

## Article

# Workflow for Window Composition Detection to Aid Energy-Efficient Renovation in Low-Income Housing in Korea

Jong-Won Lee

Department of Building Energy Research, Korea Institute of Civil Engineering and Building Technology,  
283 Goyang-daero, Daehwa-dong, Ilsanseo-gu, Goyang-si 10223, Republic of Korea; jongwonlee@kict.re.kr;  
Tel.: +82-31-910-0723

**Abstract:** Enhancing the efficiency of windows is important for improving the energy efficiency of buildings. The Korean government has performed numerous building renovation projects to reduce greenhouse gas emissions and mitigate energy poverty. To reduce the costs and manpower requirements of conventional field surveys, this study presents a deep-learning model to examine the insulation performance of windows using photographs taken in low-income housing. A smartphone application using crowdsourcing was developed for data collection. The insulation performance of windows was determined based on U-value, derived considering the frame-material type, number of panes, and area of windows. An image-labeling tool was designed to identify and annotate window components within photographs. Furthermore, software utilizing open-source computer vision was developed to estimate the window area. After training on a dataset with ResNet and EfficientNet, an accuracy of approximately 80% was achieved. Thus, this study introduces a novel workflow to evaluate the insulation performance of windows, which can support the energy-efficient renovation of low-income housing.

**Keywords:** image processing; window composition detection; energy poverty; low-income housing; energy-efficient renovation



**Citation:** Lee, J.-W. Workflow for Window Composition Detection to Aid Energy-Efficient Renovation in Low-Income Housing in Korea. *Buildings* **2024**, *14*, 966. <https://doi.org/10.3390/buildings14040966>

Academic Editor: Chaogun Zhuang

Received: 3 March 2024

Revised: 13 March 2024

Accepted: 29 March 2024

Published: 1 April 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Energy Poverty and Energy-Efficient Renovation Projects in South Korea

Energy poverty is a major concern, exacerbated by the recent surge in global energy costs. It refers to the state in which a group of households spend more than 10% of their income on energy costs for cooling, heating, and cooking [1]. Typically, three issues are considered the underlying causes of energy poverty, namely low income, high energy costs, and energy inefficiencies in housing projects [2]. Energy poverty in South Korea is mostly observed among the socially disadvantaged, such as the elderly population living alone, families headed by children, and the disabled, who constitute approximately 1.6 million households across the country. Energy-poor households typically live in old houses (aged more than 30 years) and spend a significant portion of their income on energy [3–6]. In South Korea, energy costs are a large economic burden for the energy-poor households because a large proportion of them live in old houses with low energy efficiency [7]. Low energy-efficiency dwellings are more vulnerable to heat and cold waves. Such houses also tend to have low accessibility to low-cost energy sources, such as city gas and district heating. This makes it difficult to maintain a proper indoor temperature even with sufficient use of energy. To address this energy poverty, the Korean government has performed numerous cash and commodity disbursements and building renovation projects. Particularly, the building renovation project is the most important one in the building sector for carbon neutrality, and the budget for subsidies has gradually increased as part of a key policy to tackle energy poverty. Previous studies have indicated that enhancing energy efficiency in the housing industry could reduce greenhouse gas emissions, mitigate climate

change, alleviate energy poverty, and boost regional economic growth [8–10]. Although emergency support has mostly been provided using cash and commodity programs, such as energy vouchers and electric fans, a long-term improvement in the residential environment and energy efficiency of housing is fundamentally crucial to energy poverty alleviation.

South Korea unveiled a government-supported policy referred to as the energy-efficiency improvement project for low-income households, which supports both heating and cooling improvement projects. The heating improvement project is aimed at improving the energy efficiency of walls, windows, and boilers via upgrading the insulation, windows, and floor construction, as well as boiler replacement. The cooling improvement project provides free installation of air conditioners. There are many ways to improve the overall energy performance of low-income housing. In South Korea, window replacement is often considered the fastest approach and the best way to address the issue of moving costs and alternative housing for occupants, which are often necessary when adopting other approaches. As a result, in South Korea, window replacement is the most common method used to improve energy performance through renovations of low-income buildings. The energy-efficiency improvement project for low-income households in South Korea supported approximately 609,000 households and 3035 social welfare facilities from 2007 to 2022, with a budget of approximately USD 600 million. In 2023, USD 1800 was provided, on average, per household for a total of 34,000 households. Thus, long-term building renovation for energy-poor households has been gaining prominence.

Nevertheless, the question of whether the priorities of building renovation are effectively determined for budget distribution according to the target population and purpose of use [11–14] persists. While problems caused by procedural and systemic issues exist, the most significant issue is that energy poverty in blind spots is not identified through field surveys; furthermore, the same buildings may be repeatedly identified for funding. Currently, experts visit an extremely small part of the 1.6 million households in the country and conduct field surveys. Evaluating the energy status of old buildings occupied by individuals facing energy poverty in a limited timeframe with minimal manpower is not feasible. Complex field surveys to examine the insulation performance of existing buildings in the field entail considerable cost and manpower.

Owing to the limited budget availability for building renovation, the performance of old houses in the existing project for low-income households is also diagnosed based on a simple evaluation of building designs, drawings, and the year of construction [15]. This renders the diagnosis of the performance of old housing difficult and may decrease the reliability of the renovation effort [7]. Therefore, it is important to simply measure the performance of old houses in the field and renovate those that require urgent repairs rather than perform indiscriminate renovations to achieve high efficiency on a low budget [7].

The importance of online evaluation technology has been emphasized based on the data obtained from field surveys. Numerous studies have been conducted to recognize the characteristics or components of buildings using photographs alone. Particularly, studies to recognize windows from the building envelope through photographs are frequently conducted. Accurate information regarding the windows of a building is essential for energy-efficiency ratings. As windows act as thermal bridges, they are the major source of thermal loss. Window operation has significant effects on ventilation rates as well as heating and cooling loads in building energy usage [16–19]. Therefore, analyzing the energy efficiency of a building mainly relies on its windows. In general, window replacement can improve the energy efficiency of old buildings rapidly and easily. Consequently, window replacement is a service that accounts for the largest portion of the budget allocated to energy-efficiency improvement projects for low-income households.

## *1.2. Literature Review concerning Window Detection Using Image Processing*

Detailed information on windows is included in construction plans and similar documents; however, high levels of expertise and considerable human efforts are required to extract the relevant information [20]. Accurate and reliable data pertaining to the windows

of a building may be acquired from two primary sources—from building blueprints or by personally examining the facade of the building. The initial step in obtaining construction drawings for many buildings is submitting a formal request to the land registry offices. However, in certain instances, these plans may no longer be up-to-date, particularly in the case of a limited number of buildings.

To overcome these issues, recent studies have manually and easily collected data by obtaining window images from web services (e.g., Google Street View), capturing images with cameras fixed on the ground, or using drones around the target building [21–23]. However, due to technical or privacy issues, window images constructed based on a satellite map do not have a sufficiently high resolution to extract detailed information regarding the windows. In contrast, photographs taken using drones may have a high resolution. Owing to cost and time constraints and the survey distance range that may be limited only to the target building, methods for collecting window images directly from buildings are still being commonly used [20].

Numerous previous studies have concentrated on the analysis of the relationship between the opening of windows and occupant behavior. To identify the opening of windows, sensors were directly installed in buildings, and the status of windows was recognized through photographs. In contrast to commonly used ambient sensors (e.g., air temperature), window-opening sensors should be installed in large numbers and are hence considered a hindrance to everyday life [24,25]. Furthermore, the data collected from these sensors require preprocessing in many cases because raw data can rarely be used owing to the errors that may occur during the measurement stage (e.g., data transfer failures and accidents) [26,27].

This study begins with the following question: If the insulation performance of windows can be approximately estimated using numerous collected window photographs alone before implementing the energy-efficiency improvement project for low-income households, would it not be possible to significantly reduce the cost and manpower required for actual window replacement owing to the availability of window information before the site visits?

Most of the previous studies focused on recognizing windows or their opening status from photographs [21]. In other words, it is important to obtain information on window components through field surveys for energy diagnosis and the energy-efficiency improvement project for low-income households; however, there have been no studies on the development of algorithms for recognizing the window components.

Therefore, in contrast to the previous studies, this study collected and analyzed data from the window photographs of houses inhabited by energy-poor households through a field survey. As this was conducted for numerous houses in South Korea as part of a preliminary survey, drone or camera-installation methods that require high resolutions and significant amounts of information on houses were not utilized. From this perspective, this study differs from previous ones that collected window photographs through facades and performed learning. The window photographs collected through field surveys are not taken with fixed machines, and the angle, height, and direction may vary depending on the photographer. Therefore, a practical guide for taking window photographs was also proposed. Additionally, a tool was developed to detect and label the window components from the photographs collected through the field survey. Software based on an algorithm for calculating the window area was developed so that the window area information can be directly obtained from the photographs. The results obtained expand the theoretical knowledge and could affect the determination of priorities before the field surveys as well as actual projects and policies, such as fairer and more efficient distribution methods [12].

This research introduces a novel approach to window insulation performance analysis based on crowdsourced window photographs, a method not previously explored in depth in the field of energy-efficient building renovation. Our study uniquely focuses on developing algorithms for recognizing window components from varied and unstructured photographic data, filling a major gap in current architectural and energy-efficiency re-

search. A practical guide for photographing windows is proposed, and a novel tool for window component detection and area calculation is developed. This approach significantly diverges from traditional methods, offering a more efficient and scalable solution for projects intended to alleviate energy poverty. The methodology and tools developed in this study represent a significant advancement in terms of cost efficiency and practicality, particularly for preliminary surveys in large-scale energy-efficiency improvement projects.

## 2. Materials and Methods

This research adopts a groundbreaking methodology that integrates deep learning with crowdsourced photographic data. Our approach uniquely combines the scalability of crowdsourced data with the precision of advanced computational analysis to evaluate window insulation performance. To address the potential limitations and biases of our data collection and analysis methods, we rigorously tested the model's performance under varying photographic conditions, ensuring a comprehensive evaluation of its reliability and applicability. This study comprised four distinct stages: (1) window data collection, (2) data labeling, (3) development of a method for calculating the window area from window photographs, and (4) model training and testing; subsequently, a window composition detection workflow was proposed.

