

## Research Article

# Novel Approach for Rooftop Detection Using Support Vector Machine

**Hayk Baluyan, Bikash Joshi, Amer Al Hinai, and Wei Lee Woon**

*Computing and Information Science, Masdar Institute of Science and Technology, Masdar City, Abu Dhabi, UAE*

Correspondence should be addressed to Hayk Baluyan; [bhayk@masdar.ac.ae](mailto:bhayk@masdar.ac.ae)

Received 26 September 2013; Accepted 21 November 2013

Academic Editors: A. Gasteratos, S.-J. Horng, J. M. Tavares, and C. S. Won

Copyright © 2013 Hayk Baluyan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new method for detecting rooftops in satellite images is presented. The proposed method is based on a combination of machine learning techniques, namely,  $k$ -means clustering and support vector machines (SVM). Firstly  $k$ -means clustering is used to segment the image into a set of rooftop candidates—these are homogeneous regions in the image which are potentially associated with rooftop areas. Next, the candidates are submitted to a classification stage which determines which amongst them correspond to “true” rooftops. To achieve improved accuracy, a novel two-pass classification process is used. In the first pass, a trained SVM is used in the normal way to distinguish between rooftop and nonrooftop regions. However, this can be a challenging task, resulting in a relatively high rate of misclassification. Hence, the second pass, which we call the “histogram method,” was devised with the aim of detecting rooftops which were missed in the first pass. The performance of the model is assessed both in terms of the percentage of correctly classified candidates as well as the accuracy of the estimated rooftop area.

## 1. Introduction

Automatic rooftop detection from satellite/aerial images is an important task in a variety of applications. Interesting examples include change detection in urban monitoring, the production of digital maps, land use analysis, verification, and updating GIS databases and route planning [1, 2]. For example, accurate identification and localization of rooftops in urban images are a key step in territorial planning and city modeling. Similarly, knowledge of the location, profile, and density of buildings can be very useful in estimating the distribution of a city’s population. In particular, rooftop detection can be used to analyze the size and location of human settlements in slums and other disorganized areas [2].

However, detecting rooftops from aerial or satellite images can be very challenging. One reason is that the images used often differ in terms of lighting conditions, quality, and resolution. Another reason is that buildings may have diverse and complicated shapes and structures and as such can be easily confused with similar objects such as cars, roads, and courtyards. The result of these complications is that there

are currently no algorithms or features that are universally applicable, that is, which can be used to detect roofs in all or even a majority of aerial and satellite images.

Much of the earlier work on rooftop detection has depended on computer vision and image processing techniques such as edge detection, corner detection, and image segmentation. One widely used approach is to first generate rooftop candidates using image segmentation techniques and then to identify the true rooftops within this set of candidates, where the latter process is performed using discriminative features such as intensity, shape, and area. Ren et al. and Nosrati and Saeedi [3, 4] proposed a technique for automatic polygonal rooftop extraction based on rooftop hypothesis generation and refinement. In the hypothesis generation step rooftop candidates are generated using edge and corner detection. The generated hypotheses are refined in the second step by using features that characterize rooftops such as the standard deviation of pixel intensities within rooftop candidates, relative gray level difference between rooftop surface points and outside points. In [5] a method for automatic building detection in aerial images using hierarchical feature

based image segmentation is presented. In this approach the images are first segmented using the mean shift segmentation algorithm [6] to generate candidate building regions. In the subsequent step shadow information is used to determine if a candidate region is a rooftop. Jin and Davis [7] proposed a method based on differential morphological profile to generate building hypotheses with a verification process which used shadows and spectral information.

Many modern approaches have used machine learning to perform rooftop detection. In [8] a method is presented which used machine learning algorithms for selecting or rejecting candidate rooftops. Use of machine learning techniques facilitated better identification of true rooftops from the candidates even in the presence of noise and artifacts. In [9] a method for detecting building rooftops using LIDAR data was presented (LIDAR is a remote sensing technology that measures surface elevation using a laser). Mathematical morphological filtering is first used to separate ground and nonground objects in the image. Next, buildings and trees are classified from the nonground objects. An unbalanced support vector machine is used in the methodology since this reflects the characteristics of the data to be classified. For example, in urban areas the number of buildings in the image may be considerably more than the number of trees. Unbalanced SVM made the classification more accurate and automatic. Similarly a technique of detecting trees in urban areas has been presented in [10] using the aerial image LIDAR data. The detection of trees in urban areas is eventually used to exclude tree parts from the building rooftops for 3D city modeling. Firstly, segmentation is performed using a region growing algorithm and then trees are detected by performing classification on segments using SVM. Classification performance was assessed using ROC analysis. The performance of this methodology was then compared to other traditional approaches and was found to be better. In [11] a method is presented for detecting building damage in aerial images. The method used shadow information in addition to the spectral information to perform building damage detection. Images were first segmented using an improved watershed algorithm to produce multilevel image segments. An SVM was then used to classify the segments that were generated. It was observed that the accuracy of the presented methodology was significantly higher than the benchmark methodology which only used spectral information for classification.

Other studies have used both spectral and spatial features for the classification task (e.g., [12]). The main motivation for using both spatial and spectral features was that land cover types in urban areas are spectrally similar. So, the accuracy obtained using spectral features alone is comparatively low. It was observed that the use of spatial features along with the spectral features improved the accuracy of the building damage detection task considerably.

Based on our review of the literature, a new rooftop detection system which is novel in a number of key respects was developed. The proposed method has the following key characteristics.

- (1) It uses only *panchromatic images*. In contrast, most of the approaches mentioned in the literature have used

LIDAR data [9, 10] and/or multispectral images [11, 12], both of which are more informative but also more difficult to obtain and expensive.

- (2) It based on both *spectral and spatial features* extracted from the images.
- (3) It utilizes *machine learning techniques*, namely, *k*-means clustering to segment the image into rooftop candidates and SVM to perform classification on these candidates.
- (4) classification results obtained using the SVM are subjected to a *second-pass classification stage*. For easy reference we will refer to this as the “*histogram method*.”

