Automated Rooftop Solar Panel Detection Through Convolutional Neural Networks

Pena Pereira, Simon; Rafiee, Azarakhsh; Lhermitte, Stef

DOI
10.1080/07038992.2024.2363236

Publication date
2024

Document Version
Final published version

Published in
Canadian Journal of Remote Sensing

Taylor & Francis
Taylor & Francis Group

Canadian Journal of Remote Sensing
Journal canadien de télédétection

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujrs20

Automated Rooftop Solar Panel Detection Through Convolutional Neural Networks

Simon Pena Pereira, Azarakhsh Rafiee & Stef Lhermitte

To cite this article: Simon Pena Pereira, Azarakhsh Rafiee & Stef Lhermitte (2024) Automated Rooftop Solar Panel Detection Through Convolutional Neural Networks, Canadian Journal of Remote Sensing, 50:1, 2363236, DOI: 10.1080/07038992.2024.2363236

To link to this article: https://doi.org/10.1080/07038992.2024.2363236

Published online: 04 Jul 2024.

CANADIAN JOURNAL OF REMOTE SENSING
2024, VOL. 50, NO. 1, 2363236
https://doi.org/10.1080/07038992.2024.2363236

Taylor & Francis
Taylor & Francis Group

Automated Rooftop Solar Panel Detection Through Convolutional Neural Networks

Détection automatisée de panneaux solaires sur toiture par réseaux neuronaux convolutifs

Simon Pena Pereira[a], Azarakhsh Rafiee[a], and Stef Lhermitte[b,c]

[a]Geo-Database Management Center, Delft University of Technology, Delft, The Netherlands; [b]Department Earth & Environmental Sciences, KU Leuven, Leuven, Belgium; [c]Department Geoscience & Remote Sensing, Delft University of Technology, Delft, The Netherlands

ABSTRACT

Transforming the global energy sector from fossil-fuel based to renewable energy sources is crucial to limiting global warming and achieving climate neutrality. The decentralized nature of the renewable energy system allows private households to deploy photovoltaic systems on their rooftops. However, inconsistent data on installed photovoltaic (PV) systems complicate planning for an efficient grid expansion. To address this issue, deep-learning techniques, can support collecting data about PV systems from aerial and satellite imagery. Previous research, however, lacks the consideration for ground truth data-specific characteristics of PV panels. This study aims to implement a semantic segmentation model that detects PV systems in aerial imagery to explore the impact of area-specific characteristics in the training data and CNN hyperparameters on the performance of a CNN. Hence, a U-Net architecture is employed to analyze land use types, rooftop colors, and lower-resolution images. Additionally, the impact of near-infrared data on the detection rate of PV panels is analyzed. The results indicate that a U-Net is suitable for classifying PV panels in high-resolution aerial imagery (10 cm) by reaching F1 scores of up to 91.75% while demonstrating the importance of adapting the training data to area-specific ground truth data concerning urban and architectural properties.

RÉSUMÉ

Pour limiter le réchauffement de la planète et parvenir à la neutralité climatique, il est essentiel de transformer le secteur mondial de l'énergie des combustibles fossiles aux sources d'énergie renouvelables. La nature décentralisée du système d'énergie renouvelable permet aux ménages de déployer des systèmes photovoltaïques sur leurs toits. Cependant, l'incohérence des données relatives aux systèmes photovoltaïques installés complique la planification d'une expansion efficace du réseau. Pour résoudre ce problème, les techniques d'apprentissage profond, peuvent faciliter la collecte de données sur les systèmes photovoltaïques à partir d'images aériennes et satellitaires. Les recherches antérieures ne tiennent toutefois pas compte des caractéristiques spécifiques des panneaux photovoltaïques dans les données véritables de terrain. Cette étude vise à mettre en œuvre un modèle de segmentation sémantique qui détecte les systèmes photovoltaïques dans l'imagerie aérienne afin d'explorer l'impact des caractéristiques spécifiques à une zone dans les données d'apprentissage et les hyperparamètres CNN sur les performances d'un CNN. Une architecture U-Net est donc employée pour analyzer les types d'utilization du sol, les couleurs des toits et les images à faible résolution. En outre, l'impact des données dans le proche infrarouge sur le taux de détection des panneaux photovoltaïques est analysé. Les résultats indiquent qu'un réseau U-Net est adapté à la classification des panneaux photovoltaïques dans les images aériennes à haute résolution (10 cm) en atteignant des scores F1 allant jusqu'à 91,75%, tout en démontrant l'importance d'adapter les données d'apprentissage aux données de référence spécifiques à la zone analysée, concernant ses propriétés urbaines et architecturales.

ARTICLE HISTORY
Received 10 December 2023
Accepted 27 May 2024

CONTACT Azarakhsh Rafiee a.rafiee@tudelft.nl

Introduction

Nearly three-quarters of human-caused greenhouse gas emissions that drive climate change stem from the energy sector, making climate change primarily an energy problem (ClimateWatch 2022). As a result, and driven by climate policy, the energy sector is increasingly shifting toward more renewable and sustainable energy sources in line with the Paris Agreement commitments to limit global warming to an average of well below 2 °C compared to pre-industrial levels (UNFCCC 2015). Transitioning to renewable energy technologies is key to a clean and secure energy system on the path to climate neutrality (UN 2022). In this transition, solar energy is the fastest-growing and most competitive source of renewable energy in the European Union (EC 2022).

A popular technology to convert sunlight into electricity is photovoltaic (PV) systems that rely on solar cells using the photovoltaic effect. These cells compose PV panels that can be installed in large-scale solar power plants on the ground, floating systems on lakes, or in decentralized systems on rooftops. It is worth noting that rooftop systems in aerial imagery resemble solar thermal collector (STC) systems in their shape, color, and size, which are, in contrast to PV systems, primarily used to generate heat for hot water in residential buildings (EC 2022).

The energy sector's growth is expected to continue in the upcoming decades mainly driven by PV systems, which are the most accessible sources of renewable energy for private households (EC 2022). Due to this energy sector liberalization, national agencies, such as the Federal Network Agency of Germany, demand a comprehensive and reliable data basis for planning grid expansions (MaStR 2023). All energy system operators in Germany are required to register with the agency. However, documentation for small plug-in solar units is often insufficient, which leads to inaccurate capacity estimates (EUPD 2023). Depending on the country, well to poorly documented registries of active PV systems exist, which are a hurdle for decision makers involved in the development of an efficient energy transition.

An alternative method for populating the registries or statistics with up-to-date information about installed PV systems involves leveraging remote-sensing techniques. These methods enable the identification of PV panels in satellite or aerial imagery. In recent years, a variety of methods have been employed to extract PV panels from remote sensing imagery. Traditional methods include region–line primitive association analysis and template matching (Wang et al. 2018), and machine learning algorithms, such as support vector machine classifiers (Malof et al. 2015) and random forest classifiers (Malof et al. 2016; Feng et al. 2024).

However, recent advancements in deep learning (DL) algorithms in computer vision boosted the integration of deep neural networks in remote sensing applications. Commonly used networks are convolutional neural networks (CNNs) for semantic segmentation and region-based CNNs for object detection or instance segmentation (Gui et al. 2024). As indicated by Rausch et al. (2020), DL algorithms for image classification, such as CNNs, can be useful for validating, updating, and completing PV system registries. Given the challenges of object detection in remote sensing imagery, namely the comparatively small footprint of the object and its diverse distribution in different geographical locations, it remains difficult to develop a single method that performs well for different locations and camera sensors (Gui et al. 2024). Considering this and the recent trend toward DL methods in object detection, this study will only discuss its approach and results in the context of related results from CNNs detecting PV panels in remote sensing imagery (Puttemans et al. 2016).

The comprehensive work by De Jong et al. (2020) demonstrates the ability to classify solar panels with CNNs. To optimize the effectiveness as well as the efficiency of these algorithms, research has mainly focused on the technical configurations of these networks. However, the performance of these algorithms is calculated on the basis of the network's prediction in comparison to the ground truth data. Therefore, it is also crucial to understand the impact of diverse ground truth data on the performance of the network. The importance to analyze the ground truth of PV panels and their surroundings (e.g., differences in land use or architectural characteristics) becomes evident when statistics about PV panels are created on a national or international scale. This hurdle became evident when De Jong et al. (2020) conducted validations across different geographical areas. Understanding the impact of building densities and sizes, rooftop colors and shapes, and PV systems sizes on network performance and classification results is crucial, as these factors may vary across regions and countries.

Moreover, Da Costa et al. (2021) suggest a shift from model-driven to data-driven research to detect PV panels. By comparing multiple CNN models, it became evident that their results differ insignificantly, underscoring the importance of reliable and comprehensive data sets of annotated PV panels. Knowing

the ground truth characteristics helps in the collection of appropriate ground truth data and algorithms, allowing for reasonable predictions for the object of interest.

This study aims to explore the overall effectiveness of a U-Net in detecting rooftop solar panels. Specifically, it focuses on analyzing the specific impacts of land use types, spectral bands (e.g. near-infrared (NIR)), correlations between roof and panel color, and spatial resolutions of aerial imagery on detecting rooftop solar panels using a U-Net algorithm.

The DL algorithm employed in this work is a CNN with a U-Net architecture developed by Ronneberger et al. (2015). The U-Net architecture is a straightforward CNN that has demonstrated promising results for similar applications in previous research (Castello et al. 2019; Da Costa et al. 2021). It computes semantic segmentations, which are classified images in which each pixel is associated with a target class or background information. Minor modifications in the U-Net are required to obtain a semantic segmentation with the same dimensions as the input image and label. The input consists of manually generated ground truth labels representing the target class of PV panels and aerial images at a resolution of 10 cm per pixel to compute pixel-based classifications of PV panels. Overall, the methodologies employed are compiled into a semi-automated pipeline.

