

# Understanding rooftop PV panel semantic segmentation of satellite and aerial images for better using machine learning



Peiran Li<sup>a</sup>, Haoran Zhang<sup>a,b,c,\*</sup>, Zhiling Guo<sup>a,c</sup>, Suxing Lyu<sup>a</sup>, Jinyu Chen<sup>a</sup>, Wenjing Li<sup>a</sup>, Xuan Song<sup>d</sup>, Ryosuke Shibasaki<sup>a</sup>, Jinyue Yan<sup>b</sup>

<sup>a</sup> Center for Spatial Information Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan

<sup>b</sup> Future Energy Center, Mälardalen University, Västerås 72123, Sweden

<sup>c</sup> LocationMind Inc. 3-5-2 Iwamotocho, Chiyoda-ku, Tokyo 101-0032, Japan

<sup>d</sup> SUSTech-UTokyo Joint Research Center on Super Smart City, Department of Computer Science and Engineering, Southern University of Science and Technology (SUSTech), Shenzhen, China

## ARTICLE INFO

### Keywords:

PV  
Computer vision  
Deep learning  
Satellite and aerial image  
Semantic segmentation

## ABSTRACT

The photovoltaic (PV) industry boom and increased PV applications call for better planning based on accurate and updated data on the installed capacity. Compared with the manual statistical approach, which is often time-consuming and labor-intensive, using satellite/aerial images to estimate the existing PV installed capacity offers a new method with cost-effective and data-consistent features. Previous studies investigated the feasibility of segmenting PV panels from images involving machine learning technologies. However, due to the particular characteristics of PV panel semantic-segmentation, the machine learning tools need to be designed and applied with careful considerations of the issue formulation, data quality, and model explainability. This paper investigated the characteristics of PV panel semantic-segmentation from the perspective of computer vision. The results reveal that the PV panel image data has several specific characteristics: highly class-imbalance and non-concentrated distribution; homogeneous texture and heterogeneous color features; and the notable resolution threshold for effective semantic-segmentation. Moreover, this paper provided recommendations for data obtaining and model design, aiming at each observed character from the viewpoints of recent solutions in computer vision, which can be helpful for future improvement of the PV panel semantic-segmentation.

## 1. Introduction

### 1.1. Background

With significant reduction of LCOE (Levelized Costs Of Electricity), the fast development and implementation of photovoltaic power generation, including building rooftop and utility photovoltaic [2,45,53], calls for better planning based on accurate and updated data on the installed capacity [60,63]. A field survey with manual data collection can obtain rooftop PV panel installation capacity with high precision but labor-intensive, time-consuming, and expensive. Using a satellite/aerial-image-based approach offers a new way to solve large-scale PV panel installation – segmenting solar panels from images, and has been widely discussed recently. However, the related studies were restricted to employing the ‘fashionable’ models that are well-proven in universal segmentation instead of targeting PV segmentation. Ignorance of the unique characteristics of PV segmentation might lead to a “black box” trap – only concentrating on promoting accuracy with some universal black-

box model designs but overlooking how to targeted enhance the PV segmentation performance. To this end, this study provides new insights into the PV segmentation issue, aiming to reveal the unique characteristics and targeted improvement recommendations of PV segmentation.

### 1.2. Literature review

PV related issues have been a long-standing research topic, including various fields: such as financial analysis on PV industry [54,59], a techno-economic assessment of the PV power generation in various fields [23,38], optimization of the PV installation [6,7]; Campana et al. [8], PV potential analysis [55,65], and so on. In recent years, PV capacity estimation, especially PV capacity estimation combining machine learning tech and GIS data emerged a lot.

With the GIS (Geographic Information System) technologies development, studies on rooftop PV capacity estimation based on geographic information data emerged in recent years, as summarized in Table 1. Based on easily accessible data (e.g., land use, population, and building

\* Corresponding author at: Center for Spatial Information Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan.  
E-mail addresses: [zhang\\_ronan@csis.u-tokyo.ac.jp](mailto:zhang_ronan@csis.u-tokyo.ac.jp) (H. Zhang), [jinyue@kth.se](mailto:jinyue@kth.se) (J. Yan).

**Table 1**

Previous studies for rooftop PV estimation based on geographic information data.

Method	Research Object	Data Type	Spatial Resolution	Method	Spatial scale of the study	Accuracy
[4]	Potential PV Estimation	LiDAR	unknown	Physical Modeling	Sub-city-level	-
[32]	Potential PV Estimation	LiDAR	unknown	Physical Modeling	Sub-city-level	-
[21]	Potential PV Estimation	LiDAR	2 m	Error-minimizing plane-fitting (only for roof estimation)	City-level	-
[24]	Potential PV Estimation	Satellite/Aerial Imagery	unknown	SVM	City-level	-
[14]	Potential PV Estimation	LiDAR	2 m	Neighboring buildings method (only for roof estimation)	Sub-city-level	-
[34]	Existing PV Panels Detection	Satellite Imagery	0.3 m	SVM	100 different houses	<b>Area Estimation:</b> None <b>Count Estimation:</b> CorrectDetectionRate: 94%
[58]	Existing PV Panels Detection	Aerial Imagery	0.3 m	ConvNets	City-level	<b>Area Estimation:</b> None <b>Count Estimation:</b> CorrectDetectionRate: 81%~87%
[3]	Potential PV Estimation	Vector GIS Data	-	SVM	commune level (the smallest administrative division) in Switzerland	-
[13]	Existing PV Panels Detection	Satellite Imagery	unknown	CNN	3347 low-quality Google satellite images	<b>Area Estimation:</b> None <b>Count Estimation:</b> CorrectDetectionRate: 87%
[57]	Existing PV Panels Detection	Aerial Imagery	0.15 m	CNN, VGG-1637 network	City-level (with 50 cities sample in USA)	<b>Area Estimation:</b> Precision: 93.1% for residential areas and 90.5% in non-residential areas <b>Count Estimation:</b> Recall: 88.5% in residential areas and 90.5% in non-residential areas
[18]	Potential PV Estimation	Satellite/Aerial Imagery & LiDAR	unknown	U-NET	City-level	-
[29]	Potential PV Estimation	Satellite/Aerial Imagery	unknown	DeepRoof (Original Architecture)	City-level	-
[66]	Existing PV Panels Detection	Satellite Imagery	0.3 m	U-Net (CrossNets)	City-level	<b>Area Estimation:</b> IoU (Intersection of Union): about 74% <b>Count Estimation:</b> None

density), Izquierdo et al. utilized a statistically representative stratified sample to estimate PV capacity [20]. Several similar geographic and statistical data-based methods were applied for large-scale analysis of potential PV capacity estimation [33,44]. Unlike the statistical-data-based estimation, some researchers tried to estimate potential PV capacity by modeling the solar radiance of rooftops, which is critical for extracting the rooftop shapes. Since LiDAR (Light Detection and Ranging) can easily extract the object's surface, it has been widely used for rooftop extraction. Brito et al. [4] employed LiDAR (Light Detection and Ranging) data to evaluate the photovoltaic potential in the Lisbon suburb by simply using the Solar Analyst extension for ArcGIS (a commercial GIS tool). Lukač et al. [32] utilized LiDAR to extract the rooftops by considering pyranometer measurements and multi-resolution shadowing. Jacques et al. [21] utilized low-resolution LiDAR to evaluate the PV capacity over Leeds city. Gooding et al. [14] conducted similar work in Leeds by using a neighboring buildings method. However, due to the strict restrictions of model inputs, the above studies lack flexibility when handling complex rooftop extraction issues.

With the development of machine learning technology, some studies began to employ machine learning technologies for rooftop PV capacity estimation to avoid the strict input restrictions, as summarized in Table 1. Assouline et al. [3] estimated Switzerland's potential rooftop PV capacity with vector GIS data using the SVM (support vector machine) algorithm. Joshi et al. [24] detected the rooftop for solar PV deployment of Abu Dhabi by satellite/aerial images using the SVM algorithm. Huang et al. [18] employed a deep convolutional segmentation method to extract the rooftop area for 3D model generation combining LiDAR data, then completed the PV potential estimation. Lee et al.

[29] implemented a deep learning model named *DeepRoof* method to detect roof-top for potential PV capacity assessment based on solely satellite images. Also based on machine learning technology, a series of researchers turned to focus on existing PV capacity estimation. Malof et al. [34] employed an SVM algorithm for automatic PV panel detection and succeeded in classifying PV panels based on high-resolution satellite image data from USGS (US Geological Survey). However, this method cannot perform semantic segmentation to extract the PV panels from the images. Following this study, Yuan et al. [58] performed a large-scale PV segmentation based on deep convolutional networks. Similarly, Golovko et al. [13] achieved PV segmentation based on a convolutional neural network with 3347 low-quality Google satellite images. Yu et al. [57] proposed a model called *DeepSolar* trained by datasets coming from 50 cities from the USA and argued different accuracy in residential regions and non-residential regions. In 2020, Zhuang et al. [66] proposed a cross-learning driven U-Net (CrossNets) method to segment roof-top PV panels in satellite images. However, the above studies focused on using the universal machine learning frameworks such as CNN, U-Net, DeepLabv3 and etc., lacking analyzing the characteristics of PV image data and improving the models. There exists a gap for further promoting the segmentation accuracy and robustness with machine learning methods, such as suitable model design, data processing, and their combination [11,41].

### 1.3. Motivations and contributions

To further improve the accuracy and robustness of the PV segmentation, application characteristics should be carefully considered, but not

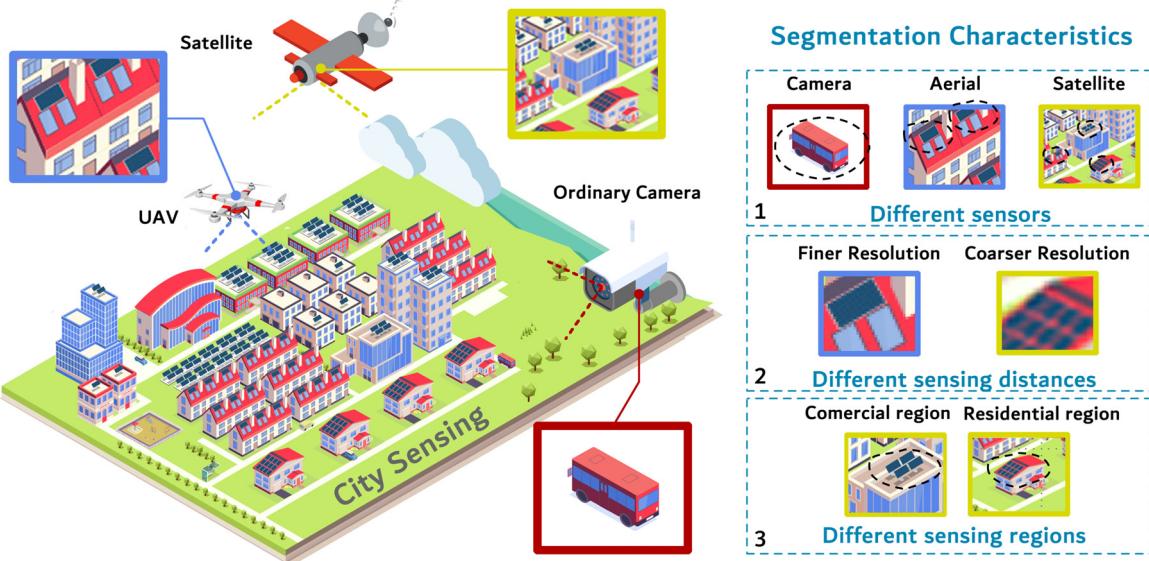


Fig. 1. Characteristics of PV segmentation.