There do not seem to be any existing studies which combine all four characteristics above, and we believe that presented together these represent a significant improvement over existing methods for performing rooftop detection. The original motivation for this study was to assess the available rooftop area in Abu Dhabi, for the deployment of photovoltaics, as such images from this area are used to evaluate this method.

## 2. Proposed Method

The proposed rooftop detection system consists of the following three main steps.

- (1) *Image Segmentation*. Each image is first divided into a set of candidate regions. This is done by first using *k*-means clustering, to divide pixels into a number of clusters based on colors and then using the flood-fill algorithm to group the pixels in each cluster into a set of connected components or regions. Each of these regions is now a rooftop *candidate*.
- (2) *Feature Extraction and SVM Classification*. For each candidate, 8 features are extracted using MATLAB's *regionprops* method. These extracted features form the dataset, where each row represents a single candidate region. The SVM classifier is then trained to distinguish between rooftop regions and nonrooftop regions.
- (3) *Histogram Method*. Although the trained SVM successfully detects many of the rooftops in the test images, in practice there were also many rooftop regions which were not detected (specific examples of these will be shown later). To detect initially rooftops that were missed initially, the histogram method was devised, which seeks to leverage the distribution of grayscale intensities of rooftop pixels that were correctly detected by the SVM. This method is based on the observation that rooftops which are in close proximity to one another also tend to have the same color.

The overall approach is shown in Figure 1. All the three steps listed above will now be discussed in greater detail.

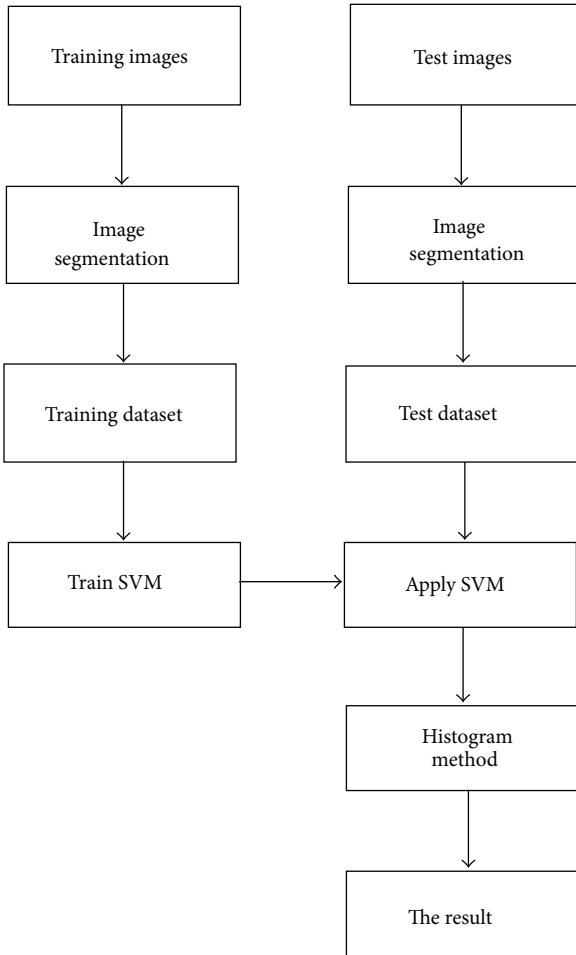


FIGURE 1: Diagram describing the overall model.

**2.1. Image Segmentation.** The goal of image segmentation is to create a set of candidate regions (segments), each of which will later be classified as rooftop or nonrooftop. To divide an image into segments we use  $k$ -means clustering [13], to divide the pixels in an image into  $k$  clusters. The clustering is based on the *color* of the pixels, where each row presented to the  $k$ -means function represents a single pixel with 3 features: the red, green, and blue component intensities.

To improve the quality of the extracted segments bilateral filtering [14] is applied prior to clustering. Bilateral filtering is a preprocessing method which seeks to remove noise while preserving edge information. It combines two filtering approaches: domain filtering, which enforces closeness by weighing pixel values with coefficients that fall off with distance, and range filtering, which averages pixel values with weights that decay with dissimilarity. The result of the bilateral filtering is shown in Figure 2(b). In the same figure the results of the  $k$ -means clustering are shown both without and with bilateral filtering (Figures 2(c) and 2(d)). It can be observed that the use of bilateral filtering results in smoother and visually “cleaner” segments. In contrast, it can be seen that many segments obtained without the use of bilateral filtering contain noticeable levels of noise.

An important consideration is the choice of an appropriate value of  $k$  (the number of clusters). As is commonly done, a range of values were tested after which it was observed that  $k = 4$  provided the best result (examples which illustrate this are shown later).

The result of applying the  $k$ -means algorithm is a labeling of each pixel in a given image into one of four different clusters (in cases where  $k = 4$ ). The next step is to convert these labeled pixels into candidate rooftops, and this is achieved by grouping them into connected regions. For this purpose, the 4-connected flood fill algorithm is applied separately to pixels from each of the four clusters—the result is a set of regions where each pixel in a region is connected to at least one other pixel in the same region via one of the four principle directions. Another option was to use the 8-connected flood fill algorithm, which permits connections via any of the 8 pixels in the immediate neighborhood of a given pixel. In practice no significant difference was observed between these two methods (an example of this is shown in Figure 3) and as such the 4-connected flood fill algorithm was used as it was computationally less demanding [15].

## 2.2. Feature Extraction and SVM Model

**2.2.1. Data Preparation.** After dividing the training images into candidate regions (segments) as described in the previous section, the dataset was constructed, in which each row represents one of the segments. Eight features were extracted to describe each segment (this is discussed in more detail in the next section). Each row is manually labeled as “1” (if it corresponds to a rooftop) or “0” (if not).

**2.2.2. Feature Extraction.** Features are numerical attributes which characterize the object to be classified. So, the extracted features are those which hold properties which can help to distinguish rooftops and nonrooftops in an image [16]. In the proposed method eight features are considered which are highly relevant to the classification task at hand. These are as follows.

