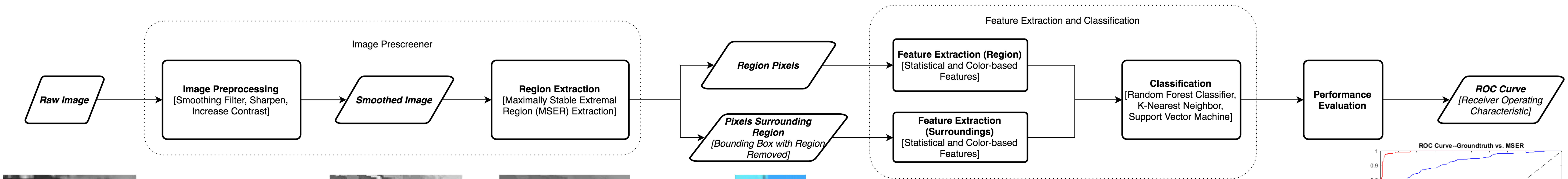
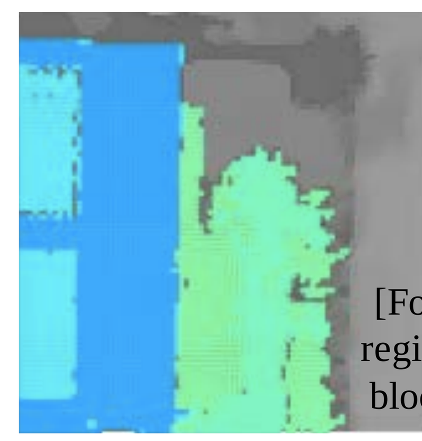
Automated Rooftop Solar PV Detection and Power Estimation through Remote Sensing

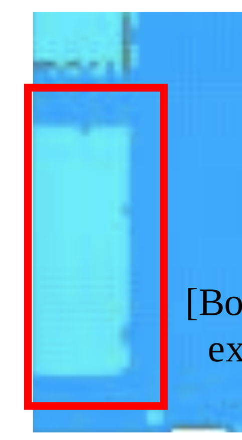
Arjun Devarajan (BS'18), Brody Kellish (BSE'17), Cassidee Kido (BSE'17), Aaron Newman (BSE'17), Nick von Turkovich (BSE'17), Wuming Zhang (BS'18)
Kyle Bradbury (Energy Initiative), Leslie Collins (Pratt), Timothy Johnson (Nicholas), Richard Newell (Nicholas, Energy Initiative)

Algorithm Outline

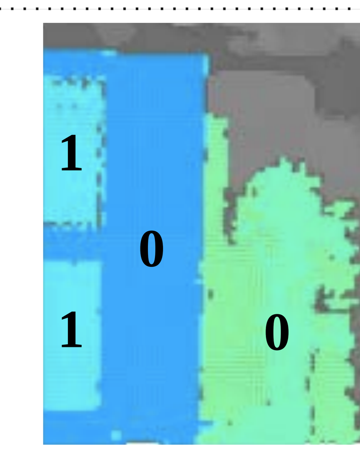




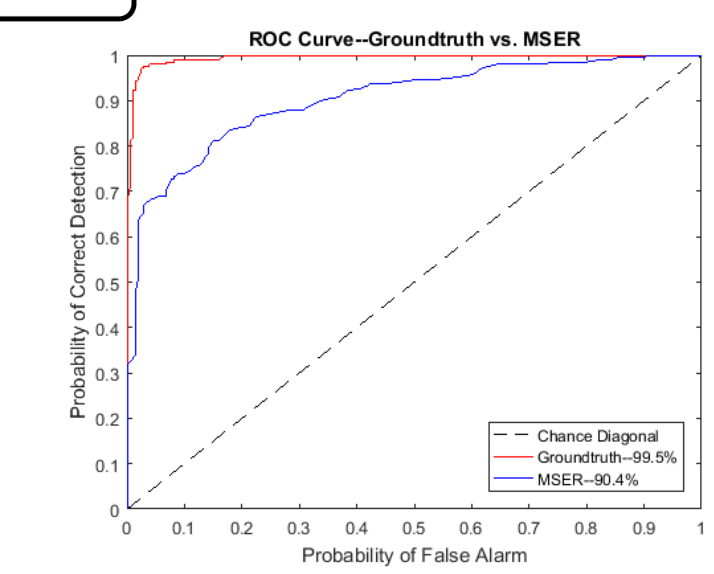
[Four extracted regions shown as blocks of color]



[Bounding box surrounding extracted positive region (on image border)]



[Labeled regions]



Abstract

As solar photovoltaics (PV) become a major sector of the energy market, there is a growing necessity for granular data regarding distributed rooftop solar PV. Solar power providers and customers, urban planners, grid system operators, and energy policymakers would vastly benefit from an imagery-based solar panel detection algorithm that can be used to form granular datasets of installations and their power capacities. A successful PV detection algorithm would also be a major contribution to the field of remote object detection. We have developed a PV identification algorithm with 90.4% accuracy when extracting regions using the maximally stable extremal regions (MSER) technique, and with 99.5% accuracy when using regions formed from the area surrounding solar arrays in our ground-truth training set. Because both methods share preprocessing, feature extraction, and classification techniques, but differ in region extraction, we have identified region extraction to be the key cause of the discrepancy, and thus, the main constraint in the overall PV identification process.

