

DeepRoof: A Data-driven Approach For Solar Potential Estimation Using Rooftop Imagery

Stephen Lee, Srinivasan Iyengar¹, Menghong Feng, Prashant Shenoy, Subhansu Maji
University of Massachusetts, Amherst, ¹Microsoft Research, Bangalore

ABSTRACT

Rooftop solar deployments are an excellent source for generating clean energy. As a result, their popularity among homeowners has grown significantly over the years. Unfortunately, estimating the solar potential of a roof requires homeowners to consult solar consultants, who manually evaluate the site. Recently there have been efforts to automatically estimate the solar potential for any roof within a city. However, current methods work only for places where LIDAR data is available, thereby limiting their reach to just a few places in the world. In this paper, we propose DeepRoof, a data-driven approach that uses widely available satellite images to assess the solar potential of a roof. Using satellite images, DeepRoof determines the roof's geometry and leverages publicly available real-estate and solar irradiance data to provide a pixel-level estimate of the solar potential for each planar roof segment. Such estimates can be used to identify ideal locations on the roof for installing solar panels. Further, we evaluate our approach on an annotated roof dataset, validate the results with solar experts and compare it to a LIDAR-based approach. Our results show that DeepRoof can accurately extract the roof geometry such as the planar roof segments and their orientation, achieving a true positive rate of 91.1% in identifying roofs and a low mean orientation error of 9.3°. We also show that DeepRoof's median estimate of the available solar installation area is within $\pm 11\%$ of a LIDAR-based approach.

CCS CONCEPTS

• **Computing methodologies** → **Image segmentation; Neural networks**; • **Hardware** → **Power and energy**;

KEYWORDS

solar energy, deep learning, semantic segmentation, satellite images

ACM Reference Format:

Stephen Lee, Srinivasan Iyengar¹, Menghong Feng, Prashant Shenoy, Subhansu Maji. 2019. DeepRoof: A Data-driven Approach For Solar Potential Estimation Using Rooftop Imagery. In *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '19)*, August 4–8, 2019, Anchorage, AK, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3292500.3330741>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

KDD '19, August 4–8, 2019, Anchorage, AK, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6201-6/19/08...\$15.00

<https://doi.org/10.1145/3292500.3330741>

1 INTRODUCTION

Advances in photovoltaic (PV) efficiencies and falling manufacturing costs continue to stimulate solar adoption around the world. The average cost of energy from solar photovoltaics has fallen to less than 12¢ per kilowatt-hour (kWh), making it comparable to generation costs from other sources. As a result, more than 70 gigawatts of solar capacity was deployed globally in 2016 alone [1]. Many countries have set ambitious goals for the percentage of renewable penetration in their overall energy mix and solar installations continue to play a dominant role in ongoing deployments.

Solar deployments vary in size, ranging from large solar farms that are deployed by utilities to small-scale installations by individuals [14]. More than half of the installed solar capacity continue to come from small-scale solar deployments, i.e., arrays with 10kW of capacity or less [18]. Most of these installations are residential in nature with deployments on rooftops of homes.

However not all roofs are suitable for solar array deployments. A clear view of the sky with no surrounding obstructions and proper orientation (e.g., south or southwest facing roofs in the northern hemisphere) are key to maximizing solar energy generation of rooftop deployments. In contrast, residential buildings surrounded by trees or other buildings that cast shadows or roofs that do not face south are considered unsuitable for rooftop deployments. The task of determining whether a particular building is well suited for rooftop solar deployment has largely been a manual process—a professional solar energy installer measures the roof area, their orientation, and uses a pyranometer and shade measurement tools¹ to assess the amount of sunlight received on the roof for different times of the day. These measurements are then used to find the ideal locations for installing solar arrays on the roof, if any, and to compute the solar generation potential of the roof. Such a process is laborious and time-consuming and certainly does not scale to large number of buildings in a city.

There have been a few recent efforts that have attempted to automate this laborious process using data-driven algorithm. For instance, Mapdwell [15] and Google's Project Sunroof [26] have both used LIDAR data to assess the solar potential of building rooftops in a city. LIDAR is a laser-based aerial mapping technology that uses airborne LIDAR sensors to extract the 3D surface structure to create a Digital Elevation Model (DEM), which can then be used to determine the geometry of the roof as well as shade from nearby objects [31]. Unfortunately LIDAR data is expensive to collect and involve flying airplanes or drones with aerial LIDAR mapping sensor, and thus, such data is not widely available for many regions in the US and the world. Consequently, current state-of-the-art techniques only offer solar potential data for select cities where LIDAR data is available, leaving large parts of the world without any coverage.

¹Solar Path Finder, SunEye 210 <http://www.solmetric.com/>

At the same time, satellite images showing rooftops of buildings are widely available for most countries through mapping services such as Google, Bing Maps or commercial ones such as DigitalGlobe. Our key hypothesis is that advances in computer vision techniques make it feasible to use 2D rooftop satellite imagery to automate the estimation of solar generation potential for any building rooftop. In our case, key research questions include whether it is feasible to (i) recognize a rooftop from its surroundings, and (ii) infer the 3D shape of the roof, and specifically the plane of each roof surface and their orientation, and (iii) estimate the solar generation potential of the roof based on its location, weather, and potential trees or other visible occlusions. In addition to widely available satellite imagery, historical solar irradiance data for various locations around the world is available from the US National Renewable Energy Lab (NREL) and public tax records in many countries provide information about the number of floors and height of a building. Consequently, we hypothesize that it is feasible to develop an automated data-driven algorithm that utilizes 2D satellite images of building roofs for solar potential estimation and that such an approach has broader applicability than LIDAR-based methods, including vast swaths of rural areas and smaller cities that are unlikely to be mapped by LIDAR in the near future.

In this paper, we present DeepRoof, a data-driven approach that uses satellite rooftop imagery for assessing the solar potential of any roof within a city. In doing so, our contributions are as follows: **Data-driven learning approach:** We design an end-to-end deep learning based approach to estimate the solar potential of a roof and determine the available area for installing solar panels. Our method estimates the size and geometry of the roof in a satellite image as well as identify nearby buildings and trees that can potentially cast a shadow and impact the generation potential. Further, it provides a per-pixel generation potential of each planar roof segment, which can be used to identify ideal locations for deploying solar panels. **System Implementation:** We implemented DeepRoof to automate the process of estimating solar potential from rooftop images. Our approach can be operated in two modes: *batch* and *interactive* mode. In the batch mode, it takes as input a list of addresses and outputs the solar potential of the roof. In the interactive mode, we implement a web interface that takes input from user and displays the solar potential for a given address.