### 2.1. Data Collection

#### 2.1.1. Crowdsourced Data Collection

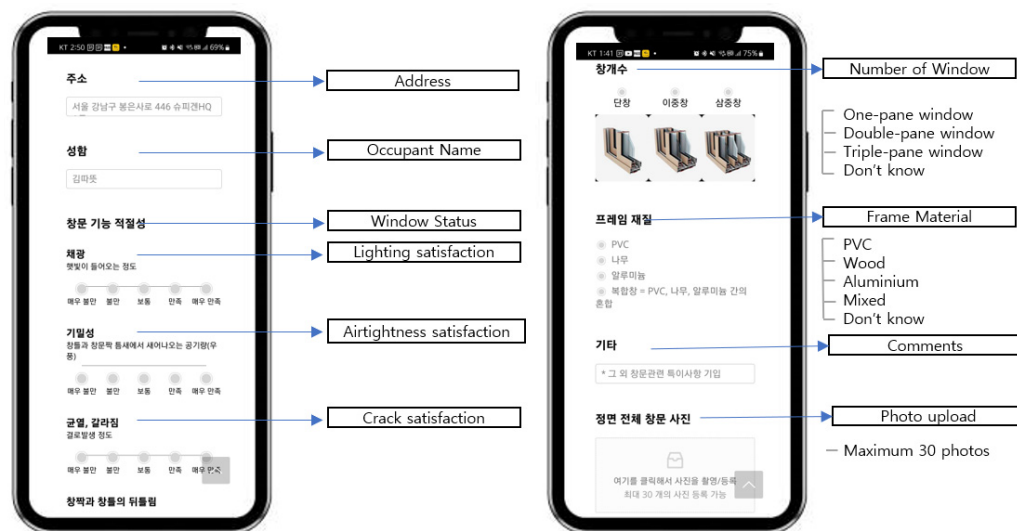
The first stage involved window data collection. The most difficult aspect of collecting photographic data of the windows installed in the energy-poor households was accessibility. Generally, it is difficult to know the address location of energy-poor households; the building or occupant information is strictly controlled by a government department for privacy protection and security reasons. Information on energy poverty can only be received through a government agency, which requires a significantly complex process. In general, creating large-scale datasets is expensive, difficult, and likely to give rise to errors [28]. There have been significant advances in the method of continuously collecting building data in terms of both quantity and contents [29], spurred by the diversity of stakeholders and their types. Today, the building data collected by volunteers as well as the government, companies, and academia are utilized in research [30–35].

This study considered the crowdsourcing method, which leverages the help of general citizens for data collection. A method was devised so that citizens benefiting from energy or housing welfare could directly participate in data collection. Considering the efficacy and reliability of crowdsourced data, we implemented stringent quality control measures, including a validation process to assess and enhance data consistency and accuracy. As an example of our method, when social workers or volunteers visited energy-poor households to deliver lunchboxes or briquettes, they were asked to conduct a survey on the living environment and obtain photographic data of aged windows. They were educated on window data collection methods and data labeling. Effective data of sufficiently high quality could be obtained using the time and labor of citizens who have easy access to energy-poor households. A smartphone application was developed and provided to the volunteers and social workers. This application allowed for the easy conduct of surveys and the immediate capture and transmission of window photographs (Figure 1).

As shown in Figure 1, the window components can be directly observed, and related details as well as data on current status, such as window deterioration and inconvenience in the use of windows, can be transmitted to a server. Although all the components can be visually identified through photographs, only window photographs and information on window components were collected with valid answers that can be used for future learning. The application was developed in such a way that anyone can easily take and transmit photographs anytime and anywhere through the internet. To address the issue of possible real-time errors, a website was also created to enable the upload of window photographs stored on a smartphone. Moreover, we established collaborative partnerships



with non-governmental organizations (NGOs) and community welfare entities to determine the locations of buildings inhabited by individuals experiencing energy poverty across various regions. Efforts were devoted to improving the quantity and quality of data on a broader level.



**Figure 1.** Smartphone application for window status survey and photograph collection.

As emphasized by Yan et al., ethical considerations, such as privacy, could arise while collecting occupant-related photographs [36]. To address such concerns, volunteers who conducted the field survey were asked to avoid taking photographs of the interiors of the residential spaces. This could be achieved by choosing an appropriate camera position while taking the photographs. However, the ethical debate over the privacy of occupants is important, and the cases among the collected data that contained the facial details or personal information of the occupants have to be carefully examined. In this regard, the authors underwent a familiarization exercise on various phrases related to personal information; the written consents of the occupants were collected regarding the use of their personal information for this research. Overall, it was presumed that the benefits of this study would outweigh the potential risks associated with privacy.

The ideal distance of a camera would be different in each study [37]. The distance determines the amount of detail for each window and the number of windows per image. As it is difficult to consider all the various environmental factors of the field survey, the volunteers were educated on the good and bad cases of window photographs. A guide was provided to them so that they could take high-quality photographs (Figure 2).

Photographs taken outside a house could be affected by numerous environmental factors, such as illumination, rain, snow, and solar glare, as described in [37]. In this study, differentiated window photographs were collected by taking window photographs from the inside through a field survey. The photographs taken from the inside of a house were less affected by the external environment; however, their quality was significantly affected by internal lighting and obstacles. Such factors adversely affect the quality of images and the accuracy of detecting window components. To address this issue, it is important to establish and develop guidelines on window photograph collection and on window photograph data refining and labeling.



**Figure 2.** Ideal and undesirable cases of window photographs for image detection. (a) Ideal case: A photo taken from the front of the window with minor obstructions. (b) Ideal case: Photo with the window as visible as possible, even if there are curtains or other obstructions. (c) Undesirable case: Dark photo with backlighting. (d) Undesirable case: Out-of-focus photo.

### 2.1.2. Window U-Value Information

The surveyed items mainly comprised an occupant survey, a window photograph satisfaction survey, window information, and photographs. In this study, the survey of information on photographs related to the detection of window components was the primary area of focus. Through questions on the state of windows, factors that affect the insulation performance of the windows were investigated. The insulation performance of a window was determined based on a parameter termed the U-value ( $W/m^2 K$ ) (also referred to as the “U-factor”) of the window, which can be obtained through an experiment if the window frame-material type (e.g., wood, PVC, and aluminum), air-layer thickness of the glass (e.g., 6, 12, and 16 mm), air type, number of panes (single, double, triple, and quadruple), width, height, and area are available. In addition, window design is

affected by other important parameters like solar factor, solar transmittance, and light transmittance. Nevertheless, the methodology employed in our study, leveraging crowd-sourced photographic data, encounters inherent limitations in accurately capturing the nuanced details requisite for determining these parameters directly from images. Only variables that can be determined through image processing methods were considered as important influencers in this study. Thus, in the field, the U-value of a window cannot be directly evaluated. Furthermore, the U-value of a window in a laboratory will not be identical to that of the window after its long-term use. In addition, extracting all the items required for the U-value entails expensive verification products and expertise. The U-values of windows can be affected by the types of external or internal shading devices. In addition, glazed surfaces, including the glazing and frame, can affect the thermal insulation of windows, and complex relationships exist among these factors. However, after consulting with other experts, we decided to limit our U-value analysis to a simple influence factor, serving as a determinant of a window's thermal performance. Therefore, this study aimed to easily and rapidly identify factors that affect the U-value of a window that are related to the insulation performance of the window through window photographs. The U-value was determined by the aforementioned factors; however, object recognition technology can extract only the information visible from photographs. This study developed a process that could select only those items that could be recognized through window photographs, perform learning using a deep-learning method, and extract the necessary information. It is acknowledged that perfectly identifying the insulation status of windows in the field or creating related formulae using attribute information that can be obtained from photographs alone is not possible. However, an attempt was made to provide help in partially determining the insulation status of the windows in the field through such additional information. Table 1 lists the information to be extracted from the window photographs for assessing the U-value.

**Table 1.** Window composition information that can be recognized from photographs.

Attribute Information	Category
Number of panes	Single/double/triple/quadruple/unknown
Frame material	PVC/wood/aluminum/composite/unknown
Opening method	Fixed/hinged/projecting/horizontally sliding/unknown
Area calculation	Glass/frame/unknown
Glass	The gas layer thickness and gas type between glass panes cannot be recognized from photographs.

The survey was conducted, and window photographs were collected in five cities (Seoul, Daejeon, Gwangju, Mokpo, and Busan). Approximately 800 households were surveyed, and 5000 photographs were collected for three years, from 2020 to 2022. However, most of the photographs included the poor condition of houses in addition to the windows. Following a thorough selection process, approximately 3000 photographs were selected. Of these, the number of finally selected photographs of adequately high quality to detect the window components was 1866.

