

DeepSolar for Germany: A deep learning framework for PV system mapping from aerial imagery

Kevin Mayer

Civil and Environmental Engineering
Stanford University
Stanford, USA
kdmayer@stanford.edu

Zhecheng Wang

Civil and Environmental Engineering
Stanford University
Stanford, USA
zhecheng@stanford.edu

Marie-Louise Arlt

Civil and Environmental Engineering
Stanford University
Stanford, USA
mlarl@stanford.edu

Dirk Neumann

Information Systems Research
Albert-Ludwigs-University Freiburg
Freiburg, Germany
dirk.neumann@is.uni-freiburg.de

Ram Rajagopal

Civil and Environmental Engineering
Stanford University
Stanford, USA
ramr@stanford.edu

Abstract—The increasing availability of high-resolution aerial imagery and the recent deep learning-based advances in computer vision have made it possible to automatically map energy systems remotely at a large scale. In this paper, we focus on optimizing the existing DeepSolar framework for photovoltaics (PV) system classification. Specifically, we propose an efficient dataset creation methodology for aerial imagery which allows us to achieve state-of-the-art results, improving the previous model’s recall score by more than eight percentage points to 98% while keeping its precision almost constant at 92%. Furthermore, we show that our optimized model extends its superior classification performance to lower image resolutions. After re-training our optimized model on lower resolution imagery, we apply it to Germany’s most populous state, North-Rhine Westphalia, and deliver a proof of concept for automatically validating, updating, and creating databases of renewable energy systems at a large scale. We conclude with a brief analysis of socio-economic factors correlating with PV system adoption.

Index Terms—Renewable energy, solar, computer vision, machine learning, deep learning, remote sensing

I. INTRODUCTION AND RELATED LITERATURE

As our world is increasingly confronted with the consequences of climate change, many countries attempt to mitigate these effects by rapidly expanding their renewable energy sources. This transformation leads to significant challenges as the decentralized nature of renewable energy systems causes a paradigm shift with respect to system operations: in contrast to conventional systems, renewables are designed as low-capacity units, are geographically distributed, and are owned by residential households and non-utility businesses instead of large specialized generation companies. Therefore, highly accurate and up-to-date information with respect to the location and capacity of renewables is needed to enable efficient system operations and planning processes.

In recent years, machine learning tools have proven to be a powerful approach to retrieve relevant information from aerial imagery [1]. This data has been used to identify various types

of structures, including buildings and roads [2]–[5]. Machine learning has been leveraged to detect rooftops and estimate the potential for photovoltaics (PV) [6]–[8].

Several publications have furthermore targeted the recognition of PV systems from satellite images and aerial imagery. The first contributions in the field have applied support vector machines (SVMs) [9] and random forests [10]. With the progress in deep learning, however, convolutional neural networks (CNNs) have become the dominant tool. With regard to classification, [11] reached a precision of 72% and a recall of 80% and [12] 80% and 95%, respectively, on a test dataset for the city of Fresno in California. [13] have proposed an Inception v3-based CNN to extend classification to the contiguous US. Other examples are [14] for Australia and [15] for parts of Germany and the Netherlands. Beyond classification, approaches for image segmentation have further enabled PV system size estimations [13], [16], [17].

Our work complements the previous line of research in the following ways. First, we achieve a precision of 92% and a recall of 98%, improving the DeepSolar framework [13] and out-performing other existing approaches. Second, we achieve this result by proposing an efficient strategy for dataset creation inspired by [18], [19]. Third, we show that our optimized model extends its superior classification performance to lower image resolutions. Finally, to the best of our knowledge, we are the first to extend our model to a European context and create a PV system database for Germany’s most populous state, North-Rhine Westphalia (NRW).

Our paper is structured as follows. In Section II, we describe our modeling framework. We introduce our approach to improving the model in Section III. Our classification results are presented in Section IV. We conclude with a discussion in Section V and a conclusion in Section VI.

TABLE I: Variables and parameters

i	Picture	TP	True positives
y_i	Label of picture i (1/0: with/without PV system)	TN	True negatives
\hat{y}_i	Estimated score ($\in [0, 1]$)	FP	False positives
w	Weights	FN	False negatives
ω	Linear classifier weights	A_i^{true}	Actual system size
κ	Cohen's kappa coefficient	A_i^{pred}	Predicted system size
p_0	observed accuracy	MAE	Mean absolute error
p_e	expected accuracy	$NMAE$	Normalized MAE
		$MAPE$	Mean absolute percentage error

We define the observed accuracy as $p_0 := \frac{TP+TN}{TP+FP+FN+TN}$ and the expected accuracy as $p_e := \frac{((TP+FP)(TP+FN)+(FN+TN)(FP+TN))}{(TP+FP+FN+TN)^2}$.

