Advances in Applied Energy 17 (2025) 100212

Contents lists available at ScienceDirect

Advances in Applied Energy

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

Advancing building facade solar potential assessment through AIoT, GIS, and meteorology synergy

Kechuan Dong [a,1], Qing Yu [b,1], Zhiling Guo [c,d,*], Jian Xu [c,d], Hongjun Tan [c,d], Haoran Zhang [b], Jinyue Yan [c,d,*]

[a] Institute of Industrial Science, The University of Tokyo, Komaba, 153-8505, Tokyo, Japan
[b] School of Urban Planning and Design, Peking University Shenzhen Graduate School, Shenzhen, 518055, Guangdong, China
[c] Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hong Kong, China
[d] The International Centre of Urban Energy Nexus, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Dataset link: https://www.google.com/maps/, https://www.mapbox.com/, https://doi.org/10.1016/j.rser.2018.03.003

Keywords:
Solar energy
BIPV
Renewable energy
Facade parsing
GIScience
Deep learning

ABSTRACT

The assessment of building solar potential plays a pivotal role in Building Integrated Photovoltaics (BIPV) and urban energy systems. While current evaluations predominantly focus on rooftop solar resources, a comprehensive analysis of building facade BIPV potential is often lacking. This study presents an innovative methodology that harnesses state-of-the-art Artificial Intelligence of Things (AIoT) techniques, Geographic Information Systems (GIS), and Meteorology to develop a model for accurately estimating spatial–temporal building facade BIPV potential considering 3 Dimension (3D) shading effect. Here, we introduce a zero-shot Deep Learning framework for detailed parsing of facade elements, utilizing cutting-edge techniques in Large-scale Segment Anything Model (SAM), Grounding DINO (Detection Transformer with improved denoising anchor boxes), and Stable Diffusion. Considering urban morphology, 3D shading impacts, and multi-source weather data enables a meticulous estimation of solar potential for each facade element. The experimental findings, gathered from a range of buildings across four countries and an entire street in Japan, highlight the effectiveness and applicability of our approach in conducting comprehensive analyses of facade solar potential. These results underscore the critical importance of integrating shadow effects and detailed facade elements to ensure accurate estimations of PV potential.

1. Introduction

The increasing urgency of addressing global environmental challenges, such as climate change and energy security, has underscored the importance of transitioning to renewable energy sources [1,2]. At Conference of the Parties 28 (COP28) in Dubai, United Arab Emirates (UAE), over 130 national governments pledged to triple the global installed renewable energy capacity to 11,000 GW by 2030 [3]. Solar energy stands out mainly through photovoltaics (PV) technology due to its vast potential and sustainability. As per the International Energy Agency (IEA), in 2023, solar PV alone constituted three-quarters of global renewable capacity additions [4]. This substantial growth not only lessens dependence on fossil fuels but also aids in curbing greenhouse gas emissions, paving the way for a cleaner, more sustainable energy landscape [5].

In urban settings, establishing extensive centralized solar power plants is frequently unfeasible due to space limitations and elevated land expenses [6]. In contrast, distributed Building Integrated Photovoltaics (BIPV) systems emerge as a practical alternative [7]. According to the requirements, BIPV solutions are available in various forms, including semi-transparent glass facades, cladding, and roof tiles [8]. By seamlessly integrating PV panels into building structures, BIPV systems not only supplement the energy needs of buildings [9] but also contribute significantly to the penetration of renewable energy in urban environments. They offer flexible installation options for both urban renewal and new construction projects, optimizing the use of typically underutilized surfaces like rooftops and facades [10,11]. Research findings indicate that BIPV can significantly contribute to distributed energy generation, promoting a more sustainable and resilient urban energy infrastructure. Specifically, a case study of a 1970s residential building undergoing refurbishment showed that integrating BIPV in renovation scenarios could lead to energy efficiency improvements, with potential gains compared to the baseline of up to 122%. These

* Corresponding authors.
E-mail addresses: zhiling.guo@polyu.edu.hk (Z. Guo), j-jerry.yan@polyu.edu.hk (J. Yan).
[1] Equal contribution.

https://doi.org/10.1016/j.adapen.2025.100212
Received 26 September 2024; Received in revised form 27 November 2024; Accepted 14 January 2025
Available online 27 January 2025

scenarios also demonstrated economic benefits, including a payback period of 14–18 years and an internal rate of return (IRR) between 5.3% and 5.9%.[12].

Current methods have limitations: they often do not consider building facades, overlook important details, and fail to account for three-dimensional characteristics. Accurately evaluating solar potential in urban environments is crucial for maximizing the performance of BIPV systems. Research indicates that the potential for facade BIPV in cities like Beijing could reach 239% of the energy generation potential of rooftop PV systems [13]. However, conventional approaches have predominantly focused on rooftop PV assessments, frequently overlooking the significant opportunities for energy generation and architectural integration that building facades offer [14,15].

Assessing the solar potential of building facades poses several challenges compared to rooftops. Facades are significantly affected by complex interactions within urban environments, including dynamic shading from neighboring buildings. Unlike rooftops, facades utilize a variety of materials, from glass to different types of cladding, each requiring a specific BIPV solution. This diversity complicates the accurate modeling of facade PV potential. Estimating solar potential has traditionally relied on simple geometric parameters, but these methods often overlook complex shading and surface interactions, leading to inaccurate energy yield predictions. Advanced 3 Dimension (3D) modeling techniques, like Light Detection And Ranging (LiDAR) and Unmanned Aerial Vehicle (UAV), improve accuracy but are costly, slow to update, and lack the semantic detail needed for selecting optimal PV installation sites [16–18]. Geographic Information Systems (GIS) and deep learning approaches show promise in rooftop assessments, yet struggle with facade evaluations due to their reliance on simplified assumptions and limited capacity for parsing crucial semantic information [19–22].

Furthermore, achieving accurate parsing of facade elements without access to a substantial training dataset is a notable challenge, leading to inaccurate PV potential estimates and suboptimal system designs [23,24]. The complexities of urban morphology – including varying shading effects influenced by seasonal and temporal changes – are frequently excluded from these assessments, further compromising the accuracy and practical value of the analyses [25]. The multifaceted nature of facade PV potential assessment requires a more advanced and context-aware approach, which traditional methods, relying heavily on remote sensing and simplified assumptions, are unable to provide (see Figs. 1 and 3).

To address these challenges, we introduce an innovative methodology, as illustrated in Fig. 1, that utilizes advanced AIoT [26], GIS [27], and meteorology to intricately estimate spatiotemporal facade BIPV potential across individual facades, buildings, city blocks, and entire streetscapes. Our approach begins with the analysis of RGB images captured by IoT devices, using Grounding DINO (Detection Transformer with improved denoising anchor boxes) [28] and zero-shot Segment Anything Model (SAM) [29] to automatically identify and mask obstructions in photographic data without requiring a large amount of training data. Subsequently, a Stable Diffusion-based inpainting model [30] is employed to reconstruct building surfaces, facilitating the precise segmentation of windows and walls to pinpoint suitable areas for PV panel installation.

A key innovation of this study is the differentiation between various facade elements – such as glass windows and opaque walls – and the application of distinct photovoltaic technologies accordingly. Specifically, we incorporate semi-transparent BIPV systems for window elements, which allow for natural light transmission, and opaque BIPV modules for walls, which provide higher energy conversion efficiency. By modeling these facade elements separately, our approach ensures a more precise assessment of the potential for solar energy harvesting, considering the unique properties of each type of surface material.

By developing a detailed 3D simulation model of urban structures integrated with multi-source weather data, geometry projection, and cumulative insolation, we conduct comprehensive analyses of PV potential that consider dynamic shading effects and material-specific efficiencies. The proposed assessment method demonstrates significant adaptability by evaluating facade parsing performance across regions with diverse architectural styles and urban configurations. By conducting evaluations at different spatial scales, ranging from individual buildings to entire streetscapes, our approach ensures robustness and scalability.

The major contributions of this study are fourfold:

1. **Framework for assessing building facade solar potential:** We developed a framework for assessing building facade solar potential that integrates AIoT, GIS, and meteorology, providing a holistic solution that goes beyond traditional rooftop-focused evaluations.
2. **Deep Learning technique for facade parsing:** We introduced a zero-shot Deep Learning technique leveraging advanced methods like Large-scale SAM, Grounding DINO, and Stable Diffusion, enabling precise parsing of facade elements without the need for extensive training data and enhancing analysis accuracy.
3. **Information fusion methodology from 2D and 3D aspects:** We created a methodology that utilizes 2D images and 3D footprints to estimate shadows on facade elements, advancing the analysis of building solar potential and offering insights into the impacts of shading on PV efficiency.
4. **3D simulation tool for BIPV potential:** We innovated a 3D simulation tool for spatial–temporal evaluation of building facade BIPV potential, incorporating urban morphology, 3D shading impacts, and multi-source weather data to enhance the depth and accuracy of our analysis.

2. Related works

In recent years, methods for assessing PV potential on buildings have been extensively researched, leading to the development of numerous techniques and tools to enhance accuracy and efficiency [31]. However, most of these studies have focused on refining the assessment of PV potential specifically for building rooftops. For instance, Singh and Banerjee [32] utilized GIS and high-resolution land use data to estimate the rooftop PV potential in Mumbai by analyzing satellite images. Hong et al. [33] employed GIS data for shadow vector modeling to simulate sunlight exposure on building rooftops at various times, combining this with weather data for comprehensive PV potential assessment. Several studies have integrated remote sensing image processing techniques, using high-resolution images and digital surface models (DSM) to analyze rooftop structures and available areas for PV potential assessment. For example, Hu et al. [34] combined this approach with the Normalized Difference Vegetation Index (NDVI) for object recognition, estimating the rooftop PV potential in Beijing. With the advent of deep learning technologies, researchers have increasingly processed remote sensing images using deep learning models to identify roof areas suitable for PV model installation and to assess their potential further. Assouline et al. [35] employed a random forest algorithm combined with GIS data to estimate the rooftop PV potential across Switzerland. This method predicted multiple variables, including available roof area, slope, orientation, and shading factors, providing a detailed assessment of PV power production potential. Walch et al. [36] integrated physical models with GIS and machine learning to evaluate the rooftop PV potential across Switzerland. Zhong et al. [37] developed a deep learning semantic segmentation model coupled with a spatial optimization sampling strategy, significantly reducing the labor cost of training roof extraction models while robustly extracting roofs of different architectural styles and land use areas for PV potential assessment. Additionally, Tan et al. [38] utilized generative AI for remote sensing imagery data augmentation, enhancing the robustness of the training model for PV potential estimation.