## 2.2. Data Labeling

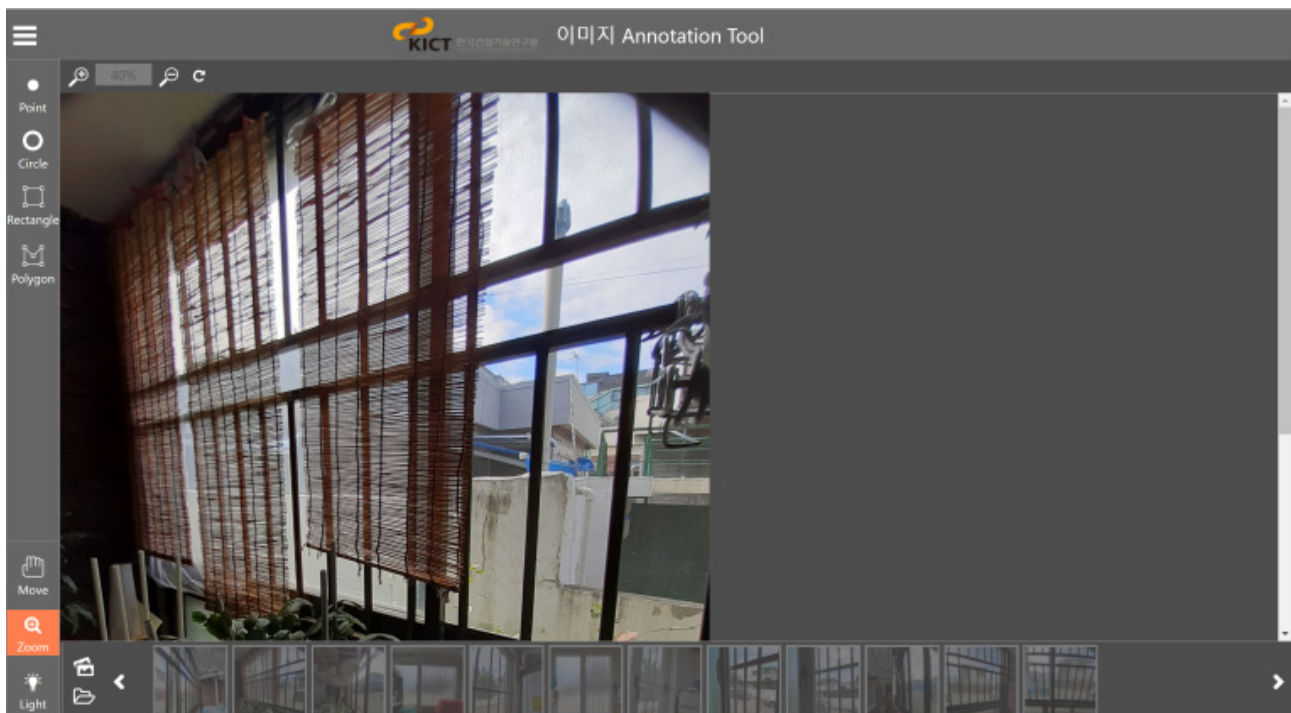
### 2.2.1. Development of Image Labeling Tool

In the second stage, annotations were manually inserted into the collected photographs for their utilization in deep learning. Previous studies have suggested that manual data processing is required to identify windows from photographs [37,38]. This problem can be resolved by designating the positions of windows and window components on the photographs. This process is called data labeling.

Given the large size of the photograph dataset, manual data labeling can be extremely laborious and time-consuming, and often requires domain knowledge and labeling skills to ensure quality. Although numerous crowdsourcing services (e.g., Amazon SageMaker

data labeling [39] and Google Cloud data labeling [40]) are available to significantly reduce the manual labor for the dataset owners, these services can be expensive for large datasets. Additionally, the manual process can be error prone. For instance, [41] reported that the ImageNet validation set contained labeling errors, including missing and incorrect labels, to the extent of 6%.

To address these problems, we created a customized, simple data labeling tool from open-source code that could label the window photographs. This was to avoid the escalating costs of commercial cloud services (Figure 3).



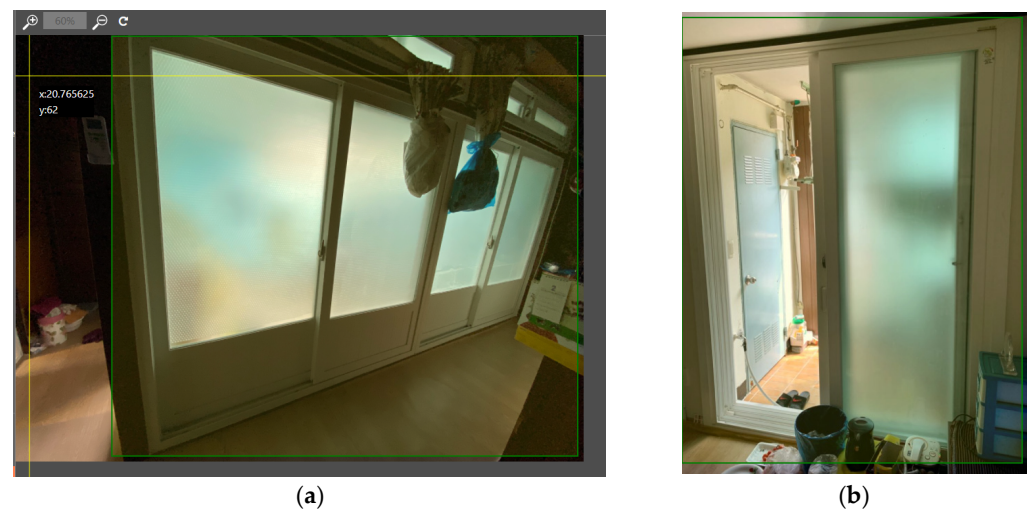
**Figure 3.** Data labeling tool for window composition detection.

The tool supports drawing points, circles, rectangles, and polygons. Appropriate shapes can be used according to the presented guidelines. In this study, only rectangles (bounding boxes) were used. Additionally, the tool has functionality for moving the labeling objects, increasing/decreasing the size of the displayed image, and changing its brightness. Briefly, the steps involved in the use of the tool are as follows. A window image to be labeled was imported through the import menu, and its size was adjusted through the increase/decrease function. Following the selection of the bounding box of the labeling tool, a box was drawn over the part that corresponded to a window on the screen. The label name, window type, frame material, and opening method corresponding to the window were entered inside the box. Finally, the label was saved as an XML file (Figure 4).

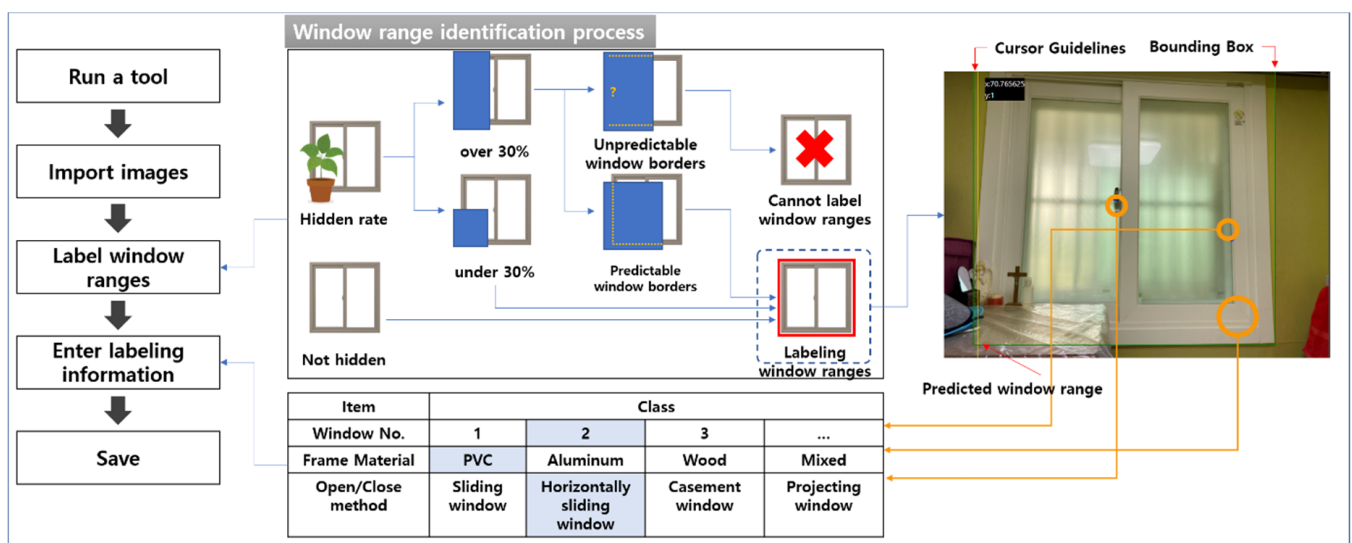
Next, we attempted to address one crucial shortcoming of the manual data labeling, which was the person-to-person variation introduced in the process. Therefore, a guideline was developed through an expert meeting, ensuring that the same procedure and criteria would be applied irrespective of the person applying the labels. The window photograph-labeling method and criteria to be adopted are as shown in Figure 5. To enhance the model's accuracy and ensure its applicability in real scenarios, our study meticulously defined the window glazing and framing quality. This was accomplished through a standardized data labeling process, as demonstrated in Figure 5, underpinned by a set of rigorous guidelines established by domain experts. The guidelines stipulate the identification and annotation of window components with high precision, accounting for material characteristics, glazing, and framing. Such granularity in data labeling minimizes intra-observer variability and



enhances the model's ability to predict the thermal performance of windows accurately, facilitating its deployment in the field for energy-efficient retrofitting projects.



**Figure 4.** Method of window image labeling. (a) Label the bounding box to encompass the window, extending up to its frame. (b) In instances where the window is positioned diagonally, ensure the bounding box encapsulates the entirety of the window.



**Figure 5.** Data labeling manual for window composition detection.

However, as the photographs inside the housing varied from case to case, it was difficult to apply the same labeling method to all the window photographs based on the guidelines in the manual. Therefore, regular discussions on labeling for each case were conducted with the volunteers. Several example cases are shown in Figure 6.

### 2.2.2. Window Area Calculation Algorithm

The window area is an important factor for determining the U-value of a window and evaluating its insulation performance. Direct measurements were generally performed using a tape measure in field surveys to identify the window area. However, cumbersome direct measurements lack reliability and accuracy. In practice, the window glass and frame sizes collected by citizens contained erroneous values and were unusable.





**Figure 6.** Various cases explaining window image labeling. (a) Despite the occlusion caused by curtains, the window's presence can be inferred from the visible frame on the left, warranting its labeling. (b) When recognizable attributes like rails and glass are obscured by laundry and appliances, labeling is confined to discernible frames. (c) Instances with obstructions are incorporated into the labeling process, provided the boundaries of such obstructions can be reasonably deduced. (d) For double-pane windows, labeling is performed based on the observable lock and frame of the posterior window.

