Applied Energy 183 (2016) 229–240

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Automatic detection of solar photovoltaic arrays in high resolution aerial imagery

Jordan M. Malof [a,*], Kyle Bradbury [b], Leslie M. Collins [a], Richard G. Newell [c]

[a] Department of Electrical & Computer Engineering, Duke University, Durham, NC 27708, United States
[b] Energy Initiative, Duke University, Durham, NC 27708, United States
[c] Nicholas School of the Environment, Duke University, Durham, NC 27708, United States

HIGHLIGHTS

- An approach for faster and cheaper PV information collection is proposed.
- An algorithm to detect photovoltaic arrays in aerial imagery is tested.
- Results demonstrate the efficacy of the PV information collection approach.
- The results are the first of their kind for solar photovoltaic array detection.
- The data is publicly available: https://dx.doi.org/10.6084/m9.figshare.3385780.v1.

ARTICLE INFO

Article history:
Received 28 June 2016
Received in revised form 29 August 2016
Accepted 30 August 2016

Keywords:
Solar energy
Detection
Object recognition
Satellite imagery
Photovoltaic
Energy information

ABSTRACT

The quantity of small scale solar photovoltaic (PV) arrays in the United States has grown rapidly in recent years. As a result, there is substantial interest in high quality information about the quantity, power capacity, and energy generated by such arrays, including at a high spatial resolution (e.g., cities, counties, or other small regions). Unfortunately, existing methods for obtaining this information, such as surveys and utility interconnection filings, are limited in their completeness and spatial resolution. This work presents a computer algorithm that automatically detects PV panels using very high resolution color satellite imagery. The approach potentially offers a fast, scalable method for obtaining accurate information on PV array location and size, and at much higher spatial resolutions than are currently available. The method is validated using a very large (135 km^2) collection of publicly available (Bradbury et al., 2016) aerial imagery, with over 2700 human annotated PV array locations. The results demonstrate the algorithm is highly effective on a per-pixel basis. It is likewise effective at object-level PV array detection, but with significant potential for improvement in estimating the precise shape/size of the PV arrays. These results are the first of their kind for the detection of solar PV in aerial imagery, demonstrating the feasibility of the approach and establishing a baseline performance for future investigations.

1. Introduction

The quantity of solar photovoltaic (PV) arrays has grown rapidly in the United States in recent years [2,3], with a large proportion of this growth due to small-scale, or distributed, PV arrays [4,5]. These small-scale installations are often found on the roofs of commercial structures, or private homes [4], and therefore are often referred to as rooftop PV.

Distributed PV offers many benefits [6], but integrating it into existing power grids is challenging. To understand and evaluate the factors driving distributed PV, and to aid in its integration, there is growing interest among government agencies, utilities, and third party decision makers in detailed information about distributed PV; including the locations, power capacity, and energy production of existing arrays. As a result, several organizations have begun collecting or publishing such information, including the Interstate Renewable Energy Council (IREC) [7], Greentech Media [8], and the US Energy Information Administration (EIA) [9,10].

Although the available information on distributed PV is expanding, it is nonetheless difficult to obtain. Existing methods of obtaining this information, such as surveys and utility interconnection filings, are costly and time consuming. They are also

* Corresponding author.
E-mail address: jmmalo03@gmail.com (J.M. Malof).

http://dx.doi.org/10.1016/j.apenergy.2016.08.191
0306-2619/

typically limited in spatial resolution to the state or national level [4,7]. For example, the EIA began reporting state-level distributed PV data at the end of 2015 [9].

1.1. The collection of solar PV information from high resolution aerial imagery

In this work we investigate a new method of collecting PV information that is potentially cheaper and faster than existing methods, and capable of obtaining much higher spatial resolution. The proposed approach uses computer algorithms to *automatically* extract useful PV information from high resolution ($\leqslant$0.3 m per pixel) color aerial imagery. This approach consists of two major steps.

The first step, and the biggest technical challenge, involves the use of computer algorithms to automatically annotate PV arrays in aerial imagery. Fig. 1 shows examples of the intended output of step one.

The second step involves using the output of step one, potentially with additional metadata, to estimate PV information over local geographic regions. For example, the total installed PV surface area (e.g., per square kilometer) can be estimated directly from the output of step one. If, for example, the algorithm detects 30 PV array pixels, and each pixel is 1 m^2, then we know there is approximately 30 m^2 of PV array surface area. More sophisticated information can also be extracted. For example, the PV surface area estimates discussed above could be augmented with additional metadata (e.g., local weather patterns, topography, and solar insolation) to yield estimates of PV capacity, or actual energy generation.

1.2. Applications and benefits of the PV information collection approach

The proposed two-step approach offers many benefits over existing methods of information collection. First, it permits the collection of PV information at very high geo-spatial resolution, which can facilitate energy infrastructure planning and decision making at state, regional, and even local levels. For example, high resolution spatial maps of PV information (e.g., installed PV capacity per square kilometer) can be integrated into geographical information systems (GIS), or Energy-GIS systems, which are increasingly used in energy infrastructure evaluation and planning [11–15].

Another benefit of our proposed approach is that it automates the process of data collection. It is relatively inexpensive to apply (i.e., run a computer program), and to reapply repeatedly as new imagery becomes available. It is consequently less invasive, faster, and cheaper than existing methods, such as compiling utility interconnection filings or conducting surveys. It also permits analyses and decision-making where more frequent estimates of the PV

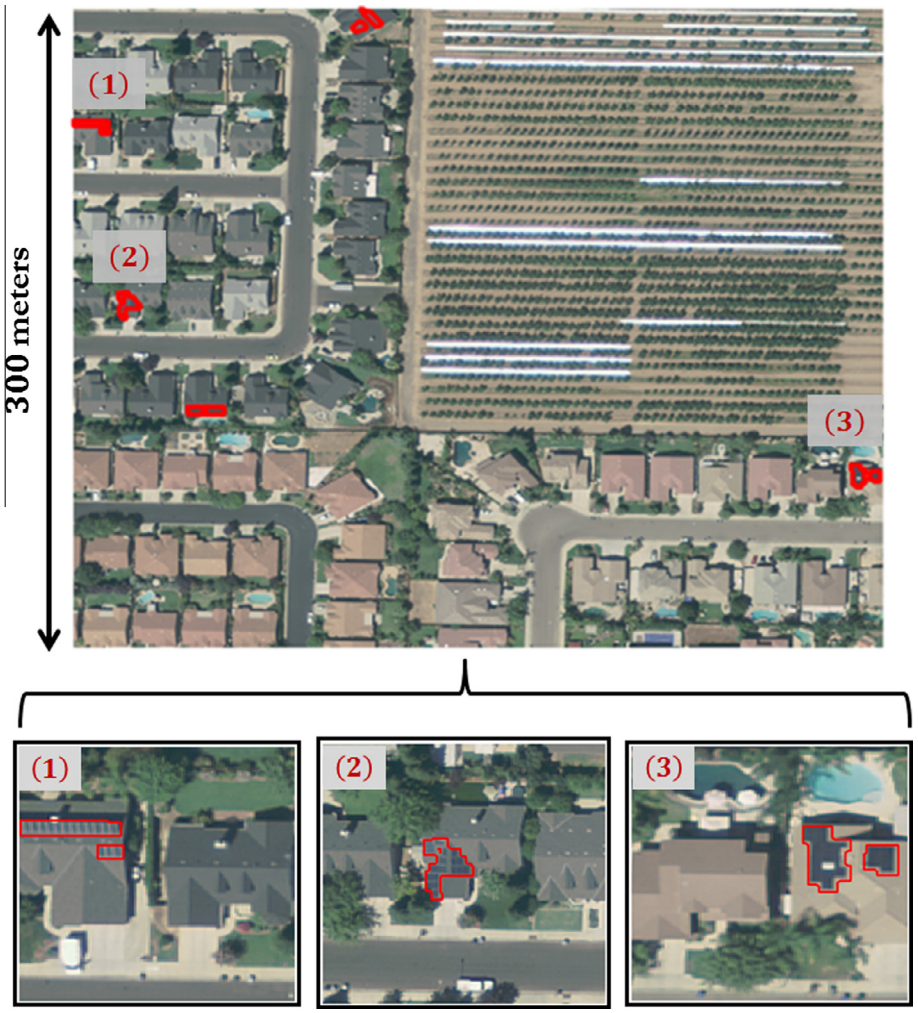

Fig. 1. An example of a color orthographic image (top) from the orthoimagery dataset, with human annotations shown in red outlines. Three of the annotations are enlarged in smaller images (bottom) so that the rooftop PV are more clearly visible. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

information may be needed. For example, if one's goal is to measure the growth of PV arrays in a city, then it is desirable to operate the algorithm frequently (e.g., every month). This would permit regular measurement of the number of new PV arrays that are installed. This information could be used by PV installers, utility companies, grid system operators, and government agencies to plan future infrastructure needs including transmission and distribution. It could also be useful for answering numerous economics and social science questions including evaluating the effectiveness of government incentives for solar PV development (e.g., in California [16]), as well as predicting where distributed solar growth is most likely to occur down to the neighborhood level. This technique could be readily adapted and applied to any location on Earth for which high resolution imagery is available, including countries without the infrastructure to measure PV deployment through other mechanisms.