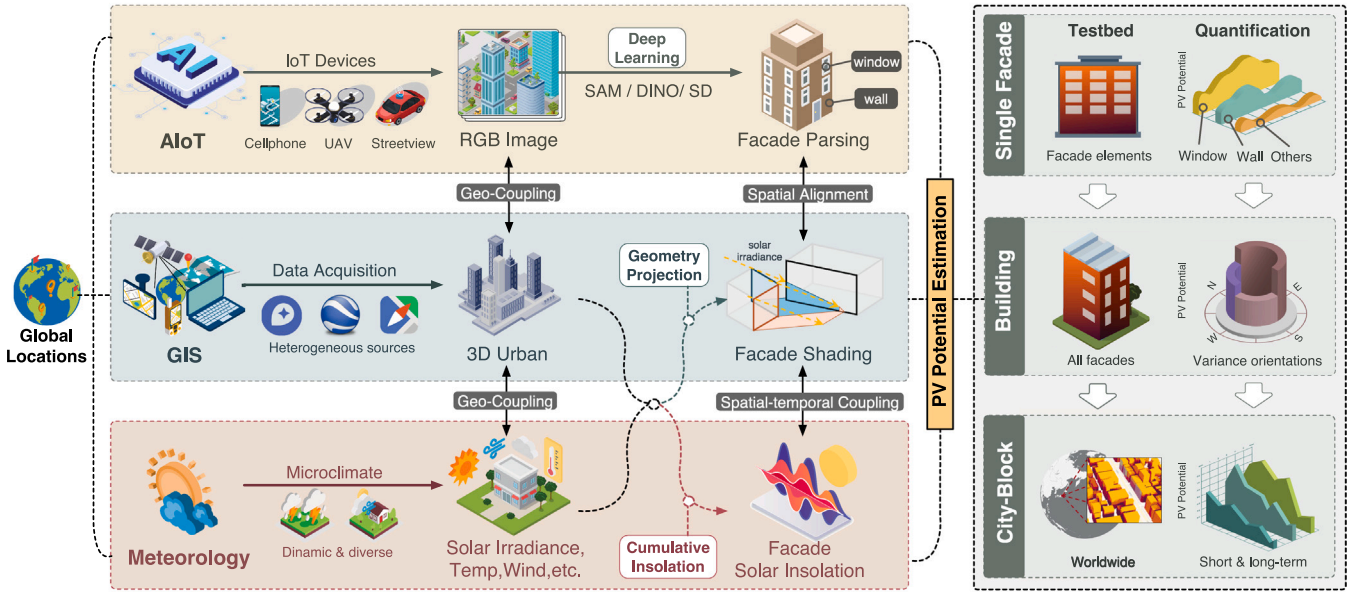

Fig. 1. Solar potential estimation on building facades through AIoT, GIS, and meteorology synergy.

Despite abundant research on rooftop PV potential, there is a notable scarcity of methods for analyzing the PV potential of building facades. This disparity is primarily due to the limited and expensive availability of 3D data sources required for facade analysis, in contrast to the various data sources available for rooftop PV potential assessment, such as remote sensing images. For instance, Prieto et al. [17] utilized LiDAR data and 3D modeling techniques to evaluate the solar potential of buildings. Similarly, Brito et al. [18] integrated LiDAR data with urban environments to assess the solar potential of buildings in Lisbon. However, the high cost and limited accessibility of LiDAR data significantly impede its widespread generalization and applicability. Cheng et al. [39] addressed this limitation by extending 2D building footprint data to 3D and employing shadow modeling for PV potential assessment. In more recent studies, Liu et al. [21] introduced a novel approach for assessing solar PV potential on both rooftops and facades in rural areas by constructing 3D building models from publicly available satellite images and GIS data. Chen et al. [22] modeling 3D building models by using remote sensing images to segment building footprints and estimate building heights to assess the PV potential of building facades and roofs. Nevertheless, LiDAR and 3D building footprint data lack semantic information about building facades, making it challenging to perform a detailed PV potential assessment of facades comparable to that of rooftops. Meanwhile, most studies continue to treat facade PV potential assessment as a byproduct of rooftop PV modeling. Even though some efforts have employed remote sensing and deep learning to expand the availability and precision of 3D data for more refined PV potential assessments, these improvements remain largely confined to rooftop evaluations [19,40,41].

Recent advancements in facade parsing leverage deep learning to enhance segmentation and detection accuracy in urban environments, tackling occlusions and facade variability. Early approaches, like the multi-scale contextual attention network by Zhang et al. [42], improved semantic segmentation by capturing fine details and multi-scale contextual information. Liu et al. [43] integrated architectural knowledge into deep learning models to detect facade elements, addressing appearance variability and enhancing detection in complex scenes. Zhou et al. [44] focused on occlusion handling, developing techniques to interpret partially visible elements accurately. Wang et al. [45] advanced this by introducing a progressive feature learning approach, incrementally refining feature understanding to manage occlusions better. Chen et al. [46] emphasized spatial relationships between elements, using contextual information to improve parsing accuracy. Past methods addressing occlusion issues often assume a certain correspondence in the structure of building facades or utilize deep neural networks to capture the contextual semantic information of facade structures for better element segmentation. However, relying on prior knowledge of facade elements to infer structural patterns frequently leads to erroneous virtual assumptions. For instance, with such assumptions, facade photos taken from oblique angles suffer from geometric distortion, causing the element arrangement to appear skewed in a 2D perspective. This distortion makes it challenging for models to make accurate inferences based on previously learned knowledge. Another issue is the diverse architectural styles of building elements, which vary significantly across countries, cities, and regions. To develop a robust and accurate facade element parsing model applicable in real-world scenarios, extensive training on a vast dataset encompassing various angles, facade styles, and forms is necessary. However, existing facade datasets do not meet these requirements, and the cost of training on such extensive data is prohibitively high. These limitations constrain the applicability of current facade parsing methods in real-world settings.

Thanks to recent developments in computer vision, driven by the development of foundational models, there has been a significant improvement in robust performance across various tasks. Grounding DINO, leveraging the Transformer architecture, excels in object detection by capturing contextual relationships within visual data, outperforming traditional CNNs in handling occlusions and varying object scales [28]. SAM, known for its semantic segmentation capabilities, has demonstrated superior performance in delineating complex scenes, including urban landscapes and natural environments [29]. Diffusion Models have introduced a novel approach to generative modeling, progressively refining noisy data to produce high-quality images, showcasing significant success in image synthesis and enhancement tasks [30]. Composite models, such as Grounded Segment Anything, integrate the strengths of these foundational models, combining object detection and segmentation to provide detailed and context-aware image analyses, proving effective in diverse environments and applications [47]. These foundational models and their composite variations underline the substantial potential for innovation in computer vision. Since these models have been trained on large-scale datasets, they can perform well in their respective tasks and have generalization characteristics in the real world, even without any further training for specific

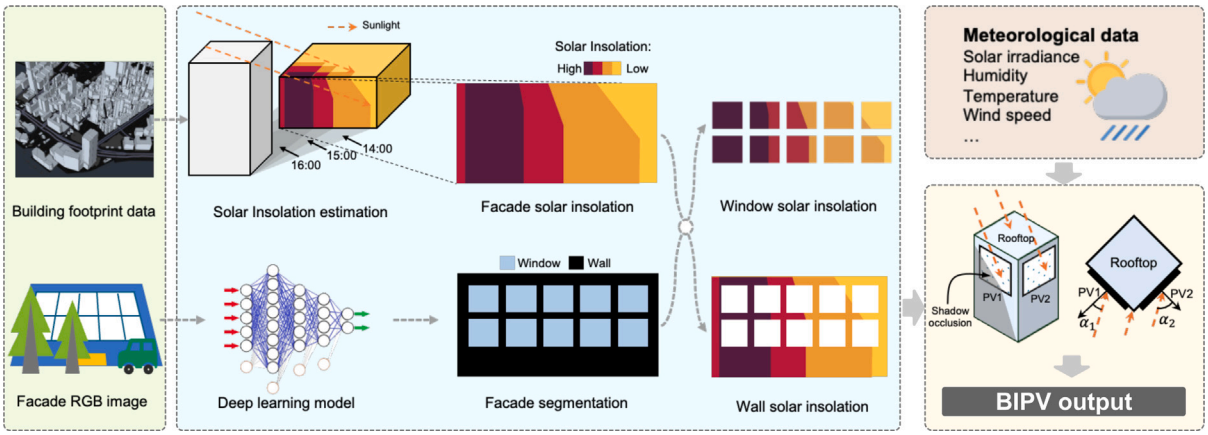

Fig. 2. Integrated methodology for estimating BIPV solar irradiance on building facade.

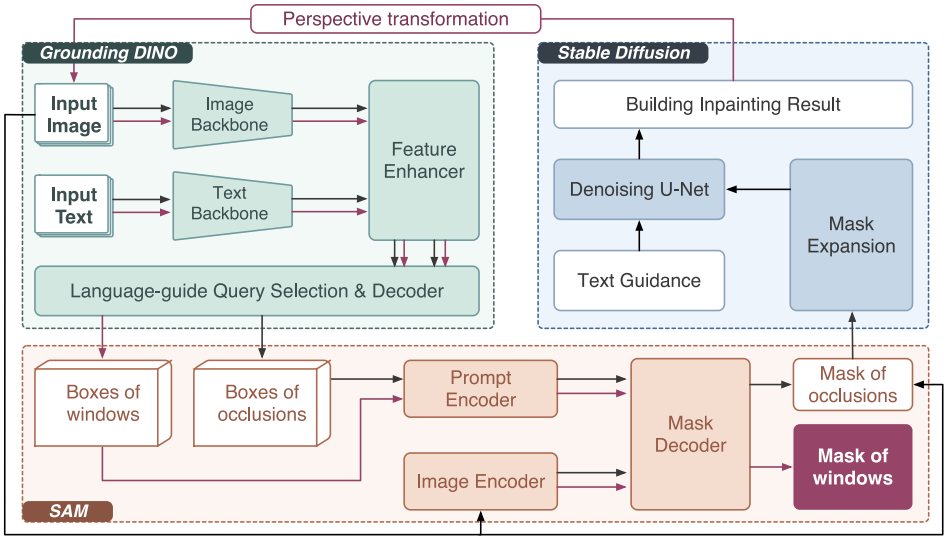

Fig. 3. Deep learning framework for building facade parsing.

tasks. Combining these models and utilizing their superior performance can achieve excellent usage effects even without additional training.