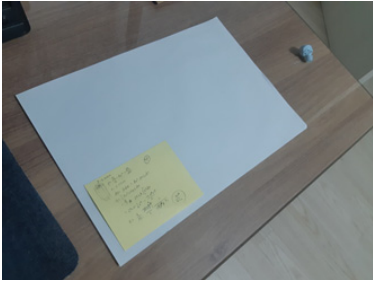
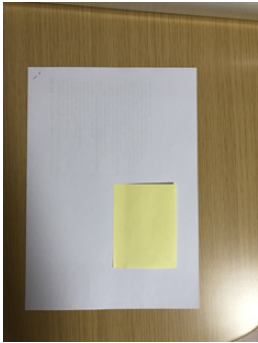
Therefore, in this study, the citizens who photographed the windows were instructed to do so with and without a Post-it™ note. Since the area of the Post-it note used is known, it would serve as a reference in the window photographs. The area of the Post-it note, as it appears in the window photographs, would help estimate the window and frame sizes. The algorithm and software created for this purpose are explained below.

The basic algorithm for calculating the image area using window photographs used perspective transformation provided by open-source computer vision (OpenCV), a real-time computer vision signal processing library. This transformation is used to convert the four points (red cross-shaped dots) designated by the user on a given window image into a rectangular photograph. To perform warping within the input photograph using the WarpPerspective() function in the OpenCV library, four points that designate the area on the input photograph to be transformed and four corresponding points on the output photograph are required. A transformation matrix for four corresponding points can be obtained using the getPerspectiveTransform() function, and the transformation is performed

using the WarpPerspective() function. In perspective transformation, all the straight lines on the original input photograph remain straight lines on the output photograph. The perspective transformation uses a  $3 \times 3$  matrix to calculate the orthogonal coordinates. If  $M_p$  is the matrix that expresses perspective transformation, the pixel coordinates of the input image  $(x, y)$  are moved by the matrix  $M_p$ , and the image pixel coordinates  $(x', y')$  are calculated as follows:

$$\begin{pmatrix} wx' \\ wy' \\ w \end{pmatrix} = M_p \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \quad (1)$$

The calculated transformation matrix values for two examples of input photographs are as shown in Figure 7.

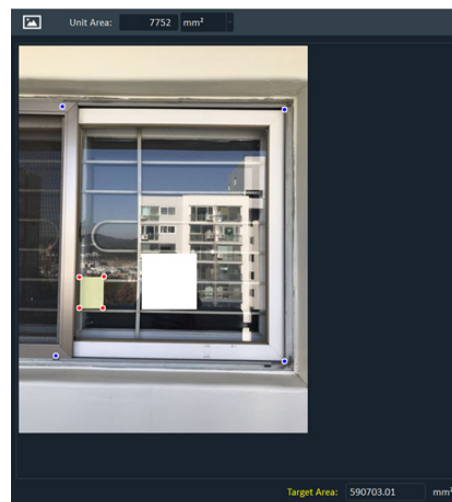
Example	Calculated transformation matrix
	<pre>matrix1 [[ 8.06921473e-01 -4.97887717e-01  5.46474691e+01]  [ 5.55383531e-01  9.96254993e-01 -4.91268224e+02]  [-5.05489801e-04  8.98220358e-04  1.00000000e+00]]  matrix2 [[ 6.53997919e-01 -3.93462163e-01  5.94659571e+01]  [ 4.12538252e-01  1.04965354e+00 -2.23356509e+02]  [-4.62925627e-04  7.33388252e-04  1.00000000e+00]]</pre>
	<pre>matrix1 [[ 1.01713424e+00  8.20269546e-03 -3.39058417e+02]  [-1.64945353e-02  1.02266119e+00 -5.45787678e+02]  [-2.32000907e-05  4.38646816e-05  1.00000000e+00]]  matrix2 [[ 1.02890352e+00  1.14005930e-02 -7.93566776e+01]  [-6.21616305e-03  1.03809923e+00 -1.98848840e+02]  [-1.67177815e-05  4.45602824e-05  1.00000000e+00]]</pre>

**Figure 7.** Calculated transformation matrix values in the window area calculation algorithm.

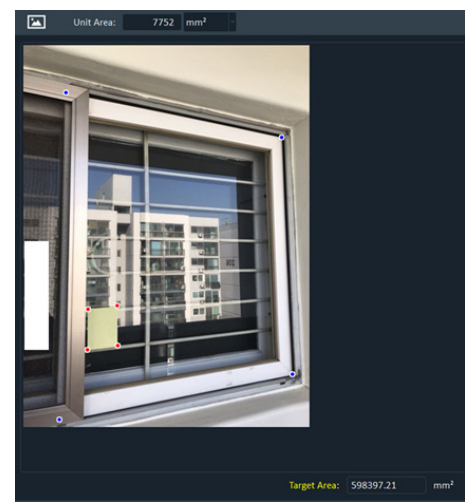
The Python programming language was used for the software implementation. The libraries required to operate the implemented software were numpy, cv2, and PyQt5. The areas, measured using a Post-it note and the developed software, were validated as follows. The reflection of the photographer on the glass was covered with a white quadrangle. The Post-it note (Model 657) used as the reference had a length and breadth of 102 and 76 mm, respectively, with a total area of 7752 mm<sup>2</sup>. As illustrated in Figure 8, four points were designated to the Post-it note by left-clicking the mouse. This would mark red circles on the photograph according to the designated coordinates. Additionally, four points were designated on the window photograph to be measured by right-clicking the mouse. This would mark blue circles on the photograph according to the designated coordinates. Once the unit area of the Post-it note was entered, the target area, that is, the window area, would be automatically calculated after designating all eight points.

As observed in the two example photographs, accuracy may differ depending on the camera angle. The errors for the photographs in Figures 8a and 8b were 1.21% and 2.53%, respectively. Additionally, it was observed that the discrepancy in the computed area remained relatively constant even when the photograph was captured from a non-perpendicular perspective, as depicted in Figure 8a.

Images taken at  
different camera  
angles



(a)



(b)

Window area  
Measured area  
Error rate

width: 725 mm; height: 805 mm; unit area: 583,625 mm <sup>2</sup>
590,703.01 mm <sup>2</sup> 598,397.21 mm <sup>2</sup>
1.21%                                      2.53%

**Figure 8.** Validation of window area calculation software. (a) A photo of the window from the front. (b) A photo of the window from a slightly sideways angle. To measure the window area, the windows are marked with blue circles and the Post-it note used as the reference is marked with red circles.

### 2.3. Data Preprocessing and Model Training

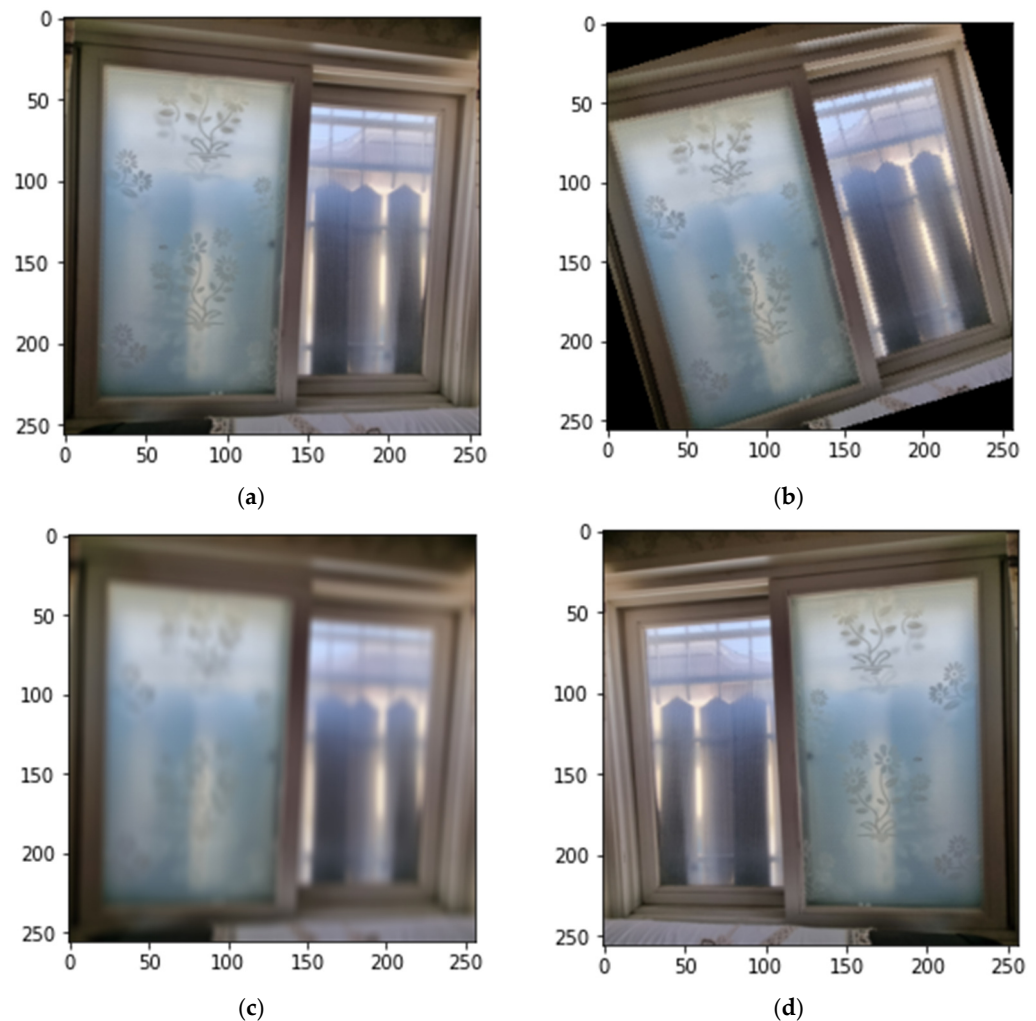
The next step involves the process of learning the window components with the photograph dataset that was subjected to labeling. The number of learning data elements that were labeled was 1866; among them, 1745 photographs were of the horizontally sliding windows, based on the window-opening method. This predominance could introduce a bias in learning. Consequently, the window-opening method was not analyzed in this study. Additionally, unclear data pertaining to windows or frame material were removed, which reduced the total number of learning data elements to 1738. The following data elements were confirmed: double-pane windows, 838; single-pane windows, 896; triple-pane windows, 4; PVC, 722; wood, 430; and aluminum, 586. For these data, the images were resized to  $256 \times 256$  to fit the pre-trained model learning input size. As the number of data elements for triple-pane windows was substantially small for learning, they were excluded during model classification.