II. DEEP SOLAR FRAMEWORK

A. Architecture

Our work is based on the *DeepSolar* CNN [13] which identifies PV systems and estimates their sizes from aerial imagery. The network has two branches:

- The classification branch determines whether a given aerial image contains a PV system or not. It leverages the Inception v3 architecture which is fine-tuned for PV system detection by transfer learning.
- The segmentation branch identifies the exact pixels representing a PV system in a positively classified image. The segmentation branch diverts from the main branch after one of its initial convolutional layers and is trained in a weakly supervised fashion.

The CNN is trained on a manually labeled dataset of aerial imagery and is optimized by the minimization of the class-weighted binary cross-entropy loss.

$$\min_{\hat{y}} -\frac{1}{N} \sum_{i=1}^N (w_1 y_i \log(\hat{y}_i) + w_0 (1 - y_i) \log(1 - \hat{y}_i)) \quad (1)$$

y_i denotes the class label of picture i (i.e. $y_i = 1$ if it contains a PV system and 0 if it does not). \hat{y}_i represents the estimated probability that picture i belongs to class 1. w_1 is the weight assigned for positive and w_0 for negative images in the cross-entropy loss. Further details on the model can be found in [13].

B. Implementation

We transferred the original Tensorflow model to PyTorch and developed the infrastructure such that aerial imagery is automatically downloaded and evaluated, with the results being stored in a database. The respective pipeline can be found on GitHub: https://github.com/kdmayer/PV_Pipeline.

III. RESEARCH APPROACH

In the following section, we describe our optimization approach with regard to data sampling and training. The steps are illustrated in Fig. 1.

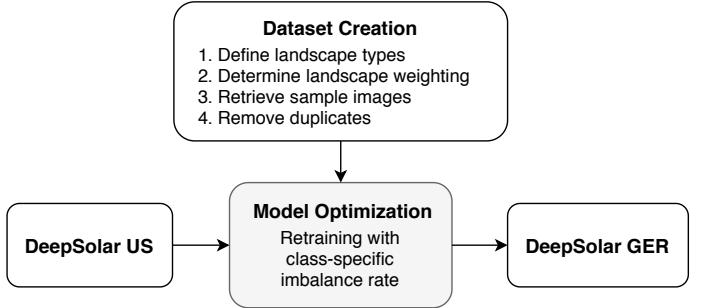


Fig. 1: Optimization Strategy for DeepSolar GER.

A. Dataset Creation Strategy

We propose a four-step strategy to compose aerial imagery datasets for efficiently optimizing PV system classification models. Our strategy complies with the three criteria of representative datasets introduced by [20]:

- 1) The dataset is significantly smaller in size compared to the original dataset.
- 2) The dataset captures most information from the original dataset when compared to other subsets of the same size.
- 3) The dataset has low redundancy among its individual representations.

Focusing on criterion (2), we choose to adapt a data sampling strategy for image scene classification [19] which increases our dataset's heterogeneity with respect to landscape types and depicted objects. This allows us to reduce the dataset size needed to optimize and test the discriminatory capabilities of our classification model.

- 1) **Define landscape types relevant to the area of interest:** We follow this approach by defining seven distinct landscape types, namely urban, rural, fields and pastures, woods, water, rock, and mining sites.
- 2) **Determine landscape weighting:** We choose to select more urban and rural settings in relation to other landscape types, as the majority of small-scale PV systems is installed on roofs. Hence, most of the difficult classification decisions are related to objects on roofs. An effective classifier must be able to distinguish between PV systems and other objects on roofs, e.g. rooftop windows. Moreover, urban and rural settings appear to contain a greater diversity of objects compared to images of non-settled regions.
- 3) **Retrieve sample images:** Based on our weighting, we specify 46 locations all over Germany, each of them belonging to one of the seven landscape types. We use each location as the center point of a regional bounding box from which 800 images are obtained by randomly manipulating their coordinates. Bounding boxes are defined by a side-width of 0.02 degrees for all landscape types except for urban areas where we used 0.04 degrees to capture a broader range of man-made structures.

