Roofpedia: Automatic mapping of green and solar roofs for an open rooftopscape registry and evaluation of urban sustainability

Abraham Noah Wu a,1, Filip Biljecki a,b,* 2

a Department of Architecture, National University of Singapore, Singapore
b Department of Real Estate, National University of Singapore, Singapore

HIGHLIGHTS

- There is a lack of open data on urban rooftop typology and current use of roofs.
- A deep learning and GIS workflow to map and quantify green and solar roofs.
- A generated dataset that covers 17 cities, scalable to include more locations.
- An index to benchmark the proliferation of green and solar roofs in cities.

GRAPHICAL ABSTRACT

ARTICLE INFO

Keywords:
Sustainable development
Convolutional Neural Network
Computer vision
Carbon neutrality
Building data
OpenStreetMap

ABSTRACT

Sustainable roofs, such as those with greenery and photovoltaic panels, contribute to the roadmap for reducing the carbon footprint of cities. However, research on sustainable urban rooftops is rather focused on their potential and is hindered by the scarcity of data, limiting our understanding of their current content, spatial distribution, and temporal evolution. To tackle this issue, we introduce Roofpedia, a set of three contributions: (i) automatic mapping of relevant urban roof typology from satellite imagery; (ii) an open roof registry mapping the spatial distribution and area of solar and green roofs of more than one million buildings across 17 cities; and (iii) the Roofpedia Index, a derivative of the registry, to benchmark the cities by the extent of sustainable rooftops in term of solar and green roof proliferation. This project, partly inspired by its street greenery counterpart 'Treepedia', is made possible by a multi-step pipeline that combines deep learning and geospatial techniques, demonstrating the feasibility of an automated methodology that generalises successfully across cities with an accuracy of detecting sustainable roofs of up to 100% in some cities. We offer our results as an interactive map and open dataset so that our work could aid researchers, local governments, and the public to uncover the pattern of sustainable rooftops across cities, track and monitor the current use of rooftops, complement studies on their potential, evaluate the effectiveness of existing incentives, verify the use of subsidies and fulfilment of...
1. Introduction

As urban population continues to grow, cities need to search for new spaces to feed the expansion. Traditionally, cities expanded horizontally to accommodate more housing, amenities, and public spaces, and this process took away forests and agricultural land which had a direct impact on the surrounding environment (Benson-Lira, Georgescu, Kaplan, & Vivoni, 2016; Du, Xie, Zeng, Shi, & Wu, 2007). According to Mursheed, Duval, Koch, and Rode (2018), compact and vertical cities are considered to be more energy-efficient and environmentally friendly, and one method of increasing the compactness of cities is to build on the roofs where there is a multitude of potential uses including civic spaces, urban farms, and solar power.

A roof serving a dual purpose is not a new invention. The flat roofs of traditional Egyptian houses were often used as an extension of their living spaces (de Garis Davies, 1929), and Corbusier (1927) believed that a roof garden is an essential element to modernist architecture. Today, luscious sky gardens with swimming pools have become a symbol of high-end residential developments, fetching considerable premiums than ordinary condominiums (Shukri & Misni, 2017) and offering residents a place to get in touch with nature, which is critically correlated to happiness in an urban setting (Han & Kim, 2019).

Besides hosting civic activities, the roofscape also has enormous potential in improving the climatic and energy performance of a building and reducing or offsetting its carbon footprint, e.g., by providing thermal insulation, which may result in decreased energy consumption (Coutts, Daly, Beringer, & Tapper, 2013; Jaffal, Ould-boukhitine, & Belarbi, 2012; Theodosiou, 2003; Wong, Tan, Kolokotsa, & Takebayashi, 2021). Yang et al. (2018) shows that green roofs can reduce 31% of a building’s heat gain on a summer day and when applied in large scale, green roofs could mitigate urban heat island (UHI) effect with an average decrease of the peak ambient temperature close to 0.9 K (Santamouris, 2014). These benefits would increase the thermal comfort of urban residents and reduce excessive air-conditioning energy usage due to the UHI effect (Wong, Chen, Ong, & Sia, 2003a; Wong et al., 2003b; Salamanca, Georgescu, Mahaloh, Moustouiali, & Wang, 2014). Additionally, green roofs provide undisturbed green cover in a city that could become habitats for migrating and breeding birds and partially mitigate the loss of habitat due to increasing urbanisation (Partridge & Clark, 2018).

Moreover, green roofs also contribute significantly to urban stormwater management, having a mean percent rainfall retention of up to 87% (VanWoert et al., 2005), and 50 to 60% of the rain falling on a green roof will be evaporated or transpired back to the atmosphere, hence easing the pressure on storm drainage systems (Jarrett, 2016). Green roofs bring further environmental advantages over their unvegetated counterparts to both buildings and their surroundings (Shafique, Kim, & Rafiq, 2018). These benefits include reduction of traffic noise in urban environments (Van Renterghem & Botteldooren, 2009; Yang, Kang, & Choi, 2012), extending the lifetime of rooftops by shielding them from solar radiation and physical damage (Kosareo & Ries, 2007; Carter & Keeler, 2008), improvement of air quality and filtering air pollutants (Getter, Rowe, Robertson, Cregg, & Andrews, 2009; Yang, Yu, & Gong, 2008), improvement of aesthetic value, appearance and creation of recreational space (Villaarle & Bengtsson, 2005; Jungels, Rakow, Allred, & Skelly, 2013), promotion of urban biodiversity (Schrad & Böning, 2006; Brenneisen, 2006), and providing space for urban agriculture (Lim & Kishnani, 2010; Whittinghill & Rowe, 2011; Beacham, Vickers, & Monaghan, 2019; Palliwal, Song, Tan, & Biljecki, 2021). When roofs are used for urban farming, they may reduce the carbon footprint by decreasing emissions from shortened food-related transport (Goldstein, Hauschild, Fernández, & Birkved, 2017).

Apart from green roofs, those that have a rooftop photovoltaic system (i.e. solar panels) mounted on them (in our paper, we call these solar roofs for the sake of simplicity) are also a popular sustainable roof typology adopted by cities around the world. In comparison to green instances, solar roofs are cheaper to install and require less maintenance which makes them accessible to more homeowners. One key benefit of a solar roof is the financial incentive it offers. In a research by Mountain and Szuster (2014), Australian PV households have saved $252 million per year between 2010 and 2012, and electricity generated from solar roofs could be half the cost than commercial providers after factoring in installation costs and government subsidies.

It is also important to note that buildings equipped with solar roofs and building-integrated photovoltaics (BIPV) on facades can reduce operational carbon emissions of buildings by up to 50% (Happle et al., 2019). When optimised to reach their maximum potential, solar roofs alone can help a building achieve net-zero energy consumption as seen in an increasing number of cases around the world (Wong, 2019; Sood, Quintana, Jayathissa, Abdelrahman, & Miller, 2019). In contrast to the literature on green roofs, studies on the effect of solar roofs in mitigating UHI remain debatable with some of them claiming that energy absorption of solar panels could reduce UHI (Masson, Bonhomme, Salagnac, Briottet, & Lemonsu, 2014) while others presenting opposing views (Wang et al., 2005; Barron-Gafford et al., 2016).

Solar roofs and green roofs can be integrated as one solution that complements each other for greater environmental benefit (Tablada et al., 2018; Ciriminna, Meneguzzo, Pecoraino, & Pagliaro, 2019). According to Hui and Chan (2011), integrated systems could generate more electricity than standalone PV systems because green roofs improve the efficiency of the solar panels by keeping the surrounding temperature at a more suitable level for energy generation.