Dataset and Evaluation: Our dataset contains a diverse set of roof types and provides a rich representation for learning different roof shape and orientation. Overall, it contains 4312 labelled planar roof segments from six different cities and 1982 unlabeled images from a city in Framingham, Massachusetts. We evaluate our approach in three distinct ways:

- We compare DeepRoof’s output with the ground truth labeled data and show that it can identify planar roof segments and determine their orientation.
- We validate DeepRoof’s results with two independent solar experts and in a scale of 1 to 10 (10 being the highest), solar experts gave an average rating of >8 for estimating the solar potential of the rooftops.
- We also validate DeepRoof’s output with Google Sunroof, a LIDAR-based technology for solar potential estimation. Our results show that the median percentage difference between

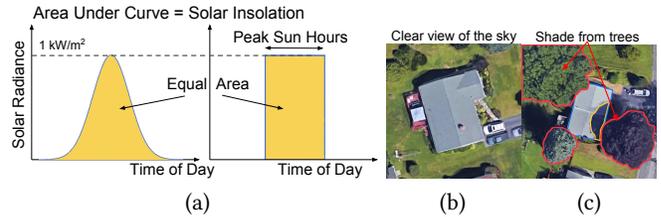


Figure 1: (a) Relationship between peak sun hours and cumulative solar irradiation. (b) Roof with clear view of the sky (c) Roof with shade from nearby structures.

Sunroof and DeepRoof for different groups of roof size is within $\pm 11\%$ of the LIDAR-based approach.

City-scale Solar Potential Analysis: We show that DeepRoof can be used to estimate the roof at city-scale. We note that our approach takes approximately 5 seconds to process a rooftop image and can scale to millions of homes using a server cluster.

2 BACKGROUND

Solar potential of a location can be defined as the amount of available solar energy over a given period. A standard measure for estimating the availability of solar is *peak sun hours*, which captures the amount of solar insolation a location receives on a typical day (see Figure 1(a)). Specifically, a peak sun hour is an hour during which the solar intensity is $1kW/m^2$. Thus, peak sun hours provide a rough estimate of the potential of an area, as it accounts for the various factors that affect available sunlight.

The amount of available sunlight depends on various factors such as the sun’s position in the sky, geography, and local climate conditions such as clouds. Figure 1 illustrates the amount of solar irradiation a surface receives varies over the day and usually peaks at solar noon. Similarly, cloudy conditions can reduce the amount of irradiance a surface gets. Geographical location also plays an important role, especially in places where days are longer during the summer season. For instance, in higher latitudes, cities have more daylight hours in summer than winter.

Estimating the solar potential of a location involves pyranometers that measure the solar irradiance falling on a surface. These pyranometers are usually placed on flat surfaces with clear view of the sky and record the solar insolation a location receives under “ideal” conditions. These solar insolation values are accessible online for several locations around the world [8].

However, assessing the *solar potential of a roof* is challenging as other local factors are involved. One such factor is a roof’s geometry. A roof’s geometry is defined by its (i) orientation — the direction the roof is facing and (ii) pitch — the slope of the roof. Intuitively, the amount of energy generated is proportional to the sunlight incident on the roof’s surface. In the northern hemisphere, a south-facing roof has more direct sunlight exposure than roofs that face in other directions. Thus, the orientation angle of the roof (i.e., the horizontal angle measured from north) determines the actual solar generation output. Similarly, the pitch of the roof also governs the amount of sunlight it receives. A surface that is perpendicular to the incident sunlight will receive more sunlight as more surface area is exposed, whereas a surface parallel to the incident sunlight will

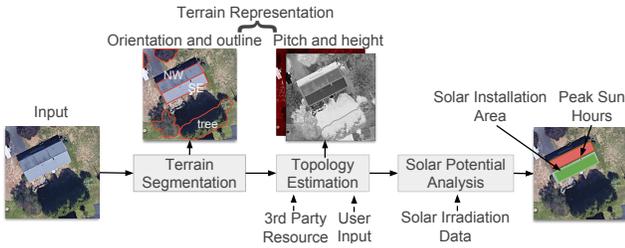


Figure 2: An overview of our DeepRoof architecture.

receive no sunlight. Thus, solar installers position the PV panels to the latitude of a location to maximize the area exposed to sunlight.

Another factor that affects the solar potential of a rooftop is their *local terrain* (see Figure 1(b) and (c)). While roofs with clear view of the sky receive maximum sunlight, buildings with shade from nearby structures such as trees can significantly reduce the amount of solar irradiance incident on its roof. Since available sunlight will be minimal, such roofs may not merit an investment in a PV installation. Thus, these local factors need to be considered to assess a roof’s solar potential.

A naive approach to assess the solar potential of a roof is to use the outline of a building that is available from online maps. Existing mapping services such as OpenStreetMap or Google Maps provide a rough outline of a building property. Assuming that the roof covers the entire building’s outline and is available for solar installation, the outline can be used to estimate the solar potential. However, the building’s outline may also include other areas such as swimming pools or lawns, and thus a naive approach may overestimate the overall potential of the roof. More importantly, information such as roof’s orientation still needs to be considered, which is necessary for estimating the solar potential of a roof.

To address these challenges, our work assumes that the satellite image of a roof is available. We also assume that real-estate property data containing information such as the available floors or height of the building and that of its surrounding buildings are known. Note that existing services already provide all the buildings for a given area [3, 13]. These building locations can be mapped to real-estate data to determine the number of floors. We further assume that the solar irradiance of the location is available.

Problem Statement: Given the location and the satellite image of a candidate building, we first seek to create a model that can extract the roof geometry of the building. We further seek to identify nearby objects such as trees and buildings in the image. Moreover, for each of the planar roof segments that are identified in the candidate building, we need to determine the available sunlight and the roof area for installing solar panels. Formally, given an image of size $w \times h$, we need to compute a set of terrain matrices $(\mathcal{T}_O, \mathcal{T}_A, \mathcal{T}_H, \mathcal{T}_P)$, each of size $w \times h$, where \mathcal{T}_O denotes the outline of all the planar roof segments of a building in the image, \mathcal{T}_A denotes the orientation of each roof planar segment of the candidate building, \mathcal{T}_H denotes the height of all the buildings in the image, and finally \mathcal{T}_P denotes the pitch of the roof planar segments of the candidate building. Further, using these terrain matrices and the geographical location of the building, we need to estimate its solar potential and identify the ideal locations on a roof for solar installation.

