Automatic Damage Detection on Rooftop Solar Photovoltaic Arrays

Qi Li, Keyang Yu, Dong Chen
Florida International University

ABSTRACT

Homeowners may spend up to ~$375 to diagnose their damaged rooftop solar PV systems. Thus, recently, there is a rising interest to inspect potential damage on solar PV arrays automatically and passively. Unfortunately, current approaches may not reliably distinguish solar PV array damage from other degradation (e.g., shading, dust, snow). To address this issue, we design a new system—SolarDiagnostics that can automatically detect and profile damages on rooftop solar PV arrays using their rooftop images with a lower cost. We evaluate SolarDiagnostics by building a lower cost (~$35) prototype and using 60,000 damaged solar PV array images. We find that pre-trained SolarDiagnostics is able to detect damaged solar PV arrays with a Matthews Correlation Coefficient of 0.95.

CCS CONCEPTS

• **Computing methodologies** → **Model development and analysis**; **Modeling methodologies**; **Model verification and validation**.

KEYWORDS

Solar Energy, Machine Learning, Deep Learning, Image Processing

ACM Reference Format:

Qi Li, Keyang Yu, Dong Chen. 2020. Automatic Damage Detection on Rooftop Solar Photovoltaic Arrays. In *The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '20), November 18–20, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3408308.3431130

1 INTRODUCTION

Solar owners may spend up to $375 on the services to maintain their "degraded" rooftop solar PV systems, including damaged solar PV panel inspection, wiring damage, annual inspection, damage localization, and solar PV array cleaning, which typically are not covered in their purchase warranty. Thus, recently, there is a rising interest to inspect potential damage on rooftop solar PV arrays automatically and passively with a lower cost. Traditional approaches [2, 6, 8], which are replying on I-V curve and P-V characteristic monitoring of solar PV system, require user expertise in measuring model parameters and hardware installation (e.g., cameras, solar radiation sensors) to collect training data. While, significant recent work focuses on data-driven approaches [3, 5, 7] require significant amount of historical solar generation data, which may not be available due to the new sites become online, to calibrate their models, and also can not accurately distinguish solar array damage from other degradation (e.g., shading, snow). To address this issue, we design a new system—SolarDiagnostics that can automatically detect and localize damage on solar PV arrays with a lower cost. Our hypothesis is that solar PV arrays are visually identifiable in their rooftop images such that we can leverage image processing and deep learning techniques to automatically profile information. Our evaluation shows that SolarDiagnostics can accurately detect damaged rooftop solar PV arrays and also learn the damage profiling information for each solar site.

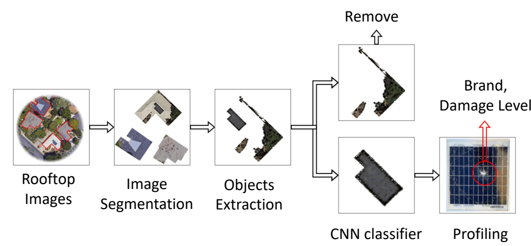

Figure 1: The pipeline overview of SolarDiagnostics.

2 PROBLEM STATEMENT

Given a solar-powered home, we first build a new approach that can automatically fetch its rooftop image. We then present a new approach that can accurately segment rooftop objects and focus on solar panel residing contours in each image. We further seek to build a deep learning classifier to accurately identify the damage on solar PV array. For each reported array, we also want to profile its damage information, such as damage location, damage level and manufacture brand, which can be used to assist solar owners to repair or replace their solar PV arrays promptly. Formally, given a solar PV array-powered home S_i, we first need to segment its rooftop objects $O_i(1 <= i <= N)$ on rooftop image I_i into small "contours"—C_i, where N is the number of objects. Then, we will identify the contours that have damaged solar panel and report their damage level, damage location, and brand information.

3 SOLARDIAGNOSTICS DESIGN

Figure 1 shows the SolarDiagnostics's pipeline of the above operations. In addition to solar PV arrays, many other "outliers" objects such as ridge, structure, chimney, shade, and window may also present on solar PV array rooftops.

Segmenting Rooftop Images. SolarDiagnostics leverages an unsupervised multi-dimensional k-Means algorithm of our SolarFinder [4] to automatically segments rooftop images to solar PV array image contours and other rooftop object contours.

BuildSys '20, November 18–20, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-8061-4/20/11.
https://doi.org/10.1145/3408308.3431130

	Model	MCC
Re-trained Approaches	CNNs	1
	SVMs-RBF	0.803
	Random Forest	0.807
	Decision Tree	0.772
	KNN	0.870
Pre-trained Approaches	CNNs	0.947
	SVMs-RBF	-0.744
	Random Forest	-0.749
	Decision Tree	-0.574
	KNN	-0.695

Table 1: The detection accuracy comparison of SolarDiagnostics when employing different classifiers.