Considering the benefits described above, it is not surprising that there is a multitude of studies that imply that the roofscape is a promising venue in supporting sustainable urban development (Hachem, Athienitis, & Fazio, 2011; Bellini & Mocci, 2019; Dubey, 2018; Zamperini & Lucenti, 2014). Many of the studies perform large-scale geographical simulations to estimate the potential and benefit of greening and installing PV panels on rooftops (Wong & Jusuf, 2008; Santos, Tenedorio, & Gonçalves, 2016; Rodríguez, Duminić, Ramos, & Eicker, 2017; Bodis, Kougiás, Jäger-Waldau, Taylor, & Szabó, 2019; Joshi, Selmi, Binard, Nys, & Teller, 2020; Shao et al., 2020; Santos, Silva, Tenedorio, & Góes, 2021). For example, according to Walsh, Castello, Mohajerizadeh, and Scartezzini (2020), solar panels can be installed on more than half of Switzerland’s 9.6 million rooftops. The resulting power would meet more than 40% of the Swiss electricity demand. Another study by the Melbourne (2016) assesses the constraints of solar and green roof installations for existing building hoping to serve as a guide for the residents interested in upgrading to a sustainable roof typology. Recently, a crowdsourced dataset of solar panel installations in the UK was collected manually, and combined with building information from other volunteered GIS platforms such as OpenStreetMap, opening opportunities in studying the patterns of solar panel deployment (Stowell et al., 2020).

However, notably scarce are studies that instead of estimating the potential, focus on the current status of the sustainable roofscape and examine the quantity and spatial distribution of current green and solar installation on rooftops at a large-scale. Such studies designed to understand the current landscape of solar and green installations in building rooftop would be useful for several purposes, including subsequent research on estimating the potential (e.g. by identifying roofs that are not yet utilised for either purpose), gauging the effectiveness of urban policies and verifying their mechanisms (e.g. subsidies by...
governments), tracking and monitoring environmental pledges by businesses, and estimating the current carbon offsetting capacity of cities and benchmarking them.

In this paper, we bridge this gap in multiple ways, by developing a three-pronged work (Fig. 1). First, we develop an automated method to identify rooftops that are vegetated and/or have photovoltaic system installations. The method relies on high-resolution satellite imagery, which is available for an increasing number of cities around the world, facilitating replication and scalability of our work. Second, we implement the method on 17 cities to create a cadastre of the roofscape, revealing the quantity and spatial distribution of rooftops that contribute towards a city’s sustainable development. Our dataset contains information on more than one million buildings, and we are making this sustainable roofscape inventory available publicly, aiding researchers, practitioners, local governments, and the public to understand the current status of the roofscape in the context of sustainable urban development and achieving carbon neutrality. Third, our registry enables us to benchmark and compare cities with regard to the utilisation of their roofscape. We establish the level of penetration of green and solar roofs in each city, and quantify it, creating the Roofpedia Index. The scores for the 17 cities in our study are computed, and the cities are compared.

Some of the aspects of the project and its name — Roofpedia — are inspired by Treepedia (Seiferling, Naik, Ratti, & Proulx, 2017; MIT Senseable City Lab, 2016), a project to measure and map the amount of street greenery in cities from the pedestrian perspective, and compare cities around the world. A notable characteristic of our modular work is that it enables replication and extensibility, thus, additional cities can be included in Roofpedia in the future — from mapping the roofscape to calculating their index for benchmarking purposes, and potentially supporting future extensions to include other typologies of rooftops such as social uses. Furthermore, this service may potentially provide a basis for a crowdsourced platform that would integrate additional information on rooftops.

The paper is organised as follows. In Section 2 we expand on the introduction and continue the literature review. Section 3 describes our methodology for automatically mapping green and solar roofs, and evaluates the results we achieved. Section 4 analyses the results, presents our database, and the derived Roofpedia Index to assess the performance of cities. In Section 5 we discuss the work, expose limitations, and map out ideas for future work. Section 6 concludes the paper.

2. Background

2.1. Sustainable roof policies and documentation

From the previous section on the benefits of a sustainable roofscape, it is evident that the roof could offer much more to a city rather than just been a plain enclosure for shelter and interior comfort. The roofscape of a city can contribute positively to a city’s environment and quality of life, especially when instruments for sustainable development such as vegetation and PV panels are applied widely.

The potential is even greater when considering that not only new buildings may be enriched with green and solar installations. For example, many existing unvegetated roofs can be converted into green roofs (Wilkinson & Reed, 2013), which may be especially appropriate

![Roofpedia Diagram](image_url)
for buildings with flat roofs and sufficient structural strength (Clark, Adriaens, & Talbot, 2008; Stovin, 2010; Castleton, Stovin, Beck, & Davison, 2010).

While green roofs are a viable expenditure from the societal, environmental and economical perspective, they come at a cost, and their installation is inhibited by high investments (Nelms, Russell, & Lence, 2007; Bianchini & Hewage, 2012; Sproul, Wan, Mandel, & Rosenfeld, 2014). To offset these costs, many local governments worldwide have introduced various incentives and guidelines to encourage the utilisation of rooftops spaces for activities that contribute to sustainable development (Claus & Rousseau, 2012; Chen, 2013). For example, sky gardens for communal uses are exempted from Floor Area calculations in Hong Kong (Hong Kong Buildings Department, 2019), and Singapore’s National Parks Board will fund up to 50% of installation costs of green roofs for developers (National Parks Board, 2009). The incentives in these two land-scarce and densely populated jurisdictions also suggest the proliferation of the role of rooftops in transcending the frontiers of space utilisation, especially in promoting carbon offsetting. In the case of solar roofs, private PV panel owners in the United States are allowed to ‘turn-back’ their meter if they use less energy than what the solar panel generates (U.S. Department of Energy, 2020). Residents in the Australian state of Victoria can enjoy a rebate for the installation of a solar panel system with a four-year, interest-free loan for the remaining amount (Solar Victoria, 2020). While the above incentives could potentially increase the adoption of sustainable roofs, publicly available data on the current state of adoption in most cities is very limited, making it difficult to assess the effectiveness of current policies or to understand the current state of adoption.

Some projects have endeavoured to tackle this issue by mapping rooftop typologies at the city-scale. There are two initiatives that are closely related to our work. First, the Melbourne (2016) created a one-of-a-kind roof registry that maps the city’s existing green and solar roofs and also potential roofs that could be converted to sustainable roofs. Another study was done in Berlin by the Senate Department for Urban Development and Housing (2017). An automatic mapping was carried out with GIS processing that returned detailed statistical data on the area and location of green roofs in the city. Both projects offer an interactive map of the dataset on sustainable roofs and provide an open database for further research.

These two projects set a commendable example in popularising roof data as they enable everyone to explore the rooftops intuitively and potentially sparking interest in the topic. However, these projects have their fair share of limitations. For example, the project in Berlin is restricted to mapping green roofs. Further, both are focused on a single city, not allowing easy comparisons on how a city fares in comparison to others when it comes to sustainable installations on their rooftops. Additionally, no statistical validation was carried out to assess the accuracy of their methodology, nor the work has been published in international scientific outlets.

In this paper, we seek to overcome their shortcomings. We advance the steps taken by these two projects further and aim to map both green and solar roofs across multiple cities to create a global registry of sustainable rooftops. We aim to assess the feasibility and accuracy of an automated mapping methodology and present the information intuitively to the public. It is important to underline that our work is not confined to a single city, as our cross-city analysis not only includes more than a dozen cities, but it is further scalable as long as satellite imagery of comparable quality as the ones in the training set is available. Such scalability has allowed us to create a cross-city benchmark (Roofpedia Index) that measures the penetration of solar and green roofs across 17 different cities. Furthermore, we have provided open access to the Roofpedia prediction pipeline where users could carry out the mapping and estimations using their own imagery. Additionally, our methodology of using satellite imagery allows us to examine the temporal evolution of sustainable rooftops, which would enable multiple use cases in the future, e.g. understanding the effectiveness of policies on a temporally fine scale. As photovoltaic capacity is dynamic and rapidly evolving, it is important to track changes through time.