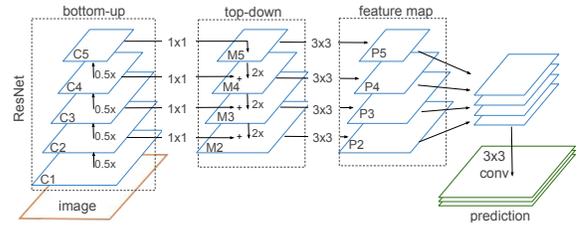


Figure 3: Overview of the FPN framework.

3 DEEPROOF DESIGN

In this section, we describe our data-driven approach to assess the solar potential of a roof. Our approach, DeepRoof, relies on the key observation that size and structure of roofs are observable in a satellite image, essential for estimating the solar system a building can support. Satellite images also indicate if there are nearby structures such as trees or buildings that can obstruct a roof segment partially or completely. These structures can be identified and used to estimate its overall impact on available solar irradiance incident on the roof. DeepRoof, illustrated in Figure 2, uses this insight to compute the solar potential of planar roof segments in a building, and identify suitable locations for installing solar panels. DeepRoof’s approach, shown in Figure 2, has three key steps:

Terrain Segmentation uses deep vision techniques to create a terrain outline of the input image by identifying all the planar roof segments and trees in the image.

Topology Estimation creates a representation of the topology using the terrain outline from the previous step. We approximate the height of the building and nearby structures using publicly available datasets that may cast shadows on the roof.

Solar Potential Analysis estimates the per-pixel solar potential of the roof using the output from the previous steps and historical solar irradiance data. Moreover, our algorithm identifies roof locations where panels will receive maximum sunlight, accounting for shade from nearby structures. Below, we describe each step in detail.

3.1 Terrain Segmentation

The first step in our pipeline is to determine all the planar roof segments, the orientation of each planar roofs and nearby structures in a satellite image. Extracting the roof segments is useful for determining the rooftop locations where solar panels can be installed. Further, trees and nearby buildings provide locations where these objects may cast shadows on the rooftop, thereby rendering them unsuitable for solar panels. Let I be the input satellite image of size $w \times h$, in this step, DeepRoof constructs the terrain matrix \mathcal{T}_A of size $w \times h$, where the pixels correspond either to the orientation of the planar roof segments or trees in the image.

Identifying objects in an image at a pixel-level, referred to as semantic segmentation, is a well-researched computer vision problem [4, 19, 21]. Since recent deep learning approaches have outperformed previous vision techniques on segmentation tasks [10, 24], we leverage these methods in our work. In our approach, we use the Feature Pyramid Network (FPN) to identify planar roof segments and nearby structures. Below we describe the key aspects of FPN, and refer the readers to the original paper [19] for more details.

Figure 3 illustrates the architecture of a feature pyramid network for the segmentation task. FPN takes as input an image and extracts

features using a convolutional neural network (CNN) architecture (e.g. ResNet [11]) augmented with a pyramid-like structure. As shown in the figure, the bottom-up pathway is augmented with a top-down pathway and lateral connections to build a multi-scale feature pyramid of the input image. Since FPN uses a standard CNN architecture for feature extraction, the network can be initialized with pre-trained weights on ImageNet [16] dataset, which allows our technique to work on relatively small datasets.

In a CNN architecture (e.g., ResNet), the bottom layers learn the low-level features such as edges, and as we move higher up, the top layers learn higher-level semantics of a real-world object such as trees, cars etc. In DeepRoof, the CNN architecture learns the planar roof segments, which are the building blocks to construct the geometry of a roof. In FPN, the ResNet layers are grouped into different network stages $\{C_1, C_2, C_3, C_4, C_5\}$, and the output map from the last layer of each stage is selected as a reference set to create the feature pyramid.

As shown, the lateral connections in the top-down pathway combines the low-resolution and the high-resolution from the convolutional network to create a multi-scale feature $\{M_2, M_3, M_4, M_5\}$ by applying a 1×1 convolution filter. A 3×3 convolution filter is applied to the output to obtain the final feature maps $\{P_2, P_3, P_4, P_5\}$. Note that the image resolution of each P_i is one-fourth the input image and has 128 channels each. Finally, $\{P_2, P_3, P_4, P_5\}$ feature maps are concatenated to create a layer with 512 channel. We then use two successive 3×3 convolution filters and batch normalization to create a feature map with channels equivalent to the number of output classes for prediction. The output is then up sampled to its original image size using bilinear interpolation and a softmax activation layer is applied to predict the final output.

We now discuss how our approach creates the roof orientation matrix \mathcal{T}_A . Our approach views each planar roof segment as an object with azimuth as its label. For instance, a planar roof segment facing north-west is labeled as *NW*. Similarly, horizontal roof surfaces are labeled as *flat* and we also label tree crowns. The model is trained using this labeled set of images. After our model is trained, the final output contains a per-pixel prediction such that each pixel is labeled with the class of the object. We then use the final output to create the roof orientation matrix \mathcal{T}_A , where each pixel label corresponds to roof orientation, trees or background.

3.2 Topology Estimation

In this step, we determine the outline matrix \mathcal{T}_O that contains all the planar roof segments of the candidate building. We also describe how we estimate the height and the pitch of the candidate roof. We assume that the outline of the candidate building is available. This is used to determine the roof segments of a candidate building from neighboring rooftops. We note that outline of a building property for a location can be easily obtained from public maps [3]. For example, in a given geographical area, OpenStreetMap provides the outline of all the buildings within a specified area, as well as their addresses [3]. Further, an outline of the candidate building can also be easily obtained as an input through an user interface. In our approach, we use the OpenStreetMap API to obtain the outline of the candidate building in our input image.

In order to recognize the planar roof segments in the orientation matrix \mathcal{T}_A , we run the *marching squares algorithm* [22] that identifies all the contours in an image. The marching squares algorithm approximates the line along the edges where the orientation value changes. The contours correspond to a planar roof segment as we expect the orientation to be similar for a given roof segment. Next, for all the contours predicted by our algorithm, we associate a contour with the candidate building if it intersects with the building’s outline. This creates an outline matrix \mathcal{T}_O , which contains all the planar roof segments of the candidate building.

We then approximate the height and the pitch of the contours identified in the candidate building as well as height of nearby structures. Currently, we rely on third-party sources to create the roof pitch matrix \mathcal{T}_P and the height matrix \mathcal{T}_H . We observe that number of floors available in real-estate dataset and Federal Emergency Management Agency (FEMA) guidelines [12] provide reasonable estimate about the height and pitch of the roof, respectively. As part of our future work, this step can be further improved by obtaining these inputs through an interactive interface from users, which can be used to fine-tune the solar potential estimation.