Preprocessing Solar PV Array Images. Although only focusing on solar residing image contours, SolarDiagnostics may still see "outliers" in their image contours. SolarDiagnostics leverages K-Means clustering approach to filter out those white rectangle contours to further removing those True Negative "outliers".
Detecting Damaged Solar Arrays. Next, we build a new Convolutional Neural Networks (CNNs)-based deep learning classifier that can accurately identify damaged solar cells/regions in each solar residing image contour. The CNNs architecture is comprised of input, convolutional layers (ReLU), max pooling, fully-connected layers (ReLU) and output.
Profiling Damaged Solar Arrays. To classify the damage level for each damaged solar PV array, we leverage the supervised machine learning approach—SVMs with linear kernel. To localize damaged "portion" on solar PV arrays, we track the longitude and latitude for each image contour's vertex. SolarDiagnostics uses pixel grayscale distribution feature to identify manufacture brand for each array.

4 IMPLEMENTATION

We implement SolarDiagnostics in `python` using Pandas, OpenCV, Scikit-learn and PyCUDA. SolarDiagnostics leverages Google Image API. In addition, we also build a SolarDiagnostics prototype. Our prototype uses down facing camera on the drone (~$35) to capture images when flying over rooftop solar PV arrays. The images are synchronized via Wi-Fi to Pi-3 based local SolarDiagnostics system.

5 EXPERIMENTAL EVALUATION

5.1 Datasets and Evaluating Metrics

Dataset 1. We use the damaged solar PV array image dataset comprised of ∼60,000 rooftop solar PV array images with the resolution as 1024x1024. We also include damage level, damage location, brand information, and other installation details for each rooftop image.
Dataset 2. We also use our drone-based SolarDiagnostics prototype to test the performance of SolarDiagnostics at 10 "mock" rooftops. The dataset has 10 "mock" residential rooftop images.
Matthews Correlation Coefficient (MCC). To quantify the accuracy of different solar PV array damage(s) detection approaches, we use the Matthews Correlation Coefficient (MCC) [1]. The expression for computing MCC is below,

$$\frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{1}$$

5.2 Experimental Results

Re-trained VS Pre-trained Approaches. All re-trained approaches can access to damaged solar PV array images from their testing sites. For CNNs approaches, we fine-tune the VGGnet using the information from the testing sites. In doing so, we are bench-marking the best performance of different approaches. All pre-trained approaches do not access to damaged solar PV array images from their testing sites. For CNNs approaches, we do not fine-tune the VGGnet using testing sites' data. In doing so, we are bench-marking the practical performance of different approaches.

Table 1 shows that the MCC reported by the pre-trained SolarDiagnostics is significantly better than that of the re-trained machine learning (ML) approaches, including SVMs-RBF, Random Forest, Decision Tree, and KNN. In addition, the pre-trained SolarDiagnostics (CNNs) approach yields the MCC (~0.95) which is slight worse than that (~1.0) of the re-trained SolarDiagnostics (CNNs). This is mainly due to the fact that the pre-trained CNNs approach can not leverage any information from testing images to fine-tune its neural network. Among all the pre-trained approaches, pre-trained CNNs approach has minimum FN as only 5.8%.
Results: *Comparing with both of the re-trained and pre-trained approaches, SolarDiagnostics is the best and stable pre-trained performing approach and it yields the best MCC as 0.95, which is almost the same as re-trained SolarDiagnostics approach.*

6 CONCLUSION AND FUTURE WORK

We design a new defense system—SolarDiagnostics that can automatically detect and localize damage on rooftop solar PV arrays using only their rooftop images. Our evaluation shows that SolarDiagnostics is able to yield an MCC as of 1.0 when detecting damage on solar PV arrays. We plan to learn the performance accuracy of SolarDiagnostics using different type of images (e.g., Tesla roof shingles). **Acknowledgements**. Partially supported by Cyber Florida Collaborative Seed Program.

REFERENCES

[1] S Boughorbel, F Jarray, and M El-Anbari. 2017. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PloS one* 12, 6 (2017).
[2] Mahmoud Dhimish, Violeta Holmes, Bruce Mehrdadi, and Mark Dales. 2017. The impact of cracks on photovoltaic power performance. *Journal of Science: Advanced Materials and Devices* 2, 2 (2017), 199–209.
[3] 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*. 1–10.
[4] Qi Li, Yuzhou Feng, Yuyang Leng, and Dong Chen. 2020. SolarFinder: Automatic Detection of Solar Photovoltaic Arrays. In *Proceedings of the 19th ACM/IEEE International Conference on Information Processing in Sensor Networks*. 100–111.
[5] Andreas Livera, Marios Theristis, George Makrides, and George E Georghiou. 2019. Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems. *Renewable energy* (2019).
[6] Adel Mellit, Giuseppe Marco Tina, and Soteris A Kalogirou. 2018. Fault detection and diagnosis methods for photovoltaic systems: A review. *Renewable and Sustainable Energy Reviews* 91 (2018), 1–17.
[7] Sunil Rao, Sameeksha Katoch, Pavan Turaga, Andreas Spanias, Cihan Tepedelenlioglu, Raja Ayyanar, Henry Braun, Jongmin Lee, U Shanthamallu, M Banavar, et al. 2017. A cyber-physical system approach for photovoltaic array monitoring and control. In *2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)*. IEEE, 1–6.
[8] Salima Sarikh, Mustapha Raoufi, Amin Bennouna, Ahmed Benlarabi, and Badr Ikken. 2019. Photovoltaic system fault identification methodology based on IV characteristics analysis. In *AIP Conference Proceedings*, Vol. 2123. AIP Publishing LLC, 020037.