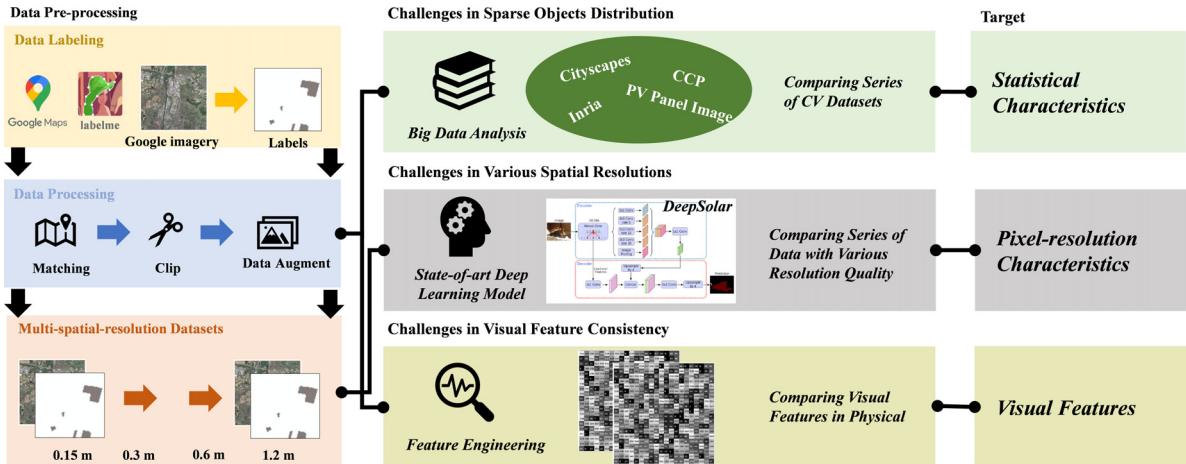


Fig. 2. Illustration of method framework.

only to pursue ‘fashionable’ machine learning methods. This paper explored and analyzed the unique characteristics of PV segmentations in a series of different perspectives to further explore possibilities for improving methods, in particular on the three contributions of this paper as follows:

- The study compared a series of representative segmentation datasets, especially for their class distribution, then pointed out the highly class-imbalance and non-concentrated distribution of PV data.
- The study investigated several PV panel image datasets with various resolution quality, then revealed that 0.3m is the threshold resolution for PV segmentation.
- The study explored the visual features of thousands of PV panels, including texture features and color features, then revealed the homogeneous texture feature and heterogeneous color feature of PV data.

## 2. Problem description and methodology

### 2.1. Problem description

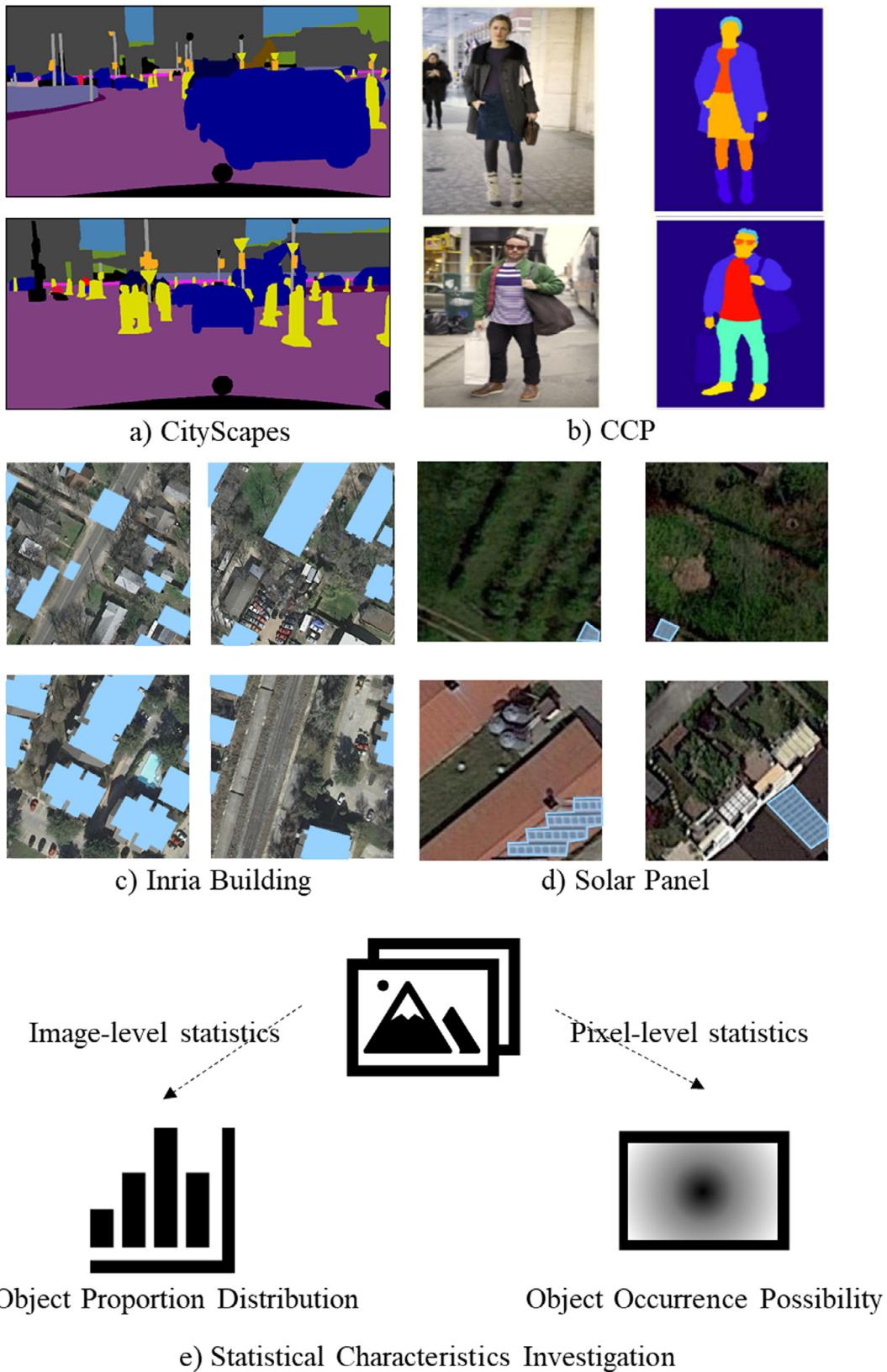
PV segmentation needs to consider the quality and characteristics of associated image data, which is significantly dependent upon sensors, sensing distance, and sensing regions. Brought from the different

sensors, sensing distances, and sensing target regions, PV segmentation (as a computer vision task) shows different characteristics, particularly when using machine learning. Fig. 1 shows three characteristics of PV segmentation, including statistical, pixel-resolution, and visual characteristics.

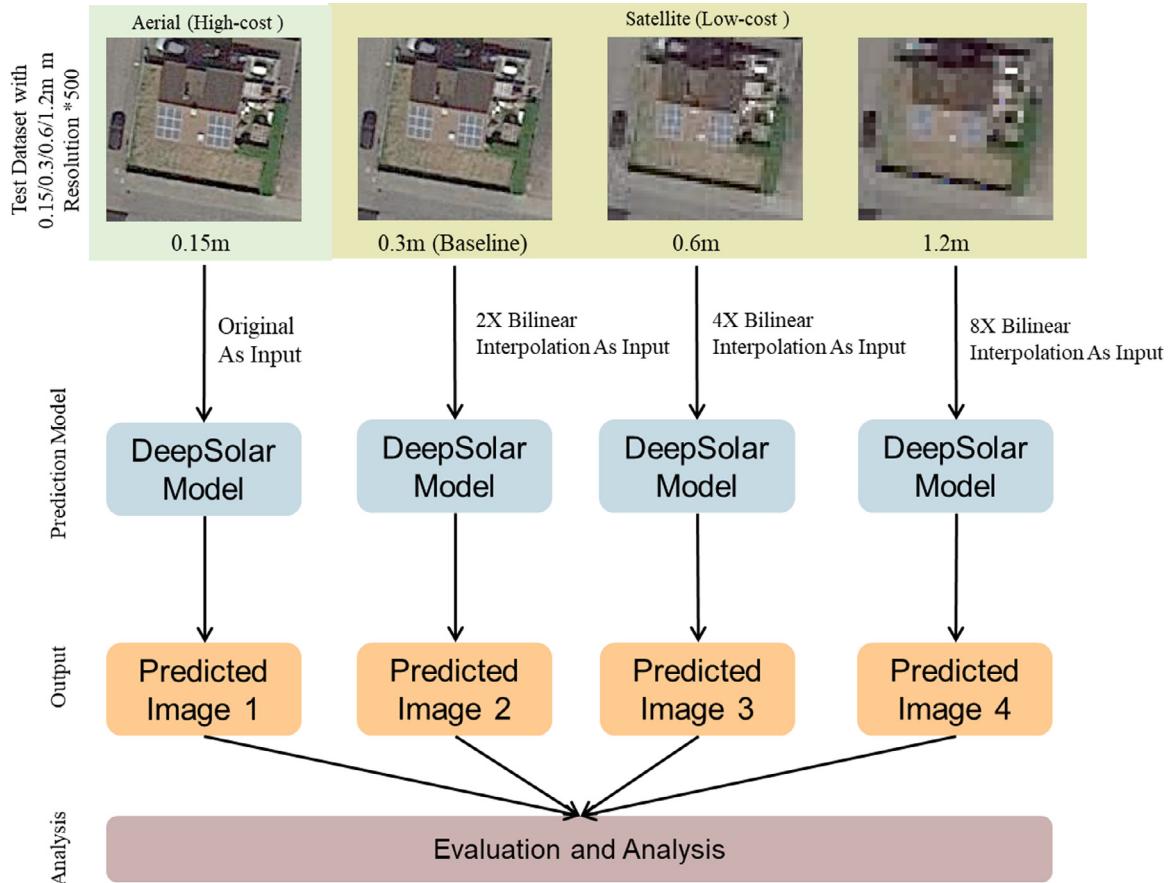
Sensing and perceiving things in cities could be conducted in various ways, such as satellite, aerial (mainly UAVs), or just an ordinary camera, depending upon the objectives to be detected. For example, a bus can be captured with an ordinary camera. However, sensing the large-scale information of building rooftops, aerial (mainly UAVs) or satellites will be the better choice, as shown in Fig. 1.

