# Real-time monitoring of occupancy activities and window opening within buildings using an integrated deep learning-based approach for reducing energy demand

Paige Wenbin Tien[1*], Shuangyu Wei[1], John Calautit[1], Jo Darkwa[1] and Christopher Wood[1]

[1]*Department of Architecture and Built Environment, University of Nottingham, Nottingham, NG7 2RD, UK*

[*]*Corresponding author. paige.tien@nottingham.ac.uk, paige.tien@gmail.com*

**Abstract**

Occupancy behaviour in buildings can impact the energy performance and operation of heating, ventilation and air-conditioning (HVAC) systems. HVAC, which uses conventional control strategies or "fixed" setpoint schedules, could not adjust to the conditioned spaces' actual requirements, resulting in building spaces being over or under-conditioned. While the unintended opening of windows can lead to substantial heat loss and consequently raises energy consumption. To optimise building operations, it is necessary to employ solutions such as demand-driven controls, which can monitor the utilisation of indoor spaces and provide the actual thermal comfort requirements of occupants. This study presents a novel vision-based deep learning framework for occupancy activity detection and recognition including the manual window operations in buildings. A region-based Convolutional Neural Network (R-CNN) model was trained and deployed to a camera for real-time detection and recognition. Based on the field experiments conducted within a case study University building, overall accuracy of 85.63% was achieved for occupancy activity detection and 92.20% for window operation detection. Building energy simulation and various scenario-based cases were used to assess the impact of such an approach on the building energy demand and provide insights into how the proposed detection method can enable HVAC systems to respond to dynamic changes within indoor spaces. Results showed that the proposed approach could reduce the over-or under-estimation of occupancy heat gains compared with the use of "fixed" or static profiles. In addition, the approach can help alert building users or managers about windows left open unintentionally, which can reduce unnecessary ventilation heat losses. Furthermore, the approach can also predict the room $CO_2$ concentration and advise occupants about a suitable natural ventilation strategy. The study highlighted the potential of the multi-purpose detection approach, but further development is necessary, including optimisation of the deep learning model, full integration with HVAC controls and further model training and field testing.

**Keywords:** Deep learning, building energy management, building ventilation, window opening, window and occupancy detection, HVAC systems

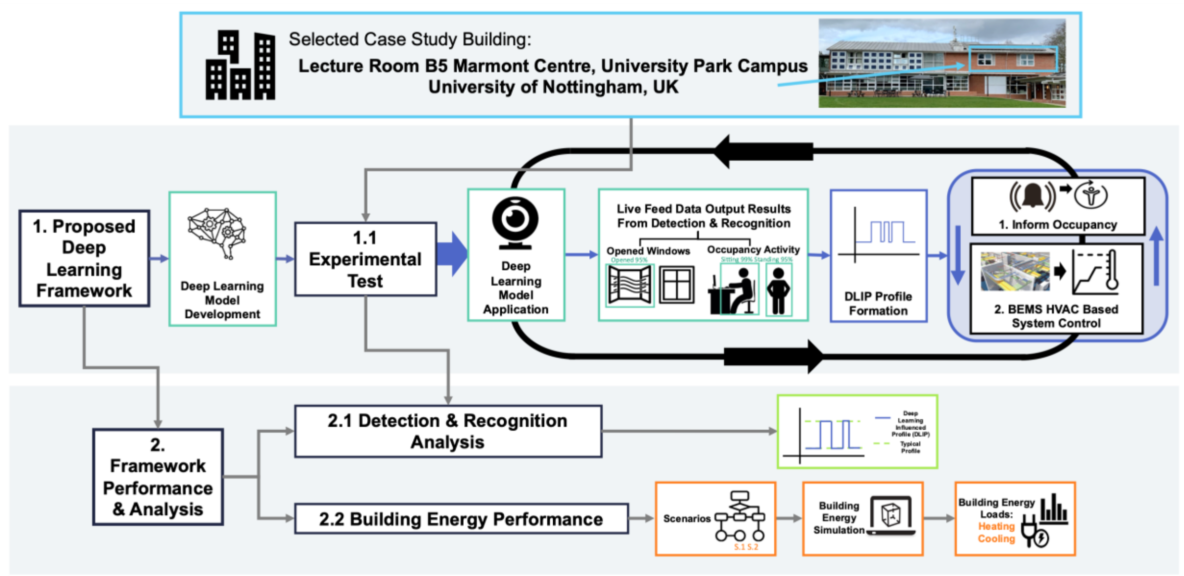
**Word count:** 8279 (excluding the abstract, references, tables and figures)

**Highlights**

- A vision-based approach was developed to detect occupancy activities and window conditions.
- A deep learning model was trained and deployed to a camera for multi-object detection.
- Field testing was carried out by performing real-time detection in a university lecture room.
- Results showed that the approach could reduce the under or overestimation of heat gains.
- The impact on heat loss, energy demand and indoor air quality was evaluated using building simulation.

