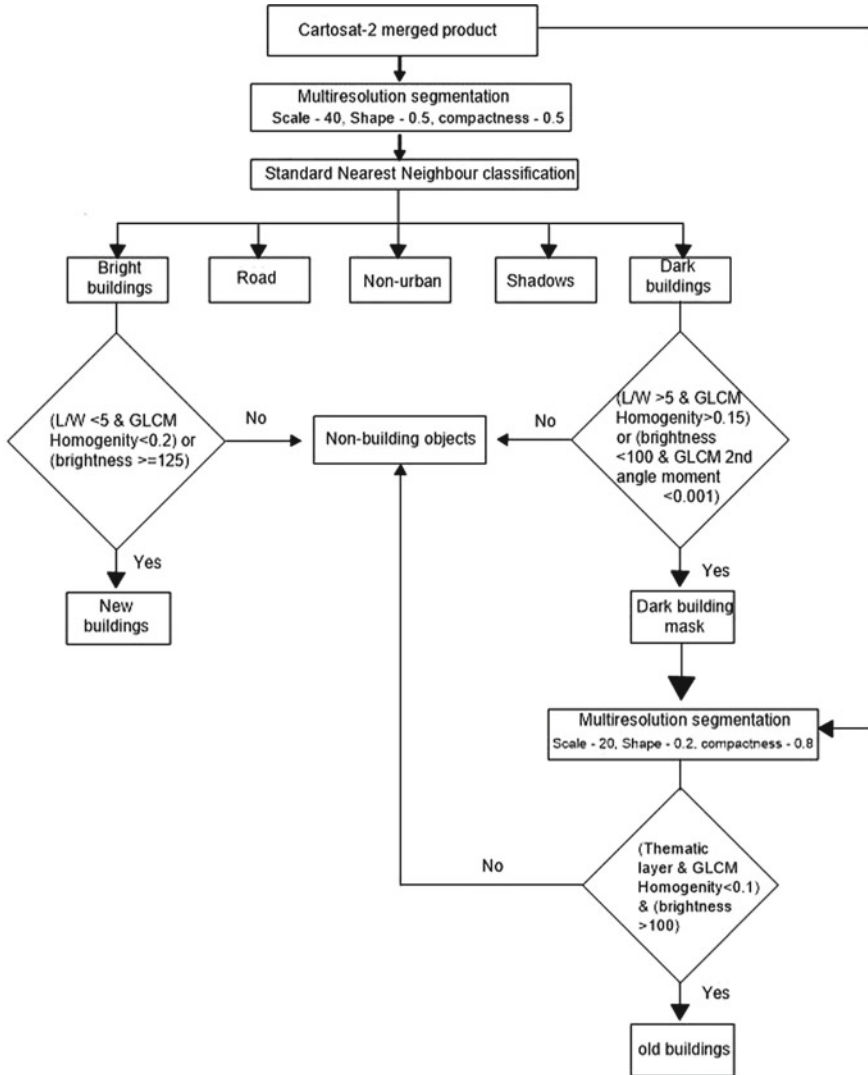
Fig. 4 Stages and algorithms used at various stages of OBIA. The output from each stage is used as input for the subsequent stage and building footprints are extracted
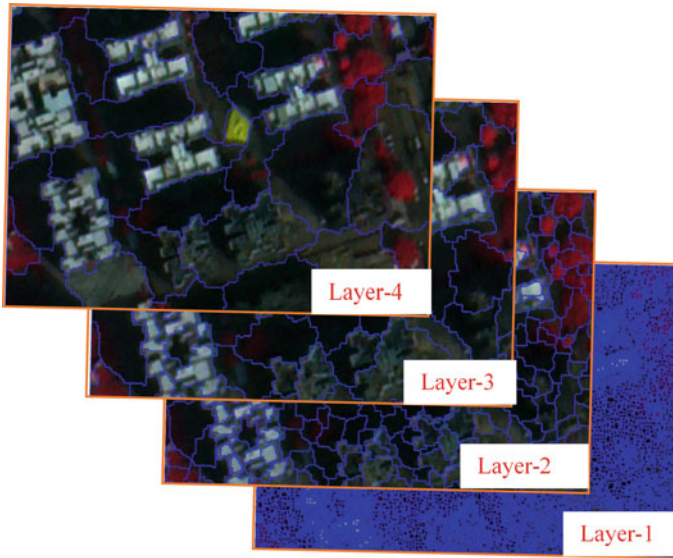
Fig. 5 The image is segmented at various scale parameters with same shape and compactness factor. Layer 1 is shown in the pixel size, layer-2 at the finer scale of 10, layer-3 at a moderate scale of 25 and layer-4 with the object scale of 50