1.3. Contributions of this work

In this work we present an algorithm for automatically identifying the location and size of solar PV arrays in aerial imagery and demonstrate its efficacy. We focus on this solar array detection problem because, as discussed in Section 1.1, solving the detection problem provides the location and size information of the solar arrays that can be translated into power capacity and energy generation estimates.

We evaluate the performance of the proposed detection algorithm using a large dataset of publicly available aerial imagery (135 km^2 area including more than 2700 PV arrays) [1]. In order to validate our results, we manually annotated all of the solar PV arrays in the aerial imagery used in this work, permitting us to compare the algorithm detections against the human annotations for performance evaluation purposes. Two qualities of the detection algorithm are evaluated: (i) its ability to detect individual PV pixels, and (ii) its ability to correctly detect PV objects (i.e., accurately identify the shape, size, and location of individual PV arrays). The results demonstrate the algorithm is highly effective on a per-pixel basis. It is likewise effective at object-level PV array detection, but with significant potential for improvement in estimating the precise shape/size of the PV arrays.

To our knowledge, this is the first time that an automated remote sensing approach has been applied to assessing installed distributed solar PV, particularly on a city-sized dataset and with a complete performance evaluation to demonstrate its efficacy. This establishes a baseline performance for future investigations and improvements. We also provide further details of our experiments (see appendix) to encourage further research.

If this technique is applied on a state or national scale, it has the potential to change the way we collect data on distributed solar PV information, and improve our understanding of this growing technology. This technique could also be readily adapted and applied to any location on Earth with high resolution imagery available.

The remainder of the paper is organized as follows. Section 2 provides a review of related work. Section 3 describes the aerial imagery data that is used for algorithm development. Section 4 presents the proposed solar PV detection algorithm. Section 5 presents the algorithm performance evaluation and Section 6 presents experimental results on the dataset. Section 7 presents our conclusions and suggestions for future work.

2. Related work

This section discusses some existing research that is relevant to our work. Section 2.1 discusses existing methods of collecting distributed PV information using remote sensing methods, and the associated energy-related applications of the collected information. Section 2.1 discusses the recognition of objects in aerial imagery, and previous work that has motivated the design of our algorithms and experiments,

2.1. Remote sensing for the collection of PV information and energy-related applications

Several remote sensing systems have been used to collect PV-related information. One popular approach has been the use of satellites to estimate ground irradiation, which is then used in models where solar PV potential is estimated [17–20]. The ArcGIS, SAGA GIS, and GRASS GIS systems are GIS-based models that estimate ground irradiation based on satellite measurements [17,18,20].

Perhaps the most well-studied remote sensing system for PV array information collection is the Light Detection and Ranging (LiDAR) system [21–25]. LiDAR is most often used to identify the locations of rooftops, as well as their angles (an important factor in PV energy production). This information is then combined with other metadata, such as solar irradiance or weather, for assessing the suitability of rooftops for the installation of small-scale solar PV systems. The LiDAR data is also often augmented with other remote sensing information such as color imagery [26] or the satellite irradiance estimates mentioned above.

The idea of automatically detecting PV arrays in aerial imagery was first investigated recently by our group [27]. In that work, we employed a less sophisticated prototype algorithm on a substantially smaller dataset (<1 km^2 area, with 53 annotated PV arrays) than the one considered here. The work presented here represents the first time that an automated remote sensing approach has been applied to assessing installed distributed solar PV on a large scale (city-sized) dataset, and with a complete performance evaluation to demonstrate its efficacy.

2.2. Recognition of objects in aerial imagery

The automatic detection of objects in aerial imagery (e.g., orthorectified imagery) has been researched extensively [28–31]. Some specific examples include the detection of roads [32–37], buildings [38–45], and vehicles [46–49]. In this published body of work, a large variety of algorithms have been proposed, employing techniques such as image processing, statistical modeling, machine learning classifiers, heuristic rules, and more.

The main component of the PV array detection algorithm proposed here is a supervised machine learning classifier called a Random Forest (RF) [50]. Supervised classifiers have previously been used for object recognition in remote sensing imagery, including the RF [51,52], support vector machine (SVM) [47] and various types of neural networks [32,49,53]. The RF in particular has been applied for land cover classification [51] and object detection [52]. In [52] it was used to classify individual pixels into one of four classes: building, street, trees, and grass. The RF takes a similar role in this work, where it is used to classify individual pixels as PV, or not PV.

An important resource in aerial imagery recognition is a labeled dataset. Such datasets consist of imagery where each instance of the target object is indicated by a bounding box, polygon, or similar annotation. Such datasets are used for (i) the development of effective detection algorithms, and (ii) an accurate assessment of their performance. Ideally a labeled dataset will cover a large surface area, and have a large number of labeled objects. Large datasets better represent practical operating conditions, which involve large volumes of data that is collected for diverse environments and imaging conditions. Datasets used in recently published research typically include a few hundred labels, and a few hundred

square meters of surface area [45,52,54,55]. Some specific recent examples include the SZTAKI-INRIA dataset for building detection, which includes 665 labeled buildings [54]. The OIRDS dataset for automotive detection consists of 1800 labels [55]. In this work we utilize a dataset of color aerial imagery that encompasses 135 km^2 of area with more than 2700 labels.

3. The aerial imagery dataset

All experiments and algorithm development in this work utilize a large dataset of color (RGB) aerial imagery collected over the US city of Fresno, California. The imagery covers a total spatial area of 135 km^2. All of the imagery was collected in the same month in 2013, using aerial photography. The imagery has a spatial resolution of 0.3 m per pixel, and all the imagery has been orthorectified. An example of the imagery is shown in Fig. 1, where the solar PV locations are annotated in red.

Further details about the data can be found at [1], where the data is also publicly available for download. The full imagery dataset is composed of 601 images that are each 5000 by 5000 pixels, across three cities, and with varying resolution. We chose to use imagery from Fresno, California because recently collected imagery was available (from 2013), with a high resolution (0.3 m), and because Fresno has a large number of solar PV installations. Over 100 images of Fresno are available in [1], from which we randomly sampled 60 of the available images for the analysis presented here. The identification tags of these images are provided in the appendix for future investigations.

3.1. Human annotations of true rooftop PV locations

In order to develop an effective computer vision algorithm, as well as accurately assess its performance, it is necessary to have the precise locations where PV installations appear in the aerial imagery. In order to obtain this information, human observers visually scanned the imagery and annotated all of the (visible) PV arrays. For improved quality, two annotators scanned each part of the imagery, and their annotations were combined by taking a union of each observer's annotations. There were a total of 2794 individual solar PV regions in the imagery after the merging process. Note again that this is a subset of the 19,863 annotations available in [1]. Some examples of annotated regions are shown at the bottom of Fig. 1 and in Fig. 4a.

To avoid a positive bias in the performance evaluation of the proposed detection algorithm, we split the available imagery into two disjoint datasets: Fresno Training and Fresno Testing. This is a common approach for validating supervised machine learning algorithms, such as the RF model used in our detection algorithm [56,57]. A summary of the imagery in each dataset is presented below in Table 1. The data was split between training and testing at a ratio of 2:1, in order to provide enough solar array examples to effectively train the RF model (see Section 4.3 for details about the RF).

4. The proposed PV detection algorithm

In this section we present the details of the proposed solar PV detection algorithm. We begin with a brief overview of the primary processing steps, followed by individual sections providing more details about each step.

Table 1
Summary of fresno color orthoimagery dataset.

Designation	Area of imagery (km^2)	Number of PV annotations
Fresno training	90	1780
Fresno testing	45	1014

4.1. Algorithm overview

The proposed rooftop PV algorithm takes RGB color aerial imagery as input and performs four major processing steps, as illustrated in Fig. 2.

