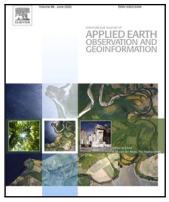
International Journal of Applied Earth Observations and Geoinformation 116 (2023) 103098

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



SolarNet: A convolutional neural network-based framework for rooftop solar potential estimation from aerial imagery

Qingyu Li [a], Sebastian Krapf [b], Yilei Shi [c], Xiao Xiang Zhu [a,*]

[a] Data Science in Earth Observation, Technical University of Munich, Munich, 80333, Germany
[b] Institute of Automotive Technology, Technical University of Munich, Garching, 85748, Germany
[c] Remote Sensing Technology, Technical University of Munich, Munich, 80333, Germany

ARTICLE INFO

Keywords:
Solar potential
Renewable energy
Roof segments and orientations
Convolutional neural network
Remote sensing

ABSTRACT

Solar power is a clean and renewable energy source. Promoting solar technology can not only offer all people affordable, reliable, and modern energy, but also mitigate energy-related emissions and pollutants. This significantly contributes to sustainable development goals. Aerial imagery can provide a cost-effective way for large-scale rooftop solar potential analysis when compared to other data sources. Existing studies mainly utilize aerial imagery and convolutional neural networks to learn the roof segmentation mask or the rooftop geometry map, which are the preliminary input for rooftop solar potential estimation. However, these methods fail to achieve precise solar potential analysis results. To address this issue, we propose a framework, which is termed as SolarNet for rooftop solar potential estimation. A novel multi-task learning network is devised in SolarNet to learn our proposed novel representation for rooftop geometry that incorporates 6 roof segments and orientations. Specifically, this network first learns a roof segmentation map, and then together with the extracted multiscale and contextual features to learn a roof geometry map. Finally, the solar potential can be estimated from the learned roof geometry map. The effectiveness of SolarNet is validated on two datasets: DeepRoof and RID datasets. Experimental results demonstrate that SolarNet can improve not only rooftop geometry prediction accuracy but also solar potential estimation precision, which significantly outperforms other competitors.

1. Introduction

In 2015, the united nations (UN) initiated an agenda with 17 sustainable development goals (SDGs), in order to promote sustainable development and prosperity of our planet. According to a report from the UN (United Nations, 2021), urban areas host 55.13% of the world's population. Moreover, it is anticipated that this proportion will increase to 60% in 2030. Solar energy provides great potential to satisfy the growing energy demand due to advanced technology and reduced costs. This also coincides with sustainable development goal (SDG) 7, which aims at offering all people affordable, reliable, and modern energy. Furthermore, solar energy is a renewable energy source and has fewer carbon emissions than fossil fuels. By gradually replacing fossil fuels with solar energy, energy-related emissions and pollutants can be mitigated. Therefore, promoting solar technology as well contributes to SDG 13, aiming to achieve global greenhouse gas neutrality.

Building rooftops are ideal locations for mounting solar energy deployments from two aspects. On one hand, rooftop solar deployments are able to leverage unused space, and overcome the limitations of certain land unavailability or high land rent in cities (Aslani and Seipel, 2022; Gassar and Cha, 2021). On the other hand, rooftop solar systems create a new role for each building, which is an active power-generator (Sánchez-Aparicio et al., 2021; Walch et al., 2020). This is helpful to establish net-zero energy buildings (Sharma et al., 2020). Therefore, the task of rooftop solar potential estimation that identifies suitable buildings for the installation of rooftop solar deployments is of great significance in offering insights for policymakers about sustainable energy plans (Fakhraian et al., 2021), e.g. energy storage and trading.

In the existing literature, mainstream strategies for rooftop solar potential analysis are to first assess the available rooftop areas (Izquierdo et al., 2008). Afterward, solar potential can be estimated from a bottom-up perspective (Korfiati et al., 2016; Assouline et al., 2017). With respect to the utilized data, existing strategies for rooftop area estimation can be categorized into four types: (1) three-dimensional (3D) building model-based, (2) statistical data-based, (3) geospatial data-based, and (4) aerial imagery-based methods. 3D building model-based

* Corresponding author.
E-mail addresses: qingyu.li@tum.de (Q. Li), sebastian.krapf@tum.de (S. Krapf), yilei.shi@tum.de (Y. Shi), xiaoxiang.zhu@tum.de (X.X. Zhu).

https://doi.org/10.1016/j.jag.2022.103098
Received 20 August 2022; Received in revised form 26 October 2022; Accepted 4 November 2022
Available online 9 December 2022

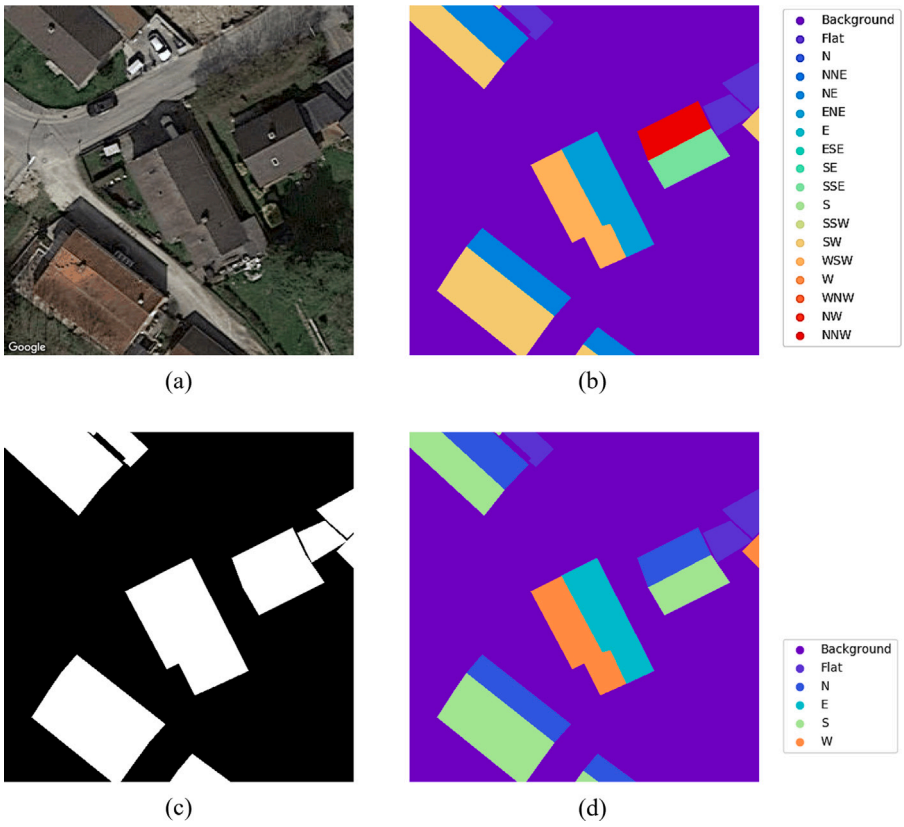


Fig. 1. Sample data for (a) aerial imagery. (b) roof geometry map with 18 classes (Lee et al., 2019). (c) roof segmentation mask. (d) roof geometry map with our proposed 6 classes.

methods (Mohajeri et al., 2018; Lingfors et al., 2017) utilize light detection and ranging (LiDAR) data or stereo photos to acquire 3D building models, which are able to capture precise rooftop geometry and quantify the shadows cast by other buildings and vegetation canopies (Fogl and Moudrỳ, 2016). With the increasing availability of open LiDAR data (Melin et al., 2017), some studies (Margolis et al., 2017; Suomalainen et al., 2017) utilize these data for solar potential analysis. However, retrieving 3D building structures is highly expensive when there are no open LiDAR data or stereo photos. In contrast, the other three types of methods are more cost-effective to provide building information for large spatial coverage. In statistical data-based methods, (Miranda et al., 2015; Ordóñez et al., 2010), statistical data such as census data is exploited to estimate rooftop area with the assumption of the available roof area per capita. Geospatial data-based methods (Sun et al., 2013; Izquierdo et al., 2008) aim at obtaining available rooftop areas from building cadastral maps. Nevertheless, both two types of methods suffer from a common limitation that the utilized data is usually out-of-date (Li et al., 2022b) and is of the coarse resolution, leading to poor accuracy (Joshi et al., 2021). In contrast, aerial imagery can break through this limitation and is able to allow up-to-date and high-resolution rooftop area quantification.

