Detection of Solar Photovoltaic Panels using Satellite Imagery

Raghav Saboo

Duke UNIVERSITY

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.

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. 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.

Acknowledgements

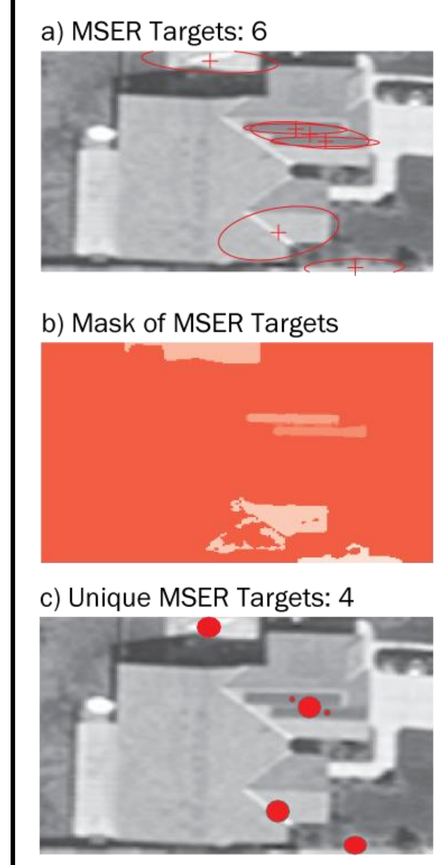
Dr. Kyle Bradbury & Dr. Jordan Malof.

Author Affiliations

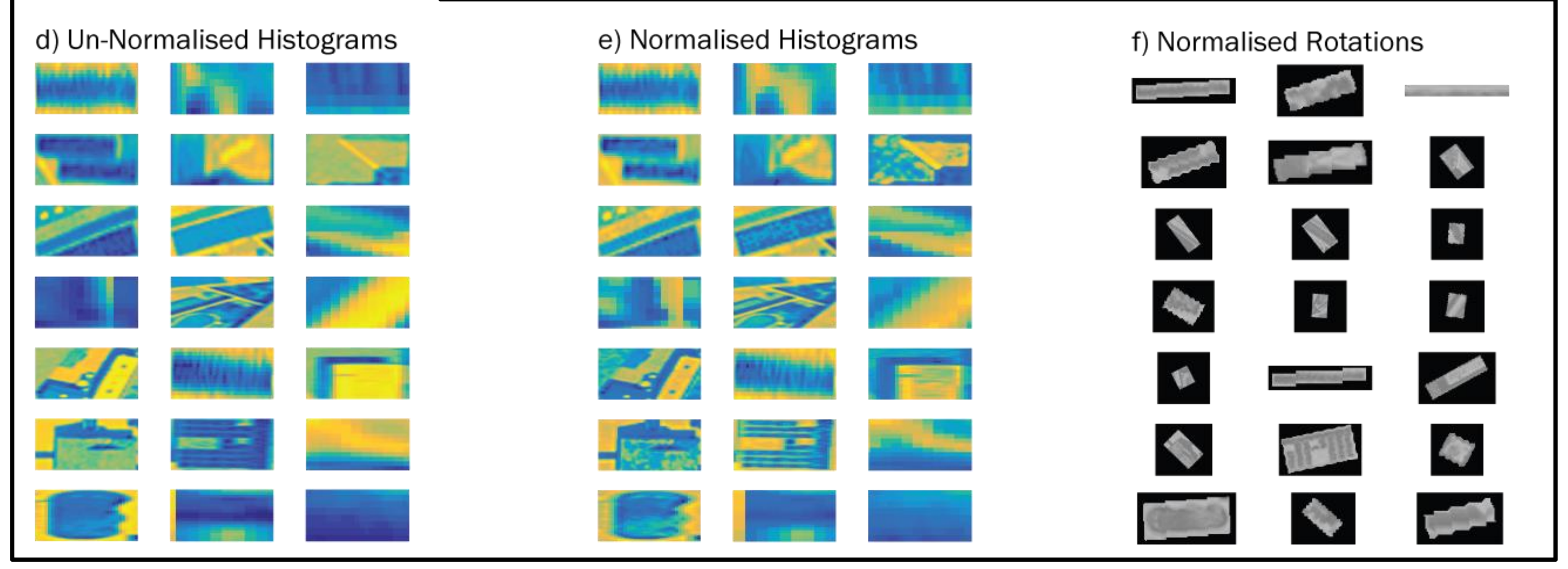
Department of Economics, Duke University, Durham, NC 27708
(Email: rs322@duke.edu)