As for the Baseline model, two commonly used models, which are known to have high performance efficiency as image classification models, were utilized [42]. One is the ResNet model, which introduced the residual learning method for stable learning in a deep-learning model. ResNet series (e.g., ResNet50 and ResNet101) have been proposed to alleviate the vanishing gradient problem; in fact, they significantly reduce the difficulty of training in deeper layers of neural networks [43]. The second is EfficientNet, which significantly improves the efficiency of the structure compared to the existing CNN models [44]. In this study, this model was implemented with Pytorch (2019) [45]. The training was performed using cross entropy loss for the loss and the Adam optimizer. A learning rate of 0.001 was used.

Among the structures of the Baseline model, ResNet50 and EfficientNet-B0 with a small number of parameters were used, considering the size of the model, and the pre-trained model was used. The fully connected layer of the pre-trained model was replaced with two types of fc layers that fit the number of classes corresponding to the number of windows and frame-material type. Learning was performed by dividing the dataset into training, validation, and testing sets in the ratio of 80:10:10. The batch size was set to 64. Data augmentation was applied during learning. HorizontalFlip, rotation, and



GaussianBlur were used, as shown in Figure 9. The ratio of data-augmented images was used to the original one as one to three. This approach was chosen after preliminary experiments to balance the diversity of data without overwhelming the original dataset characteristics. Regarding the impact of data augmentation on model accuracy, we found that incorporating augmented images improved the model's robustness, increasing the overall accuracy by approximately 5%. This enhancement is attributed to the model's improved generalization capability from the varied training examples.



**Figure 9.** Data augmentation. (a) Original. (b) Rotation. (c) GaussianBlur. (d) HorizontalFlip.

### 3. Results

In this study, we utilized both ResNet and EfficientNet architectures, applying 100 epochs learning process to assess the accuracy of window number and frame material detection. The accuracy evaluation of these models was limited to the testing data sets and did not include the training data sets. The results, as illustrated by the learning curve of EfficientNet-B0 (Figure 10), reveal a noteworthy trend in the learning process, with the loss significantly decreasing and stabilizing at low values while maintaining an accuracy consistently exceeding 73%. This stability demonstrates the effectiveness of our deep-learning approach. With a precision of 82%, a recall of 78%, and an F1-score of 80%, our findings confirm the efficacy and reliability of our deep-learning approach for window component detection. The insights gained from the confusion matrix further guided the efforts to enhance model accuracy, particularly in distinguishing between various window components. The learning process was rigorously tested across five different data splits to ensure robustness; the aggregated accuracies are listed in Table 2. These findings not only

demonstrate the efficacy of our proposed method but also mark a significant advancement in the application of deep learning to architectural research.

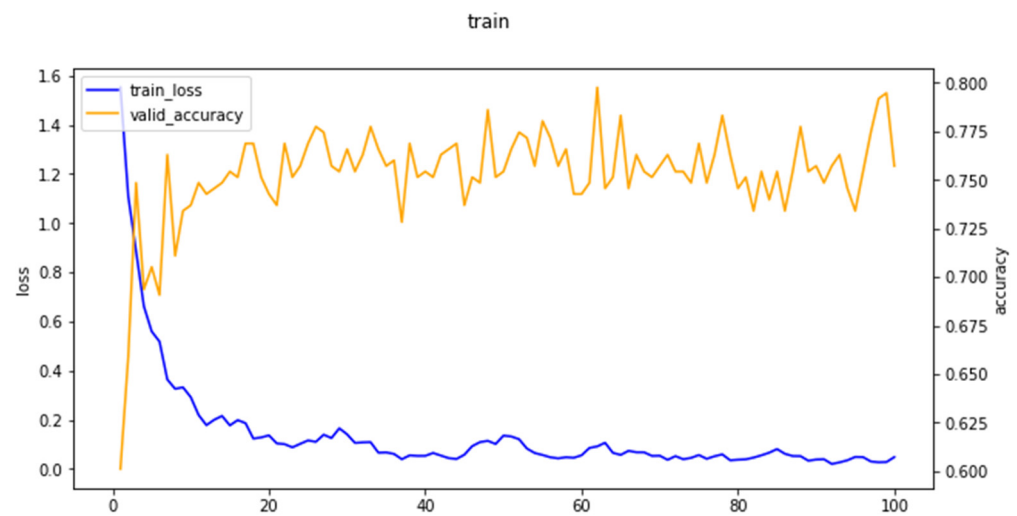


Figure 10. EfficientNet-B0 learning results.

Table 2. Model training accuracy results.

	Accuracy for the Number of Windows	Accuracy for the Frame Material
Resnet50	73.99%	78.73%
EfficientNet-B0	76.16%	79.88%

#### 4. Discussion

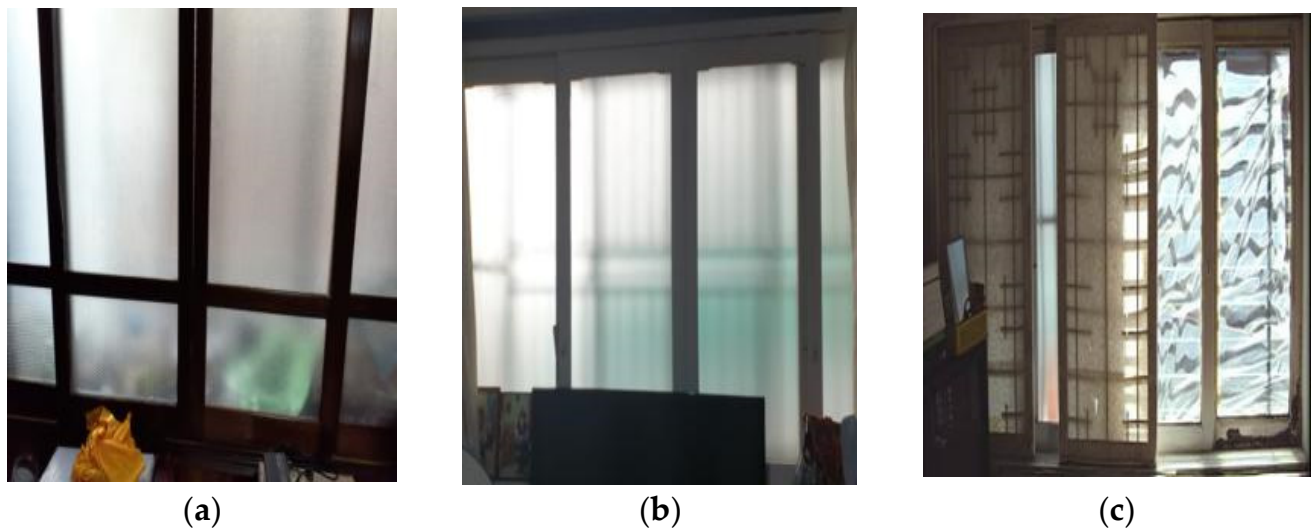
The accuracy of the Resnet50 and EfficientNet-B0 models for the number of windows and frame material was found to be approximately 80%. This accuracy may not be high enough to be used in the field; however, it is judged that window components were identified with high accuracy from window photographs with various obstacles. Additionally, the model training results showed that it would be difficult to identify environmental conditions, such as lighting, obstacles, and reflection, and achieve an appropriate general-purpose accuracy. Herein, we discuss four additional developments aimed at improving the performance of the algorithms.

First, it is necessary to analyze which photographs were less or more amenable to learning window components from. Testing the accuracy by training the models based on various types of window photographs provided insight into the limitations of photographs that may decrease the accuracy of or hinder learning. These limitations were related to specific structures in individual window photographs, obstacles, image noise, the inappropriateness of the labeling used to filter the definition conditions, and environmental conditions related to lighting and reflection. As the survey was conducted to obtain correct answers during the collection of window photographs, those for which the values expected by the trained models were different from the correct answers were analyzed.

For the photographs in Figures 11a and 11b, a strong backlight made it difficult to identify the surface texture. Possibly, only the smooth surface rendered by the backlight was detected, and the material was predicted to be PVC when, in fact, it was wood. Particularly, in Figure 11b, the frame material was wood, but was painted. Therefore, it might be difficult to determine the material. Because it is also difficult to identify the number of window rails on the photograph, only one window was visible, and thus the triple-pane window was detected as a single-pane window. In Figure 11c, Korean traditional and modern windows were both present, which caused a problem in selecting the window to be classified by the trained models. This caused certain windows to be identified as single-pane windows. In the case of the material, the white paint on the wood was detected as PVC. Therefore, when



there were different types of windows in one photograph, they could decrease the accuracy of the model, indicating that a separate manual classification method is required during data collection or preprocessing.