Most existing methods (Lee et al., 2019; Huang et al., 2019; Krapf et al., 2021) assess rooftop solar potential by detecting the roof segmentation mask (c.f. Fig. 1(c)) from aerial imagery. Nevertheless, the detailed structures (e.g., tilts and orientations of roof segments) are neglected, leading to an overestimation of solar potential. Recently, a study from (Lee et al., 2019) proposes to take individual roof segments and their orientation into consideration, as the solar irradiation received by roof segments is dependent on their orientations. Therefore, the orientations of roof segments will also exert a great influence on the generated solar power. By taking orientations of roof segments into account during the solar potential analysis, not only ideal locations for solar deployments can be identified, but also overestimation of solar potential can be avoided. Specifically, in Lee et al. (2019), convolutional neural networks are implemented to learn 18 classes (c.f. Fig. 1(b)), including roof segments and their orientations. However, it is time-consuming and laborious to prepare such complicated annotation data for network training. Furthermore, the imbalanced class distribution also results in a biased network learning (Krapf et al., 2021). In this work, we want to learn a simplified representation of geometrical structures for building rooftops. Therefore, we propose a novel representation of rooftop geometry that is composed of six classes (c.f. Fig. 1(d)). By doing so, not only the class imbalance issue is alleviated, but also the annotation workload can be largely reduced.

In this study, we propose a framework, namely SolarNet, to first learn this novel representation of rooftop geometry and then carry out a solar potential analysis. Moreover, we note existing networks neglect the building roof semantic information when learning rooftop geometry. Therefore, we devise a novel multi-task learning network that learns both roof segmentation masks (c.f. Fig. 1(c)) and roof geometry maps (c.f. Fig. 1(d)). Roof segmentation masks indicate the semantic information of rooftops, enabling the network to not only preserve the overall information of the building roof but also better suppress the background clutters (e.g., road, car, and tree). By allocating segmentation maps on multiscale and contextual features, rooftop geometry could be better learned.

The contributions of this article are threefold:

(1) In this study, a novel framework, termed as SolarNet, is proposed for rooftop solar potential analysis. SolarNet shows great potential in the large-scale rooftop solar potential estimation tasks.

(2) We propose a novel representation for rooftop geometry, which consists of 6 different classes, including background, roof segments, and their orientations. By doing so, not only the estimation accuracy of solar potential can be improved, but also the annotation workload can be alleviated.

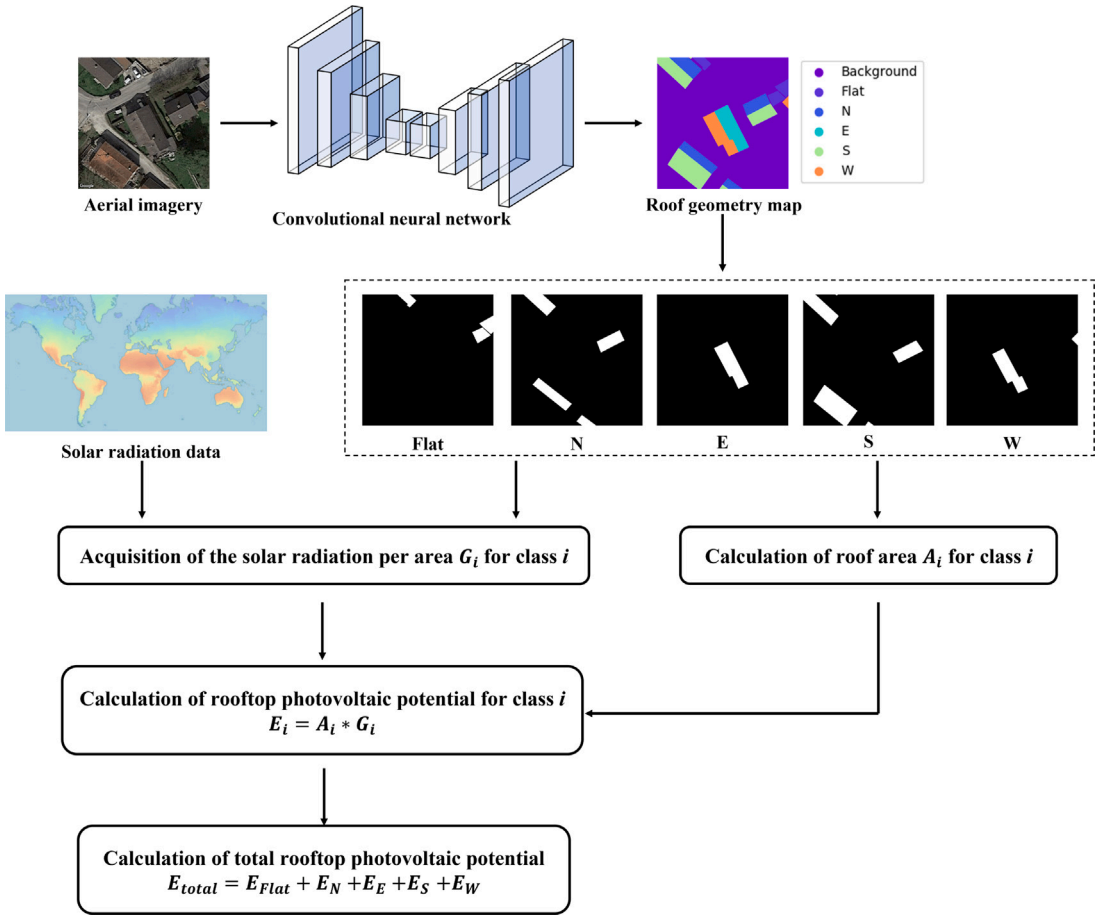


Fig. 2. The flowchart of the proposed framework.

(3) We propose a novel multi-task learning network in the SolarNet to learn this novel representation of rooftop geometry, which is further utilized to carry out rooftop solar potential analysis. Specifically, roof segmentation masks are first learned and then combined with multiscale contextual features for the learning of rooftop geometry.

2. Study area and data

To our knowledge, only two benchmark datasets for the semantic segmentation of roof segments are released in the existing literature, i.e., the DeepRoof (Lee et al., 2019) and the RID (Krapf et al., 2022) datasets.

The DeepRoof dataset collects aerial images from six cities in the United States, and more than 85% of images come from Framingham, Massachusetts, Pinellas Park, and Florida. The RID dataset captures aerial imagery from a German city, WartenBerg .