Data Set & Preprocessing

FIGURE 1



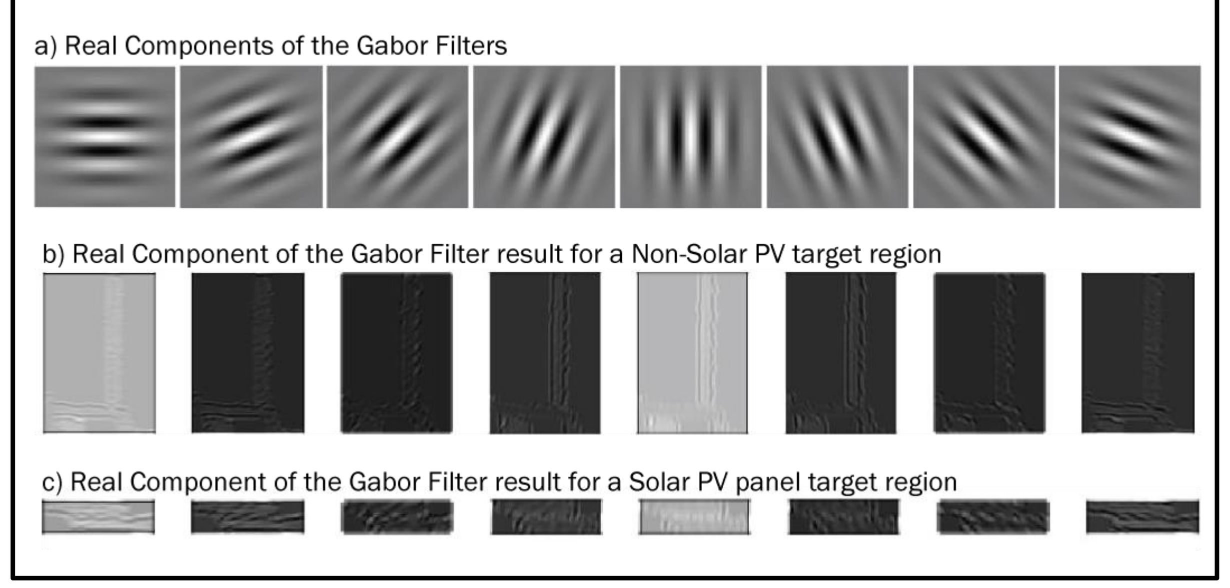
The data set consists of 100 satellite orthoimages of rooftops, with 50 of the rooftops having solar panels on them. An example is seen in Figure 1a). In order to determine the target regions that needed to be classified in each image a Maximally Stable Extremal Regions (MSER) algorithm was implemented on each image. This resulted in target regions as shown in both Figure 1a) and b), and in total 606 target regions were detected. However many of these regions were duplicates or overlaps, and in order to eliminate these duplicates a Mean Shift algorithm was implemented on the centroids of each of the target regions. This produced only unique MSER target regions as seen in Figure 1c) 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 as seen in Figures 1d) to f).



Feature Set

1. Gabor Filter K-Mediods Clustering

FIGURE 2



A sample of 10 Solar PV regions and 10 Non-Solar PV regions were taken and each region had a 8 directional Gabor Filter applied to its grayscale version. A map of magnitude and phase of the results was created for Solar PV regions and Non-Solar PV regions. Each map was further run through a k-mediods clustering algorithm to divide them up into 4 clusters each and then template histograms of the distance from the centroid clusters were created for Solar PV Regions and Non-Solar PV regions. The same process was then run on the remaining dataset and the chi-squared value of the histogram of each region was matched to the template histograms to generate features measuring match to Solar PV or Non-Solar PV.

2. Gabor Filter GMM Clusters

Similarly two Gaussian Mixture Models were trained based on the same clusters on the map for a Solar PV region and a Non-Solar PV region. Features as the posterior probability of the Gabor Filter map of the test regions were then computed for the GMM of a Solar PV region and Non-Solar PV region as features.