(1) Feature extraction. This step consists of extracting image statistics, or features, around each pixel that characterize the colors, textures, and other patterns surrounding the pixel. The feature extraction step effectively maps the 3-channel RGB image into an M-channel image, where M is the number of features extracted around each pixel location.
(2) Random Forest Classifier. The image statistics computed in the feature extraction stage are the input to a trained RF classifier. The RF is a machine learning classification model that assigns a probability, or "confidence", to each pixel in the imagery. The confidence value indicates how likely the pixel is to correspond to a PV array. The output of this step is a single channel image, or spatial map, of where PV arrays are likely to be located. An example image and associated confidence map are shown in Fig. 4.
(3) Post-processing. This step is designed to improve the accuracy of the confidence map that was generated in the RF classification step. This process consists of identifying high confidence individual pixels (i.e., local maxima locations) and then growing regions of pixels around them. All pixel confidence values outside of these grown regions are then set to zero.
(4) Object detection. This step identifies groups of contiguous high confidence pixels that are likely to correspond to a single PV array. Each identified group of contiguous pixels is

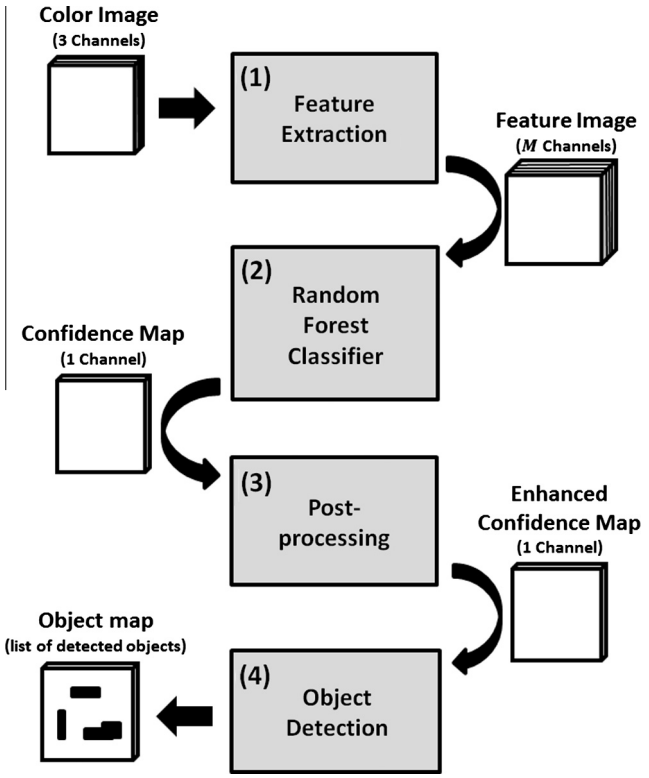

Fig. 2. A flowchart of the PV detection algorithm. Each gray block corresponds to a major processing step. Additionally, the input and output of each stage is also shown on the right or left of each block.

returned from this step as a detected object, and the confidence of that object is set to the value of the maximum pixel confidence value in that object. The output of this step is a list of objects and their confidence values, which is used to perform object-based scoring.

4.2. Feature extraction

The feature extraction process takes a 3-channel RGB aerial image as input, and returns an M-channel "feature image". In the feature image, each pixel is now represented by an *M*-dimensional vector of feature values that are computed using the original RGB image. The goal of this new representation is to encode information about each pixel which might be helpful for deciding whether that pixel corresponds to a PV array. This is accomplished by extracting useful information about the imagery surrounding each pixel, such as nearby colors and textures. This information will make it possible for the RF classifier (the next processing step) to better decide whether a given pixel is a PV array, based solely on its M-dimensional feature vector.

There is a large body of published research regarding good image features, and many effective approaches have been developed [58,59]. The choice of a particular feature representation depends on the goals of the application. An important consideration in this work (aside from classification accuracy) is computational efficiency, so that the proposed algorithm can be applied at a national scale. It should therefore be possible to extract features quickly, with few computations. One class of image features that achieves this objective is the local color statistic (LCS) feature [60,61]. In this feature representation, the feature vector for a given pixel consists of the means, μ, and variances, σ^2, of pixel intensities in several nearby windows. For a given window, these statistics are computed separately in each of the three color channels of the image (so six total feature values for each window). Here we adapt this approach for the purpose of solar PV detection.

In our adaptation of this approach we used a window size of 3×3 pixels, because it roughly corresponds to the size of individual PV panels in the aerial imagery. A larger window size would yield statistics that mix background pixels and PV pixels, making the statistics less useful for identifying PV arrays. A smaller window would be too small to compute any statistics.

In our approach we also organize the windows into concentric rings, denoted by S_r. The subscript r refers to the (approximate) radius of the ring around its center point, denoted p_0, which is the pixel we wish to represent. Fig. 3 illustrates how features are extracted for a ring with $r = 4$ [62].

Our final feature representation will combine statistics from several rings, with varying radii. This is accomplished by concatenating the statistics from each of the rings together into one large feature vector. Although using many rings allows us to encode more image information, and potentially better classification accuracy, it can also cause problems. Therefore, the choice of features must be made carefully.

Our choice of features is based on two well-known feature selection criteria from the machine learning research: each feature should (i) encode information that is useful for the problem, and (ii) not be redundant (i.e., encode the same information as other features) [56,57,63]. Based on these two criteria we chose to use three rings: S_0, S_2, and S_4, due to the following rationale. Since all 9 windows in S_0 are located at the same position, and so we only used one of them. We only used even-numbered values of r to avoid redundancy in the features, due to overlap between many windows. We only included windows where $r \leqslant 4$ because more distant windows (from p_0) would contain little relevant information for classification.

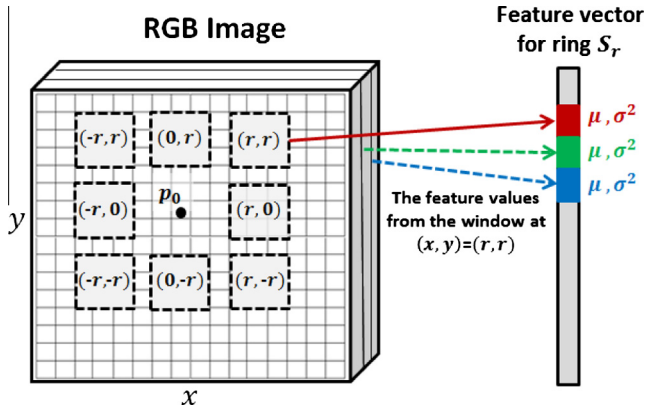

Fig. 3. Illustration of the features extracted for a single ring of windows, S_r. The subscript r refers to the (approximate) radius of the windows from the center position, denoted p_0, which is the pixel we wish to represent with features. Each window is labeled with its (x,y) coordinates, with respect to p_0. In this particular example $r = 4$, although the coordinates of each window are given in terms of the parameter r. Each window contributes six statistics to the feature vector: a mean, μ, and a variance, σ^2 from each color channel. This is illustrated for one window, located at (r, r).

4.3. The random forest classifier

Random Forests (RFs) [50] are a state-of-the-art supervised statistical classification method. They have been successfully applied to a variety of problems, such as image processing [64], medical diagnosis [65], pose recognition [66,67], and remote sensing [51,52]. In this work we use the RF to classify each pixel in the imagery, similar to the way it was used in some other contexts [52,67]. The RF receives the feature vector for each pixel and assigns it a probability, or confidence, indicating the likelihood that it corresponds to a PV array.

Although the RF has many advantages, there are two primary reasons the RF was employed in this work. First is the ability of the RF to learn complex nonlinear relationships between input and output variables. This is important because the relationship between the image features and the pixel labels (i.e., PV or non-PV) are very complex. The second motivation for the RF is its known computational speed during training and testing [51,52,67]. It can also be implemented on graphics processing units for further speed improvements [68]. This computational efficiency is important for handling the massive amounts of data common in high resolution aerial imagery applications. For example, our datasets consists of 1.5 billion pixels, however this encompasses only 135 km^2 of the United States' nearly 9.857×10^6 km^2 area.

The RF actually consists of an ensemble of T simpler supervised classifiers called decision trees [69]. An illustration of an RF is provided in Fig. 5. Each tree consists of a series of decision nodes which terminate (at the bottom of the tree) in a leaf node. To classify a new feature vector, x, it is presented to the top decision node, and it is subsequently directed down the tree, based on the values in the feature vector, until it reaches a leaf node. At the leaf node a probability is assigned to the vector indicating to which class (e.g., PV or non-PV) it belongs.

