Quantification of the suitable area for rooftop solar panel installation from overhead imagery using Convolutional Neural Networks

Roberto Castello, Alina Walch, Raphaël Attias, Riccardo Cadei, Shasha Jiang and Jean-Louis Scartezzini

Solar Energy and Building Physics Laboratory, EPFL, Lausanne, Switzerland

Corresponding author: roberto.castello@epfl.ch

Abstract. The integration of solar technology in the built environment is realized mainly through rooftop-installed panels. In this paper, we combine state-of-the-art Machine Learning and computer vision techniques together with high-resolution overhead images to provide a geo-localization of the available rooftop surfaces for solar panel installation. We further associate them to the corresponding buildings by means of a geospatial post-processing approach. The stand-alone Convolutional Neural Network used to segment suitable rooftop areas reaches an intersection over union of 64% and an accuracy of 93%, while a post-processing step using building database improves the rejection of false positives. The model is applied to a case study area in the canton of Geneva and the results are compared with another recent method used in the literature to derive the available area.

1. Introduction

The built environment has the highest energy demand due to its intense human activity: buildings are the main drivers, with their continuous need for energy supply to perform routine operations. Nonetheless, buildings also offer an ideal frame for deploying renewable energy technologies, such as rooftop-mounted photovoltaic (RPV) panels, in order to generate local clean electricity. In this context, the large-scale deployment of RPV has attracted increasing attention in recent years. For a strategic integration of RPV in the built environment, an accurate assessment of their potential electricity generation is essential.

While this assessment can be very precise locally, at larger scale (regional and national) it poses a major challenge. This is mostly due to the lack of a scalable approach for accurately estimating the available area for RPV panel installation. For individual roofs, panels can be custom-fitted based on aerial photographs taking into account all obstructing objects, a practice widely applied for creating offers for RPV installations. Custom-fitting however is not scalable to 9.6 million rooftops in Switzerland. Automatic panel fitting algorithms based on roof shape data have hence been developed to estimate the available area for RPV installations at regional and national scale [1,2]. While these algorithms are scalable, they depend on the availability and accuracy of the roof shape data. Such data however rarely includes information on rooftop obstructing objects. Using Machine Learning (ML) techniques, the impact of these objects on the available roof area can be extrapolated from areas with detailed roof shape data [1,2]. While such detailed roof data include superstructures (dormers, chimneys, etc.), other obstructing objects, such as windows, remain unknown. In other studies, the

fraction of available roof area for large-scale RPV potential studies is estimated using constant coefficients and expert knowledge [3] or sampling techniques [4]. Aerial photographs, which are used for the custom fitting of PV panels, are rarely used in large-scale approaches due to their large volume and processing requirements. The only notable attempt is the work of [5] where public geographical data are used to extract roof slopes and available areas using edge detection techniques. However, the algorithms employed fail in 27% of the analyzed buildings, classifying them as flat instead of considering their actual slope.

In this paper, we leverage the power of Convolutional Neural Networks (CNNs) for pixel-wise semantic segmentation in combination with high spatial resolution aerial imagery to provide a quantification of the available area for RPV installation. This approach has already shown promising results in detecting existing RPV installations [6], having the advantages of being scalable and requiring a relatively small training set of images.

2. Data and Methods

2.1. Dataset

For the training of the CNN network, we use ortho-rectified 8-bit RGB images of Switzerland provided by the Swiss Federal Office of Topography (swisstopo) in TIF format, collected during the year 2015 [7]. They come in tiles of 17500x12000 pixels, with a spatial resolution of 0.25x0.25m^2. After splitting the images in smaller tiles of 250x250 pixels we selected 524 images covering the center of Geneva and the north and south neighborhood. Our approach is fully supervised, therefore we needed to provide the CNN with a set of labeled images in which the available area on the rooftop has been segmented by humans according to visual criteria. The labelling of the data was done manually and it took 12 hours. Available rooftop surfaces were labelled by means of closed polygons. Pixels belonging to visible structures like solar panels, windows, pipes, electrical and HVAC installations have been excluded from the roof labels.