3.3 Solar Potential Analysis

We now discuss how we compute the solar potential of a roof and the available area for installing solar panels using the terrain matrix.

3.3.1 Solar irradiation on a roof. We note that the solar potential of a roof is the combined potential of all its planar roof segments. We determine the amount of solar irradiance for each planar roof surface in a candidate building for different time of the day in a year, accounting for shade from nearby objects. We now describe how the solar irradiance is computed for a tilted roof surface. The power output of a solar panel depends on the angle sunlight is incident on the PV module, which is maximum when the PV surface is perpendicular to the sun. Thus, the solar irradiance of a roof plane having an orientation $\psi \in \mathcal{T}_A$ and roof pitch $\beta \in \mathcal{T}_P$ are dependent on two components – beam and diffused irradiance. While the beam irradiance S_B is the direct radiation received from the sun, the diffused radiation S_D is received from radiations scattered by particles in the atmosphere. Assuming an isotropic model for diffused irradiance[20], the total solar irradiance of a tilted roof surface is given by:

$$S(\beta, \psi) = \underbrace{S_B \cdot R_B(\beta, \psi)}_{\text{beam irradiance}} + \underbrace{S_D \cdot R_D(\beta, \psi)}_{\text{diffused irradiance}}$$

$$R_B(\beta, \psi) = \cos \alpha \sin \beta \cos(\psi - \theta) + \sin \alpha \cos \beta$$

$$R_D(\beta) = \frac{1 + \cos(\beta)}{2}$$

where, α is the solar elevation angle and varies with the time of the day and θ is the solar azimuth angle and dependent on the latitude of the location. Past values of S_B and S_D are publicly available from various sources [8], and can then be used to compute the total irradiance S for different time of the day in a year.

We consider objects that are roughly within 100 meters from the building for analyzing shadows. We compute the periods when shadows are cast from nearby objects, and subtract the direct radiation S_B from our calculation, i.e., direct sunlight. Note that the

diffused irradiance is still received through scattering, and hence it is not ignored. We then compute the annual peak sun hours at a pixel-level, i.e. number of hours with $1kW/m^2$.

3.3.2 Solar installation size. Finally, we estimate the number of solar panels that can be installed on the roof. The planar roof segments are already available in the terrain matrices. The general procedure is to pack as many solar panels on each of the planar roof segments in a candidate roof. Our problem is similar to the 2D bin packing, where the objective is to maximize the number of 2D shapes that can fit into a rectangular bin. Here, the planar roof segments represent an irregular shaped bin and the 2D object is the solar panel. Since the computational complexity of 2D bin packing is known to be NP-hard, we use a greedy algorithm to determine the number of panels that can fit on the roof. The greedy algorithm outputs the overall number of panel that fits the roof, and we determine the install capacity by multiplying the total number of panels with the rated power output per panel.

4 DEEPROOF IMPLEMENTATION

We have implemented DeepRoof as a system to automate the process of solar potential estimation. DeepRoof can operate in two modes: batch and interactive. In the batch mode, our system takes a list of addresses, or GPS coordinates, as input and computes the solar generation potential for each building in the specified list. The batch mode is useful when computing the solar generation potential of all homes in a neighborhood or an entire city. Our system takes the batch of addresses and first computes the GPS coordinates of each address. It then queries a mapping service, currently set to Google Maps in our implementation, to download the satellite rooftop imagery for each location in the list. The batch of roof images is then provided as input to our DeepRoof model, which outputs the planar roof segments. Our system then uses the approach outlined in Section 3 to output the per-pixel solar potential as well as available roof area. These results can then be viewed by clicking on each address in the list. Figure 4 (a) shows the process for computing the solar potential for a batch of buildings.

Our system can also operate in interactive mode via a web interface. In this case, the user specifies an address or the GPS coordinates of a location. Our system then invokes the backend of DeepRoof, which are the same for both the batch and interactive modes. After computing the results, the per-pixel solar potential is overlaid on the satellite imagery. The interactive interface displays the overall potential of the rooftop at a pixel-level. Figure 4 (b) shows the overall energy potential as well as the available installation area as shown in interactive mode. The bright region indicates the location which receives maximum sunlight.

Overall, our system implementation consists of three components: (i) an interface that allows a user to input an address of a building and visualize its overall solar potential (ii) our deep learning model that identifies the planar roof segments and nearby structures in a rooftop imagery and (iii) a set of APIs that implements our approach in Section 3 to query Google Maps for rooftop imagery and compute the solar potential. DeepRoof’s user interface is implemented using flask, a light-weight web framework in python. Our DeepRoof’s CNN model is implemented using the keras library, which invokes Google’s Tensorflow in the backend.

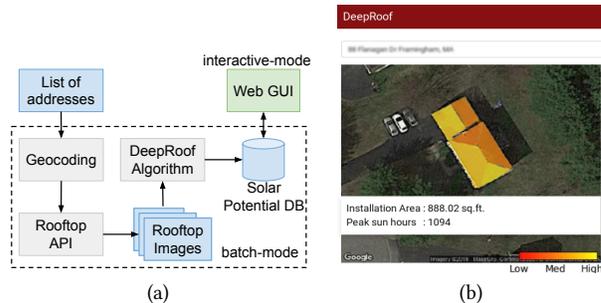


Figure 4: (a) Key components in DeepRoof’s implementation. Our approach supports two modes: batch and interactive mode (using a Web GUI) (b) Screenshot of the web interface that help visualize the solar potential of a building.

Table 1: Dataset 1. Summary of the labeled dataset.

| City | #images | #buildings | #roof segments |
|-------------------|---------|------------|----------------|
| Framingham, MA | 279 | 1161 | 1722 |
| Pinellas Park, FL | 122 | 944 | 2121 |
| Fresno, CA | 43 | 69 | 171 |
| Seattle, FL | 8 | 46 | 158 |
| Denver, CO | 7 | 44 | 90 |
| Indianapolis, IN | 5 | 10 | 50 |
| Total | 464 | 2274 | 4312 |

Table 2: Dataset 2. Key characteristics of the unlabelled Framingham, MA dataset used in our city-scale case study.

| roof types | #buildings | #floors | land area(acres) |
|------------------|------------|---------|------------------|
| Gable, Flat | | | |
| Hip, Complex hip | 1982 | 1 - 6 | 0.031 - 2.92 |

Finally, our system has the ability to parallelize its TensorFlow computations on a cluster of nodes when processing a batch of buildings—in order to scale the computations to a larger number of homes in a region or city.