2.2. Automated roof mapping with deep learning

In recent years, deep learning, especially Convolutional Neural Networks (CNN) (Shelhamer, Long, & Darrell, 2017) offer new opportunities for object detection in optical remote sensing images (Cheng & Han, 2016) and it has been used for various applications in the built environment (Middel, Lukasczyk, Zakrzewski, Arnold, & Maciejewski, 2019; Chen, Stouffs, & Biljecki, 2021). Basu et al. (2015) demonstrated that deep learning algorithms can classify satellite image tiles to an accuracy of 97.95%. Going further in granularity, Chen, Xiang, Liu, and Pan (2014) created a pipeline for vehicle detection in satellite images using Deep Convolutional Network.

There exists a multitude of research on the feasibility of mapping the rooftopscape with CNNs with satellite or aerial images. Chen and Li (2019) showed that CNNs can detect and segment building boundaries of small detached houses from aerial images in the city of Christchurch. Castello, Roquette, Esguerra, Guerra, and Scarpettini (2019) demonstrated a deep learning method that predicts solar panels at pixel level with an accuracy of about 94% and an Intersection over Union of up to 64% using high-resolution aerial photos, which is a significant leap compared to traditional machine learning approaches (Malof, Bradbury, Collins, & Newell, 2016). However, the research was only an evaluation of a small region, and its scalability remains to be tested.

While aerial images are still limited to selected areas, satellite images have a wider coverage, and previous studies had demonstrated that deep learning methods on satellite images can still provide reliable predictions of the shapes and the sizes of buildings roofs or building footprints (Nosrati & Saeedi, 2009; Lee, Iyengar, Feng, Shenoy, & Maji, 2019). Therefore, deep learning performed on satellite images will not be limited to a specific area but has the potential to scale up to cover myriads of cities around the globe.

Using a CNN architecture called Inception-v3 and a large-scale training set of 366,467 images, Yu, Wang, Majumdar, and Rajagopal (2018) created a solar installation database for the contiguous U.S. This project demonstrates the scalability of computer vision algorithms in mapping solar panels from satellite images, and their results are encouraging for the development of our project. However, the study provides visualisation only at the scale of census tracts, unlike the Melbourne study where the resolution is at the building level. It is also difficult to use the same method on green roofs as there is yet to be a large labelled green roof dataset available. Further, while appreciating the scope of the study given the large extent of the U.S., the project is still focusing on only one country.

Nevertheless, compared to traditional spectral classification, CNNs can be used on lower quality images that are widely available in the public realm. Using a modified CNN pipeline with U-Net architecture, in the next section, we present an automated method to identify both green and solar roofs from satellite images and assign corresponding labels to building datasets derived from OpenStreetMap to create a roof registry at the building level for spatial visualisation and further analyses.

Besides the novelties and contributions discussed in Section 2.1, our method has several advantages compared to existing methods mentioned in this section. Firstly, we present work at a substantially higher spatial resolution than the current state of the art (i.e. census tract level to building level). Secondly, the model is created with international scalability in mind and it can be used in multiple cities around the world with different urban morphology and building typology. Thirdly, our method identifies both green and solar roofs, which adds another dimension to understanding the sustainability of the rooftopscape across cities. Finally, as the cities are assessed with the same model and within the same system, comparisons between cities would be fairer and easier than comparing results from different research projects. This advantage enables us to create a sustainable index that
ascribes standardised scores to the state of sustainable roof practice in each city. Therefore, one of the outcomes of the work is not only a database of such roofs across many cities, but also gauging how sustainable and effective cities are when it comes to unlocking the space provided by rooftops.

3. Automated roof classification

3.1. Data and study areas

To demonstrate the scalability and feasibility of our methodology, 17 morphologically diverse and geographically distributed cities spanning Europe, North America, Australia, and Asia are selected as our study areas. These are: Berlin, Copenhagen, Las Vegas, Los Angeles, Luxembourg City, Melbourne, New York, Paris, Phoenix, Portland, San Diego, San Francisco, San Jose, Seattle, Singapore, Vancouver, and Zurich.

The satellite imagery used in this study is retrieved with the Mapbox static tiles API, except for Luxembourg for which the data is acquired from the open data by the Luxembourgish Land Registry and Topography Administration, affirming that our method is not constrained to one data source. Mapbox has a generous free tier usage limit enabling obtaining imagery of several cities. According to Mapbox, the images available for access are from a combination of different sources, including Maxar’s Vivid products for much of the world, Nearmap and USDA’s NAIP 2011–2013 in the contiguous United States, and open aerial imagery from Denmark, Finland, Germany, and other regions. These datasets offer resolutions of about 50 cm and higher in major cities in the world, and would be an effective way to assess the scalability and adaptability of our methodology. Data on building footprints have been extracted from OpenStreetMap, which is being increasingly used in research (Cerri, Steinhausen, Kreibich, & Schröter, 2020; Feldmeyer, Nowak, Jamshed, & Birkmann, 2021), and which has a very good level of completeness for the study areas we have selected (Fan, Zipf, Fu, & Neis, 2014; Biljecki, 2020). The building footprints serve a dual purpose: they filter out greenery and solar panels not installed on rooftops, and they are used to integrate the results of our work, enriching the building data. The visualisations of the final result are rendered on Mapbox for public access and both the dataset and code are available openly on Github.

In our research, we have included a few more cities besides the listed ones (e.g. Washington D.C. and Marseille). However, eventually we did not proceed forward with including them in our database and index, due to a deviating performance of the model to identify sustainable roofs, which was largely caused by external factors such as quality of imagery and presence of particular features in these cities (e.g. a large density of skylights that tend to be misidentified as solar panels). Nevertheless, considering these cities was vital for understanding the data requirements and exposing limitations of the method, and further examples are given in Section 3.3.

3.2. Methodology

Our methodology can be broken down into four steps (Fig. 2). First, we label the training images across multiple diverse locations and tile the satellite images to a uniform shape for the deep learning algorithm. The polygonal labels required for the training are manually labelled across multiple cities to create a wide range of examples in different urban contexts and image conditions. Secondly, the processed dataset is trained in a Convolutional Neural Network based on a modified U-net architecture. The network is initialised with a pre-trained model (ResNet50) to improve the accuracy of the prediction. Thirdly, the building footprints of the predicted areas are passed in as new inputs and the probabilities predicted by the CNN are converted into georeferenced polygons to tag the respective building footprints with either ‘Solar’ or ‘Green’ or both. Lastly, the resulting, semantically enriched building footprints are analysed quantitatively and visualised spatially to create the Roofpedia Registry and the Roofpedia Index.

3.2.1. Labelling and tiling

Fig. 3 shows a sample polygon label for rooftop solar panels and rooftop greener in multiple cities. In total, 1812 polygons have been labelled across eight cities to capture a wide range of rooftop greener and solar panels. Specifically, there are 545 unique labelled rooftop greener summing up to an estimated 208,116 square metres in area and 1,254 solar panels summing up to an estimated 175,698 square metres in area. Converted into individual tiles for training, validation and testing, there are a total of 1,517 labelled tiles for rooftop greener and 1,380 labelled tiles for solar panels.

