



**Journal of Engineering  
Sciences Assiut University  
Faculty of Engineering  
Vol. 48, No. 5  
September 2020  
PP. 869-887**



## BUILDING EXTRACTION FROM VERY HIGH-RESOLUTION SATELLITE IMAGES FOR MAP UPDATING IN EGYPT

**Mostafa H. Shoaib<sup>1</sup>, Yasser Mostafa<sup>1</sup>, Yousef A. Abbas<sup>2</sup>**

<sup>1</sup> *Civil Eng. Dept., Faculty of Engineering, Sohag University, Sohag, Egypt*

<sup>2</sup> *Civil Eng. Dept., Faculty of Engineering, Assiut University, Assiut, Egypt*

*Received 21 April 2020; Revised 16 May 2020; Accepted 31 May 2020*

### ABSTRACT

Robust building detection from satellite images has been a subject of interest for several decades. Very High Resolution (VHR) satellite images support the efficient extraction of manmade objects. The main aim of this paper is to present an approach for building extraction from VHR satellite images for map updating in Egypt. To achieve this aim, a comparison of pixel and object-based classification techniques has been applied. Then, different refinement processes based on shadow, context, shape, and Digital Surface Model (DSM) data are carried out. Two study areas from the VHR satellite images for Assuit and Sohag cities are used. A comparison of the classification techniques shows that the Maximum Likelihood Classifier (MLC) for pixel-based technique and Support Vector Machine (SVM) for object-based technique give the highest overall accuracy results. Refinement based on shadow, context, shape, and DSM information improves the overall accuracy with an average of 18%. Thus, the building extraction results can contribute significantly to update maps in Egypt .

**Keywords:** Building extraction, Pixel-based, Object-based, Classification, Accuracy assessment, Map updating.

### 1. Introduction

The automatic extraction of buildings from remote sensing images is a hot topic and an active field of research. This is due to an increasing need for building detection in a variety of applications such as; urban planning, change detection, disaster monitoring, emergency services, transportation

urban monitoring, and map updating, etc. [1]. Automatic detection of buildings from satellite images based on rooftops is a difficult task for many reasons. Building roofs are composed of different materials. The diverse materials used for the rooftops play an important role in heterogeneity as these materials may have a similar texture and spectral properties with the surrounding objects, which makes the segmentation process very difficult. Because of these reasons, some additional features (like area, shape, etc.) are also required for increasing the accuracy of extracted buildings from satellite images.

The recent availability of commercial VHR satellites such as IKONOS, QuickBird, WorldView, and GeoEye provides a new data source for accurate building extraction. The high spatial resolution of the imagery helps to reveal very fine details in urban areas and hence, facilitates the classification and extraction of buildings. Many researchers studied developing robust building detection algorithms. Ghaffarian [2] proposed a fully automatic technique for buildings detection from single high-resolution Google Earth images using supervised classification. This method has the availability of building detection without influencing their geometry characteristics but can't separate non-building from building areas when they have similar spectral properties. Belgiu and Drăguț [3] compared between supervised and unsupervised segmentation methods to classify buildings. The results showed that unsupervised segmentation can extract buildings successfully but supervised segmentation still overcomes it.

Attarzadeh and Momeni [4] proposed a novel object-based technique for automatic detection of building from satellite images. This method provided promising results, where neighborhood, contextual and geometrical features were used in addition to spectral features. An automatic classification algorithm based on the object-based method has been developed by Shedlovska and Hnatushenko [5]. The results showed that the method gives good results when the mean-shift method was applied in the segmentation step. Elsharkawy et al [6] classified an urban area to six classes based on combining of pixel and object-based methods. They concluded that the method gives good results and applying curvelet transform increases the accuracy of building detection. Mostafa et al [7] compared pixel and object-

based classification to extract features in the Egyptian environment. They found that the object-based method achieves the highest accuracy and as the planning degree increases as the results are more accurate. The object-oriented building extraction was applied by Jiang et al [8]. They used DSM to extract high objects like trees and buildings. Normalized Digital Vegetation Index (NDVI) was used for distinguishing between trees and buildings.

One class SVM was used by Manandhar et al [9] to determine the man-made structures (buildings, roads, etc.). They proceed with the texture segmentation technique using a conditional threshold value to extract buildings. They concluded that their method can detect buildings with different colors and shapes. Turker and San [10] presented a method for extracting the buildings from high-resolution IKONOS images through binary SVM classification. In addition to original spectral bands, the bands NDVI, normalized DSM, Principal Component 1 (PC1), PC2, PC3, and PC4 were also included in the classification. The results showed that using additional bands NDVI, normalized DSM, and PCs increase the classification accuracy for extracting the buildings. Benarchid et al [11] presented an automatic building detection using SVM and shadow information. They concluded that using shadow increases the achieved accuracy.

Senaras [12] proposed an automated building extraction technique using fuzzy k-Nearest Neighbors (KNN). He classified each pixel according to its Hue Saturation Brightness (HSB) values and concluded that the applied technique gives satisfying results in the urban areas. Noi and Kappas [13] compared the performances of SVM, KNN, and Random Forest for land cover classification. The SVM classifier produced the highest accuracy. A comparison of SVM and MLC techniques was applied by Mondal et al [14] to identify the accurate method for land-use classification. They concluded that the accuracy of SVM was better than MLC.

The main aim of this research is to extract buildings from VHR satellite images for map updating in Egypt. A comparison of pixel and object-based classification techniques has been applied. The classification accuracy assessment has been calculated for each technique. The best classification method result is refined based on shadow, context, shape, and DSM data. The remainder of this paper is organized as follows. Section 2 shows the data

used. The methodology is reported in section 3. Accuracy assessment and refinement are illustrated in sections 4 and 5, respectively. Section 6 shows the map updating procedure. Finally, conclusions are discussed in section 7.