During training, each tree is "grown" independently of the other trees, in a top-down manner, using a random bootstrap sample of pixels from the training data. The decision nodes are learned such that they best separate the training data according to some performance measurement (e.g., the Gini impurity index or information gain). In this work we use the Gini index. Each node of each tree considers only a random subset of the input features (of size m) when inferring how to split the data. The parameter m is often

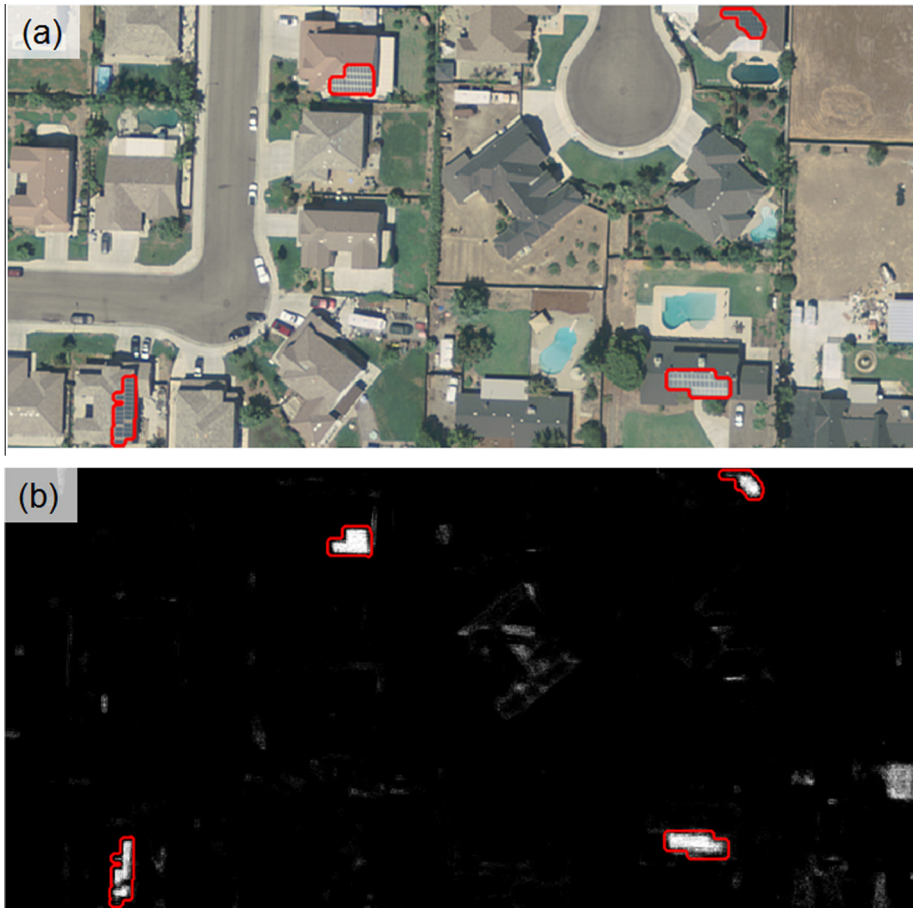

Fig. 4. An example of an aerial image (top) and its corresponding confidence map (bottom). In both images the true solar PV locations have been annotated in red. The confidence map is the output of stage two in the detection algorithm (see Fig. 2). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

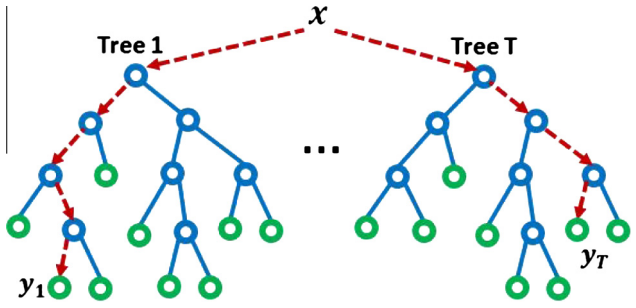

Fig. 5. Illustration of an RF classifier architecture. The RF consists of several decision trees, where each tree consists of decision nodes (blue circles) and leaf nodes (green nodes). To classify an input feature vector, x, it is presented to each tree independently. For a given tree, x, is passed down the decision nodes based on its values until it reaches a leaf node. At the leaf node it receives a probability indicating to which class it belongs. Here y_t indicates the probability of belonging to the PV class. The RF output is creating by averaging the probabilities returned from each tree. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

cited as the only major adjustable parameter of the RF, and a conventional setting that usually works well is $m = \sqrt{M}$, where M is the number of feature dimensions [51].

Decision nodes are created until (i) splitting no longer improves the Gini index or (ii) it would result in fewer than 5 observations in a leaf node. The parameter settings for the RF, and other algorithms in this work, are presented in Table 3.

4.4. Post-processing

The goal of the post-processing (PP) step is to improve the pixel-wise classification accuracy of the raw confidence maps, as well as to make them better suited for the object detection step. The algorithm for PP is outlined in Table 2, and an example of the input and output of this process is shown in Fig. 7.

Broadly speaking, the PP algorithm identifies individual high confidence pixels (local maxima) and then grows a new smooth region (in terms of confidence values) around them. Local maxima are retained for region growing only if they (i) exceed some minimum threshold, c_0, and (ii) they are the largest maxima in a surrounding square window of length L_s. This filtering criterion is designed to remove maxima locations that are likely to be false alarms. Regions are grown around each remaining maxima location using Otsu's method [70], which automatically segments pixels into foreground (high confidence) and background (low confidence) regions. The region growing takes place in a square window of length L_g that is centered on each maxima location.

The output of this PP operation, I', consists of many small connected regions that all have the same confidence value. This makes object extraction easier. The final step of post-processing is to apply morphological closing and dilation [58], respectively, in order to smooth the grown regions. These operations are performed with a disk with radius r_1 and r_2, respectively. The parameter values of the PP algorithm are provided in Table 3. These values were set by optimizing the performance of the algorithm on the Fresno Training data (see Section 5.3).

Table 2

The Post-processing Algorithm
Data: $I \equiv$ Confidence Map, **Parameters**: L_g, L_s, c_0, r_1, r_2
Result: $I' \equiv$ Enhanced Confidence Map
Initialize: $I' \leftarrow$ image of zeros, with size equal to I
1. Apply non-maximum suppression to I, using filters that are $L_s \times L_s$ in size
2. Remove all maxima below global threshold, c_0
3. **For** each remaining maxima location, i, **do**
(a) Crop an L_g x L_g region around i
(b) Use Otsu's method [70] to find foreground pixels in the cropped region
(c) Find all connected components
(d) Retain only the connected component, C, that contains the maxima location, i
(e) Place C into image I', with all of the pixel values in C set to $I(i)$
end
4. Apply a morphological closing and dilation [58] to I', with a disk of radius r_1 and r_2, respectively

Table 3
Detection algorithm parameters.

Symbol	Processing step	Quantity	Value
T	RF	Number of RF trees	30
M	RF	Number of features in RF	102
m	RF	Number of variables to sample at each tree node in RF	$\sqrt{M} \cong 10$
L_s	PP	Length of the of square non-maximum suppression region	9
c_0	PP	Global confidence threshold after non-max suppression	0.375
L_g	PP	Length of the square region around each maxima location used for Otsu's method	19
r_1, r_2	PP	Disk radius used for morphological closing, and dilation, respectively	5,2

4.5. Object detection

The object detection phase identifies groups of adjacent, or neighboring, high confidence pixels and identifies them as detected objects. This is achieved by first thresholding the confidence map: any pixel confidence greater than zero is set to one, and all others are set to 0. The result is a binary image, which is used for finding contiguous groups of pixels. The resulting connected regions are all taken as detected objects. The confidence of each region is given by the maximum confidence pixel in that region. An example output from this process is shown in Fig. 7.

5. Experimental design

This section begins with an overview of the experimental design used for the experiments in Section 6, followed by a description of our performance evaluation methodology. Note that Section 2.2 provides a review of relevant image recognition literature, which provides motivation for our experimental design.

We conducted two experiments, with the primary goal of measuring the performance of the proposed PV array detection algorithm. The first experiment measures how well the algorithm identifies individual PV pixels: pixel-based classification performance. The second experiment measures how well the algorithm can identify objects (groups of pixels) that correspond to PV array annotations, as well as their precise shape and size. The experiments are conducted on two datasets of aerial imagery denoted as Fresno Training, and Fresno Testing (see Table 1). As described in Section 3, all PV arrays visible in the imagery were annotated by humans to provide ground truth pixels/objects for use in scoring the detector.

In order to estimate the performance of the algorithm we split the data into two (disjoint) datasets: the Fresno Training data, and the Fresno Testing data. The primary role of the Fresno Training dataset was to train the RF classifier, as well as optimize other parameters associated with the detection algorithm. The Fresno Testing dataset was used for validation to obtain an unbiased performance estimate for the detector. This approach is similar to the popular cross-validation scheme [56,57]. If the dataset is very large, as is the case here, cross-validation can be impractically slow, and the approach we chose can be used as an alternative to cross-validation [56]. The performance metric used to evaluate the performance of the algorithm is the precision recall (PR) curve. The PR curve is a popular performance metric for object detection in aerial imagery [32,38,71,72], and therefore it is adopted here.

Note that in practice, the detection algorithm requires human annotations only for training. Once the algorithm is trained, it can be applied to new imagery without any annotations. The additional testing data annotations are only needed for the purpose of estimating the performance of the algorithm in practical operating conditions (i.e., when it is applied to new data).