- 4) **Remove image duplicates:** Finally, images with overlapping areas are removed to reduce the dependencies between our dataset splits.

B. Model Optimization Strategy

1) *Optimizing DeepSolar for classification:* To optimize the classification branch, we rely on the following three established training procedures.

- **Threshold-tuning:** The critical probability is iteratively adjusted above which a softmax-based classification score indicates a positive sample.
- **Fine-tuning:** All weights in all layers are trainable, but initialized with the pre-trained model weights from the original study, i.e. *DeepSolar US*.
- **Feature extraction:** Only one or a few of the final fully connected layers are trainable and are typically initialized with random weights. The remaining weights remain frozen and are copied from *DeepSolar US*.

Importantly, given the high class imbalance between positive and negative samples in the dataset, the *DeepSolar* framework as introduced by [13] extends the cross-entropy loss by a class-specific imbalance rate [21]. The imbalance rate describes the relative weight w_1 given to positive samples. By assigning a value of five to the imbalance rate, we particularly emphasize the loss caused by false negative classifications which would otherwise have a very limited impact on the loss as pictures with PV systems are rare.

Furthermore, for feature extraction and fine-tuning we choose stochastic gradient descent with a mini-batch size of 64 and a learning rate scheduler which decreases the initial learning rate of 0.0001 by 30% every five epochs. We adopt adaptive moment estimation, i.e. Adam, as an optimizer [22] and a weight decay term of 0.0001 which effectively penalizes large weights, thereby reducing the risk of overfitting [23]. Eventually, for the same reason, training is early-stopped when the model’s generalization performance in terms of Cohen’s κ fails to improve for 10 epochs on the validation set.

2) *Optimizing DeepSolar for segmentation:* As introduced by [13], *DeepSolar*’s segmentation branch is trained in a weakly supervised manner. Diverting from the main branch after one of its initial layers, the segmentation branch relies on linear classifier weights ω during training to classify average-pooled feature map vectors as positive if the respective input image depicts a PV system and as negative otherwise.

After training, the linear classifier weights ω are used to produce one class activation map (CAM) for each input image, as illustrated in Fig. 2. As each input image results in a stack of 512 feature maps, CAMs are generated by applying ω to each pixel location across all feature maps separately.

By introducing a tunable threshold value, we can turn CAMs into binary segmentation masks as each pixel’s classification score either falls above or below the respective threshold. Multiplying the number of pixels which exhibit above-threshold classification scores with the area covered by each pixel, we can estimate the PV system size for a given image. The

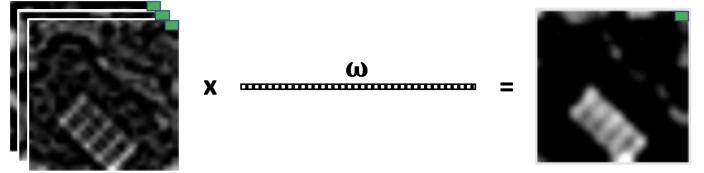


Fig. 2: Generating CAMs from feature maps for segmentation.

threshold value is tuned to minimize the area-wise difference between the predicted segmentation masks generated from our CAMs and their manually segmented counterparts in the validation set.

We perform greedy layer-wise training of the segmentation branch as introduced by [13] to improve the pixel-level classification scores of our CAMs. All training parameters, such as imbalance rate, learning rate, or weight decay, remain as described in the classification setting above.

C. Performance metric

Table II lists the performance metrics we use to evaluate our results. Cohen’s κ relates to the model’s classification performance while the others relate to the model’s segmentation performance.

TABLE II: Performance metrics.

Metric	Definition	Description
Cohen’s κ	$\kappa = \frac{p_o - p_e}{1 - p_e}$	Measure of observed vs. expected accuracy
MAE	$\frac{1}{n} \sum_i^n A_i^{true} - A_i^{pred} $	Absolute mean deviation from true PV system size
NMAE	$\frac{MAE}{\frac{1}{n} \sum_i^n A_i^{true}}$	Normalized absolute mean deviation
Bias	$\frac{\sum_i^n A_i^{true} - A_i^{pred}}{\sum_i^n A_i^{true}}$	Measure of systematic under- or over-estimation
MAPE	$\frac{1}{n} \sum_i^n \left \frac{A_i^{true} - A_i^{pred}}{A_i^{true}} \right $	Mean absolute percentage deviation from true size

MAE, NMAE, and bias: all true negative CAMs are ignored.