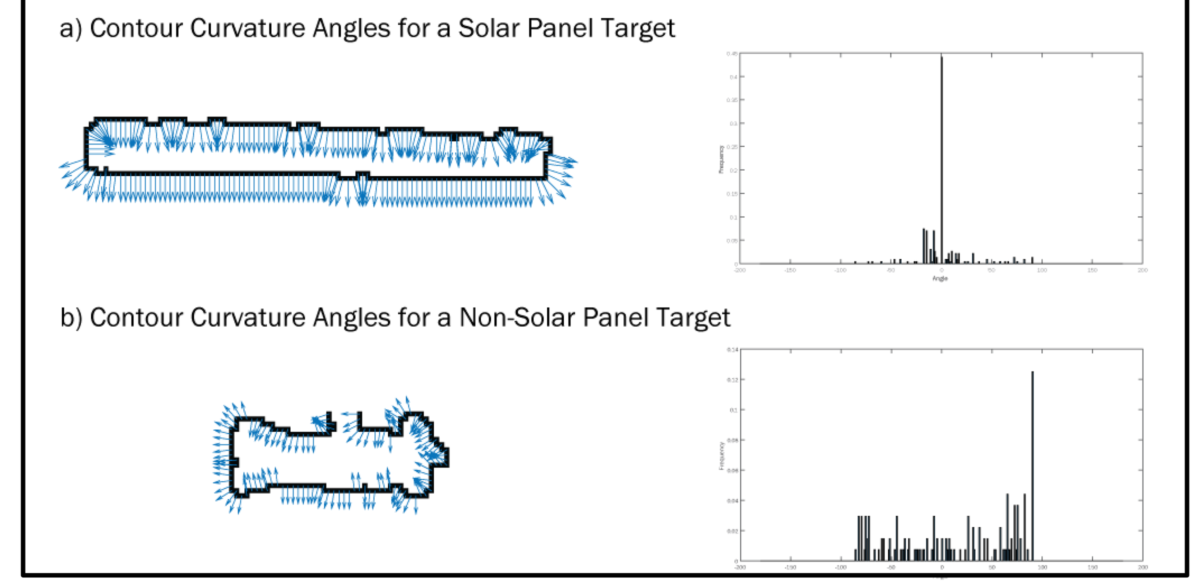
3. Solidity & Convex Hull Extent

Solidity here describes the extent to which the shape is convex or concave. The convex hull extent is a scalar that specifies the ratio of pixels in the region to pixels in the total bounding box.

3. Contour Curvature

Here the distribution of the normal angles to the contour of the regions was used as a feature. Most Solar PV regions had a impulse response type histogram of angles, whereas most Non-Solar PV regions had an approximately uniform distribution of angles.

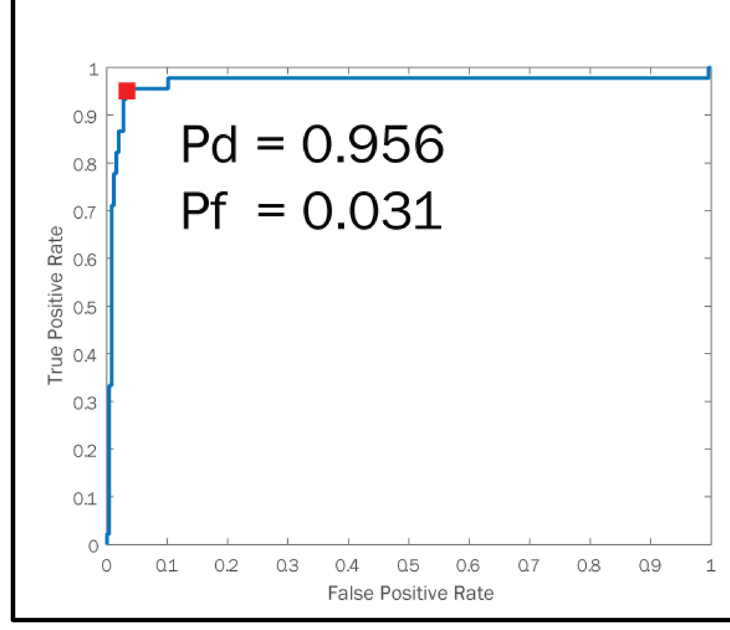
FIGURE 3



Classification & Results

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



Future Work & Conclusion

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.