The next few sections describe PR curves, the Jaccard index (used to measure the degree of overlap between detected objects with human annotations), and how the algorithm parameters were optimized.

5.1. Performance metrics

PR curves measure the performance tradeoff between making correct detections and false detections, as the sensitivity of a detector, or classifier, is varied. An illustration of a PR curve is shown in Fig. 6. The x-axis of a PR curve is the recall, R, which is the proportion of all true target objects (e.g., PV arrays) in the data that were returned by the algorithm as detections. The y-axis is the precision, P, which is the proportion of all detected objects (i.e., both true and false) which are true targets. An effective detector will tend toward the top right corner of the PR curve, thereby maximizing both recall and precision. A detector that detects objects randomly (i.e., it is ineffective) will achieve a precision that is equal to the proportion of objects in the dataset that are targets. For example, in the pixel-based PV detection experiments, roughly 0.07% of the Fresno Testing pixels correspond to PV arrays. Therefore a random detector would achieve $P = 0.0007$, for all values of R.

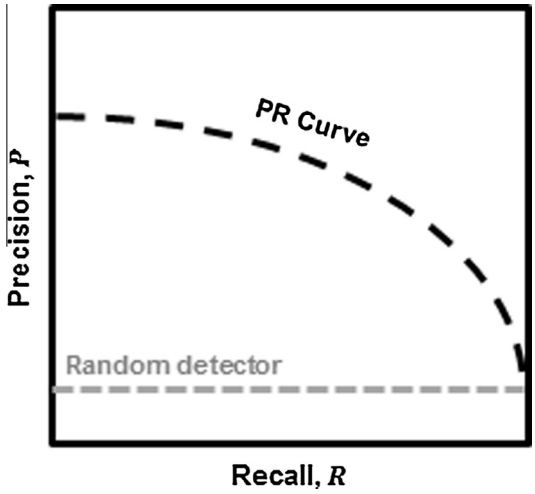

Fig. 6. Illustration of a PR curve. A good detector will obtain a curve that is closer to the top right corner of the graph. Random guessing results in a line that achieves constant precision for all values of R, where the precision is equal to the proportion of total detections returned by the algorithm that are from the target class (e.g., PV arrays).

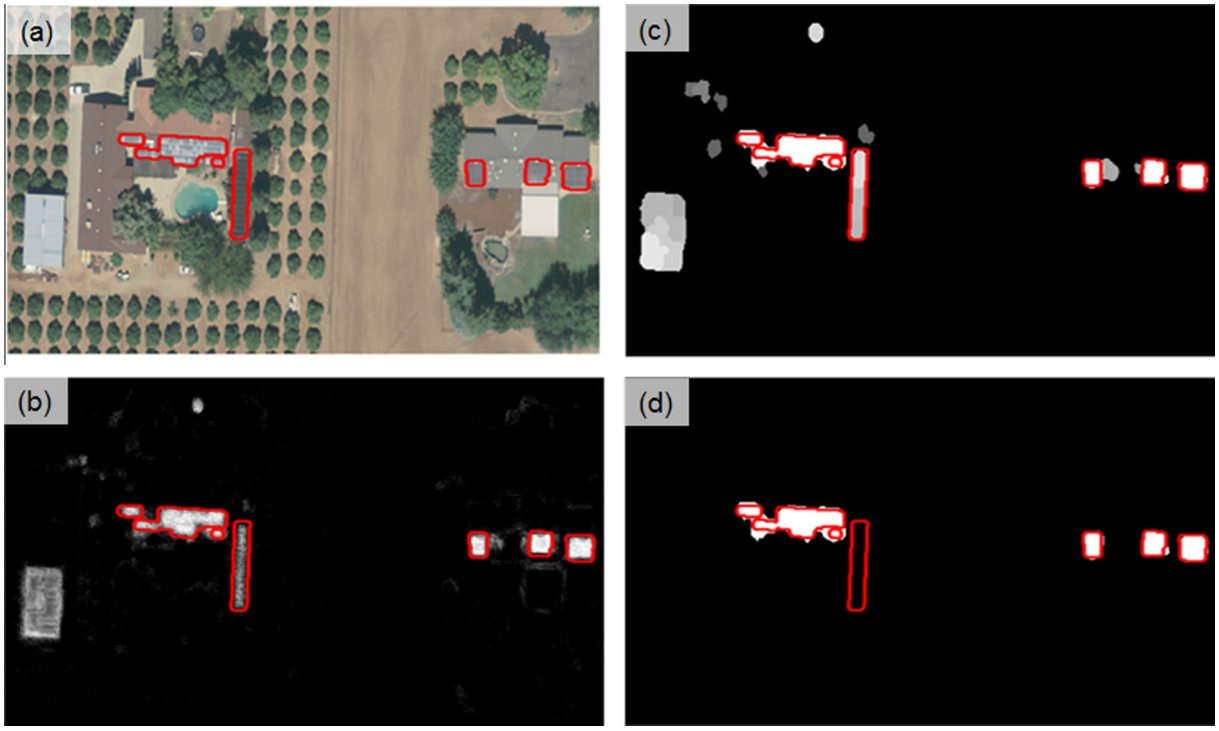

Fig. 7. Example output of the rooftop PV detection algorithm after several of the major processing steps. Four different images are shown, (a–d), and each image shows the human PV annotations in red. (a) is the original RGB image. (b) is the confidence map output from the Random Forest classifier, without post-processing; brighter pixels indicate higher confidence. (c) is the confidence map after post-processing. (d) shows the objects detected after the object detection stage of processing. Given the detection rate and false alarm rate employed in this example, the detector correctly removes all of the false alarms, while losing one of the true panel regions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The sensitivity for a given detection algorithm can be varied by raising or lowering a threshold, t_0, that is applied to the confidence values of the list of potential detections (e.g., pixels or objects). All potential detections above t_0 are accepted as detections, and all potential detections below t_0 are rejected. P and R are then computed based on the group of accepted detections.

5.2. Linking detections to human annotations

One issue that arises with object-based scoring is determining when a detected object should be considered a correct detection. A detected object (i.e., a region labeled as a PV array) may overlap with a PV annotation from the ground truth data, but how much overlap should be required to declare it as detected correctly? This problem is apparent in Fig. 7d, where none of the detected objects match *perfectly* with the human annotations, but they might reasonably be considered correct detections. To address this issue, a metric is needed to measure the shape/size similarity between two objects (i.e., groups of pixels).

One metric that has been utilized for this purpose is the Jaccard index [73]. The Jaccard index, J, for two image objects (groups of pixels), denoted A and B respectively, is given by

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}. \quad (1)$$

Fig. 8 below shows the Jaccard index for two objects as their level of overlap varies. The Jaccard index allows us to measure precisely how similar a detected object is to a human annotation. A threshold can then be applied where only detected objects above the threshold (with respect to a human annotation) can be considered a correct detection. Ultimately, the choice for this threshold should depend on the final application of the detector, and the corresponding level of shape/size accuracy that is needed. Therefore,

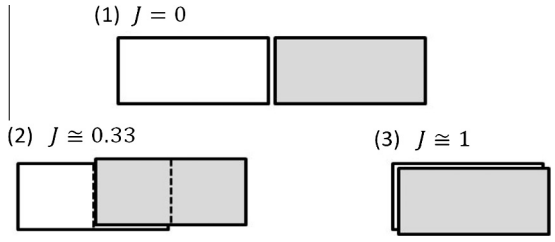

Fig. 8. Illustration of the Jaccard Index, J. The gray and white boxes represent two sets of pixels, such as a detected object from the solar PV detector, and a true solar PV annotation. As the degree of overlap of the two sets increases, J increases from 0 toward 1.

we report object-based performance for multiple thresholds. A similar approach was recently adopted in [74] for building detection.

In some instances a detected object will overlap with multiple ground truth annotations (see the left-most three annotations in Fig. 7d for an example). In this case the multiple annotations are treated as one annotation composed of the union of the individual annotations. If the union of the annotations has a sufficiently high Jaccard index with respect to a detection, all three annotations are considered to be detected by the detector.

5.3. Algorithm training and optimization

All training and parameter optimization was performed on the Fresno Training dataset. The final set of chosen parameters for the algorithm is shown in Table 3. These parameters are used in all experiments.

The RF classifier itself was trained using five million pixels from the Fresno Training dataset. This subset of pixels was chosen by

first selecting all of the available solar PV pixels (roughly 500,000), and then randomly sampling the remaining non-PV pixels from the training imagery. Using increasing numbers of pixels improves performance but at the cost of increasing the computation time of the RF. Five million was found to be a good tradeoff between performance and computation time on the training data.