The DeepRoof dataset labels 2274 buildings, while the RID dataset has 1880 annotated buildings. For both two datasets, 18 classes are defined to depict the rooftop geometry, which includes a background class, a flat roof class, and 16 azimuth classes. A flat roof has a slope of 0°, and its azimuth is not defined. The categorization into azimuth classes depends on the roof segment's orientation. There are four main directions east (E), south (S), west (W), and north (N) as well as four intermediate directions, i.e. south-west (SW), south-east (SE), north-west (NW), and north-east (NE). In addition, there are eight orientations in between, for example, north-north-east (NNE) or south-west-west (SWW). In Section 3.2.1, the azimuth angles and corresponding classes are described in more detail. Note that for both RID and DeepRoof datasets, the ground reference solar potential is estimated based on the ground reference rooftop geometry classes.

3. Methodology

In this section, the pipeline of SolarNet is first presented. Then, the proposed representation of rooftop geometry and multi-task learning network is introduced in detail. Afterward, the procedure of rooftop solar potential estimation is described. Finally, the selected metrics for the evaluation of SolarNet are introduced.

3.1. Pipeline

The workflow of SolarNet is illustrated in Fig. 2. A convolutional neural network (CNN) is first adopted to learn the rooftop geometry from aerial imagery, where roof geometry map encodes roof segments and their orientations. Next, individual classes are extracted for the calculation of rooftop solar potential. Specifically, the class-specific roof area is calculated from the predicted roof geometry map. Then, for each class, the roof area is multiplied by its geographic solar potential, which is acquired from a solar radiation database for each azimuth class. Finally, the total solar potential is derived as the summation of potential from all classes.

3.2. Identification of roof segments and orientations

3.2.1. Definition of roof geometry

Roof orientations play an important role in solar potential analysis. A function of slope and azimuth (c.f. Fig. 3) can be utilized to describe the mean annual solar irradiation received by roof segments. For instance, in regions of the northern hemisphere, the yearly energy generation of a north-oriented segment can be 50% lower than a south-oriented segment.

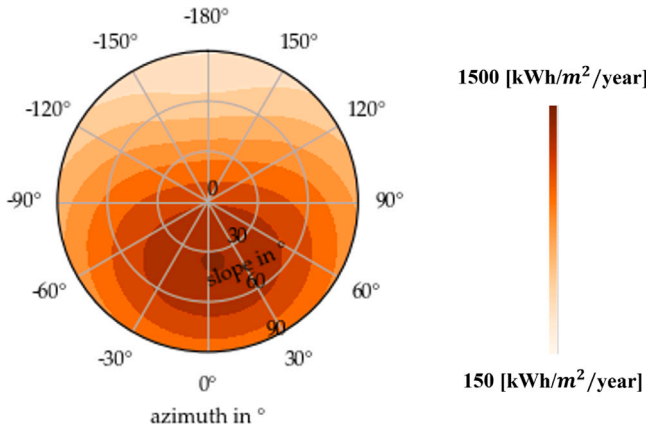


Fig. 3. Mean annual solar irradiation received by rooftops for different ranges of azimuth and slope.

It is obvious that a greater number of azimuth classes are able to obtain a more precise depiction of rooftop geometry. However, a larger number of classes might impede network learning (e.g. the class imbalance in RID and DeepRoof datasets). Fig. 4 displays the class distributions of RID and DeepRoof datasets, which are given as the frequency of pixels per azimuth class and the additional flat roof class. For instance, in the DeepRoof dataset, flat roofs cover more than 10% of the images, and this proportion is much higher than that of other classes. Besides, both datasets display a bias towards four azimuth classes. In the DeepRoof dataset, roof orientations (N, S, E, and W) have similar proportions, which are larger than other azimuth classes. By contrast, the buildings in RID are mostly NNW, SSE, ENE, and WSW oriented. The skewed class distribution in both datasets is due to architectural reasons, as buildings are usually constructed to increase the inlet of daylight.

Most existing methods (Lee et al., 2019; Krapf et al., 2021, 2022) learns 18 classes, consisting of 16 azimuth classes (S, SE, SSE, etc.) as well as a flat and background class, which not only suffer from the imbalanced distribution among classes but also require a heavy workload for data annotation. Therefore, a novel representation is proposed in this research to depict rooftop geometry. Specifically, we propose to use 4 azimuth classes rather than the 16 azimuth classes used in the existing literature. Fig. 5 visualizes, how classes are defined with respect to the azimuth angle. The 16 classes span over a range of 22.5° for each class, while the 4 classes are defined as ranges of 90°. It can be seen, that some of the classes are split up during the transformation from 16 classes to 4 classes, e.g. SW which is distributed to S and W. Figs. 1(b) and 1(d) visualize ground reference masks with 16 and 4 azimuth classes for a sample of an aerial image from the RID dataset, respectively. With our proposed representation of rooftop geometry, the issues related to class imbalance can be eased (see Fig. 6).

3.2.2. Multi-task learning network

The identification of roof segments and orientations has two goals. One is to extract roof segments from aerial images, and the other is to assign a class with respect to the segment's orientation. In this study, a multi-task learning network is devised to first learn building roof segmentation masks, which are then utilized to learn rooftop geometry map. This is because building roof masks can not only suppress background clutters but also make roof regions more discernible, which guarantees the high accuracy of segmentation results of rooftop geometry.

The proposed multi-task learning network is an encoder–decoder architecture (Fig. 7). In the encoder, the spatial resolution of the input is downsampled to generate semantic features. Afterward, the features are upsampled to a full-resolution prediction result in the decoder. Specifically, the encoder is a ResNet34 model (He et al., 2016) that consists of four stages. Each stage is appended with one basic ResNet block (BRB), where the goal is to learn residual connections across layers. The outputs of the four BRBs are upsampled to the same size of the aerial image, respectively, and then concatenated for the learning the roof segmentation map. An astrous spatial pyramid pooling (ASPP) (Chen et al., 2017) module is linked between the encoder and decoder, and this module can capture multiscale contextual information, which helps to upgrade the capability of the network in perceiving targets at varying scales. Next, the roof segmentation mask-enhanced features can be obtained by allocating the learned roof segmentation map to the features that are extracted from the last block of the decoder. To produce the final rooftop geometry map, the enhanced feature maps are fed into a convolutional layer. Moreover, features from both the encoder and decoder are concatenated by skip connections, allowing signals to propagate in a more efficient way.

Our network aims to learn both the roof segmentation map and the roof geometry map. The global loss function L of this multi-task learning network is defined as follows:

$$L = \lambda \cdot L_S + L_G\,, \tag{1}$$

where L_S and L_G are two cross-entropy loss functions for optimizing the roof segmentation map and the roof geometry map, respectively. λ is a hyperparameter representing the importance of the first term. In this research, λ is empirically set as 0.1.

3.3. Analysis of rooftop solar potential

In this research, we carry out rooftop solar potential analysis by estimating the geographical solar potential for each roof segment, i.e., the solar radiation received on a tilted plane at a given location. PVGIS (Huld et al., 2012) is a solar energy database, which provides the annual solar radiation data (Wh/m^2/year) by a given slope and azimuth at a certain geographic location. We acquire the solar radiation per area G_i for the class i with a look-up table from PVGIS, with i referring to the flat roof class or azimuth classes. Note that for non-flat classes, the slope is defined as 30°. To determine the solar potential per roof segment, the generated rooftop geometry map from the multi-task learning network is first transformed into a vector representation, and the area A_i of the class i is calculated. Combined the solar radiation per area G_i (Wh/m^2/year) with area A_i, the total amount of rooftop solar potential E_i (Wh/year) for the class i is derived by a multiplication,

$$E_i = G_i * A_i\,. \tag{2}$$

Finally, the total solar potential E_{total} is calculated by summing up each class's potential. Fig. 8 is an example to exhibit the relation between annual global solar irradiation incident on building rooftops and their corresponding solar potential.

$$E_{total} = E_{Flat} + E_N + E_E + E_S + E_W\,. \tag{3}$$

3.4. Experiment setup

For a comprehensive evaluation, SolarNet is compared with state-of-the-art approaches from two perspectives, rooftop geometry prediction accuracy, and rooftop solar potential estimation precision. Specifically, as to the former, we compare the proposed multi-task learning network with four semantic segmentation models, U-Net (Ronneberger et al., 2015), Efficient-UNet (Baheti et al., 2020), FC-DenseNet (Jégou et al., 2017), and DeepLab V3+ (Chen et al., 2018). Regarding the task of rooftop solar potential estimation, we perform comparisons with the representation of rooftop geometry that is defined by 18 classes (Lee et al., 2019). Note that here for a fair comparison in rooftop solar potential estimation, we take the multi-task learning network as the backbone for learning different representations of rooftop geometry.