which segments the image on the basis of pixel homogeneity [14]. Figure 5 shows the objects formed with different parameters of segmentation.

The main parameters considered for the multiresolution segmentation are scale, size and compactness factor. At different scales, the weighted values ranging from 0 to 1 are assigned to compactness and shape factor. To effectively control the homogeneity of the object, the summation of shape and compactness factor should be one. The scale parameter depends on the level of detail required for accurate classification, larger the scale (e.g. 100) larger the object size. Compactness factor maintains the smoothness of the object boundaries, whereas shape factor adjusts the spatial and spectral homogeneity of the objects. Generally, a larger optimal segmentation scale is used for lower spatial resolution and vice versa [15].

Multiresolution segmentation is tested at different scale, shape and compactness factor. We consider the scale levels range from 10 to 100, with the shape and compactness factor ranges from 0.2 to 0.8 for the analysis. The scale level of 10 gives a higher level of detail compared to scale 100. After trial and error process, the most suitable parameters for segmentation are selected.

In the whole OBIA process, image segmentation is executed in two stages. In the first stage, multiresolution segmentation is processed with scale level 40, shape and compactness factor 0.5 to execute the basic level of classification. During the second stage, a particular class of interest is further segmented with scale level of 20 to obtain a higher level of details and shape factor is maintained as 0.2 to give more attention to spectral homogeneity.

3.2.2 Image Classification

Object-based image classification on the segmented objects can be executed using nearest-neighbor classifier and membership function. To classify through standard nearest-neighbor (SNN) classifier, the feature space like bands, indices, geometry, and texture are defined for each class. A set of objects for each class are selected as a training sample, in order to assign the classes for the segmented object. In SNN classifier, each object in the image is marked either 0 or 1 [16]. The classifier will check each and every object based on the training samples whether it belongs to the particular class or not.

The image segmented with the scale of 40, the shape and compactness factor of 0.5 is classified through SNN classifier. The image objects are classified into 7 basic classes such as bright buildings (Buildings with higher spectral reflectance), dark buildings (Buildings with lower spectral reflectance), vegetation, soil, road, water bodies, and shadows. The common feature space like brightness, mean, texture is defined for each class. The additional feature class, geometry is used for delineation of the road using its linear characteristic. To increase the accuracy of SNN classification, buildings are classified into two classes because the newly constructed buildings reflectance are higher compared to old buildings.

Membership function classifies the image based on fuzzy logic, where each object in the image has a value between the ranges 0–1. Based on the knowledge and spectral reflectance curve, threshold values are assigned in each stage of classification and image objects are checked whether it satisfying the threshold value or not. The ruleset and constraints are defined after several sets of trial and error process. eCognition provides the several feature spaces to assign the threshold value for classification and it also permits the users to define their own indices. Grey Level Co-occurrence Matrix (GLCM) is used to acquire features like homogeneity, heterogeneity, dissimilarity, etc., by estimating the statistics of pixels. The features derived from texture information are used to differentiate the pixel having the same spectral reflectance.

In the first stage of classification, the object class bright building is used. There are some errors in SNN classification, because the reflectance of soil and concrete payment is similar to new building reflectance. The errors are eliminated by applying threshold values to geometry and texture. Length to width ratio is an important parameter for road classification. Homogeneity of soil and building are quite different, so those parameters are used for building delineation. Even though some of the commercial building object geometry is similar to the road object geometry, so the brightness value above 125 is classified as a new building. Similarly, the dark buildings (building 2) is taken as a class of interest and the condition is applied vice versa and classified as non-building objects. The objects which are not satisfying the conditions are retained in the original class (obtained from SNN classification). Further, object class dark building is exported as a vector layer.