5 EVALUATION METHODOLOGY

Below we describe our dataset, experimental setup and metrics used to evaluate our approach.

5.1 Dataset

Dataset 1. We collected satellite images from six different cities using Google Maps API (Table 1). We labeled the images using a modified VIA annotator tool [7]. Each roof plane in the image was annotated (including adjacent buildings) and assigned an orientation angle from the north, or labeled as a flat roof. However, we did not label some of the small roof segments, where solar panels cannot be installed. We also labeled nearby trees with visible tree crowns. Apart from the satellite images, we also downloaded the outline and height of the building from OpenStreetMap API [3]. We augmented the height of the building with third-party real-estate datasets in cases where the height was not available.

Dataset 2. For our city-scale case study, we selected a total of 1982 buildings from the city of Framingham, MA (see Table 2). The dataset contains real-estate information such as number of floors, roof type (e.g., hip and gable). We downloaded the satellite images and building outline from Google Maps and OpenStreetMaps. Further, we also collected the solar installation area and available sun hours from Google’s Project Sunroof to compare our approach with a LIDAR-based approach. For estimating the peak sun hour, we used the solar irradiation data from National Solar Radiation Database (NSRDB) [8]. The dataset contains the diffused and direct solar irradiation as well as the azimuth and elevation of the sun for a given location at a granularity of 1 hour. Our dataset is available for download at UMass Trace Repository (<https://bit.ly/2JE8LON>).

5.2 Experimental Setup

We augmented our dataset by rotating the images at different angles. Further, we categorized the orientation directions (0° to 360°) to one of the 16 orientation (i.e., N, NNE, NE, etc.), assigning each roof segment to its closest orientation. In addition to using FPN in DeepRoof, we used other baseline segmentation models – namely UNet [27] and MaskRCNN [10]. Further, the segmentation models were trained using two different CNN architectures namely ResNet 50 and ResNet101, resulting in a total of six models. Since FPN, UNet, and MaskRCNN can use ResNet architecture for feature extraction, pre-trained weights from ImageNet were used as per the literature [10, 16].

5.2.1 Training and model selection. We split our dataset into three disjoint sets: train (60%), validation (20%) and test (20%). The datasets are split before augmenting the dataset to prevent the model from seeing the hold-out set. To prevent overfitting, we trained our models until their performance doesn’t improve further on the validation dataset. Further, we used stochastic gradient descent optimizer, with a learning rate of 0.001 and a momentum of 0.9. For training the model, we ran our neural network model for 240k iterations and reduced the learning rate by a factor of 10 at the 100k and 160k iteration. We report our result on the unseen test dataset.

5.3 Metrics

We note that standard error metrics such as the mean absolute error are not an ideal evaluation metric for capturing the performance of the model in predicting orientation. For instance, if the predicted orientation is NNW (337.5°) and the ground truth orientation is N (0°), the error in prediction is 22.5°. However, mean absolute error will report an error of 337.5°. Thus, we introduce *mean orientation error* (MOE) as a metric to capture the per-pixel error between the predicted and the actual azimuth angle.

$$mean_orientation_err = \frac{1}{M} \sum_i \frac{\sum_j p_{ij} * degree_separation(o_i, o_j)}{t_i}$$

where, M is the number of classes (i.e., azimuths), o_i is the azimuth angle, p_{ij} denotes the number of pixels of azimuth j classified as azimuth i , and t_i is the total number of pixels in class i . Finally, the *degree_separation* is a function that returns the azimuth angle difference between the two azimuths. The value of MOE is between 0°(perfect prediction) and 180°(opposite prediction).

| | | | |
|------|------|------|------|
| BG | 98.2 | 1.2 | 0.6 |
| Roof | 8.7 | 91.1 | 0.2 |
| Tree | 17.0 | 1.7 | 81.3 |
| | BG | Roof | Tree |

(a) DeepRoof (101)

| | | | |
|------|------|------|------|
| BG | 98.0 | 1.2 | 0.8 |
| Roof | 7.9 | 91.9 | 0.2 |
| Tree | 18.1 | 1.8 | 80.2 |
| | BG | Roof | Tree |

(b) UNet (101)

| | | | |
|------|------|------|------|
| BG | 97.2 | 2.0 | 0.8 |
| Roof | 13.6 | 86.3 | 0.2 |
| Tree | 22.2 | 1.7 | 76.2 |
| | BG | Roof | Tree |

(c) MaskRCNN (101)

Figure 5: Normalized confusion matrix of roof classification.

| | | |
|------|------|------|
| Flat | 99.3 | 0.7 |
| Tilt | 4.0 | 96.0 |
| | Flat | Tilt |

(a) DeepRoof (101)

| | | |
|------|------|------|
| Flat | 99.3 | 0.7 |
| Tilt | 5.2 | 94.8 |
| | Flat | Tilt |

(b) UNet (101)

| | | |
|------|------|------|
| Flat | 98.6 | 1.4 |
| Tilt | 2.6 | 97.4 |
| | Flat | Tilt |

(c) MaskRCNN (101)

Figure 6: Normalized confusion matrix of slope type.

6 EXPERIMENTAL RESULTS

In this section, we validate our results with the ground truth including a LIDAR-based approach. We also validate our output with solar experts and show that our data-driven approach can be used to analyze a city-scale dataset.

6.1 Roof classification

We first evaluate the performance of DeepRoof in identifying roofs since these locations are potential sites where solar panels can be installed. To do so, we classify a pixel as a roof if the segmentation model predicts an orientation for the pixel or has labeled the roof as flat. We present our results for all the segmentation models used in our *terrain segmentation* step. Figure 5 shows the normalized confusion matrix of roof classification across all models using ResNet101 architecture on Dataset 1. The values in the confusion matrix are normalized by the number of pixels in each class. We observe that the ResNet50 architecture yields a lower true positive rate² compared to ResNet101 (not shown in the figure). This is because the deeper network enables it to learn the feature subspaces better. In particular, we observe that the true positive rate of ResNet101 reaches 91.1%, 91.9% and 86.3% for DeepRoof, UNet and MaskRCNN respectively. Further, the results of DeepRoof and UNet are comparable with the difference in the true positive rate within $\pm 1\%$ of each other in classifying roofs and nearby structures.