Different sensing methods result in different *statistical characteristics*. In most cases, the object is the center of the image and occupies the central part of the pixels of the image. Most of the ‘fashionable’ semantic segmentation machine learning methods were developed based on these concentrated datasets. However, the objects (such as PV panels) in satellite/aerial are more deconcentrated and only occupy a small part of the images.

Additionally, different sensing distance generates different *pixel-resolution characteristics*. Satellites can acquire images faster and cheaper than aerial ways, with a relatively lower resolution. As shown in part-2 of Fig. 1, the aerial image can show much finer resolution information



**Fig. 3.** Illustration of different semantic segmentation scenes in computer vision. (a) illustrates the examples in Cityscapes. (b) illustrates the examples of CCP. (c) illustrates the examples of Inria Aerial Image Labeling Dataset. (d) illustrates the examples of our dataset with 0.15 m spatial resolution. (e) illustrates the two parts of statistical characteristics investigation: object proportion distribution in image-level, objects-occurrence possibility in pixel-level.



**Fig. 4.** Illustration for datasets with different spatial resolutions.

than the satellite image when we zoom to the same scale for observing PV panels. In the worst case, the PV panels in the satellite image even cannot be identified via eyes. Pixel-resolution of images affects the segmentation accuracy, but to what extent it affects is unclear, which shall be further investigated to improve data quality and model.

Moreover, different target areas may bring different *visual features*. The same type of object may present different visual features in different regions. Yu et al. [57] took different experiments in residential and non-residential areas, resulting in different segmentation accuracy, which may imply different features of different areas, as shown in part-3 of Fig. 1. However, this hypothesis has not been experimentally proven. An in-depth investigation on this issue will help the model robustness promotion.

## 2.2. Methodology

The methodology framework consists of four parts: data processing, statistical characteristics analysis, pixel-resolution characteristics analysis, and visual features analysis, as shown in Fig. 2.

### 2.2.1. Data pre-processing

This study collected patch satellite/aerial images with 0.15 m/0.3 m/0.6 m/1.2 m spatial resolution in Heilbronn (a city in Europe) from Google, then composited all the patch images to form the entire imagery and manually labeled all PV panels. By a random-sliding window and data augmentation, the entire imagery was cut into numerous source images and label images to form a dataset. Like the training dataset generation, a test dataset was created for evaluation (Details shown in Appendix.2).

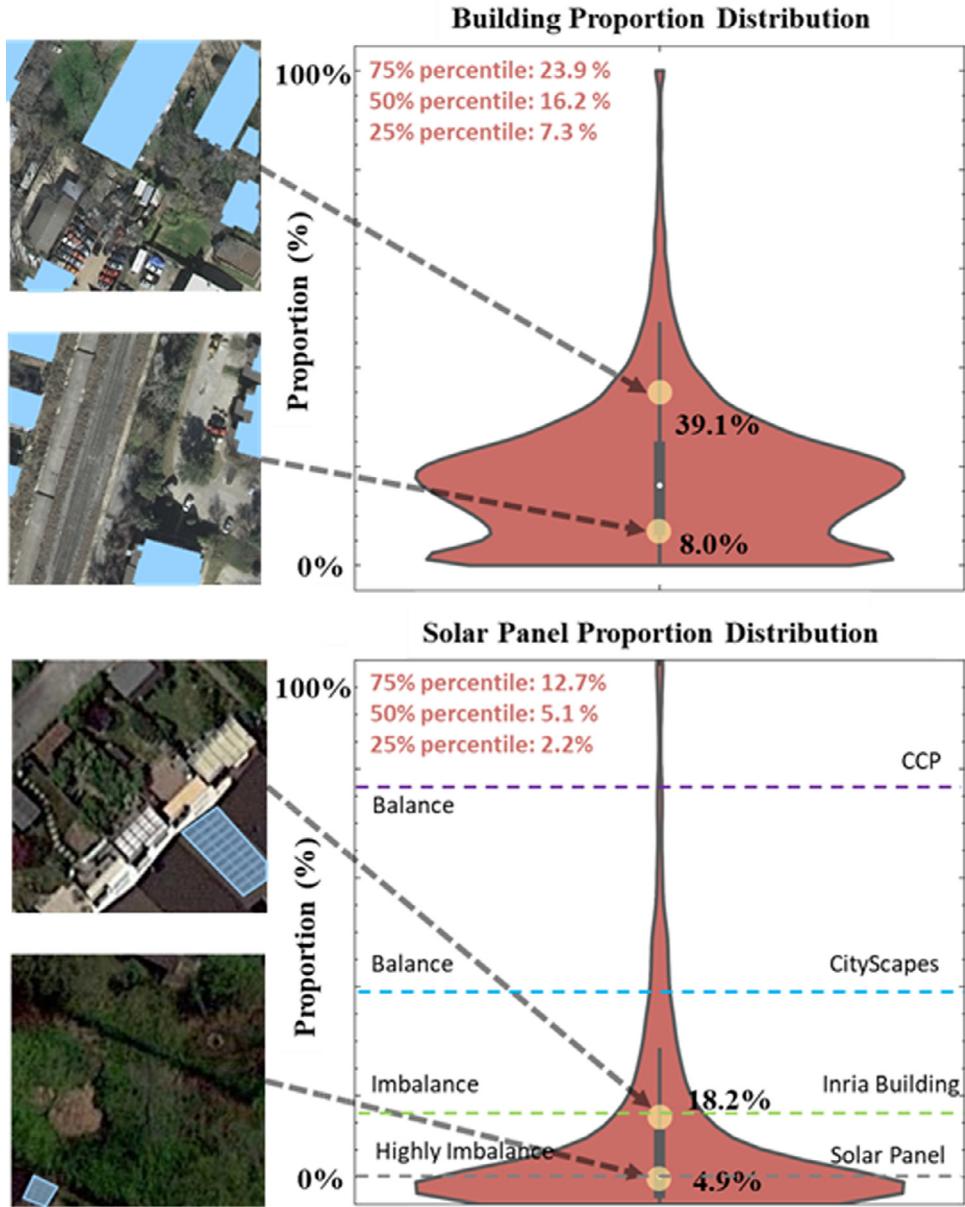
### 2.2.2. Statistical characteristics of PV segmentation

To analyze the *statistical characteristics* of PV segmentation, we compared a series of representative computer vision datasets, discussed the statistical characteristics of PV segmentation, revealed the challenges in sparse objects distribution of PV panels. The representative datasets are Cityscapes, CCP, and Inria building datasets (Fig. 3, *details shown in Appendix.1*). This study utilized two indicators to indicate the statistical characteristics: object proportion distribution in image-level and object occurrence possibility at pixel-level. (1) The ratio of objects to background is taken as the object proportion for an image. Then, from numerous images in a dataset, a statistical distribution of object proportion could be figured out. (2) In each image, every pixel value indicates its property of object or background. An object-occurrence possibility heatmap can be mapped by counting the occurrence frequency from numerous images of a dataset.

### 2.2.3. Pixel-resolution characteristics of PV segmentation

We compared a series of PV datasets with various resolution qualities to analyze the pixel-resolution characteristics. Specifically, 0.15 m, 0.3 m, 0.6 m, and 1.2 m resolutions are analyzed, cover the typical resolution of mainstream satellite/aerial images (Fig. 4). Since the highest spatial resolution of current mainstream non-military optical remote sensing satellites is 0.31 m [28] - a higher resolution has to be obtained by aerial ways, we took 0.3 m as the baseline resolution. The datasets of the four different resolutions are unified by bilinear interpolation to a resolution of 0.15 m as input data. A state-of-art PV segmentation model (*DeepSolar*) is employed to segment PV panels from the source images.

For pixels in a predicted image – result of segmentation, four situations were labeled by comparing predicted value and ground-truth value:



**Fig. 5.** Object/background proportion distribution of the slice images (building object in Inria Dataset, and PV panel in PV dataset). The left side shows several sliced images, and the right side shows the violin plots of numerous sliced images. Within this violin plot, the white marker represents the median of the data; the black bar indicates the interquartile range. In addition, within the violin plot of the PV panel dataset, the mean object/background proportion of four datasets is shown for comparison.

- TP(True Positive) means positive (is PV panel) for prediction and positive (is PV panel) in ground-truth.
- FP(False Positive) means positive (is PV panel) for prediction but negative (not PV panel) in ground-truth.
- FN(False Negative) means negative (not PV panel) for prediction but positive (is PV panel) in ground-truth.
- TN(True Negative) means negative (not PV panel) for prediction and negative (not PV panel) in ground-truth.

For each predicted image, the four most common metrics including *Precision*, *Recall*, *F1score*, and *IoU*(Intersection Over Union), are used to evaluate the PV segmentation:

$$\text{Precision} = \sum \frac{TP}{TP+FP} / N \quad (1)$$

$$\text{Recall} = \sum \frac{TP}{TP+FN} / N \quad (2)$$

$$F1 = \sum \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} / N \quad (3)$$

$$IoU = \sum \frac{TP+TN}{TP+FP+FN+TN} / N \quad (4)$$

#### 2.2.4. Visual features of PV panels

From the perspective of computer vision, the visual features should be from color, texture, or shape analyses [25,51]. Since there is no significant difference in shape features among PV panel images, here, we focused on texture features (a set of metrics calculated in image processing designed to quantify the perceived texture of an image) and color features (a set of metrics calculated in image processing designed to quantify the color vision of an image). To analyze the visual features of solar panels, we clipped thousands of PV panel images (without background), extracted the texture feature and color features, then performed the clustering analysis.

For texture feature extraction, Local Binary Pattern (LBP) [56] was employed to extract the texture feature. In LBP, each pixel of an image is compared with the pixel values of its neighbors, providing a feature pattern that measures the neighborhood relationships between pixels, thus effectively extracting texture features of an image with low computational cost, which is defined as

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \quad (5)$$

where  $(x_c, y_c)$  represents the center pixel of the neighborhood window, its pixel value is  $i_p$ ,  $i_c$  is the value of other pixels in the neighborhood, and  $s(x)$  is the symbolic function.

For color feature extraction, this study computed nine color features derived from the color histogram. The color histogram is one of the most widely used color features in many image retrieval systems [46,50,61]. It represents the composition of colors in an image, i.e., which colors are present and the probability that each color is present. We calculated the histogram, then took the mean, standard deviation, the variation range, of the red channel, blue channel, and green channel, to form a vector with nine feature values, as the color feature of a solar panel image.