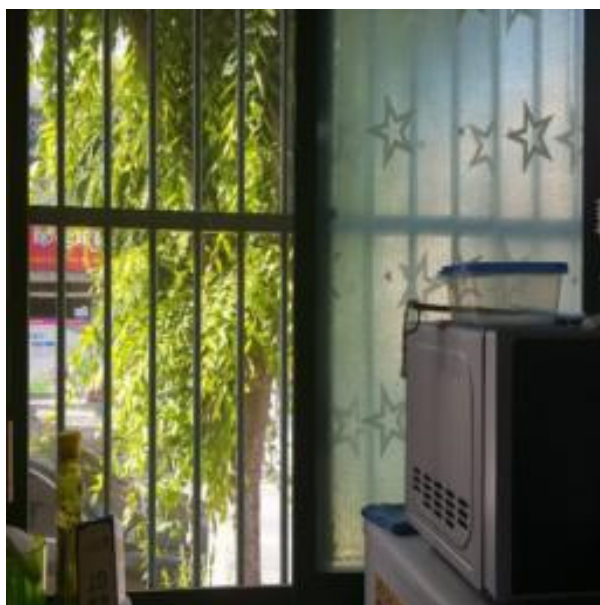
**Figure 11.** (a) Case where wood was detected as PVC. (b) Case where wood was detected as PVC and triple-pane window as single-pane window. (c) Case where wood was detected as PVC and triple-pane window as single-pane window.

The material of the windows shown in Figures 12a and 12b was aluminum; however, the metallic texture of aluminum could not be detected owing to the low illumination in the photograph. Consequently, the material was detected as PVC, which represented the largest proportion of the training data. Particularly in Figure 12b, the number of windows could not be determined owing to the opaqueness of the glass. The photographs in Figures 12c and 12d show a combination of various elements, including the windows. In Figure 12c, the windows and walls of external buildings became indistinguishable from the windows in the photograph. This indicates that external as well as internal obstacles significantly affected the learning accuracy of the model, which must be considered during the data collection or refining. Eventually, obstacles present in the window photographs were highly likely to lead to an incorrect identification of the number of windows or frame material because they hindered the process of distinctly identifying the windows.

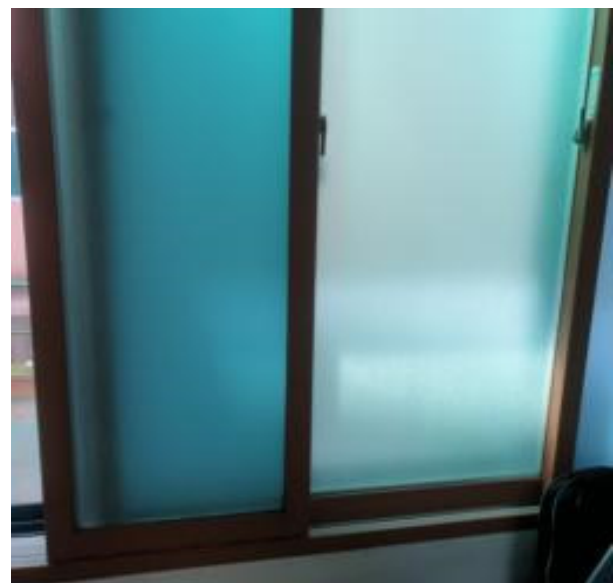
The photograph in Figure 13a shows stains on the PVC window frame. Such stains could be mistaken for wood texture. Similarly, the photograph in Figure 13b shows low illumination and backlight, which could cause the PVC window frame to be mistaken for a wooden frame. Two windows, one single-pane and another double-pane, were recognized as triple-pane windows because of the difficulty in identifying the number of window rails, and learning the number of windows was degraded by the opaque stickers attached to the window (Figure 13a). In Figure 13b, the front window is a double-pane window, which was determined to be a triple-pane window because of another window on the outside. This photograph also shows that an obstacle outside a window affects the learning of the window.

The four photographs in Figures 14a and 14d show that the window frames were particularly dark. They could have been easily detected as PVC frames if they had been visually inspected. They were wrongly detected as aluminum owing to the black or brown stickers on the PVC frame or the dark-colored paint. Furthermore, PVC was mistaken for aluminum owing to the illumination or window-composition problem. In the case of the photograph in Figure 14a, the texture of the material might have been detected as that of aluminum owing to the old and unmaintained condition of the PVC frame and window guard outside the window. From the photograph in Figure 14b, it was difficult to detect

the texture or shape of PVC owing to low illumination. The window guard could also affect the identification of the material and number of windows. From the photograph in Figure 14c, the material was not accurately identified because only a part of the window was included. It was also difficult to determine the number of windows for the same reason. For the photograph in Figure 14d, detection was hindered by an obstacle that resembled a carpet. Low illumination could also have resulted in the frame being detected as aluminum. For the photograph in Figure 14e, the accuracy was reduced by various elements outside the window. The number of windows could not be identified owing to the insect-screen frame or safety railing between the window frames. For the photograph in Figure 14f, the window frame was detected as aluminum owing to the influence of the window guard that occupies a large proportion in the photograph. Additionally, the biased composition or confusing image components, such as the landscape outside a window and the carpet attached to the window, appear to have rendered an accurate classification difficult.



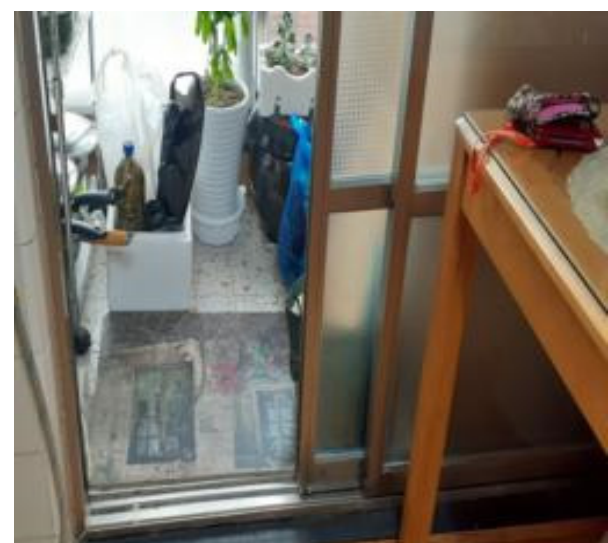
(a)



(b)



(c)

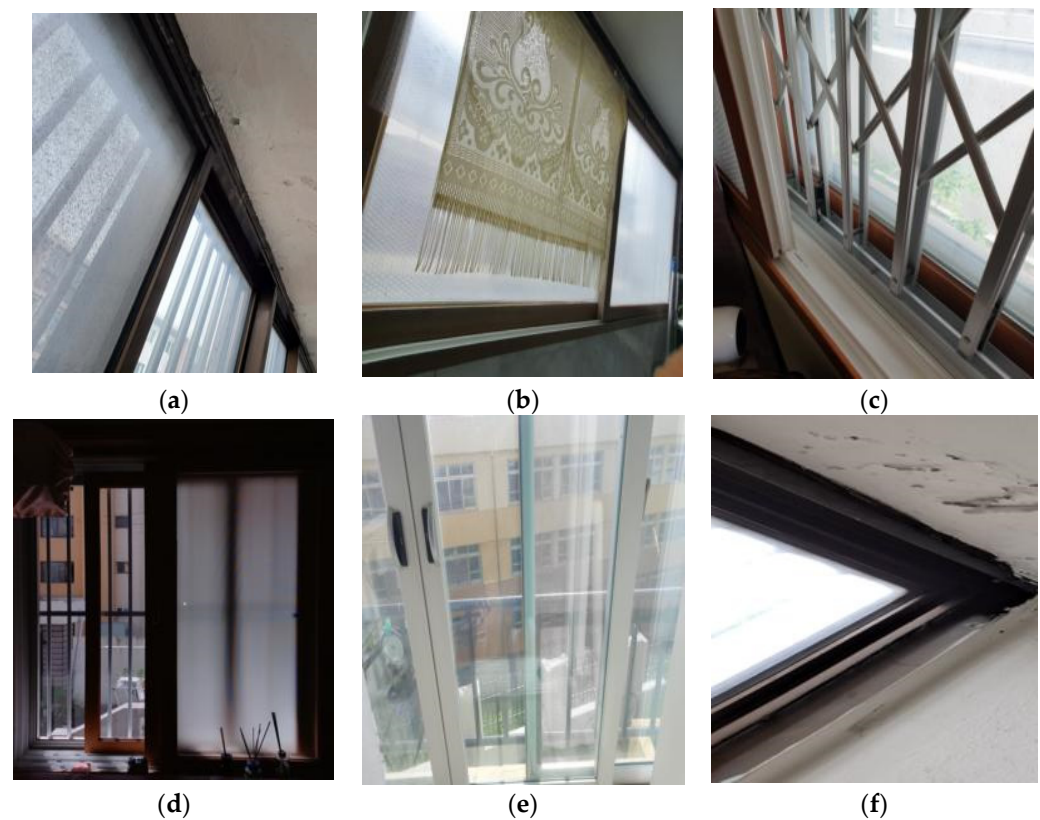


(d)

**Figure 12.** The upper row (a,b) shows the cases where wood was detected as aluminum. The lower row (c,d) shows the cases where a single-pane window as a double-pane window.



**Figure 13.** (a) Case where PVC was detected as wood and a single pane or double pane window as a triple-pane window. (b) Case where aluminum was detected as wood and a single pane or double pane window as a triple-pane window.

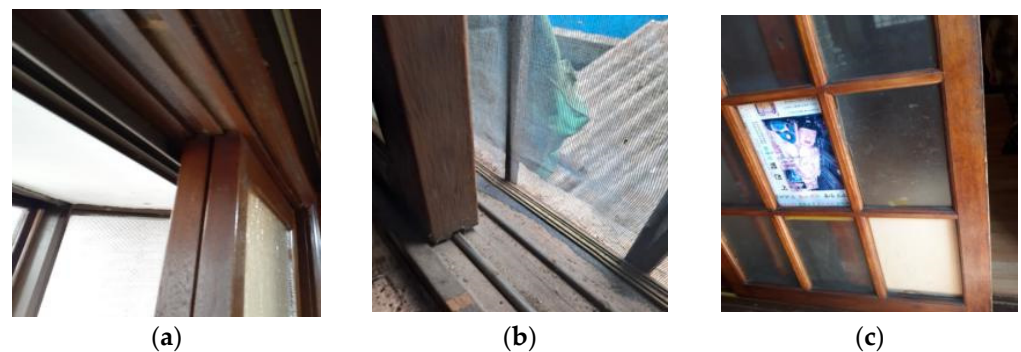


**Figure 14.** The upper row (a–c) shows the cases where PVC was detected as aluminum. The lower row (d–f) shows the cases where PVC was detected as aluminum and a single-pane window was detected as a double-pane window.