Therefore, while significant advances have been made in estimating BIPV potential through the integration of GIS, remote sensing, and deep learning technologies, the primary focus remains on rooftops, with very little on the facade. Although innovative methods have also been applied to assess the solar potential of building facades using 3D data, the scarcity of accessible, cost-effective, and semantically rich 3D data presents a major obstacle. This highlights the need for further development of new technologies and approaches to enable more comprehensive and accurate evaluations of facade BIPV potential.

3. Methodology

3.1. Methodology framework

This study introduces a novel method for assessing the PV potential of building facades by integrating advanced computer vision and 3D shadow modeling techniques, as depicted in Fig. 2. The methodology encompasses a series of processes aimed at estimating solar insolation and leveraging it for BIPV, primarily on facades.

The framework combines building footprint data, which in this context refers to the geometric representation of a building's outline as seen from above, with facade RGB images. Building footprint data provides the two-dimensional shape of the building at ground level, capturing its perimeter and spatial extent. When combined with building height information, these footprints enable the creation of three-dimensional building models. This data is foundational for accurately aligning facade RGB images with building geometry, as well as for modeling the shading interactions between urban building clusters. Leveraging 3D building footprint data, the method analyzes the interaction of sunlight with different segments of the building facade across various temporal scales, offering a detailed perspective on the facade's exposure to solar energy over time. By applying a deep learning model to parse facades in RGB images, we segment them into windows and walls. It is worth noting that in this research, we consider BIPV materials. Glass curtain walls and other glass-based materials are classified as windows, enabling a comprehensive evaluation of the window-to-wall ratio for each building. The estimation of solar insolation is then integrated with the facade segmentation results to ascertain the incident solar energy on different parts of the facade. By incorporating meteorological data into the analysis alongside the detailed building profile, the BIPV output from the facade is estimated, providing a comprehensive understanding of the solar energy potential of building facades. The following section of the methodology part will introduce the specific methods for data acquisition, facade parsing, shadow casting simulation, and BIPV output estimation.

3.2. Data acquisition and coupling

To validate the methodology proposed in this paper, we meticulously selected distinct building cases from diverse regions worldwide for in-depth analysis. Our dataset comprised 18 buildings situated in various cities and regions spanning Japan, Australia, Germany, Sweden, and the United States. Specifically, we conducted detailed estimations for individual buildings in Sydney, Australia; Göttingen, Germany; Stockholm, Sweden; and Lexington, United States. In the case of Japan, our analysis focused on the facades of buildings lining both sides of a street in central Tokyo.

The RGB images capturing the facades of these buildings were sourced from Google Street View, while the 3D building footprint data was extracted from Mapbox. To collect facade imagery, we utilized the Google Street View API to gather street-level images across diverse urban settings. This approach ensures a wide coverage of different architectural styles and environments. For skyscrapers and heavily occluded facades, we adjusted the camera angles and positions to obtain high-quality usable images for facade analysis. Moreover, alternative data collection methods can be employed for skyscrapers where street-level images are insufficient. Drone-based imagery acquisition offers a flexible solution, allowing for multi-angle coverage of high-rise facades, particularly those at altitudes that are challenging to capture otherwise. In addition, handheld mobile devices can be used to collect high-resolution localized facade segments, particularly in areas where drone usage is restricted due to regulations. These alternative methods ensure that the data collected is comprehensive and suitable for detailed facade parsing and solar potential assessment across diverse building types.

This data was instrumental in contextualizing the buildings within their surrounding environments and in estimating solar irradiation patterns, factoring in inter-building shadows. Meteorological data essential for the analysis was sourced from the National Solar Radiation Database (NSRDB) [48]. Through geo-positioning techniques, we seamlessly integrated the acquired RGB images, 3D urban data, and meteorological information in a unified spatial–temporal framework for comprehensive analysis.

3.3. Deep learning building facade parsing from RGB images

This study introduces a novel deep learning framework for building facade parsing, as shown in Fig. 3, that integrates the SAM, Grounding DINO, and Stable Diffusion models. This integrated approach enables efficient segmentation, detection, and reconstruction of building facades, addressing challenges such as occlusions and the need for precise semantic understanding.

Our framework employs the SAM for robust facade element segmentation and occlusion segmentation in urban street scenes. Developed by Meta Al Research, SAM is based on the Vision Transformer (ViT) architecture and utilizes the Segment Anything 1B (SA-1B) dataset. This dataset features over one billion high-quality masks derived from 11 million diverse, privacy-respecting images. The extensive and varied nature of the SA-1B dataset enables SAM to exhibit remarkable zero-shot segmentation capabilities, adapting effortlessly across a variety of urban scenes.

SAM's training incorporates a unique "data engine" that progresses through stages of assisted manual annotation, semi-automatic annotation, and fully automatic annotation. This iterative training approach not only enhances SAM's segmentation accuracy but also its efficiency in realworld applications, allowing it to handle complex segmentation tasks such as separating building facades from occluding objects in dense urban environments. This makes SAM exceptionally suited for practical deployment where versatility and adaptability are crucial. Specifically, SAM produces binary segmentation masks M for an input image I based on point prompts p or bounding boxes B:

$$M = f_{\text{SAM}}(I, p, B) \tag{1}$$

In this equation, f_{SAM} represents the segmentation function, where the model segments the input image I into regions of interest (e.g., windows or occlusions) based on user-provided annotations such as points or bounding boxes. While SAM is effective for general segmentation, it lacks the ability to distinguish between different instances, limiting its applicability in facade element analysis, where multiple distinct objects (e.g., windows, walls) must be identified.

To overcome this limitation, our framework incorporates Grounding DINO, an innovative model that reformulates the object detection task into a phrase grounding task. This model adeptly combines extensive linguistic descriptions with visual data, facilitating the recognition of a diverse array of objects across datasets such as COCO [49], O365 [50], and OpenImage [51]. Grounding DINO is pre-trained on a comprehensive corpus that includes not only image-text pairs but also category names sampled from various conventional object detection datasets to simulate diverse textual inputs during training. Additionally, the model leverages grounding data from sources like GoldG, which contains images in Flickr30k entities [52,53] and Visual Genome [54], and RefC data and semantically-rich caption data to improve performance on novel categories.

The robust pre-training approach of Grounding DINO encompasses traditional detection data, grounding data preprocessed by MDETR [55], and pseudo-labeled caption data, providing the model with rich semantic contexts. This methodology significantly enhances the model's ability to generalize from textual descriptions to visual representations, making it exceedingly valuable for applications such as dynamic urban scene analysis, where the detection of emerging object types is crucial.

The model employs separate feature extraction backbones for both the input image I and the text prompt T, resulting in image features F_{img} and text features F_{text}:

$$F_{\text{img}} = f_{\text{img\_backbone}}(I), \quad F_{\text{text}} = f_{\text{text\_backbone}}(T) \tag{2}$$

These features are fused using a cross-attention mechanism, aligning the image and text features to enhance the model's understanding of the objects described in the text:

$$A = \text{CrossAttention}(F_{\text{img}}, F_{\text{text}}) \tag{3}$$

Grounding DINO utilizes a language-guided query selection process to generate bounding boxes for regions of interest such as occlusions and windows. The top N query indices are selected from the fused feature space to generate the final bounding boxes $B_{\text{occlusion}}$:

$$B_{\text{occlusion}} = f_{\text{decoder}}(A) \tag{4}$$

These bounding boxes are then passed to SAM to refine the segmentation of the occluded areas and window regions, resulting in more accurate and context-aware segmentation. To prevent residual occlusion information from affecting the subsequent image reconstruction process, we expand the segmentation mask to account for boundary inaccuracies:

$$M_{ex} = \text{dilate}(M) \tag{5}$$

After segmentation, we apply Stable Diffusion, which has been trained on LAION-5B and subsets thereof [56], for image inpainting. Stable Diffusion leverages a denoising process within the latent space, allowing for efficient high-resolution image synthesis while preserving fine architectural details. The model first encodes the input image into a lower-dimensional latent representation, and through a series of iterative denoising steps, it progressively refines this latent representation, guided by the semantic information from the input text prompt. The cross-attention mechanism ensures that the generated image remains consistent with the described facade elements, such as windows and

walls.

Upon completing the denoising process, the decoder maps the refined latent code back to the image space, generating a newly inpainted image where occlusions are reconstructed based on architectural context:

$$I_{\text{new}} = f_{\text{StableDiffusion}}(I) \tag{6}$$

Finally, Grounding DINO and SAM are employed once more to analyze the restored building facade image, providing a detailed and precise segmentation of facade elements such as windows, walls, and occlusions. The distinct structural components of the facade are represented as a set of elements, denoted by $E = \{e_1, e_2, \ldots, e_n\}$:

$$E = f_{\text{Grounding DINO+SAM}}(I_{\text{new}}, T_{\text{element}}) = \{e_1, e_2, \ldots, e_n\} \tag{7}$$

In this equation, I_{new} represents the newly inpainted building facade image, and T_{element} refers to the textual prompt used by Grounding DINO to guide the object detection process. The function $f_{\text{Grounding DINO+SAM}}$ denotes the combined operations of Grounding DINO and SAM, which extract and identify individual facade elements, such as windows and walls.

3.4. Simulation of shadow occlusion on building facades from 3D building footprint data

In densely built urban environments, solar radiation often cannot fully reach building facades due to mutual shading caused by adjacent structures. Accurately assessing the PV potential of facades, therefore, requires considering shadow casting by surrounding buildings. To address this, we employ the *pybdshadow* package (developed by our research team [57]) to simulate shadow conditions on building facades at different times of day, allowing for more precise solar radiation estimates and, consequently, more reliable assessments of PV potential. The influence of varying incident angles on solar radiation is dynamically modeled using the pybdshadow package, ensuring that both direct and reflected solar energy components are accurately represented based on the time-specific orientation of each facade.