Further, the K-means algorithm, one of the most widely used clustering analyses, was applied to investigate the aggregation of extracted features. And an indicator - *Silhouette Coefficient*, was introduced to evaluate the k-means clustering result, which is defined as:

$$\text{Silhouette Coefficient} = \frac{(b - a)}{\max(a, b)} \quad (6)$$

The *Silhouette Coefficient* was calculated using the mean intra-cluster distance  $a$  and the mean nearest-cluster distance  $b$  for each sample [43], where  $b$  is the distance between a sample and the nearest cluster. The best value is 1 and the worst value is -1. Negative values generally indicate that a sample has been assigned to the wrong cluster.

### 3. Finding and discussion

#### 3.1. Statistical characteristics of PV segmentation

##### 3.1.1. Finding 1: highly class imbalance

The results on object/background proportion distribution are shown in Fig. 5, where building proportion distribution (top) and solar panel proportion distribution (bottom) are detailed presented. The result indicated that the CCP dataset and Cityscapes dataset keep relatively balanced between main-object with background, with 76.2 and 39.6% mean proportion. On the contrary, the Inria building dataset and PV dataset present imbalanced (mean proportion shown as the dotted line in Fig. 5). The median object proportion of the Inria building dataset is

23.9%, which is relatively lower than the balanced cases. While the PV dataset shows a highly class imbalance, most images (75% of images) have an object proportion lower than 12.7%.

A serious class imbalance situation for PV data poses a challenge for the training of deep learning models [37]. As mentioned above, in most images, PV panels only take a small part of the whole extent. In over 90% of images of the test dataset, PV panels only take no more than 10% of the whole extent. Moreover, the number of PV panel pixels is far less than the background. From the computer vision perspective, this is a typical class imbalance situation when a class is overrepresented (Table 2), i.e., having much more examples than others in the dataset.

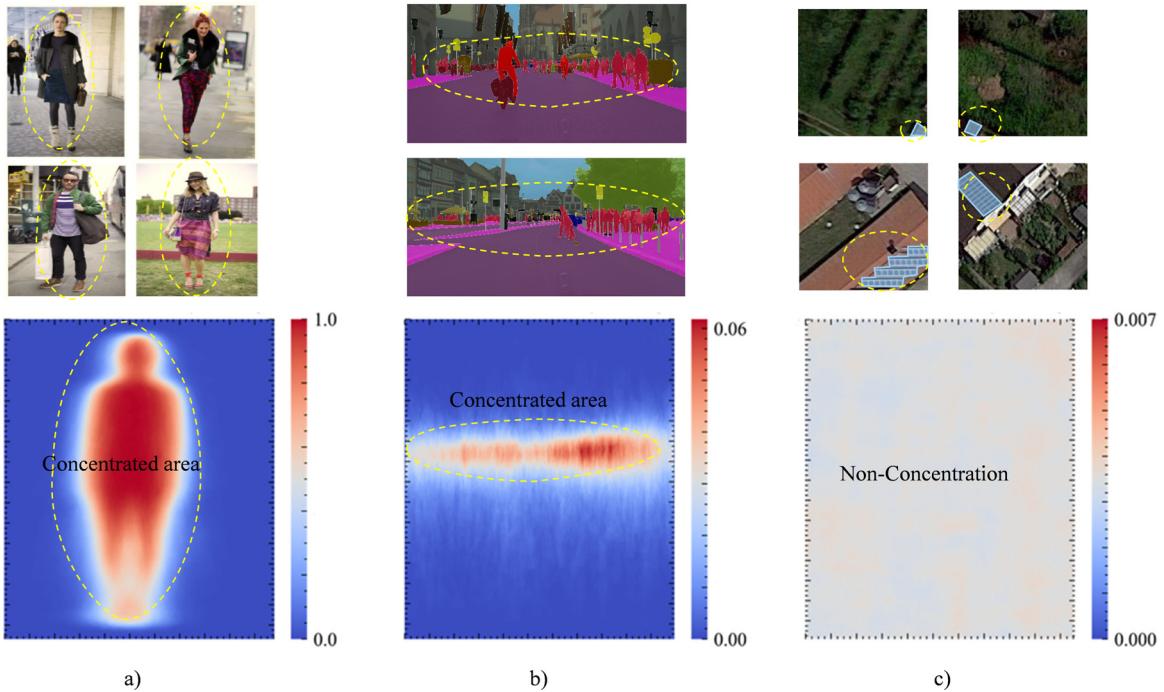
As a consequence of the unbalanced dataset, two problems remain to settle: (1) Inefficient training since the most area is accessible negative, which contributes little proper learning signal, and (2) The easy negatives can overwhelm training and lead to degenerate models [30].

The solutions to this issue can be sorted into four methods: hard sampling, soft sampling, generative and sampling-free methods.

- Hard sampling method addresses imbalance by selecting a proper subset of positive and negative examples: sampling before training [12,40] and sampling while training [10,27,31,40].
- Soft sampling method addresses imbalance by adjusting the contribution of different samples by defined criteria: set constant weights of different classes' samples [39], or set dynamically weight of different classes' samples [30,64].
- Generative methods address the imbalance by directly producing and injecting artificial samples into the training dataset [52].
- Different from the above sample-adjusting methods, sampling-free methods address imbalance beyond sampling. For example, it has been reported that the model can be well trained without any sampling mechanism if hyperparameters are set appropriately [9].

##### 3.1.2. Finding 2: non-concentrated distribution

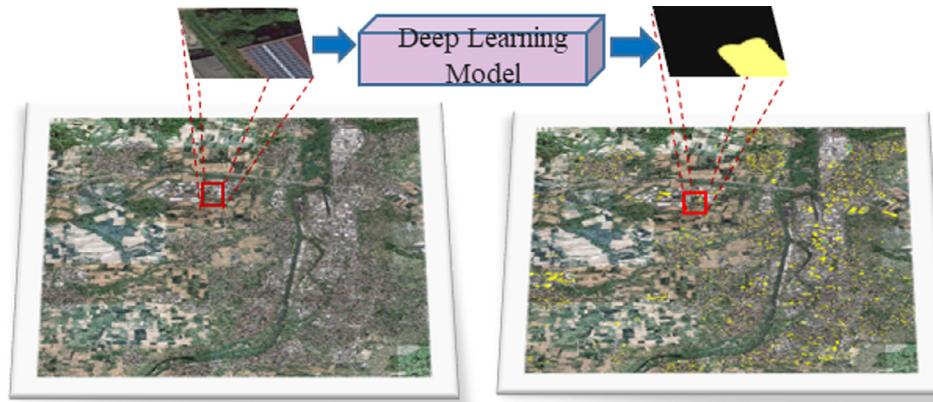
The pixel-level objects appearing probability in the CCP dataset, Cityscapes dataset, and the PV dataset are mapped (shown in Fig. 6). The



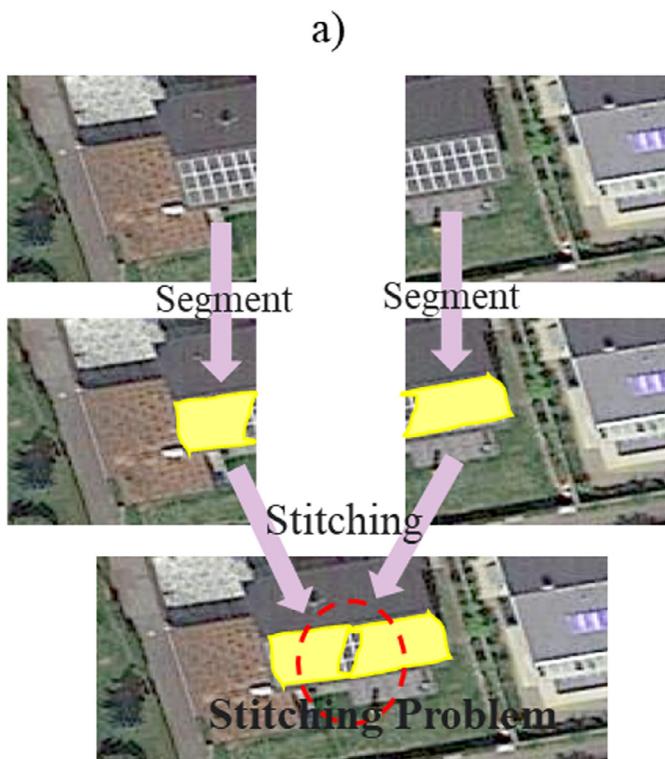
**Fig. 6.** Illustration for the appearing probability density of the target on the sliced image. (a) shows the probability density of the target (person) of the CCP dataset. (b) shows the probability density of the personal target of the CityScapes dataset. (c) shows the probability density of the target (PV panel) of our dataset.

**Table 2**  
Targeted better using machine learning for PV segmentation.

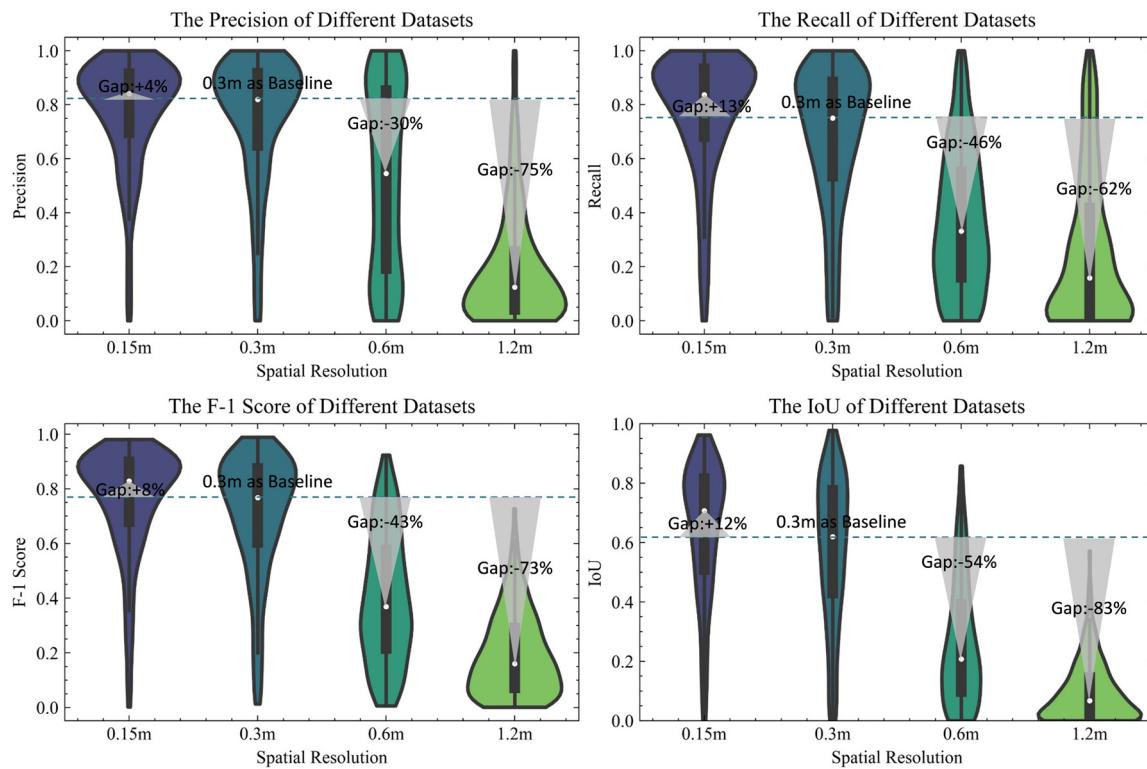
PV Segmentation Characteristics	Problems Caused	Recommendation
Highly Class Imbalance	Training Bias	Solving the imbalance in training: e.g. hard sampling and soft sampling etc.
Non-Concentration Discontinuity and nonlinearity of resolution quality's effects	Stitching Error • Data Acquisition Way • Model Generality	Improving the strategy for tiling and stitching • For data choice: finer resolution helps, efficiency-cost trade-off point is 0.3 m, resolution higher than 1.2 is completed unacceptable. • Deal multi-resolution situation with FPN architecture may help generality in more solutions.
Homogenous texture and Heterogenous Color	Model Generality	• Model combining manual feature engineering may help. • Perform radiance calibration in pre-processing or take corresponding measure in model.