Introduction and Previous Work

Increased attention to the global warming crisis has led to the rapid adoption of distributed rooftop solar photovoltaics (PV) across the United States. Simultaneously, object detection in imagery is now a common research tool with varying applications ranging from military surveillance to facial recognition on social media. With the recent proliferation of solar panels across the U.S., remote object detection has become an increasingly attractive tool to help track solar installations nationwide. Similar studies have been conducted previously for imagery-based building [1] and solar panel [2] detection, both with promising results. We previously amassed a ground-truth dataset of over 19,000 solar arrays in Fresno, Modesto, Oxnard, and Stockton, California. We are using this dataset to train an algorithm that identifies solar PV in high resolution satellite orthoimagery. Once our algorithm is robust, we can apply it to various imagery datasets to create a database that tracks the location and power capacity of all solar arrays across the U.S. Such an algorithm progresses the field of remote object detection and is necessary for creating accurate PV datasets, which are vital to reliably integrating solar panels into existing grid systems, informing public policy, and promoting the propagation of future solar PV installations.

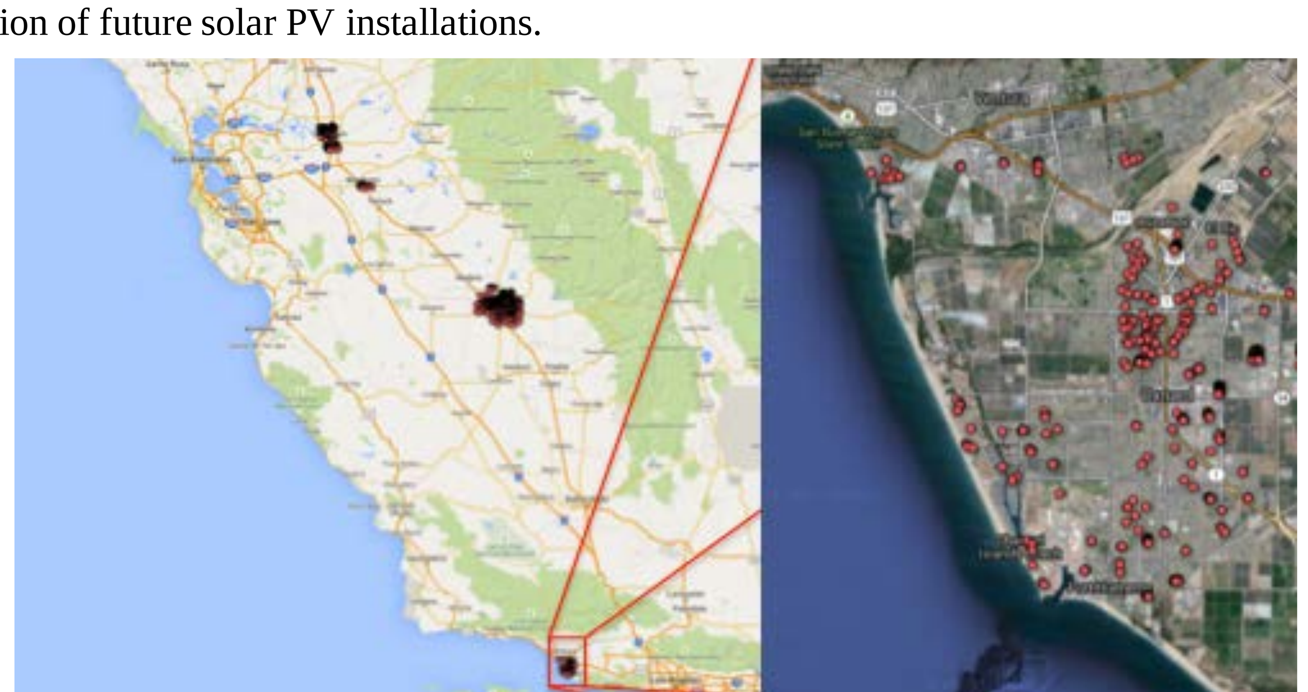


Figure 1: Ground-truthed solar panel training set (red circles indicate sample array locations)

Preprocessing through Region Extraction

Region extraction is performed to reduce the scope of the feature extraction and classification problem by identifying and isolating candidate regions. In order to determine an appropriate strategy for extracting solar panel regions from satellite orthoimagery, we first visually assessed the performance of maximally stable extremal region (MSER) extraction given a raw image by visualizing the extracted regions. We chose the MSER algorithm based on preliminary successes in previous solar panel detection studies [2] and because it is a fast method.