The parameters were chosen in order to optimize performance on the training data. This parameter optimization was done by measuring the performance of the algorithm (on the Fresno Training dataset) as the parameters were varied over a coarse grid of potential values. Note that the parameter m was set to the conventional value of $\sqrt{M}$, rather than being optimized.

6. Experimental results

This section describes the experimental results obtained using the experimental design discussed in Section 5. First, pixel-based performance results are presented, followed by object-based performance.

6.1. Pixel-based performance

The pixel-based performance for the PV detection algorithm, on both the training and testing data, is shown in Fig. 9. Results are shown for the RF, and the RF after PP has been applied (RFPP). The primary goal of this experiment was to demonstrate that the RFPP algorithm can effectively detect PV array pixels. The results on the Fresno Testing dataset provide an unbiased estimate of the performance of the RF and RFPP algorithms. The results indicate that the solar PV detector is very effective at discriminating non-panel pixels from panel pixels. This is made most clear by considering how well a random detector (i.e., a completely ineffective detector) would perform. Recall from Section 5.1 that, because PV arrays constitute only 0.07% of the pixels in Fresno Testing, the random detector achieves $P = 0.0007$ for all values of R. Both the RF and RFPP detectors achieve performance far above this.

Further insight can be obtained from the results in Fig. 9 by comparing the performance of the detectors on the training data and testing data, respectively. As is expected, the results indicate that there is an overall performance drop between the training data and the testing data. Quantitatively this means that, for each value of R, the algorithm typically obtains a lower P on the testing data than it does on the training data. One exception to this occurs for the RFPP algorithm when R is below 0.6, however the testing and training performance is similar at these operating sensitivities.

The results also suggest that the main contributor to the performance loss incurred on the testing data is the RF classifier (as opposed to RFPP). This is because the RF algorithm performance drops between the training and testing dataset, however, the RFPP algorithm offers the same advantages on both the training and testing dataset; relative to the performance of the RF alone. This suggests that the RF is overfitting to the training data, or in other words, the RF learned to recognize patterns that are too unique to the training data, and as a consequence it less effectively identifies previously unseen PV arrays in the testing data. This can be addressed in many ways, and is an important consideration for future work.

6.2. Object-based performance

The primary goal of this experiment is to demonstrate the effectiveness of the detector for individual PV arrays. Further, we want to examine how well the detector can identify the precise shape/-size of individual PV arrays. As a result, we measure the object-based performance of the detector on the Fresno Testing dataset for varying settings of the Jaccard index during scoring. These resulting PR curves are shown in Fig. 10.

The results indicate that the object-based performance of the detector is once again well above that of the baseline random detector performance. Although this is true for all values of J, the performance of the detector decreases rapidly as J increases. As a specific example, when $J = 0.1$ the detector achieves $R = 0.7$ with $P = 0.6$, while at $J = 0.5$, $R = 0.55$ at the same value of P. When $J = 0.7$, the detector never reaches $P = 0.6$. This outcome is expected because, as J is increased, many of the objects detected that are near true PV locations are no longer considered correct detections. This also results in more PV annotations remaining undetected, even when the detector is operated with high sensitivity. This is why the maximum R obtained for each detector decreases as J increases.

Different values for J are likely to be appropriate depending on the intended purpose of the detector. For example, lower J values (e.g., $J = 0.1$) are appropriate for applications where only the gen-

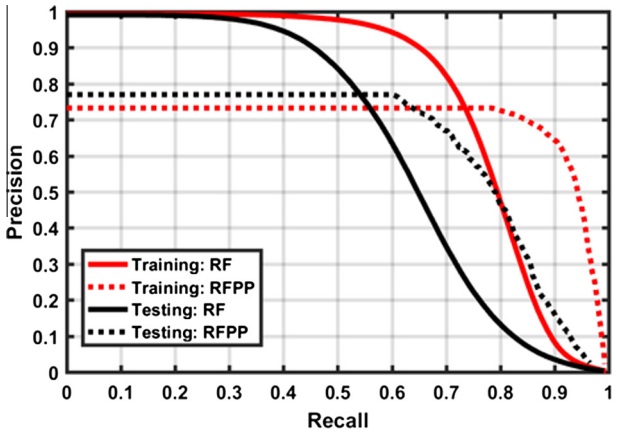

Fig. 9. PR curves for the pixel-wise performance of the PV detector on the Fresno Training dataset (red), and the Fresno Testing dataset (black). For each dataset, the performance of the detector is shown before post-processing (solid lines), and after post-processing (dashed lines). The random detector for this problem achieves $P = 0.0007$ for all values of R, but this is not shown due to its small magnitude. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

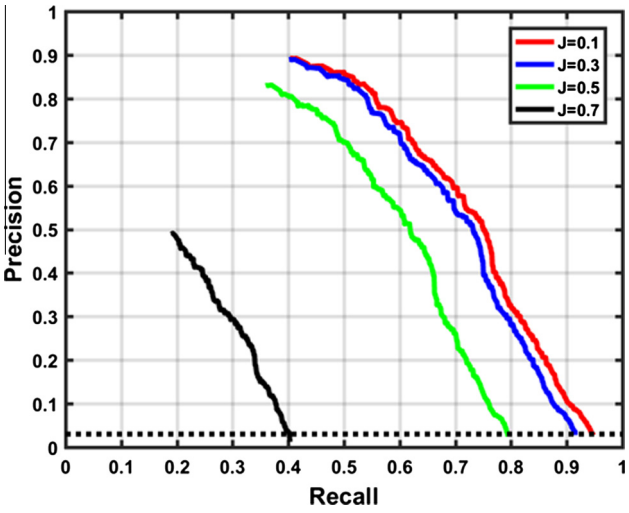

Fig. 10. PR curves for the object-based performance of the rooftop PV detector on the Fresno Testing dataset. Each PR curve corresponds to a different setting of the Jaccard index, J, during scoring. The left-most point of the curves represents the performance when classifying every object with confidence of one (i.e., the maximum RF output) to a detection. With object-based scoring, the detectors are not guaranteed to place objects over all true PV array locations, and indeed, none of the detectors reach $R = 1$.

eral location of target objects is important, and obtaining the precise shape/size is not. In the context of solar PV array detection, this may be the case if the detector is used as a preprocessing step for further, and more sophisticated (but slower), detection algorithms. Note that when operated with $J = 0.1$ the detector is capable of detecting roughly 90% of the targets, with $P \cong 0.1$. Since there are roughly 1000 PV arrays in the testing data, this corresponds to roughly 10,000 total detections returned by the detector (900 true detections and 9100 false detections) over the 45 km^2 testing area. This dramatically reduces the amount of image locations that must be considered for further processing, facilitating the use of more sophisticated subsequent processing. The detector proposed here is designed to operate quickly on large datasets, and therefore could be used in this role.

In contrast to lower J values, a higher value (e.g. $J = 0.7$) is appropriate for detection applications where it is important to accurately estimate the size and shape of target objects. In the context of solar PV array detection, this may be the case, for example, if the goal is to estimate the power capacity of individual solar PV arrays. Setting J to higher values will lead to a performance measure that better reflects the capability of a given detector to achieve that goal, which is a much more difficult task than simply detecting the likely presence of an object (using, e.g., $J = 0.1$). This difficulty is reflected in the much poorer performance of the proposed detector on this task (e.g., see Fig. 10 with $J = 0.7$). Looking forward, the performance reported here for $J = 0.7$ establishes a baseline for future improvement in achieving this type of goal.

7. Conclusions and future work

We investigated a new approach for the problem of collecting information for small-scale solar PV arrays over large areas. The proposed approach employs a computer algorithm that automatically detects solar PV arrays in high resolution ($\leqslant$0.3 m) color (RGB) imagery data. A detection algorithm was developed and validated on a very large collection of aerial imagery ($\geqslant$135 km^2) collected over the city of Fresno, CA. Human annotators manually scanned and annotated solar PV locations to provide ground truth for evaluating the performance of the proposed algorithm. Performance was measured in a pixel-based and object-based manner, respectively, using PR curves. In the case of object-based scoring, the algorithm was also scored based on how well it can identify the shape and size of the true panel object.

The results demonstrate that the algorithm is highly effective on a per-pixel basis. The PR measures indicate it can detect most of the true PV pixels while removing the vast majority of the non-PV pixels. The object-based PR curves indicated that the algorithm was likewise effective at object detection, however, it was far less effective at estimating the precise shape/size of the PV arrays.

The results presented here are the first of their kind for distributed PV detection in aerial imagery, and demonstrate the feasibility of collecting distributed PV information over large areas using aerial or satellite imagery. This may ultimately yield a faster, cheaper, and more scalable solution for the large scale collection of distributed PV information, and potentially information for other aspects of energy production and consumption as well. While the results here demonstrate the promise of this approach to information collection, several challenges remain as opportunities for future work.