In the second stage of classification, the vector layer is used as input along with a raster image for classification. Few of road and shadow pixel reflectance is similar to that of the dark building, some of them are wrongly classified in dark building (Fig. 6). To remove the unwanted pixel, the objects are subjected to further segmentation.
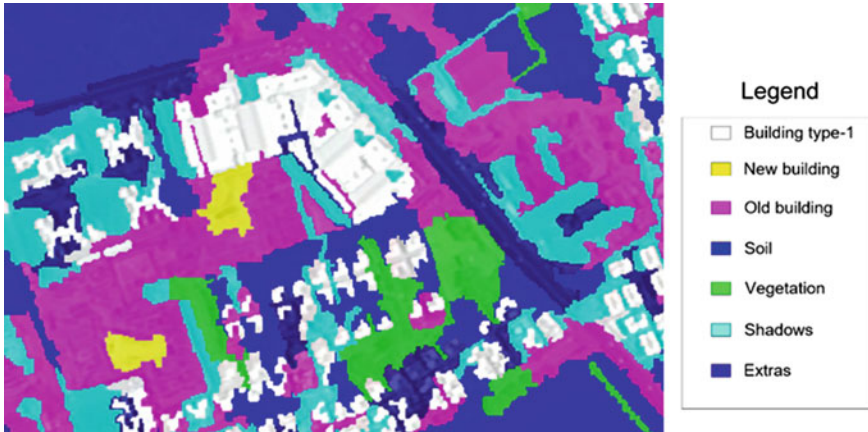
Fig. 6 Classified image; yellow color shows the omitted non-building objects from the building class

Segmentation is carried out only on the thematic layer with the shape factor of 0.2, scale 20, and compactness 0.8 [17]. Spectral and texture properties are used for further classification of old buildings. In the thematic layer, Brightness >100 and GLCM Homogeneity <0.1 is classified as Old buildings). The new and old building are exported and dissolved for further analysis.

4 Results and Discussion

The OBIA method defined above automatically delineate the building footprints from high-resolution satellite images. The created ruleset can be used for all sets of the high-resolution image with some minor modification in threshold value. Three subsets of the study area are taken, with discrete building type and density to check the accuracy of building footprint extraction (Fig. 7). In the subset region, the building footprints are also extracted manually to compare the accuracy of automatically extracted building footprint.