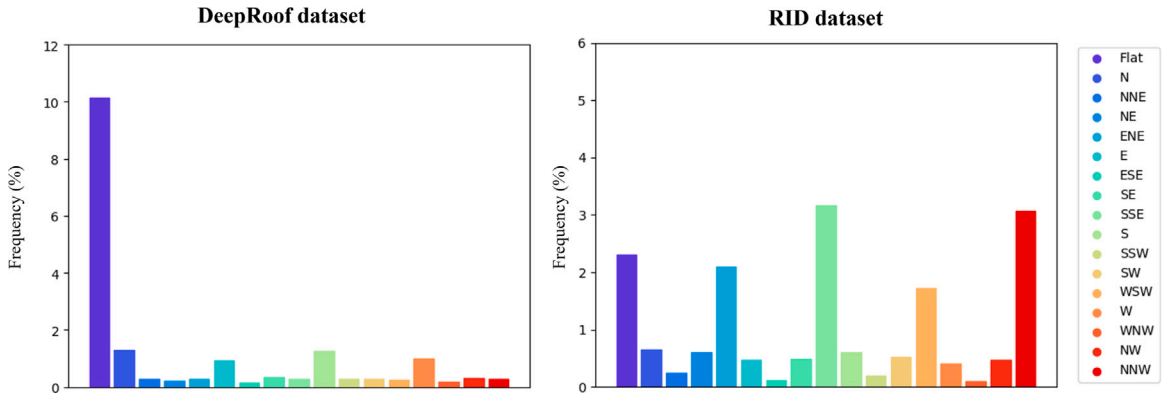


Fig. 4. The class distribution of roof geometry defined by 18 classes in DeepRoof and RID datasets, respectively.

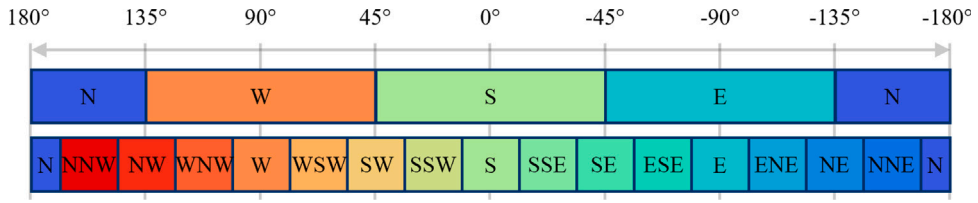


Fig. 5. Azimuth classification into 4 and 16 classes.

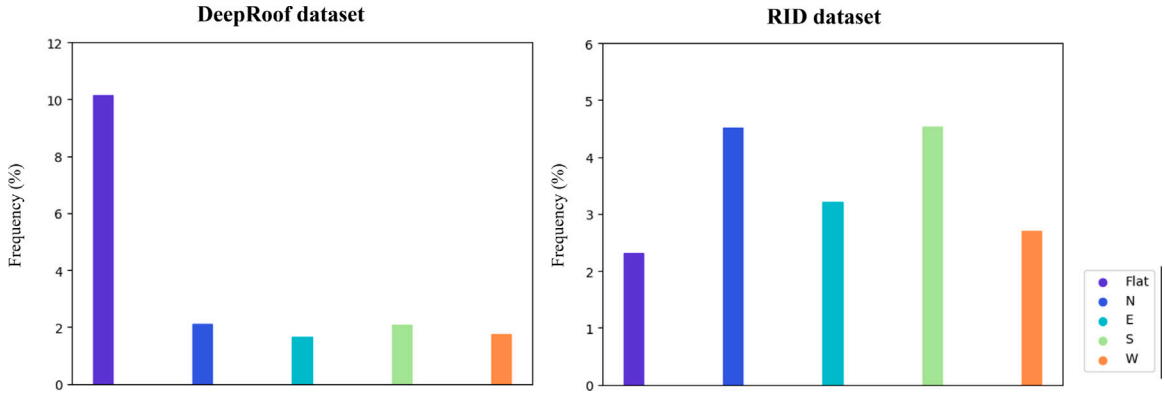


Fig. 6. The class distribution of roof geometry defined by 6 classes in DeepRoof and RID datasets, respectively.

3.5. Accuracy assessment

The performance of models is evaluated from two aspects, i.e., rooftop geometry prediction accuracy and solar potential estimation precision.

3.5.1. Rooftop geometry prediction accuracy

To evaluate the predicted roof segments and orientations, intersection over union (IoU) is first computed for each category with the following equation:

$$\text{IoU} = \frac{TP}{TP + FP + FN}, \tag{4}$$

where TP, FP, and FN indicate the numbers of true positives, false positives, and false negatives, respectively. Afterward, the mean IoU (mIoU) is computed by averaging the IoU metrics of all classes. Notably, a large IoU score indicates high accuracy. Moreover, overall accuracy (OA), the percentage of pixels that are correctly classified, is also derived for the assessment of prediction results.

3.5.2. Rooftop solar potential estimation precision

To demonstrate the effectiveness of the proposed framework in solar potential estimation, percent root mean squared error (%RMSE) is exploited as the evaluation criterion. It is defined as:

$$\%\text{RMSE} = \frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}(E_j - \hat{E}_j)^2}}{\frac{1}{n}\sum_{j=1}^{n} E_j}, \tag{5}$$

where $\hat{E}_j$ and E_j are solar potential values of jth image in estimated results and ground reference, respectively, and n denotes the number of images for evaluation.

4. Results

4.1. Results of rooftop geometry prediction

For the data split in both RID and DeepRoof datasets, the ratio of train:validation:test is set as 7:1:2. Numerical results of rooftop geometry prediction on the test set of two datasets are shown in Table 1. We can observe that SolarNet achieves improved results in terms of mIoU and OA. Specifically, compared to DeepLab V3+, SolarNet obtains increments of 14.84% in mIoU on the RID dataset and 19.48% in mIoU on the DeepRoof dataset. Compared to the commonly used U-Net (Ronneberger et al., 2015), our approach also reaches improvements of 4.73% in mIoU on the RID dataset and 2.58% in mIoU on the