- (1) *Area.* This is the area of a given segment in terms of the number of pixels. This feature can help filter out objects such as trees and cars which are simply too small to be a rooftop.
- (2) *Ratio of Minor Axis to Major Axis Lengths.* This is the ratio of width to length of a given region. In Figure 4, the major and minor axes of a building are shown in red and blue, respectively. As can be seen the lengths of the minor and major axes of the highlighted building are comparable—in comparison, objects such as roads are elongated and have very low minor to major axis ratios.
- (3) *Visible Vegetation Index (VVI).* The VVI gives an indication of the presence of vegetation in an image [10]. VVI is frequently calculated for multispectral

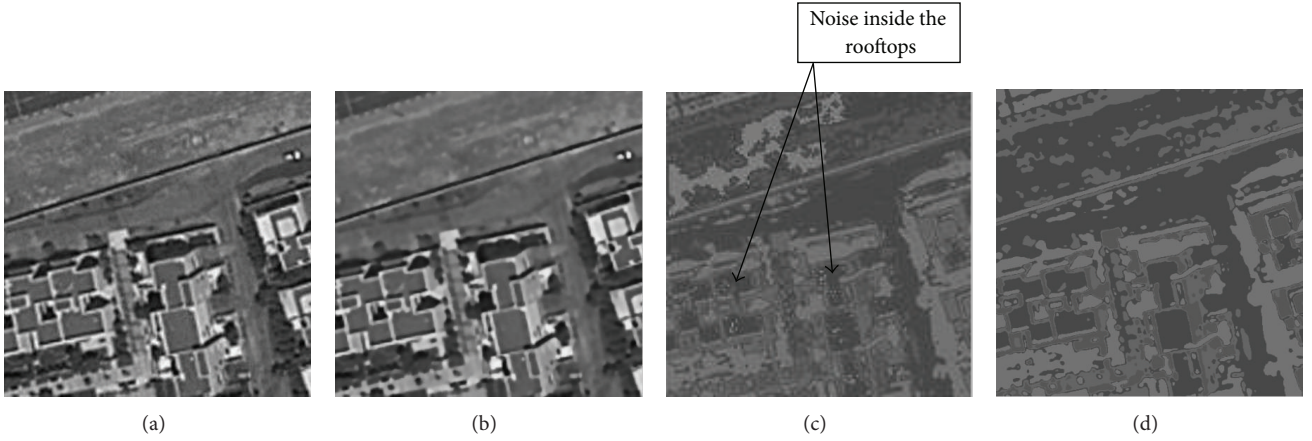


FIGURE 2: The original image (a). The result after applying bilateral filtering on the original image (b). The result of the  $k$ -means clustering with 4 clusters on the original image (c). The result of the  $k$ -means clustering with 4 clusters on the image obtained after bilateral filtering (d).

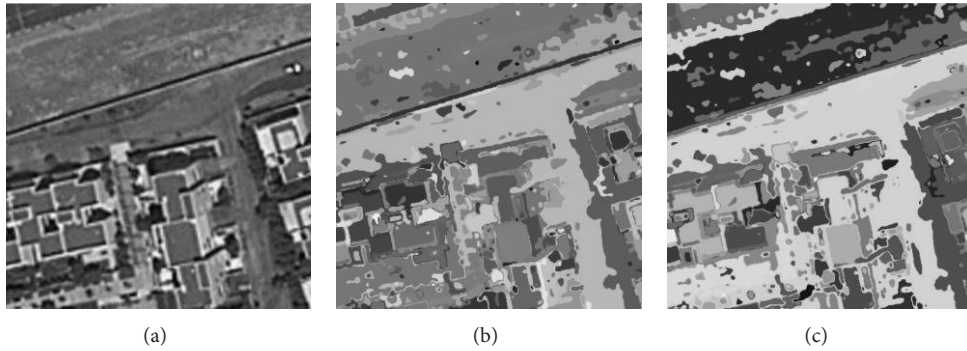


FIGURE 3: The original image (a). The result of 4-connected flood-fill algorithm (b). The result of 8-connected flood-fill algorithm (c).

images, but in the case of an RGB image it can be approximated using

VVI

$$= \left[ \left( 1 - \left| \frac{R - R_0}{R + R_0} \right| \right) \left( 1 - \left| \frac{G - G_0}{G + G_0} \right| \right) \left( 1 - \left| \frac{B - B_0}{B + B_0} \right| \right) \right]^{1/W}. \quad (1)$$

Here,  $R$ ,  $G$ , and  $B$  denote the red, green, and blue intensities in the image, whereas  $R_0$ ,  $B_0$ , and  $G_0$  are the values of red, blue and green used to reference the green color.  $W$  is used to adjust the sensitivity of the scale and is known as weight component [17].

- (4) *Solidity*. Solidity can be calculated as the ratio of the total area of a region to the area of the convex hull of the region [18]. Because most rooftops are rectangular in shape, rooftop-related regions in an image are likely to have higher values of solidity.
- (5) *Mean Intensity*. This is the mean of all the grayscale intensity values present in a region [18]. Usually mean intensity values of the rooftops are similar. As in Figure 4 the rooftops are of similar intensity which is

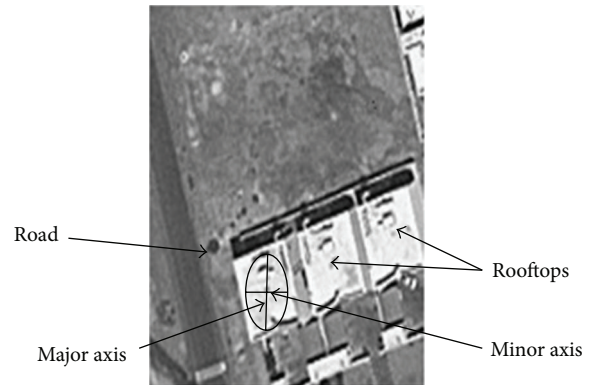


FIGURE 4: The image showing rooftop and nonrooftop objects together with the features.

different from that of other objects such as roads and vegetation.

- (6) *Variance in Intensity*. This is the variance of the pixel intensities within a segmented region. A rooftop would tend to be fairly homogeneous in appearance, and as such the corresponding region would also have

a lower variance of intensity when compared to a nonrooftop region.