The short version of the paper was presented at The 12th International Conference on Applied Energy ICAE2020, Dec 1-10, 2020. This paper is a substantial extension of the short version of the conference paper.

## Graphical Abstract




## 1. Introduction and Literature Review

Heating, ventilation, and air-conditioning (HVAC) systems are designed to ensure comfortable indoor thermal conditions within a building space. The operations of large commercial HVAC systems are typically monitored and controlled by a building energy management system (BEMS) to optimise building performance and reduce energy use. Several HVAC scheduling techniques have been proposed and developed for maximising energy savings while maintaining the occupant's thermal comfort [1]. This includes basic scheduling technique, which involves controlling the 'ON' and 'OFF' states of the HVAC system, conventional scheduling, which involves controlling the setpoint temperatures and advanced scheduling – an enhanced technique based on the two scheduling techniques or the combinations of both [1].

The most popular technique implemented in buildings would be the second technique which assumes a 24 hour per day HVAC operation, and the setpoint temperatures are varied according to the occupancy level. The setpoint temperatures and schedules are based on guidelines such as ASHRAE 90.1 [2] and ASHRAE 55 [3]. For example, during occupied hours, a temperature range of 22 – 27°C is recommended for cooling and 17 – 22°C for heating. For unoccupied hours, temperatures of 27 – 30°C for cooling and 14 – 17°C for heating are advised. Such guidelines suggest setpoint temperatures depending on the purpose of the building space or room type.

For example, CIBSE [4] suggests office buildings in the UK be maintained at an operative room temperature of 21 – 23°C during the winter and 22 – 25°C for summer. However, more building spaces are being designed to be multifunctional, with variations in the room's function or space. In addition, some building spaces do not remain to be utilised as initially designed or planned due to changes in occupancy patterns and more flexible working hours. Hence, an HVAC system that employs fixed or predefined schedules cannot adapt to such changes and could result in over or under conditioned rooms. With the largest proportion of the total energy consumption of buildings due to the operations of HVAC systems (up to 35% [5]), solutions that optimise the operation to improve the HVAC systems' energy efficiency while maintaining indoor thermal comfort are necessary.

Advanced control strategies such as in [6] can be used to operate HVAC system by considering the real-time data of the indoor environment such as temperature and $CO_2$ concentration. The proposed strategy effectively reduced the energy demand of the HVAC system of irregularly occupied spaces. However, many buildings are still not equipped with such advanced control technology despite its clear advantage. Based on the US energy information administration (EIA) [7], a large proportion of U.S. households still utilise thermostatic controls based on ON/OFF or manual thermostat-based systems. Reasons for preventing occupants from employing such systems include the complexity and uncertainties of the approach and the unwillingness of occupants to change existing systems and adopt novel solutions. Kontes et al. [8] suggest that the lack of adoption of advanced control techniques in the industry, such as model predictive control, is due to the increased cost of developing and identifying models for predicting future states of the building.

Occupancy behaviour can significantly affect the operations of an HVAC system [9]. According to Chen et al. [10], there are several categories of occupancy behaviour that influences building energy consumption. It includes occupancy presence and the number of occupants in a space and their interactions (occupancy activities, interactions with window opening and building energy systems). Therefore, to enable effective operations of HVAC systems, occupancy-driven control is necessary. Esrafilian-Najafabadi and Haghighat [11] indicated that HVAC control systems based on predefined occupancy schedules provided inadequate energy-saving and thermal comfort in many cases. They highlighted the need to develop solutions that focus on reactive control, predictive control, and rule-based controls to avoid a mismatch between predefined schedules and the actual daily occupancy patterns.