## 2. Data used

### 2.1. Satellite images

Two satellite images are used in this paper. The first was acquired for Assuit city using WorldView-2 satellite as shown in Figure 1-A (1000 m by 500 m) with 0.5 meters resolution for panchromatic and 2 meters for multispectral images. It is composed of eight multispectral bands (Coastal, Blue, Green, Yellow, Red, Red Edge, Near Infra-Red 1, and Near Infra-Red 2) with a radiometric resolution of 11 bits per pixel. The acquisition date is 4/3/2016 with an angle of 13.57 degrees. The second was acquired for Sohag city using GeoEye satellite as shown in Figure 1-B (530 m by 530 m) with 0.5 meters resolution for panchromatic and 2 meters for multispectral images. It is composed of four multispectral bands (Blue, Green, Red, and Near Infra-Red) with a radiometric resolution of 11 bits per pixel. The acquisition date is 24/5/2014 with an angle of 19.61 degrees.

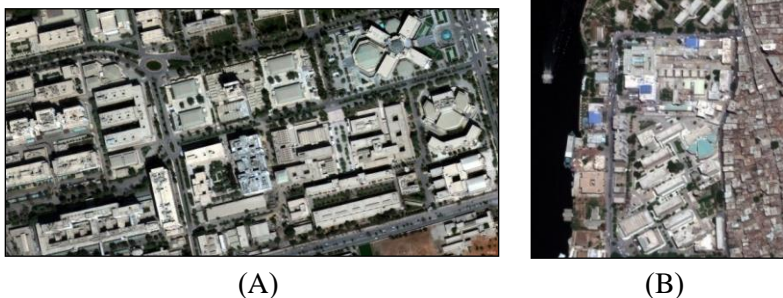


Fig. 1. Used image data, (A) Satellite image for Assuit city, and (B) Satellite image for Sohag city.

### 2.2. Topographic map

Topographic map covering Sohag city had been acquired from aerial photographs in 2006, with a large scale at 1:5000 (Figure 2-A). It had been produced by the Egyptian Survey Authority and used as a basis for comparison.

### 2.3. Ground data

#### 2.3.1. Ground Control Points (GCPs)

Twenty-five Ground Control Points (GCPs) were affixed in Sohag study area to be used for image geometric correction. The coordinates of the GCPs had been determined from the field observations using the GPS receiver SOKKIA GRX2 with a horizontal accuracy of  $\pm (10 \text{ mm} + 1 \text{ ppm})$ . Differential Global Positioning System (DGPS) was the applied technique. Figure (2-B) shows the configuration of these points in the study area.

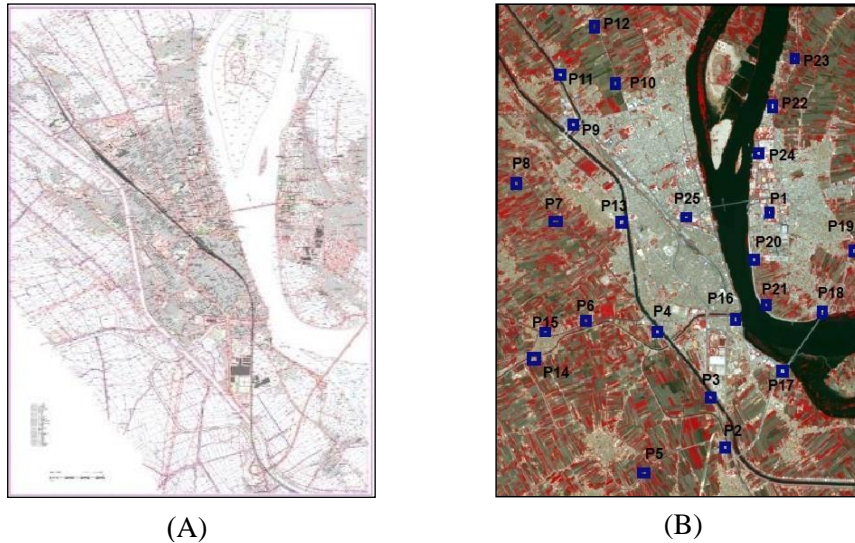


Fig. 2. Used image data, (A) Satellite image for Assuit city, and (B) Satellite image for Sohag city.

### 2.3.2. Digital Surface Model (DSM)

DSM is generated for both study areas with the help of shadow information. Firstly, the total station had been used to calculate the height of one of the buildings in the field. Secondly, shadow length had been measured from satellite images for the observed building. Finally, relatively and proportionately, the heights of the remaining buildings and trees were calculated according to their shadow lengths [15]. Figure (3) shows the DSM results in a raster format.

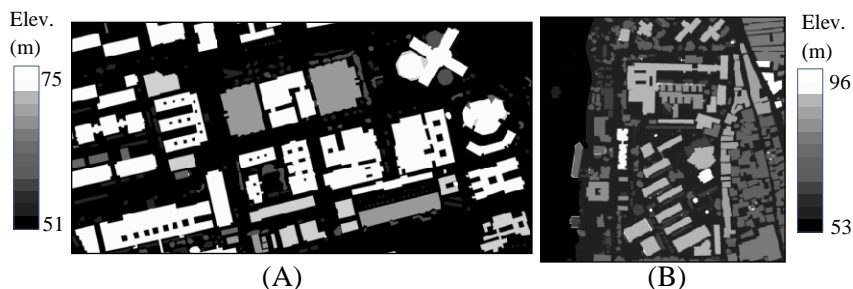


Fig. 3. DSM band, (A) Assuit city, and (B) Sohag city.