Next, we evaluate the performance of DeepRoof in differentiating between a flat roof and a pitched roof. We classify the slope of a pixel as tilted if the model predicts an orientation for the pixel else we classify it as a flat roof segment. Figure 6 shows the normalized confusion matrix of slope type prediction on Dataset 1. Both ResNet50 and ResNet101 shows high accuracy in differentiating between the types of slope, i.e., flat or tilted, with accuracy >98% and >93% for flat and pitched roofs respectively. As again, ResNet101 architecture yields better result compared to ResNet50 owing to the deep architecture. Although MaskRCNN performs relatively poor in classifying roofs, among the pixels classified as roofs, it has a high true positive rate of >97.4% in differentiating between

²True positive rate is the ratio of true positives identified from all the positive cases.

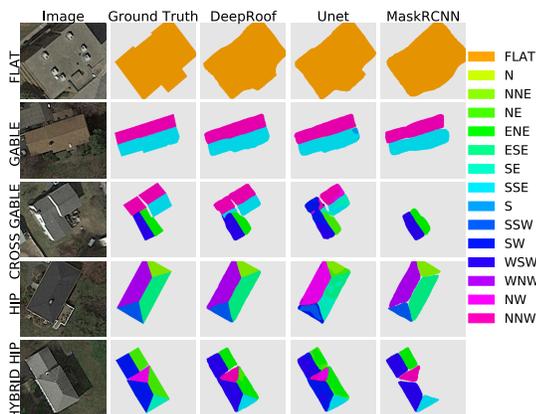


Figure 7: Segmentation output on different roof types.

Table 3: Mean orientation error of the predicted roofs for different architectures on Dataset 1.

| Backbone | DeepRoof | UNet | MaskRCNN |
|-----------|------------|-------|----------|
| ResNet50 | 10.63 | 10.94 | 11.43 |
| ResNet101 | 9.3 | 11.94 | 9.37 |

flat and pitched roof. We also note that DeepRoof has a higher true positive rate compared to UNet and yields better results in identifying slope types. In particular, DeepRoof has a true positive rate of 96% compared to UNet’s 94.8%.

6.2 Roof orientation

Figure 7 shows several examples of the segmentation output of different roof types for all models. Each color in the figure depicts a roof plane and the orientation of the roof plane. The figures illustrate that the models can identify the planar roof segments and also determine their orientation that are visually close to the ground truth. ResNet101 shows better visual improvements in segmentation output compared to ResNet50. Also, we observe that MaskRCNN fails to identify some of the roof segments, which corroborates the lower true positive rate discussed above. As seen in the figure, both DeepRoof and UNet shows a close resemblance to the ground truth, even when the buildings have different geometry.

Next, we evaluate the performance of predicting the orientation of the rooftop. Table 3 summarizes the mean orientation error on Dataset 1. A lower mean orientation error is better, where a zero value indicates that the predicted pixel orientation matches the ground truth. Similarly, a mean orientation error of 180° indicates that the pixels were predicted opposite to the ground truth orientation. Overall, we observe that all the models yield an MOE of less than 12° . Since the orientation labels are 22.5° apart, it indicates that the models are able to predict a rooftop’s orientation correctly in most cases. It is interesting to note that the models are able to learn the spatial relationship between the planar roof segments and predict their orientation. As shown, DeepRoof yields the lowest MOE compared to other segmentation models with MOE of 10.6 and 9.3 using ResNet50 and ResNet101 respectively.

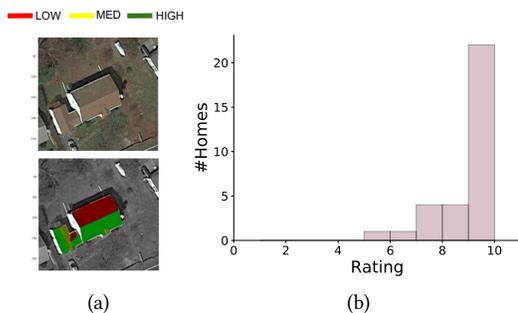


Figure 8: (a) Sample image provided to experts for validation (b) Average rating distribution of the response on a scale of 1 to 10 (10 being the highest).

6.3 Expert validation of our results

For this experiment, we asked two independent solar experts with experience in installing PV panels to rate our solar estimation output from DeepRoof. Our objective with this study was to address the following questions (i) How well does DeepRoof estimate the solar potential of each planar roof segments? (ii) Are there locations on the roof that our approach fails to identify as possible locations where experts would consider installing solar panels?

To answer the above questions, we selected only buildings with pitched roofs from the test dataset and omitted flat roofs that are relatively easier to estimate solar potential. We highlighted all the areas on the roof where solar panels can be installed. Further, we discretized the sunlight received in each planar roof segment into high, medium and low, and presented the image to the experts for analysis. We asked them to rate DeepRoof’s result on a scale of 1 to 10 (10 being the highest rating). We also asked them to consider nearby structures such as trees that may affect the solar potential while rating DeepRoof’s output. In total, we randomly selected 30 images from the test Dataset 1 for our evaluation. Figure 8 (a) shows a sample image from the study that was provided to the experts for validation. The image on the top shows the ground truth, and the image on the bottom shows the output estimated by our algorithm.

Figure 8 (b) shows the average rating distribution of the expert’s response to DeepRoof’s output. The graph shows that DeepRoof estimates the solar potential with high accuracy, as seen from the high rating received for most homes. Overall, we observe that both the experts gave a rating of 8 and above to 22 of the 30 homes. For these homes, DeepRoof not only predicted the orientation correctly but also considered shade from nearby trees to estimate the solar potential. For homes that received a rating lower than 8, in most cases, DeepRoof failed to identify the surrounding trees. We note that these images had fall trees without leaves and, thus, was classified as background pixels. On an average, the experts gave a rating of 8.8 for a typical home. We also received responses on whether DeepRoof identified all the roof segments where solar panels can be installed. Both the experts validated that DeepRoof didn’t miss out any candidate roof segment suitable for solar installation.

6.4 Comparison with a LIDAR-based approach

Google’s Sunroof project, a LIDAR based approach, estimates the solar potential of a roof as follows. For a given address, the Sunroof

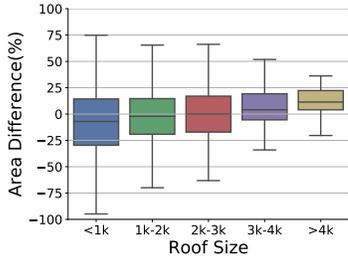


Figure 9: Difference in installation area estimated by Sunroof and DeepRoof for different roof sizes.

provides the total solar installation area and a pixel-level sunlight available on the roof. These estimates are calculated using LIDAR and NREL’s solar irradiance data. To compute the available solar installation area, Sunroof uses a greedy algorithm that maximizes the number of solar panels that can fit on a planar roof segment [26]. Since the pixel-level solar potential is not accessible via Sunroof, we cannot meaningfully compare the results, and hence we only compare the solar installation area of the roof.