- (7) *Extent*. The extent is the ratio of pixels in a given region to the total number of pixels of the bounding box. This is similar in concept to the solidity feature.
- (8) *Eccentricity*. The value of eccentricity ranges from 0 to 1. A segment having eccentricity 0 is a circle whereas segment with eccentricity 1 is a line segment. This feature can help the classifier to reject objects which are overly elongated.

Each feature was normalized by subtracting the mean of the feature and dividing it by the standard deviation.

**2.2.3. SVM Model.** The support vector machine (SVM) is a machine learning technique which finds the decision boundary (or “hyperplane”) that optimally separates the data points of one class from those of the other class, where a “hyperplane” is optimal if it maximizes the margin of separation between the two classes. Like most kernel methods, the performance of an SVM is heavily dependent on the choice of kernel function. Because of its good classification performance on our data, we used the Gaussian radial basis function kernel:

$$K(x, y) = e^{-(x-y)^2/2\sigma^2}. \quad (2)$$

Different values of sigma were tried and it was found that setting  $\sigma = 1$  produced the best results (illustrative examples are shown later).

**2.3. Histogram Method.** As already mentioned, it is likely that the SVM will not be able to detect all the rooftop regions in an image. To help address this problem, color information from the detected rooftops was subsequently used to find the “missing” rooftops.

The main idea is to use the information from the regions which were classified by the SVM as rooftops in order to detect the misclassified segments. This is based on the observation that rooftops within a single image tend to have the same pixel intensities. Hence, the idea is to use the intensity information of the segments which were classified as rooftops by the SVM, to affect a “second-pass” of classification. An example is shown in Figure 5. In Figure 5(b) the segments which have been labeled black are the ones which were classified as rooftops by the SVM and the ones which were colored by black boundary are some of rooftop segments which were misclassified as nonrooftops. As can be seen, the grayscale intensities of these misclassified regions are similar of those of the detected rooftops, which suggest the histogram method could be very useful for these situations.

Two histograms were used: one for the intensities of pixels which were classified as rooftops and another for pixels which were classified as nonrooftops. Each histogram consisted of 10 bins, which represented a reasonable balance between computational requirements, good results, and adequate coverage of each bin (in terms of pixels). The two histograms are shown in Figures 5(c) and 5(d).

From the first distribution (shown in Figure 5(c)) 2 bins were chosen which contained the most number of pixels. At this point we applied the heuristic that misclassified rooftop pixels should fall into either of these two bins or into one of the immediate neighbors. In this way we ended up examining up to 6 bins out of 10 bins (in boundary cases this can be as low as 3 bins); thus the likelihood of the misclassified pixels falling into one of these 6 bins is very high.

Considering only 2 bins also avoids adding too much noise to the model, since considering too many bins can significantly increase both true positive and false positive rates. An example is shown in Figure 6, where it can be seen that taking 3 bins with the most number of pixels helps to detect brown rooftops; however it also resulted in an increase in false positives (Figure 6(c)).

Another issue related to the histogram method is having different objects (roads, cars, and so on), which are of the same color as the rooftops (as an example see Figure 7(a)). In such cases the histogram method cannot effectively distinguish nonrooftop from rooftop regions. A similar problem is encountered when there are no rooftops on the image at all (see Figure 7(b)). In such cases the histogram method will admit large numbers of false positives.

In order to avoid the situations discussed above a thresholding scheme was applied. The scheme adopted is based on the fact that the aim of the histogram method is to complement SVM classification; if the number of nondetected pixels in a bin is significantly higher than that of detected pixels in the same bin, there would be little sense in using that bin. For example, from Figure 5(d) it can be observed that there are 10000 nondetected pixels in the fifth bin; however from Figure 5(c) we have only less than 200 detected pixels in the fifth bin. Thus, the number of nonrooftop pixels (based on the SVM classification) is greater than the number of rooftop pixels by factor 50. It was found that applying a threshold to this ratio was very useful in avoiding cases like this. In our case setting a threshold of 15 proved to be the best choice for our datasets, though this remains as a tunable parameter which needs to empirically set when used with different datasets.

The results will be discussed in greater detail in the next section, but briefly our observation was that the “histogram” method performed very well for one of our datasets, where it resulted in a big increase in performance. Unfortunately for the second dataset this method did not perform as well; however, even in this case it still produced a slight improvement in performance. Suggested reasons for this will be presented later on in the paper.

### 3. Results and Discussion

**3.1. Data.** As explained earlier, one of the aims of research was for the proposed method to be able to work using only panchromatic data. Such data can be obtained from a variety of commercial sources, but for this study images that were manually collected from Google Maps were used. Since this paper was focused on finding the total amount of rooftop area for deployment of photovoltaics in Abu Dhabi, UAE, we use