For training, we have used the imagery of 8 instead of all 17 cities in our study to determine whether the trained models will be scalable in new locations. In each of the cities, areas are randomly selected during labelling and thus the labelled data do not represent all green and solar roofs in the city. Nevertheless, they are sufficiently representative of the green and solar roof typology in the city, and as the continuation of this paper will demonstrate, the model trained on these locations scales

![Fig. 2. Overview of the methodology of Roofpedia.](image-url)
Fig. 3. Labelling solar panels and rooftop greenery across cities.

Fig. 4. Tiling of satellite images and labels.
successfully in new areas without training data, facilitating the extension of the work to easily and accurately include new cities after the publication of this paper.

After creating training labels, we slice the tiles and the vector polygon masks of the satellite images using a zoom level of 19 according to the Slippy Map tiling convention described by OpenStreetMap. This zoom level enables the model to focus on roof objects and minimise confusion with objects outside of the building boundaries. The images and masks are further sliced into 256 x 256-pixel datasets for training. Labelled polygons are similarly sliced and converted into image masks for training (Fig. 4). In total, 2,897 images are prepared with 80% serving as training data and 20% as test data to validate the accuracy of the model.

3.2.2. Image segmentation with deep learning

The deep learning model used is a type of Fully Convolutional Neural Network (Shelhamer et al., 2017), which excels at image boundary identification and instance segmentation. The model is modified from U-Net (Ronneberger, Fischer, & Brox, 2015) by replacing the encoding section of the U-Net with a pre-trained Resnet50 network (He, Zhang, Ren, & Sun, 2015). The U-Net architecture is chosen because of its optimisation on augmented data which uses available training samples more efficiently. As such, a U-Net can be trained from very few images and still provide reliable results. By replacing the pre-trained Resnet50 as the encoding layer for the U-Net, we leverage the power of transfer learning to reduce false negatives (Shi et al., 2019), and further improve the validation accuracy of the model (Shin et al., 2016; Hussain, Bird, & Faria, 2018).

We use PyTorch (Paszke et al., 2019) for building the deep learning model and we have borrowed a preprocessing pipeline from Robosat, a satellite feature extraction toolbox (Ng & Hofmann, 2018). We created our own pipeline integrating these packages, which is available publicly for reproducibility.

To avoid confusion in shared features between greenery and solar panels, the model is trained on the solar dataset and green dataset separately, learning the weights independently for each feature. During training, the model learns the respective features that make up a patch rooftop greenery or a solar panel from the input image and mask and validates its accuracy against the test set.

Fig. 5 shows the results of the prediction model against the ground truth in the evaluation set for the green and solar models, respectively. The performance of the model is measured by Mean Intersection over Union (mIoU) which measures the number of pixels common between the ground truth and prediction masks divided by the total number of pixels present across both masks. The model for identifying solar portions of roofs achieved a mIoU score of 0.784 while the green roof counterpart achieved a mIoU of 0.396. Compared to solar panels which have a well defined edge, some green roofs contain trees and vegetation that do not have a defined edge. This fuzziness has resulted in lower mIoU score for the green roof model as it struggles to find these edges. Nevertheless, this shortcoming does not have a major influence over the quality of the final result as only the locations of the predicted polygons are used in GIS operations.

3.2.3. Classification with geospatial engineering

Even though the training labels of the deep learning model only contain solar panels and greenery on building rooftops, the entire image tile is passed through the model during training and inference. This process causes features that look like solar panels or greenery outside of the building rooftops to be picked up. For example, green patches in a park or trees on the streets could potentially be identified since they share similar features as rooftop gardens. Open-air carparks also have a similar grid-like feature that could be confused as solar panels. While this issue can be technically solved by complex image pre-processing procedures within the model architecture (i.e. creating a cropped image for each building), it will increase the size of the data and the complexity of the model, and thus increase the computational complexity for training and prediction.

To avoid this issue, we added a post-processing algorithm to remove false positives and restrict the predictions only within the building polygons. The prediction results from the deep learning model are converted into georeferenced polygons for GIS processing. When converting the pixel masks into predicted polygons for GIS operations, there is a small degree of rectification to simplify the polygons to reduce the number of edges while maintaining the general shape of the prediction. In addition, a denoiser is added to remove ‘speckles’ to reduce the number of False Positives. The setting of the denoiser is fine-tuned to provide the best balance between accuracy and precision.

Following the conversion, the resultant polygons that do not intersect building polygons are removed, and polygons that intersect building polygons but are too irregular or insignificant are ignored since we have noticed that these are predominantly false positives. Finally, building footprints that contain or intersect with the cleaned prediction set are assigned a label depending on whether they have a solar or green

![Fig. 5. Model prediction results compared with ground truth.](image-url)
roof, or both (Fig. 6).

### 3.3. Results and evaluation

To determine the scalability and accuracy of the prediction, 11 neighbourhoods in different cities are selected for evaluation. These areas are not included in the original training set and their ground truths are labelled manually. As such, the results obtained is unbiased and effective in measuring the adaptability of the model across cities. Furthermore, some of the neighbourhoods are in cities excluded from training data which will help to measure the adaptability of the model when applied to a new city where it has not been trained. The size of the neighbourhoods varies with a range of 475 to 3425 buildings, and all buildings in the selected areas are carefully labelled as the ground truth, which is then compared against the predicted results.

#### 3.3.1. Metrics for evaluation

We evaluate the performance of the methodology in both the number of buildings (Count) and the area of rooftops (Area) identified with green or solar roofs. A model that predicts Count accurately allows us to analyse the distribution of the identified rooftops in the urban fabric whereas a model that predicts Area accurately allows us to analyse how extensive a city has adopted sustainable roof typologies amongst its total roof area. Learning from the performance metrics proposed by Wentz and Zhao (2015), the following variables are used for calculating the performance of the method. These variables are also shown as columns in Tables 1–4.

- Total Count/Area (Total): Total number/area of building footprints in the study region
- Truth: Total number/area of building footprints in the Ground Truth set
- Prediction (TP + FP): Total predicted number/area of building footprints predicted by the model
- True Positive (TP): Both predicted and ground truth data classify roof as green or solar.
- False Positive (FP): Only the predicted data classify roof as green or solar.

Accuracy of the prediction is measured by the Percentage of True Positives (%Matching) in terms of Count or Area. This value is obtained by dividing the Count or Area of the True Positives against the Total Count or Area of the Ground Truth in Eq. (1). Precision is measured by the Percentage of False Positives (%FP) in terms of Count or Area. This value is obtained by dividing the Count or Area of the False Positives (FP) against the Count or Area of all buildings (Total) in the dataset in Eq. (2). The proportion of roofs that contains greenery or solar panels both in terms of Count or Area is represented by %Cover. This metric helps to interpret the extensiveness of the coverage of solar and green roofs in the regions according to the true positives in the prediction in Eq. (3).

\[
\text{%Matching} = 100 \times \frac{TP}{Truth} \tag{1}
\]

\[
\text{%FP} = 100 \times \frac{FP}{Total} \tag{2}
\]

\[
\text{%Cover} = 100 \times \frac{Matching(TP)}{Total} \tag{3}
\]

#### 3.3.2. Green roof results

Tables 1 and 2 present the prediction results of green roofs against the manually labelled ground truth. Using absolute Count as the metric, the accuracy of the model reached 77.45% of the ground truth, and the