Data and methods

Study area

The study area is located in Cologne, North Rhine-Westphalia (NRW), Germany, and is divided into three subareas to analyze the impact of different land use types on the detection of PV panels (see Figure 1). The subareas were carefully defined to cover the most common forms of built environments where rooftop-installed PV systems are found. The environments differ not only in land use types but also in their urban or rural characteristics. The northern subarea (1) is a commercial area in the Ossendorf district. The second subarea primarily consists of residential and mixed land use and comprises the city

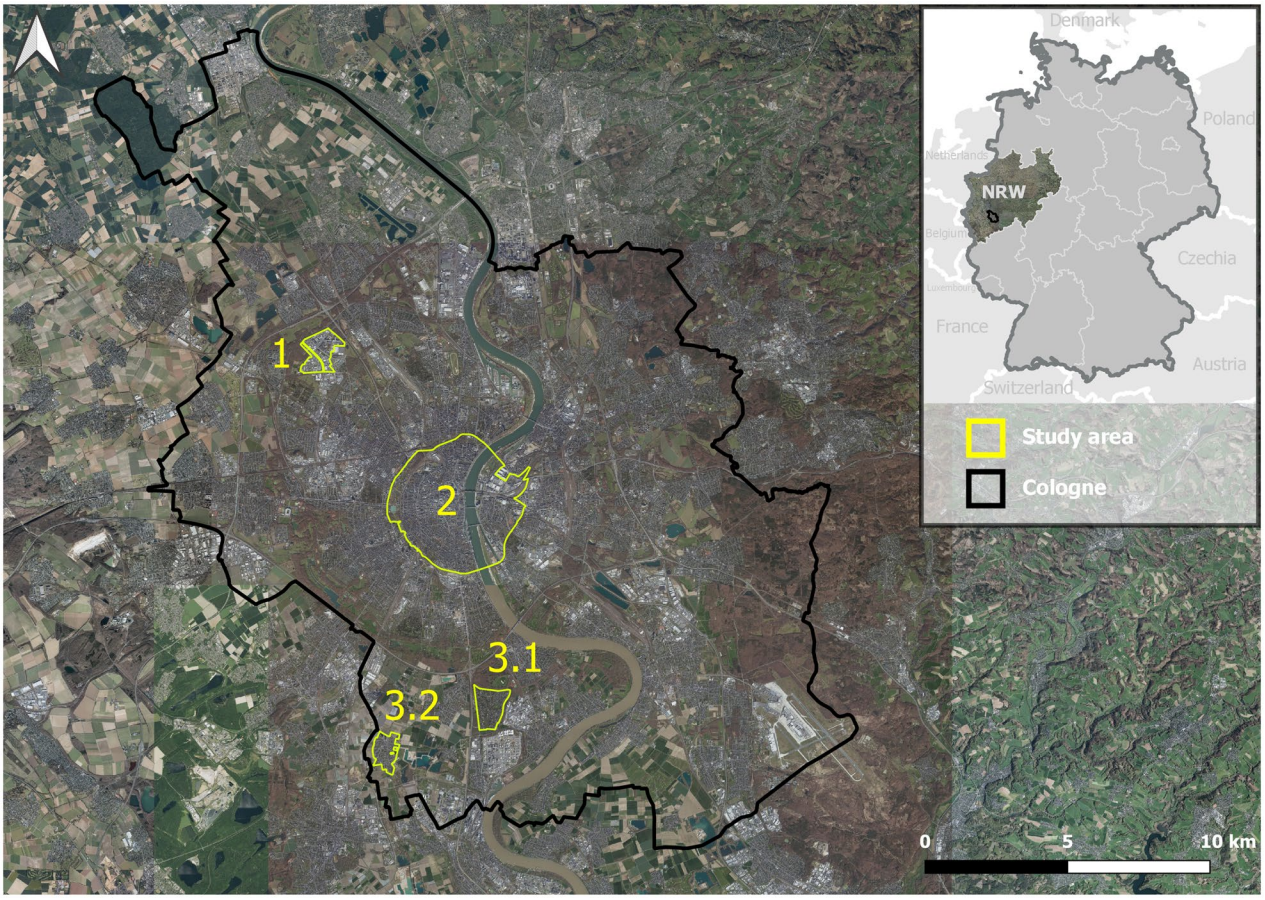

Figure 1. Study area overview: (1) commercial area, (2) city center, (3.1) suburb (Hahnwald), (3.2) suburb (Meschenich); Subareas outlined in yellow; Cologne outlined in black.

center. The southern neighborhoods of Hahnwald (3.1) and Meschenich (3.2) represent residential areas in the outer suburbs of Cologne. These areas have a rural character reflected in the single-family detached homes with spacious gardens, lower building density, and open spaces with fields and meadows.

True digital orthophotos

The aerial images are provided by the open spatial data infrastructure GeobasisNRW (2023a) as tiles with dimensions of 1×1 km in the format of JPG2000.

Each tile has a resolution (ground sample distance) of 10 cm and an average position accuracy of 2 to 3 pixels (20–30 cm). Furthermore, they consist of four spectral channels, namely red, green, blue (RGB) and NIR, with a radiometric resolution of 8 bits and a temporal resolution of 2 years. The images underlie the projected coordinate system ETRS89/UTM32 (EPSG 25832).

The aerial images were processed to distortion-free and true to scale images called digital orthophotos (DOPs). In an additional step, DOPs were rectified to TrueDOPs by adjusting tilting objects, e.g., buildings (GeobasisNRW 2023b). Therefore, TrueDOPs allow a vertical view on the image by eliminating blind spots[1] while preserving the geometric and radiometric qualities of DOPs.

Ground truth labels

Below, a comprehensive overview of ground truth labels and their characteristics is given. In total, the data set contains around 12,508 PV panels spread over 171 buildings (manually counted). Table 1 indicates that the city center is representing an intermediate subarea in terms of PV panels per building and the mean building size. In general, in the commercial subarea, larger rooftops allow significantly larger PV systems than smaller rooftops in the suburbs. The position accuracy of the manually drawn ground truth labels is relatively high due to the basis of TrueDOPs for which the height of a building is not distorting the location of a PV panel.

Table 1. Overview of PV panels per subarea in the entire study area.

	Commercial	City center	Suburbs	Total
Buildings with PV panels	31	62	78	171
PV panels	7,994	2,431	2,083	12,508
Mean PV panels/building	258	39	26	73
Buildings (with PV panel) mean size (m^2)	1,364	418	140	410

A strong variation in PV system sizes affects a balanced presence of target class pixels in the label patches. Table 2 shows the mean percentage of pixel values labeled as PV panels per image patch of 256×256 pixels in each subarea. In aerial imagery, the small size of PV panel's results in a higher percentage of background pixels compared to PV panel pixels. This leads to an average imbalance between the target class and background pixels in all labeled patches. Furthermore, the table highlights a disparity in the average percentage of target class pixels among all subareas, particularly between the commercial (19.16%) and the other two subareas (5.38% and 3.98%). The effect amplifies for lower-resolution images of 20 cm when the patch dimensions of 256×256 pixels remain the same.

To gain a better understanding of the local characteristics in the subareas, the colors of rooftops and PV panels were documented for all collected ground truth labels. This information is summarized in Figure 2, which also provides details on the predominant colors and the ratio between blue and black PV panels per subarea. For instance, more than 60% of the commercial rooftops with a PV system installed are white, while there is a wide variation in rooftop colors in the suburbs. Additionally, the suburbs have the highest percentage of black PV panels. The rooftop color might have different effects on the detection of PV panels depending on the PV panel color. It needs to be mentioned that blueish PV panels tend to appear in light grey in TrueDOPs when exposed in directed orientation toward the sun as well as dark blueish when they are opposed to the sun. Reflections can vary depending on the horizontal and vertical angles between the sun, the PV panel, and the airborne camera capturing the images.

Figure 3 shows the variety of rooftop colors by providing one example image for each color detected in the subareas.

Generating DOP and ground truth label patches

As emphasized by Jiang et al. (2021), CNNs are sensitive to the patch size, highlighting the importance of finding the appropriate patch dimensions based on the spatial resolution and PV system size.

Table 2. Average percentage of pixels associated with PV panels per label patch for each subarea at 10 and 20 cm resolutions.