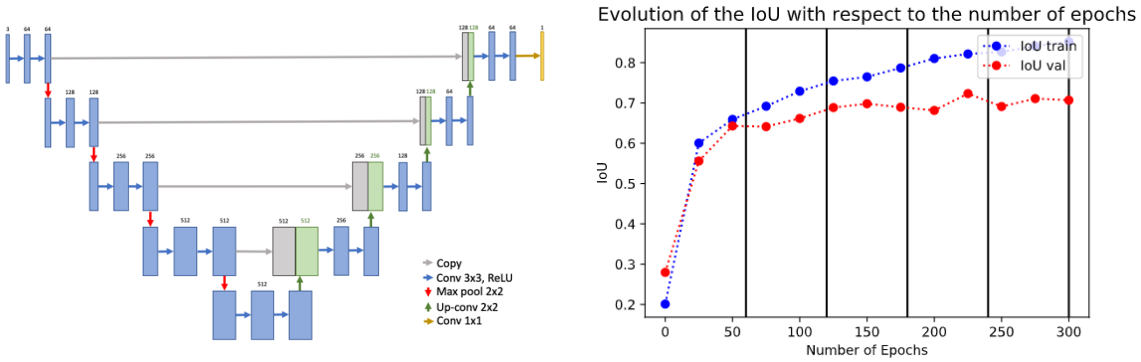
In order to link the detected pixels to the existing building stock and to remove false positives pixels (those identified by the CNN model outside the building footprints), we combine the segmented images with the national rooftop database from swisstopo [8]. It contains around 9.6 million vector polygons representing all roofs of the Swiss 3D building cadaster with their tilt, aspect and area.

2.2. Convolutional Neural Network

2.2.1. Network architecture

We adapt the U-Net architecture [9], a CNN developed for biomedical images segmentation, to our task and inputs. The network originally comes with two parts. A contracting path (*encoder*) extracts features at different levels through a sequence of convolutions, filters, activation functions and pooling layers, allowing to capture the context of each pixel. Then a symmetric expanding path (*decoder*) up-samples the results, increasing the resolution of the detected features. The output of each stage of the down-sampling phase is fed directly to the corresponding up-sampling phase to avoid separate training of encoder and decoder. U-Net has shown great performances with only a limited number of training images. In order to detect the available rooftop area to install RPV modules, we slightly modify the original architecture. We restrict the output of the model to a pixel-wise binary classification (suitable rooftop area or not suitable) and we adapt the number of filters and the input size in order to match the 250x250 pixels. Figure 1a shows the modified architecture, which has in total 14'788'929 parameters.

For each node we choose the Rectified Linear Unit activation function followed by batch normalisation. Since we are interested in the segmented area of an image, the output vector has a sigmoid activation function, which sets for each pixel the probability that it belongs to an area in the image suitable for RPV installation (positive class). Each pixel is then attributed to the positive class when its value exceeds a probability threshold, which is set at 50%.



(a) (b)

Figure 1. Modified U-Net architecture for the pixel classification task (a). Evolution of the Intersection Over Union score on training and validation sets during the training over 300 epochs (b).

2.2.2. Data preparation

The colour gradient plays an important role in segmentation tasks of RGB images. We increase the saturation in order to emphasize the contrast of the rooftops with respect to the surrounding environment (roads, trees, pools, etc.). Besides that, we also tested the adjustment of the brightness of images to support the detection of cast shadows. This however did not result in any improvement of the segmentation performance. We standardized each image with the mean and the variance of the whole dataset to make sure that all data points have the same scale. We apply diverse data augmentation techniques on the loaded training images. First, we randomly flip it horizontally, vertically or rotate it by 90 degrees: this increased by eight times the dimension of our dataset. Second, we randomly crop each image to a 248x248 size, which helps in detecting rooftops lying at the edges of the image. Third, we randomly add uniform and Gaussian noise to avoid model overfitting.

