Automatic Rooftop Detection Using a Two-Stage Classification

Bikash Joshi
Hayk Baluyan
Amer Al Hinai
Wei Lee Woon

Technical Report DNA #2013-02

December 2013

Data & Network Analytics Research Group (DNA)
Electrical Engineering and Computer Science,
Masdar Institute of Science and Technology,
PO Box 54224, Abu Dhabi, UAE.

Automatic Rooftop Detection using a Two-stage Classification

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

Masdar Institute of Science and Technology - Department of Electrical Engineering and Computer Science (EECS), Abu Dhabi, UAE

Abstract. This paper presents a novel application of machine learning techniques to automatically detect building rooftops in images obtained from Google Earth. The proposed technique is based on two separate classification stages. First, the image is segmented into homogeneous regions using the k-means algorithm. These segments are then treated as candidate rooftop regions; features are extracted from each segment and submitted to a MLP which serves as the first stage of the classification procedure. New features are then extracted from the outputs of the MLP and these are presented to an SVM which then performs the second classification pass. In this way, the first classification stage acts as a preprocessing step which, in combination with the SVM stage significantly reduces the number of false-positives. To establish the efficacy of the proposed method, its results are compared with those obtained using other approaches presented in the literature.

1 Introduction

Automatic rooftop detection from satellite/aerial images is important in a variety of applications. Some examples include change detection in urban monitoring, the production of digital maps, land use analysis and route planning. However some of the challenges associated with rooftop detection include different lighting conditions, quality and resolution of the images. The result of these complications is that there are currently no algorithms or features that are universally applicable, i.e. which can be used to detect roofs in all or even a majority of aerial and satellite images.

Computational solutions for rooftop detection are based on image processing operations such as edge detection, corner detection and image segmentation. The basic approach is to first generate rooftop candidates using image segmentation techniques and then to identify true rooftops using features like shape, area and the presence of shadow [1][2]. Machine learning methods have also proved popular in recent studies, where Artificial Neural Networks (ANN) and Support Vector Machines (SVM) have been widely used. In [3] an ANN is used to facilitate effective rooftop identification in the presence of noise and artifacts. In [4] a review is presented of ANN based approaches in various related areas including shape segmentation in medical images, biometric patterns and gestures extraction, letters and characters detection, edge and contour detection. For each of these applications ANNs performed reasonably well, thus demonstrating its potential for solving image processing tasks. Neural network based methods have also been used to segment images into parts that meet certain criteria

[5]. Most of these approaches are based on pixel-level segmentation, which assign each pixel to a given segment based on features generated for the pixel in question. However this approach does not perform well for rooftop detection; in particular these methods fail to detect rooftops when applied to test images containing objects (e.g. cars, roads) which are of the same color as rooftops.

To address these apparent shortcomings, we propose here a novel method which is based on two consecutive classification stages. First, an MLP is used to classify candidates segments into rooftop and non-rooftop. New features are then extracted from the outputs of the MLP which are based on rooftop properties inferred using results gathered over the entire image. The second classification stage is then performed on these features using an SVM. While this is not unique, we also note that there appear to be relatively few studies in which SVMs are used for object detection (notable examples are [6] and [7]). Dividing the classification process into two separate stages appears to reduce the number of false positives, resulting in a significant improvement in the performance of the model in comparison with traditional approaches.

This approach of taking information from the classification result and using it for final prediction worked successfully in our previous work [8] as well. In this new approach we extend our previous work and perform the final prediction based on another pass of classification.

2 The Methodology

The proposed methodology consists of three main steps: image segmentation, first-stage then second-stage classification. We will discuss each of these steps in greater detail in the following subsections:

2.1 Image Segmentation

Prior to the segmentation we apply bilateral filtering to the image in order to enhance it.Then we divide the image into a set of segments using k-means clustering. Each of these segments further is considered as a candidate region for being a rooftop or part of rooftop. The clustering of color images is done on the basis of red, green and blue component intensities.

By trying different values of k, it was determined that k=4 gave the best segmentation result for our images. After all the pixels had been divided into 4 clusters, candidate regions were generated by finding connected-regions - i.e. regions where pixels of the same cluster were adjacent to each other. The 4-connected flood fill algorithm was used for this purpose.