	Commercial (%)	City center (%)	Suburbs (%)
10 cm	19.16	5.38	3.98
20 cm	10.03	1.75	1.40

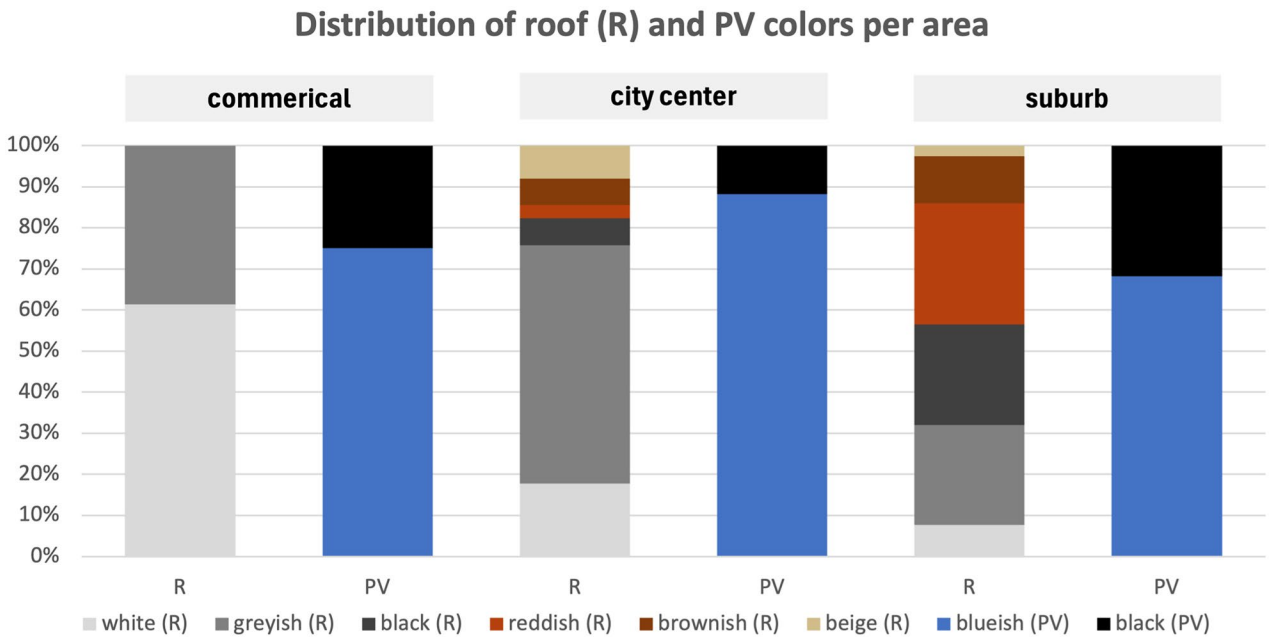

Figure 2. Distribution of roof colors (of rooftops with PV panel) and PV panel color ratio per subarea.

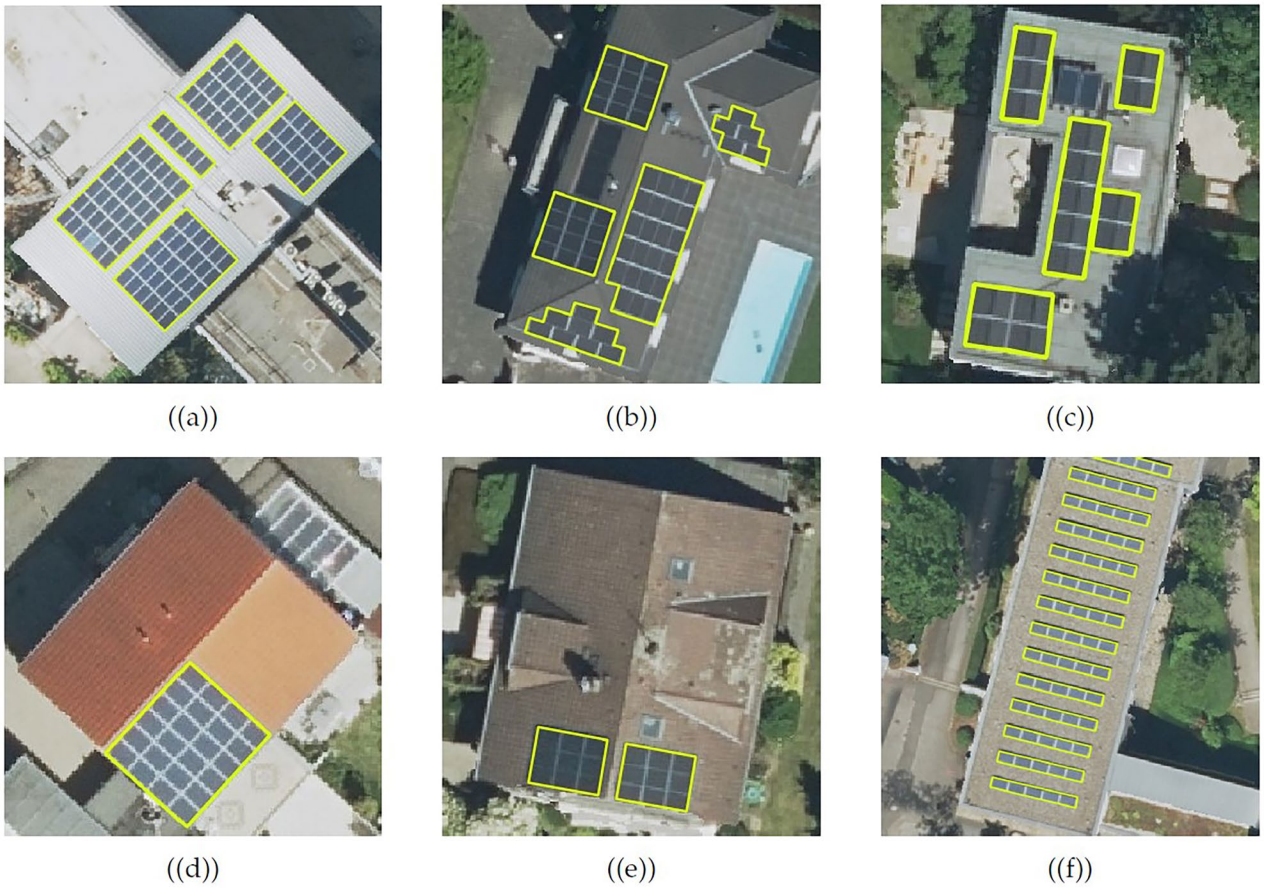

Figure 3. Rooftop color comparison: (a) white, (b) black, (c) grayish, (d) reddish, (e) brownish, (f) beige (sandy ground); PV systems outlined in yellow.

To find the appropriate patch dimensions, the footprint sizes of buildings (with PV system) incorporated in the classification were analyzed. Those footprint sizes ranged from 6 to 5,413 m^2 with a mean of 410 m^2. Consequently, having objects of around 20×20 m, an input dimension of 256×256 pixels at a resolution of 10 cm (25.6×25.6 m) seems appropriate to capture building rooftops. For that reason, all input data are tiled image patches of 256×256 pixels. Depending on the input resolution, the patches' sizes were 25.6×25.6 m (10 cm resolution) and 51.2×51.2 m (20 cm resolution). To consider only patches that contain target class pixels, only those patches were extracted that intersect with the ground truth data.

Data split

The concept of data splitting enables an unbiased evaluation of a network's performance. Evaluating the network's performance during and after the training process helps to choose the appropriate network hyperparameters. For instance, unbiased evaluation is crucial to prevent overfitting the network to the training data, making it unsuitable for predictions on an independent testing set. Therefore, splitting the data set into training, validation, and testing sets is necessary. Typically, most data is assigned to the training data set, which propagates through the network. An independent validation data set, not seen by the network before, is important to evaluate the network's performance during the training process. This allows for observation of the training progress based on an evaluation after each epoch. Finally, the network's performance can be evaluated through the prediction of the test dataset. The data set is divided according to a common split ratio of 70% for training data, 20% for validation data, and 10% for testing data (Castello et al. 2019; Da Costa et al. 2021; Kingma and Ba 2014).

However, adjacent patches covering the same PV system, which may be part of the training, validation, or test dataset, can lead to correlations in the accuracy assessment. The potential influence is addressed by preventing patches from overlapping and using data augmentation to reduce the impact of spatial similarities before feeding data into the network. Data augmentation is used to enrich the variety of patches to the model. The augmentation is based on horizontal and vertical flips of patches and has proven to enhance the robustness of models by mitigating the effect of overfitting (De Jong et al. 2020).

Modified U-Net architecture and hyperparameter definition

The employed U-Net architecture is a slightly modified version (see Figure 4) of the original architecture developed by Ronneberger et al. (2015). Its main differences lie in the additional batch normalization layer between the convolutional layer and the ReLU activation function. It aims to make the model more robust and to increase the computational speed of training by allowing higher learning rates (Ioffe and Szegedy 2015). Additionally, the dimensions of the feature maps do not decrease after convolutional operations due to implemented zero-padding operations. This allows for the input and output dimensions to be identical, simplifying the processing and analysis of the data. Therefore, the original copy and crop operations are replaced by a simple copy operation to reconstruct the same spatial resolution at each level of the U-Net.

Additionally, the U-Net is modified to accept patches of 256×256 pixels as input. This is an essential adjustment to meet the requirement for appropriate patch dimensions. The selected input dimensions meet the requirement of having tiles with even x- and y-dimensions, which allow seamless tiling after each max-pooling operation with a kernel size of 2×2 (Ronneberger et al. 2015). The third dimension, which defines the number of image channels, can be manually changed from 3 to 4 channels to either process RGB or RGB plus NIR images. Nevertheless, basic hyperparameters defined in the original U-Net architecture remain the same such as the ReLU activation function. The same applies to convolutional kernels keeping the size of 3×3 and a stride of 1, which was also utilized in similar studies (Castello et al. 2019; Malof et al. 2017). Similarly, the max-pooling operation with a kernel of 2×2 pixels and a stride of 2 is adapted. Further, the number of convolutional layers doubles with each max-pooling operation from 64 up to 1024 layers.

The weights are initialized with the He uniform variance scaling initializer (He et al. 2015). Transfer learning with pre-trained weights is not employed due to the NIR band, which is denoted as the fourth image channel, and most pre-trained weights are based on RGB channels only.

