Detection of Solar Panels from Satellite Imagery

Raghav Saboo

December 7, 2015

Abstract

Due to the fast falling prices of Solar PV Panels over the past 5 years, the number of household and commercial scale PV installations have grown exponentially. This trend is expected to continue over the coming decades and there is now an urgent need for a robust system to collect information on such installations. The research presented here provides an automated method for quickly developing a granular database of Solar PV installations using readily available satellite orthoimagery. The method is based on the latest in image recognition research, with feature engineering tailored for the problem of detecting solar panels with varying orientations and sizes from complex images with multiple objects as usually found in satellite images. The accuracy of the method using a small dataset achieved a high true positive rate of 95.6% and a low false positive rate of 3.1% . The next steps are to implement the method on a larger dataset, and adding the capability of estimating capacity of the installed systems.

1 Introduction

With the growth of Solar PV installations at the household and commercial scale, small scale solar power generation is quickly becoming a viable alternative to conventional sources of electricity. This is evident, for example, from the success story of the California Solar Initiative (CSI) which resulted in estimated installed capacity to rise by 12 times between 2006 and 2014.[1] However, as many of the installations are not followed up by strict reporting or documentation requirements, electricity utilities currently lack complete information about the installations of solar panels at households and commercial building they supply. This makes load planning difficult for them. Similarly policy makers and urban planners require more granular data on the growth of Solar PV installations across USA to understand how to improve current systems in place to promote renewable energy adoption as well as make decisions on the development of the electricity grid infrastructure.

The information that is currently available on Solar PV installations is largely based on surveys done by the Solar Energy Industries Association (SEIA) and the Energy Information Administration (EIA). The methods used are usually manual in the form of self-reported surveys or tax rebate forms. This means that many installations go unreported and even when reported the data can be sparse with limited information on the exact location of the installation.

To overcome this hurdle an automated solution was developed based on image recognition techniques leveraging satellite orthoimagery (aerial imagery taken orthogonal to the surface of the Earth) to remotely detect Solar PV panels. The approach uses various texture based features as well as shape based features in order to account for the variance in size and orientation of solar panels as well as the noise created by the complex range of objects found in satellite imagery. In order to create an initial dataset to train and test the algorithm, images from Google Maps were taken of 100 rooftops and hand labeled for the presence of Solar PV panels. The images were then preprocessed to obtain target regions to classify, and these target regions were normalized by rotation and color space before features were extracted for classification using an SVM classifier which was scored under a 100-fold cross validation.

2 Dataset and Preprocessing

The data set consists of 100 satellite orthoimages of rooftops, with 50 of the rooftops having solar panels on them. In order to determine the target regions that needed to be classified in each image a Maximally Stable Extremal Regions (MSER) algorithm was implemented on each image. This resulted in target regions, and in total 606 target regions were detected.

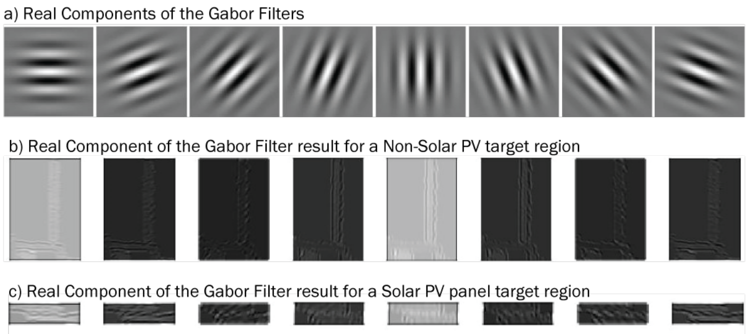
However, it was noticed that the prescreener retains duplicate or overlapping regions. In order to remove these regions, a Mean Shift algorithm was implemented utilizing the centroid of each detected region. This preprocessing step produced only unique MSER target regions and in total there were 318 unique target regions identified. These regions were then cropped, the histogram of the RBG image normalized, and finally they were also rotation normalized around the center of mass. After this step, with each potential region containing the identified solar panels, the algorithm looks to classify the regions based on the set of features extracted from each of the regions as explained in the next section.

3 Features Engineering

3.1 Gabor Filters based Texture Features

Gabor filters are linear filters whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Gabor filters are very commonly used to obtain texture features as they are robust to texture orientation and scale.[2] As it is predictable that solar photovoltiac panels will have very unique textures as compared to other objects commonly found in satellite imagery such as roads, roof tiling, or foilage. Thus an ensemble of Gabor filters were used, as shown in the figure below, and they were then convolved with the crops of the unique MSER regions. This produced arrays of high dimensional complex number features which were then used to generate template cluster maps for regions that are solar panels and regions that are non-solar panel regions.

Figure 3.1: Gabor filter ensemble and the examples of the results of convolution with MSER regions



3.1.1 Clusters using K-Mediods

After the creation of the filter responses for all the MSER regions, a subset of 10 solar panel regions and 10 non-solar panel regions were randomly selected as a training set to generate template cluster maps of the filter responses. Two filter response maps, one for solar panel regions and one for non-solar panel regions, were created using the phase and magnitude of the responses. Then a K-Mediods algorithm was run on these maps to generate 4 clusters in each map, and the feature generated from this was the histogram of the distance of points on the map from their cluster centroids. What was evident, after repeating the procedure above on multiple random subsamples of MSER regions, was that the map for solar panel regions regularly generated two distinct peaks on the histogram whereas the map for non-solar panel regions did not. Thus using a chi-squared comparison test,[3] the gabor filter response map histogram of the test regions were compared to each of the template histograms. Based on this statistic a classification score based on match with solar panel histogram template was generated as a feature.