---

**Table 1**

<table>
<thead>
<tr>
<th>Region</th>
<th>Area (km²)</th>
<th>Truth Count</th>
<th>Pred. Count</th>
<th>Matching</th>
<th>%Matching</th>
<th>%FP</th>
<th>%Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin 1</td>
<td>1.20</td>
<td>754</td>
<td>153</td>
<td>158</td>
<td>121</td>
<td>79.08</td>
<td>4.91</td>
</tr>
<tr>
<td>Berlin 2</td>
<td>1.21</td>
<td>1009</td>
<td>93</td>
<td>108</td>
<td>86</td>
<td>92.47</td>
<td>2.18</td>
</tr>
<tr>
<td>Melbourne</td>
<td>4.50</td>
<td>1079</td>
<td>26</td>
<td>44</td>
<td>24</td>
<td>92.31</td>
<td>1.85</td>
</tr>
<tr>
<td>New York 1</td>
<td>1.02</td>
<td>757</td>
<td>145</td>
<td>89</td>
<td>81</td>
<td>55.86</td>
<td>1.06</td>
</tr>
<tr>
<td>New York 2</td>
<td>2.05</td>
<td>1481</td>
<td>73</td>
<td>56</td>
<td>52</td>
<td>71.23</td>
<td>0.27</td>
</tr>
<tr>
<td>Paris</td>
<td>1.70</td>
<td>3425</td>
<td>83</td>
<td>78</td>
<td>64</td>
<td>77.11</td>
<td>0.41</td>
</tr>
<tr>
<td>Seattle</td>
<td>1.80</td>
<td>598</td>
<td>35</td>
<td>32</td>
<td>29</td>
<td>82.86</td>
<td>0.50</td>
</tr>
<tr>
<td>Zurich</td>
<td>1.13</td>
<td>475</td>
<td>115</td>
<td>99</td>
<td>79</td>
<td>68.70</td>
<td>4.21</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>77.45</strong></td>
<td><strong>1.92</strong></td>
<td><strong>8.04</strong></td>
<td><strong>1.10</strong></td>
<td><strong>8.67</strong></td>
<td><strong>1.50</strong></td>
<td><strong>3.85</strong></td>
</tr>
</tbody>
</table>

---

Fig. 6. Geospatial engineering of the predicted result: tagging buildings with sustainable rooftops based on detected greenery and solar panels on them.
Table 2
Green roof result by Area.

<table>
<thead>
<tr>
<th>Region</th>
<th>Area (km²)</th>
<th>Bldg. Area (m²)</th>
<th>Truth (m²)</th>
<th>Pred. (m²)</th>
<th>Matching (m²)</th>
<th>%Matching</th>
<th>%FP</th>
<th>%Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin 1</td>
<td>1.20</td>
<td>673152</td>
<td>234785</td>
<td>234616</td>
<td>180603</td>
<td>76.92</td>
<td>8.02</td>
<td>26.83</td>
</tr>
<tr>
<td>Berlin 2</td>
<td>1.21</td>
<td>533421</td>
<td>76766</td>
<td>113235</td>
<td>73418</td>
<td>95.64</td>
<td>7.46</td>
<td>13.76</td>
</tr>
<tr>
<td>Melbourne</td>
<td>4.50</td>
<td>1042319</td>
<td>61436</td>
<td>108072</td>
<td>60071</td>
<td>97.78</td>
<td>4.61</td>
<td>5.76</td>
</tr>
<tr>
<td>New York 1</td>
<td>1.02</td>
<td>364340</td>
<td>123435</td>
<td>101075</td>
<td>91310</td>
<td>73.97</td>
<td>2.68</td>
<td>25.06</td>
</tr>
<tr>
<td>New York 2</td>
<td>2.05</td>
<td>889578</td>
<td>263599</td>
<td>245660</td>
<td>245431</td>
<td>93.11</td>
<td>0.03</td>
<td>27.59</td>
</tr>
<tr>
<td>Paris</td>
<td>1.70</td>
<td>866545</td>
<td>91143</td>
<td>104551</td>
<td>79547</td>
<td>87.28</td>
<td>2.89</td>
<td>9.18</td>
</tr>
<tr>
<td>Seattle</td>
<td>1.80</td>
<td>771931</td>
<td>76402</td>
<td>70469</td>
<td>68408</td>
<td>89.54</td>
<td>0.27</td>
<td>8.86</td>
</tr>
<tr>
<td>Zurich</td>
<td>1.13</td>
<td>221836</td>
<td>97946</td>
<td>92136</td>
<td>84077</td>
<td>86.45</td>
<td>3.36</td>
<td>38.17</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>87.59</strong></td>
<td>3.66</td>
<td>19.40</td>
</tr>
</tbody>
</table>

Table 3
Solar roof result by Count (building classification).

<table>
<thead>
<tr>
<th>Region</th>
<th>Area (km²)</th>
<th>Bldg. Count</th>
<th>Truth</th>
<th>Pred.</th>
<th>Matching</th>
<th>%Matching</th>
<th>%FP</th>
<th>%Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin 1</td>
<td>2.31</td>
<td>2251</td>
<td>87</td>
<td>90</td>
<td>72</td>
<td>82.76</td>
<td>0.80</td>
<td>3.20</td>
</tr>
<tr>
<td>Berlin 2</td>
<td>1.22</td>
<td>754</td>
<td>31</td>
<td>43</td>
<td>28</td>
<td>90.32</td>
<td>1.99</td>
<td>3.71</td>
</tr>
<tr>
<td>Copenhagen</td>
<td>1.64</td>
<td>1397</td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>77.78</td>
<td>0.36</td>
<td>0.50</td>
</tr>
<tr>
<td>Marseille</td>
<td>3.51</td>
<td>5639</td>
<td>73</td>
<td>79</td>
<td>72</td>
<td>98.63</td>
<td>0.12</td>
<td>1.28</td>
</tr>
<tr>
<td>Melbourne</td>
<td>4.52</td>
<td>1079</td>
<td>36</td>
<td>59</td>
<td>36</td>
<td>100.00</td>
<td>2.13</td>
<td>3.34</td>
</tr>
<tr>
<td>New York 1</td>
<td>1.54</td>
<td>1414</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>100.00</td>
<td>0.00</td>
<td>1.84</td>
</tr>
<tr>
<td>New York 2</td>
<td>5.52</td>
<td>3079</td>
<td>28</td>
<td>27</td>
<td>24</td>
<td>85.71</td>
<td>0.10</td>
<td>0.78</td>
</tr>
<tr>
<td>Washington D.C.</td>
<td>2.10</td>
<td>562</td>
<td>37</td>
<td>51</td>
<td>37</td>
<td>100.00</td>
<td>2.49</td>
<td>6.58</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>91.90</strong></td>
<td>1.00</td>
<td>2.65</td>
</tr>
</tbody>
</table>

Table 4
Solar roof result by Area.