2.2.3. Model training

The whole dataset is split into three sets: 80% of images for the training, 10% for validation, and 10% for testing. We train the U-Net architecture starting from a random set of weights. The model is trained on small batches of two images at a time. Since the goal is to classify each pixel either as available or not for solar RPV installation, the binary cross entropy loss is the natural choice. The proportion of pixels that are labelled as available is low compared to non-available ones, creating an unequal distribution of the two classes. We overcome this potential bias by applying weighted binary cross entropy loss, setting the weight for the less frequent class to 4. Accuracy is the standard metric for classification tasks. Given the uneven class frequency in images, we apply Intersection Over Union (IoU or Jaccard Index), which is a more suited metric for unbalanced datasets. We rely on IoU in order to evaluate the training performance of our model. In the literature an IoU larger than 0.5 is considered a good prediction. As gradient-based optimization algorithm, we use Adam, commonly known in computer vision tasks to speed up the convergence, setting the default first and second moment estimates to 0.9 and 0.999. We start from a learning rate equal to 0.0008 and we further reduce it by a factor 0.8 every 60 epochs. As observed in Figure 1b, 225 epochs are sufficient to allow the IoU on the validation set to converge. The final metrics obtained on training, validation and test set are reported in Table 1.

Table 1. Performance of the CNN evaluated on the training, validation and test sets.

	IoU	**Accuracy**	**Recall**	**Precision**
Training	0.8823	0.9794	0.9299	0.9437
Validation	0.7211	0.9464	0.8360	0.8508
Test	0.6420	0.9307	0.7522	0.7874

2.3. Post-processing algorithm

Post-processing the segmented images by overlaying them with roof shapes is necessary to link the detected areas to the existing building stock. It further allows to add contextual information, such as the size, tilt and orientation of the underlying roof, and to remove false positive pixels outside buildings. The post-processing is performed through a series of geospatial processing steps: First, the detected pixels are intersected with the roof shapes. The shapes are buffered by 3 m to account for the frequent misalignments between the roof shapes and the images. Second, the pixels within roofs/buildings are clustered to groups of at least 3 m^2 of surface to eliminate further scattered pixels. Third, the clusters are aggregated per roof, converted from pixels to polygons (shapes) and projected onto the roof by dividing their area by the cosine of the roof tilt.

To further estimate the number of RPV panels that could be installed on the rooftop available area, we apply a geospatial algorithm developed in [2], which virtually installs RPV panels by projecting rectangular polygons of size 1.6 m^2 onto the rooftops. In contrast to [2], RPV panels are (virtually) installed only on the newly estimated shapes representing the available area, without leaving any buffer to the edge of these shapes. On flat roofs, panels are installed along the main aspect of the roof, simulating a "triangular" panel arrangement, namely rows of opposite orientation that are tilted at 15°.

3. Results and discussion

The CNN-based model is applied over two areas of the city of Geneva, where buildings are mostly residential and of relatively small size. Buildings with rooftop surfaces smaller than 10 m^2 have nearly zero available area and thus are excluded as considered not economically viable for the installation of panels. The total number of roof surfaces considered is 2391.

3.1. Available area for RPV installation

Figure 2 shows the result of the detection of rooftop available area for two different residential building types (flat and pitched), selected in the case-study area. Beside detecting most of the obstructing objects on rooftops, including the already installed RPV panels and roof areas covered by trees, surprisingly the CNN-based model excludes as "by-product" also shaded areas on the roofs, where the actual installation of RPV would be unsuitable. Furthermore, most roof terraces are classified as unsuitable by the model, since these often contain several objects which reduce the available area and consequently do not allow for the installation of RPV panels. Limitations of the current models are in detecting dormers, which are usually built from the same material as the roof itself and hence hard to distinguish from an overhead perspective, as well as areas on "green roofs" with vegetation.

(a) (b) (c) (d)

Figure 2. Example of area detected by the CNN model (blue shades in a, c) and virtually installed panels (light blue shapes in b, d) for large flat (a, b) and small pitched (c, d) rooftops.