Accuracy assessment

The accuracy metric (1) is defined by the number of correctly predicted image pixels divided by the total number of predictions. Correctly predicted image

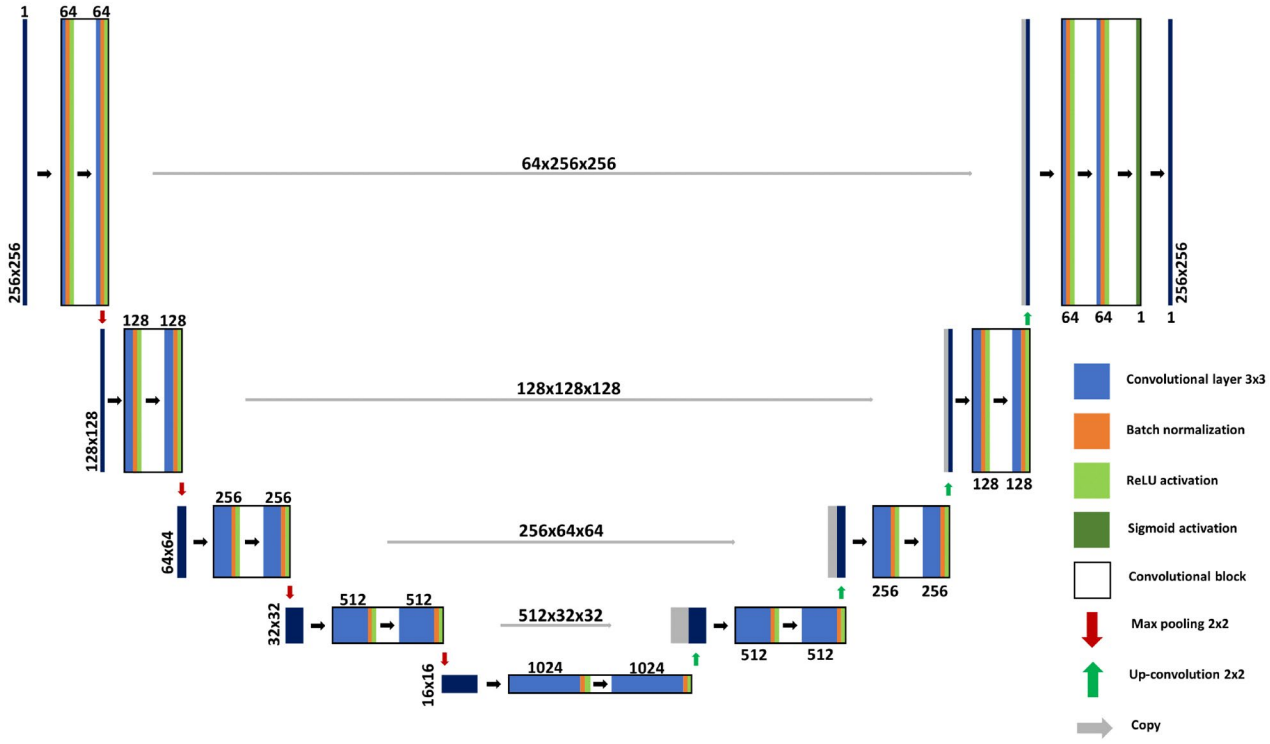

Figure 4. Modified U-Net architecture; dark blue boxes represent input or output features.

pixels include true positives (TP), in which PV panels are correctly identified as well as true negatives (TN) (absence of PV panels is correctly identified).

$$Accuracy = \left(TP+TN\right)/\left(TP+TN+FP+FN\right) \quad (1)$$

The precision metric (2) calculates the proportion of TP to the total number of actual PV panels, including TP and not identified PV panels or false positives (FP). To express the proportion of correctly identified PV panels to all predictions of PV panels, the recall metric (3) was applied. It is calculated by the number of TP divided by the number of TP and FN.

$$Precision = TP/\left(TP+FP\right) \quad (2)$$

$$Recall = TP/\left(TP+FN\right) \quad (3)$$

Further, there is the F1-score (4) which expresses the harmonic average of precision and recall. It computes the overlap between ground truth data and prediction and divides it by the total number of pixels.

$$\begin{aligned} F1 &= \left(2\times Precision\times Recall\right)/\left(Precision+Recall\right) \\ &= \left(2\times TP/2\times TP+FP+FN\right) \end{aligned} \quad (4)$$

The Jaccard index or intersection over union (IoU) (5) is a coefficient to measure the similarity between two samples. It is calculated by dividing the intersection between the label and prediction by the union of both samples. The IoU penalizes under- and oversegmentation more than the F1-score, which is based on the greater impact of FN and FP (Müller et al. 2022).

$$IoU = TP/TP+FP+FN \quad (5)$$

Training and testing experiments

The first experiment analyzes the impact of training a model on different subareas corresponding to different land use types. For this experiment, the U-Net is trained and tested on each subarea separately as well as on the combined subareas. Further, cross-validations across various land use types are carried out by assessing the performance of each subarea model using the respective test data set from all other subareas. Further experiments are conducted to analyze the impact of including NIR data in the training process. The last experiment explores different input resolutions.

Figure 5 below summarizes the framework and method implementation of this research.

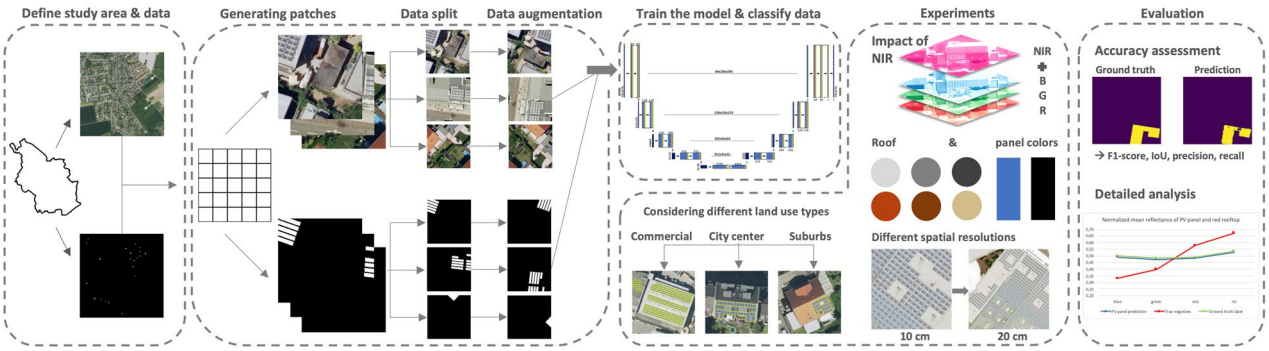

Figure 5. Framework for PV panel detection in aerial imagery using a U-Net and considering the following factors: land use types, NIR, roof and panel color, and spatial resolutions.

Table 3. Evaluation of U-Net based on Focal Loss (FL) and BCE with different learning rates using RGB TrueDOPs of the city center (**highest scores in bold**); Epochs = 100.

loss	LR	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	IoU (%)
BCE	1e-2	96.91	84.74	22.63	35.72	59.31
BCE	1e-3	98.11	90.49	55.93	69.13	75.45
BCE	1e-4	**99.23**	93.97	85.10	**89.31**	**89.95**
FL	1e-2	96.21	0.00	0.00	0.00	48.10
FL	1e-3	98.04	**98.39**	49.17	65.57	73.39
FL	1e-4	98.90	84.57	**86.76**	85.65	86.88

Table 4. Classification results of each subarea and all areas combined (**highest scores in bold**).

Area	Precision (%)	Recall (%)	F1-score (%)	IoU (%)
Commercial	89.40	**91.50**	**90.44**	88.96
City center	89.10	85.59	87.31	88.25
Suburbs	**97.86**	60.66	74.89	78.96
All areas	91.64	88.74	90.16	**90.36**

Results and analysis

Optimal loss function and learning rate

The imbalance between target class pixels and background information presented in Table 2 results in uncertainty about the appropriate loss function to implement. Based on the results of Table 3 showing the best U-Net performance for the combination of the binary cross-entropy (BCE) loss function and a learning rate of 0.0001, both hyperparameters are determined. Furthermore, it becomes evident that the accuracy metric is not suitable for assessing the performance of the model considering the disparities between precision and recall while having high accuracies.

Given the relatively limited number of 100 patches per subarea, a batch size of 5 is selected. Further, a fixed number of 60 epochs is picked based on an observation of the model's performance. This observation reveals that the validation loss begins to stabilize at 60 epochs while the training loss continues to increase (indicating overfitting).

Quantitative evaluation of RGB classifications

Considering the F1-score and the recall the best performance is achieved by the U-Net trained on patches of the commercial subarea (see Table 4). Both models trained on commercial areas and all subareas show the most consistent performance scores around 90% (± 1.7%). The model based on city center patches achieved approximately high and constant results between 85 and 90%. Except for achieving the highest precision score, the model trained on suburb patches yields the poorest scores in the recall, F1-score, and IoU. Consequently, the model tends to predict PV panels only at those locations where PV panels are actually installed while it is prone to omit PV panels in the classification.

It becomes evident that subareas with a higher average of target pixels per patch achieve better F1-scores and start to stabilize at earlier stages in the training process (see Figure 6). This is the result of larger PV system sizes in commercial areas, which provide uniform patterns of long panel arrays and homogeneous colors to learn by the U-Net.

Visual evaluation of RGB classifications

Figure 7a and b reveal difficulties in the prediction of the darker PV panels in the city center, similar to Figure 7d showing a PV system with a homogeneous black surface, installed on a dark rooftop in the suburbs. Contrary Figure 7c shows a rather successful classification of a black PV system, which is accentuated due to the sharp contrast to the light grey rooftop facilitating the detection of almost all PV panels. Further, the presence of distinctive PV frames likely facilitates the detection of PV panels, as CNNs learn patterns in imagery. Most PV systems analyzed in this study have silvery white aluminum frames with rectangular shapes, as shown in Figure 7a, b, and d.

Figure 6. U-Net's training and validation performance according to F1-score per subarea: Commercial (left), city center (middle), and suburbs (right).