7.1. Future work

7.1.1. Improved detection algorithms

Because the results here are the first of their kind for this problem they establish a baseline performance, or benchmark, for future algorithm development. To facilitate such efforts, the data used in this work is freely available for download [1], and the exact images used in our experiments are listed in the supplemental materials. It is our hope that others will build upon these results, and develop increasingly effective detection algorithms.

7.1.2. Inferring capacity and energy production

Another important line of future work is the inference of PV array capacity, energy production, and other characteristics from the imagery. Recall from the Section 1, that this is the second major step for creating a complete system for extracting PV information from aerial imagery. This second challenge could be pursued using the imagery detected from the PV detector, or otherwise using the ground truth annotations in the aerial imagery dataset.

7.1.3. Establishing practical performance needs

While the results here demonstrate the ability of an algorithm to discriminate between PV and non-PV imagery (as compared to a random detector), it is unclear what levels of performance would be needed for different practical applications. Creating a complete system for inferring distributed PV information would help reveal what level of performance is needed from the detection stage in order to obtain practically useful energy information; with qualities and reliability that is similar, or better than, current estimation strategies (e.g., the EIA [9]).

7.1.4. Information collection for other energy resources

Finally, we hope this work also motivates the collection of other energy information from aerial imagery, in addition to distributed PV. Other examples might include inferring the energy consumption of individual households, and (from that information) counties, or cities. This could be pursued, for example, by estimating the volume of a household from aerial imagery. Other elements of the energy system could also potentially be detected, such as power lines, or power plants.

Acknowledgements

This work was supported in part by the Alfred P. Sloan Foundation and the Wells Fargo Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Alfred P. Sloan Foundation or the Wells Fargo Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.apenergy.2016.08.191.

References

[1] Bradbury K, Saboo R, Malof J, Johnson T, Devarajan A, Zhang W, et al. Distributed solar photovoltaic array location and extent data set for remote sensing object identification. Figshare; 2016. http://dx.doi.org/10.6084/m9.figshare.3385780.v1 [accessed June 1, 2016].
[2] Alam MJE, Muttaqi KM, Sutanto D. An approach for online assessment of rooftop solar PV impacts on low-voltage distribution networks. IEEE Trans Sustain Energy 2014;5:663–72. http://dx.doi.org/10.1109/TSTE.2013.2280635.
[3] Chersin A, Ongsakul W, Mitra J, Member S. Improving of uncertain power generation of rooftop solar PV using battery storage. In: Int Conf Util Exhib Green Energy Sustain Dev. IEEE; 2014. p. 1–4.
[4] Electric power monthly. US Energy Information Administration; 2016.
[5] Net generation from renewable sources: total (all sectors), 2006-February 2016; 2016.
[6] Singh GK. Solar power generation by PV (photovoltaic) technology: a review. Energy 2013;53:1–13. http://dx.doi.org/10.1016/j.energy.2013.02.057.
[7] Sherwood L. U.S. solar market trends 2013; 2013.
[8] Association SEI. U.S. solar market prepares for biggest quarter in history; 2015. http://www.seia.org/news/us-solar-market-prepares-biggest-quarter-history.

[9] EIA electricity data now include estimated small-scale solar PV capacity and generation. EIA (US Energy Inf Adm; 2015. https://www.eia.gov/todayinenergy/detail.cfm?id=23972.

[10] Electric power monthly: with data for January 2015. US Energy Information Administration; 2015.

[11] Hiremath RB, Shikha S, Ravindranath NH. Decentralized energy planning; modeling and application-a review. Renew Sustain Energy Rev 2007;11:729–52. http://dx.doi.org/10.1016/j.rser.2005.07.005.

[12] Bazmi AA, Zahedi G. Sustainable energy systems: role of optimization modeling techniques in power generation and supply - a review. Renew Sustain Energy Rev 2011;15:3480–500. http://dx.doi.org/10.1016/j.rser.2011.05.003.

[13] Sánchez-Lozano JM, Teruel-Solano J, Soto-Elvira PL, Socorro García-Cascales M. Geographical information systems (GIS) and multi-criteria decision making (MCDM) methods for the evaluation of solar farms locations: case study in south-eastern Spain. Renew Sustain Energy Rev 2013;24:544–56. http://dx.doi.org/10.1016/j.rser.2013.03.019.

[14] Mirakyan A, De Guio R. Integrated energy planning in cities and territories: a review of methods and tools. Renew Sustain Energy Rev 2013;22:289–97. http://dx.doi.org/10.1016/j.rser.2013.01.033.

[15] Huang Z, Yu H, Peng Z, Zhao M. Methods and tools for community energy planning: a review. Renew Sustain Energy Rev 2015;42:1335–48. http://dx.doi.org/10.1016/j.rser.2014.11.042.

[16] California solar initiative n.d. http://www.gosolarcalifornia.ca.gov/csi/index.php.

[17] Hofierka J, Suri M, Šúri M. The solar radiation model for Open source GIS: implementation and applications. Proc Open Source GIS-GRASS Users Conf 2002:11–3.

[18] Šúri M, Hofierka J. A new GIS-based solar radiation model and its application to photovoltaic assessments. Trans GIS 2004;8:175–90. http://dx.doi.org/10.1111/j.1467-9671.2004.00174.x.

[19] Hammer A, Heinemann D, Hoyer C, Kuhlemann R, Lorenz E, Müller R, et al. Solar energy assessment using remote sensing technologies. Remote Sens Environ 2003;86:423–32. http://dx.doi.org/10.1016/S0034-4257(03)00083-X.

[20] Neteler M, Bowman MH, Landa M, Metz M. GRASS GIS: a multi-purpose open source GIS. Environ Model Softw 2012;31:124–30. http://dx.doi.org/10.1016/j.envsoft.2011.11.014.

[21] Lukač N, Žlaus D, Seme S, Žalik B, Štumberger G. Rating of roofs' surfaces regarding their solar potential and suitability for PV systems, based on LiDAR data. Appl Energy 2013;102:803–12. http://dx.doi.org/10.1016/j.apenergy.2012.08.042.

[22] Gooding J, Crook R, Tomlin AS. Modelling of roof geometries from low-resolution LiDAR data for city-scale solar energy applications using a neighbouring buildings method. Appl Energy 2015;148:93–104. http://dx.doi.org/10.1016/j.apenergy.2015.03.013.

[23] Lukač N, Seme S, Dežan K, Žalik B, Štumberger G. Economic and environmental assessment of rooftops regarding suitability for photovoltaic systems installation based on remote sensing data. Energy 2016;107:854–65. http://dx.doi.org/10.1016/j.energy.2016.04.089.

[24] Martin AM, Dominguez J, Amador J. Applying LIDAR datasets and GIS based model to evaluate solar potential over roofs: a review. AIMS Energy 2015;3:326–43. http://dx.doi.org/10.3934/energy.2015.3.32.

[25] Lukac N, Seme S, Laus D, Tumberger G, Alik B. Buildings roofs photovoltaic potential assessment based on LiDAR (light detection and ranging) data. Energy 2014;66:598–609. http://dx.doi.org/10.1016/j.energy.2013.12.066.

[26] Szabó S, Enyedi P, Horváth M, Kovács Z, Burai P, Csoknyai T, et al. Automated registration of potential locations for solar energy production with light detection and ranging (LiDAR) and small format photogrammetry. J Clean Prod 2016;112:3820–9. http://dx.doi.org/10.1016/j.jclepro.2015.07.117.

[27] Malof JM, Hou R, Collins LM, Bradbury K, Newell R. Automatic solar photovoltaic panel detection in satellite imagery. In: Int Conf Renew Energy Res Appl. IEEE; 2015. p. 1428–31.

[28] Mayer H. Object extraction in photogrammetric computer vision. ISPRS J Photogramm Remote Sens 2008;63:213–22. http://dx.doi.org/10.1016/j.isprsjprs.2007.08.008.

[29] Blaschke T. Object based image analysis for remote sensing. ISPRS J Photogramm Remote Sens 2010;65:2–16. http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004.

[30] Toshev A, Taskar B, Daniilidis K. Shape-based object detection via boundary structure segmentation. Int J Comput Vis 2012;99:123–46.

[31] Baltsavias EP. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. ISPRS J Photogramm Remote Sens 2004;58:129–51. http://dx.doi.org/10.1016/j.isprsjprs.2003.09.002.

[32] Mnih V, Hinton GE. Learning to detect roads in high-resolution aerial images. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2010:210–23. http://dx.doi.org/10.1007/978-3-642-15567-3_16. 6316 LNCS.

[33] Mokhtarzade M, Zoej MV. Road detection from high-resolution satellite images using artificial neural networks. Int J Appl Earth Obs Geoinf 2007;9:32–40.

[34] Bhattacharya U, Parui SK. An improved backpropagation neural network for detection of road-like features in satellite imagery. Int J Remote Sens 1997;18:3379–94. http://dx.doi.org/10.1080/014311697216937.