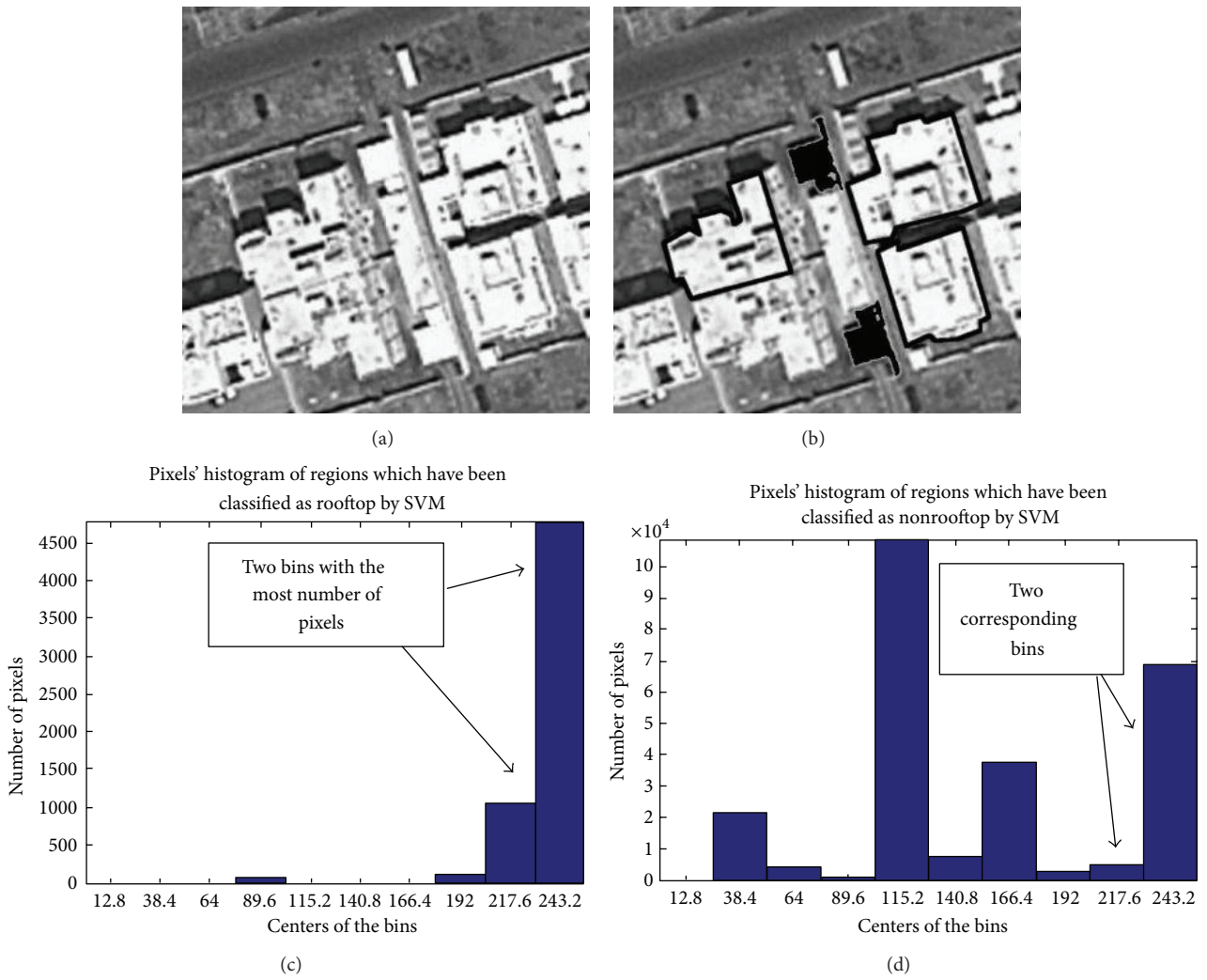


FIGURE 5: The original image (a) .The result after the SVM (b) .The distribution of intensities of the detected rooftop pixels (c) .The distribution of intensities of the nondetected rooftop pixels (d).



FIGURE 6: The original image (a) .The result when considering 2 bins (b) .The result when considering 3 bins (c).

images gathered from selected residential areas in Abu Dhabi city. To ensure the generality of our model, it was tested on two separate datasets, “Raha” and “Khalifa,” which consist of images gathered from Al Raha Gardens and Khalifa City A.

For the segmentation process to work properly, the  $k$ -means algorithm had to be provided with images of an appropriate size. For this study, satellite images were divided into small tiles with  $512 \times 512$  pixels, which corresponded to a plot of land measuring  $70 \text{ m} \times 70 \text{ m}$ . This size was chosen because it provided a pragmatic balance between being small enough such that the  $k$ -means algorithm could be effective, while still being large enough such that each tile typically contained a number of houses and hence roofs. The second issue was important as it meant that rooftops were rarely split between neighboring tiles.

14 such images were collected for each dataset, out of which 8 were used for training and the remaining 6 images were used for testing and validation. In addition rooftops in each of these images were manually labeled and these labels were subsequently used to label the regions extracted during the segmentation process, where each rooftop region is labeled “1” and nonrooftop regions “0.”

Figure 8 shows examples of images from both datasets and also an example of a manually labeled image. As can be seen, many objects (such as cars and roads) have pixel intensities which are very similar to rooftops and as such our model needs to be able to distinguish these objects from true rooftops. For example for “Raha” images it is obvious that roads have almost the same color as most of the rooftops (see Figure 8(a)) and for “Khalifa” images there are many brown regions which look like rooftops; however they are not (see Figure 8(b)).

**3.2. Experimental Results.** Commonly adopted performance metrics were used to evaluate the performance of the system. These are Precision, Recall and  $F_1$  score, which are defined as shown here:

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\%, \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\%, \\ F_1 &= \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} * 100\%. \end{aligned} \quad (3)$$

Here TP, TN, FP, FN are, respectively, true positive, true negative, false positive, and false negative rates.

As mentioned, to determine the optimal value of  $\sigma$ , the performance in terms of  $F_1$  score was calculated for a range of values of  $\sigma$ . Results for Raha and Khalifa datasets are shown in Figures 9(a) and 9(b). As we mentioned previously, it can be seen that  $\sigma = 1$  results in the best performance.

As might be expected, it can be seen that the accuracy of the SVM grows with the size of the training dataset. The relationship between the  $F_1$  score for the SVM on the training dataset and the number of training images used is shown in Figure 10.

TABLE 1: The results for trained SVM on Al Raha Gardens (a) and Khalifa City A (b) validation datasets.

(a)					
Number of clusters	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
Precision (%)	60.5	68.19	79.66	73.5	67.39
Recall (%)	86.62	84.9	83.71	72.4	68.22
$F_1$ score (%)	71.24	75.6	81.63	72.94	67.8
(b)					
Number of clusters	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
Precision (%)	64.5	68.2	77.2	75.5	69.12
Recall (%)	69.2	73.72	88.55	73.4	69.4
$F_1$ score (%)	66.76	70.85	82.48	74.43	69.25

TABLE 2: The numbers of detected rooftops before and after applying the histogram method.