### **3. Methodology**

The building extraction methodology includes six steps (Figure 4). Firstly, satellite images had been corrected geometrically and fitted to a known coordinate system. Secondly, corrected panchromatic and multispectral images had been fused to obtain pan-sharpened images. Thirdly, pan-sharpened images had been classified using pixel and object-based approaches to obtain building and non-building. For the pixel-based approach, unsupervised and supervised classification had been applied to the images. The different techniques Maximum Likelihood, Mahalanobis Distance, Minimum Distance, Spectral Angle Mapper, and Spectral Correlation Mapper had been used for supervised classification. For the object-based approach, classification had been applied using SVM, KNN, and threshold (Intensity and Saturation) methods. Fourthly, building and non-building classes obtained had been assessed to determine the achieved accuracy for each one. Fifthly, refinement had been applied to the best-classified image to improve its accuracy. Finally, the raster layer obtained from the refinement process had been converted to vector format to be used in map updating.

#### ***3.1. Data preprocessing***

Data preprocessing includes the required tasks to present the satellite images in a suitable form which can lead to better interpretation and detecting the maximum information from these images [16]. This stage had been carried out according to the following steps:

##### ***3.1.1. Image geo-referencing***

Sohag satellite image had been rectified to Egypt Transverse Mercator coordinate system using twelve ground control points as well as thirteen checkpoints. The third-order polynomial transformation method had been applied. The Root Mean Square Error (RMSE) obtained in the checkpoints is 0.57 m. The resampling process had been carried out using the nearest neighbor method. Geo-referencing and resampling were implemented in ERDAS IMAGINE 2013 software.



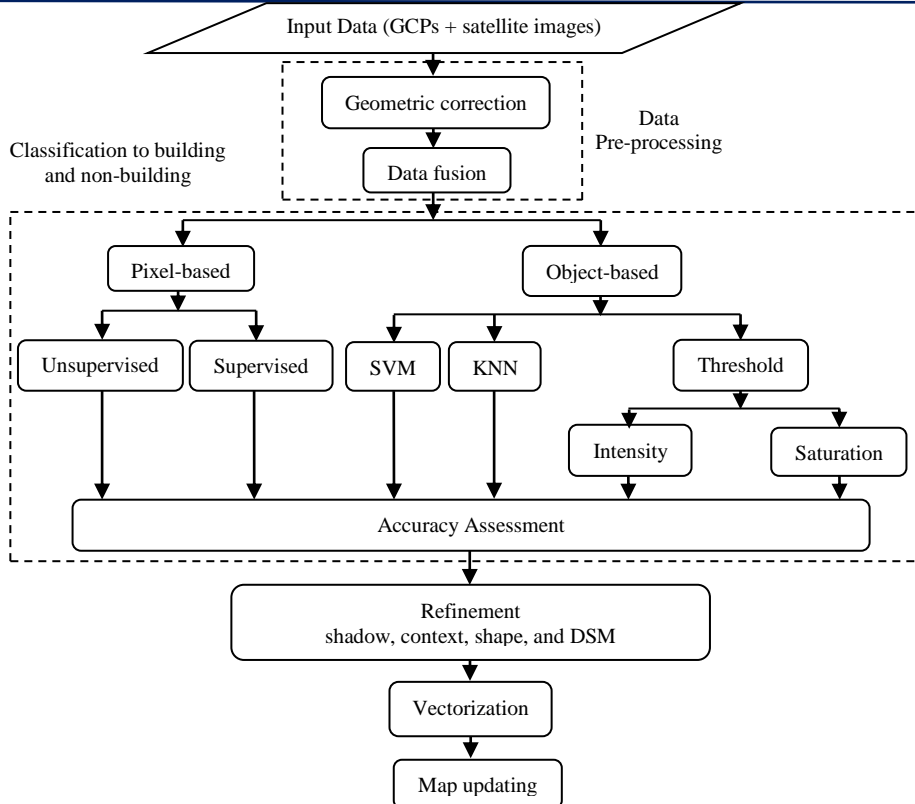


Fig. 4. Flowchart for building extraction.

### 3.1.2. Data fusion

Pan-sharpened of WorldView-2 and GeoEye images had been produced by the fusion of high-resolution panchromatic and multispectral images using the Intensity-Hue-Saturation technique that was implemented in ERDAS IMAGINE 2013 software. This step is very important to improve multi-spectral resolution to 0.5 m. The nearest neighbor method had been selected for resampling to keep the color content of the original multispectral image unchanged.

### 3.2. Pixel-based technique

Pixel-based classification technique was always the primary tool to extract land cover classes from digital remotely sensed data. In this paper, unsupervised and supervised classifications have been applied using ERDAS IMAGINE 2013 software.

### *3.2.1 Unsupervised pixel-based classification*

Unsupervised classification, in essence, reverses the supervised classification process. Spectral classes are assembled first, based solely on the numerical information in the data, and are then matched by the analyst to information classes. Clustering algorithms are used to specify the natural (statistical) groupings in the data. The main idea of clustering is to partition the sample data set into smaller sets so that each subset contains a similar type of data in itself and also different from the other subsets.

### *3.2.2 Supervised pixel-based classification*

Supervised classification is a process of classifying pixels into a number of classes based on training samples which are manually selected [2]. In this work, ten samples were selected for each class in a polygon shape. Classification had been executed by five popular pixel-based statistical classifiers commonly used in the supervised classification; Maximum Likelihood, Mahalanobis Distance, Minimum Distance, Spectral Angle Mapper, and Spectral Correlation Mapper [17]. After that, the building class had been separated from other classes. In this paper, the image had been classified into four classes for the study area (A) and five classes for the study area (B).

### **3.3. Object-based technique**

The object-based classification technique is one of the most popular methods. The basic processing units of object-based image analysis are image objects or segments, not single pixels. It uses important information (shape, texture, and contextual information) that is present only in meaningful image objects and their mutual relationships [18], [19]. In this paper, the Multiresolution segmentation algorithm is used to create image objects with scale 25, shape 0.2, and compactness 0.5 (the parameters were specified by trial and error method). Then, a set of training samples (50 samples per class in a polygon shape) are manually assigned for buildings and non-buildings. After that, classification is performed using Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) algorithms. In addition, threshold is considered as an object-based technique for building extraction. It is executed after the segmentation process to obtain buildings and non-buildings objects. The



object-based classification technique has been applied using eCognition software.