MAPE: only true positive and false negative CAMs.

IV. RESULTS

A. Datasets

We created two distinct datasets. The first dataset is based on [24] and created according to our strategy laid out in Section III. This dataset (‘Google-based dataset’) is used to optimize the existing model on imagery with high spatial resolution. It consists of 38,304 images with a visually estimated ground sampling distance (GSD) of 5 $\frac{cm}{pixel}$, which are distributed over different landscape types as displayed in Table III. We extend the original Google-based dataset used by [13] by only 8.1% (original DeepSolar: 472,953).

In contrast, the second dataset is based on [25] and randomly sampled from within the state of NRW. This dataset (‘OpenNRW-based dataset’) is geared toward testing and extending *DeepSolar*’s applicability to lower spatial resolutions.

As this imagery contains four times fewer pixels per area than Google-based imagery, we download upsampled images which mimic Google’s aerial image characteristics with respect to GSD, i.e. 320 by 320 pixels depicting an area of 16 by 16 meters. Both datasets are enriched with positive samples from Germany’s existing PV system registry (MaStR) [26]. The positive samples are distinct between both datasets. Table IV shows the splits for our training, validation, and test sets. All aerial images appear to be taken around midday and have been independently labeled by three research assistants, classified according to majority vote.

TABLE III: Overview and frequency of landscape types in the Google-based dataset.

Landscape Type	# of locations	# of images
Settlements (Urban)	16	13,200 ¹
Settlements (Rural)	10	8,000
Open Land	11	8,800
Woods	4	3,200
Water	2	1,600
Rock	2	1,600
Strip Mining	1	800

¹ We randomly sampled 800 images from each of the 46 rectangles specified. Due to different initial parameter settings, 400 additional images were sampled from the rectangle covering the city of Berlin which we kept as they represent additional information.

TABLE IV: Final statistics for GoogleMaps and OpenNRW-based datasets.

		Positive	Negative	Total	Share [%]
Google Maps	Training	1,198	25,613	26,811	70.0
	Validation	155	3,643	3,798	9.9
	Test	311	7,384	7,695	20.1
	Total	1,664	36,640	38,304	100.0
Open NRW	Training	1,486	16,868	18,354	26.0
	Validation	231	11,688	11,919	16.9
	Test	400	40,000	40,400	57.2
	Total	2,117	68,556	70,673	100.0

B. DeepSolar GER for aerial imagery with a GSD of $5 \frac{\text{cm}}{\text{pixel}}$

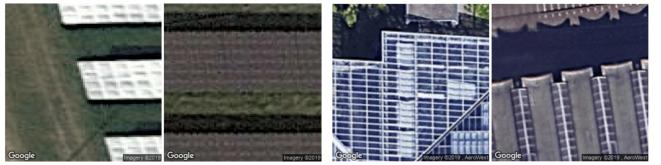
1) *Classification branch*: By fine-tuning *DeepSolar US* with our Google-based dataset, we are able to achieve state-of-the-art precision and recall scores amounting to 92.15% and 98.07% respectively on the Google-based test set. When comparing the performance of *DeepSolar GER* to the results of the original study¹, *DeepSolar GER* increases the original model's average recall score by more than 8.5 percentage points while keeping an almost constant precision score on the test set.

The fine-tuned classification branch is chosen to evaluate the Google-based test set since it proves to be superior to the classification branches trained by threshold-tuning and feature extraction as illustrated by their respective validation set results in Table V.

¹On its US-based test set, [13] reported precision and recall scores of 93.1% and 88.5% in residential areas and of 93.7% and 90.5% in non-residential areas.

TABLE V: Classification performance on the validation set of 3,798 images with a GSD of $5 \frac{cm}{pixel}$.

Training Approach	Original Model	Threshold Tuning	Feature Extraction	Fine-Tuning
Threshold	0.5	0.075	0.5	0.5
Cohen's κ	84.33%	89.10%	87.41%	94.16%
Precision	93.08%	90.73%	80.65%	91.02%
Recall	78.06%	88.39%	96.77%	98.06%



We find that fine-tuning the classification branch greatly improves the recall score. While the threshold-tuned version of *DeepSolar US* misclassifies 18 out of 155 positive images as false negatives, *DeepSolar GER* does so in only three cases. Comparing the subset of false negatives which depict solar farms, it becomes clear that image characteristics linked to poor frame prevalence, uncommon color schemes, and strong reflections are particularly pronounced among the false negatives as illustrated by Figure 3a. However, in contrast to the threshold-tuned version of *DeepSolar US*, *DeepSolar GER* is less prone to misclassifications from reflections.