Figure 15 shows that the composition of the photographs was biased. As a result, the probability of wrongly identifying the material increased. It appears that the material was classified as aluminum because most of the photographs of windows with aluminum frames had biased structures. This implies that it is important to determine the same



photograph frame without biased structures among the classes when photographs are collected for each class.



**Figure 15.** (a) Case where a triple-pane window was detected as a double-pane window. (b) Case where wood was detected as aluminum. (c) Case where wood was detected as aluminum, and a triple-pane window was detected as a double-pane window.

In Figure 15a, the frame at the bottom left was possibly wrongly recognized as the window frame to be analyzed. It appears that the learning rate was decreased by the presence of several window members in one window photograph. In Figure 15b, elements other than a window, including those outside the window, occupied a large area, and the insect screen was made of aluminum. This case shows that the accurate detection rate might decrease when various window material elements are present in one photograph. A double-pane window was determined by the number of window rails. This indicates that using photographs that do not show the accurate number of rails for training would decrease the correct answer rate. In Figure 15c, the paper attached to the window was recognized as an obstacle, and the material could not be accurately determined between aluminum and wood. As the window rail could not be identified, the window appeared to have been detected as a double-pane window, based on the window frame on the back side of the opaque glass.

Based on the analysis, it was found that errors occurred mostly in images with low light, deep shadows, and partially covered windows. Upcoming research will plan to conduct a more thorough analysis of these conditions to better understand the limitations of our current model and explore possible improvements. The influential factors behind photographs being identified incorrectly with the trained data of window photographs are summarized as follows. These factors affect model training accuracy. They must be considered during the collection, preprocessing, and labeling of window photograph data.

1. When taking window photographs, environmental conditions related to lighting and reflection should be considered. Thus, taking photographs without a backlight is required during window photograph collection. It is also important to take and collect photographs after securing sufficient brightness using both natural and artificial light.
2. Biased compositions or photographs that emphasize only certain areas must be avoided during collection or labeling. It was confirmed that photographs with a biased window composition decreased the recognition rate during model learning. It is proposed that photographs that contained the entire window structure should be considered and that the training dataset should be constructed by excluding biased photographs during data refining or labeling. For severely tilted photographs, the use of the data refined using the rectifying function is proposed.
3. It was inferred that the presence of surrounding obstacles other than the window decreased the window recognition rate. Therefore, it is important to organize the surrounding environment to the extent possible while taking photographs; furthermore, while taking photographs, the focus should be on the window alone. Additionally, as the accuracy was affected by both the internal and external environments of a window, the background photograph would need to be obscured using the out-focusing

method. Moreover, as the window guard affected the frame material and number of windows, it would be necessary to perform training on the window guard separately. If the results of recognizing the window guard alone were excluded during the window component learning, the correct answer rate would have been higher.

Second, the amount of high-quality data needs to be increased. Numerous data elements are required for deep learning. In this study, however, data were collected at a low cost from five cities in the country through crowdsourcing, and only 1800 data elements were used for learning. A higher level of accuracy and reliability could be achieved using a greater amount of high-quality data. In this study, data augmentation was employed to enhance the learning dataset, aiming to mitigate the potential insufficiency of the collected data for deep learning. Despite the extensive collection of window photographs, the dearth of data with specific window annotations decreased the efficacy of the deep-learning models. Moreover, achieving a balanced representation for each class proved challenging. For example, although both horizontally sliding and hinged windows were photographed, the visited houses predominantly contained windows of the former type, posing challenges to data learning. To address this disparity, future field surveys will prioritize the collection of photographs representing diverse window-opening methods. In addition, while this study utilized  $256 \times 256$  resolution images to match the input size of our pre-trained models, we hypothesize that higher resolution images could potentially improve the accuracy by providing more detailed features for the model to learn from. We propose to investigate this in future work, as it extends beyond the scope of the current study due to computational and resource constraints.

Third, it is possible to increase the amount of learning data by generating sufficient synthetic data through a new artificial intelligence (AI) method. To address the problem of insufficient data with annotations, many studies have been conducted to reduce the effort required to construct learning datasets. In recent years, attempts have been made to automatically and cost-efficiently generate synthetic data in controllable and computable virtual environments to augment the amount of learning data [28,46,47]. In virtual environments, all elements can be edited, including virtual cameras (unique parameters, locations, and poses), object or scene models (geometry, material, and texture), and environment (e.g., weather and lighting conditions). Thus, users can capture various images in an unlimited manner and insert annotations into such images flexibly and automatically [47,48]. According to these studies, the related deep-learning models trained with synthetic images or the actual image dataset augmented using synthetic images can exhibit higher or similar performance compared to the models trained with actual images alone [49]. In future research, methods to further increase the accuracy of the models using synthetic data will be explored.

Fourth, as the performance of AI models has been rapidly improving, it is necessary to implement optimal pre-trained models, as well as the Resnet50 and EfficientNet-B0 models used in this study. In the future, it is also necessary to create optimal algorithms that can identify window components by upgrading the algorithms to fit hyperparameter-tuning or the corresponding domain. Upgraded algorithms will be applied to actual building energy-efficiency improvement projects, and further research will be continued to reduce energy poverty.

Finally, in future research, it is necessary to create an evaluation method that can determine the insulation status of windows in the field as well as improve the performance of algorithms. The information extracted in this study is helpful in partially estimating the insulation status of windows; however, it did not comprise any formula or weight in any form, rendering it only an element of subjective judgment. It is necessary to create methods to determine the grade or score for windows in the field with extracted information on the window frame-material type, number of windows, and area, and to comprehensively evaluate the information related to the insulation performance of different windows. This was the motivation for this study, and diverse image processing methods through window photographs are expected to be the basis for further development.



This research marks a significant advancement in the field through the use of deep-learning models, specifically Resnet50 and EfficientNet-B0, to accurately analyze window components, a methodology previously unexplored in building energy-efficiency studies. Our study's application of advanced AI techniques to the complex task of identifying window numbers and materials from photographs presents a complete workflow for energy-efficient renovation in buildings. The distinct framework used in this research, which can overcome obstacles like unfavorable lighting and reflections, showcases the potential of image-based analysis in the decision-making process of building renovation assessments.

## 5. Conclusions

This study marks a significant stride toward enhancing energy efficiency, with a special focus on low-income households, by introducing a methodology for window component detection. Globally, numerous initiatives aim to retrofit buildings, ranging from residential homes to commercial offices, to improve their energy performance. Particularly in South Korea, the scope of energy-efficiency improvement projects has been broadened to assist energy-poor households, which have limited means to upgrade their living spaces. The urgency to enhance the thermal performance of windows, a critical component of a building's envelope, is evident, especially when operating under financial and temporal constraints. This research illustrates a comprehensive methodology for gathering, preprocessing, and analyzing window photographs to facilitate the identification of window components in need of replacement or upgrade. In conclusion, our novel workflow not only reduces cost and manpower requirements for field surveys but also facilitates more informed decision making based on precise data analysis. By streamlining the initial stages of building renovation surveys through this approach, we can expand the reach of renovation projects to include a greater number of energy-vulnerable households.

Building on this groundwork, the study has unveiled the promising potential of leveraging advanced image processing and deep-learning techniques to refine the assessment of window insulation performance. The inclusion of crowdsourced data has not only enhanced the scalability of our methodology but also its applicability across diverse environments, showcasing a novel way to engage communities in energy-efficiency efforts. Collaborative efforts with government bodies, non-profit organizations, and industry stakeholders will be pivotal in translating our research findings into actionable strategies that can be integrated into existing building renovation frameworks. Such partnerships will ensure that the advancements made through this study can have a tangible impact on the lives of those most affected by energy poverty, aligning with broader goals of sustainability and climate resilience.

Looking ahead, the pursuit of higher-resolution imagery stands out as a promising research direction, potentially unlocking finer details and features critical for assessing window performance more accurately. This endeavor will require a careful balance between computational demands and the depth of analysis achievable with higher-quality images. Additionally, adapting our models to better handle various environmental conditions remains a key challenge. Overcoming these obstacles will not only enhance the reliability of our methodology but also its versatility in real-world applications.

Therefore, this study not only proposes a novel workflow that could revolutionize the preliminary stages of energy-efficient renovations but also lays a solid foundation for future advancements in the field. Through continued innovation and collaboration, we aspire to significantly improve the energy efficiency of buildings, particularly for low-income and energy-poor households, thereby contributing to a more sustainable and equitable future.

**Funding:** This work was supported by the Korea Institute of Civil Engineering and Building Technology (KICT) [Grant number 20230088-001, 20240094-001].