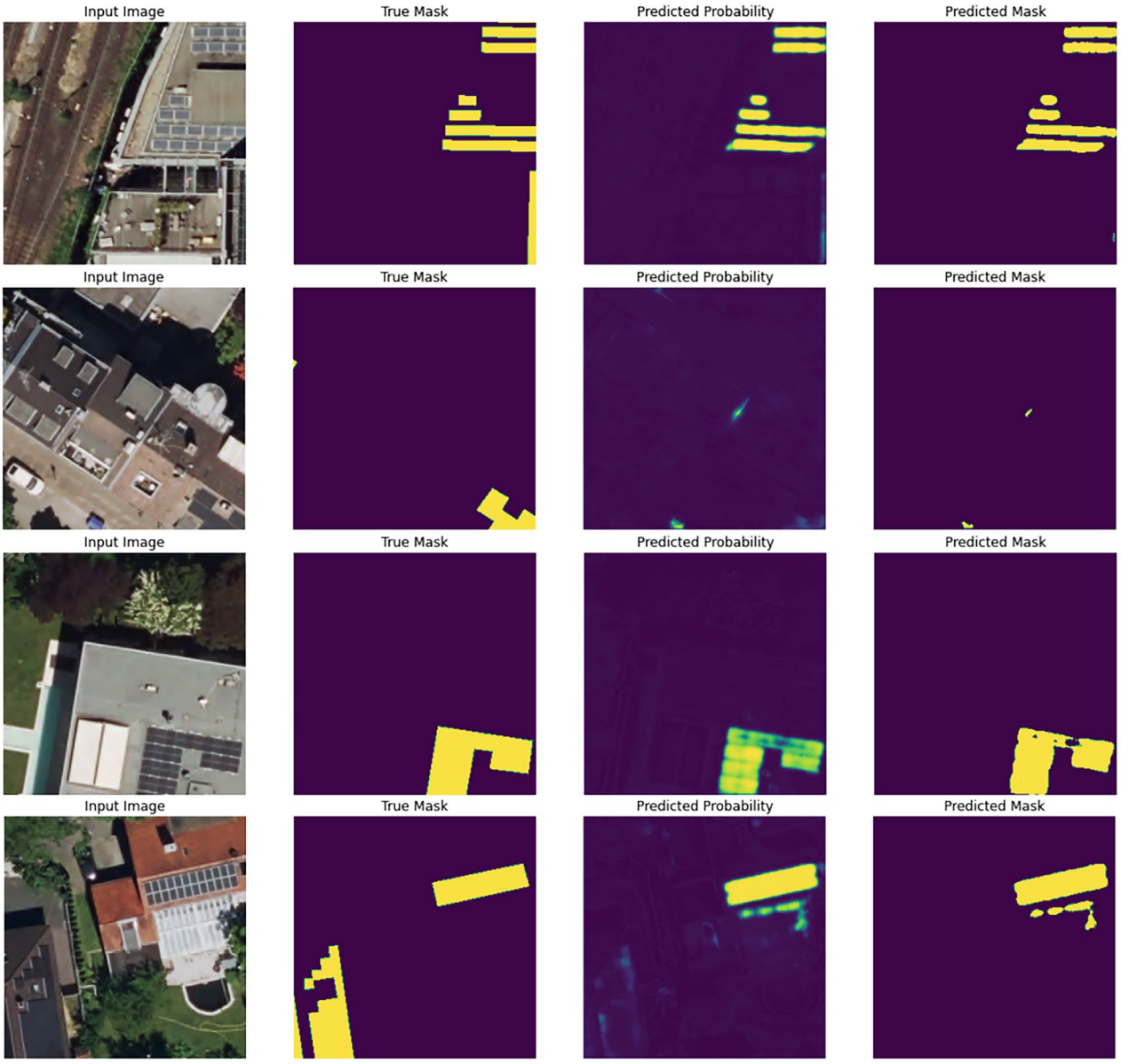

Figure 7. Top-down (a)–(d): (a) and (b) Testing sample located in the city center; (c) and (d) Testing sample located in suburbs.

PV systems without such frames are also present in Figure 7b and d, where the model cannot detect black PV panels. It can be assumed that frameless PV panels may hinder their detection, particularly when they have low contrast with the rooftop.

In contrast, the presence of these characteristics at different objects can lead to FP predictions. This effect

can be observed in Figure 7d, where the white rectangular structure of a conservatory caused a misclassification. Additionally, in Figure 7b, two more FP instances were likely caused by a combination of reflectance and structural patterns. The first misclassification involved a single-axle trailer on the road, covered with a blue canvas, which was nearly the same size as a PV panel. Secondly, there is a FP prediction in the center of the image. It can be concluded that the classification is not solely determined by the reflectance, as the spot only covers a fraction of a homogeneous rooftop in terms of colors. The prediction is likely based on the combination of the dark roof color (of the shadowed northeast-orientated side) and the pattern of parallel bright edges from a dormer and the boundary to the next rooftop. This pattern corresponds to the silvery white frame that typically bounds PV panels. Although the same pattern can also be found on the other side of the rooftop, it appears to have a different reflectance due to its sunlit southwest orientation. However, the presented FN predictions are all located at the patches' edges. To analyze whether this could be an artifact (a systematic anomaly produced by the model), all FN predictions are overlapped within one patch (see Figure 8). This overlap creates a heat map of errors, so systematic errors can be identified by their location in the patch. The heat map shows one hotspot of errors in the center-left and one hotspot in the lower right corner of the patch. Overall, it proves that there is no artifact at the patches' edges, which reinforces the previous assumption that the model has difficulties in detecting black PV panels without bright frames, that are located on dark rooftops.

Cross-validation across land use types

Cross-validation is used to assess the suitability of different subareas for training a model capable of classifying other subareas. It helps to identify the most suitable training subarea, which can also be considered as an intermediate subarea comprising key features of different subareas. Each model is expected to perform best on the subarea on which it was trained (see Table 5). Therefore, this analysis focuses on model performances outside their training areas.

Overall, training images of the city center suit best for classifying commercial areas or suburbs. Vice versa, the model trained on images of the commercial area achieves the best classification of the city center. The poorest F1-score result of less than 50% is achieved by a model trained on suburb images classifying images of commercial areas. The results demonstrate the greatest discrepancies between commercial areas and suburbs while showing that a model trained on images of the city center serves as the best

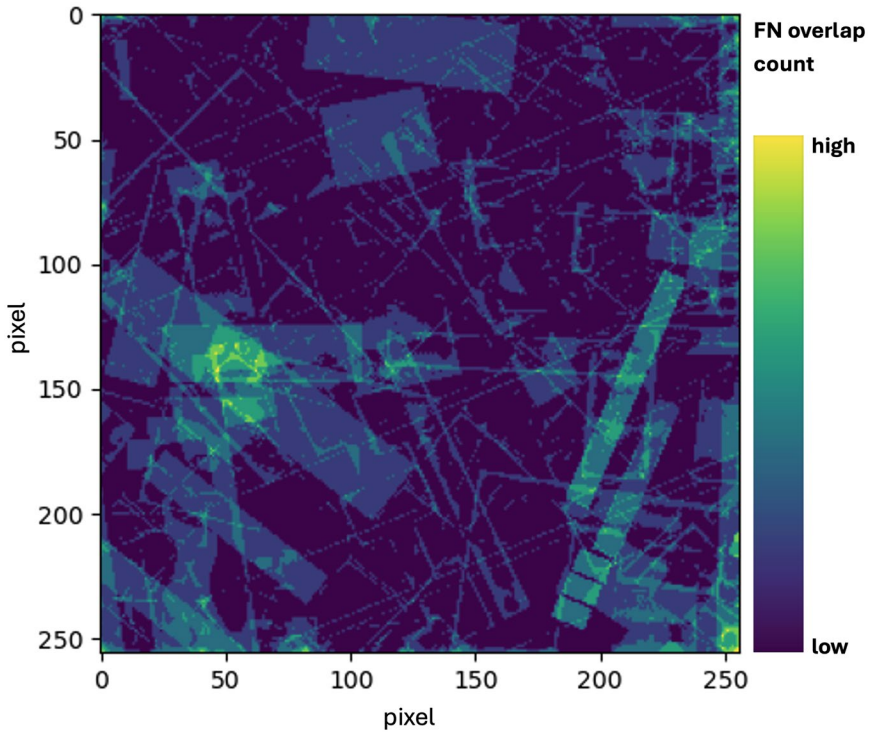

Figure 8. Heat map of all FN predictions computed from 56 RGB testing images of all areas and each subarea (from purple = no FN, to yellow = multiple FN).

Table 5. F1-scores of cross predictions (**highest scores in bold**).

Trained/predicted	Commercial (%)	City center (%)	Suburbs (%)
Commercial	(90.44)	**72.89**	**61.85**
City center	**59.82**	(87.31)	**77.73**
Suburbs	**48.52**	63.49	(74.89)

Table 6. Evaluation of RGB-NIR image classification (**highest scores in bold**).

Area	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
Commercial	93.91	**84.07**	**88.72**	87.35
City center	92.07	83.41	87.53	88.45
Suburbs	**96.81**	52.65	68.21	74.71
All areas	94.06	**89.55**	**91.75**	**91.81**

all-rounder for classifying different land use areas. The discrepancy is also reflected by their differences in terms of predominant roof colors, variations of roof color, the number of PV panels installed, and the mean size of buildings with PV systems.

Classification based on TrueDOPs including NIR data

Both training and evaluation are conducted based on the same patches used in previous sections. In addition to the 3-channel RGB images, the NIR channel is included, resulting in four-channel images.

Despite the high precision of U-Nets trained on images of the suburbs, the best classification performance is achieved by a model trained on images of all subareas (see Table 6). Similar to the previous results (see Table 4) the suburb model obtains the greatest gap of around 44% between precision and recall scores.