3.2. Comparison with other methods

To set this work in context with existing studies on the available area for RPV installations, we compare the suitable roof surface extracted by the CNN-based method with a recent estimation from a large-scale estimation of the solar RPV potential for Switzerland [2]. For this comparison, we divide

the rooftops in three categories based on the individual roof surfaces: small roofs (surfaces of 10-100 m^2), medium roofs (100-500 m^2), and large roofs (>500 m^2). We further exclude roofs where the CNN-based model detects less than 20% of the available area, assuming that for these roofs there might have been an incomplete detection of the surface, given the limit in the model sensitivity. Figure 3 shows the comparison for three estimations: the suitable fraction of roofs as detected by the CNN model (*CNN*), the area covered by virtually installed panels (*Panels*, see Section 2.3) and the available area for RPV installation as estimated in [2] (*RPV*), which also employs a panel fitting algorithm in combination with ML to extrapolate the impact of superstructures on the available area.

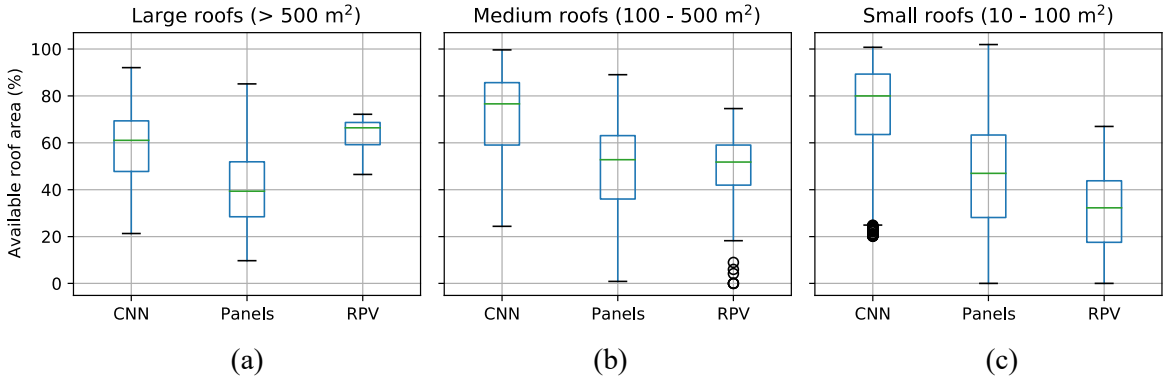


Figure 3. Estimated available surface for RPV installation as a fraction of the roof area. Results are separated in three building categories according to the roof size (a, b, c).

The latter two estimations (*Panels* and *RPV*) are hence directly comparable, indicating that the CNN model in combination with panel fitting yields a smaller median fraction of available area (39%) than the *RPV* approach (66%) for large roofs, similar fractions (51-52%) for medium-sized roofs and a larger fraction (47% vs. 32%) for small roofs. The large interquartile ranges and whiskers show that these fractions vary largely between roofs, which also causes large uncertainties for the estimations in [2]. Assuming the CNN model yields more realistic results than the *RPV* estimate as it is based on aerial images, the results hence suggest that the available area on large roofs is over-estimated in [2], presumably due to the large areas covered by superstructures such as HVAC equipment on many of these roofs. As the case study only covers primarily residential areas, no conclusion can be drawn for industrial areas. By contrast, the rooftop available area on small roofs may be underestimated in [2]. In the formulation of large-scale strategies for PV integration, these roofs however are of lower priority, since they mostly belong to single-family homes and represent small individual potentials.

Table 2. Statistical comparison of the three estimations of available area (*CNN*, *Panels*, *RPV*) for all roofs as well as for large, medium and small roofs.

	All roofs			**Large (> 1000 m^2)**			**Med. (100-500 m^2)**			**Small (10 - 100 m^2)**		
	CNN	*Panels*	*RPV*	*CNN*	*Panels*	*RPV*	*CNN*	*Panels*	*RPV*	*CNN*	*Panels*	*RPV*
Total area (10^3 m^2)	98.5	65.8	65.4	19.2	13.4	27.3	25.8	17.3	23.3	56.2	35.9	30.2
Mean % of roof	74	45.4	33.1	57.4	40	63.5	70.9	49	49.5	74.8	45.1	30.4
Std. % of roof	19.6	22.5	17.9	19.3	17.3	7.4	19.3	19.6	14	19.5	22.9	17.6
Median % of roof	79	47.7	35	61	39.4	66.4	76.6	52.8	51.8	80	47	32.2