Occupancy actions that impact the indoor environment and thermal comfort include directly interacting with thermostats, opening or closing windows, and changing their clothing [12]. These actions are mainly due to dissatisfaction with the thermal and/or indoor air quality conditions. Hence, developed control strategies for HVAC systems must present solutions for energy consumption, thermal comfort and air quality. However, the occupancy-based control strategies evaluated by [11] suggest that up to 58% of the studies focused on reducing energy consumption and only 26% on providing thermal comfort.

More guidelines are being put in place to encourage the use of natural ventilation strategies in buildings such as the UK [13] due to its clear benefits in terms of indoor air quality and minimising the energy use and cost associated with mechanical ventilation and cooling [14]. On the other hand, providing adequate thermal comfort conditions could become more challenging as actions such as window opening to enable natural ventilation in a room can significantly impact the indoor conditions and are difficult to monitor/control. With natural ventilation, the internal condition will depend on the indoor-outdoor conditions [15], seasonal environmental factors [16], window opening patterns [17] and personal preferences [18]. The correct use of natural ventilation can lead to significant energy saving [19]; however, its incorrect usage can also lead to unnecessary energy demand. In mild or temperate climates such as the UK, windows left open can lead to significant heat losses. Hence, the interaction of occupants with natural ventilation devices should also be considered when developing control strategies. Based on the evaluation made by [20], significant reductions in energy demand between 8.5 % and 44.8 % could be potentially achieved if such solutions can be implemented.

Recently, many studies have been focusing on employing data and demand-driven solutions for enhancing the operation and performance of HVAC [21]. These solutions use real-time and historical data, such as occupancy count and presence. Sensor-based approaches are typically employed to obtain information about occupancy behaviour within buildings. This includes technology such as infrared [22], Wi-Fi [23] and Radio Frequency Identification (RFID) [24] to provide data about occupancy count and location. Khalilnejad et al. [25] proposed a data-driven method to identify savings opportunities using whole building meter data to change thermostat setpoint temperature and reschedule the HVAC operations. The method detects occupancy patterns and quantifies the baseload of the HVAC operation. Energy savings of up to 2.1% was achieved from HVAC rescheduling. These data can also be used to determine the energy effectiveness of the commissioning of HVAC systems, as detailed in [26]. Results achieved suggests 2–6% of the total energy could be reduced by implementing the HVAC commissioning.

Furthermore, indirect methods such as environmental-based sensors can also be used [27] to monitor the changes within the space. The data collected can be used for demand response-based solutions for more effective system controls [28], energy optimisation [29], and building energy management [30]. However, such techniques also have several limitations, such as requiring multiple sensors distributed across the room and not detecting occupancy activities in real-time, impacting indoor conditions. With diverse and varying occupancy patterns and indoor-outdoor conditions, such sensors can suffer from time delays of measurements and diluted representation due to the mixing of air in spaces, leading to detection error [31]. Vision-based detection offers a promising solution and can provide a higher occupancy resolution level. It can provide information about the presence, count and activity of the occupants in the space, enabling better control and flexible management of HVAC. Such detailed information can help predict how much heat, $CO_2$, and contaminants are produced by the occupants and how they interact with appliances and lighting, producing heat and windows and openings that affect the air and heat exchange [32].

### *1.1 Literature Gap*

The explored solutions and approaches suggest that many works used direct and indirect detection methods, which may not be able to provide fine resolution levels of occupancy information. A potential solution is to use AI-based techniques such as computer vision and deep learning that can help detect and monitor the usage of building spaces. To provide accurate detection and recognition, computer vision and deep learning approaches can be utilised [33]. This has typically been used for the development of detectors that focuses on the detection and recognition of occupancy behaviour [34], window conditions [35] and also issues related to the diagnostics of both damage and faults [36]. For example, Chahyati et al. [37] and Ahmad et al. [38] employed it for people tracking to assist surveillance, while Kajabad and Ivanov [39] detected people's behaviour for the understanding of the share of overall traffic in areas within shops and museums. Despite the ability to provide real-time detection of occupancy and behaviour [40], there are limited developments and studies that focus on its application for assisting systems and operations within the built environment, in particular HVAC and natural ventilation.