Facade geometries, particularly window placements, are derived from image analysis and perspective transformations. These transformations rectify images of facades taken at oblique angles into frontal views, ensuring the correct relative positions of windows and facade elements. In real-world scenarios involving complex architectural forms, we combine automatic perspective transformation with interactive point selection, enabling users to manually define key points when necessary. For complex cases where automatic transformation is insufficient, manual point selection is used to identify the four corners of the facade, allowing a more accurate transformation to a frontal perspective.

Then, the segmented window masks are fine-tuned to correct small errors, trim irregular contours, and conform to the actual window structure. Edge points of the window masks are then projected onto the corresponding three-dimensional building footprint data, providing a more detailed and accurate geometric representation of facade elements.

Using the *pybdshadow* model, we then simulate shadow conditions for each building facade over time. To incorporate the newly added window elements, we enhance *pybdshadow*'s shadow calculation capabilities, allowing it to differentiate between various facade components. While the original model only estimated overall facade shading, our improved method performs 3D intersection analysis to estimate individual shadow coverage for different facade elements. For a set of facade elements $E = \{e_1, e_2, \ldots, e_n\}$ and shadow S at a given time, the shadowed portion of each element e_i is represented as:

$$e_{si} = e_i \cap S \tag{8}$$

where e_{si} denotes the part of the facade element e_i covered by shadow. Consequently, the set of shadow-covered areas on the facade can be expressed as:

$$E_{\text{shadow}} = \{e_{s1}, e_{s2}, \ldots, e_{sn}\} \tag{9}$$

where set E_{shadow} represents the shadow coverage across different facade elements, providing a detailed temporal and spatial assessment of facade solar exposure.

3.5. Estimation of facade BIPV output by integrating building and meteorological data

Our approach considers not only the effects of shadows cast by surrounding buildings and the structural features of the building facade on the usability of BIPV systems but also incorporates variations in the energy conversion efficiency of different BIPV models. The BIPV model parameters are simulated using the settings provided by the *pvlib* library to predict output.

For building facades with varying orientations and positions, solar radiation intensity is influenced by three main components: Direct irradiance, diffuse irradiance, and reflected irradiance. These components are represented by the following equation:

$$I_r = G_{\text{Dir}} + G_{\text{Dif}} + G_{\text{Ref}} \tag{10}$$

where the direct irradiance G_{Dir}, diffuse irradiance G_{Dif}, and reflected irradiance G_{Ref} are calculated as:

$$G_{\text{Dir}} = D_n \cdot \cos(\text{AOI}) \tag{11}$$

$$G_{\text{Dif}} = D_h \cdot \frac{1 + \cos(\text{tilt})}{2} \tag{12}$$

$$G_{\text{Ref}} = G_h \cdot \text{Albedo} \cdot \frac{1 - \cos(\text{tilt})}{2} \tag{13}$$

In these equations, the Angle of Incidence (AOI) represents the angle between incoming sunlight and the normal to the panel. The tilt refers to the orientation of the facade relative to the ground, which influences the surface's exposure to sunlight. Albedo represents the ground's reflectivity, contributing to the reflected irradiance. Direct Normal Irradiance (DNI) is the radiation received on a plane perpendicular to the sunlight. In contrast, Diffuse Horizontal Irradiance (DHI) measures scattered radiation received on a horizontal surface, and Global Horizontal Irradiance (GHI) is the total horizontal irradiance, including direct and scattered components. We utilize time-specific DNI, DHI, and GHI values to estimate irradiance on both shadowed and illuminated areas of building facades. The direct, diffuse, and reflected irradiance components are calculated considering the AOI for each facade. This ensures that the solar exposure model accurately captures how the incident angles vary over time and their effect on solar energy availability, thereby preventing potential overestimations.

To further refine the PV potential estimations, we account for the influence of temperature on BIPV efficiency. Ambient temperature and wind speed data from the NSRDB are used to estimate the temperature of the PV cells, following the Sandia Array Performance Model (SAPM). The SAPM calculates the module back surface temperature T_m and cell temperature T_C based on irradiance, ambient temperature, and wind speed, as shown in the following equations:

$$T_m = E \times \exp(a + b \times WS) + T_a \tag{14}$$

$$T_C = T_m + \frac{E}{E_0} \Delta T \tag{15}$$

Here, T_m is the back surface temperature, T_C is the cell temperature, E is the total incident irradiance, and E_0 is the reference irradiance of 1000 W/m^2. WS denotes wind speed, and a and b are model-specific parameters. Once the PV cell temperature and irradiance are estimated,

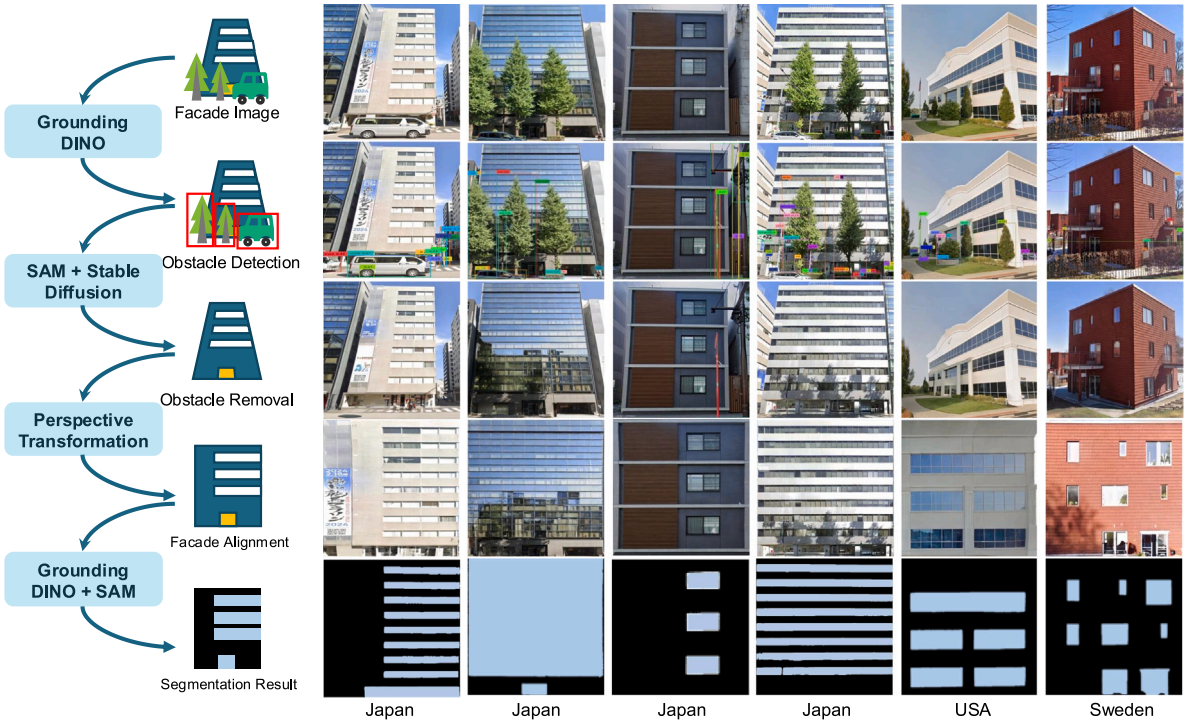

Fig. 4. Workflow and result for building facade RGB images parsing for study cases.

the potential power output is calculated using *pvlib*'s single-diode model. Since PV systems installed on windows typically have lower efficiency than those installed on building facades, a simple correction factor of 70% is applied to window PV potential estimates compared with PV of the wall in the experiment to account for this disparity.

4. Results

This section delves into a detailed analysis of building facade parsing and the estimation of building-level facade BIPV potential, taking into account influences such as shadows, windows, and seasonal variations. To substantiate the robustness and versatility of our methodologies, we extend our validation to multiple global locations. These diverse geographic validations are pivotal for confirming the generalizability of our approach across varied urban architectures and climatic conditions, ensuring that our findings are comprehensive and applicable on a global scale.

4.1. Building facade parsing

Fig. 4 shows the workflow and results of the proposed methodology for detailed facade element parsing across various study cases. The facades showcased in this figure originate from Japan, the United States, and Sweden, each exhibiting distinctive facade features, window arrangements, and varying degrees of regularity.

The proposed method demonstrates consistent performance even in scenarios with significant obstructions, including cases where critical architectural details such as wall edges are obscured. By leveraging the visible portions of the facade and incorporating contextual environmental cues, the model is capable of accurately inferring and reconstructing the hidden sections, effectively restoring the facade with a high degree of precision. In addition, the method shows remarkable versatility when applied to different types of obstructions. By modifying the input text prompts, the model, aided by the SAM model, can accurately identify and segment a wide range of obstacles in front of the facades. Once the obstacles are removed, a perspective transformation is applied to correct the images, aligning them to a frontal view of the facades. This transformation is crucial for rectifying the inevitable perspective distortion when capturing building facades from non-parallel angles, thus ensuring that the reconstructed facade elements can be precisely mapped onto 3D building models. Subsequently, the same model used for obstacle detection is employed for parsing facade elements. By adjusting the text prompts, the model accurately detects and segments windows without requiring additional training. This method effectively identifies window positions and edge geometries, producing accurate segmentations across a variety of window-to-facade ratios, architectural styles, and facade regularities.

4.2. Single facade and building level BIPV potential estimation

So far, estimates of BIPV potential on facades often neglect key factors such as shadowing effects and window-to-wall ratios, leading to inaccuracies. To address this, four scenarios were analyzed: (1) no consideration of shadows or windows ("None"), (2) accounting only for shadows ("Shadow"), (3) considering only the window-to-wall ratio ("Window"), and (4) factoring in both shadows and windows ("Both").

Figs. 5, 6, and 7 illustrate the BIPV potential assessment for building facades. Subfigure **a** depicts the 3D model of the building, **b** shows the simulated daily solar insolation along with segmented window areas, **c** presents the parsed RGB image of each facade, **d** outlines the PV output of the facades under the four scenarios, **e** estimates daily PV potential, and **f** displays the monthly potential.