### *3.3.1. Support Vector Machine (SVM)*

SVM is a supervised non-parametric classification technique derived from statistical learning theory. There are three steps in the SVM classification. First, the training of the classes is represented as feature vectors. Next, feature vectors are mapped into a feature space using the kernel function. Finally, an n-dimensional hyperplane that optimally separates the classes is created. Kernel type rbf, C=10, and gamma=5 parameters were selected by trials in eCognition Developer 9.0.

### *3.3.2. K-Nearest Neighbor (KNN)*

KNN is a supervised non-parametric classification technique which is widely used in pattern recognition. The main idea for this classification method is to assign a sample point by examining the class labels of the nearest k-point. Then, the training phase locates feature vectors of the sample objects into the sample space and labels their class information. In the classification phase, for each new object whose label is not known, its k-nearest neighbor is detected in the feature vector space. Then, by applying a majority voting for these neighbors, its class label is predicted.

### *3.3.3. Threshold*

Threshold is an effective part of image segmentation, where the main objective is to isolate objects from the background. It had been used to isolate building objects from non-building objects and there is no need to assign training samples manually. To extract building objects, a color model is used where Red, Green, Blue (RGB) layers are converted to Intensity, Hue, Saturation (IHS) layers. Image objects which belong to buildings have distinctive properties in both intensity and saturation layers because building's roofs are mostly made of bright materials. So, using an appropriate threshold, buildings can be extracted.

## **4. Accuracy assessment**

To assess the accuracy of each technique, the reference image is created by manual interpretations (figure 5). Then, a comparison has been made between

the obtained results and the reference image. Three types of accuracy are used in the accuracy assessment, namely, the producer's accuracy, the user's accuracy, and the overall accuracy. The producer's accuracy refers to the measure of omission errors that correspond to those pixels belonging to a class of interest that the classifier has failed to recognize. The user's accuracy, on the other hand, represents the measure of commission errors that correspond to those pixels from other classes that the classifier has labeled as belonging to the class of interest. The overall accuracy has the advantage of being directly interpretable as the proportion of pixels being classified correctly [20].

1) Producer's Accuracies:

$$\text{a) Building: } \eta_B = TP/(TP + FN). \quad (1)$$

$$\text{b) Non-building: } \eta_N = TN/(FP + TN). \quad (2)$$

2) User's Accuracies:

$$\text{a) Building: } P_B = TP/(TP + FP). \quad (3)$$

$$\text{b) Non-building: } P_N = TN/(TN + FN). \quad (4)$$

$$\text{3) Overall Accuracy: } \tau = (TP + TN)/(TP + TN + FP + FN) \quad (5)$$

where True Positive (TP) refers to the number of building pixels correctly identified, False Negative (FN) refers to the number of building pixels identified as non-buildings, False Positive (FP) refers to the number of non-building pixels identified as buildings, True Negative (TN) refers to the number of non-building pixels correctly identified, and  $TP + TN + FP + FN$  stands for the total number of pixels in the image.

The accuracy measurement results of all study areas are summarized in Table 1. In the pixel-based image unsupervised classification, the study area (A) had been classified into four classes (building, roads, vegetation, and shadow). The overall accuracy for building detection is 78.21%. The study area (B) had been classified into five classes (building, roads, vegetation, water, and shadow). The overall accuracy for building detection is 76.15%. For supervised classification, the overall accuracy of the study areas (A) and (B) ranges from 77.49% to 80.49% and 73.34% to 76.54%, respectively. The Maximum Likelihood classifier outperforms other classifiers in both study areas with an overall accuracy of 80.49% and 76.54%, respectively.

**Table 1.** Accuracy assessment results.

Method / Study area		Producer's Accuracy		User's Accuracy		Overall Accuracy ( $\tau$ )%		
		Building $\eta_B$ (%)	Non-Building $\eta_N$ (%)	Building $P_B$ (%)	Non-Building $P_N$ (%)			
Unsupervised Pixel Based Technique		A	55.68%	90.63%	76.63%	78.76%	78.21%	
		B	40.95%	93.21%	74.51%	76.51%	76.15%	
Supervised Pixel Based Technique	Maximum Likelihood	A	83.22%	79.47%	69.10%	89.57%	80.49%	
		B	82.07%	73.86%	60.34%	89.48%	76.54%	
	Mahalanobis Distance	A	84.92%	77.48%	67.53%	90.30%	80.12%	
		B	87.90%	67.79%	56.94%	92.04%	74.35%	
	Minimum Distance	A	75.38%	83.84%	72.01%	86.06%	80.10%	
		B	69.37%	75.27%	57.61%	83.53%	73.34%	
	Spectral Angel Mapper	A	80.75%	78.49%	67.43%	88.09%	79.29%	
		B	74.67%	74.16%	58.34%	85.80%	74.33%	
Spectral Correlation Mapper	A	81.45%	75.30%	64.52%	88.04%	77.49%		
	B	49.86%	87.21%	65.38%	78.21%	75.02%		
Object Based Technique	SVM	A	76.10%	83.25%	71.47%	86.33%	80.71%	
		B	51.01%	90.56%	72.36%	79.23%	77.65%	
	KNN	A	77.35%	80.03%	68.11%	86.50%	79.08%	
		B	65.15%	82.53%	64.38%	83.02%	76.86%	
	Color Model	Intensity	A	84.87%	78.27%	68.29%	90.36%	80.61%
			B	74.31%	78.64%	62.77%	86.34%	77.23%
		Saturation	A	75.93%	82.60%	70.65%	86.15%	80.23%
			B	81.30%	70.02%	56.78%	88.54%	73.70%

In object-based image classification for the study area (A) and (B), overall accuracy ranges from 79.08% to 80.71% and 73.70% to 77.65%, respectively. SVM classifier outperforms other classifiers in both study areas with an overall accuracy of 80.71% and 77.65%, respectively. Results of the study area (A) are better than study area (B) because there is a large area at study area (B) not well planned with different spectral reflectance from other well planned one.