Previous works employed different types of AI and occupancy-based techniques to solve building energy-related problems, such as prediction methods [41] and energy management [42, 43]. Many works focused on strategies to enhance controls [28] to optimise HVAC system operations [44, 45] and [46, 47] thermal comfort management. However, studies on the integration of vision-based approaches with HVAC controls are limited. Several studies, such as [35], have attempted to estimate the impact of the vision-based detection and recognition approach on the building energy demand and thermal comfort. The initial works explored detection methods for occupancy activities, usage of equipment and window operation, which can be used to estimate internal heat gains and ventilation heat loss. One of the advantages of such an approach is the capability to carry out multiple types of detection in a space using a single sensor or camera. However, the integration of these methods has not been explored previously. There is a concern that combining all strategies into one could cause a reduction in detection performance; this will be explored in this study. Furthermore, many of the works have been carried out in office-type buildings or environments, and hence its capability to adapt and be utilised in different conditions [48, 49] must be explored. This work will address this by applying the method in a classroom environment. One of the main challenges of building controls is the conflict between reducing energy consumption and improving comfort [50] and air quality. Hence, this work will also explore how a multi-objective system based on computer vision approaches can be achieved.

Focusing on the detection of window conditions (open/close), while there are many methods available, there is limited research on the use of window detection to aid demand-driven control solutions for energy and comfort management in buildings. For example, Zheng et al. [51] proposed a non-intrusive measurement (from outdoor) method to identify window positions and their opening proportion. An image recognition-based approach was proposed. Photographs of windows from various angles were collected and then processed to further understand the window opening state. In conjunction with this, data of the internal conditions were also collected, which enabled the analysis of the direct impact of window openings on indoor temperature. However, the influence on energy performance and integration with HVAC was not discussed. Based on the review of relevant works, there are limited studies on computer vision-based window detection methods that could provide real-time information of the window state or condition for building occupants and building control systems.

Furthermore, many of the previous works are mainly focused on enhancing the model's performance, such as its accuracy, speed, etc., in detecting the number and distribution of occupants or equipment in a space. To date, only limited studies attempted to demonstrate the usage of the detected information to control the operation of HVAC for optimising energy efficiency, thermal comfort and air quality. In addition, only limited studies [52, 53] employed such solutions to detect and recognise the occupant's activities and predict the heat emitted to the space by the occupants (sensible and latent heat gains) and usage of equipment. Furthermore, the impact of the approach implementation on energy consumption and practicality should be explored. This will be addressed in this work by carrying out simulations and analysis of different scenarios.

### *1.2 Aims and objectives*

The present work will build on the novel approaches introduced in [52, 54] to develop and evaluate a framework (Figure 1) that enables the real-time detection and recognition of the occupancy activities and conditions of windows being opened or closed by occupants within a building space. The proposed approach can provide real-time prediction of the internal heat gains and detection of the status of windows (open/close) for building control systems. This can enable the adjustment of the operations of building HVAC systems to ensure that adequate indoor thermal conditions and air quality are achieved while minimising unnecessary building energy loads. A model based on a faster region-based convolutional neural network (Faster R-CNN) will be trained for the detection and recognition of occupancy activities and window status using a camera. Validation of the approach will be conducted using a set of testing data, and the accuracy and suitability for live detection will also be evaluated. Field experiments will be carried out within a case study university lecture room to test the capabilities of the proposed approach. Using building energy simulation (BES), the case study building was simulated with different scenario-based operation profiles to assess the indoor air quality and potential energy savings that can be achieved.

## 2. Method

The proposed research approach is given in Figure 1. As highlighted, a case study lecture room within a university building was selected to assist the testing and evaluation of the application of such an approach. Furthermore, Figure 1 outlines the key stages of the method. This includes the steps of developing and implementing the proposed vision-based deep learning framework in Part 1 and the analysis of the utilisation of the deep learning model for real-time detections from the experimental test and under scenario-based situations using building energy simulation (BES) in Part 2. Further details of each of the steps highlighted in Figure 1 are presented in the following sub-sections.