**Fig. 7.** Illustration for the problem of stitching segmentation. (a) illustrates the sliding patches. (b) presents a typical example for problems when stitching segmentation results.



b)



**Fig. 8.** The overall segmentation results of test datasets with different spatial resolutions. (a) the precision of different test datasets. (b) the recall of different test datasets. (c) the f1-score of different test datasets. (d) the IoU of different test datasets

CCP dataset – a camera-created dataset, shows a concentrated area that occupies the central part of images (Fig. 6 (a)). Similarly, the Cityscapes dataset - a car-on-board-camera created dataset, also shows a concentrated area for person objects (Fig. 6 (b)). Unlike the camera-created datasets, PV panels are randomly distributed on images (Fig. 6 (c)) since satellite/aircraft sensing just untargeted scanned the ground while flying in.

This non-concentration characteristic causes ‘titling-stitching’ challenges (Table 2). Satellite/aerial images are usually stored as huge images, which could be called ‘tiles’, and they are too large to be segmented directly (Huang et al. [17]). Thus, as shown in Fig. 7, during label inference, smaller images called ‘patches’ are processed separately and then ‘stitched’ together to create the entire output map.

This procedure brings to error when stitching ‘patches’ to the entire map, and there are two basic methods to address this issue:

- Make the patches overlapped when cutting the entire image into patches, and average the overlapping area when stitched them together so that the stitched error could be partly eliminated [5,35,47].
- Another method is to cut off the edges of the output label patches and join the remaining patches without overlapping or averaging (Huang et al. [15]; Huang et al. [16,19]).
- It is worth noting that Huang et al. [17] discussed the nature of stitching strategies: all approaches are fundamentally motivated by the translational variance of segmentation networks, i.e. the label predicted for a particular pixel depends upon its relative position in the input patch. And based on the above theory and a series of experiments, they give three specific and useful recommendations: increasing the output size during label inference helps; there is little benefit to averaging labels; clipping the 20–40 pixels that are closest to the edges of label patches (their experiments include PV image dataset, thus here we keep the accurate recommended clipping range value).

### 3.2. Pixel-resolution characteristics of PV segmentation

#### 3.2.1. Finding 3: discontinuity and nonlinearity of resolution’s effects

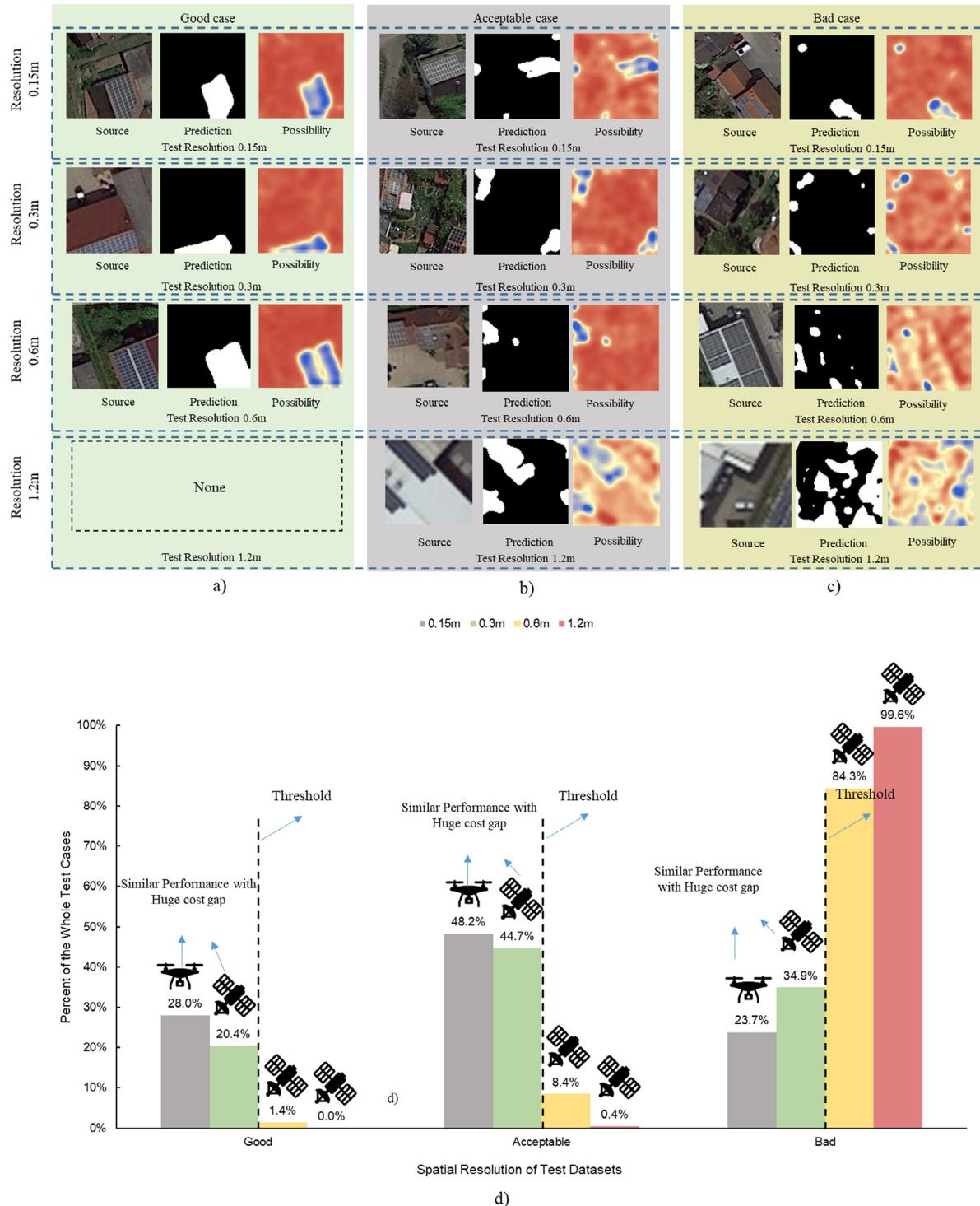
In terms of the pixel-resolution characteristics of the PV segmentation, the accuracy of different resolutions was calculated by various metrics introduced in Section 2.2. The results show that 0.3 m is the threshold resolution for PV segmentation: the segmentation accuracy decreases along with the lower resolution, but when the spatial resolution changes from 0.15 m to 0.3 m, the accuracy changes little, when the resolution decreases from 0.3 m to 0.6 m or 0.6 m to 1.2 m, the accuracy decreases sharply, as Fig. 8 shows.

Based on the IoU, three types of results can be classified as shown in Fig. 9, namely, Good case with  $\text{IoU}$  higher than 0.8; Acceptable case with  $\text{IoU}$  within 0.5 to 0.8; and Bad case with  $\text{IoU}$  less than 0.5.

The results show that the accuracy gap between 0.15 m and 0.3 m resolution is not significant since they have a similar proportion of good case (28.0% for 0.15 m and 20.8% for 0.3 m) and acceptable case (48.2% for 0.15 m and 44.7% for 0.3 m). However, when it comes to 0.6 m’s spatial resolution, only less than 1.5% of cases show good, and less than 10% cases show acceptable. Moreover, when it comes to 1.2 m resolution, nearly no case shows acceptable accuracy. Even for the highest accuracy case of 1.2 m resolution, the result is also unacceptable (Fig. 9(d)). Therefore, it could be concluded that 0.3~0.6 m is the trade-off resolution for PV segmentation, and 1.2 m is the minimum for the resolution requirements of data.

Making the pixel-resolution characteristics clear enable us to discuss the data acquisition ways, and model’s generality with different resolutions, of PV segmentation (Table 2):

- In terms of data acquisition ways, satellite way has better cost-effectiveness than aerial way: (1) The PV segmentation could perform well with 0.3 m resolution images, and some remote sensing satellite images have been able to reach 0.31 m resolution [28].