The best results for pixel-based and object-based techniques are shown in Figure (5). SVM classifier which gives the best classification results will be used in the refinement process.

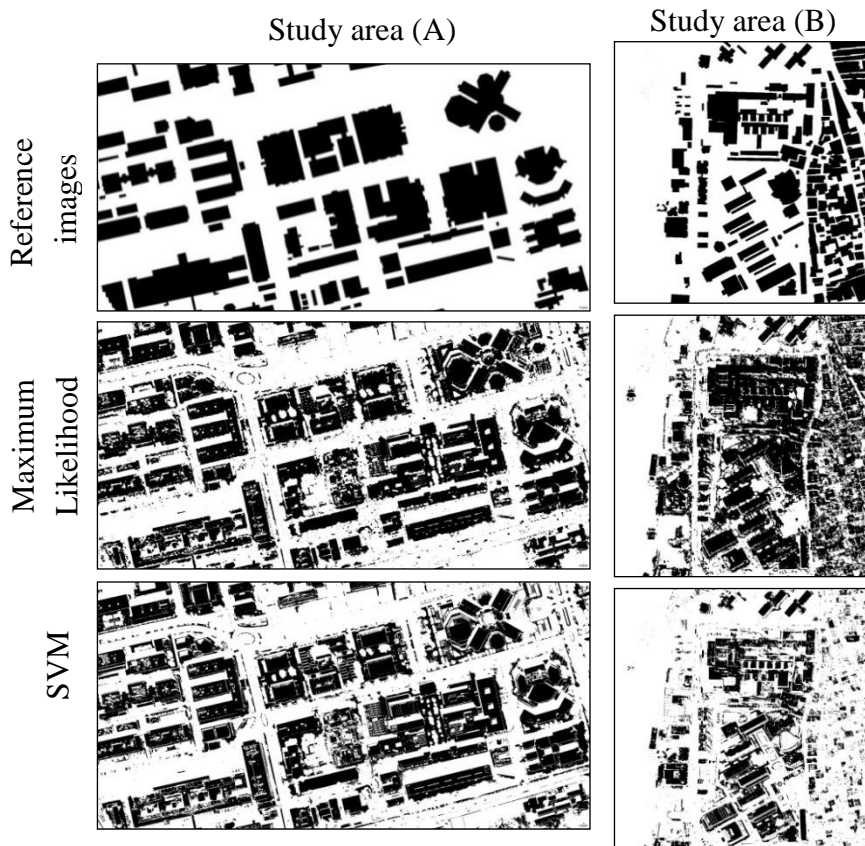


Fig. 5. Image classification results.

## 5. Refinement

After assessing the accuracy of the proposed techniques, it has been noted that some patches are misclassified. So, the refinement process had been applied to improve the achieved results. Refinement had been applied to SVM images result based on shadow, context information, shape information, and DSM data.

### 5.1. Refinement based on shadow detection

A common characteristic of urban areas in images is the presence of shadows, which are connected to buildings. Shadow length depends on the height of the buildings and the illumination conditions. Shadow regions have been detected (Figure 6) using Shadow Detector Index (SDI) formula proposed by Mostafa and Abdelhafiz [21]:

$$SDI = \frac{(1-PC1)+1}{((G-B)*R)+1} \quad (6)$$

where R, G, and B are normalized components of red, green, and blue bands,

respectively. PC1 is a normalized component of the first principal component.

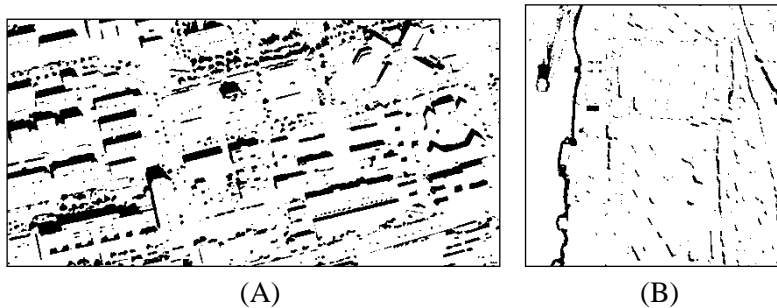


Fig. 6. Shadow mask, (A) Assuit city, and (B) Sohag city.

If the building patches are not connected to shadow regions, these patches are deleted because they are considered false-positive objects (Figure 7).

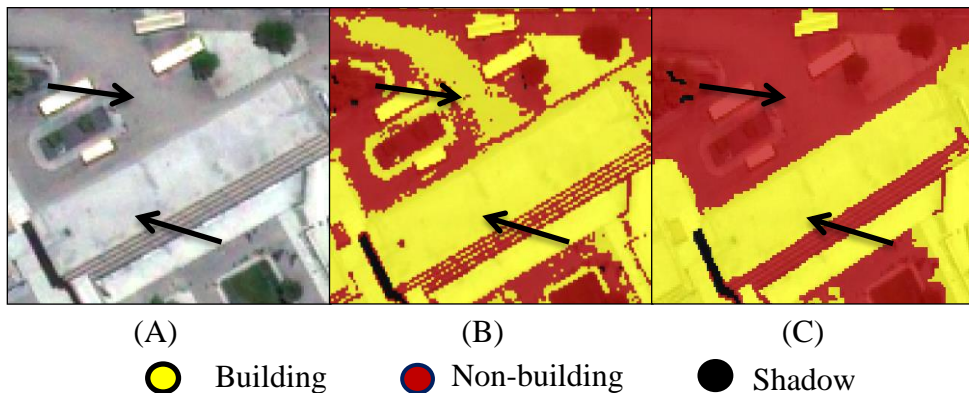


Fig. 7. Refinement result based on shadow, (A) Original image, (B) Classified image, and (C) Refined image.