The comparison of classification scores between RGB and RGB-NIR-based models (see Figure 9) shows a performance drop when the NIR channel is included in the classification of suburbs. The classification of RGB-NIR images of the commercial area depicts a minor performance drop, while the performance in the city center remains relatively constant with a barely noticeable increase. Despite the predominant negative impact on performance, the model trained on images from all areas shows an overall increase in performance.

Figures 10 and 11 provide a detailed analysis on PV panel detections and omissions, including the potential impact of NIR data. The analysis compares normalized mean reflectance (MR) of PV panel predictions, ground truth labels, FN, and TN. The MR values are calculated solely using the pixels within the rooftops' outline. The black and bluish PV systems have a consistent reflectance in the RGB spectrum, with only a slightly higher reflectance of blue light. The uniform pattern of the bluish PV panels in Figure 11 is due to their grayish appearance in direct sunlight. In contrast, the red rooftop has a diverse reflectance pattern due to the absorption of blue light, the higher reflectance of red light, and an even stronger reflectance of NIR. Overall, the PV system is easily distinguishable from the rooftop due to the rectangular pattern of frames and the difference in reflectance. The opposite effect is visible in Figure 10, showing the failed detection of black PV panels on a dark rooftop. Similarly to the grayish blue PV panels, the most notable difference between rooftop and PV panels lies in the reflectance of NIR. The NIR reflectance of the dark rooftop is stronger than the RGB reflectance, however, the discrepancy for black PV panels is even greater.

Classification of lower-resolution TrueDOPs

The classification scores in Table 7 show that the commercial area model performs better with lower-resolution images than models trained on city center or suburb images. Most notable is the gap between precision and recall of city center and suburb classifications. It indicates a low number of FP predictions, but an even higher number of FN predictions represented by missing PV panel predictions.

Despite the constant classification scores of the commercial area, Figure 12 demonstrates performance drops for lower-resolution images, particularly in the case of the city center and suburb models.

Discussion

The study focuses on investigating the impact of different land use types, the addition of NIR data to aerial images, the correlation between roof and panel color, and the sensitivity of the U-Net models toward lower resolutions with regard to the panel size. It highlights how heterogeneous environments and varying building characteristics can impact model predictions. Additionally, the impact of different image resolutions on the model's performance is explored, showing that decreasing resolution affects the model's ability to differentiate between small objects. Further, the study introduces a novel use of NIR data in aerial imagery and its potential for improving PV panel detections. Although the impact is not evaluated as extensively as other factors, the results indicate mixed effects on model performance.

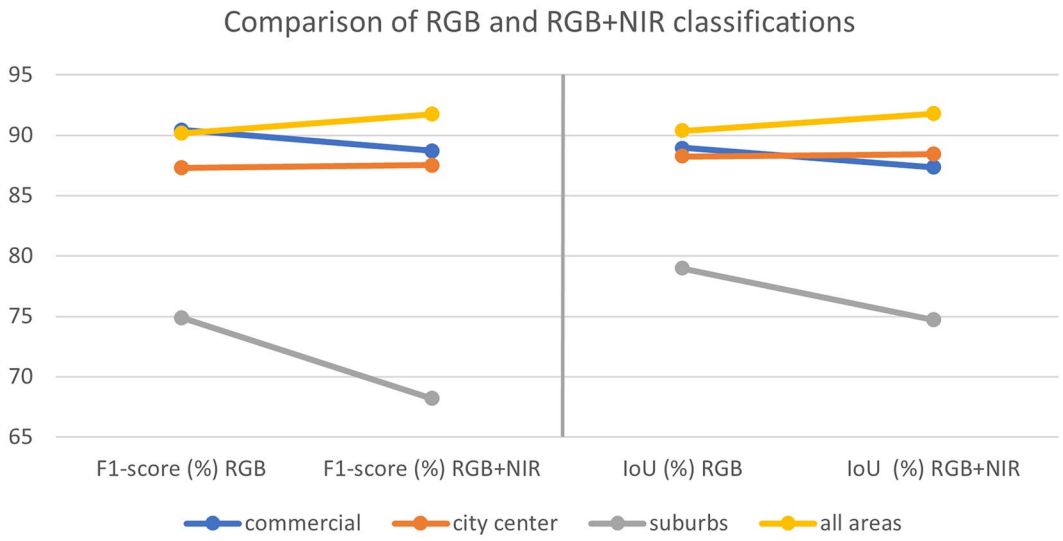

Figure 9. Comparison of RGB and RGB-NIR-based classifications assessed with F1-score.

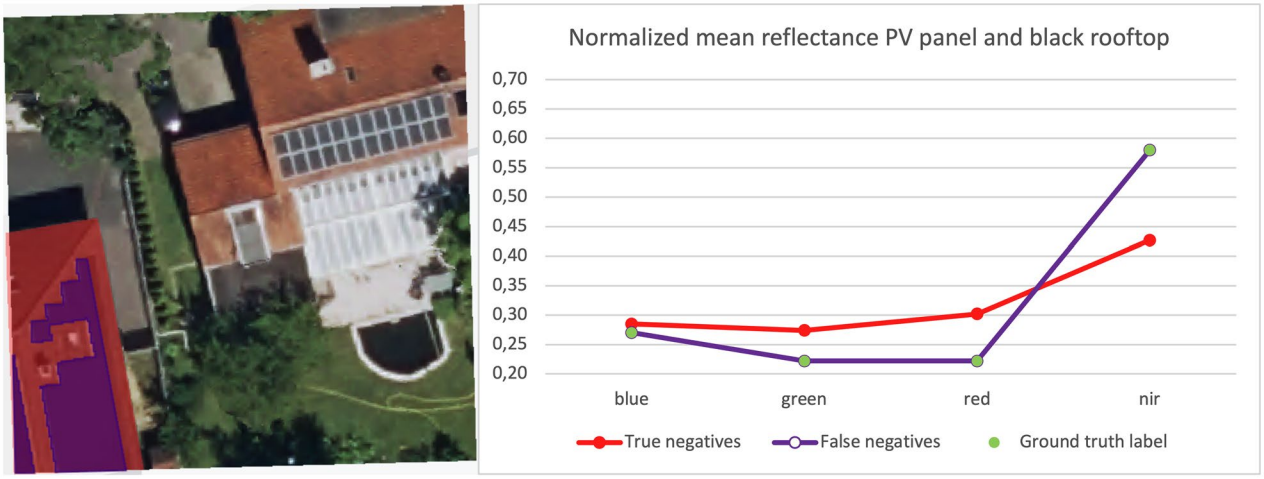

Figure 10. Mean reflectance values of black PV panel and black rooftop.

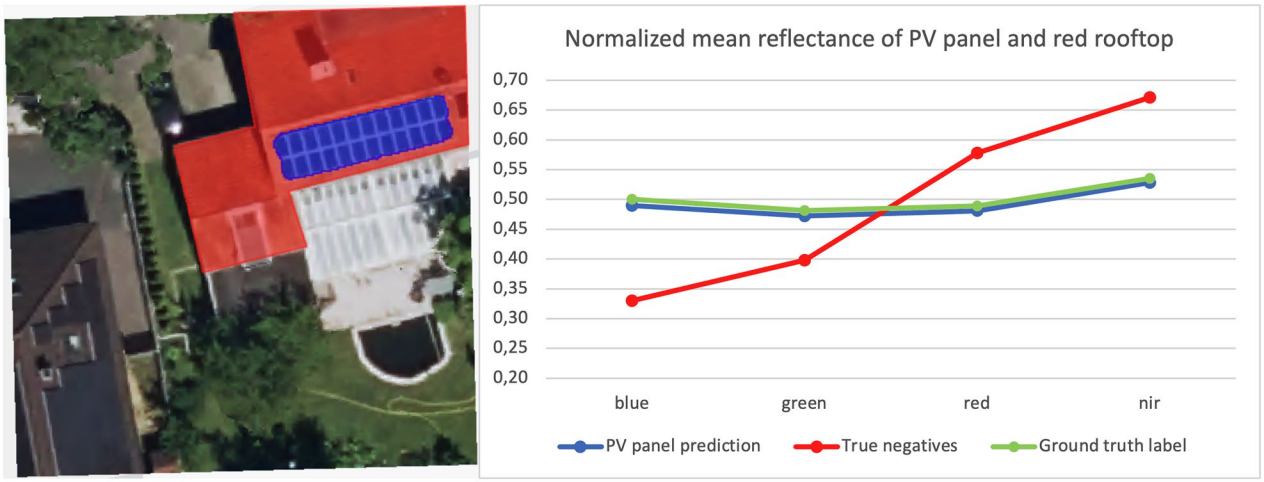

Figure 11. Mean reflectance values of bluish PV panels and red rooftop.

The comparison of quantitative results between this model and models that were trained on millions of images (De Jong et al. 2020; Malof et al. 2017) should be done cautiously. However, similar trends can be observed in related projects. In summary, the performance of the model trained on RGB images (10 cm) is quantified by an F1-score of 90.16% and an IoU of 90.36%. The results of subareas vary between an F1-score of 74.89% for the suburbs and 90.44% for the commercial area and IoU scores between 78.96% and 88.96% for the suburbs (lower score) and the commercial area (higher score), respectively. In comparison, this performance is significantly better than the U-Net's performance of Castello et al. (2019) achieving an F1-score of 80% and an IoU of 64%. Nevertheless, their model was applied to a greater variety of urban and rural settings from different regions. In contrast, Da Costa et al. (2021) achieved a better performance (F1-score: 95.38%; IoU: 91.17%), given that their project focuses solely on one type of PV system, namely large-scale solar plants.