Furthermore, both models exhibit almost identical behavior when it comes to false positives. While the threshold-tuned version of *DeepSolar US* misclassifies 14 out of the 3,643 negative images in our validation set as false positives, *DeepSolar GER* does so in 15 cases. Moreover, both models arrive at the same misclassifications in eight cases, illustrating that the majority of false positives remains the same. We observe that the driving factor for false positives are window-like structures, such as in Fig. 3b. They often adhere to a PV system's size and the dark blue color scheme, seem to be composed in a modular fashion, are nestled in a white frame, are installed in a roof-like environment, and are characterized by an angular, sometimes even rectangular, shape.

2) *Segmentation branch*: On our test set of 7,695 images, the optimized segmentation branch with threshold 0.2 achieves a precision of 92.66% and a recall of 97.43% for image-level classifications, which translates into a κ statistic of 94.77%. Deploying our re-trained segmentation branch to produce pixel-level segmentation masks, we achieve MAE, NMAE, MAPE, and bias statistics of 17.47 m^2 , 23.34%, 21.73%, and 0.54% on the test set, respectively.

According to Section III-B2, we choose a threshold of 0.2 for the re-trained segmentation branch to evaluate the test set, as it achieves the lowest MAE of 18.88 m^2 per non-negative CAM on the validation set, complemented by an NMAE of 25.14%, a bias estimate equal to 0.53%, and a MAPE of 23.84% as illustrated by Fig. 4

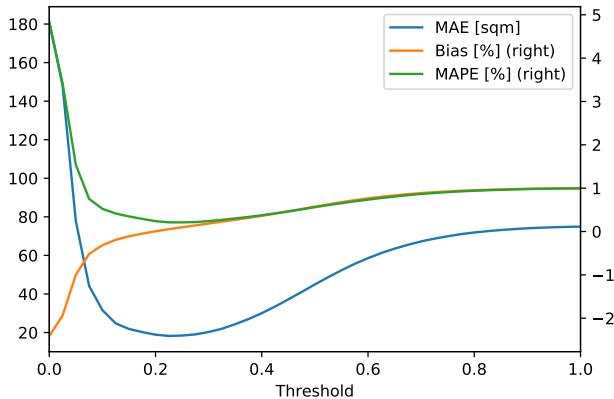


Fig. 4: MAE, bias, and MAPE with respect to different threshold values, measured on the Google-based validation set.

TABLE VI: Comparing the classification performance of different *DeepSolar* versions on the OpenNRW validation set.

	<i>DeepSolar US</i>	<i>DeepSolar GER</i>	<i>DeepSolar NRW</i>
Training set ¹	Google US	+ Google GER	+ OpenNRW
Threshold	0.075	0.025	0.375
Cohen's κ	56.8%	75.5%	82.9%
Precision	96.9%	80.6%	79.8%
Recall	40.7%	71.9%	87.0%

¹ Each model version is trained on a different training set, with "+" indicating the extension of a training set with respect to the one(s) on the left.

C. DeepSolar NRW for aerial imagery with a GSD of $10 \frac{\text{cm}}{\text{pixel}}$

By fine-tuning *DeepSolar GER* on our OpenNRW-based dataset, we are able to achieve precision and recall scores of 63.96% and 86.69% on the OpenNRW test set.

DeepSolar GER is chosen to be fine-tuned for OpenNRW as it outperforms *DeepSolar US* by almost 19 percentage points with respect to Cohen's κ on the OpenNRW validation set, as illustrated in Table VI. This difference of 19 percentage points in κ indicates that the increase in recall by 31.2 percentage points more than compensates the drop in precision of 16.3 percentage points on the validation set. Apart from that, the difference of almost 16 percentage points between the validation and test set precision scores can be attributed to the fact that we artificially enriched both dataset splits with PV system imagery. As the test set ratio of positive to negative samples is much closer to reality than the ratio of the validation set, the test set performance should be considered a more reliable estimate and has also been validated during model deployment in our case study for openNRW.