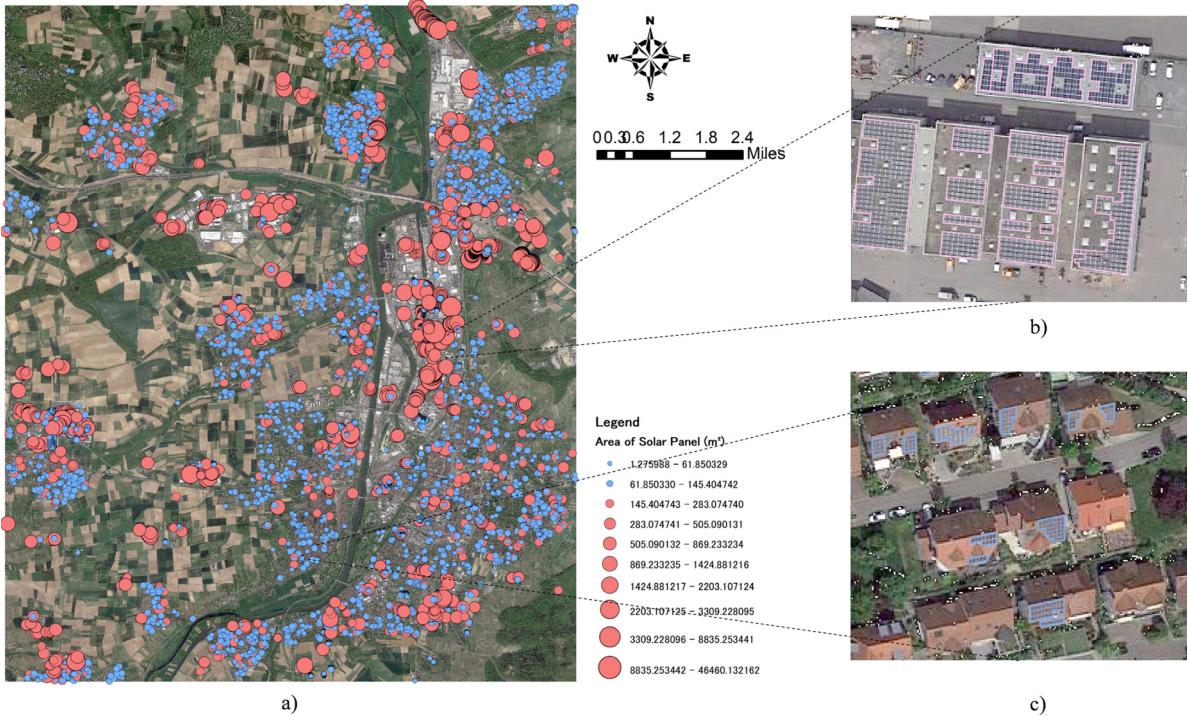


**Fig. 9.** Illustration of semantic segmentation result and the distribution of result accuracy. (a) presents three good cases for four different spatial resolution test datasets. (b) presents four acceptable cases for four different spatial resolution test datasets. (c) presents four bad cases for four different spatial resolution test datasets. d) shows the proportion of different accuracy classes of four different spatial resolution test datasets.

Aerial way can reach a higher resolution but the accuracy benefit is not significant. (2) Acquiring images from satellites can reach a lower cost while providing greater coverage, compared with aerial way. For example, *Quickbird*, as one of the common-used remote sensing satellites, can obtain panchromatic images with 0.61 m's spatial resolution of totally 210,000 km long and 16.5 km wide every day, which surpasses aerial photogrammetry far away. (3) Aerial

photogrammetry (UAVs or planes) are strictly restricted by government policy especially in the urban area, while satellite remote sensing is freer of it.

- In terms of model's resolution generality, it indicates that multi-resolution issues should be carefully considered. The pyramid methods, including feature pyramid and image pyramid, are the classical targeted solution for multi-resolution issues. An image pyramid is to



**Fig. 10.** Illustration for PV panels distribution with different sizes. (a) illustrates the distribution of PV panels with different sizes in our training dataset. (b) presents a typical example for large-size PV. (c) presents a typical example for a small-size PV panel.

construct image pyramids from the original one [48,49], while the feature pyramid is to employ extracted feature layers to format the pyramid [26,31,62].

### 3.3. Visual features of PV segmentation

#### 3.3.1. Finding 4: homogeneous texture and heterogeneous color

As Fig. 10 below shows, intuitively, large-size PV panels (usually located in commercial/non-residential areas) look different from small-size PV panels (usually in residential areas). Yu et al. took PV segmentation experiments in residential and non-residential areas and resulted in different segmentation accuracy [57]. However, from the theoretical perspective of the computer vision field, the visual features should be from color, texture, or shape analyses [25,51], but not only from intuitive.

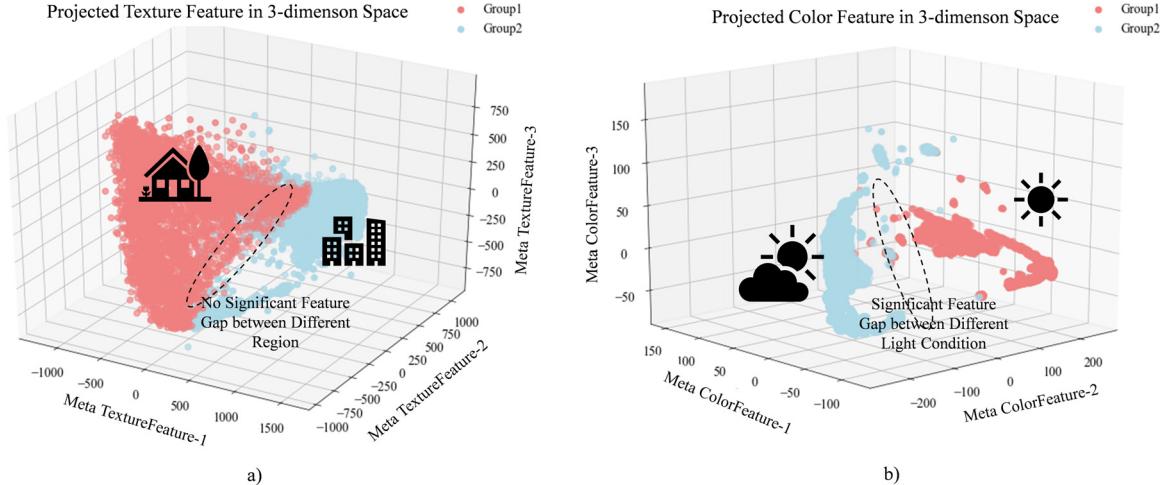
As the results of the visual feature analysis, the different clustering performance (indicated by *Silhouette Score*) shows the homogeneous texture and heterogeneous color of PV panels. The *Silhouette Score* of texture clustering of PV panels is 0.138, which means that there is no enough difference for type-separate as the texture perspective. Projecting the visual feature vector of every panel into a 3-dimension space, we found that the feature vectors are highly aggregated without any gap, as shown in Fig. 11(a). The results conclude that PV panels show a high degree of texture consistency, i.e., they have homogeneous texture. The *Silhouette Score* performance of color features of PV panels is 0.621, which means the divisibility is much better than the texture feature. The significant feature gap between different light conditions of PV panels is clearly shown in Fig. 11 (b). However, almost all the PV panels show a highly similar color histogram pattern - most of the pixel color values distributes around 0 or 255 (Fig. 12 (a)). This phenomenon indicates that the typical color feature of PV panels is a pattern of monotonous color with contrasting colors. The critical point for the gap of different groups of color features is the brightness instead of color pattern, as shown in Fig. 12 (a). PV panels of *Group2* have more pixels in color 255 (black side) than 0 (white side) compared to those in *Group1*. Since there is no physical texture difference between different PV panels, it highly

implies that the lighting conditions, i.e., sunning or shade facing, results in this phenomenon, instead of the inner difference of PVs themselves (Fig. 12 (b)).

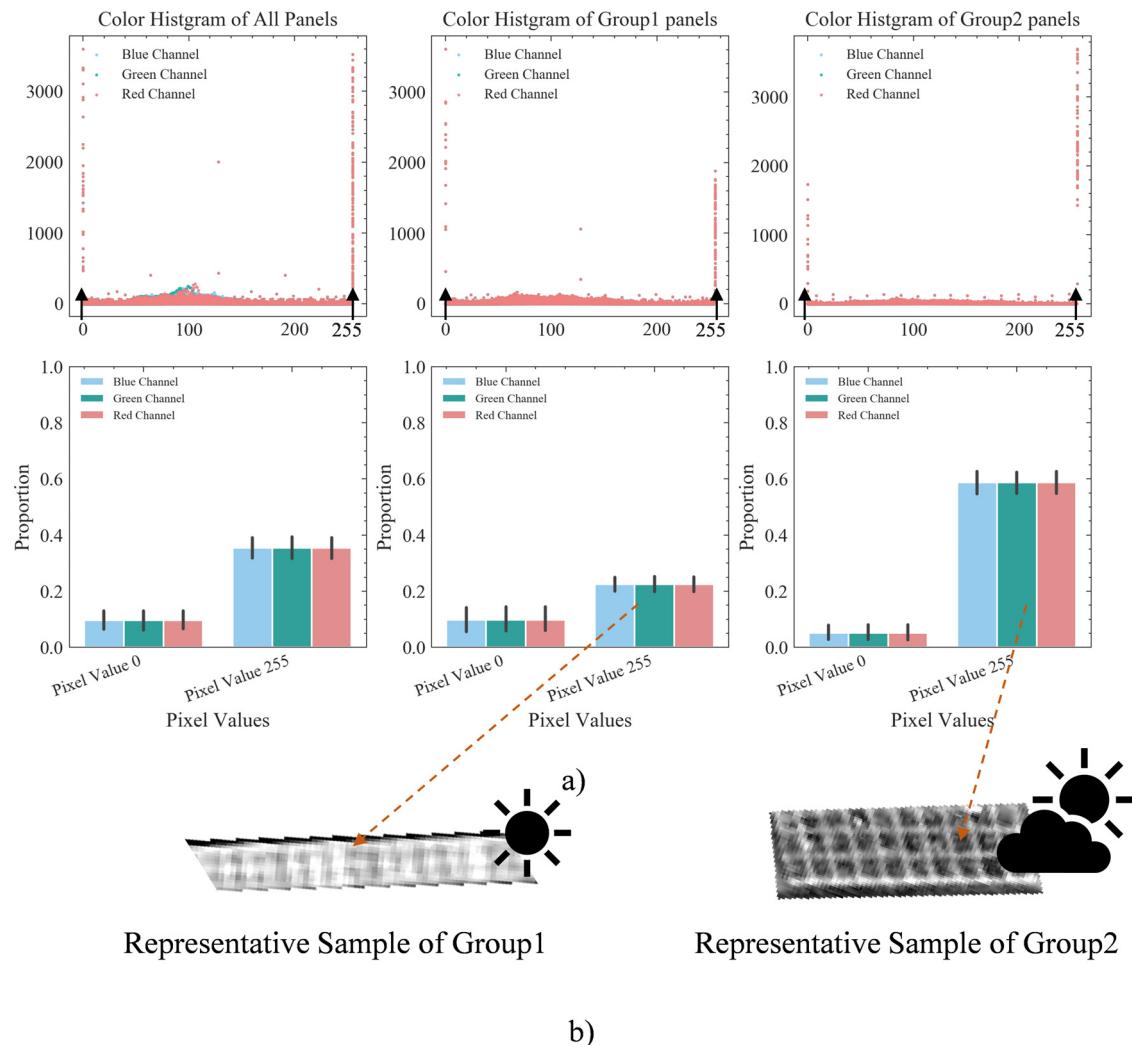
The above results concluded that the color nature of PV in vision is monotonous images with white-black colors but lighting conditions resulting in different clustering groups in color clustering. On the other hand, for the texture feature, it is found that nearly all the PV panels show highly consistent texture characteristics.