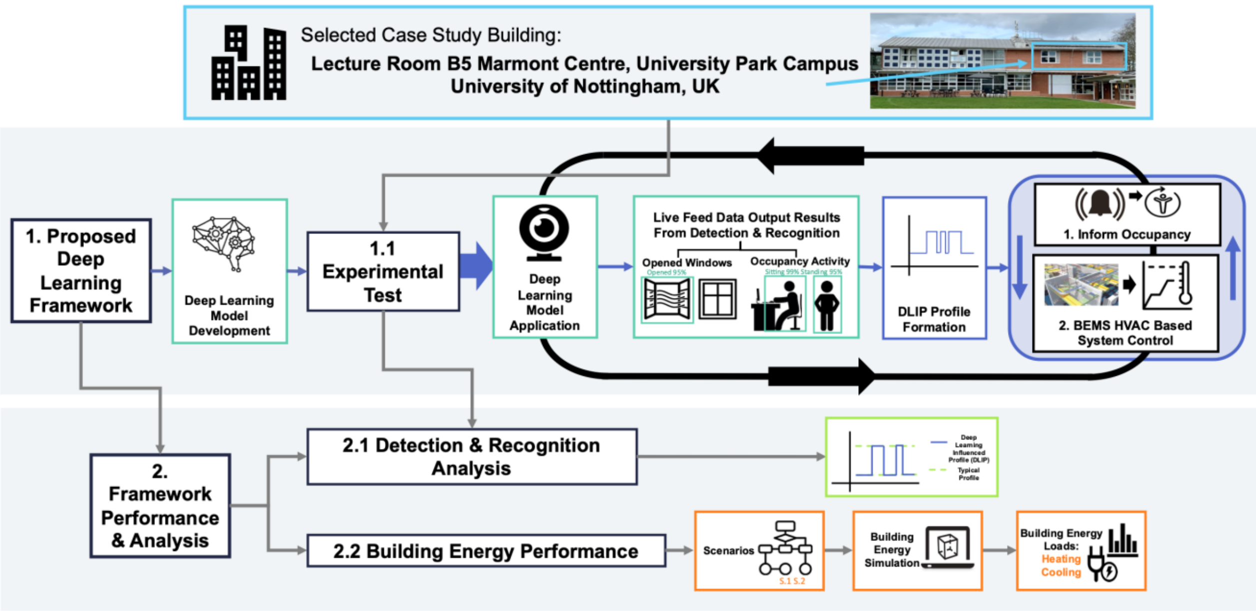



Figure 1. Overview of the research method.

### 2.1. Case Study Building

The lecture room located on the first floor of the Marmont Centre at the University of Nottingham (University Park Campus, Nottingham, UK, Figure 2a) was selected as the case study building and used to support the testing of the proposed deep learning vision-based approach. As shown in Figure 1, where it indicated that when the vision-based models for the detection and recognition of occupancy activities and window conditions were successfully developed, and when the models were deployed to form an AI-powered camera, it allowed the camera to perform real-time based detections and recognitions. Hence, this room was used to perform the initial detection during a typical afternoon on a winter day. Figure 2b presents the experimental test setup, and the corresponding floor plan is shown in Figure 2c.

As shown in Figure 2d, the room consists of four sets of windows with two different configurations. The north-facing windows had an arrangement of 2 x 3 with a total of six 0.915m x 0.416m (0.38m$^2$) glazing panels. The two south-facing windows had an arrangement of 4 x 4 with a total of 8 0.835m x 0.657m (0.55m$^2$) glazing panels. All windows had the same properties where they were double glazed with a U-value of 2.20 W/m$^2$K, and that they all had the same opening strategy (top-hung opening). Hence, to enable initial evaluation of the detection and recognition performance, the detection camera was positioned so that the detection was focused on a section of the room, and specifically towards the 'South Facing Windows 1'. Furthermore, the camera used was a standard 1080p camera with a wide 90-degree field of view, and it was positioned at a height near the ceiling of the room, replicating the position for a typical room ceiling sensor.