<table>
<thead>
<tr>
<th>Region</th>
<th>Area (km²)</th>
<th>Bldg. Area (m²)</th>
<th>Truth (m²)</th>
<th>Pred. (m²)</th>
<th>Matching (m²)</th>
<th>%Matching</th>
<th>%FP</th>
<th>%Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin 1</td>
<td>2.31</td>
<td>809753</td>
<td>74151</td>
<td>113425</td>
<td>69314</td>
<td>93.48</td>
<td>5.45</td>
<td>8.56</td>
</tr>
<tr>
<td>Berlin 2</td>
<td>1.22</td>
<td>669506</td>
<td>75217</td>
<td>106371</td>
<td>72403</td>
<td>96.26</td>
<td>5.07</td>
<td>10.81</td>
</tr>
<tr>
<td>Copenhagen</td>
<td>1.64</td>
<td>465472</td>
<td>18885</td>
<td>21211</td>
<td>15386</td>
<td>81.47</td>
<td>1.25</td>
<td>3.31</td>
</tr>
<tr>
<td>Marseille</td>
<td>3.51</td>
<td>612603</td>
<td>63316</td>
<td>67446</td>
<td>61976</td>
<td>97.88</td>
<td>0.89</td>
<td>10.12</td>
</tr>
<tr>
<td>Melbourne</td>
<td>4.52</td>
<td>1042319</td>
<td>76263</td>
<td>184071</td>
<td>76263</td>
<td>100.00</td>
<td>10.34</td>
<td>7.32</td>
</tr>
<tr>
<td>New York 1</td>
<td>1.54</td>
<td>449264</td>
<td>41442</td>
<td>41442</td>
<td>41442</td>
<td>100.00</td>
<td>0.00</td>
<td>9.22</td>
</tr>
<tr>
<td>New York 2</td>
<td>5.52</td>
<td>1746588</td>
<td>136615</td>
<td>123048</td>
<td>113901</td>
<td>83.37</td>
<td>0.52</td>
<td>6.52</td>
</tr>
<tr>
<td>D.C.</td>
<td>2.10</td>
<td>905803</td>
<td>161260</td>
<td>268043</td>
<td>161260</td>
<td>100.00</td>
<td>11.79</td>
<td>17.80</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>96.25</strong></td>
<td>4.71</td>
<td>10.20</td>
</tr>
</tbody>
</table>

Fig. 7. Examples of green roof model confusions.
average %FP reached 1.92%. Comparing the Area of the prediction with the
ground truth, accuracy increased to 87.59% but %FP also increased,
to 3.66%.

In both cases, %FPs are under 5% for most of the areas except Berlin
1 (Stadtmitte) and Berlin 2 (Friedrichshain). The percentage of matching
(%Matching) is generally above 80% in terms of Area except for New
York 1 (Hudson Square) and above 70% in terms of Count except for
Zurich (Alstetten) and New York City 1 (Hudson Square).

Looking closely into these regions, we observed that some degree of
confusion had occurred for the model. As shown in Fig. 7, some roofs in
Berlin (7a), have a light green and ‘furry’ appearance which the model
confuses as a green roof. In the case of Zurich (7b, 7c), the model failed
to recognise seasonal changes in green roof colour (browning of leaves
in autumn) in the presence of solar panels. It successfully picked up some
badly maintained green roofs (dark brown in colour with a furry texture)
but confused it with dark brown pitches roofs in the vicinity. Finally, the
model also failed to identify some green roofs in smaller buildings such
as in New York City (7d). This issue might be caused by the removal of
noise in the prediction pipeline where some true positives are sacrificed
to reduce the percentage of false positives (in the case of New York City,
removing noise contributed to an average 67.66% reduction of false
positives (4.17% to 1.35%) at the cost of 5.48% decrease of accuracy
(88.38% to 83.54%)).

3.3.3. Solar roof results

Tables 3 and 4 present the prediction results of solar roofs against the
manually labelled ground truth. Using the same method of calculation as
that of the green roofs, the accuracy of the solar roof model in terms of
Count reached 91.90% and an average %FP of 1.00%. Accuracy in terms
of calculating the Area increased to 94.06% while %FP also increased to
4.42%. This performance suggests that larger buildings might have
higher chances of having features that confuse the model into catego-
rising it as a solar roof building.

There is some confusion between building skylights and solar panels as
they have almost identical features and difficult to distinguish also by
humans. This issue is the major cause of the rise of %FP in Berlin (8a)
and Washington D.C. (8b) as illustrated in Fig. 8. Another factor that
causes confusion is the angle of projection of tall buildings. When the
perspective of the satellite photo is not ortho-rectified, the image may
show the building facade where the glass windows can be confused with
solar panels as seen in Marseille (8c). Lastly, objects that share similar
features with solar panels such as dark rectangles with white border
might also be counted as false positives. These objects could be rooftop
sunshades or stadium seats such as the example in Melbourne (8d).

Fig. 8. Examples of solar roof model confusions.

3.3.4. Evaluation conclusion

The evaluation results suggest that the percentage of false positives
for both solar and green roofs are within acceptable levels, and the two
models are sufficiently accurate to provide insightful city-wide spatial
and quantitative analyses, both in the location and size of sustainable
instruments on rooftops. In future work, it would be worthwhile to
expand our approach by integrating other forms of data. For example,
point clouds obtained from airborne Light Detection and Ranging
(LiDAR) might provide further information that would aid the classifi-
cation (Zhao, Myint, Wentz, & Fan, 2015), such as 3D geometry and
intensity of the returns, which may hint at the characteristics of the
surface.

4. Roofpedia registry and index

The results from the predictions on the 17 cities (Section 3) are used
to create (i) a prototype of a sustainable roof registry, i.e. open dataset
and web viewer (Section 4.1); and (ii) an index to assess the penetration
of sustainable roof typologies in cities (Section 4.2).

4.1. Roof registry

The roof registry is derived for the 17 cities and it is released as an
openly accessible interactive map (Fig. 9) that visualises the spatial
distribution of green and solar roofs in two styles. At smaller scale levels,
the centroids of the buildings are shown as yellow (solar roofs) or green
(green roofs) dots for easy identification of the city-wide distribution
of sustainable roofs and addresses can be searched. Further, the generated
data can be visualised as a heatmap to understand the density of sustainable
rooftops.

4.2. Roofpedia Index — a measure for rooftop sustainability across cities

The idea of the index is comparable to the one described by the
Treepedia project (MIT Senseable City Lab, 2016), which measures the
street canopy cover in different cities with a Green View Index (Li et al.,
2015; Seiferling et al., 2017). In a similar fashion, we create a city-scale
summarised index for the roofscape, which assigns scores based on the
density and area coverage of green and solar roofs in a city. We hope to
raise curiosity and awareness in rooftop utilisation among the public and
also provide planners and researchers with a barometer to access the
effectiveness of rooftops policies in different cities. As indicated in the
previous section, the percentage cover of green and solar roofs are
calculated based on Count and Area separately. Calculating the percentage by Count helps us to understand the degree of adoption by individual owners while calculating the percentage by Area indicates the extensiveness of the cover. For example, the evaluation subset in New York City (NYC2) has a low coverage of green roofs in terms of Count (3.51%) but a much higher coverage in terms of Area (27.59%). Such discrepancy suggests that green roofs in NYC are mostly installed on several large buildings while smaller buildings have yet to adopt green roofs. Conversely, although the evaluated subsets Marseille and Berlin 2 have a similar percentage of the cover by area, Berlin 2 has 2.9 times higher % by Count (3.72% vs 1.28%). This difference suggests that the density of buildings with solar roofs is higher in Berlin and more homeowners and developers have adopted solar panels on their roofs.

The Roofpedia Index recognises that both factors are important and takes them into consideration. %Count (Eq. (4)) and %Area (Eq. (5)) for both solar and green roof penetration are calculated for each city and normalised with Eqs. (6) and (7). A combined score is calculated for both solar and green roofs with Eqs. (8) and (9). An overall score that aggregates both the Solar Score and Green Score is then calculated by taking the average of the sum of two scores as in Eq. (10).

\[
\%\text{Count} = 100 \times \left( \frac{\text{Matching Count (TP)}}{\text{Total Count}} \right)
\]  

\[
\%\text{Area} = \left( \frac{\text{Cover Area (PA)}}{\text{Project Area (PA)}} \right)
\]  

\[
\%\text{Solar} = \left( \frac{\text{Solar Score}}{\text{Solar Score} + \text{Green Score}} \right) \times 100
\]  

\[
\%\text{Green} = \left( \frac{\text{Green Score}}{\text{Solar Score} + \text{Green Score}} \right) \times 100
\]  

\[
\text{Overall Score} = \frac{\text{Solar Score} + \text{Green Score}}{2}
\]  

(a) Mapping buildings with sustainable rooftops (example for New York City).