The understanding of visual features of PV panels helps us to enhance the model generality of PV segmentation (Table 2):

- The homogenous feature can serve as an additional part for some easily confused cases (e.g., rooftop glass shown as **Appendix. S2**) to improve the robustness of the model. Deep learning has hugely extended the boundaries in the field of computer vision. However, it is challenging to adjust the hidden parameters, when a deep learning model performs poorly outside the training data [36]. For example, if a deep learning model for PV segmentation misidentifies similar objects such as rooftop glass, it's hard to targeted adjust the 'black-box'. In this case, separating them by extracted texture feature (e.g. extracted texture feature by LBP) could serve as a complementary way of deep learning model: as Alhindi et al. [1] reported that SVM could achieve higher accuracy using LBP as features that surpasses the accuracy obtained by deep learning way.
- The heterogeneous color feature suggests the demand for the radiance calibration of input images. The variation in radiosity between the source and target datasets affects the model performance. For example, when employing a model trained with a dataset collected during sunny days to deal with a dataset collected during cloudy days, the accuracy might be limited. As reported in 2015, a commonly used strategy is data augmentation, which will resample the original input in various cases, including geometric and radiosity transformations, to extend the sample space, thus extending the generalization capability of the trained model [42]. Further, Ji et al. combined relative radiometric calibration and radiometric resampling to improve the model generalization [22].



**Fig. 11.** Illustration of clustering result of texture features and color features of solar panel images. (a) the 3-D visualization of texture feature clustering result: the cluster group 1 refers to the solar panels in residential area and the cluster group 2 refers to the solar panels in commercial/urban area. (b) the 3-D visualization of color feature clustering result: the cluster group 1 refers to the solar panels with more light pixels (implies acquisition in sunning situation) and the cluster group 2 refers to the solar panels with less light pixels (implies acquisition in shade facing situation).



**Fig. 12.** Illustration of the color histogram of PVs. (a) illustrates the color histogram of all panels, panels of group1 and group2. Moreover, the bar chart illustrates the mean proportion of all panels, panels of group1 and group2. (b) illustrates two representative samples of group 1 and group 2.

#### 4. Conclusion

This paper investigated the characteristics of PV segmentation as a computer vision issue, revealing a series of challenges, giving a series of targeted better using recommendations for PV segmentation.

First, it is concluded that the PV satellite and aerial images characterize by the highly class-imbalance and non-concentrated distribution. The PV panel image data shows a highly class imbalance: most images (75% of images) have an object proportion lower than 12.7% since the PV panels behaved as small objects in satellite and aerial images. Class imbalance leads to inadequate training and negative-sample-bias of a semantic segmentation model. A series of related studies have been reviewed to recommend possible solutions such as hard sampling and soft sampling methods. Besides, PV panels appear randomly in space with no concentrated area in slice images. The non-concentrated distribution will raise the “tiling-stitching” problem, which brings errors when stitching patch results. Also, some recommendations were raised to reduce this type of error.

Secondly, a notable resolution threshold for effective PV segmentation has been found - 0.3 m is the threshold resolution for a PV segmentation issue, and 1.2 m is the lowest limitation for PV segmentation. The accuracy stays similar from 0.15 m to 0.3 m but decreases sharply when it comes to 0.6 m’s spatial resolution. Moreover, when it comes to 1.2 m resolution, nearly no case shows acceptable accuracy. It means satellite image (can reach 0.31 m) is enough for PV segmentation. Besides the data resolution choice, we recommend FPN architecture to help robustness in a multi-resolution situation regarding model robustness and generality.

Finally, this study proved the homogeneous texture feature and heterogenous color feature. Most PV panel images behave highly similar texture features but two different color features. It implies that different optical situations affect their color features. Through the review of a series of related works, it is recommended that homogeneous texture can serve as a stable feature to help the improvement of a PV segmentation model, and heterogeneous color implies the necessity of image radiance calibration or corresponding generality model improvement.

Despite these findings, uncertainties remain with the limitation of compared datasets, which may lead to the ignorance of some facts about PV segmentation. Notwithstanding this uncertainty, with the multiple numerical experiments, it can be well believed for the generality of our findings.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.adapen.2021.100057](https://doi.org/10.1016/j.adapen.2021.100057).

#### References

- [1] Alhindhi, T.J., Kalra, S., Ng, K.H., Afrin, A., & Tizhoosh, H.R. (2018). Comparing LBP, HOG and deep features for classification of histopathology images. In arXiv. <http://kima.uwaterloo.ca/>.
- [2] Amabile L, Bresch-Pietri D, Hajje GE, Labb   S, Petit N. Optimizing the self-consumption of residential photovoltaic energy and quantification of the impact of production forecast uncertainties. *Adv Appl Energy* 2021;100020. doi:[10.1016/j.adapen.2021.100020](https://doi.org/10.1016/j.adapen.2021.100020).
- [3] Assouline D, Mohajeri N, Scarzetti JL. Quantifying rooftop photovoltaic solar energy potential: a machine learning approach. *Sol Energy* 2017. doi:[10.1016/j.solener.2016.11.045](https://doi.org/10.1016/j.solener.2016.11.045).
- [4] Brito MC, Gomes N, Santos T, Tened  rio JA. Photovoltaic potential in a Lisbon suburb using LiDAR data. *Sol Energy* 2012. doi:[10.1016/j.solener.2011.09.031](https://doi.org/10.1016/j.solener.2011.09.031).
- [5] Camilo, J., Wang, R., Collins, L.M., Bradbury, K., & Malof, J.M. (2018). Application of a semantic segmentation convolutional neural network for accurate automatic detection and mapping of solar photovoltaic arrays in aerial imagery. In arXiv.
- [6] Campana PE, Leduc S, Kim M, Olsson A, Zhang J, Liu J, Kraxner F, McCallum I, Li H, Yan J. Suitable and optimal locations for implementing photovoltaic water pumping systems for grassland irrigation in China. *Appl Energy* 2017;185:1879–89. doi:[10.1016/j.apenergy.2016.01.004](https://doi.org/10.1016/j.apenergy.2016.01.004).
- [7] Campana PE, Li H, Zhang J, Zhang R, Liu J, Yan J. Economic optimization of photovoltaic water pumping systems for irrigation. *Energy Convers Manag* 2015;95:32–41. doi:[10.1016/j.enconman.2015.01.066](https://doi.org/10.1016/j.enconman.2015.01.066).
- [8] Campana, P.E., Landelius, T., Andersson, S., Lundstr  m, L., Nordlander, E., He, T., Zhang, J., Stridh, B., & Yan, J. (2020). A gridded optimization model for photovoltaic applications. *Sol Energy*, 202, 465–484. doi:[10.1016/j.solener.2020.03.076](https://doi.org/10.1016/j.solener.2020.03.076)
- [9] Chen, J., Liu, D., Xu, T., Wu, S., Zhang, S., Luo, B., Peng, X., & Chen, E. (2019). Is sampling heuristics necessary in training deep object detectors? ArXiv. <http://arxiv.org/abs/1909.04868>
- [10] Dai J, Li Y, He K, Sun J. R-FCN: object detection via region-based fully convolutional networks. *Advances in neural information processing systems* 2016. doi:[10.5555/3157096.3157139](https://doi.org/10.5555/3157096.3157139).
- [11] Doll  r P, Wojek C, Schiele B, Perona P. Pedestrian detection: an evaluation of the state of the art. *IEEE Trans Pattern Anal Mach Intell* 2012. doi:[10.1109/TPAMI.2011.155](https://doi.org/10.1109/TPAMI.2011.155).
- [12] Girshick, R., Donahue, J., Darrell, T., Malik, J., & Berkeley, U.C. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. 580–587. doi:[10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [13] Golovko V, Bezobrazov S, Kroshchanka A, Sachenko A, Komar M, Karachka A. Convolutional neural network based solar photovoltaic panel detection in satellite photos. In: Proceedings of the 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017, 1; 2017. p. 14–19. doi:[10.1109/IDAACS.2017.8094501](https://doi.org/10.1109/IDAACS.2017.8094501).
- [14] Gooding J, Crook R, Tomlin AS. Modelling of roof geometries from low-resolution LiDAR data for city-scale solar energy applications using a neighbouring buildings method. *Appl Energy* 2015. doi:[10.1016/j.apenergy.2015.03.013](https://doi.org/10.1016/j.apenergy.2015.03.013).
- [15] Huang B, Collins LM, Bradbury K, Malof JM. Deep convolutional segmentation of remote sensing imagery: a simple and efficient alternative to stitching output labels. In: Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS); 2018.
- [16] Huang B, Lu K, Audebert N, Khalel A, Tarabalka Y, Malof J, Boulch A, Sauv BLE, Collins L, Bradbury K, Lef  vre S, El-Saban M. Large-scale semantic classification: outcome of the first year of inria aerial image labeling benchmark. In: Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS); 2018.
- [17] Huang B, Reichman D., Collins, L.M., Bradbury, K., & Malof, J.M. (2018). Tiling and stitching segmentation output for remote sensing: basic challenges and recommendations. In arXiv.
- [18] Huang Z, Mendis T, Xu S. Urban solar utilization potential mapping via deep learning technology: a case study of Wuhan, China. *Appl Energy* 2019;250:283–91 September 2018. doi:[10.1016/j.apenergy.2019.04.113](https://doi.org/10.1016/j.apenergy.2019.04.113).
- [19] Iglovikov, V., Mushinskiy, S., & Osin, V. (2017). Satellite imagery feature detection using deep convolutional neural network: Kaggle competition. In arXiv.
- [20] Izquierdo S, Rodrigues M, Fueyo N. A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations. *Sol Energy* 2008. doi:[10.1016/j.solener.2008.03.007](https://doi.org/10.1016/j.solener.2008.03.007).
- [21] Jacques DA, Gooding J, Giesekam JJ, Tomlin AS, Crook R. Methodology for the assessment of PV capacity over a city region using low-resolution LiDAR data and application to the City of Leeds (UK). *Appl Energy* 2014. doi:[10.1016/j.apenergy.2014.02.076](https://doi.org/10.1016/j.apenergy.2014.02.076).
- [22] Ji S, Wei S, Lu M. A scale robust convolutional neural network for automatic building extraction from aerial and satellite imagery. *Int J Remote Sens* 2019. doi:[10.1080/01431161.2018.1528024](https://doi.org/10.1080/01431161.2018.1528024).
- [23] Jiang M, Lv Y, Wang T, Sun Z, Liu J, Yu X, Yan J. Performance analysis of a photovoltaics aided coal-fired power plant. *Energy Proc* 2019;158:1348–53. doi:[10.1016/j.egypro.2019.01.330](https://doi.org/10.1016/j.egypro.2019.01.330).
- [24] Joshi B, Hayk B, Al-Hinai A, Woon WL. Rooftop detection for planning of solar PV deployment: a case study in Abu Dhabi. *Lect Notes Comput Sci Incl Subser Lecture Notes Artif Intell Lect Notes Bioinf* 2014. doi:[10.1007/978-3-319-13290-7\\_11](https://doi.org/10.1007/978-3-319-13290-7_11).
- [25] Kavitha JC, Surulandi A. Texture and color feature extraction for classification of melanoma using SVM. In: Proceedings of the 2016 International Conference on Computing Technologies and Intelligent Data Engineering, ICCTIDE 2016; 2016.
- [26] Kirillov A, Girshick R, He K, Dollar P. Panoptic feature pyramid networks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2019. p. 6392–401. 2019-June. doi:[10.1109/CVPR.2019.000656](https://doi.org/10.1109/CVPR.2019.000656).
- [27] Kong T, Sun F, Yao A, Liu H, Lu M, Chen Y. RON: Reverse connection with objectness prior networks for object detection. In: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017; 2017.
- [28] Kwan C, Budavari B, Bovik AC, Marchisio G. Blind Quality Assessment of Fused WorldView-3 Images by Using the Combinations of Pansharpening and Hypersharpener Paradigms. *IEEE Geosci Remote Sens Lett* 2017. doi:[10.1109/LGRS.2017.2737820](https://doi.org/10.1109/LGRS.2017.2737820).
- [29] Lee S, Iyengar S, Feng M, Shenoy P, Maji S. Deeprof: a data-driven approach for solar potential estimation using rooop imagery. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2019. p. 2105–13. doi:[10.1145/3292500.3330741](https://doi.org/10.1145/3292500.3330741).
- [30] Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision; 2017.
- [31] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., & Berg, A.C. (2016). SSD : single shot MultiBox detector.
- [32] Luk  a N, Z  aus D, Seme S, Z  alik B,   tumberger G. Rating of roofs’ surfaces regarding their solar potential and suitability for PV systems, based on LiDAR data. *Appl Energy* 2013. doi:[10.1016/j.apenergy.2012.08.042](https://doi.org/10.1016/j.apenergy.2012.08.042).