The analysis reveals that incorporating shadows is crucial for accurate PV potential estimates. Buildings in both hemispheres typically have one or two facades with significant sunlight exposure, and ignoring shadow effects leads to substantial overestimations. For instance, in the Sweden case (Fig. 5), accounting for shadows alone reduces PV potential by 8% for the first facade while considering both shadows and

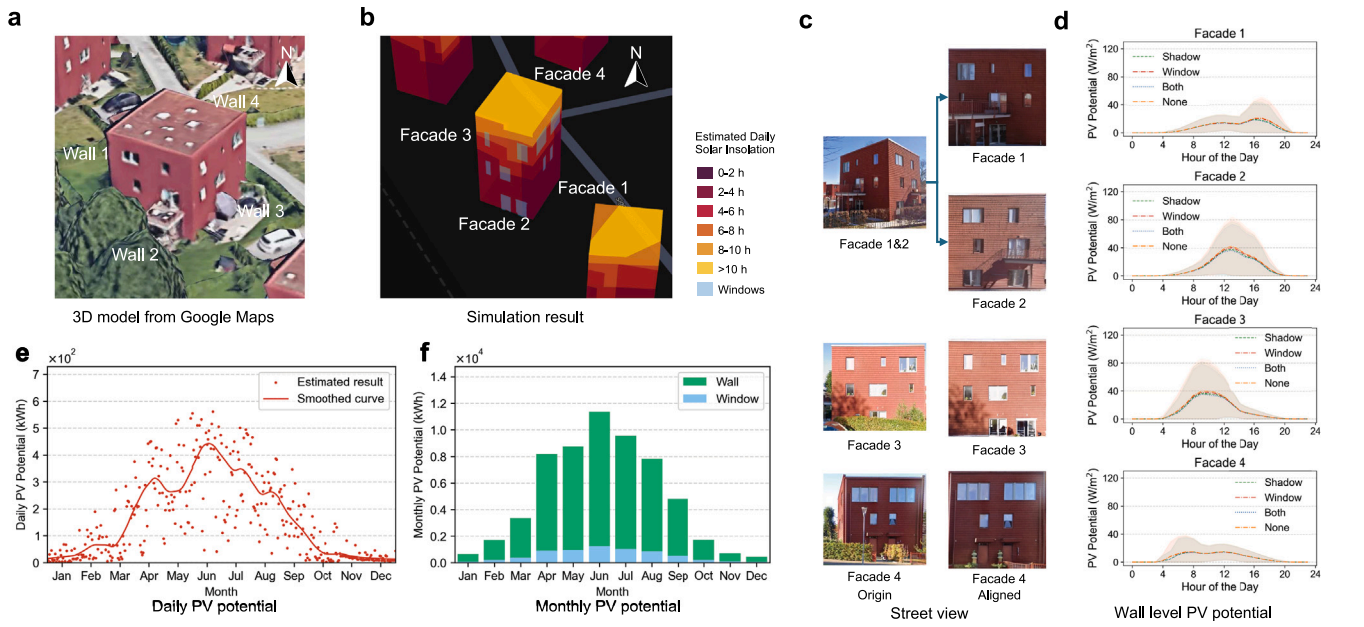

Fig. 5. Trends in annual PV potential changes for architectural cases in Stockholm, Sweden (59.1946° N, 18.0407° E).

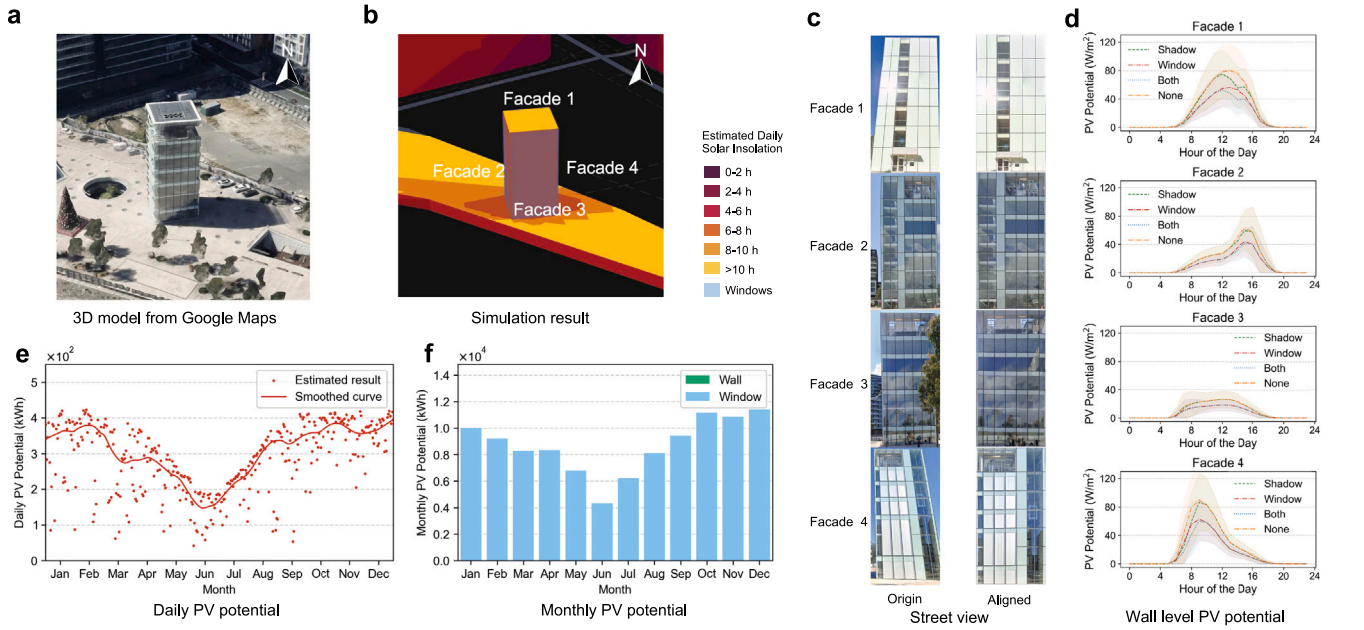

Fig. 6. Trends in annual PV potential changes for architectural cases in Sydney, Australia (33.5204° N, 151.5236° E).

windows further increases this reduction to 12%. In Sydney (Fig. 6), the influence of windows becomes more pronounced, with PV potential reduced by up to 30% due to increased window area. This highlights the importance of accounting for windows, as most buildings have substantial portions of their facades that do not receive direct sunlight. In the United States case (Fig. 7), the proposed method achieves a maximum correction of 26% compared to estimates that neglect shadowing and window effects. This demonstrates that including both shadows and window-to-wall ratios leads to significantly more accurate assessments of facade-level PV potential.

Subfigure **e** and **f** of Fig. 5, 6, and 7 show the seasonal variations in daily and monthly PV potential throughout the year for buildings located in Sweden, Australia, and the United States. Geographical location and climatic conditions result in distinct differences in PV potential across these regions. In Sweden, seasonal variations are particularly pronounced. Winter months, especially January and February, exhibit significantly low daily PV potential due to reduced daylight hours and low solar angles, with values of 636.9 kWh and 1715.1 kWh, respectively. As daylight hours increase during spring, PV potential rises, peaking in July and August at 11367.0 kWh and 9579.4 kWh, driven by higher solar angles. Following this peak, PV potential declines sharply in autumn, reaching 457.6 kWh in December. In the United States, seasonal variation is noticeable but less extreme compared to Sweden. PV potential increases steadily from January (4412.0 kWh), reaching a peak during July and August, with values of 12429.5 kWh and 11401.9 kWh, respectively. Although the potential decreases towards the year's end, the overall variation is more moderate, owing to the country's temperate climate and relatively stable solar irradiance throughout the year.

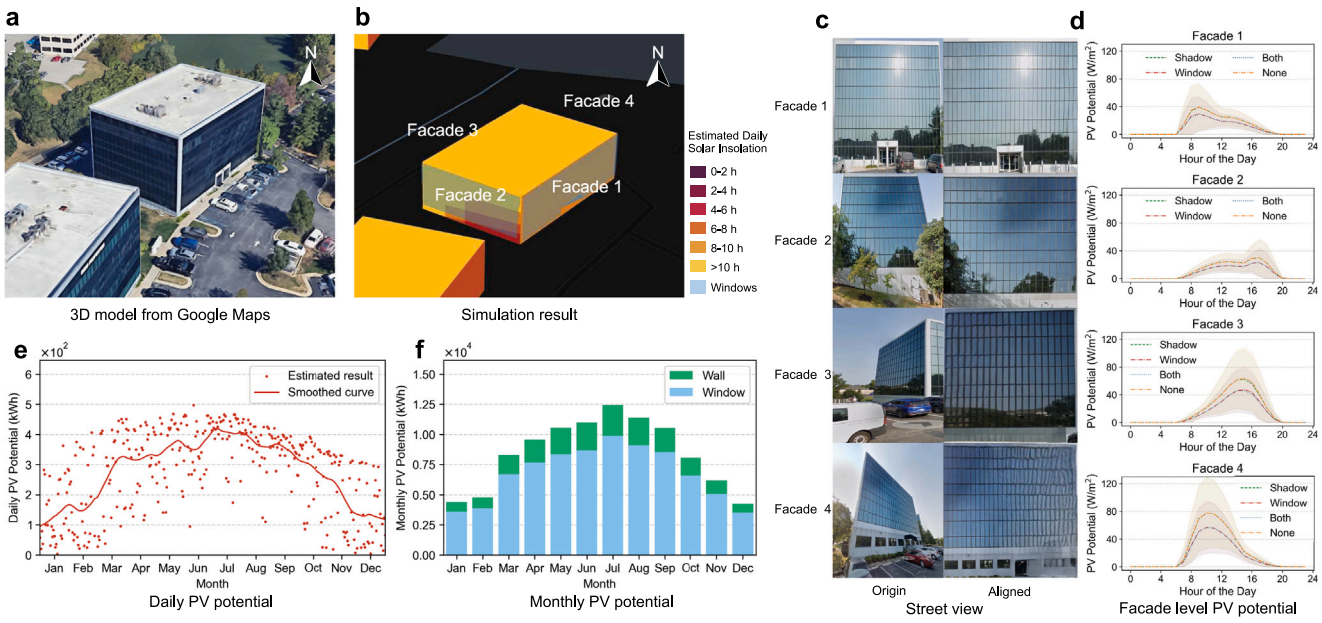

Fig. 7. Trends in annual PV potential changes for architectural cases in Lexington, The United States (38.0247° N, 84.2949° E).