(b) A visualisation for understanding the spatial density of solar roofs (example for Singapore).

Fig. 9. Visualisation of the spatial distribution of buildings with sustainable rooftops in different styles. Basemap: OpenStreetMap contributors and Mapbox.
Green Roof Index.

Table 5

<table>
<thead>
<tr>
<th>Rank</th>
<th>City</th>
<th>Bldg.</th>
<th>Green Roofs</th>
<th>%Count</th>
<th>%Area</th>
<th>Score by Count</th>
<th>Score by Area</th>
<th>Green Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zurich</td>
<td>18440</td>
<td>5760</td>
<td>31.2</td>
<td>41.6</td>
<td>100</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>Berlin</td>
<td>28077</td>
<td>3899</td>
<td>13.6</td>
<td>24.8</td>
<td>43</td>
<td>58</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>New York</td>
<td>34985</td>
<td>1924</td>
<td>5.6</td>
<td>17.2</td>
<td>17</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>Copenhagen</td>
<td>15505</td>
<td>725</td>
<td>4.7</td>
<td>13.1</td>
<td>14</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>Paris</td>
<td>74014</td>
<td>2766</td>
<td>3.7</td>
<td>11.2</td>
<td>11</td>
<td>25</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>San Diego</td>
<td>28303</td>
<td>373</td>
<td>1.3</td>
<td>11</td>
<td>4</td>
<td>24</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>San Jose</td>
<td>182314</td>
<td>2650</td>
<td>1.5</td>
<td>10.2</td>
<td>4</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>Phoenix</td>
<td>15217</td>
<td>245</td>
<td>1.6</td>
<td>9.9</td>
<td>4</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Melbourne</td>
<td>16809</td>
<td>258</td>
<td>1.5</td>
<td>8.4</td>
<td>4</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>Las Vegas</td>
<td>20389</td>
<td>192</td>
<td>0.9</td>
<td>7.8</td>
<td>2</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>Seattle</td>
<td>81044</td>
<td>347</td>
<td>0.4</td>
<td>5.6</td>
<td>1</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>Los Angeles</td>
<td>50978</td>
<td>419</td>
<td>0.8</td>
<td>4.7</td>
<td>2</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>Luxembourg City</td>
<td>11131</td>
<td>125</td>
<td>1.1</td>
<td>2.7</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>Portland</td>
<td>122900</td>
<td>302</td>
<td>0.2</td>
<td>3.7</td>
<td>0</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>San Francisco</td>
<td>165814</td>
<td>389</td>
<td>0.2</td>
<td>2.6</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>Vancouver</td>
<td>163818</td>
<td>108</td>
<td>0.1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

%Area = 100 × (Matching Area (TP)/Total Area) (5)

Score by Count = 100 × (%Count – min(%Count))/(max(%Count) – min(%Count)) (6)

Score by Area = 100 × (%Area – min(%Area))/(max(%Area) – min(%Area)) (7)

Solar Score = (Score by Count\textsubscript{Solar} + Score by Area\textsubscript{Solar})/2 (8)

Green Score = (Score by Count\textsubscript{Green} + Score by Area\textsubscript{Green})/2 (9)

Overall Score = (Solar Score + Green Score)/2 (10)

Tables 5 and 6 show the ranking of Solar and Green Roof coverage of the 17 cities and Table 7 shows the combined ranking of both scores, deriving a holistic index indicating the proliferation of sustainable roofscapes in cities.

The results offer material for discussion and further deliberations. Looking at the ranking for green roofs, Zurich tops the chart with 41.6% coverage in Area and 31.2% cover in Count. This impressive lead in green roof coverage affirms the efforts taken by the Zurich City Government in making Green Roofs mandatory for all new buildings since 1991 ( Tierversorgungsdepartement, 2020 ). Berlin comes in second place with approximately 24.8% in Area and 13.6% in Count. This result also presents a substantial lead compared to the rest of the cities in the study and echoes the long tradition of green roofs in Berlin as early as the beginning of the 19th century ( Ahrendt, 2007 ).

New York ranks highest in the U.S. with the largest green roof cover. With the new Climate Mobilization Act ( NYC Mayor’s Office of Sustainability, 2019 ), which made solar or green roofs mandatory on all new construction as well as buildings undertaking major roof renovations in the cities, it is highly likely that New York will continue to maintain its edge in a sustainable roofscape. For example, although Vancouver did not perform well in our index, it is nevertheless consistently ranked as one of the most sustainable cities in the world. According to the Sustainable Cities Index 2016 ( Arcadis, 2016 ), Vancouver ranks highest among North American cities in terms of environmental sustainability. Further, according to Treepedia, Vancouver is one of the cities highest on the list in terms of green canopy. Finally, Vancouver has access to plenty of hydropower, which provides 25% of the city’s energy need alone, and has plans to derive 100% of the energy used from renewable sources before 2050 ( City of Vancouver, 2015 ). Therefore, we believe that the Roofpedia Index complements existing sustainability indices (e.g., ( Phillis, Kouikoglou, & Verdugo, 2017 ) ) by adding a new dimension of consideration in assessing the overall sustainability of a city.

5. Discussion and limitations

We have demonstrated the feasibility and reliability of a scalable method to identify sustainable roofscapes using computer vision and geospatial techniques. The result from the 17 cities has also revealed insights into the distribution and penetration of sustainable roofscapes in these urban areas. Besides the perennial importance of space utilisation in the built environment and the positive role of greenery and solar panels, our work is also important at a time of the increased collocation of photovoltaic and farming systems, and the growing importance of urban agriculture ( Tablada et al., 2018; Ciriminna et al., 2019; Langemeyer, Madrid-Lopez, Beltran, & Mendez, 2021 ).

We have shown that a Convolutional Neural Network could be an alternative approach in roof typology detection apart from traditional classification methods on satellite images. The advantage over the traditional methods is that CNNs are more forgiving of image quality. Where traditional classification methods require multi-spectral satellite imagery that are only available for select cities, our method can cover more cities with RGB images that are available at a reasonable degree of resolution. As such data is becoming available for more locations,
However, one challenge we have encountered in the process is the impact of inconsistent image quality across cities on prediction accuracy. Running the prediction pipeline on cities with lower quality images results in excessive noise, rendering the prediction unreliable. For example, when predicting rooftop greenery in Singapore, the targets were too blurry even for humans to discern. Fortunately, due to their darker colours and more defined edges, solar panels were much more perceptible than green roofs, thus, we have included Singapore in the predicted dataset to make sure the results are not exceptionally skewed.

Our work has further benefits for use cases in different domains, predominantly in energy and urban planning. For example, BIPV technology have created panels that mimic the appearance of traditional building materials (Verbeene et al., 2014). Therefore, even if such panels are installed on building rooftops, it is difficult for both human and the model to detect their presence. Another consideration in interpreting the result of the predictions is the correlation between the percentage of cover area and the actual roof area covered by greenery and solar panels. While this is largely true, there are cases where the actual coverage of rooftop greenery and solar panels only covers a small proportion of the roof area.

On the note of building footprints, it is important to keep in mind that they are sourced from OpenStreetMap, a volunteered geospatial dataset with global coverage. While the quality is sufficiently good, the crowdsourcing nature and myriads of contributors across the world may have different approaches in mapping buildings. For example, a set of adjacent buildings (e.g. terraced houses) might be mapped as a single building by one mapper, while another one might map them as multiple individual buildings. These different modelling approaches might have an impact on the results.