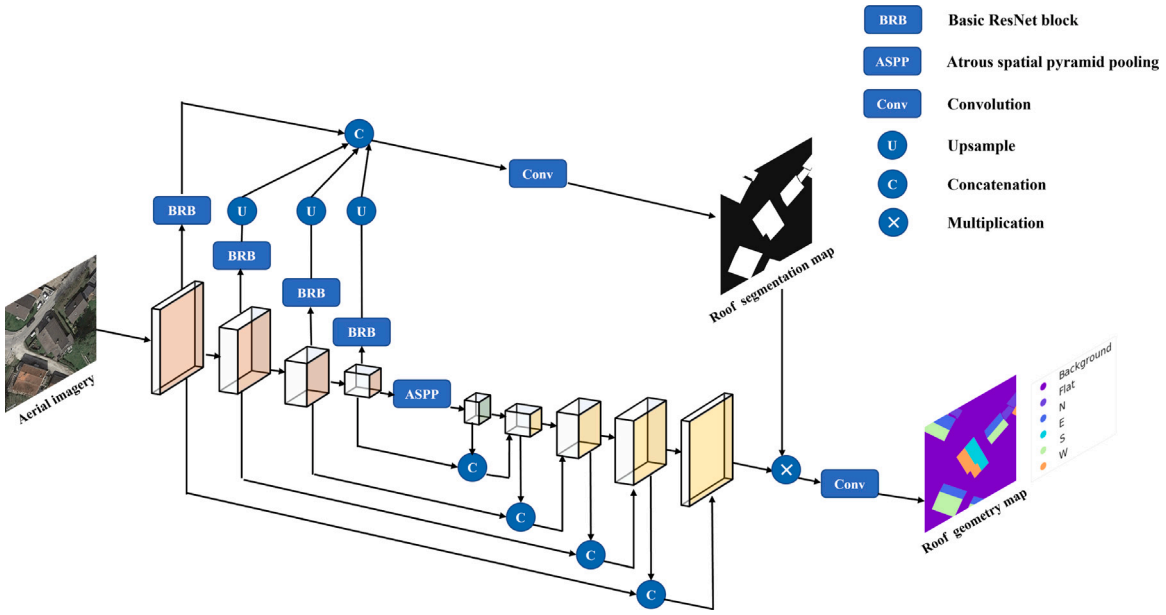


Fig. 7. The flowchart of the proposed multi-task learning network.

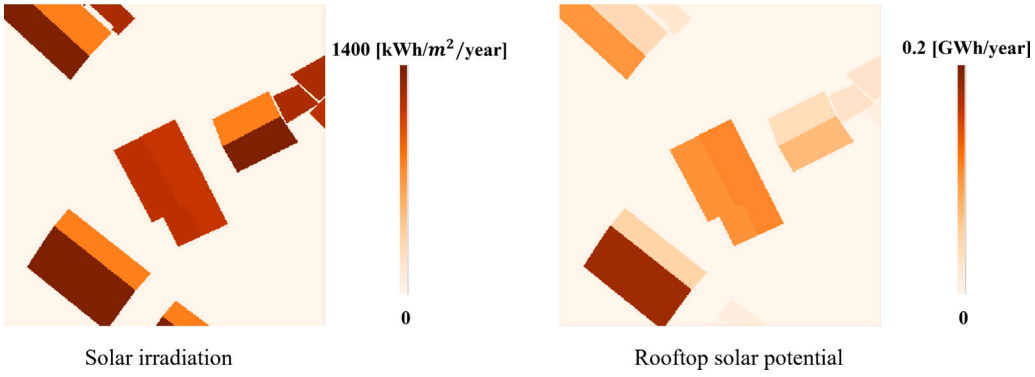


Fig. 8. Annual global solar irradiation incident on building rooftops and their corresponding solar potential.

Table 1
Accuracy metrics of different semantic segmentation networks on both the RID and DeepRoof datasets. (%)

Method	RID dataset		DeepRoof dataset	
	mIoU	OA	mIoU	OA
U-Net (Ronneberger et al., 2015)	63.86	94.66	57.12	91.90
FC-DenseNet (Jégou et al., 2017)	56.58	93.59	35.93	87.69
Efficient-UNet (Baheti et al., 2020)	57.66	93.66	47.36	90.54
DeepLab V3+ (Chen et al., 2018)	53.75	92.05	40.22	87.96
SolarNet	**68.59**	**95.55**	**59.70**	**92.96**

DeepRoof dataset. This indicates that the proposed multi-task learning network is effective in extracting roof segments and orientations.

The accuracy of individual classes on RID and DeepRoof datasets are shown in Tables 2 and 3, respectively. It can be observed that differences exist in the IoU of various classes. For instance, the class of "Flat" has the lowest IoU among all classes on the RID dataset. And on the DeepRoof dataset, all methods show relatively worse results for the classes of "E" and "W". One possible reason might be that these classes have relatively smaller numbers of pixels in the training data (see Fig. 6). In this case, the model performance is largely affected due to the limited supervisory information for network learning.

Figs. 9 and 10 present several examples of rooftop geometry maps generated by various networks on both RID and DeepRoof datasets. Other networks easily misclassify the class of "W" as other categories, whereas SolarNet suffers less from such errors. Moreover, SolarNet is able to alleviate more false alarms and recover more precise outlines of roof segments in comparison with other approaches. This is due to its multi-task learning network, which can leverage the roof segmentation map for learning precise rooftop geometry. By utilizing the roof masks related information, background clutters can be suppressed and roof regions can be more discernible.

4.2. Results of solar potential estimation

In order to provide a rough idea of the trade-off between less accurate estimates of solar potential and higher accuracy of rooftop geometry, we have investigated the difference between estimated solar potential using ground reference data with 18 and 6 roof segments and orientations. Specifically, we first estimate the solar potential by using the ground reference data with 6 roof segments and orientations. Afterward, we calculate the %RMSE of it with respect to the ground reference solar potential. According to the calculation on both two datasets, the RID dataset shows 1.62% in %RMSE, while %RMSE for DeepRoof is 0.55%. This negligible difference suggests that although a greater number of azimuth classes can acquire a more precise depiction of rooftop geometry, its superiority in solar potential estimation results is not so evident.