The research's significance is evident in its contributions to scientific knowledge in several areas. The results highlight the impact of diverse land use types on PV panel detection accuracy, contributing novel insights into the influence of urban and architectural variations within a city. This expands the understanding of data-driven approaches for CNN-based PV panel detection. For instance, the results prove that the proportion of target class pixels per image patch is of great importance for the loss function and the learning rate. Both hyperparameters are required to be selected with care regarding the area characteristics. It turned out that the combination of the BCE and a learning rate of 0.0001 works best for the area-specific data sets, in comparison to the FL. Furthermore, this research reflects the large-scale cross-validations conducted in the DeepSolaris project by De Jong et al. (2020) on a local level. In the DeepSolaris project, the model's performance remained constant when the training area was nearby the validation area, while performance drops of the model were noticed for cross-validations within an entire state as well as in a cross-border context. Similarly, this work proved that differences in architectural and urban characteristics can already have an impact on the model's predictions within a city. In particular, this is the case between the suburbs and commercial areas. Another similarity in related research concerns the recall scores. The scores are lower than the precision scores in all experiments, except for the commercial area, which achieved the highest recall of 91.5%. Da Costa et al. (2021) received a similar gap, with a recall score of 93.1% and a precision score of 88.5%. Considering the large-scale solar plants in their work, it can be assumed that this gap was determined by the size of the PV systems (in commercial areas) in the images, resulting in fewer target class pixels being omitted. To overcome these challenges, the patch sizes can be carefully chosen to account for the specific characteristics of each land use type and its PV system sizes. For instance, smaller PV systems in

Table 7. Evaluation of U-Net based on RGB TrueDOPs at 20 cm resolution (**highest scores in bold**).

Area	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
Commercial	87.29	**85.17**	**86.22**	**86.89**
City center	**93.12**	12.47	22.00	55.40
Suburbs	85.46	28.12	42.32	62.89
All areas	77.09	62.09	68.78	75.04

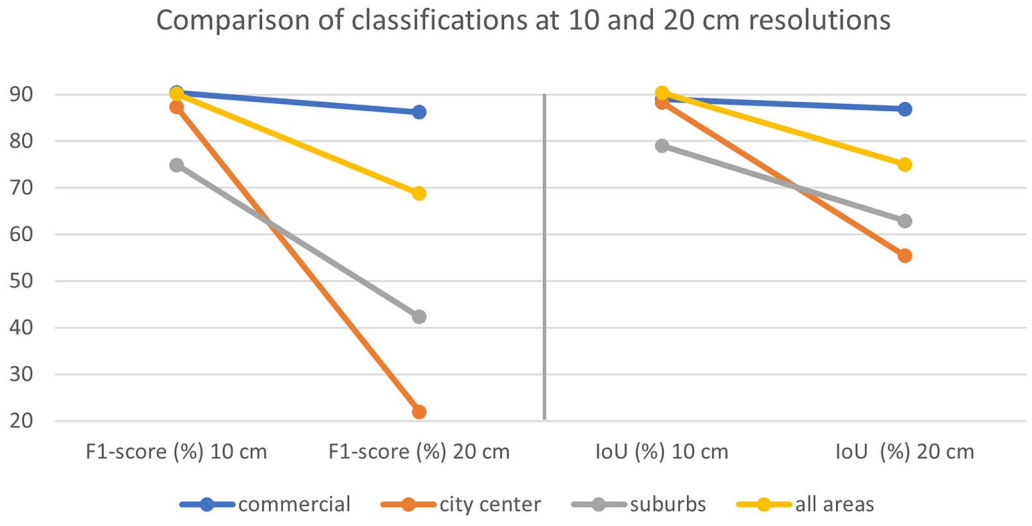

Figure 12. Comparison of classifications at 10 and 20 cm resolutions assessed with F1-score.

the suburbs require smaller patch sizes to address a potential class imbalance by capturing more detail. Conversely, larger patch sizes may be more appropriate in commercial areas where PV systems are typically larger. Overall, tailoring the input data dimensions to these unique characteristics can potentially improve the model's ability to accurately detect PV panels. Also, the use of additional data augmentation techniques could enrich the dataset and improve the model's generalization capabilities, as augmenting the data exposes the model to a wider range of scenarios and helps it learn more robust and adaptive representations. This could be particularly beneficial in the suburbs, where target class pixels are underrepresented.

The research provides a nuanced understanding of how rooftop colors correlate with the detection of PV panels. The results draw attention to the significance of contrast between PV panels and rooftops for detecting PV panels. The relevant research does not provide a detailed exploration of the impact of rooftop colors on the detection of PV panels in aerial imagery, making it difficult to compare research results. A rather general observation concerns a higher precision than recall score. The precision scores of classifications at 10 cm resolution vary between 89.1% and 97.86%. The gap between both scores is reflected by more FN than FP predictions, meaning that most PV panel classifications are correct, while few PV panels are not detected at all. A comparable discrepancy is observed in the results of Malof et al. (2017) achieving a recall of 80% and precision of 95%. The results of this research demonstrated that heterogeneous rooftops and PV systems in terms of rooftop sizes, shapes, and colors, as well as PV panel types, cause more false negatives which affect the recall score.

An essential part of this study is exploring the integration of NIR data in the semantic segmentation of PV systems. It is important to note that the impact of the NIR channel on detecting PV panels cannot be set in the context of related research as Da Costa et al. (2021) did not evaluate the impact of the NIR band in their study. Moreover, the research has revealed that the use of the NIR channel in openly available aerial imagery is a novel approach for detecting PV panels, as no additional studies have been found to investigate this method. However, the results indicate mixed effects. Minor improvements and decreases in the detection rate were noticed in the images of commercial areas and the city center, as well as in all images combined. In the case of the suburbs, the NIR rather caused a performance drop of the model than an improvement of the performance. Nevertheless, the detailed analysis proved the potential using NIR data to distinguish between black PV panels and dark rooftops.

The findings about how changing the image resolution affects the detection accuracy emphasize the importance of maintaining a balanced relationship between image dimensions, resolution, and the sizes of target objects. It became clear that the number of epochs significantly depends on the characteristics of the ground truth data, but also on the size of the PV systems, in proportion to the image resolution and the image dimensions. Latter defines the proportion of target class pixels per image, which varies depending on different land use types. As indicated by De Jong et al. (2020), the minor resolution difference between 10 and 20 cm affects the model's ability to differentiate small objects, such as skylights, from PV panels. Similar results were obtained when the spatial resolution was decreased to 20 cm while keeping the same hyperparameters. In addition to false positives that misclassified skylights or glass roofs with PV panels, the impact of heterogeneous environments (e.g., various types of urban objects, building shapes, and heights) on the classification became evident. This impact is reflected by significant performance drops of models that classified images of the city center.

By contextualizing the findings within the scope of related research, this study contributes to the broader understanding of CNN-based semantic segmentations for PV panel detection. Overall, the results of this study prove the importance of defining the ground truth data set and hyperparameters to match area-specific characteristics and PV systems, while offering insights into land use effects, rooftop colors, the utilization of NIR data, and the impact of different image resolutions.

The limitations of the results can be outlined by the generation and availability of ground truth data. The quality of the manually collected ground truth data depends on the annotator's ability to accurately identify PV panels. The selective annotation strategy, aiming to include only confidently identified PV panels, could result in an underrepresentation of poor or ambiguous instances. This limitation could potentially be addressed with accurate PV panel location information, allowing for more comprehensive ground truth data. However, this might impact the precision score by potentially introducing more false positives. Further, the constrained data availability should be taken into account when interpreting the results. The limited amount of training and validation data, coupled with small batch sizes, can lead to fluctuations in the learning process and longer training to reach

the point of convergence. This scarcity of data affects the model's training and may require more epochs for adequate convergence.

Future research directions and recommendations can be proposed based on the limitations and findings outlined above. To improve the classification of PV panels, utilizing additional data sources is suggested. For example, incorporating height data or building footprints could help focus on rooftops and reduce confusion with other urban objects. The effectiveness of integrating thermal infrared imagery in the semantic segmentation of PV panels within complex surroundings was demonstrated by recent studies of Wang et al. (2019) and Buerhop et al. (2022). Therefore, integrating georeferenced thermal imagery as an extra image channel could enhance the accuracy of classifying PV panels in urban areas. While the focus of this study is on PV panel detection using binary classification, it is recommended to consider a multi-class classification approach. This would involve incorporating ground truth data of both PV panels and STC to enable more precise learning and differentiation between these two types of panels. To improve the detection of black PV panels on dark rooftops, it may be helpful to split the class of PV panels into bluish and black panels. This will allow the model to be more precise for each type of panel. Collecting ground truth data is time-consuming and relies heavily on annotator expertise. To overcome this limitation, leveraging synthetic data generated by artificial intelligence (AI) can be beneficial. As previous research by Kriese et al. (2022) and Liu et al. (2020) has shown, adding synthetic data to the training set can enhance the performance of CNNs. Using AI-generated synthetic data could enable training CNNs on bigger and more diverse data sets of PV panels.

Conclusion

This research demonstrates various aspects emerging from different land use types that need to be considered when compiling an appropriate training data set. Commercial areas stand out due to their homogeneity in terms of low variations in rooftop characteristics, such as roof colors, slopes, and sizes. This homogeneity in combination with predominantly large PV systems facilitates the training of the model. However, the opposite effect emerges from residential areas in the suburbs having small PV systems installed on flat or pitched roofs with up to 6 different roof colors resulting in a low recall score caused by falsely classified PV panels indicating a class imbalance. Since a city center represents a mix of commercial and residential characteristics, it is most suitable as a training area for a model predicting PV panels in commercial areas and suburbs.

Regardless of the land use type, the results indicate that models are prone to failure when black PV panels are installed on dark rooftops. In particular, this applies to black PV panels without a bright frame, which affects their detectability since they compose a continuous surface rather than clear patterns that are easier to recognize for CNNs.