	Before the histogram method	After the histogram method
Image 1 from Raha	12 out of 17	14 out of 17
Image 2 from Raha	9 out of 14	10 out of 14
Image 3 from Raha	7 out of 12	12 out of 12
Image 1 from Khalifa	4 out of 13	13 out of 13
Image 2 from Khalifa	6 out of 15	14 out of 15
Image 3 from Khalifa	4 out of 12	10 out of 12

While accuracy increases with the number of images used, this seems to level off after around 8 images and this was hence deemed to be sufficient amount of training data.

Finally, there was also the issue of the suitable value of  $k$  to be used when performing segmentation. As was already mentioned we have tried different values of  $k$  and found  $k = 4$  to be the most suitable in our case. The performance of the SVM with parameter  $\sigma = 1$  for different number of clusters on the validation dataset is shown in Table 1, where the best result ( $F_1$  score equal 82%) for both validation sets can be seen.

We evaluate the overall performance of our method based on two criteria: the number of detected rooftops and the overall area covered by detected rooftops. We compare the results before and after applying the histogram method. In Table 2 the results for all 6 testing images are given.

It can be observed that the SVM performs quite well on “Raha” images even without using the histogram method. However for “Khalifa” images the performance of the basic SVM is weak and the histogram method produces a huge improvement for “Khalifa” datasets. One possible reason for this is that rooftops on “Raha” images are well separated from each other by white boundaries (see Figure 8(a)). Hence the image segmentation step often results in candidate regions which in general represent single rooftop. Since all rooftops

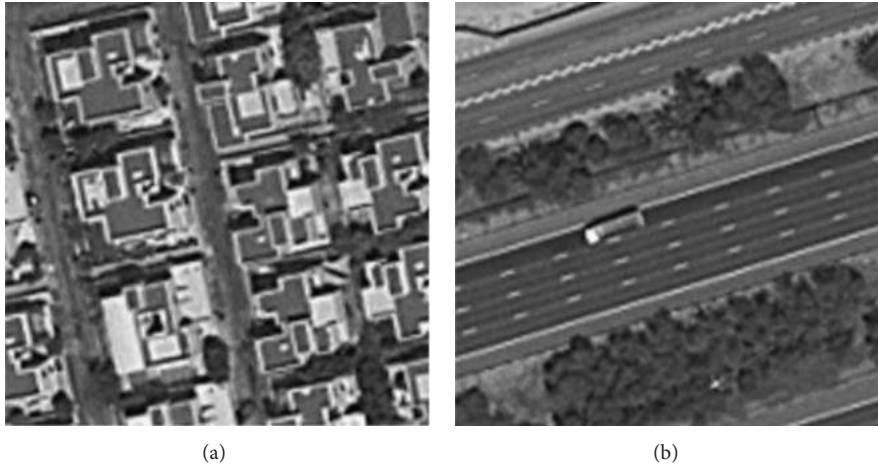


FIGURE 7: The image with roads of the same intensity as the rooftops (a). The image without any rooftop (b).

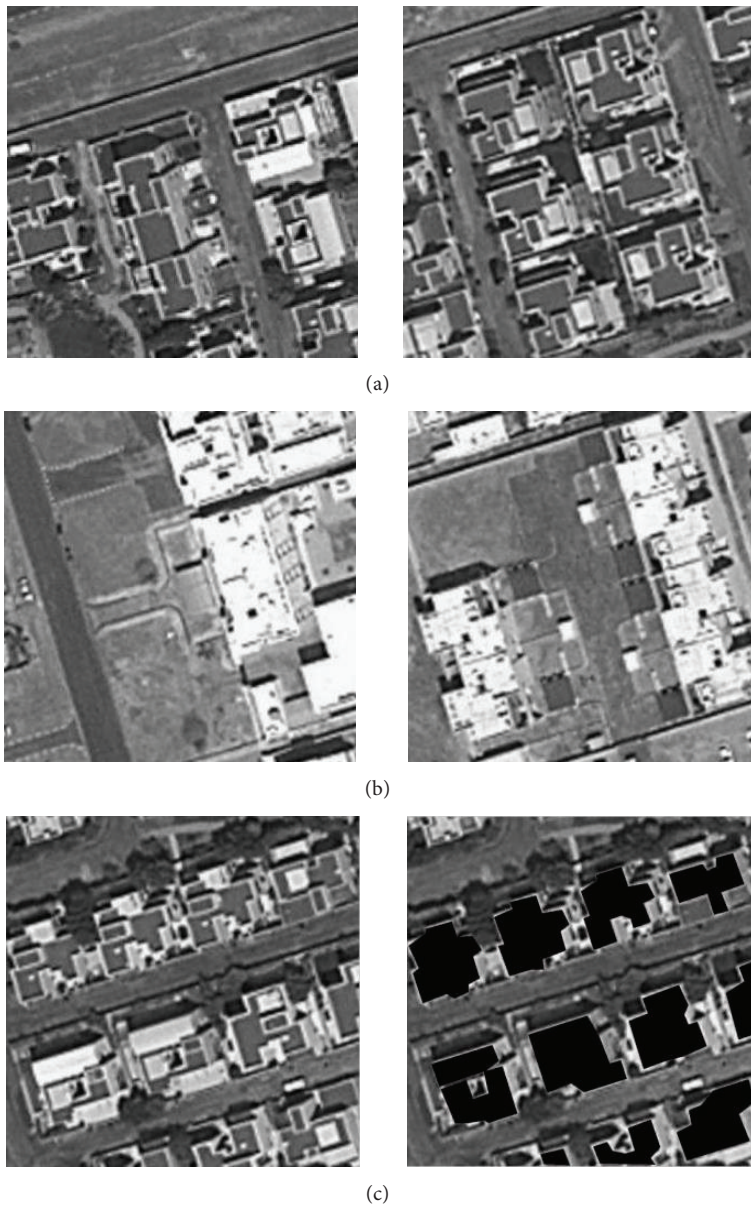


FIGURE 8: Sample images from Al Raha Gardens (a). Sample images from Khalifa City A (b). An example of a manually labeled image (c).