**Data Availability Statement:** The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The author would like to acknowledge the support and advice provided by NGOs, social workers, and volunteers who participated in this project.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- González-Eguino, M. Energy poverty: An overview. *Renew. Sustain. Energy Rev.* **2015**, *47*, 377–385. [\[CrossRef\]](#)
- Henry, C.L.; Baker, J.S.; Shaw, B.K.; Kondash, A.J.; Leiva, B.; Castellanos, E.; Wade, C.M.; Lord, B.; Van Houtven, G.; Redmon, J.H. How will renewable energy development goals affect energy poverty in Guatemala. *Energy Econ.* **2021**, *104*, 105665. [\[CrossRef\]](#)
- IEA. *Energy Access Outlook 2017: From Poverty to Prosperity*; International Energy Agency: Paris, France, 2017.
- Herrero, S.T.; Ürge-Vorsatz, D. Trapped in the heat: A post-communist type of fuel poverty. *Energy Pol.* **2012**, *49*, 60–68. [\[CrossRef\]](#)
- Boardman, B. Fuel Poverty. In *International Encyclopedia of Housing and Home*; Smith, S.J., Ed.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 221–225. [\[CrossRef\]](#)
- Boute, A. Modernizing the Russian district heating sector: Financing energy efficiency and renewable energy investments under the new federal heat law. *Pace Environ. Law Rev.* **2012**, *29*, 746–810. [\[CrossRef\]](#)
- Shin, D.H.; Kim, S.H.; Kim, J.H.; Kim, S. Experimental analysis of low-cost energy retrofit strategies for residential buildings to overcome energy poverty. *Case Stud. Therm. Eng.* **2022**, *32*, 101874. [\[CrossRef\]](#)
- Karásek, J.; Pavlica, J. Green investment scheme: Experience and results in the Czech Republic. *Energy Pol.* **2016**, *90*, 121–130. [\[CrossRef\]](#)
- Cole, J.C.; McDonald, J.B.; Wen, X.; Kramer, R.A. Marketing energy efficiency: Perceived benefits and barriers to home energy efficiency. *Energy Effic.* **2018**, *11*, 1811–1824. [\[CrossRef\]](#)
- Figus, G.; Lecca, P.; McGregor, P.; Turner, K. Energy efficiency as an instrument of regional development policy? The impact of regional fiscal autonomy. *Reg. Stud.* **2019**, *53*, 815–825. [\[CrossRef\]](#)
- Turcu, C. Unequal spatial distribution of retrofits in Bucharest’s apartment buildings. *Build. Res. Inf.* **2017**, *45*, 892–909. [\[CrossRef\]](#)
- Perez-Bezos, S.; Grijalba, O.; Irulegi, O. Proposal for prioritizing the retrofitting of residential buildings in energy poverty circumstances. *Environ. Clim. Technol.* **2020**, *24*, 66–79. [\[CrossRef\]](#)
- Willand, N.; Moore, T.; Horne, R.; Robertson, S. Retrofit poverty: Socioeconomic spatial disparities in retrofit subsidies uptake. *Build. Cities* **2020**, *1*, 14–35. [\[CrossRef\]](#)
- Egner, L.E.; Klöckner, C.A.; Pellegrini-Masini, G. Low free-riding at the cost of subsidizing the rich. Replicating Swiss energy retrofit subsidy findings in Norway. *Energy Build.* **2021**, *253*, 111542. [\[CrossRef\]](#)
- Han, H.S. Energy welfare status and cases. *J. Electr. World Month. Mag.* **2020**, *12*, 34–39.
- Gartland, L.M.; Emery, A.F.; Sun, Y.S.; Kippenhan, C.J. Residential energy usage and the influence of occupant behavior. In Proceedings of the ASME Winter Annual Meeting, The American Society of Mechanical Engineers, New Orleans, LA, USA, 28 November–3 December 1993; pp. 1–9.
- Jack, R.; Loveday, D.; Allinson, D.; Lomas, K. Quantifying the effect of window opening on the measured heat loss of a test House. In *Sustainable Ecological Engineering Design*; Dastbaz, M., Gorse, C., Eds.; Selected Proceedings from the International Conference of Sustainable Ecological Engineering Design for Society (SEEDS); Springer International Publishing: New York, NY, USA, 2016; pp. 183–196.
- Howard-Reed, C.; Wallace, L.A.; Ott, W.R. The effect of opening windows on air change rates in two homes. *J. Air Waste Manag. Assoc.* **2002**, *52*, 147–159. [\[CrossRef\]](#)
- Fabi, V.; Andersen, R.V.; Corgnati, S.P.; Olesen, B.W. Occupants’ window opening behavior: A literature review of factors influencing occupant behavior and models. *Build. Environ.* **2012**, *58*, 188–198. [\[CrossRef\]](#)
- Luong, D.; Richman, R.; Touchie, M. Towards window state detection using image processing in residential and office building facades. *Build. Environ.* **2022**, *207-B*, 108486. [\[CrossRef\]](#)
- Neuhausen, M.; König, M. Automatic window detection in facade images. *Autom. Constr.* **2018**, *96*, 527–539. [\[CrossRef\]](#)
- Lee, S.C.; Nevatia, R. Extraction and integration of window in a 3D building model from ground view images. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; p. 113. [\[CrossRef\]](#)
- Neuhausen, M.; Obel, M.; Martin, A.; Mark, P.; König, M. Window detection in facade images for risk assessment in tunneling. *Vis. Eng.* **2018**, *6*, 1. [\[CrossRef\]](#)
- Cheng, K.; Yao, J.; Zheng, R. Energy performance of occupant behaviors on windows: A green building based study. *Energies* **2023**, *16*, 2209. [\[CrossRef\]](#)
- Tien, P.; Wei, S.; Calautit, J.K.; Darkwa, J.; Wood, C. Detection of window opening using a deep learning approach for effective management of building ventilation heat losses. In Proceedings of the 5th IBPSA-England Conference on Building Simulation and Optimization (Virtual), Loughborough, UK, 21–22 September 2020.
- Schnelle, L.; Lichtenberg, G.; Warnecke, C. Using low-rank multilinear parameter identification for anomaly detection of building systems. *IFAC-Pap.* **2022**, *55*, 470–475. [\[CrossRef\]](#)
- Erhan, L.; Ndubuaku, M.; Di Mauro, M.; Song, W.; Chen, M.; Fortino, G.; Bagdasar, O.; Liotta, A. Smart anomaly detection in sensor systems: A multi-perspective review. *Inf. Fusion* **2021**, *67*, 64–79. [\[CrossRef\]](#)

28. Ying, H.; Sacks, R.; Degani, A. Synthetic image data generation using BIM and computer graphics for building scene understanding. *Autom. Constr.* **2023**, *154*, 105016. [CrossRef]
29. Biljecki, F.; Chow, Y.S.; Lee, K. Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes. *Build. Environ.* **2023**, *237*, 110295. [CrossRef]
30. Yuan, J.; Chowdhury, P.K.R.; McKee, J.; Yang, L.; Weaver, J.; Bhaduri, L.S. Exploiting deep learning and volunteered geographic information for mapping buildings in Kano, Nigeria. *Sci. Data* **2018**, *5*, 180217. [CrossRef]
31. Zhang, Z.; Qian, Z.; Zhong, T.; Chen, M.; Zhang, K.; Yang, Y. Vectorized rooftop area data for 90 cities in China. *Sci. Data* **2022**, *9*, 66. [CrossRef] [PubMed]
32. Sirko, W.; Kashubin, S.; Ritter, M.; Annkah, A.; Bouchareb, Y.S.; Dauphin, Y.; Keysers, D.; Neumann, M.; Cisse, M.; Quinn, J. Continental-scale building detection from high resolution satellite imagery. *arXiv* **2021**, arXiv:2107.12283. [CrossRef]
33. Huang, X.; Wang, C. Estimates of exposure to the 100-year floods in the conterminous United States using national building footprints. *Int. J. Disaster Risk Reduct.* **2020**, *50*, 101731. [CrossRef]
34. Peters, R.; Dukai, B.; Vitalis, S.; van Liempt, J.; Stoter, J. Automated 3D reconstruction of LoD2 and LoD1 models for all 10 million buildings of the Netherlands. *Photogramm. Eng. Remote Sens.* **2022**, *88*, 165–170. [CrossRef]
35. Dukai, B.; Peters, R.; Wu, T.; Commandeur, T.; Ledoux, H.; Baving, T.; Post, M.; van Altena, V.; van Hinsbergh, W.; Stoter, J. Generating, storing, updating, and disseminating a country-wide 3D model. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *44*, 27–32. [CrossRef]
36. Yan, D.; O'Brien, W.; Hong, T.; Feng, X.; Gunay, H.B.; Tahmasebi, F.; Mahdavi, A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **2015**, *107*, 264–278. [CrossRef]
37. Bourikas, L.; Costanza, E.; Gauthier, S.; James, P.A.; Kittle-Davies, J.; Ornaghi, C.; Rogers, A.; Saadatian, E.; Huang, Y. Camera-based window-opening estimation in a naturally ventilated office. *Build. Res. Inf.* **2016**, *46*, 148–163. [CrossRef]
38. Zheng, H.; Li, F.; Cai, H.; Zhang, K. Non-intrusive measurement method for the window opening behavior. *Energy Build.* **2019**, *197*, 171–176. [CrossRef]
39. Amazon SageMaker Data Labeling Pricing. Available online: <https://aws.amazon.com/sagemaker/data-labeling/pricing/> (accessed on 10 February 2023).
40. AI Platform Data Labeling Service. Available online: <https://cloud.google.com/ai-platform/data-labeling/docs> (accessed on 10 February 2023).
41. Northcutt, C.G.; Athalye, A.; Mueller, J. Pervasive label errors in test sets destabilize machine learning benchmarks. In Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks, Virtual, 6 December 2021. [CrossRef]
42. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
43. Qiu, Q.; Lau, D. Real-time detection of cracks in tiled sidewalks using YOLO-based method applied to unmanned aerial vehicle (UAV) images. *Autom. Constr.* **2023**, *147*, 104745. [CrossRef]
44. Tan, M.; Le, Q.V. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (ICML), Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
45. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process.* **2019**, *32*. [CrossRef]
46. Siu, C.; Wang, M.; Cheng, J.C.P. A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection. *Autom. Constr.* **2022**, *137*, 104213. [CrossRef]
47. Hwang, J.; Kim, J.; Chi, S. Site-optimized training image data-base development using web-crawled and synthetic images. *Autom. Constr.* **2023**, *151*, 104886. [CrossRef]
48. Tobin, J.; Fong, R.; Ray, A.; Schneider, J.; Zaremba, W.; Abbeel, P. Domain randomization for transferring deep neural networks from simulation to the real world. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada, 24–28 September 2017; pp. 23–30. [CrossRef]
49. Assadzadeh, A.; Arashpour, M.; Brilakis, I.; Ngo, T.; Konstantinou, E. Vision-based excavator pose estimation using synthetically generated datasets with domain randomization. *Autom. Constr.* **2022**, *134*, 104089. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.