In our approach, we select a fixed solar panel size of $250W^3$, and align it based on the orientation and pitch of the roof segments. Our greedy algorithm then uses the panel dimensions as input to analyze the number of panels that can fit on each of the predicted planar roof segments. We ran our algorithm on the test dataset in Dataset 1 and all the buildings in Dataset 2. We had a total of 2073 buildings after combining the datasets.

Figure 9 plots the distribution of percentage difference in solar installation area predicted by Sunroof and DeepRoof for different roof sizes. A negative percentage difference indicates under-prediction and a positive difference indicates over-prediction. As seen in the figure, the median percentage difference between Sunroof and DeepRoof for different roof sizes varies from -7% to 11%. The median percentage difference is 0.5% for roof size between 2000 to 3000 sq.ft. This indicates that on an average DeepRoof’s estimation tends to be close to Sunroof’s estimated area. We also note that the variance decreases with an increase in roof sizes. This is because planar roof segments in small rooftops are comparatively difficult to identify. However, we note that for the first and the third quartiles are within 25% of Google Sunroof’s estimate. Thus, DeepRoof estimates the solar installation area using rooftop images that are close to LIDAR based approaches.

6.5 City-scale solar estimation

We first provide a breakdown of the time it takes for DeepRoof to estimate the potential of a rooftop image. The computation-heavy tasks in DeepRoof involve semantic segmentation of the image, estimating the solar potential and the solar installation area. We note that the DeepRoof model achieves an inference time of 169 ms on an NVIDIA M40 GPU per rooftop imagery. Separately, depending on the number of planar segments identified and nearby structures, calculating the solar potential and the installation area takes on an average 3 to 5 seconds to complete on a single machine. Assuming 5 seconds of processing time per image, a single machine running

³The standard dimensions of a 250W PV panel is 1m×1.65m. Google Sunroof project also uses a 250W panel to analyze the potential [26].

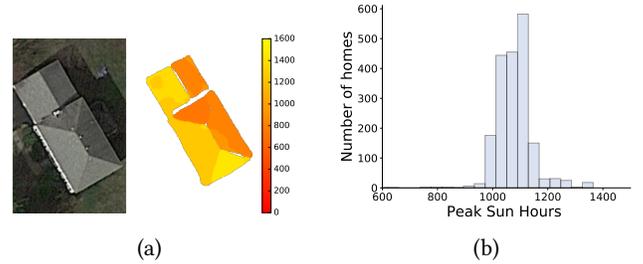


Figure 10: (a) Per-pixel peak sun hours of a building. The bright colored region indicates higher solar potential. (b) Average peak sun hours distribution in Dataset 2.

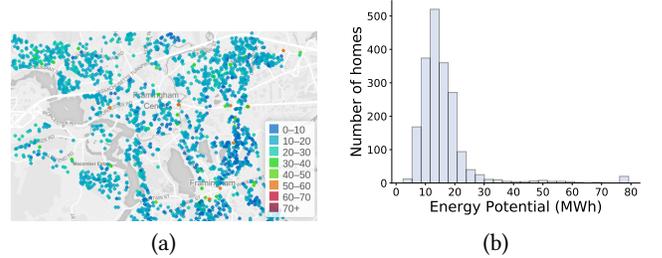


Figure 11: (a) Spatial representation of the annual solar energy generation potential (b) Energy potential distribution.

DeepRoof can process 10,000 buildings in approximately 14 hours. Since processing rooftop images is an embarrassingly parallel, a server cluster can be deployed to speed up the process further. Thus, DeepRoof can easily scale to millions of homes. We now analyze the output of DeepRoof on the citywide buildings in Dataset 2 and present our results below.

6.5.1 Peak Sun Hours. Figure 10(a) shows the rooftop image and the peak sun hours received by each pixel in a sample building. As seen in the image, our technique can identify south-facing and south-west facing roofs with higher solar potential (depicted by the brighter yellow color). Similarly, north and northeast facing roofs have lower energy yield. Thus, DeepRoof’s output can provide custom insights to homeowners on the optimal placement of panels.

Figure 10(b) shows the distribution of the average peak sun hours of all the building in Dataset 2. We observe that the peak sun hours range from 607 to 1471.7 hours. The variation in the average peak sun hours is due to the different orientation and pitch of the roof segments along with the shadows caused by nearby structures. Further, the median peak sun hours available is 1077.95 hours. We observe that all but 5 locations receive a minimum of 800 hours of peak sunlight — indicating significant solar potential among these buildings.

6.5.2 Energy Generation Potential. We compute the energy generation potential assuming the entire roof can generate electricity [28]. Figure 11 (a) shows the spatial representation of the overall solar potential in MWh of the buildings and Figure 11 (b) shows the solar energy distribution of all the buildings. As seen in the figure, the median energy potential of a home is approximately 14.03 MWh. Assuming the national US average energy consumption of 10.9MWh for a typical residential customer [2], our results indicate

that most households can become completely energy self-sufficient using rooftop solar. We also find that on an average, except for homes that have significant foliage or obstruction, the total annual solar production of all the buildings to be 31248.3 MWh, which is at least 1.44 times the annual energy needs of a typical home.

7 RELATED WORK

There has been significant work on estimating the global irradiance at ground level [5, 9]. Previous studies have used satellite data on the earth-atmosphere system and ground pyranometer to measure the variability in solar irradiance for a location [5]. These provide reasonable estimates on how much sunlight is available for a location over a given period [8]. Prior work has also studied the sunlight available on tilted surface [9]. We use these estimation models in our work to study the solar potential of a roof.

Automatically estimating a roof’s solar potential requires identifying buildings and trees. Various methods have been proposed to automatically identify buildings in satellite and aerial images [30, 31]. Most techniques rely on LIDAR based approaches for detecting and modeling buildings [31]. Separately, there have been studies that combine street and aerial images to detect street trees and identify its species [30]. However, most of the existing approaches use LIDAR data for modeling roofs and extracting its geometry. In contrast, our approach provides an alternative to LIDAR-based approaches and uses satellite images for solar potential estimation.