Australia experiences the least seasonal variation in PV potential due to its Southern Hemisphere location and stable climate. The highest values are recorded between November and December, with a peak of 11704.7 kWh11704.7 kWh during summer. In contrast, winter months such as June and July show reduced potential, at 4615.0 kWh4615.0 kWh and 6501.8 kWh6501.8 kWh, respectively. These minor seasonal variations are due to Australia's high and consistent solar irradiance levels throughout the year.

4.3. Street level facade BIPV potential estimation

Fig. 8 illustrates the case result of a street in Tokyo, Japan, where subfigure **a** shows the 3D model of the 6 buildings on the street, **b** shows the facade image of the buildings corresponding with their BIPV potential daily, **c** shows the simulation result of solar insolation for 4 different seasons from building footprint data. Additionally, **d** and **e** show the estimated daily and monthly PV output of the whole street, with annual data further highlighting the overall energy production potential.

The results indicate that the primary contribution to PV power generation comes from the building facades. Notably, glass surfaces in buildings 1, 2, 3, 5, and 6 contribute significantly to the overall PV potential. The consistent orientation of the facades leads to a relatively uniform daily variation in PV generation potential across the street.

Shadows have a minimal effect on PV potential calculations, as most facades receive limited direct sunlight due to their orientation. Additionally, accounting for facade materials, particularly glass, provides a more accurate estimate of PV potential. This correction is particularly important in urban areas, where buildings with large windows would otherwise result in overestimated PV power generation if facade materials were not carefully considered.

5. Discussion

5.1. Impact of shading effect and facade element on BIPV potential and accuracy enhancement

The shading effect and facade element significantly influence the PV potential of building facades. Our analysis encompassed four scenarios: considering shadows, considering windows, both, and neither (none), as shown in Fig. 9.

Due to the lower efficiency of semi-transparent PV glass compared to opaque wall materials, which are more adept at capturing solar energy, data from numerous buildings consistently indicated that higher window ratios were linked to lower PV efficiencies, this trend persisted regardless of the presence of shadows. For instance, in Australia's Building, the annual average daily integral of efficiency reached approximately 401.3 kWh when only considering the shadow effect but dropped to around 280.9 kWh when windows were included, which decreased the PV potential from 31% to 36% for different facades. This finding aligns with previous studies that demonstrated facades with higher glass ratios result in lower overall PV performance. Therefore, quantifying the window-to-wall ratio is critical for realistic PV potential assessments.

Moreover, methods considering windows yield more accurate results across all scenarios. The hourly efficiency graphs for different walls highlight this accuracy. The figure shows significantly lower efficiency when windows are accounted for, even without considering shadows. This consistency underscores the importance of including windows in the analysis.

Additionally, shadow effects primarily impact sun-facing facades, reducing efficiency during peak sun hours. As shown in USA 3 building in Fig. 9, shadowing effect reduces PV potential by 38%. However, window considerations still substantially lower efficiency for non-sun-facing walls, emphasizing that windows should always be included in comprehensive assessments. This approach ensures that energy models do not overestimate the building's PV potential, as noted in similar research by [13].

The analysis underscores the importance of factoring in shading effects and windows when evaluating PV potential for both individual buildings and entire street layouts. For instance, in the street scenario (Fig. 10), efficiency notably drops when considering shading and windows, regardless of facade orientation. Examining multiple buildings reveals a substantial decrease in total efficiency with the inclusion of windows. This trend holds true across various walls and buildings, emphasizing the risk of overestimation when neglecting windows. Integrating windows into models offers a more accurate forecast of energy output, aligning with modern research methodologies.

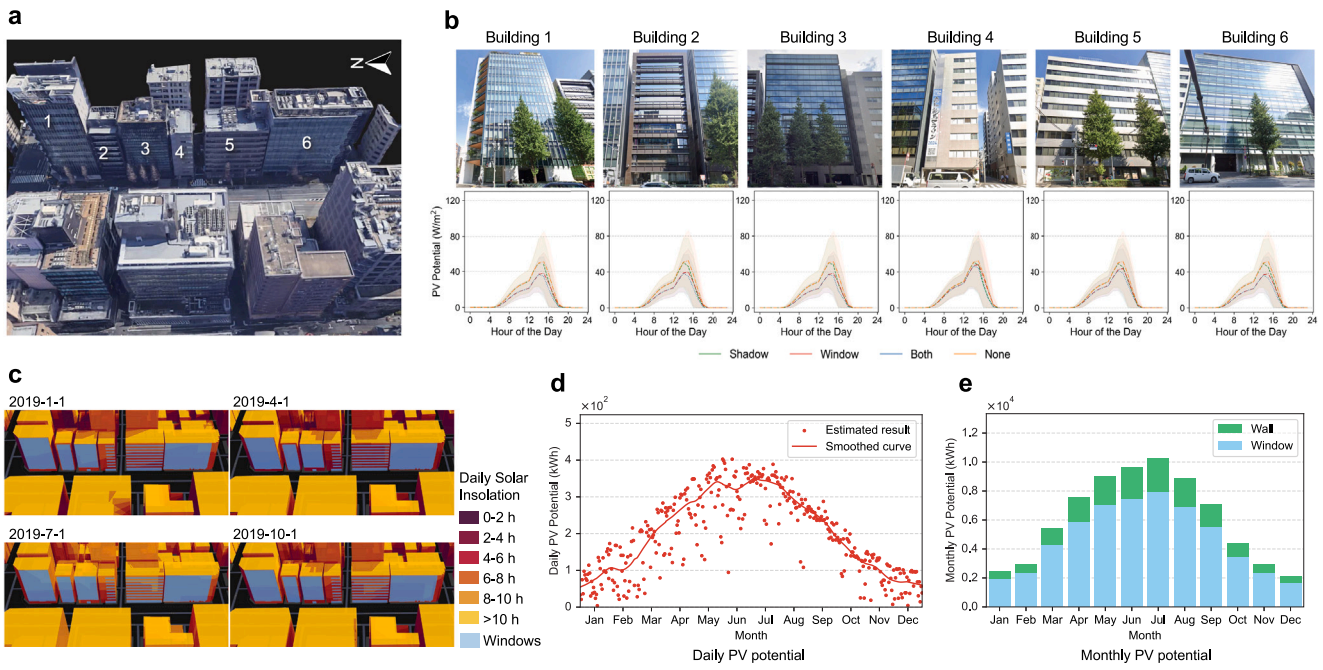

Fig. 8. Comparative analysis of the facade BIPV potential of a certain street in Tokyo, Japan (35.4123° N, 139.4132° E).

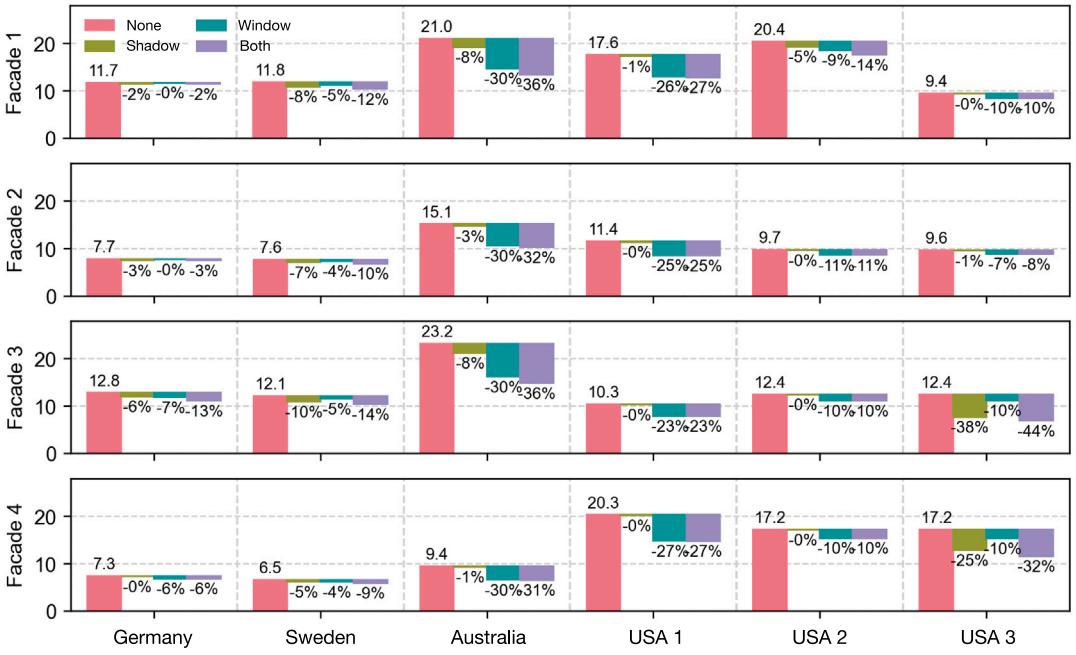

Fig. 9. The impact of different factors on the PV potential of building facades.

5.2. Temporal analysis: Daily, monthly, and seasonal

The analysis of daily variations in PV potential reveals significant disparities among building facades. One notable factor contributing to these differences is the shading effect, which can reduce the PV potential at various times of the day. Another critical factor is the orientation of the facades; more than half of a building's facades often face directions that receive insufficient direct sunlight due to structural constraints. Consequently, these facades exhibit significantly lower PV potential than those that predominantly face the sun and receive ample direct sunlight. Figs. 9 and 10 all reveal that even different walls of a building or two rows of walls on a street can show significant differences in PV potential due to variations in shadows and orientation.

Daily PV potential variations indicate that the PV potential of building facades is highly sensitive to weather conditions. Weather variations cause substantial fluctuations in both direct and diffuse solar radiation. Case studies from buildings in Sweden and Sydney (subfigure **e** in Figs. 5 and 6) demonstrate that external factors can significantly reduce PV potential on several days throughout the year, resulting in values far below the period's average. Therefore, the installation and planning of PV systems on building facades in different regions must consider the impact of these weather-induced instabilities, which are crucial for ensuring a robust and sustainable energy supply.