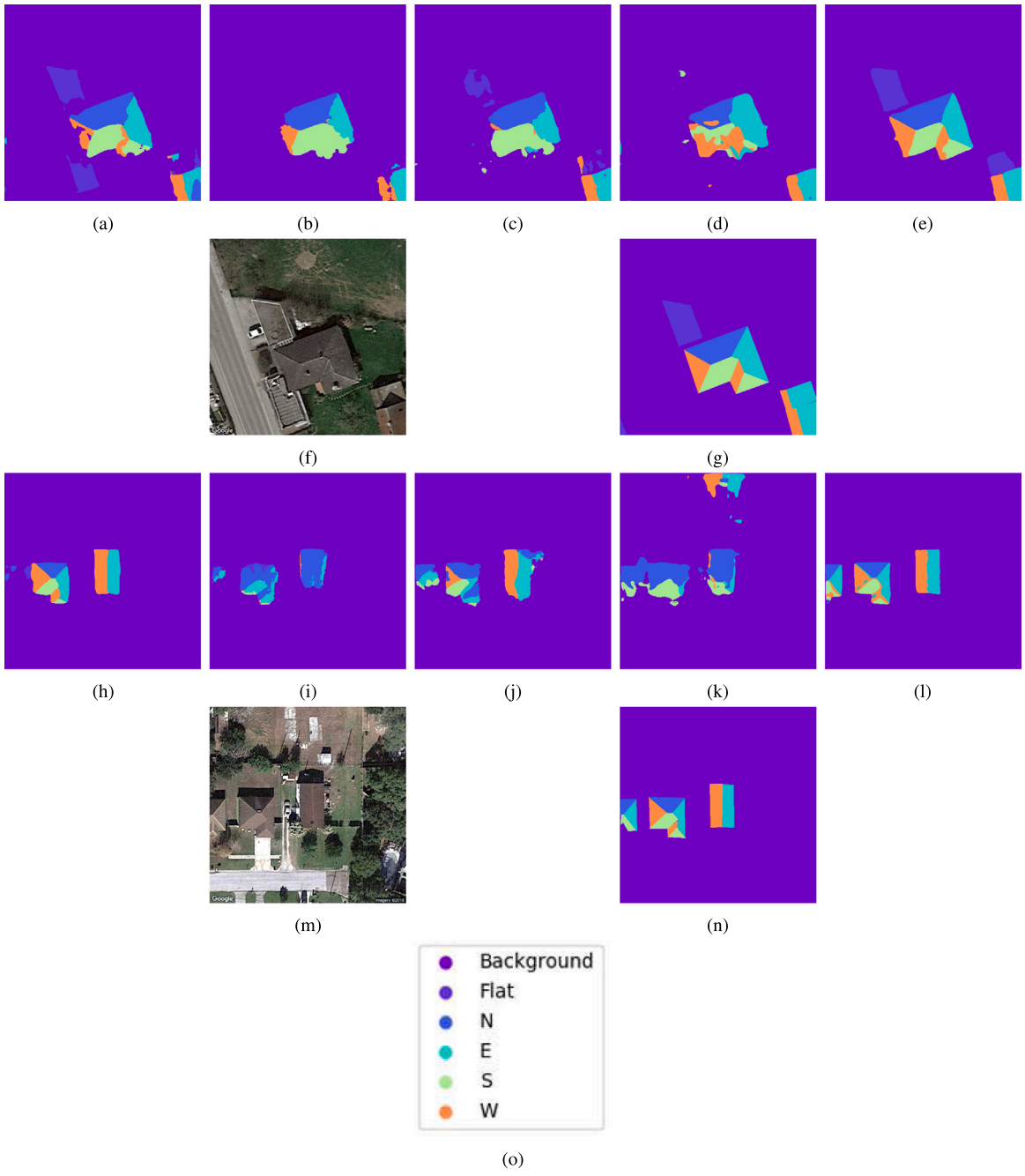


Fig. 9. Examples of rooftop geometry segmentation results obtained by different methods. (a) U-Net (Ronneberger et al., 2015). (b) FC-DenseNet (Jégou et al., 2017). (c) Efficient-UNet (Baheti et al., 2020). (d) DeepLab V3+ (Chen et al., 2018). (e) SolarNet. (h) U-Net (Ronneberger et al., 2015). (i) FC-DenseNet (Jégou et al., 2017). (j) Efficient-UNet (Baheti et al., 2020). (k) DeepLab V3+ (Chen et al., 2018). (l) SolarNet. (f) and (g) are aerial imagery and ground reference from the RID dataset. (m) and (n) are aerial imagery and ground reference from the DeepRoof dataset. (o) is the legend for different classes of roof segments and orientations.

Table 2
IoU of individual classes that are obtained from different semantic segmentation networks on the RID dataset. (%)

Method	Background	Flat	N	E	S	W
U-Net (Ronneberger et al., 2015)	96.32	25.86	76.16	59.33	71.09	54.40
FC-DenseNet (Jégou et al., 2017)	95.89	22.60	69.27	45.64	61.42	44.65
Efficient-UNet (Baheti et al., 2020)	95.79	19.69	71.13	49.54	64.89	44.91
DeepLab V3+ (Chen et al., 2018)	94.67	24.65	62.81	40.57	59.86	39.95
SolarNet	**97.15**	**34.63**	**81.19**	**65.27**	**74.33**	**59.00**

To further validate the performance of the proposed framework in solar potential estimation, we report experimental results on both datasets (c.f. Table 4). When compared to the representation of 18 roof segments and orientations, that of 6 roof segments and orientations contributes to the reduction in %RMSE of solar potential estimation on both datasets. Especially for the DeepRoof dataset, our proposed

Table 3
IoU of individual classes that are obtained from different semantic segmentation networks on the DeepRoof dataset. (%)

Method	Background	Flat	N	E	S	W
U-Net (Ronneberger et al., 2015)	92.58	56.18	56.73	42.69	50.54	44.00
FC-DenseNet (Jégou et al., 2017)	90.13	45.91	35.37	10.82	27.27	6.09
Efficient-UNet (Baheti et al., 2020)	91.94	52.37	45.34	27.42	38.58	28.52
DeepLab V3+ (Chen et al., 2018)	89.47	49.73	33.95	18.54	30.56	19.09
SolarNet	**93.94**	**61.30**	**60.32**	**45.25**	**56.71**	**40.65**