While interpreting the prediction results, we did not take false positives into consideration. As shown in Section 3.3, %FP fluctuates across cities and could potentially reach 11.79% in terms of Area and 2.49% in terms of Count. Although the average %FP is well below 2% in terms of count and 4% in terms of Area, a city with a high %FP could be placed at a higher position in the index mistakenly, affecting interpretation of results. Therefore, quick manual inspections are still required on the predicted dataset to make sure the results are not exceptionally skewed.

On a broader scope, this line of research also grants an opportunity for the public to participate in the mapping process. Just like OpenStreetMap, an interactive platform can be created for the public to verify the result of the AI models to eliminate false positives and the models could be then retrained with a larger dataset to improve its accuracy and precision iteratively.

It is important to recognise that this work is focused on rooftops, and it is geared towards elevating the prominence of the roofscape in the sustainability context. However, they are not the only venue for greener and PV panels. For example, solar panels may be installed on walls of buildings (Saretta, Caputo, & Frontini, 2020), but also on other platforms in cities, e.g. noise barriers (Zhong et al., 2021) and vehicles (Brito, Santos, Moura, Pera, & Rocha, 2021). The same goes for urban greening efforts, i.e. vertical spaces of buildings such as walls (Tan, Wong, & Junuf, 2014; Song, Tan, & Tan, 2018; Palliwal et al., 2021; Huang, Tan, Lu, & Wong, 2021). Furthermore, innovations in BIPV technology have created panels that mimic the appearance of traditional building materials (Verbeene et al., 2014). Therefore, even if such panels are installed on building rooftops, it is difficult for both human and the model to detect their presence. A possible direction for future work is to complement our work with other forms of urban data that provide another perspective and cover further venues, e.g. using street view imagery (Tang & Long, 2019; Ding, Fan, & Gong, 2021).

Our work has further benefits for use cases in different domains, predominantly in energy and urban planning. For example, the city-wide dataset that we have generated may aid researchers to focus on additional cities can be added to the registry and index. Our modular pipeline does not only allow new cities to be added, but also new roof typologies. Adding such requires training custom models, which can be effortlessly and create labels without inferring from the surrounding context, the model would be able to automate the detection accurately.

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Solar Roof Index.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>City</td>
</tr>
<tr>
<td>1</td>
<td>Las Vegas</td>
</tr>
<tr>
<td>2</td>
<td>Zurich</td>
</tr>
<tr>
<td>3</td>
<td>Singapore</td>
</tr>
<tr>
<td>4</td>
<td>Phoenix</td>
</tr>
<tr>
<td>5</td>
<td>Melbourne</td>
</tr>
<tr>
<td>6</td>
<td>Berlin</td>
</tr>
<tr>
<td>7</td>
<td>Copenhagen</td>
</tr>
<tr>
<td>8</td>
<td>New York</td>
</tr>
<tr>
<td>9</td>
<td>Paris</td>
</tr>
<tr>
<td>10</td>
<td>San Diego</td>
</tr>
<tr>
<td>11</td>
<td>Los Angeles</td>
</tr>
<tr>
<td>12</td>
<td>Seattle</td>
</tr>
<tr>
<td>13</td>
<td>San Jose</td>
</tr>
<tr>
<td>14</td>
<td>Portland</td>
</tr>
<tr>
<td>15</td>
<td>San Francisco</td>
</tr>
<tr>
<td>16</td>
<td>Luxembourg City</td>
</tr>
<tr>
<td>17</td>
<td>Vancouver</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7</th>
<th>Roofpedia Index (overall ranking).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>City</td>
</tr>
<tr>
<td>1</td>
<td>Zurich</td>
</tr>
<tr>
<td>2</td>
<td>Berlin</td>
</tr>
<tr>
<td>3</td>
<td>Las Vegas</td>
</tr>
<tr>
<td>4</td>
<td>Phoenix</td>
</tr>
<tr>
<td>5</td>
<td>Melbourne</td>
</tr>
<tr>
<td>6</td>
<td>New York</td>
</tr>
<tr>
<td>7</td>
<td>Copenhagen</td>
</tr>
<tr>
<td>8</td>
<td>Paris</td>
</tr>
<tr>
<td>9</td>
<td>San Diego</td>
</tr>
<tr>
<td>10</td>
<td>Los Angeles</td>
</tr>
<tr>
<td>11</td>
<td>San Jose</td>
</tr>
<tr>
<td>12</td>
<td>Seattle</td>
</tr>
<tr>
<td>13</td>
<td>Portland</td>
</tr>
<tr>
<td>14</td>
<td>Luxembourg City</td>
</tr>
<tr>
<td>15</td>
<td>San Francisco</td>
</tr>
<tr>
<td>16</td>
<td>Vancouver</td>
</tr>
</tbody>
</table>
A.N. Wu and F. Biljecki

6. Conclusion

As urban population continues to grow, cities need effective measures to improve their environment for both the residents and the surrounding ecology. The roofscape has a demonstrated potential in contributing to a city’s sustainability, such as doubling its use by greening them and having solar panels installed on them. While most studies so far have been focused on estimating the potential of rooftops, Roofpedia enables us to understand the current and actual status of the rooftops, i.e. how many of them are currently occupied with solar panels and for extensive greeneries. Such information advances the state of the art, complements existing studies (including those estimating the potential), and it is useful for several purposes, e.g. gauging the efficiency of government policies, tracking and monitoring pledges by businesses, verifying the use of subsidies, estimating the current carbon offsetting capacity of cities and benchmarking them, and determining how much of the potential has already been realised.

By mapping the distribution of sustainable rooftops across cities in the world, we hope to raise public awareness on the importance of the role of rooftops in supporting sustainable development. We expect that interest in rooftops will be gaining more traction in urban analytics, and thus, our work and the generated data may provide a solid basis for follow-up studies. While our work is focused on green roofs and those with solar panels, we believe that it paves the way for future work in large-scale identification of further rooftop typologies using computer vision.

Besides the ideas discussed in Section 5, further opportunities for future work are many. For example, our method can be used to study the temporal evolution of sustainable rooftops, as its landscape is increasingly dynamic thanks to decreasing costs of technology and widely available satellite images (Zhang et al., 2019). The presented approach can be used to analyse satellite imagery from different epochs to provide the difference and evolution in green and solar installations over time, adding the temporal aspect to this project. Such research line would be welcomed by different use cases, such as understanding the effect of urban policies. In fact, our work suggests that cities that rank highly in the index usually have incentives for green or solar roofs in place, indicating that Roofpedia could be a useful instrument for urban planners and policymakers.

Another idea we have in mind is to expand our work with a crowdsourced platform, bringing it closer to the public and enable enriching our dataset with additional information that might not be available from satellite imagery, and with information that may widen the work beyond the sustainability context. For example, it might be beneficial to expand the work into providing information about social uses of rooftops through a web mapping service for the public who would like to explore the roofscape around them, e.g. to find rooftops in their neighbourhood that have public access and a nice view.

Further options include enrichment of building datasets using the images that we generated. For example, as OpenStreetMap and 3D building models support modelling and integrating features such as solar panels (Biljecki et al., 2021; Stowell et al., 2020), it might be worthwhile exploring whether our project could serve as a source for tagging buildings with green or solar installations and/or modelling their geometry.

References


CRediT authorship contribution statement

Abraham Noah Wu: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. Filip Biljecki: Methodology, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

The authors gratefully acknowledge the sources of the input datasets and the discussions with Yoong Shin Chow (NUS Urban Analytics Lab). The reviewers’ insightful and helpful comments improved the quality of the article and are highly appreciated. This research is part of the project Large-scale 3D Geospatial Data for Urban Analytics, which is supported by the National University of Singapore under the Start-Up Grant R-295-000-171-133.