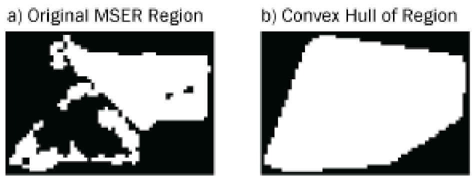
3.1.2 Clusters using Gaussian Mixture Models

Using the same filter response maps generated above, another measure of pattern match was generated using the Gaussian Mixture Models algorithm. The filter maps were again clustered into 4 regions, and each region was assigned a bivariate gaussian distribution based on being a solar panel region or being a non-solar panel region. These template likelihoods generated from the training regions were then used to compute the posterior likelihood of the filter response maps of the test images to belong to either of the template gaussian distributions. This posterior probability was then used as a feature.

3.2 Convex Hull Extent and Solidity

Convexity is an important shape feature and is defined as the ratio of perimeters of the convex hull over that of the original contour. The figure below shows an example of a convex hull compared to the original region.[4, 5]

Figure 3.2: An Example of Convex Hull obtained from a binary mask of a MSER Region

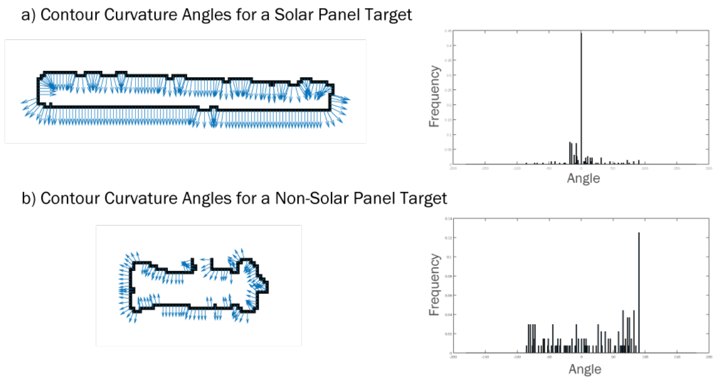


Solidity describes the extent to which the shape is convex or concave and is defined by $Solidity = A_s/H$ where $A_s$is the area of the shape region and H is the convex hull area of the shape. Thus the solidity of a convex shape is always 1. In the context of solar panels, the solidity would be expected to be close to 1 as most panels are square or rectangular, whereas other type of regions have irregular and non-convex shapes.[5]

3.3 Contour Curvature

Curvature is an important feature for distinction of shapes with several approaches that have proven to be very useful for object recognition.The approach used in this project was to look at the distribution of the angles normal to the contour of the MSER regions. As can be seen from the figure below, solar panel regions had largely right angle normal vectors whereas non-solar panel regions showcased a much more uniform distribution of angles for the normal vectors.[5]

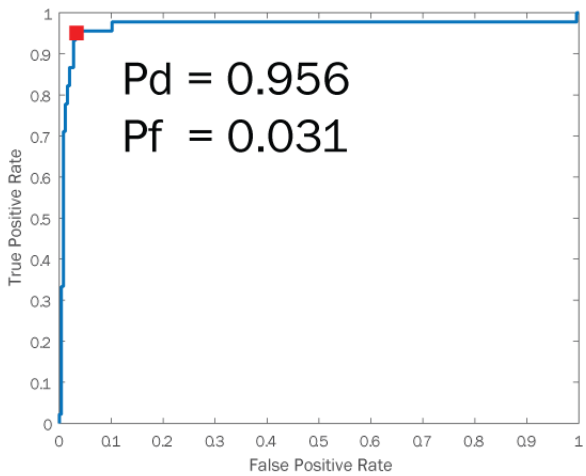
Figure 3.3: Examples of the Normal Angles of two regions



4 Diagnostics

With these features implemented along with a support vector machine classifier, a 100 k-fold cross-validation was run to obtain the receiver operating characteristic (ROC) curve, as pictured in Figure 4. The test data included 53 regions of solar panels and 265 regions without solar panels, and the algorithm was able to obtain a False Positive Rate of 3.1% and True Positive Rate of 95.6%.

Figure 4.1: ROC of a 100-fold cross validation SVM classifier run on the pre-screened regions



5 Conclusion and Future Work

These results provide a good starting point to running on larger unlabeled satellite images to get a database of Solar PV installations. Future work will revolve around improving performance by fine-tuning the prescreener, and utilizing the larger ground-truth data set to train the classifier. Additionally different filters might be explored to obtain better texture based features, and other shape based features such as chord distribution, shock graphs and boundary moments will be considered.

References

[1] Hallock, Lindsey, and Michelle Kinman. "California's Solar Success Story: How the Million Solar Roofs Initiative Transformed the State's Solar Energy Landscape." Frontier Group and Environment California Research & Policy Center, April 2015. http://www.environmentcalifornia.org/sites/environment/files/reports/CA_Solar_Success_scrn_FINAL_7-7-2015.pdf.

[2] Crosier, Michael, and Lewis D. Griffin. "Texture Classification with a Dictionary of Basic Image Features." In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 1–7. IEEE, 2008. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4587663.

[3] Varma, Manik, and Andrew Zisserman. "A Statistical Approach to Texture Classification from Single Images." International Journal of Computer Vision 62, no. 1–2 (2005): 61–81.

[4] Toshev, Alexander, Ben Taskar, and Kostas Daniilidis. "Shape-Based Object Detection via Boundary Structure Segmentation." International Journal of Computer Vision 99, no. 2 (2012): 123–46.

[5] Yin, Peng-yeng. Pattern Recognition Techniques, Technology and Applications. Rijek, Crotia: InTech, 2008. http://www.intechweb.org/books/show/title/pattern_recognition_techniques_technology_and_applications.