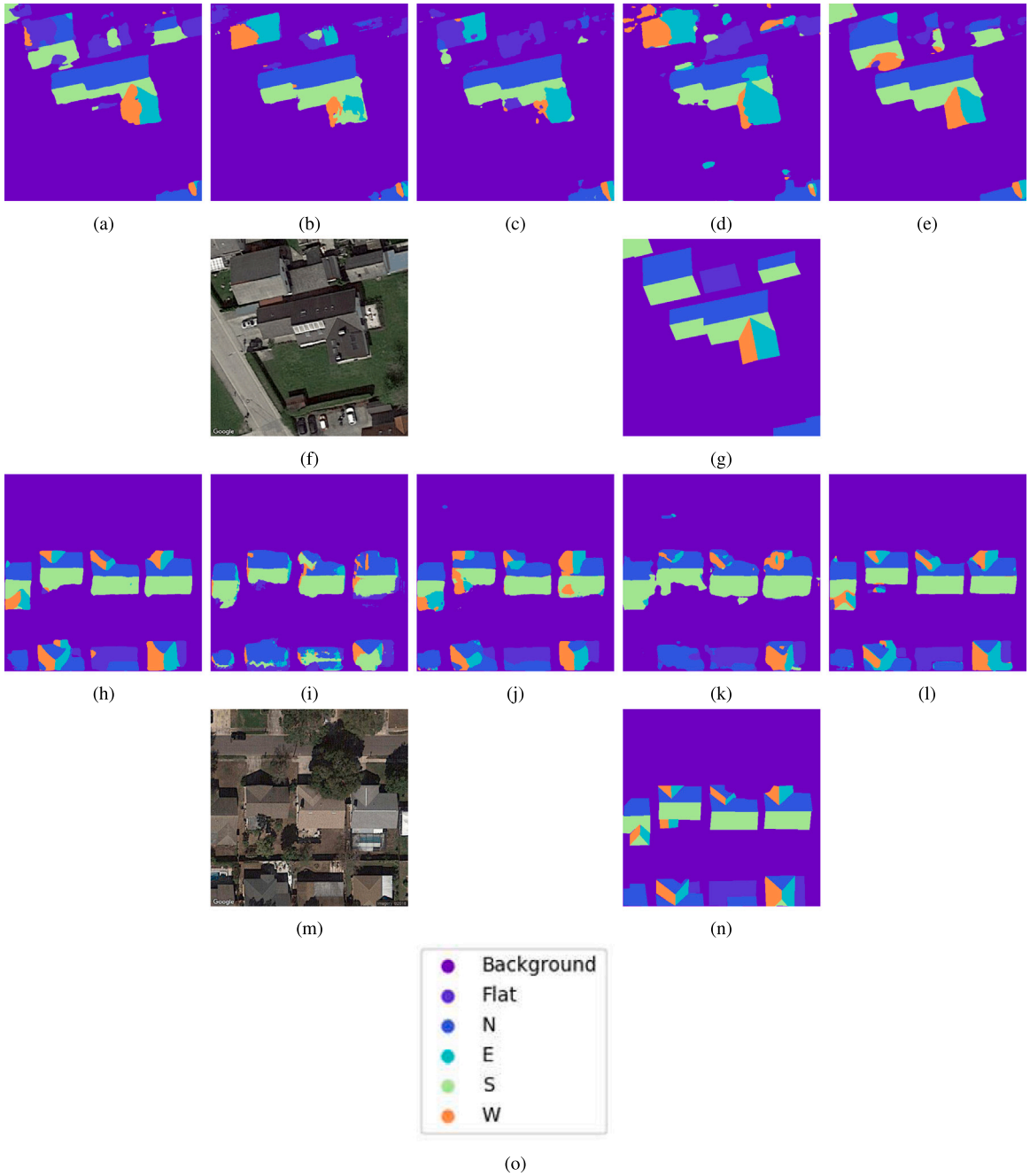


Fig. 10. Examples of rooftop geometry segmentation results obtained by different methods. (a) U-Net (Ronneberger et al., 2015). (b) FC-DenseNet (Jégou et al., 2017). (c) Efficient-UNet (Baheti et al., 2020). (d) DeepLab V3+ (Chen et al., 2018). (e) SolarNet. (h) U-Net (Ronneberger et al., 2015). (i) FC-DenseNet (Jégou et al., 2017). (j) Efficient-UNet (Baheti et al., 2020). (k) DeepLab V3+ (Chen et al., 2018). (l) SolarNet. (f) and (g) are aerial imagery and ground reference from the RID dataset. (m) and (n) are aerial imagery and ground reference from the DeepRoof dataset. (o) is the legend for different classes of roof segments and orientations.

representation (6 roof segments and orientations) obtains a decrease of 4.71% in %RMSE. This again demonstrates the effectiveness and robustness of our approach in the task of rooftop solar potential estimation. Although fewer classes are defined to depict the rooftop

Fig. 11. Zoomed-in results of rooftop solar potential estimation for a sample urban area.

Table 4
Numerical results of rooftop solar potential on the RID dataset and the DeepRoof dataset. (%)

Representation	RID dataset	DeepRoof dataset
	%RMSE	%RMSE
18 roof segments and orientations	17.02	25.82
6 roof segments and orientations	**15.30**	**21.10**

geometry, the proposed representation of 6 roof segments and orientation is able to achieve higher solar potential estimation precision in comparison with the other representation. The superiority of the proposed representation might be benefited from more accurate rooftop geometry prediction results. Fig. 11 illustrates a zoom-in visual example where rooftop solar potential estimation results are obtained by our framework. The roof segments in darker colors represent high solar potential. This confirms that our framework is promising to support large-scale rooftop solar potential analysis.

5. Discussion

In this study, a novel framework, namely SolarNet, is proposed to take advantage of aerial imagery and a CNN model for rooftop solar potential estimation. Specifically, we propose to define rooftop geometry with six classes, which can be learned by a novel devised multi-task learning network. By doing so, we are able to not only achieve more accurate rooftop geometry maps but also improve the solar potential estimation precision.

In the following, we critically discuss these main outcomes of our study from four perspectives — data, methodology, time, and geography:

(1) **Data**: In this study, we make use of two public datasets including aerial imagery and the ground reference rooftop geometry, which permits satisfactory performance in rooftop geometry prediction and solar potential estimation. Nevertheless, the collection of such datasets is a dominant obstacle in large-scale applications. To address this issue, open access remote sensing imagery, e.g., Google Earth aerial imagery with sub-meter resolution (Qi and Wang, 2014) can be leveraged as an alternative data source, as they cover the whole globe and have a high update speed. Still, for the ground reference data preparation, we can utilize CityGML (Kolbe, 2009) data where the geometry of 3D building models are well characterized. This is helpful to acquire roof segments and orientations as the ground reference data for network learning.

(2) **Methodology**: In this research, the solar potential estimation relies heavily on rooftop geometry maps that are provided by the CNN. Therefore, related methodological challenges arising from CNN models need to be considered. According to the quantitative assessment of the accuracy of the rooftop geometry prediction, the proposed multi-task learning network achieves the mIoU of 68.59% and the mIoU of 59.70% on RID and DeepRoof datasets, respectively. Of course, uncertainties remain despite the high accuracies, however, the proposed multi-task learning network is still superior to other semantic segmentation networks. With respect to our work, any misclassification of roof segments and orientations introduced by the CNN will lead to questionable solar potential estimation results. For instance, it can be observed that misclassifications are present around the edges of predicted roof segments and orientations. In future research, regularized loss (Li et al., 2022c) can be incorporated in the objective function to guide network learning, improving edge details of the rooftop geometry map. Moreover, one challenge for CNN models on large-scale applications is collecting a large number of annotated samples, which is laborious and time-consuming. In the future, we aim to explore semi-supervised learning

strategies (Li et al., 2022a) that are capable of compensating for the limited supervisory information due to insufficient annotated data.

(3) **Time**: For the task of rooftop solar potential estimation, this study concentrates on acquiring an accurate representation of rooftop geometry shapes. This is because we aim to identify suitable rooftops that can mount solar energy deployments in the local neighborhoods. In other words, the rooftop geometry shape is a more important factor that needs to be taken into account if we want to find rooftops with higher solar energy potential. Nevertheless, in other applications, the effect of time on rooftop solar potential estimation results is undoubtedly significant (Moudrỳ et al., 2019). This is because that variability in weather conditions between years exerts a great influence on solar irradiation. For instance, the accurate estimation of daily or hourly solar potential can avoid excessive fossil fuel consumption and emergency power purchases (Bhola and Bhardwaj, 2016), helping to achieve a balance between renewable and fossil fuel-based energy generation. In future research, we plan to investigate the magnitude of estimated errors that are induced by time.

(4) **Geography**: In our research, we applied the workflow to cities in Germany and the United States, and acquire satisfactory results in both rooftop geometry prediction and solar potential estimation. In this regard, the high transferability of our framework to other geographical regions is expected. However, if structural appearances in building rooftops are more complex, new studies are in demand to test the applicability. Therefore, we encourage the use of our proposed framework for solar potential analysis in less-developed regions where energy and electricity are restricted.

6. Conclusion and outlook

Rooftop-mounted solar systems are able to provide low-cost and clean renewable energy, which helps to alleviate environmental degradation and energy crisis. To this end, we have proposed a framework, namely, SolarNet, which can estimate rooftop solar potential from aerial imagery. Specifically, we have established a novel representation to depict the rooftop geometry and propose a multi-task learning network to learn rooftop geometry information. Finally, rooftop solar potential can be estimated from the learned roof geometry map. We investigate the proposed framework on two datasets: the RID and DeepRoof datasets. Results suggest that the multi-task learning network can achieve a more accurate roof geometry map than other competitors. Moreover, the proposed representation of rooftop geometry can not only improve solar potential estimation precision but also alleviate annotation workload. In the future, we plan to conduct research in various countries around the world to further promote our framework.

CRediT authorship contribution statement

Qingyu Li: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Sebastian Krapf:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – review & editing, Visualization. **Yilei Shi:** Writing – review & editing, Supervision. **Xiao Xiang Zhu:** Resources, Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

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.

Data availability

The authors do not have permission to share data.