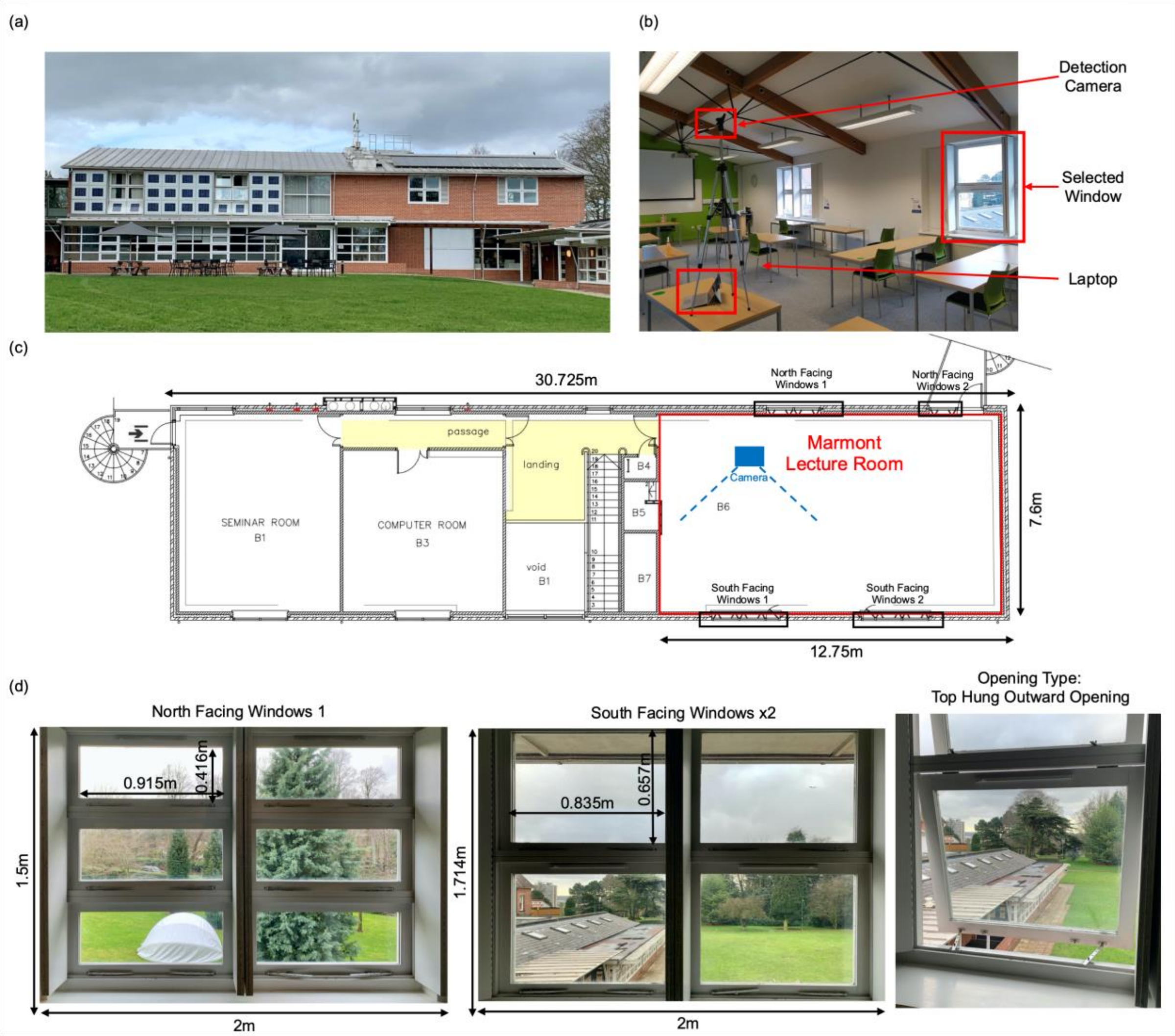



Figure 2. (a) Marmont Centre at the University of Nottingham, UK. (b). Set up for the experimental test. (c). 1st Floor plan. (d) Design and configuration of windows within the Marmont lecture room.

The room was also modelled using Building Energy Simulation (BES) tool IES VE [57] to further assess this framework's potential and the impact of the method on building energy loads. The building is naturally ventilated and is integrated with a central heating system. The selected room has a floor area of 96.9m$^2$ with dimensions of 12.75m x 7.6m and a floor to ceiling height of 2.5m. Furthermore, for BES, the maximum opening area was set to 50% for an opened window. The wall, roof, ground and doors U-values were 0.33, 0.22, 0.32 and 3.00 W/m$^2$K.

Both lectures and tutorial sessions occur within the room for architectural engineering students. The room is also accessible for students during the weekends; hence the building operational hours of 09:00 – 18:00 would be assumed for all days of the week. Furthermore, Nottingham, UK weather data file was used for the simulation. Heating profiles were set to maintain an indoor temperature of 21°C during occupied hours [3,4]. Details about the associated profiles for windows and occupancy assigned and any modifications to the building heating and cooling profiles were given in the corresponding building energy and simulation section. The infiltration rate value was assumed to be constant and set at 0.5 air changes per hour.