2.2 First-stage classification

2.2.1 Data Preparation and Feature Extraction

A set of 10 training images was prepared by manually labelling rooftops present in these images. Each of these training images were then segmented into regions from which a set of 14 features were extracted. Features are numerical attributes which allow rooftop and non-rooftop regions to be distinguished from each other. In this study, 14 features which were relevant to the classification task at hand were selected. These are: (1) Area (2) Minor to major axis ratio (3) Mean intensity (4) Solidity (5) Variance (6) Entropy (7) Roundness (8) Rectangularity (9) Contrast (10) Correlation (11) Entropy (12) Homogeneity (13) Number of Corners (14) Energy. Each row in the dataset hence corresponds to one image segment and is manually labeled as "1" (if it corresponds to a rooftop) or "0" (if not).

2.2.2 Classification using Multi-layer Perceptron(MLP)

A Multi-layer Perceptron is a feedforward artificial neural network with at least one "hidden" layer, nonlinear activation functions and the capability to approximate arbitrary decision boundaries. A MLP is first trained using the training dataset then is presented with each test image. The output will then be the predicted labels and class probabilities for each candidate region.

2.3 Second Classification Stage

In practice, the MLP will not be able to detect all of the rooftop regions in the image. In particular we noticed the presence of many false positives in this first pass of labeled regions. To address this problem, a second classification stage is performed which uses information from the results of the first classification stage. This also helps to identify some of the rooftops which were missed in the first-stage classification.

2.3.1 Feature "Re-extraction"

The following features are used for the second stage classification:

- Class Probability: We use the class probability given by MLP model for each segments of the test image as one feature.
- Confidence Value: Rooftops in one region are similar to each other. So, the result of first-stage classification shows the most prevalent intensity value of rooftop. So, we take a histogram of the rooftops identified by first-stage classification by dividing it into 8 bins. Then we provide each bin a confidence value based on the number of pixels belonging to that bin. So, for second-stage classification we calculate the confidence value for each segment by checking to which bin it belongs.
- Shadow: We also use building shadows as a feature for the second-stage classification. It is obvious that a building rooftop region should be accompanied by its shadow.

In particular, note that the second item in the list above serves to incorporate information from the entire image into the classification of individual regions. It was felt that this was of particular importance in improving the performance of our method.

2.3.2 Classification using Support Vector Machine(SVM)

SVM is a supervised learning technique which finds an optimal decision boundary to separate the feature space into the corresponding classes. One important consideration during the use of SVM is the choice of kernel function. Preliminary investigations found that the polynomial kernel function gave the best results and this was used for all subsequent experiments.

First, we trained a SVM using the training images. Then we test each of our test images using the trained SVM model. Thus, we get our final rooftop and non-rooftop labels.

3 Result and Discussion

3.1 Data

For the experiments described here, images gathered from Google Earth were used. The research presented here was part of a broader effort to determine the total rooftop area in Abu Dhabi which was available for solar PV installation; as such, images depicting residential areas in Abu Dhabi were used. To ensure the generality of our model it was tested with two image datasets with some differing characteristics; these are referred to as "Khalifa" and "Raha", which are names of the corresponding regions in Abu Dhabi.

To keep the number of clusters manageable and thus facilitate proper segmentation, the satellite images of these areas were divided into 512 x 512 pixel sized tiles, which corresponds to an area of approximatelly 70m x 70m. The Khalifa dataset consists of 25 such images of which 10 are used for training and 15 for testing; the Raha dataset consists of 23 such images of which 10 are used for training and 13 for testing. Rooftops in these images are labeled manually and these labels are subsequently used to label each segment as either a rooftop("1") or a non-rooftop("0").

3.2 Experimental Results

We used commonly adopted performance metrics for evaluation. The used metrics are Precision and Recall, which can be defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{Recall} = \frac{TP}{TP + FN}$$

Where TP,FP,TN and FN are True Positive, False Positive, True Negative and False Negative respectively.

Sample result of our algorithm is shown in Fig. 1 It can be seen that the result of first-stage classification contains many false positive regions. This is what our second-stage classification helps to improve. Second-stage classification reduces false positives largely, which in turn significantly improves the precision.