### 5.2. Refinement based on context and shape information

Some of the misclassified objects are highly enclosed by the building's objects. So, class-related (context) features had been used for refinement. If misclassified objects have a high common border to building objects, they should also belong to the class buildings because they are considered false-negative objects (Figure 8). Shape information such as the building's area also had been used for the refinement process, where small misclassified objects are deleted.

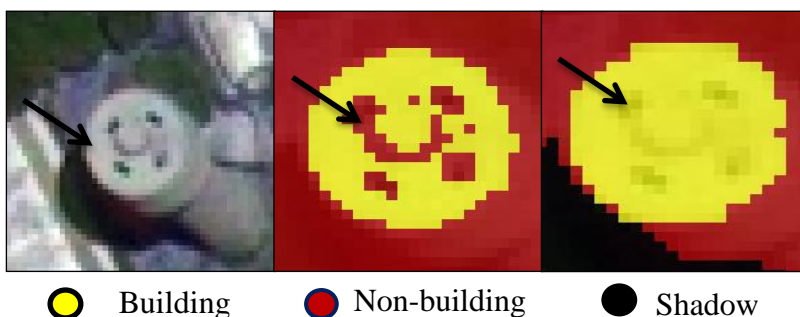


Fig. 8. Refinement result based on context information.



### 5.3. Refinement based on DSM data

A common characteristic of buildings objects in images is that they have a different height from the surrounding objects. This feature had been utilized to correct FN and FP objects. It helps to detect the edge pixels for each building which makes the extraction process more accurate. It also helps to detect buildings with low brightness materials that can't be detected using the SVM classifier.

### 5.4. Refinement results

According to refinement based on shadow, context information, and shape information, overall accuracy improved from 80.71% to 88.54% for the study area (A) and from 77.65% to 81.63% for the study area (B). Refinement of the previous results based on DSM data, improved overall accuracy for the study area (A) from 88.54% to 97.56% and from 81.63% to 96.73% for the study area (B). The study area (A) refinement results are better than the study area (B) because there are buildings at study area (B) that have low heights under the threshold value. Building extraction results after refinement for both study areas are shown in Figure (9).

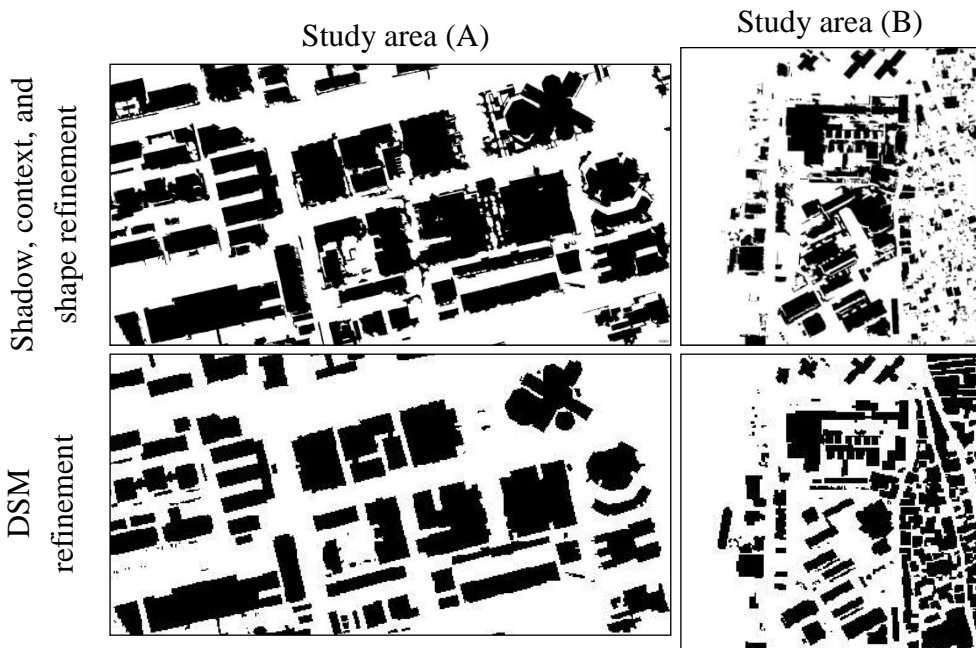


Fig. 9. Building extraction results after refinement.



## 6. Map Updating

After the refinement process of the classification results, the raster objects are converted to the vector domain as a polygon. This step has been applied using ERDAS IMAGINE 2013 software. Figure (10) shows the conversion from raster format to vector format for the extracted buildings. More effort has been paid to complete the vectorized polygon layer into a feasible condition. Many editing operations had been performed on a vector layer to eliminate unnecessary vertices in the lines such as generalizing to remove steps from the line and reshaping to modify the shape of a single element after editing its nodes (Figure 11). In practice, conversion to practicable vector format and further completion by manual editing increases the smoothing of building boundaries.

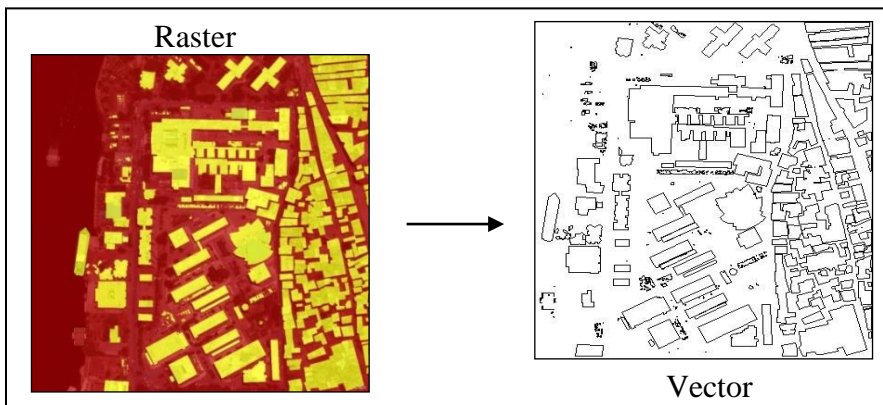


Fig. 10. Raster to vector conversion.