Buildings in a well-distributed built-up area are extracted very well with high accuracy. Old buildings in the study are not completely extracted by the developed ruleset. If ruleset is modified to extract the old buildings, it adversely increases the commission error. The accuracy of the automatic method with respect to the manual is 82.5% (Fig. 8a). Red color indicates buildings, which are common in both manual and automatic method. The green color shows the unidentified buildings in the automatic method and blue color shows non-building objects. The density of buildings is higher in Fig. 8b compared to Fig. 8a. An organized construction is done in Fig. 8b. So, building footprints are easily delineated. The accuracy of buildings from OBIA technique (Fig. 8b) compared with the manual method is 80%. In Fig. 8c, the build-
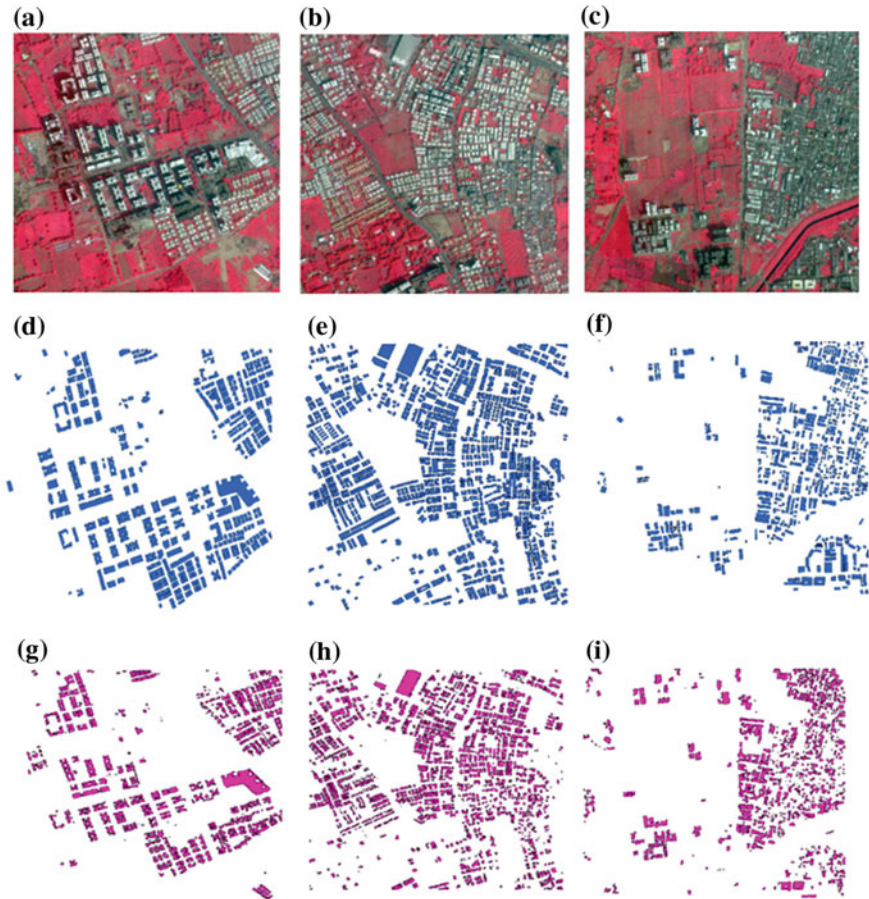
Fig. 7 **a–c** Subset of study area at various building type and density. **d–f** are manually extracted building footprint through visual interpretation. **g–i** are automatically extracted buildings through OBIA

ings (overcrowded slum areas) are very congested and constructed in an unplanned manner. Moreover, it is very difficult to point to individual building visually. However, the created ruleset is executed to validate the algorithm. The accuracy of the automatic method in overcrowded slum areas (Fig. 8c) compared to visual interpretation is around 69.2%. The accuracy of building footprint extracted automatically at a well-distributed built-up area, crowded built-up area and overcrowded built-up area are given in Table 2. The result shows that the accuracy of the automatic method for well-distributed built-up area is comparatively higher than obtained in crowded built-up and overcrowded informal settlements.
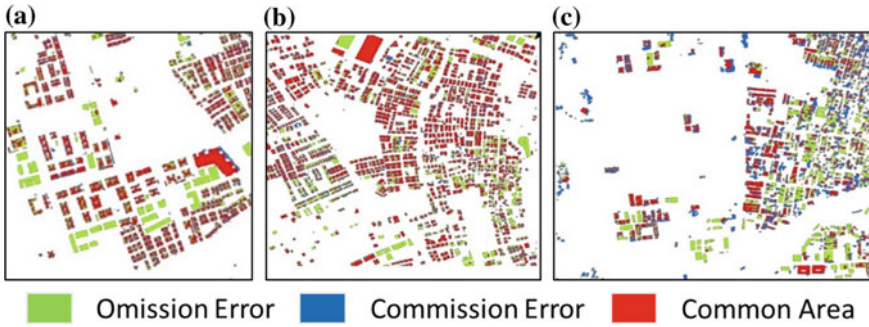
Fig. 8 Accuracy assessment of buildings extraction in (**a**) well-defined built-up area; **b** crowded built-up area; and **c** overcrowded areas