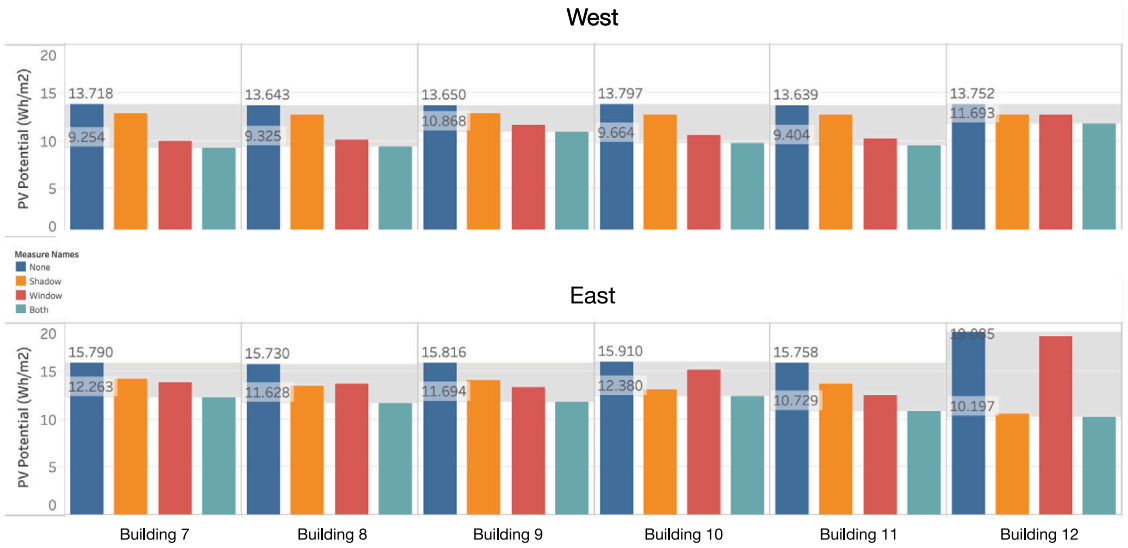

Fig. 10. The impact of different factors on the average daily PV potential of building facades on a certain street in Tokyo, Japan.

Monthly and seasonal variations also exhibit significant regional differences in the PV potential of building facades, as illustrated in subfigure f of Fig. 5, 6, and 7. For instance, buildings in the Northern Hemisphere, such as those in Sweden (Fig. 5 **f**) and the United States (Fig. 7 **f**), display a distinct trend of lower PV potential in winter (January and February) and higher potential in summer (July and August). In contrast, buildings in Sydney, located in the Southern Hemisphere (Fig. 6 **f**), show an opposite trend. Furthermore, apart from the variation in PV potential due to the distance from the equator, Sweden's monthly PV potential exhibits substantial relative changes; the extremely low PV potential during winter magnifies the proportionate increase in summer PV potential (see Fig. Fig. 5 **f**). On the other hand, Sydney's proximity to the equator ensures a relatively high PV potential even during the lowest months (June and July), resulting in more minor relative differences throughout the year, as depicted in Fig. 6 **f**.

6. Conclusion

This study presents a universal approach that integrates AIoT, GIS, and meteorological data for facade parsing and solar potential assessment. By incorporating zero-shot deep learning techniques and utilizing both 2D images and 3D building footprints, we achieve precise and efficient facade analysis. Additionally, a 3D simulation tool is developed for spatial–temporal evaluations of building facade BIPV potential.

The proposed method offers several distinct advantages. Its robust generalization capabilities enable accurate facade parsing and solar potential assessments in diverse locations, eliminating the need for additional training data. This makes it adaptable to various architectural styles and urban environments, even in complex settings with occlusions. Additionally, the method bypasses the need for costly LiDAR data by relying on affordable 3D footprint data and local weather information, enhancing the accuracy of solar assessments. Improvements in 3D shadow modeling allow for more refined PV potential simulations across different facade areas, applying tailored parameters to each region. This adaptability to different building types and shading conditions makes the method a powerful tool for advancing sustainable and energy-efficient urban infrastructure development.

One limitation of the current study is that it does not differentiate between different types of buildings when assessing the potential for PV installations. Buildings with cultural, aesthetic, or historical significance, such as museums, heritage sites, or specific residential structures, may not be appropriate for PV integration due to preservation or architectural concerns. Our methodology focuses on a generalized estimation of facade solar potential, which provides a theoretical energy yield without consideration of building-specific constraints or practical feasibility. We acknowledge that this generalization may limit the applicability of our findings to certain building types.

Scaling this approach to the city level also introduces challenges. The complexity of collecting high-quality and consistent facade imagery across diverse urban areas remains a significant hurdle. Variations in image quality, incomplete coverage, and accessibility issues could impact the overall scalability of the method. Moreover, dealing with large-area occlusions, such as trees or adjacent buildings that block facades, is an ongoing challenge. Although altering camera perspectives or using additional data sources could help mitigate these issues, such solutions are resource-intensive and may require further refinement for efficient application at a city-wide scale.

Future work will encompass experimenting with a greater variety of regions and diverse building scenarios, employing a broader array of authentic data for validation purposes. On the modeling front, upcoming studies will focus on utilizing more detailed 3D models, such as real 3D models, and identifying additional facade elements to conduct more refined analyses. Furthermore, we recognize the importance of differentiating between different building types when assessing the suitability of PV installations. Future research will extend our framework to incorporate building type considerations, evaluating the feasibility of PV integration based on cultural, historical, and aesthetic significance. This will enable us to provide a more nuanced assessment that balances solar energy potential with the appropriateness of PV deployment on architecturally significant structures, such as museums and heritage buildings. In addition, we plan to employ lightweight models utilizing edge computation to design real-time detection applications for AIoT-based facade solar estimation, enabling practical implementation of our platform in diverse urban environments. This approach will allow for efficient and context-sensitive deployment of PV systems, ensuring that solar energy strategies are tailored to both energy needs and the preservation of cultural heritage.

Code availability

The code for shadow casting estimation from building footprint data is provided by the open source python package *pybdshadow* developed by the author, which is accessible from https://github.com/ni1o1/pybdshadow. Other codes are available on request from the corresponding author.

CRediT authorship contribution statement

Kechuan Dong: Writing – review & editing, Writing – original draft, Visualization, Methodology, Conceptualization. **Qing Yu:** Writing – original draft, Visualization, Methodology, Conceptualization. **Zhiling Guo:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Conceptualization. **Jian Xu:** Writing – original draft. **Hongjun Tan:** Visualization, Data curation. **Haoran Zhang:** Supervision. **Jinyue Yan:** Supervision, Conceptualization.

Declaration of competing interest

This manuscript has not been published or presented elsewhere in part or in entirety and is not under consideration by another journal. The study design was approved by the appropriate ethics review board. We have read and understood your journal's policies, and we believe that neither the manuscript nor the study violates any of these. There are no conflicts of interest to declare.

Acknowledgments

We are grateful for the support and funding provided by the P0043885 - Flexibility of Urban Energy Systems (FUES), Japan project, P0047700 - International Centre of Urban Energy Nexus, Hong Kong and P0052733 - RISUD: Cutting-edge Solar Synergies Integrated with 3D Urban Environments towards a Carbon-Neutral City, Japan.