D. Comprehensive PV system database for NRW

Applying *DeepSolar NRW* to the whole state of NRW, we are able to identify 1,147,048 unique geo-referenced aerial images depicting PV systems. Additional effort will be needed to merge connected PV systems to unique PV installations as a single input image to *DeepSolar NRW* can either depict

multiple PV units or cut through them. However, we are able to use the number of positively classified images as an indicator for the frequency of PV installations and explore its correlation with important socio-economic data provided by [27].

TABLE VII: Correlation between PV frequency and selected socio-economic factors for NRW's 53 counties [27].

Socio-economic factor	Correlation w.r.t. PV image count
Gross Value Added ¹ by agriculture	+0.75
Unemployment rate	-0.61
GDP per person	-0.13
Disposable income per person	+0.20
Population density	-0.64

On a county-level, the socio-economic analysis reveals a strong positive correlation between the frequency of PV systems and the agricultural sector size. Furthermore, we observe moderate negative correlations between the frequency of PV systems and the unemployment rate as well as between the frequency of PV systems and the population density. These results indicate that counties with a low unemployment rate and a relatively high agricultural sector activity appear to adopt PV systems more frequently.

V. DISCUSSION

DeepSolar GER achieves state-of-the-art classification results with a precision of 92.15% and a recall of 98.07% on our Google-based test set. With high-resolution imagery becoming ever more available, these results mark an important step towards remotely mapping energy infrastructure, like PV systems, on a large scale. We think that further improvements for PV system classification are hardly possible as misclassified images are even hard for humans to identify, e.g. due to reflections, shadows, or other building structures closely resembling PV systems [1].

We achieve these results by oversampling difficult sceneries and emphasizing the impact of false negatives on the cross-entropy loss by choosing a large imbalance rate. Yet, while an oversampled training set helps to focus training efficiently, validation and test sets should ideally be drawn at random to obtain an unbiased generalization estimate. Therefore, we validate our results as reported in Section IV-B1 on a randomly sampled test set of 30,326 images, for which we find a κ of 94.18% with a recall of 97.43% and a precision of 91.27%. This demonstrates that our model can be used to evaluate unknown aerial images with a given resolution without a loss in performance.

We show that our optimized framework, *DeepSolar GER*, extends its superior classification performance to imagery with four times fewer pixels per area than the original study. Using additional imagery with lower spatial resolution for training, we continue to improve the classification results of *DeepSolar GER* on low-resolution imagery. Future research should improve and extend *DeepSolar NRW*'s capabilities to imagery with even lower spatial resolutions and extend the

model's applicability to areas suffering from a lack of high-resolution imagery, including developing countries.

Furthermore, we create a comprehensive solar system database for Germany's most populous state. We thereby enable system operators to compile, verify, and update their solar system registries even in the absence of a central database. Future research will need to organize the classification data and provide additional information regarding other factors which are relevant from a system operator's perspective. As such, the actual PV capacity installed could be approximated by the size of the PV system as estimated by *DeepSolar*'s segmentation branch. Future work should therefore aim to improve the current segmentation results with a MAE of 17.47 m^2 and a MAPE of 21.73%. Leveraging 3D building data, future research could also aim to model the tilt angle and orientation of PV systems, incorporating an important step towards effective solar nowcasting.

VI. CONCLUSION

In our work, we optimize *DeepSolar*, a CNN capable of detecting and estimating PV systems from aerial imagery, and extend it to low-resolution imagery and a European context. Our work illustrates the potential that concerted data sampling and class-specific loss strategies can offer. This approach allows us to significantly improve the performance of our classification model while drastically reducing the time and resources needed for re-training.