Table 2 Summary of OBIA accuracy

Urban areas	Reference area (m^2)	Correctly classified area (m^2)	Unidentified buildings (m^2)	Wrongly classified area (m^2)	Total classified built-up area (m^2)
Well-distributed built-up area	121,021.06	99,825.39 (82.5%)	21,195.67 (17.5%)	22,420.09 (18.5%)	122245.48
Crowded built-up area	200,202.02	160,161.67 (80%)	40,040.35 (20%)	57,769.89 (11.9%)	217931.56
Over-crowded slum area	72,170.76	49,942.16 (69.2%)	2,222.86 (30.8%)	1,726.36 (23.92%)	51668.52

5 Conclusions

The study proposed, a blend of object-based nearest-neighbor classification and rule-based classifier for building footprint extraction from very-high-resolution satellite images. The study developed a process-ruleset for object-based image analysis, which can be used for building footprint extraction from the very-high-resolution sensor with appropriate morphology-based modifications. The combination of supervised nearest-neighbor classification with decision-rules resulted in an accuracy of over 82.5% in the extraction of building footprints. The accuracy of feature extraction in densely populated areas is low due to the effect of spectral mixing of features and smaller size of objects. The results of the study will be useful in developing a 3D city model for Ahmedabad city and for assessment of changes in the built-up volume in the city.

References

1. Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. Environ Res Lett 4:044003
2. Dahiya S, Garg PK, Jat MK (2013) A comparative study of various pixel-based image fusion techniques as applied to an urban environment. Int J Image Data Fusion 4:197–213
3. Jensen JR, Cowen DC (1999) Remote sensing of urban/suburban infrastructure and socio-economic attributes. Photogram Eng Remote Sens 65:611–622
4. Kpalma K, El-Mezouar MC, Taleb N (2014) Recent trends in satellite image pan sharpening techniques. In: 1st International conference on electrical, electronic and computing engineering. Vrniacka Banja, Serbia
5. Blaschke T, Burnett C, Pekkarinen A (2004) Image segmentation methods for object-based analysis and classification. In: Jong SMD, Meer FDV (eds) Remote sensing image analysis: including the spatial domain. remote sensing and digital image processing, vol 5. Springer, Dordrecht, pp 211—236
6. Lia, HT., Gua, HY., Hana, YS., Yanga, JH., Hanb, SS (2008) An efficient multi-scale segmentation for high-resolution remote sensing imagery based on statistical region merging and minimum heterogeneity rule. In: International archives of the photogrammetry, remote sensing and spatial information sciences, vol. XXXVII, Part B4. Beijing, pp 1257–1262
7. Im J, Jensen JR, Tullis J (2008) A object-based change detection using correlation image analysis and image segmentation. Int J Remote Sens 29:399–423
8. World cities population data. http://www.worldpopulationreview.com
9. Belfiore OR, Meneghini C, Parente C, Santamaria R (2016) Application of different pan-sharpening methods on worldview-3 IMAGES. J Eng Appl Sci 11:490–496
10. Myint SW, Gober P, Brazel A, Grossman-Clarke S, Weng Q (2011) Per-pixel versus object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens Environ 115:1145–1161 (2011)
11. Lang S, Tiede D (2007) Definiens developer. GIS Bus 9(2007):34–37
12. Baatz M, Schape A (2000) Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: Proceedings of the Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT Symposium. Salzburg, Austria
13. Hay GJ, Castilla G (2008) Geographic object-based image analysis (GEOBIA): a new name for a new discipline. In: Blaschke T, Lang S, Hay GJ (eds) Object-based image analysis. Lecture notes in geoinformation and cartography, pp 75–89. Springer, Berlin, Heidelberg
14. Willhauck G (2000) Comparison of object oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos. Int Arch Photogram Remote Sens, vol XXXIII, Part B3, pp 214–221
15. Ma L, Li M, Ma X, Cheng L, Du P, Liu Y (2017) A review of supervised object-based land-cover image classification. ISPRS J Photogram Remote Sens 130:277–293
16. Shackelford AK, Davis CH (2003) A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. Geosci Remote Sens IEEE Trans 41:2354–2363
17. Sharma SA, Agrawal R, Jayaprasad P (2016) Development of 3D city models using irs satellite data. Indian Soc Remote Sens 44(2):187–196