Adding a NIR channel to RGB imagery has indicated different effects on the detection of PV panels in different areas. While there is less to no effect on the detection process in the city center, there is only a marginal performance drop in the case of the commercial area and a slight increase in detection performance of all areas combined. However, there are negative impacts on the detection rate of PV panels in the suburbs when using the NIR channel. Nevertheless, the results demonstrate the potential of using NIR data to enhance the detection of PV panels with low contrast to their surroundings.

The DL model is particularly sensitive to lower-resolution images of areas in which relatively small PV systems are located. Decreasing the spatial resolution and decreasing the size of the PV panels in the image patches causes a greater imbalance between target class pixels and background information, making it more challenging for the model to learn and increasing the risk of misclassification.

The study enhances scientific knowledge by explaining how land use types, their architectural characteristics, rooftop colors, NIR integration, and resolution changes affect semantic segmentations of PV panels using CNNs. Furthermore, adaption of hyperparameter in association with these urban characteristics will improve the classification performance. In addition to the research implications, this knowledge can be valuable for potential applications in urban planning or energy policy by supporting the assessment of renewable energy adaptation at different administrative levels. In addition, mapping the exact location and size of PV systems can help validate and enrich existing databases.

Future research should explore additional input data, such as height data or thermal infrared imagery, and consider utilizing synthetic data generation to enrich the training data. Multi-class classification of PV panels should also be considered. Further, research in the direction of vision transformers has shown promising results in various computer vision tasks

and may offer novel insights for PV panel detection in high resolution aerial imagery. Similarly, the exploration of state-of-the-art unsupervised segmentation models could support the labeling of ground truth data.

Note

1. Refers to not visible areas in the image due to buildings appearing tilted in the image.

Acknowledgments

This study is based on my master's thesis supervised by Azarakhsh Rafiee and Stef Lhermitte. I want to thank my co-reader Roderik Lindenbergh for his valuable feedback on my thesis.

Disclosure statement

No conflict of interest was reported by the author(s).

ORCID

Simon Pena Pereira http://orcid.org/0009-0006-1158-9975
Azarakhsh Rafiee http://orcid.org/0000-0002-2609-2674
Stef Lhermitte http://orcid.org/0000-0002-1622-0177

References

Buerhop, C., Bommes, L., Schlipf, J., Pickel, T., Fladung, A., and Peters, I.M. 2022. "Infrared imaging of photovoltaic modules: A review of the state of the art and future challenges facing gigawatt photovoltaic power stations." *Progress in Energy*, Vol. 4(No. 4): 1. doi:10.1088/2516-1083/ac890b.

Castello, R., Roquette, S., Esguerra, M., Guerra, A., and Scartezzini, J.-L. 2019. "Deep learning in the built environment: Automatic detection of rooftop solar panels using convolutional neural networks." *Journal of Physics: Conference Series*, Vol. 1343(No. 1): 012034. doi:10.1088/1742-6596/1343/1/012034.

ClimateWatch. 2022. "Climate watch historical country greenhouse gas emissions data." Last modified January 04, 2023, https://www.climatewatchdata.org/ghg-emissions?end_year=2019&start_year=1990.

Da Costa, M.V.C.V., de Carvalho, O.L.F., Orlandi, A.G., Hirata, I., de Albuquerque, A.O., Silva, F.V.e., Guimarães, R.F., Gomes, R.A.T., and de Carvalho Júnior, O.A. 2021. "Remote sensing for monitoring photovoltaic solar plants in Brazil using deep semantic segmentation." *Energies*, Vol. 14(No. 10): 2960. doi:10.3390/en14102960.

De Jong, T., Bromuri, S., Chang, X., Debusschere, M., Rosenski, N., Schartner, C., Strauch, K., Boehmer, M., and Curier, L. 2020. "Monitoring spatial sustainable development: Semiautomated analysis of satellite and aerial images for energy transition and sustainability indicators." *CoRR*, abs/2009.05738.

EC. 2022. "Solar energy." Last modified January 04, 2023, https://energy.ec.europa.eu/topics/renewable-energy/solar-energy_en.

EUPD. 2023. "FEHLERHAFTE REGISTRIERUNGEN VON BALKON-PV-ANLAGEN ERSCHWEREN DETAILLIERTE ANALYSE DES PHOTOVOLTAIK-ZUBAUS IN DEUTSCHLAND." Last modified February 11, 2024, https://www.eupd-research.com/fehlerhafte-registrierungen-von-balkon-pv-anlagen-erschweren-detaillierte-analyse-des-photovoltaik-zubaus-in-deutschland/.

Feng, Q., Niu, B., Ren, Y., Su, S., Wang, J., Shi, H., Yang, J., and Han, M. 2024. "A 10-m national-scale map of ground-mounted photovoltaic power stations in China of 2020." *Scientific Data*, Vol. 11(No. 1): 198. doi:10.1038/s41597-024-02994-x.

GeobasisNRW. 2023a. "Digitale orthophotos (10-fache kompression) – paketierung: Einzelkacheln." Last modified January 11, 2023, https://www.opengeodata.nrw.de/produkte/geobasis/lusat/dop/dop_jp2_f10/.

GeobasisNRW. 2023b. "Digitale orthophotos." Last modified August 22, 2023, https://www.bezreg-koeln.nrw.de/geobasis-nrw/produkte-und-dienste/luftbild-und-satellitenbildinformationen/aktuelle-luftbild-und-0.

Gui, S., Song, S., Qin, R., and Tang, Y. 2024. "Remote sensing object detection in the deep learning era—A review." *Remote Sensing*, Vol. 16(No. 2): 327. doi:10.3390/rs16020327.

He, K., Zhang, X., Ren, S., and Sun, J. 2015. "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification." 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Santiago, Chile, 1026–18. doi:10.1109/ICCV.2015.123.

Ioffe, S., and Szegedy, C. 2015. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 448–456.

Jiang, H., Yao, L., Lu, N., Qin, J., Liu, T., Liu, Y., and Zhou, C. 2021. "Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery." *Earth System Science Data*, Vol. 13(No. 11): 5389–5401. doi:10.5194/essd-13-5389-2021.

Kingma, D. P., and Ba, J. L. 2014. "Adam: A method for stochastic optimization." 3rd International Conference on Learning Representations, ICLR 2015 – Conference Track Proceedings, San Diego, CA, USA. doi:10.48550/arxiv.1412.6980.

Kriese, J., Hoeser, T., Asam, S., Kacic, P., Da Ponte, E., and Gessner, U. 2022. "Deep learning on synthetic data enables the automatic identification of deficient forested windbreaks in the paraguayan chaco." *Remote Sensing*, Vol. 14(No. 17): 4327. doi:10.3390/rs14174327.

Liu, W., Liu, J., and Luo, B. 2020. "Can synthetic data improve object detection results for remote sensing images?" doi:10.48550/arXiv.2006.05015.

Malof, J. M., Hou, R., Collins, L. M., Bradbury, K., and Newell, R. 2015. "Automatic solar photovoltaic panel detection in satellite imagery." 2015 International Conference on Renewable Energy Research and Applications (ICRERA), Palermo, 1428–1431. doi:10.1109/ICRERA.2015.7418643.

Malof, J. M., Bradbury, K., Collins, L. M., Newell, R., Serrano, A., Wu, H., and Keene, S. 2016. "Image features for pixel-wise detection of solar photovoltaic arrays in aerial imagery using a random forest classifier." 2016

IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, UK, 799–803. doi:10.1109/ICRERA.2016.7884446.

Malof, J. M., Collins, L. M., and Bradbury, K. 2017. "A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery." 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Fort Worth, TX, USA, 874–877. doi:10.1109/IGARSS.2017.8127092.

MaStR. 2023. "Zielsetzung des Marktstammdatenregister." Last modified January 05, 2023, https://www.marktstammdatenregister.de/MaStRHilfe/subpages/GrundlagenZielsetzung.html.

Müller, D., Soto-Rey, I., and Kramer, F. 2022. "Towards a guideline for evaluation metrics in medical image segmentation." *BMC Research Notes*, Vol. 15(No. 1): 210. doi:10.48550/ARXIV.2202.05273.

Puttemans, S., Van Ranst, W., and Goedemé, T. 2016. "Detection of photovoltaic installations in RGB aerial imaging: A comparative study." GEOBIA 2016: Solutions and Synergies, Enschede, The Netherlands. doi:10.3990/2.429.

Rausch, B., Mayer, K., Arlt, M., Gust, G., Staudt, P., Weinhardt, C., Neumann, D., and Rajagopal, R. 2020. "An enriched automated PV registry: Combining image recognition and 3d building data." *CoRR*, abs/2012.03690.

Ronneberger, O., Fischer, P., and Brox, T. 2015. "U-net: Convolutional networks for biomedical image segmentation." *CoRR*, abs/1505.04597.

UN. 2022. "What is renewable energy?" Last modified January 04, 2023, https://www.un.org/en/climatechange/what-is-renewable-energy.

UNFCCC. 2015. "Paris agreement." Paris Climate Change Conference, Paris, France.

Wang, J. L., Li, A. Y., Huang, M., Ibrahim, A. K., Zhuang, H., and Ali, A. M. 2019. "Classification of white blood cells with PatternNet-fused Ensemble of Convolutional Neural Networks (PECNN)." 2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018, Louisville, KY, USA, 325–330. doi:10.1109/ISSPIT.2018.8642630.

Wang, M., Cui, Q., Sun, Y., and Wang, Q. 2018. "Photovoltaic panel extraction from very high-resolution aerial imagery using region-line primitive association analysis and template matching." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 141: 100–111. doi:10.1016/j.isprsjprs.2018.04.010.