Recently there has been significant interest in estimating the potential of roofs for installing solar panels [17, 25, 26]. While some studies have proposed manual methods [17], others have proposed automated methods for estimating potentials [6, 26]. Manual estimation requires expensive instruments [29] and professionals to reasonably assess a roof’s suitability. On the other hand, automated approaches require LIDAR data, which are not readily available for all cities or remote locations [23]. Unlike prior work, we use satellite images that are readily available from mapping services. Recent advances in deep vision techniques make detection of objects in aerial images feasible [24]. Our work leverages the state-of-the-art vision techniques to approximate both the orientation and the roof’s geometry using only rooftop images and publicly available irradiance datasets. Thus, our approach provides a scalable approach for estimating solar potential in locations where LIDAR is not available.

8 CONCLUSION

Solar potential estimation of a roof can substantially benefit homeowners deciding to adopt solar. In this paper, we proposed DeepRoof, a data-driven approach to estimate the solar potential of a roof using satellite images. We extensively evaluated our approach using available ground truth roof dataset having diverse roof shapes and sizes. We also validated our results with solar experts and compared DeepRoof’s output to a LIDAR-based approach. Our results showed that DeepRoof can accurately extract the roof geometry such as the planar roof segments and their orientation, and achieved a true positive rate of 91.1% in identifying roofs and a low mean orientation error of 9.3°. Further, we also analyzed the solar potential of a city-scale dataset and showed that installing solar panels can lead to energy self-sufficiency in these homes.

Acknowledgements: This research is supported by NSF grants

CNS-1645952, IIP-1534080, CNS-1405826, CNS1505422, IIS-1749833, and the Massachusetts Department of Energy Resource.

REFERENCES

- [1] 2016. Solar Energy Industries Association, Q4 2016 U.S. Solar Market Insight. <http://www.seia.org/research-resources/us-solar-market-insight>. (Dec 2016).
- [2] 2018. How Much Electricity on Average Do Homes in Your State Use? (Ranked by State). <https://goo.gl/j49Wnp>. (May 21st 2018).
- [3] 2018. *OpenStreetMap Overpass API*. https://wiki.openstreetmap.org/wiki/Overpass_API.
- [4] V Badrinarayanan, A Kendall, and R Cipolla. 2015. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561* (2015).
- [5] C Demain, M Journée, and C Bertrand. 2013. Evaluation of different models to estimate the global solar radiation on inclined surfaces. *Renewable Energy* (2013).
- [6] F Dornaika, A Moujahid, Y El Merabet, and Y Ruichek. 2016. Building detection from orthophotos using a machine learning approach: An empirical study on image segmentation and descriptors. *Expert Systems with Applications* (2016).
- [7] A. Dutta, A. Gupta, and A. Zissermann. 2016. VGG Image Annotator (VIA). <http://www.robots.ox.ac.uk/vgg/software/via/>. (2016).
- [8] R George, S Wilcox, M Anderberg, and R Perez. 2007. *National Solar Radiation Database (NSRDB)–10 Km Gridded Hourly Solar Database*. Technical Report. National Renewable Energy Laboratory (NREL), Golden, CO.
- [9] M Gulini, M Vašák, and M Baotic. 2013. Estimation of the global solar irradiance on tilted surfaces. In *17th International Conference on Electrical Drives and Power Electronics*.
- [10] K He, G Gkioxari, P Dollár, and R Girshick. 2017. Mask R-CNN. In *ICCV*.
- [11] K He, X Zhang, S Ren, and J Sun. 2016. Deep residual learning for image recognition. In *CVPR*.
- [12] ICC IBC. 2006. International building code. *International Code Council, Inc.(formerly BOCA, ICBO and SBCCI)* (2006).
- [13] Srinivasan Iyengar, Stephen Lee, David Irwin, Prashant Shenoy, and Benjamin Weil. 2018. WattHome: A Data-driven Approach for Energy Efficiency Analytics at City-scale. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.
- [14] Srinivasan Iyengar, Stephen Lee, Daniel Sheldon, and Prashant Shenoy. 2018. Solarclique: Detecting anomalies in residential solar arrays. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*.
- [15] J A Jakubiec and C F Reinhart. 2012. Towards validated urban photovoltaic potential and solar radiation maps based on lidar measurements, GIS data, and hourly daysim simulations. *IBPSA-USA Journal* (2012).
- [16] A Krizhevsky, I Sutskever, and G Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *NIPS*.
- [17] V Kuthanazhi, S Jois, P Jadhav, K Kumar, A Magal, A Pimpalkhare, J Vasi, A Kottantharayil, K Ramamritham, NC Narayanan, et al. 2016. Estimating Mumbai’s rooftop PV potential through mobilization of IEEE student community. In *Photovoltaic Specialists Conference (PVSC)*.
- [18] Stephen Lee, Srinivasan Iyengar, David Irwin, and Prashant Shenoy. 2017. Distributed rate control for smart solar arrays. In *Proceedings of the Eighth International Conference on Future Energy Systems*.
- [19] T Lin, P Dollár, R Girshick, K He, B Hariharan, and S Belongie. 2017. Feature pyramid networks for object detection. In *CVPR*.
- [20] B Liu and R Jordan. 1961. Daily insolation on surfaces tilted towards equator. *ASHRAE Journal* (1961).
- [21] J Long, E Shelhamer, and T Darrell. 2015. Fully convolutional networks for semantic segmentation. In *CVPR*.
- [22] W E Lorensen and H E Cline. 1987. Marching cubes: A high resolution 3D surface construction algorithm. In *ACM SIGGRAPH computer graphics*.
- [23] R Margolis, P Gagnon, J Melius, C Phillips, and R Elmore. 2017. Using GIS-based methods and lidar data to estimate rooftop solar technical potential in US cities. *Environmental Research Letters* (2017).
- [24] V Mnih and G E Hinton. 2012. Learning to label aerial images from noisy data. In *Proceedings of the 29th International conference on machine learning (ICML)*.
- [25] NREL. 2017. NREL PVWatts Calculator. <http://pvwatts.nrel.gov/>. (Mar 2017).
- [26] Project 2017. Project Sunroof. <https://bit.ly/2YnQ71E>. (March 2017).
- [27] O Ronneberger, P Fischer, and T Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*.
- [28] SolarRoof 2018. Solar Roof. <https://www.tesla.com/solarroof>. (May 2018).
- [29] SunEye 2018. SunEye 210 Shade Tool. <http://www.solmetric.com/>. (March 2018).
- [30] J D Wegner, S Branson, D Hall, K Schindler, and P Perona. 2016. Cataloging public objects using aerial and street-level images-urban trees. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [31] Y Wei, Z Zhao, and J Song. 2004. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. In *IEEE International Geoscience and Remote Sensing Symposium*.