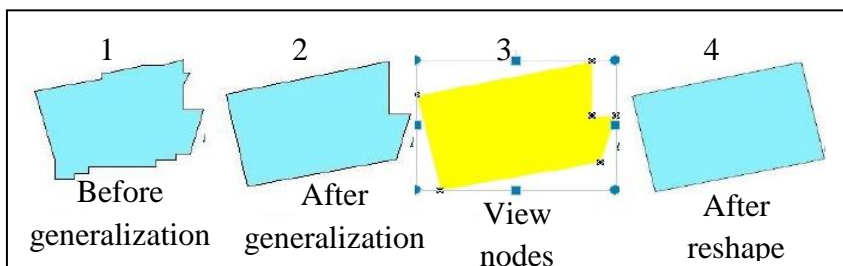


Fig. 11. Post-processing on vector layer.

The extracted buildings from the study area (B) are compared with the old map to distinguish the changes that happen in the real world. The most common quality measures for building extraction are Building Extraction

Rate (BER) (eq.7) and Roof Area Coverage (RAC) (eq.8) [22]. BER and RAC are counted considering the 25% coverage threshold.

$$\text{Building Extraction Rate (BER)} = \frac{BCE}{BCE+BNE} = 274 / (274+28) = 90.73\% \quad (7)$$

where BCE is the “Building Correctly Extracted” and BNE is the number of “Buildings Not Extracted”.

$$\text{Roof Area Coverage (RAC)} = \frac{EBA}{RPA} = 89927 / 91679.5 = 98.08\% \quad (8)$$

where EBA is the total “Extracted Building Areas” and RPA is total “Reference Polygon Areas”.

The correct buildings area extracted from the GeoEye image was 8.35% more than that in the reference map. It is worth mentioning that, the satellite image was acquired in 2014 and the reference map had field revision in 2006. The increase in the area of buildings is due to the presence of new buildings after the reference map date. Figure (12) shows the updated areas in the Sohag city 1:5000 map.

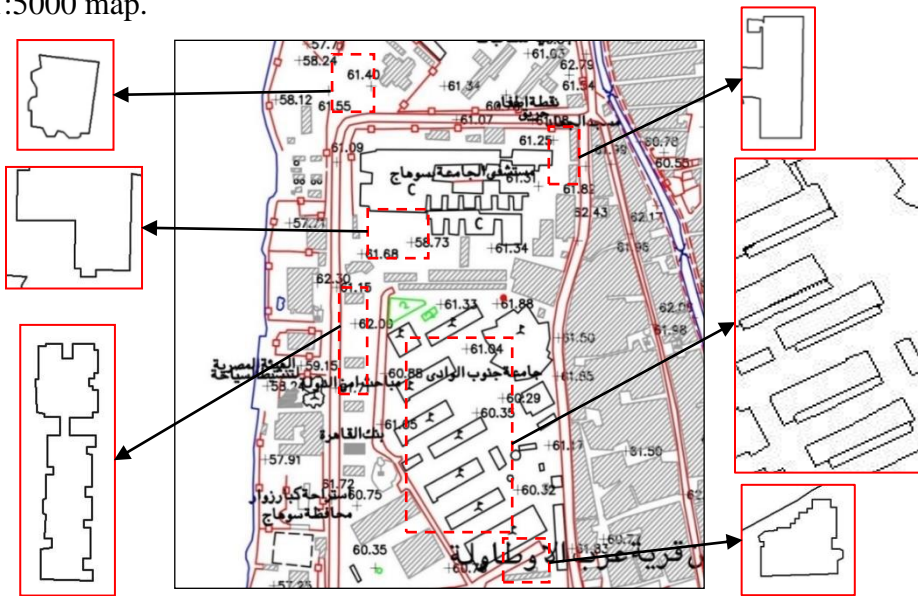


Fig. 12. The reference map showing the updated areas.

## 7. Conclusions

The main aim of this paper is to extract building from VHR satellite images for map updating in Egypt. A comparison of pixel and object-based classification were carried out. Then, the refinement process based on shadow, context, shape, and DSM is used. The Maximum Likelihood as a pixel-based classifier gives better accuracy than Mahalanobis Distance,

Minimum Distance, Spectral Angle Mapper, and Spectral Correlation Mapper. Its overall accuracy for the study area (A) and (B) was 80.49% and 76.54%, respectively. SVM as an object-based classifier gives a higher accuracy than KNN and threshold. Its overall accuracy for the study area (A) and (B) was 80.71% and 77.65%, respectively. A comparison of the results of pixel and object-based methods showed that SVM outperforms the Maximum Likelihood method. Using shadow, context information, shape information, and DSM data improve the results of SVM classification significantly. The overall accuracy for the study area (A) and (B) increased to 88.54% and 81.63%, respectively. It also increased using DSM data to 97.56% and 96.73%, respectively. The acquired building extraction results are good enough to be used for updating Egyptian maps. For future work, it's recommended to use an automatic method for generalization and reshape of the obtained vector layer to avoid manual editing.