Data availability

The data of street view images are available from Google Maps (https://www.google.com/maps/). The building footprint data are available from Mapbox (https://www.mapbox.com/). The meteorological data including solar radiation and weather data used in this study on a global scale are collected from the NSRDB (https://doi.org/10.1016/j.rser.2018.03.003). The other datasets that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Gernaat David EHJ, de Boer Harmen Sytze, Daioglou Vassilis, Yalew Seleshi G, Müller Christoph, van Vuuren Detlef P. Climate change impacts on renewable energy supply. Nature Clim Change 2021;11(2):119–25.
[2] Jin Yubin, Zeng Zhenzhong, Chen Yuntian, Xu Rongrong, Ziegler Alan D, Chen Wenchuang, Ye Bin, Zhang Dongxiao. Geographically constrained resource potential of integrating floating photovoltaics in global existing offshore wind farms. Adv Appl Energy 2024;13:100163.
[3] UEA COP28. Global renewables and energy efficiency pledge. 2023.
[4] IEA. Renewables 2023. IEA, Paris; 2024, https://www.iea.org/reports/renewables-2023. Licence: CC BY 4.0.
[5] Shahsavari A, Akbari M. Potential of solar energy in developing countries for reducing energy-related emissions. Renew Sustain Energy Rev 2018.
[6] Li Peiran, Zhang Haoran, Guo Zhiling, Lyu Suxing, Chen Jinyu, Li Wenjing, et al. Understanding rooftop PV panel semantic segmentation of satellite and aerial images for better using machine learning. Adv Appl Energy 2021;4:100057.
[7] Clarke Joanna, Searle Justin. Active building demonstrators for a low-carbon future. Nat Energy 2021;6(12):1087–9.
[8] Mohammad Abdalrahman Khaled, Garrd Aydan, Ghosh Aritra. Do building integrated photovoltaic (BIPV) windows propose a promising solution for the transition toward zero energy buildings? A review. J Build Eng 2023;107950.
[9] Wu Hongyu, Pratt Annabelle, Munankarmi Prateek, Lunacek Monte, Balamurugan Sivasathya Pradha, Liu Xuebo, et al. Impact of model predictive control-enabled home energy management on large-scale distribution systems with photovoltaics. Adv Appl Energy 2022;6:100094.
[10] Makrides G, Zinsser B, Norton M, Georghiou G, Schubert M, Werner J. Potential of photovoltaic systems in countries with high solar irradiation. Renew Sustain Energy Rev 2010;14:754–62.
[11] Ramos A, Chatzopoulou M, Guarracino I, Freeman James, Markides C. Hybrid photovoltaic-thermal solar systems for combined heating, cooling and power provision in the urban environment. Energy Convers Manage 2017;150:838–50.
[12] Aguacil S, Duque S, Lufkin Sophie, Rey Emmanuel. Designing with building-integrated photovoltaics (BIPV): A pathway to decarbonize residential buildings. J Build Eng 2024;110486.
[13] Liu Zhaoru, Liu Xue, Zhang Haoran, Yan Da. Integrated physical approach to assessing urban-scale building photovoltaic potential at high spatiotemporal resolution. J Clean Prod 2023;388:135979.
[14] Guo Zhiling, Lu Jiayue, Chen Qi, Liu Zhengguang, Song Chenchen, Tan Hongjun, et al. Transpv: Refining photovoltaic panel detection accuracy through a vision transformer-based deep learning model. Appl Energy 2024;355:122282.
[15] Guo Zhiling, Zhuang Zhan, Tan Hongjun, Liu Zhengguang, Li Peiran, Lin Zhengyuan, et al. Accurate and generalizable photovoltaic panel segmentation using deep learning for imbalanced datasets. Renew Energy 2023;219:119471.
[16] Kim Sungkyum, Jang Seungsoo, Choi Chanho, Shin Sung Gyun, Lee Minjae, Cho Woosung, et al. Local intensity estimation of solar radiation with vision-based 3D modeling of local areas through unmanned aerial vehicle. Environ Prog Sustain Energy 2022;41(1):e13726.
[17] Prieto Iñaki, Izkara Jose Luís, Usobiaga Elena. The application of LiDAR data for the solar potential analysis based on urban 3D model. Remote Sens 2019.
[18] Brito M, Redweik P, Catita C, Freitas S, Santos M. 3D solar potential in the urban environment: A case study in lisbon. Energies 2019.
[19] Li Qingyu, Krapf Sebastian, Mou Lichao, Shi Yilei, Zhu Xiao Xiang. Deep learning-based framework for city-scale rooftop solar potential estimation by considering roof superstructures. Appl Energy 2024;374:123839.
[20] Ni Haozhan, Wang Daoyang, Zhao Wenzhuo, Jiang Wolin, Mingze E, Huang Chenyu, et al. Enhancing rooftop solar energy potential evaluation in high-density cities: A deep learning and GIS based approach. Energy Build 2024;309:113743.
[21] Liu Jiang, Wu Qifeng, Lin Zhipeng, Shi Huijie, Wen Shaoyang, Wu Qiaoyu, et al. A novel approach for assessing rooftop-and-facade solar photovoltaic potential in rural areas using three-dimensional (3D) building models constructed with GIS. Energy 2023;282:128920.
[22] Chen Zhe, Yang Bisheng, Zhu Rui, Dong Zhen. City-scale solar PV potential estimation on 3D buildings using multi-source RS data: A case study in wuhan, China. Appl Energy 2024;359:122720.
[23] Hong Taehoon, Lee Minhyun, Koo Choongwan, Jeong Kwangbok, Kim Jimin. Development of a method for estimating the rooftop solar photovoltaic (PV) potential by analyzing the available rooftop area using hillshade analysis. Appl Energy 2017;194:320–32.
[24] Redweik P, Catita C, Brito M. Solar energy potential on roofs and facades in an urban landscape. Sol Energy 2013;97:332–41.
[25] Trzmiel G, Głuchy D, Kurz D. The impact of shading on the exploitation of photovoltaic installations. Renew Energy 2020;153:480–98.
[26] Dong Bowei, Shi Qiongfeng, Yang Yanqin, Wen Feng, Zhang Zixuan, Lee Chengkuo. Technology evolution from self-powered sensors to aIoT enabled smart homes. Nano Energy 2021;79:105414.
[27] Zhu Rui, Kwan Mei-Po, Perera Amarasinghage Tharindu Dasun, Fan Hongchao, Yang Bisheng, Chen Biyu, et al. GIScience can facilitate the development of solar cities for energy transition. Adv Appl Energy 2023;10:100129.
[28] Liu Shilong, Zeng Zhaoyang, Ren Tianhe, Li Feng, Zhang Hao, Yang Jie, et al. Grounding dino: Marrying dino with grounded pre-training for open-set object detection. 2023, arXiv preprint arXiv:2303.05499.
[29] Kirillov Alexander, Mintun Eric, Ravi Nikhila, Mao Hanzi, Rolland Chloe, Gustafson Laura, et al. Segment anything. In: Proceedings of the IEEE/CVF international conference on computer vision. 2023, p. 4015–26.
[30] Rombach Robin, Blattmann Andreas, Lorenz Dominik, Esser Patrick, Ommer Björn. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022, p. 10684–95.
[31] Tan Hongjun, Guo Zhiling, Zhang Haoran, Chen Qi, Lin Zhenjia, Chen Yuntian, et al. Enhancing PV panel segmentation in remote sensing images with constraint refinement modules. Appl Energy 2023;350:121757.
[32] Singh Rhythm, Banerjee R. Estimation of rooftop solar photovoltaic potential of a city. Sol Energy 2015;115:589–602.
[33] Desthieux G, Carneiro C, Camponovo R, Ineichen P, Morello E, Boulmier Anthony, et al. Solar energy potential assessment on rooftops and facades in large built environments based on LiDAR data, image processing, and cloud computing. Methodological background, application, and validation in Geneva (Solar Cadaster). Front Built Environ 2018;4:1–22.
[34] Hu Mengjin, Liu Zhao, Huang Yaohuan, Wei Mengju, Yuan Bo. Estimation of rooftop solar photovoltaic potential based on high-resolution images and digital surface models. Buildings 2023.
[35] Assouline D, Mohajeri Nahid, Scartezzini J. Large-scale rooftop solar photovoltaic technical potential estimation using random forests. Appl Energy 2018;217:189–211.
[36] Walch A, Castello R, Mohajeri Nahid, Scartezzini J. Big data mining for the estimation of hourly rooftop photovoltaic potential and its uncertainty. Appl Energy 2020;262:114404.
[37] Zhong Teng, Zhang Zhixin, Chen Min, Zhang Kai, Zhou Zixuan, Zhu Rui, et al. A city-scale estimation of rooftop solar photovoltaic potential based on deep learning. Appl Energy 2021;298:117132.

[38] Tan Hongjun, Guo Zhiling, Lin Zhengyuan, Chen Yuntian, Huang Dou, Yuan Wei, et al. General generative AI-based image augmentation method for robust rooftop PV segmentation. Appl Energy 2024;368:123554.
[39] Cheng Liang, Li Shuyi, Xu Hao, Liu Xiaoqiang, Mao Junya, Ruan Xiaoguang, et al. RETRACTED: Calculating potential of solar energy and CO2 emissions reduction for cityscale buildings based on 3D remote sensing technologies. Remote Sens Environ 2018;209:612–32.
[40] Li Qingyu, Krapf Sebastian, Shi Yilei, Zhu Xiao Xiang. SolarNet: A convolutional neural network-based framework for rooftop solar potential estimation from aerial imagery. Int J Appl Earth Obs Geoinf 2023;116:103098.
[41] Liu Jiang, Wu Qifeng, Lin Zhipeng, Shi Huijie, Wen Shaoyang, Wu Qiaoyu, et al. A novel approach for assessing rooftop-and-facade solar photovoltaic potential in rural areas using three-dimensional (3D) building models constructed with GIS. Energy 2023;282:128920.
[42] Liu Hantang, Xu Yinghao, Zhang Jialiang, Zhu Jianke, Li Yang, Hoi CH Steven. DeepFacade: A deep learning approach to facade parsing with symmetric loss. IEEE Trans Multimed 2020.
[43] Liu Hantang, Li Wentong, Zhu Jianke. Translational symmetry-aware facade parsing for 3-D building reconstruction. IEEE MultiMedia 2021;29:38–47.
[44] Ma Wenguang, Ma Wei, Xu Shibiao. Deep facade parsing with occlusions. KSII Trans Internet Inf Syst 2022;16:524–43.
[45] Wang Bowen, Zhang Jiaxing, Zhang Ran, Li Yunqin, Li Liangzhi, Nakashima Yuta. Improving facade parsing with vision transformers and line integration. 2023, http://dx.doi.org/10.48550/arXiv.2309.15523, ArXiv arXiv:2309.15523.
[46] Chen Dong, Xiang Guiqiu, Peethambaran Jiju, Zhang Liqiang, Li Jing, Hu Fan. AFGL-Net: Attentive fusion of global and local deep features for building façades parsing. Remote Sens 2021;13:5039.
[47] Ren Tianhe, Liu Shilong, Zeng Ailing, Lin Jing, Li Kunchang, Cao He, et al. Grounded sam: Assembling open-world models for diverse visual tasks. 2024, arXiv preprint arXiv:2401.14159.
[48] Sengupta Manajit, Xie Yu, Lopez Anthony, Habte Aron, Maclaurin Galen, Shelby James. The national solar radiation data base (NSRDB). Renew Sustain Energy Rev 2018;89:51–60.
[49] Lin Tsung-Yi, Maire Michael, Belongie Serge, Hays James, Perona Pietro, Ramanan Deva, et al. Microsoft coco: Common objects in context. In: Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6-12, 2014, proceedings, part v 13. Springer; 2014, p. 740–55.
[50] Shao Shuai, Li Zeming, Zhang Tianyuan, Peng Chao, Yu Gang, Zhang Xiangyu, et al. Objects365: A large-scale, high-quality dataset for object detection. In: Proceedings of the IEEE/CVF international conference on computer vision. 2019, p. 8430–9.
[51] Krasin Ivan, Duerig Tom, Alldrin Neil, Ferrari Vittorio, Abu-El-Haija Sami, Kuznetsova Alina, et al. Openimages: A public dataset for large-scale multi-label and multi-class image classification. 2, (3):2017, p. 18, Dataset Available from https://Github.Com/Openimages.
[52] Plummer Bryan A, Wang Liwei, Cervantes Chris M, Caicedo Juan C, Hockenmaier Julia, Lazebnik Svetlana. Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models. In: Proceedings of the IEEE international conference on computer vision. 2015, p. 2641–9.
[53] Plummer Bryan A, Wang Liwei, Cervantes Chris M, Caicedo Juan C, Hockenmaier Julia, Lazebnik Svetlana. Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models. In: Proceedings of the IEEE international conference on computer vision. 2015, p. 2641–9.
[54] Krishna Ranjay, Zhu Yuke, Groth Oliver, Johnson Justin, Hata Kenji, Kravitz Joshua, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. Int J Comput Vis 2017;123:32–73.
[55] Kamath Aishwarya, Singh Mannat, LeCun Yann, Synnaeve Gabriel, Misra Ishan, Carion Nicolas. Mdetr-modulated detection for end-to-end multi-modal understanding. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021, p. 1780–90.
[56] Schuhmann Christoph, Beaumont Romain, Vencu Richard, Gordon Cade, Wightman Ross, Cherti Mehdi, et al. Laion-5b: An open large-scale dataset for training next generation image-text models. Adv Neural Inf Process Syst 2022;35:25278–94.
[57] Yu Qing, Li Ge. pybdshadow - a python package to analyze building shadow.