Acknowledgments

The work is jointly supported by the TUM Innovation Network Earth Care, by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [ERC-2016-StG-714087], Acronym: *So2Sat*), by the Helmholtz Association, Germany through the Framework of the Helmholtz Excellent Professorship "Data Science in Earth Observation - Big Data Fusion for Urban Research"(grant number: W2-W3-100), by the German Federal Ministry of Education and Research (BMBF), Germany in the framework of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond" (grant number: 01DD20001) and by German Federal Ministry for Economic Affairs and Climate Action, Germany in the framework of the "national center of excellence ML4Earth" (grant number: 50EE2201C). This work is a part of the project "Investigation of building cases using AI" funded by Bavarian State Ministry of Finance and Regional Identity (StMFH) and the Bavarian Agency for Digitization, High-Speed Internet and Surveying, Germany.

References

Aslani, M., Seipel, S., 2022. Automatic identification of utilizable rooftop areas in digital surface models for photovoltaics potential assessment. Appl. Energy 306, 118033.
Assouline, D., Mohajeri, N., Scartezzini, J.-L., 2017. Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. Sol. Energy 141, 278–296.
Baheti, B., Innani, S., Gajre, S., Talbar, S., 2020. Eff-unet: A novel architecture for semantic segmentation in unstructured environment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 358–359.
Bhola, P., Bhardwaj, S., 2016. Solar energy estimation techniques: A review. In: 2016 7th India International Conference on Power Electronics. IICPE, IEEE, pp. 1–5.
Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans. Pattern Anal. Mach. Intell. 40 (4), 834–848.
Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision. ECCV, pp. 801–818.
Fakhraian, E., Forment, M.A., Dalmau, F.V., Nameni, A., Guerrero, M.J.C., 2021. Determination of the urban rooftop photovoltaic potential: A state of the art. Energy Rep. 7, 176–185.
Fogl, M., Moudrỳ, V., 2016. Influence of vegetation canopies on solar potential in urban environments. Appl. Geogr. 66, 73–80.
Gassar, A.A.A., Cha, S.H., 2021. Review of geographic information systems-based rooftop solar photovoltaic potential estimation approaches at urban scales. Appl. Energy 291, 116817.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.
Huang, Z., Mendis, T., Xu, S., 2019. Urban solar utilization potential mapping via deep learning technology: A case study of wuhan, China. Appl. Energy 250, 283–291.
Huld, T., Müller, R., Gambardella, A., 2012. A new solar radiation database for estimating PV performance in Europe and Africa. Sol. Energy (ISSN: 0038092X) 86 (6), 1803–1815. http://dx.doi.org/10.1016/j.solener.2012.03.006.
Izquierdo, S., Rodrigues, M., Fueyo, N., 2008. A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations. Sol. Energy 82 (10), 929–939.
Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y., 2017. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 11–19.
Joshi, S., Mittal, S., Holloway, P., Shukla, P.R., Ó Gallachóir, B., Glynn, J., 2021. High resolution global spatiotemporal assessment of rooftop solar photovoltaics potential for renewable electricity generation. Nature Commun. 12 (1), 1–15.
Kolbe, T.H., 2009. Representing and exchanging 3D city models with CityGML. In: 3D Geo-Information Sciences. Springer, pp. 15–31.
Korfiati, A., Gkonos, C., Veronesi, F., Gaki, A., Grassi, S., Schenkel, R., Volkwein, S., Raubal, M., Hurni, L., 2016. Estimation of the global solar energy potential and photovoltaic cost with the use of open data. Int. J. Sustain. Energy Plan. Manag. 9, 17–30.
Krapf, S., Bogenrieder, L., Netzler, F., Balke, G., Lienkamp, M., 2022. RID – roof information dataset for computer vision-based photovoltaic potential assessment. Under Final Review.
Krapf, S., Kemmerzell, N., Khawaja Haseeb Uddin, S., Hack Vázquez, M., Netzler, F., Lienkamp, M., 2021. Towards scalable economic photovoltaic potential analysis using aerial images and deep learning. Energies 14 (13), 3800.

Lee, S., Iyengar, S., Feng, M., Shenoy, P., Maji, S., 2019. Deeproof: A data-driven approach for solar potential estimation using rooftop imagery. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2105–2113.

Li, Q., Shi, Y., Zhu, X.X., 2022a. Semi-supervised building footprint generation with feature and output consistency training. IEEE Trans. Geosci. Remote Sens.

Li, Q., Taubenböck, H., Shi, Y., Auer, S., Roschlaub, R., Glock, C., Kruspe, A., Zhu, X.X., 2022b. Identification of undocumented buildings in cadastral data using remote sensing: Construction period, morphology, and landscape. Int. J. Appl. Earth Obs. Geoinf. 112, 102909.

Li, Q., Zorzi, S., Shi, Y., Fraundorfer, F., Zhu, X.X., 2022c. RegGAN: An end-to-end network for building footprint generation with boundary regularization. Remote Sens. 14 (8), 1835.

Lingfors, D., Bright, J.M., Engerer, N.A., Ahlberg, J., Killinger, S., Widén, J., 2017. Comparing the capability of low-and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis. Appl. Energy 205, 1216–1230.

Margolis, R., Gagnon, P., Melius, J., Phillips, C., Elmore, R., 2017. Using GIS-based methods and lidar data to estimate rooftop solar technical potential in US cities. Environ. Res. Lett. 12 (7), 074013.

Melin, M., Shapiro, A., Glover-Kapfer, P., 2017. LiDAR for ecology and conservation - WWF conservation technology series (3). http://dx.doi.org/10.13140/RG.2.2.22352.76801.

Miranda, R.F., Szklo, A., Schaeffer, R., 2015. Technical-economic potential of PV systems on Brazilian rooftops. Renew. Energy 75, 694–713.

Mohajeri, N., Assouline, D., Guiboud, B., Bill, A., Gudmundsson, A., Scartezzini, J.-L., 2018. A city-scale roof shape classification using machine learning for solar energy applications. Renew. Energy 121, 81–93.

Moudrỳ, V., Beková, A., Lagner, O., 2019. Evaluation of a high resolution UAV imagery model for rooftop solar irradiation estimates. Remote Sens. Lett. 10 (11), 1077–1085.

Ordóñez, J., Jadraque, E., Alegre, J., Martínez, G., 2010. Analysis of the photovoltaic solar energy capacity of residential rooftops in Andalusia (Spain). Renew. Sustain. Energy Rev. 14 (7), 2122–2130.

Qi, F., Wang, Y., 2014. A new calculation method for shape coefficient of residential building using Google Earth. Energy Build. 76, 72–80.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Sánchez-Aparicio, M., Martín-Jiménez, J., Del Pozo, S., González-González, E., Lagüela, S., 2021. Ener3DMap-SolarWeb roofs: A geospatial web-based platform to compute photovoltaic potential. Renew. Sustain. Energy Rev. 135, 110203.

Sharma, P., Kolhe, M., Sharma, A., 2020. Economic performance assessment of building integrated photovoltaic system with battery energy storage under grid constraints. Renew. Energy 145, 1901–1909.

Sun, Y.-w., Hof, A., Wang, R., Liu, J., Lin, Y.-j., Yang, D.-w., 2013. GIS-based approach for potential analysis of solar PV generation at the regional scale: A case study of Fujian Province. Energy Policy 58, 248–259.

Suomalainen, K., Wang, V., Sharp, B., 2017. Rooftop solar potential based on LiDAR data: Bottom-up assessment at neighbourhood level. Renew. Energy 111, 463–475.

United Nations, 2021. Sustainable Development Goal 11: Make cities inclusive, safe, resilient and sustainable. https://www.un.org/sustainabledevelopment/cities/ (Accessed: 2021-12-16).

Walch, A., Castello, R., Mohajeri, N., Scartezzini, J.-L., 2020. Big data mining for the estimation of hourly rooftop photovoltaic potential and its uncertainty. Appl. Energy 262, 114404.