[35] Laptev I, Mayer H, Lindeberg T, Eckstein W, Steger C, Baumgartner A. Automatic extraction of roads from aerial images based on scale space and snakes. Mach Vis Appl 2000;12:23–31. http://dx.doi.org/10.1007/s001380050121.

[36] Bajcsy R, Tavakoli M. Computer recognition of roads from satellite pictures. Syst Man Cybern IEEE Trans 1976:623–37.

[37] Hu J, Razdan A, Femiani JC, Cui M, Wonka P. Road network extraction and intersection detection from aerial images by tracking road footprints. IEEE Trans Geosci Remote Sens 2007;45:4144–57. http://dx.doi.org/10.1109/TGRS.2007.906107.

[38] Chaudhuri D, Kushwaha NKK, Samal A, Agarwal RCC. Automatic building detection from high-resolution satellite images based on morphology and internal gray variance. Sel Top Appl Earth Obs Remote Sensing, IEEE J 2015:1–13. http://dx.doi.org/10.1109/JSTARS.2015.242565.

[39] Youssef MMS, Mallet C, Chehata N, Le Bris A, Gressin A. Combining top-down and bottom-up approaches for building detection in a single very high resolution satellite image. In: IEEE Int Geosci Remote Sens Symp (IGARSS). p. 4820–3. http://dx.doi.org/10.1109/IGARSS.2014.6947573.

[40] Jabari S, Zhang Y, Suliman A. Stereo-based building detection in very high resolution satellite imagery using IHS color system. In: IEEE Int Geosci Remote Sens Symp (IGARSS). p. 2301–4. http://dx.doi.org/10.1109/IGARSS.2014.6946930.

[41] Ghaffarian S, Ghaffarian S. Automatic building detection based on Purposive FastICA (PFICA) algorithm using monocular high resolution Google Earth images. ISPRS J Photogramm Remote Sens 2014;97:152–9. http://dx.doi.org/10.1016/j.isprsjprs.2014.08.017.

[42] Wang M, Yuan S, Pan J. Building detection in high resolution satellite urban image using segmentation, corner detection combined with adaptive windowed hough transform. In: IEEE Int Geosci Remote Sens Symp (IGARSS). p. 508–11. http://dx.doi.org/10.1109/IGARSS.2013.6721204.

[43] Izadi M, Saeedi P. Automatic building detection in aerial images using a hierarchical feature based image segmentation. In: 20th Int Conf Pattern Recognit. p. 472–5. http://dx.doi.org/10.1109/ICPR.2010.123.

[44] Nosrati MS, Saeedi P. A novel approach for polygonal rooftop detection in satellite/aerial imageries. In: 16th IEEE Int Conf Image Process. p. 1709–12. http://dx.doi.org/10.1109/ICIP.2009.5413641.

[45] Khoshelham K, Nardinocchi C, Frontoni E, Mancini A, Zingaretti P. Performance evaluation of automated approaches to building detection in multi-source aerial data. ISPRS J Photogramm Remote Sens 2010;65:123–33. http://dx.doi.org/10.1016/j.isprsjprs.2009.09.005.

[46] Holt AC, Seto EYW, Rivard T, Gong P. Object-based detection and classification of vehicles from high resolution aerial photography. Photogramm Eng Remote Sens 2009:871–80.

[47] Madhogaria S, Baggenstoss P, Schikora M, Koch W, Cremers D. Car detection by fusion of HOG and causal MRF. IEEE Trans Aerosp Electron Syst 2015;51:575–90. http://dx.doi.org/10.1109/TAES.2014.120141.

[48] Kembhavi A, Harwood D, Davis LS. Vehicle detection using partial least squares. IEEE Trans Pattern Anal Mach Intell 2011;33:1250–65. http://dx.doi.org/10.1109/TPAMI.2010.182.

[49] Chen X, Xiang S, Liu C-L, Pan C-H. Vehicle detection in satellite images by parallel deep convolutional neural networks. In: 2nd IAPR Asian Conf Pattern Recognit. p. 181–5. http://dx.doi.org/10.1109/ACPR.2013.33.

[50] Breiman L. Random forests. Mach Learn 2001;45:5–32. http://dx.doi.org/10.1023/A:1010933404324.

[51] Gislason PO, Benediktsson JA, Sveinsson JR. Random forests for land cover classification. Pattern Recognit Lett 2006;27:294–300. http://dx.doi.org/10.1016/j.patrec.2005.08.011.

[52] Tokarczyk P, Montoya J, Schindler K. An evaluation of feature learning methods for high resolution image classification. ISPRS Ann Photogramm Remote Sens Spat Inf Sci 2012;I-3:389–94. http://dx.doi.org/10.5194/isprsannals-I-3-389-201.

[53] Mokhtarzade M, Zoej MJV. Road detection from high-resolution satellite images using artificial neural networks. Int J Appl Earth Obs Geoinf 2007;9:32–40. http://dx.doi.org/10.1016/j.jag.2006.05.001.

[54] Benedek C, Descombes X, Zerubia J. Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. Pattern Anal Mach Intell IEEE Trans 2012;34:33–50.

[55] Tanner F, Colder B, Pullen C, Heagy D, Eppolito M, Carlan V, et al. Overhead imagery research data set - an annotated data library & tools to aid in the development of computer vision algorithms. Proc - Appl Imag Pattern Recognit Work 2009:3–10. http://dx.doi.org/10.1109/AIPR.2009.5466304.

[56] Bishop CM. Pattern recognition and machine learning, vol. 1. New York: Springer; 2006.

[57] Theodoridis S, Koutroumbas K. Pattern Recogn 1999.

[58] Forsyth D, Ponce J. Computer vision: a modern approach. 2nd ed. Pearson; 2012.

[59] Szeliski R. Computer vision: algorithms and applications. Springer; 2010.

[60] Clinchant S, Csurka G, Perronnin F, Renders J, Europe C, Maupertuis D. XRCE ' s participation to ImagEval n.d.

[61] Sánchez J, Perronnin F, Mensink T, Verbeek J. Image classification with the fisher vector: theory and practice. Int J Comput Vis 2013;105:222–45. http://dx.doi.org/10.1007/s11263-013-0636-x.

[62] Vijayaraj V, Bright Ea, Bhaduri BL. High resolution urban feature extraction for global population mapping using high performance computing. IEEE 2008;1:278–81.

[63] Duda RO, Hart PE, Stork DG. Pattern classification. John Wiley & Sons; 2001.

[64] Lepetit V, Fua P. Keypoint recognition using randomized trees. IEEE Trans Pattern Anal Mach Intell 2006. http://dx.doi.org/10.1017/CBO9781107415324.004.

[65] Lempitsky V, Verhoek M, Noble JA, Blake A. Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography. Lect Notes Comput Sci 2009;5528:447–56. http://dx.doi.org/10.1007/978-3-642-01932-6_48 [including Subser Lect Notes Artif Intell Lect Notes Bioinformatics].
[66] Fanelli G, Gall J, Van Gool L. Real time head pose estimation with random regression forests. IEEE Conf Comput Vis Pattern Recognit 2011:617–24. http://dx.doi.org/10.1109/CVPR.2011.5995458.
[67] Shotton J, Fitzgibbon aW, Cook M, Sharp T, Finocchio M, Moore R, et al. Real-time human pose recognition in parts from single depth images. CVPR 2011:1297–304. http://dx.doi.org/10.1007/978-3-642-28661-2-5.
[68] Sharp T. Implementing decision trees and forests on a GPU. Lect Notes Comput Sci 2008;5305 LNCS:595–608. http://dx.doi.org/10.1007/978-3-540-88693-8-4 [including Subser Lect Notes Artif Intell Lect Notes Bioinformatics].
[69] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC Press; 1984.
[70] Otsu N, Smith PL, Reid DB, Environment C, Palo L, Alto P, et al. A threshold selection method from gray-level histograms 1979;20:62–6.
[71] Cheriyadat aM. Unsupervised feature learning for aerial scene classification. IEEE Trans Geosci Remote Sens 2014;52:439–51. http://dx.doi.org/10.1109/TGRS.2013.224144.
[72] Manno-Kovacs A, Ok AO. Building detection from monocular VHR images by integrated urban area knowledge. IEEE Geosci Remote Sens Lett 2015;12:2140–4. http://dx.doi.org/10.1109/LGRS.2015.2452962.
[73] Levandowsky M, Winter D. Distance between sets. Nature 1971;234:34–5. http://dx.doi.org/10.1038/234034a0.
[74] Ok AO, Senaras C, Yuksel B. Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery. IEEE Trans Geosci Remote Sens 2013;51:1701–17. http://dx.doi.org/10.1109/TGRS.2012.2207123.