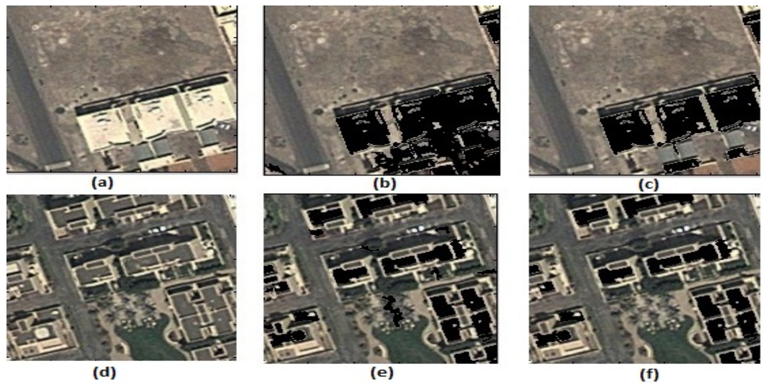

Fig. 1: The original image from Khalifa City A (a) The result after first-stage classification (b) The result after the second-stage classification (c) The original image from Raha Gardens (d) The result after first-stage classification (e) The result after the second-stage classification (f)

	Result of First-Stage			Result of Second-Stage		
	Precision	Recall	FP	Precision	Recall	FP
Khalifa	89.95	82.39	9.21	92.24	81.16	6.82
Raha	87.78	86.69	12.16	92.29	83.08	7.07

Table 1: Results using Khalifa and Raha test dataset

Even though there is slight decrease in recall values as some of the rooftops identified in the first-stage may be lost, the improvement in precision and false positives is significant.
The overall performance of our method in the "Khalifa" and "Raha" datasets is shown Tab. 1. This shows the overall precision, recall and FP values for the test datasets.

3.3 Comparison with other similar works

We have compared our method with other approaches presented in the comparison section of [9]. The comparison is shown in Tab. 2. As can be seen, our approach produced broadly better performance than the approaches presented in the comparison table. However, while promising, no definitive conclusion can be drawn from this as the scores reported were obtained using different types and numbers of images.

Metrics	Lefevre	Muller	Persson	Shorter	Our Methodology	
					Khalifa	Raha
Precision	79.4 %	79.5 %	93.0 %	51.6 %	92.24 %	92.29 %
Recall	63.6 %	77.3 %	53.0 %	78.7 %	81.16 %	83.08 %

Table 2: Comparison with Existing Approaches

4 Conclusion and Future Work

This paper presented a new approach for rooftop detection using a novel two-stage classification procedure. We showed that this methodology was able to detect a high percentage of rooftops and the second-stage classification really helped to reduce the false positives returned by the first classification stage.

One weakness of this approach was that a small number of cases where the first-stage classification was very accurate, the second classification stage actually removed some of the true rooftops. For future work we plan to test this method on images from different geographical locations and also to test the method with more complex images and over a larger geographical area.

References

[1] M.S. Nosrati and P. Saeedi. A novel approach for polygonal rooftop detection in satellite/aerial imageries. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1709–1712, 2009.

[2] M. Izadi and P. Saeedi. Automatic building detection in aerial images using a hierarchical feature based image segmentation. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 472–475, 2010.

[3] M. A. Maloof, P. Langley, T. O. Binford, R. Nevatia, and S. Sage. Improved rooftop detection in aerial images with machine learning. In *Machine Learning*, 2003.

[4] Juan A Ramírez-Quintana, Mario I Chacon-Murguia, and Jose F Chacon-Hinojos. Artificial neural image processing applications: A survey. *Engineering Letters*, 20(1):68, 2012.

[5] Michael Egmont-Petersen, Dick de Ridder, and Heinz Handels. Image processing with neural networks a review. *Pattern recognition*, 35(10):2279–2301, 2002.

[6] John Secord and Avideh Zakhor. Tree detection in urban regions using aerial lidar and image data. *Geoscience and Remote Sensing Letters, IEEE*, 4(2):196–200, 2007.

[7] Peijun Li, Benqin Song, and Haiqing Xu. Urban building damage detection from very high resolution imagery by one-class svm and shadow information. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 1409–1412, 2011.

[8] Hayk Baluyan, Bikash Joshi, Amer Al Hinai, and Wei Lee Woon. Novel approach for rooftop detection using support vector machine. *ISRN Machine Vision*, (In Press), 2013.

[9] Nicholas Shorter and Takis Kasparis. Automatic vegetation identification and building detection from a single nadir aerial image. *Remote Sensing*, 1(4):731–757, 2009.