Comparing the absolute available areas for the three estimations (Table 2), both approaches yield similar areas across all roofs, whereby a larger total area of *RPV* for large roofs is offset by a larger total area of *Panels* for small roofs. One drawback of the automatic panel fitting as shown in Figure 2b/d is that it yields suboptimal configurations and hence a lower number of installed panels compared to manual panel placement on images. This effect is particularly noticeable for small roofs, causing the large difference between the *CNN* and *Panels* estimation. However, the panel fitting may yield an

overall more realistic estimation of available roof surface, as it allows to account for roof geometry and because it is unlikely that the whole available area would be exploited on all roofs.

4. Conclusions

In this work we present a novel method to estimate the available rooftop surfaces for solar panel installation by combining state-of-the-art ML and computer vision techniques together with high-resolution aerial images. The application of the method to the case study areas in Geneva demonstrates that an image segmentation approach can reliably detect and exclude several types of roof superstructures. Assuming this CNN-based model yields realistic results, as it is based on aerial images, we compare it with an existing approach for large-scale estimate applied to the same group of buildings. The results suggest that the median fraction of available area on large roofs seems to be over-estimated by the large-scale approach, while the potential of small surfaces is under-estimated. To further enhance the discriminative power of the CNN-based model, a fourth channel (infrared) may be added to the images to include information on the emissivity of roofs materials. Furthermore, higher-resolution aerial images from more recent campaigns may allow to include more recent buildings. These ameliorations would make possible the extension of the CNN-based model for the quantification of rooftop areas at the national scale. Such an estimate may provide valuable insights for the ongoing public discussion on the realizable solar potential across Switzerland, as the output of the model would be available for individual roofs and could be integrated in existing databases and visualization tools.

Acknowledgments

This research has been financed and supported by the Swiss National Science Foundation (SNSF) under the National Research Program 75 (Big Data) for the 'HyEnergy' project, No. 407540_1167285

References

[1] D. Assouline, N. Mohajeri, J.-L. Scartezzini, Large-scale rooftop solar photovoltaic technical potential estimation using Random Forests, Applied Energy, Volume 217, 2018, Pages 189-211, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2018.02.118

[2] A. Walch, R. Castello, N. Mohajeri, J.-L. Scartezzini, Big data mining for the estimation of hourly rooftop photovoltaic potential and its uncertainty, Applied Energy Volume 262, 2020, 114404, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2019.114404

[3] M. Portmann, D. Galvagno-Erny, P. Lorenz, D. Schacher. Sonnendach.ch: Berechnung von Potenzialen in Gemeinden, Bundesamt für Energie BFE, October 2016

[4] Wiginton LK, Nguyen HT, Pearce JM., Quantifying rooftop solar photovoltaic potential for regional renewable energy policy, Computers, Environment and Urban Systems 2010; 34:345–57 https://doi.org/10.1016/j.compenvurbsys.2010.01.001

[5] Kai Mainzer, Sven Killinger, Russell McKenna, Wolf Fichtner, Assessment of rooftop photovoltaic potentials at the urban level using publicly available geodata and image recognition techniques, Solar Energy, Vol.155, 2017, Pag. 561-573, ISSN 0038-092X, https://doi.org/10.1016/j.solener.2017.06.065

[6] R. Castello; S. Roquette; M. Esguerra; A. Guerra; J.-L. Scartezzini, Deep learning in the built environment: automatic detection of rooftop solar panels using Convolutional Neural Networks, J. Phys.: Conf. Ser. 1343 012034, 2019, https://doi.org/10.1088/1742-6596/1343/1/012034

[7] Swisstopo, Geoportal 2020, https://shop.swisstopo.admin.ch/en/products/free_geodata

[8] D. Klauser, Solarpotentialanalyse für Sonnendach.ch, Schlussbericht, Bundesamt für Energie BFE, 2016, http://www.bfe.admin.ch/geoinformation/06409/

[9] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28