- [33] Mainzer K, Fath K, Mckenna R, Stengel J, Fichtner W, Schultmann F. A high-resolution determination of the technical potential for residential-roof-mounted photovoltaic systems in Germany. *Sol Energy* 2014. doi:[10.1016/j.solener.2014.04.015](https://doi.org/10.1016/j.solener.2014.04.015).
- [34] Malof JM, Hou R, Collins LM, Bradbury K, Newell R. Automatic solar photovoltaic panel detection in satellite imagery. In: Proceedings of the 2015 International Conference on Renewable Energy Research and Applications, ICRERA 2015, 5; 2015. p. 1428–31. doi:[10.1109/ICRERA.2015.7418643](https://doi.org/10.1109/ICRERA.2015.7418643).
- [35] Marmanis D, Schindler K, Wegner JD, Galliani S, Datcu M, Stilla U. Classification with an edge: improving semantic image segmentation with boundary detection. *ISPRS J Photogr Remote Sens* 2018. doi:[10.1016/j.isprsjprs.2017.11.009](https://doi.org/10.1016/j.isprsjprs.2017.11.009).
- [36] O'Mahony N, Murphy T, Panduru K, Riordan D, Walsh J. Real-time monitoring of powder blend composition using near infrared spectroscopy. In: Proceedings of the International Conference on Sensing Technology, ICST, 2017-Decem; 2018. p. 1–6. doi:[10.1109/ICSenT.2017.8304431](https://doi.org/10.1109/ICSenT.2017.8304431).
- [37] Oksuz K, Cam BC, Kalkan S, Akbas E. Imbalance problems in object detection: a review. *IEEE Trans Pattern Anal Mach Intell* 2020;88:28:1 c–1. doi:[10.1109/tpami.2020.2981890](https://doi.org/10.1109/tpami.2020.2981890).
- [38] Qi L, Jiang M, Lv Y, Zhang Z, Yan J. Techno-economic assessment of photovoltaic power generation mounted on cooling towers. *Energy Convers Manag* 2021;235:113907. doi:[10.1016/j.enconman.2021.113907](https://doi.org/10.1016/j.enconman.2021.113907).
- [39] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2016.
- [40] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN : towards real-time object detection with region proposal networks. 39(6), 1137–1149.
- [41] Risi S, Togelius J. Increasing generality in machine learning through procedural content generation. *Nat Mach Intell* 2020. doi:[10.1038/s42256-020-0208-z](https://doi.org/10.1038/s42256-020-0208-z).
- [42] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci Incl Subser Lecture Notes Artif Intell Lect Notes Bioinf* 2015. doi:[10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [43] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987. doi:[10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [44] Schallenberg-Rodríguez J. Photovoltaic techno-economical potential on roofs in regions and islands: the case of the Canary Islands. Methodological review and methodology proposal. *Renew Sustain Energy Rev* 2013. doi:[10.1016/j.rser.2012.11.078](https://doi.org/10.1016/j.rser.2012.11.078).
- [45] Schleifer AH, Murphy CA, Cole WJ, Denholm PL. The evolving energy and capacity values of utility-scale PV-plus-battery hybrid system architectures. *Adv Appl Energy* 2021;2:100015. doi:[10.1016/j.adapen.2021.100015](https://doi.org/10.1016/j.adapen.2021.100015).
- [46] Sergyán S. Color histogram features based image classification in content-based image retrieval systems. In: Proceedings of the SAMI 2008 6th International Symposium on Applied Machine Intelligence and Informatics; 2008. p. 221–4. doi:[10.1109/SAMI.2008.4469170](https://doi.org/10.1109/SAMI.2008.4469170).
- [47] Sherrah, J. (2016). Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery. 1–22. <http://arxiv.org/abs/1606.02585>.
- [48] Singh B, Davis LS. An analysis of scale invariance in object detection - SNIP. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2018. p. 3578–87. doi:[10.1109/CVPR.2018.00377](https://doi.org/10.1109/CVPR.2018.00377).
- [49] Singh B, Najibi M, Davis LS. Sniper: efficient multi-scale training. *Ad Neural Inf Process Syst* 2018.
- [50] Suhasini PS, Sri Rama Krishna K, Krishna M, I V. Content based image retrieval based on different global and local color histogram methods: a survey. *J. Inst Eng India Ser B* 2017;98(Issue 1):129–35 Springer India. doi:[10.1007/s40031-016-0223-y](https://doi.org/10.1007/s40031-016-0223-y).
- [51] Tian DP. A review on image feature extraction and representation techniques. *Int J Multimed Ubiquitous Eng* 2013.
- [52] Wang, X., Shrivastava, A., & Gupta, A. (2017). A-fast-RCNN: hard positive generation via adversary for object detection.
- [53] Yan J, Yang Y, Elia Campana P, He J. City-level analysis of subsidy-free solar photovoltaic electricity price, profits and grid parity in China. *Nat Energy* 2019;4(8):709–17. doi:[10.1038/s41560-019-0441-z](https://doi.org/10.1038/s41560-019-0441-z).
- [54] Yan J, Yang Y, Elia Campana P, He J. City-level analysis of subsidy-free solar photovoltaic electricity price, profits and grid parity in China. *Nat Energy* 2019;4(8):709–17. doi:[10.1038/s41560-019-0441-z](https://doi.org/10.1038/s41560-019-0441-z).
- [55] Yang Y, Campana PE, Stridh B, Yan J. Potential analysis of roof-mounted solar photovoltaics in Sweden. *Appl Energy* 2020;279:115786. doi:[10.1016/j.apenergy.2020.115786](https://doi.org/10.1016/j.apenergy.2020.115786).
- [56] Yi X, Eramian M. LBP-based segmentation of defocus blur. *IEEE Trans Image Process* 2016;25(4):1626–38. doi:[10.1109/TIP.2016.2528042](https://doi.org/10.1109/TIP.2016.2528042).
- [57] Yu J, Wang Z, Majumdar A, Rajagopal R. DeepSolar: a machine learning framework to efficiently construct a solar deployment database in the United States. *Joule* 2018;2(12):2605–17. doi:[10.1016/j.joule.2018.11.021](https://doi.org/10.1016/j.joule.2018.11.021).
- [58] Yuan J, Yang HHL, Omaitaou OA, Bhaduri BL. Large-scale solar panel mapping from aerial images using deep convolutional networks. In: Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016; 2016. p. 2703–8. doi:[10.1109/BIGDATA.2016.7840915](https://doi.org/10.1109/BIGDATA.2016.7840915).
- [59] Zhang C, Campana PE, Yang J, Yan J. Analysis of distributed photovoltaic financing: a case study approach of crowd-funding with photovoltaic water pumping system in microgrids. *Energy Proc* 2016;103:387–93. doi:[10.1016/j.egypro.2016.11.304](https://doi.org/10.1016/j.egypro.2016.11.304).
- [60] Zhang H, Yan J, Yu Q, Obersteiner M, Li W, Chen J, Zhang Q, Jiang M, Wallin F, Song X, Wu J, Wang X, Shibasaki R. 1.6 Million transactions replicate distributed PV market slowdown by COVID-19 lockdown. *Appl Energy* 2021. doi:[10.1016/j.apenergy.2020.116341](https://doi.org/10.1016/j.apenergy.2020.116341).
- [61] Zhang Z, Li W, Li B. An improving technique of color histogram in segmentation-based image retrieval. In: Proceedings of the 5th International Conference on Information Assurance and Security, IAS 2009, 2; 2009. p. 381–4. doi:[10.1109/IAS.2009.156](https://doi.org/10.1109/IAS.2009.156).
- [62] Zhao Y, Han R, Rao Y. A new feature pyramid network for object detection. In: Proceedings of the 2019 International Conference on Virtual Reality and Intelligent Systems, ICVRIS 2019; 2019. p. 428–31. doi:[10.1109/ICVRIS.2019.00110](https://doi.org/10.1109/ICVRIS.2019.00110).
- [63] Chen Z, Liu J, Zhang Z, Wang F, Chai H, Yu Y, Lu X, Wang T, Lin Y. Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image. *IEEE Trans Ind Appl* 2020;56(4):3385–96. doi:[10.1109/TIA.2020.2984617](https://doi.org/10.1109/TIA.2020.2984617).
- [64] Zheng J, Zhang H, Yin L, Liang Y, Wang B, Li Z, Song X, Zhang Y. A voyage with minimal fuel consumption for cruise ships. *J CleanProd* 2019;215:144–53. doi:[10.1016/j.jclepro.2019.01.032](https://doi.org/10.1016/j.jclepro.2019.01.032).
- [65] Zhong T, Zhang K, Chen M, Wang Y, Zhu R, Zhang Z, Zhou Z, Qian Z, Lv G, Yan J. Assessment of solar photovoltaic potentials on urban noise barriers using street-view imagery. *Renew Energy* 2021;168:181–94. doi:[10.1016/j.renene.2020.12.044](https://doi.org/10.1016/j.renene.2020.12.044).
- [66] Zhuang L, Zhang Z, Wang L. The automatic segmentation of residential solar panels based on satellite images: a cross learning driven U-Net method. *Appl Soft Comput* J 2020;92:106283. doi:[10.1016/j.asoc.2020.106283](https://doi.org/10.1016/j.asoc.2020.106283).