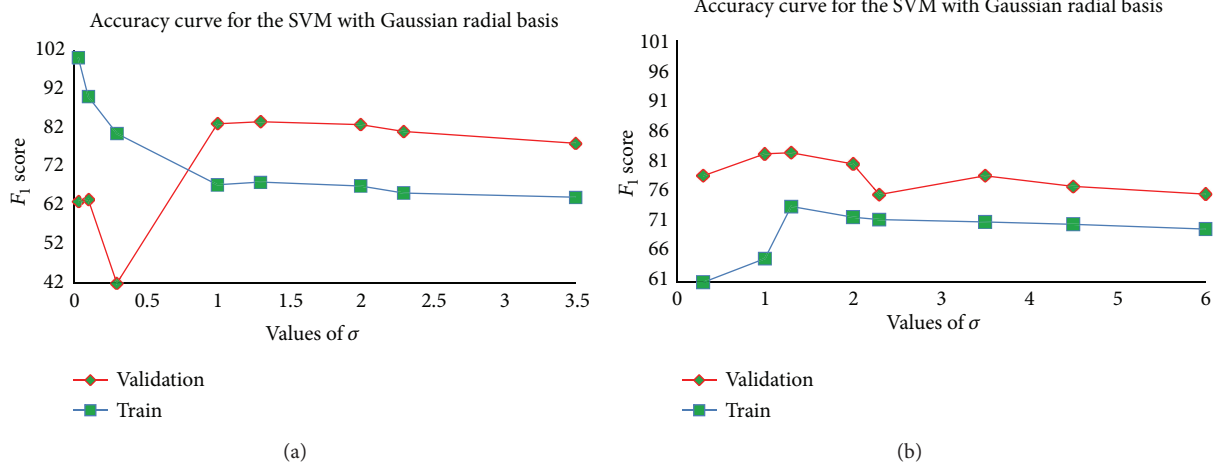


FIGURE 9:  $F_1$  score for different values of sigma for “Raha” (a) and “Khalifa” datasets (b).

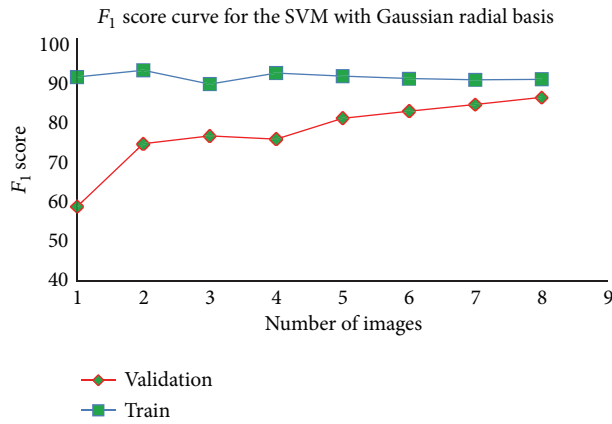


FIGURE 10:  $F_1$  score for different number of training images for “Khalifa” dataset.

TABLE 3: The amount of detected rooftop pixels before and after applying the histogram method for Khalifa City A (a) and Al Raha Gardens test images (b).

(a)		
	Before the histogram method	After the histogram method
Precision (%)	92.8	93.16
Recall (%)	52.4	70.01
$F_1$ score (%)	66.9	79.9
(b)		
	Before the histogram method	After the histogram method
Precision (%)	88.5	97.6
Recall (%)	7.1	79.7
$F_1$ score (%)	13.1	87.7

have almost similar values for the extracted features, it makes the job of the SVM to make better classification easier.

In contrast in “Khalifa” images rooftops are not separated from each other very well (see Figure 8(b)); hence after the image segmentation step it is possible that 2 and more rooftops will be represented in a single candidate region, which forces the SVM to consider such candidate regions as outliers. Since rooftops in the Al Raha Gardens region have almost the same intensity as roads, cars, and other nonrooftop objects, the “histogram” method was frequently unable to detected rooftops that were not already detected using the other features. In contrast the rooftops in the Khalifa City A images are quite distinct in terms of the intensities of the corresponding regions, and this allowed the “histogram” method to make a significant contribution.

To better evaluate the overall performance of the model, results based on the correctly classified rooftop and non-rooftop pixels are presented in Table 3.

Again it can be observed that in contrast to the “Raha” dataset, the histogram method significantly improves the performance of the system on the “Khalifa” dataset (though we still see a slight improvement in the case of the “Raha” dataset).