REFERENCES

- [1] J. de Hoog, S. Maetschke, P. Ilfrich, and R. R. Kolluri, "Using satellite and aerial imagery for identification of solar pv: State of the art and research opportunities," in *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*. Association for Computing Machinery, 2020, p. 308–313.
- [2] X. Jin and C. H. Davis, "Automated building extraction from high-resolution satellite imagery in Urban areas using structural, contextual, and spectral information," *Eurasip Journal on Applied Signal Processing*, no. 14, pp. 2196–2206, 2005.
- [3] V. Mnih and G. E. Hinton, "Learning to detect roads in high-resolution aerial images," in *Proceedings of the 11th European Conference on Computer Vision: Part VI*, 2010, p. 210–223.
- [4] J. Yuan and H.-h. L. Yang, "Large-Scale Solar Panel Mapping from Aerial Images Using Deep Convolutional Networks," in *2016 IEEE International Conference on Big Data (Big Data)*, 2016, pp. 2703–2708.
- [5] N. Audebert, B. Le Saux, and S. Lefèvre, in *Semantic Segmentation of Earth Observation Data Using Multimodal and Multi-scale Deep Networks*, 2017, pp. 180–196.
- [6] K. Mainzer, D. Schlund, S. Killinger, R. McKenna, and W. Fichtner, "Rooftop PV Potential Estimations: Automated Orthographic Satellite Image Recognition based on Publicly Available Data," 2017. [Online]. Available: <http://publica.fraunhofer.de/documents/N-426206.html>
- [7] Google Inc., "Project Sunroof," 2020, last access on 2020-07-12. [Online]. Available: <https://www.google.com/get/sunroof>
- [8] S. Lee, S. Iyengar, M. Feng, P. Shenoy, and S. Maji, "Deepprof: A data-driven approach for solar potential estimation using rooftop imagery," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, 2019, p. 2105–2113.
- [9] J. M. Malof, R. Hou, L. M. Collins, K. Bradbury, and R. Newell, "Automatic solar photovoltaic panel detection in satellite imagery," in *2015 International Conference on Renewable Energy Research and Applications*, vol. 5, 2015, pp. 1428–1431.
- [10] J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied Energy*, vol. 183, pp. 229–240, 2016.
- [11] —, "A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery," in *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pp. 650–654, 2016.
- [12] J. M. Malof, L. M. Collins, and K. Bradbury, "A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 874–877.
- [13] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, "DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States," *Joule*, vol. 2, no. 12, pp. 2605–2617, 2018.
- [14] K. Stamatou, "How We Used Deep Learning to Identify Solar Panels on 15 Million Buildings," 2018, last access on 2020-07-10, [Online]. Available: <https://blog.maxar.com/earth-intelligence/2018/how-we-used-deep-learning-to-identify-solar-panels-on-15-million-buildings>.
- [15] CBS DeepSolaris, "Automatically detect solar panels with aerial photos," 2020, last access on 2020-07-10. [Online]. Available: <https://www.cbs.nl/en-gb/about-us/innovation/project/automatically-detect-solar-panels-with-aerial-photos>
- [16] J. Camilo, R. Wang, L. M. Collins, K. Bradbury, and J. M. Malof, "Application of a semantic segmentation convolutional neural network for accurate automatic detection and mapping of solar photovoltaic arrays in aerial imagery," 2018. [Online]. Available: <http://arxiv.org/abs/1801.04018>
- [17] X. Hou, B. Wang, W. Hu, L. Yin, A. Huang, and H. Wu, "SolarNet: A Deep Learning Framework to Map Solar PowerPlants In China From Satellite Imagery," 2020. [Online]. Available: <https://arxiv.org/pdf/1912.03685.pdf>
- [18] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870, 2007.
- [19] G. Cheng, J. Han, and X. Lu, "Remote Sensing Image Scene Classification: Benchmark and State of the Art," in *Proceedings of the IEEE*, vol. 105, no. 10, 2017, pp. 1865–1883.
- [20] Feng Pan, Wei Wang, A. Tung, and Jiong Yang, "Finding Representative Set from Massive Data," in *5th IEEE International Conference on Data Mining (ICDM'05)*, 2003, pp. 338–345.
- [21] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–449, 2002.
- [22] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," pp. 1–15, 2014. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [23] "A Simple Weight Decay Can Improve Generalization," in *Conference on Neural Information Processing Systems*, 1992.
- [24] Google Inc., "Google Maps Static API."
- [25] Beauftragter der Landesregierung NRW fuer Informationstechnik (CIO), "openNRW," 2017, last access on 2019-06-10. [Online]. Available: <https://open.nrw/geobasisdaten-des-landes-nrw-gebuehrenfrei-als-open-data-verfuegbar>
- [26] BNetzA, "Marktstammdateinregister," 2019, last access on 2019-04-05. [Online]. Available: https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_{_}Institutionen/MaStR/MaStR.html
- [27] Federal Statistical Office (Destatis), "Regionalatlas," last access on 2020-07-14. [Online]. Available: <https://www.regionallstatistik.de/genesis/online/logon>