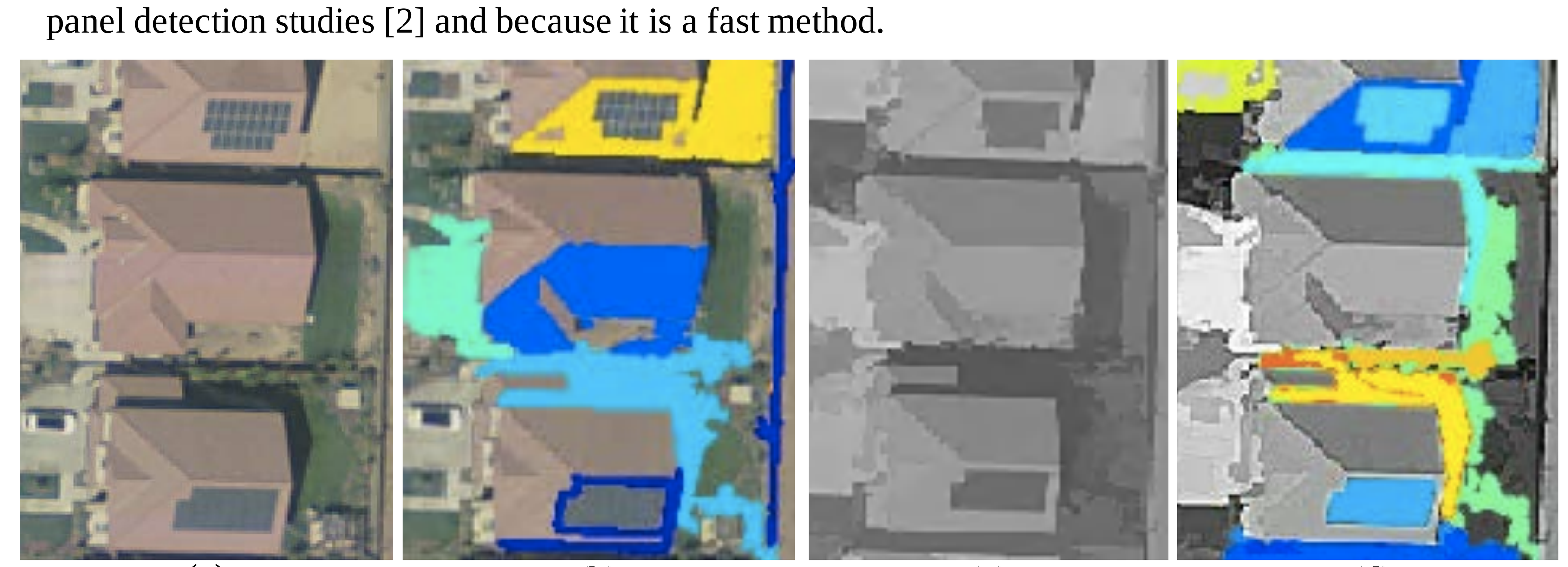


Figure 2: Significant improvements visualized with pre-processing
(a) Raw image with solar panel, (b) MSER applied to raw image with solar panel,
(c) Smoothing filter applied to raw image, (d) MSER applied to pre-processed image

Through visualization it was apparent that MSER alone was not effectively detecting solar panels in raw imagery (figure 2a), and as a result we concluded that it would be necessary to apply a smoothing filter to the image in order to reduce pixel intensity variation (figure 2b). Specifically, our aim in developing a pre-processing filter was to remove white gridlines that are characteristic of solar panels. This task was accomplished through opening and closing by reconstruction, a non-linear smoothing technique, which was shown in previous studies to be effective at removing thin lines from imagery. Applying MSER to the filtered and restructured image yielded improvements in comparison to MSER alone (figure 2c), but detection rates remained low regardless.

Figure 4: Smoothing filter has the unintentional side effect of reducing contrast and introducing blur in the image. To counteract these effects we sharpened the image and increased contrast.
(a) Original image, (b) Loss of clarity due to smoothing, (c) Restored image after smoothing

References
[1] S. Jabari and Y. Zhang, "Building Detection in Very High Resolution Satellite Image Using IHS Model", in ASPRS 2014 Annual Conference, Louisville, Kentucky, 2014.
[2] J. Malof, R. Hou and L. Collins, "Automatic Solar Photovoltaic Panel Detection in Satellite Imagery", in ICRERA, Palermo, Italy, 2015.
[3] K. G. Bryan Bollinger, "Peer Effects in the Diffusion of Solar Photovoltaic Panels," Mark. Sci. Provid. RI, vol. 31, no. 6, pp. 900–912, 2012.