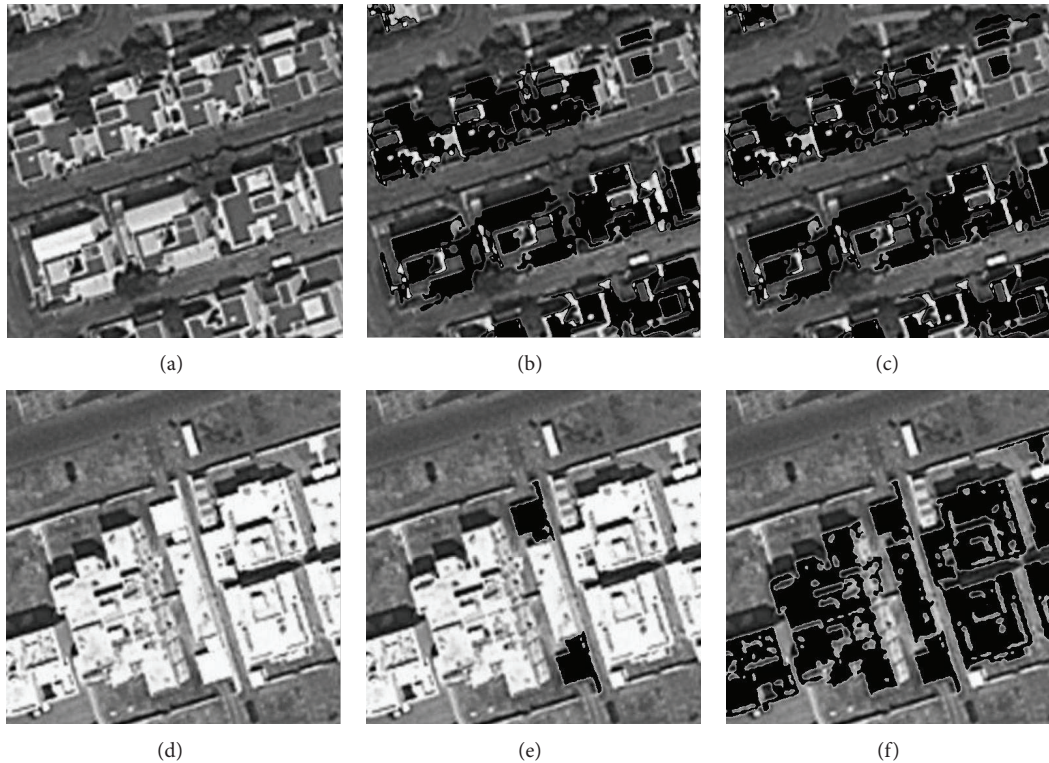


FIGURE 11: The original image from Raha Gardens (a). The result after the SVM (b). The result after the “histogram method” (c). The original image from Khalifa City A (d). The result after the SVM (e). The result after the “histogram method” (f).

More results of the performance of our model can be seen in Figure 11. It can be seen that there is not a big difference between Figures 11(b) and 11(c) since the basic SVM already performs well and the task of the histogram method in this case is to avoid the inclusion of additional noise. However a significant difference can be observed between Figures 11(e) and 11(f), which shows how the histogram method significantly improves the performance of the system.

#### 4. Conclusion

The paper presented a new approach for detecting rooftops using machine learning techniques like  $k$ -means and SVM. While the results are still preliminary, we showed that the proposed method was able to retrieve a very high percentage of the rooftops present in an image while at the same time maintaining a low false positive rate. The method gives especially good results when all the rooftops in the image are of a similar color or gray level intensity. A unique feature of this method is its use of the “histogram method” to find rooftops which were initially missed by the SVM.

However there are still some situations in which the method does not perform well. For example rooftops which are very big relative to the image size were sometimes classified as nonrooftop by the SVM, which tended to consider such rooftops as outliers.

Another weakness of the method is poor performance when rooftops of many different colors are encountered. Also,

when there is a single “dominant” rooftop color, it renders the system less sensitive to rooftops with less common colors.

For future work we intend to extend the system along three main directions:

- (1) improvements to the classification process via additional feature engineering to discover more informative features and screening and testing alternative classifiers, such as the unbalanced SVM used in [9];
- (2) the addition of a higher-order classification stage. Rooftops which are in close proximity to each other tend to have similar characteristics (color, design, orientation, density, etc). While the histogram method is a step in this direction there are other characteristics beyond simply the grayscale intensity;
- (3) testing and extension of the method to larger geographical areas.

#### References

- [1] H. V. Guducu, “Building Detection from Satellite Images Using Shadow and Color Information,” 2008.
- [2] H. G. Akçay and S. Aksoy, “Building detection using directional spatial constraints,” in *Proceedings of the 30th IEEE International Geoscience and Remote Sensing Symposium (IGARSS '10)*, pp. 1932–1935, July 2010.
- [3] K. Ren, H. Sun, Q. Jia, and J. Shi, “Building recognition from aerial images combining segmentation and shadow,” in *Proceedings of the IEEE International Conference on Intelligent*

- Computing and Intelligent Systems (ICIS '09)*, pp. 578–582, chn, November 2009.
- [4] M. S. Nosrati and P. Saeedi, “A novel approach for polygonal rooftop detection in satellite/aerial imageries,” in *Proceedings of the 16th IEEE International Conference on Image Processing (ICIP '09)*, pp. 1709–1712, November 2009.
  - [5] M. Izadi and P. Saeedi, “Automatic building detection in aerial images using a hierarchical feature based image segmentation,” in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*, pp. 472–475, August 2010.
  - [6] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
  - [7] X. Jin and C. H. Davis, “Automated building extraction from high-resolution satellite imagery in Urban areas using structural, contextual, and spectral information,” *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 14, pp. 2196–2206, 2005.
  - [8] M. A. Maloof, P. Langley, T. O. Binford, R. Nevatia, and S. Sage, “Improved rooftop detection in aerial images with machine learning,” *Machine Learning*, vol. 53, no. 1-2, pp. 157–191, 2003.
  - [9] W. Xin, “A new classification method for LIDAR data based on unbalanced support vector machine,” in *Proceedings of the International Symposium on Image and Data Fusion (ISIDF '11)*, August 2011.
  - [10] J. Secord and A. Zakhor, “Tree detection in urban regions using aerial lidar and image data,” *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 2, pp. 196–200, 2007.
  - [11] P. Li, B. Song, and H. Xu, “Urban building damage detection from very high resolution imagery by One-Class SVM and shadow information,” in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '11)*, pp. 1409–1412, July 2011.
  - [12] P. Li, H. Xu, S. Liu, and J. Guo, “Urban building damage detection from very high resolution imagery using one-class SVM and spatial relations,” in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '09)*, pp. VII2–VII4, July 2009.
  - [13] K. Alsabti, S. Ranka, and V. Singh, “An efficient  $k$ -means clustering algorithm,” in *Proceedings of the IPDS/SPDP Workshop on High Performance data Mining*, April 1998.
  - [14] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proceedings of the 1998 IEEE 6th International Conference on Computer Vision*, pp. 839–846, January 1998.
  - [15] A. H. Al-Fayadh, H. R. Mohamed, and R. S. Al-Shimsah, “CT angiography image segmentation by mean shift algorithm and contour with connected components image,” *International Journal of Scientific and Engineering*, vol. 3, no. 8, pp. 4–9, 2012.
  - [16] Feature Selection, Wikipedia, 2013, [http://en.wikipedia.org/wiki/Feature\\_selection](http://en.wikipedia.org/wiki/Feature_selection).
  - [17] Visible Vegetation Index, <http://phl.upr.edu/projects/visible-vegetation-index-vvi>.
  - [18] Mathworks, 2013, <http://www.mathworks.com/help/images/ref/regionprops.html#bqkf8jf>.