## References

- [1] Wang, S., Hou, X., and Zhao, X. (2020). "Automatic Building Extraction From High-Resolution Aerial Imagery via Fully Convolutional Encoder-Decoder Network With Non-Local Block". *IEEE Access*, 8, 7313-7322.
- [2] Ghaffarian, S. (2014). "Automatic building detection based on supervised classification using high resolution Google Earth images". *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(3), 101.
- [3] Belgiu, M., and Drăguț, L. (2014). "Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery". *ISPRS Journal of Photogrammetry and Remote Sensing*, 96, 67-75.
- [4] Attarzadeh, R., and Momeni, M. (2012). "Object-based building extraction from high resolution satellite imager". *International Archives of the Photogrammetry*", *Remote Sensing and Spatial Information Sciences*, 39, B4.
- [5] Shedlovska, Y. I., and Hnatushenko, V. V. (2018, April). "A very high resolution satellite imagery classification algorithm". In *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)* (pp. 654-657). IEEE.
- [6] Elsharkawy, A., Elhabiby, M., and El-Sheimy, N. (2012). "New combined pixel/object-based technique for efficient urban classification using WorldView-2 data". *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39, B7.

- [7] Mostafa, F. A., Mostafa, Y., and Yousef, M. A. (2014). "The optimal method for classifying high resolution satellite images in Egypt environment". *Journal of Engineering Sciences Assiut University*, vol. 42, no. 4, pp. 1106-1121.
- [8] Jiang, N., Zhang, J. X., Li, H. T., and Lin, X. G. (2008). "Object-oriented building extraction by DSM and very high-resolution orthoimages". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37, 441-446.
- [9] Manandhar, P., Aung, Z., and Marpu, P. R. (2017, July). "Segmentation based building detection in high resolution satellite images". In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 3783-3786). IEEE.
- [10] Turker, M., and San, K. (2010). "Building detection from pan-sharpened IKONOS imagery through support vector machines classification". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 28(8), 841-846.
- [11] Benarchid, O., Raissouni, N., El Adib, S., Abbous, A., Azyat, A., Achhab, N. B., and Chahboun, A. (2013). "Building extraction using object-based classification and shadow information in very high resolution multispectral images, a case study: Tetuan, Morocco". *Canadian Journal on Image Processing and Computer Vision*, 4(1), 1-8.
- [12] ŞENARAS, Ç. (2007). "An automated building extraction model using fuzzy K-NN classifier from monocular aerial images" Doctoral dissertation, middle east technical university.
- [13] Thanh Noi, P., and Kappas, M. (2018). "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery". *Sensors*, 18(1), 18.
- [14] Mondal, A., Kundu, S., Chandniha, S. K., Shukla, R., and Mishra, P. K. (2012). "Comparison of support vector machine and maximum likelihood classification technique using satellite imagery". *International Journal of Remote Sensing and GIS*, 1(2), 116-123.
- [15] Zeng, C. (2014). "Automated Building Information Extraction and Evaluation from High-resolution Remotely Sensed Data". Doctoral dissertation, University of Western Ontario.
- [16] Abd Elwahed, A., Farrag, A., and Mostafa, Y., (2011). "Evaluation of Information Content and Feature Extraction Capability from Egypt Sat-1 Images". *Journal of Engineering Sciences, Assiut University*, 39(3), 529-538.
- [17] Perumal, K., and Bhaskaran, R. (2010). "Supervised classification performance of multispectral images". arXiv preprint arXiv:1002.4046.
- [18] Darwish, A., Leukert, K., and Reinhardt, W. (2003, July). "Image segmentation for the purpose of object-based classification". In *International Geoscience and Remote Sensing Symposium* (Vol. 3, pp. III-2039).
- [19] Nady, B., Mostafa, Y., Abas, Y., and Enieb, M., 2020 "Using Of VHR Satellite Images for Road Network Extraction in Egypt" *Journal of Engineering Sciences, Assiut University, Faculty of Engineering, Vol. 48, No. 1, pp. 20 - 31.*

- [20] Stehman, S. V. (1997). "Selecting and interpreting measures of thematic classification accuracy". *Remote sensing of Environment*, 62(1), 77-89.
- [21] Mostafa, Y., and Abdelhafiz, A. (2017). "Accurate shadow detection from high-resolution satellite images". *IEEE Geoscience and Remote Sensing Letters*, 14(4), 494-498.
- [22] Mostafa, F. A., Mostafa, Y., Yousef, M., and Abas Y., (2014), "Using of high resolution satellite images for updating large scale mapping in Egypt". *J. Eng. Sci*, 42, 1122-1137.

## استخراج المباني من مرئيات الأقمار الصناعية عالية الدقة لتحديث الخرائط في مصر

### الملخص العربي

لقد كانت عملية الكشف عن المباني من صور الأقمار الصناعية موضوع اهتمام لعدة عقود. حيث تدعم صور الأقمار الصناعية عالية الدقة الاستخراج الفعال لهذه المباني. يهدف هذا البحث الي تقديم طريقة لاستخراج المباني من صور الأقمار الصناعية عالية الدقة لتحديث الخرائط في مصر. ولتحقيق هذا الهدف، تم عمل مقارنة بين طريقتي التصنيف (object-based classification) , (Pixel-based classification). ثم تم اجراء عمليات التحسين المختلفة بالاعتماد على الظل والسياق والشكل ونموذج السطح الرقمي. تم استخدام منطقتين للدراسة من صور الاقمار الصناعية عالية الدقة لمدينتي أسيوط وسوهاج. وبالمقارنة بين طرق التصنيف اتضح أن طريقة (Maximum likelihood) لتقنية (Pixel-based) وطريقة (Support vector machine) لتقنية (object-based) حققوا أعلي النتائج للدقة الكلية. وتبين أن التحسين المعتمد على معلومات الظل والسياق والشكل ونموذج السطح الرقمي بإمكانه تحسين الدقة الكلية بمتوسط ١٨%. وبالتالي، يمكن أن تساهم نتائج استخراج المباني بشكل كبير في تحديث الخرائط في مصر.