Feature Extraction and Classification

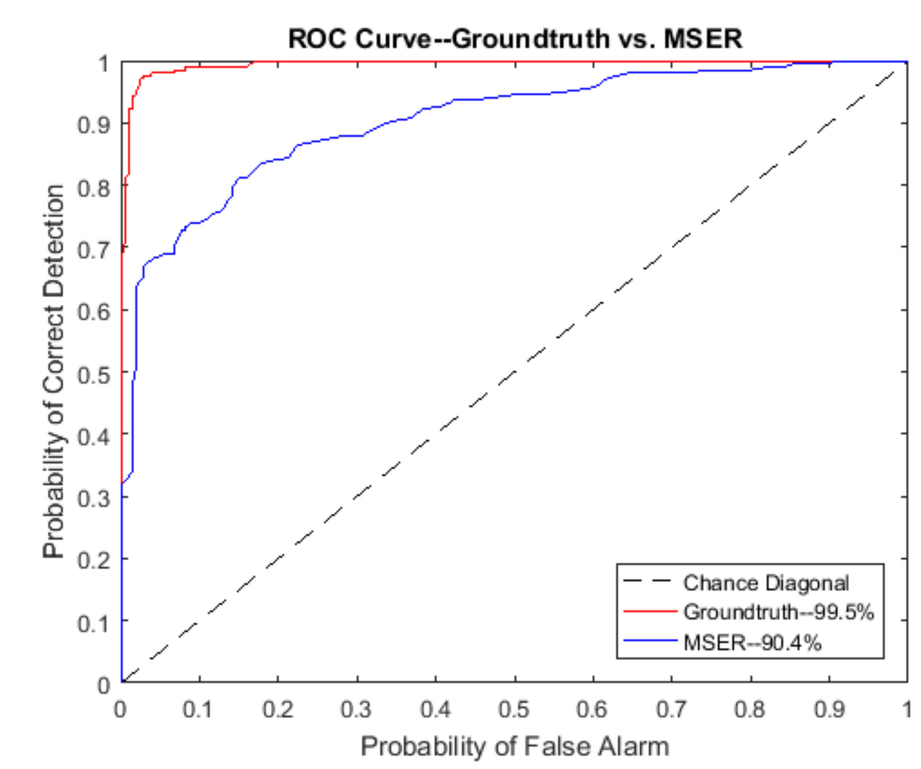


Figure 5: ROC curves corresponding to ground-truthed regions and regions extracted through MSER

To set up the feature extraction and classification experimental framework, we first selected 200 cropped images with solar arrays and 200 without solar arrays from the ground-truth dataset to serve as positive and negative examples, respectively. We also extracted 200 candidate regions using a baseline region extraction method (MSER), and let these serve as our comparison positive examples. We formed a bounding box about each of the regions in both sets and grew them by a uniform amount. By extracting several basic statistics from the regions and their backgrounds within the bounding boxes, we used a k-folds validation method (with 10 partitions) and tested multiple classifiers, of which random forest was the most successful. The discrepancy between our ground-truth and candidate region (MSER) classifications can be seen in figure 5, further emphasizing the vitality of ideal region extraction methods.

Conclusions and Future Research

In completing these experiments we aimed to reduce the global information gap regarding the distribution and capacity of installed rooftop solar PV by laying the foundation for an algorithm that detects and characterizes solar arrays in satellite orthoimagery. Using our novel ground-truth dataset, we demonstrated the vitality of effective prescreening for the problem of detecting residential rooftop solar arrays in imagery by comparing the performance of a random forest classifier trained and tested on ground-truth regions and on regions extracted through a baseline prescreener. Using ground-truth, ideal regions, we achieved highly accurate classification results, while we were much less successful in classifying regions obtained using our baseline prescreener. We also developed an algorithm that estimates the size and power capacity of a ground-truthed solar array within a reasonable error threshold. In future studies we plan to improve upon the baseline preprocessing algorithm with the goal of increasing the Jaccard index of the extracted regions by investigating other region extraction algorithms. Additionally, we would like to assess the application of our approach to other object identification problems, such as asset management, and apply our techniques to the challenge of estimating power capacity and energy generation of distributed PV resources across the United States.

Acknowledgements
We would like to give our sincere thanks to Dr. Jordan Malof, Dr. Guillermo Sapiro, Dr. Steve Sexton, Rui Hou, and Yijun Li, for their great support in satellite imagery detection. The poster could not have been completed without the effort and cooperation from our group members. We also sincerely thank our team leader Dr. Kyle Bradbury as well as our team advisors Dr. Timothy Johnson, Dr. Leslie Collins and Dr. Richard Newell, for their guidance and encouragement.