Furthermore, the modelling of the windows for the building energy simulation cases consisted of an exposed wall type exposure, with a top hung window opening. The windows which were assumed to be opened were assigned with an openable area of 50% and a maximum openable angle of 45°. The degree of the opening was assigned with a modulating profile corresponding to the window profiles created for each simulation case.

### 2.2. Vision-Based Deep Learning Framework

Computer vision is one of the most common fields of artificial intelligence (AI). It utilises a technique that enables computers to understand and recognise the present objects and their characteristics that include shapes, textures, colours, sizes, spatial arrangement, among other things, to provide a description as complete as possible of the image [58]. Recently, deep learning methods have been extensively used to assist the field of computer vision. As part of our daily lives, deep learning with computer vision has been commonly used to provide applications that include object detection [59], face recognition [60], and also machine vision in self-driving cars [61].

The most common deep learning algorithm used for various computer vision tasks is the convolutional neural network (CNN) [62]. It requires input data in the form of further processed images. Based on the desired responses required from the detector, the model must be configured prior to training. Next, once the model is sufficiently trained (as shown in Figure 1), the model can be deployed to an AI-powered camera. It should be noted that the model development stage can be an iterative process as models can be continuously improved to provide a better, more accurate detector. This includes repeating the training process with the enhancement of the input image dataset and the CNN model architecture. In the present study, only one model for occupancy activity and window conditions were trained and assessed, however, this can be further explored in future works. To develop the proposed vision-based detector, a deep learning framework was employed. The details for each of the steps are detailed within the following subsections.

#### 2.2.1. Image Datasets and Pre-Processing Stages

Table 1 presents the description of the datasets in terms of the number of images and labels assigned. Overall, the same workflow process as [54, 55] was applied. This consists of gathering the images to form the datasets and manually labeling images using the software, LabelImg. Figure 3 presents an example of the types of images gathered and how they were manually labelled to highlight each image's specific region of interest. As indicated by [54, 55], the number labels assigned to each individual image were also solely based on the content of each image. For most cases (such as the images shown in Figure 3), multiple labels were assigned by highlighting a bounding box around each occupant and on each of the gaps of the windows across all sides of the window.

Table 1. The number of images of occupancy activities and windows within the training and testing dataset.

| Category | Number of Images | | | Number of Labels | | |
|---|---|---|---|---|---|---|
| | Training | Testing | Total | Training | Testing | Total |
| | | | **Occupancy Activities** | | | |
| Sitting | 400 | 100 | 500 | 753 | 149 | 902 |
| Standing | 400 | 100 | 500 | 701 | 134 | 835 |
| Walking | 400 | 100 | 500 | 1000 | 177 | 1177 |
| Total | 1200 | 300 | | 2454 | 460 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Windows** | | | | | | |
| Open | 666 | 160 | 826 | 1398 | 318 | 1716 |


Figure 3. Example images gathered from Google Images to form the image datasets (training and testing) for both categories of occupancy activities and windows, along with examples of how images were manually labelled to highlight the specific region of interest.

For the occupancy activity dataset, ‘sitting, standing and walking’ activities were selected as the desired model detection responses [54, 55]. However, in the present study, the category napping was removed due to the minimal number of occupants performing this type of activity in the room. The category of ‘none’, which represented no person present, was also removed. The present model assumes that no occupant is present within the space when no detection is made. Furthermore, the dataset used in training the model in [54], which had an average detection accuracy of 97.32% for the similar types of occupancy activities, was used and enhanced with up to 400 training and 100 testing images for each category (occupancy activity); giving a total of 1,500 images as compared to 600 images in [54]. As shown in Figure 3, similar images to [55] were collected, and the same labelling method was adopted.

Previously, the window dataset used to develop a window detector in [64] consisted of two categories: ‘open’ and ‘closed’ windows. However, the results suggested that low detection accuracy of 77.78% was achieved when using both categories, leading to a large percentage of incorrect detections and false predictions. Therefore, in the present study, only one category was considered, ‘open window’, and no detection means the windows in the view were closed. Hence, the images gathered for this dataset were different to the dataset used previously. As shown by the example images in Figure 3, the images within the dataset do not have to include the whole (full) window. Instead, it consisted of images that only presents ‘opened windows’ which presented opening types/designs of side-hung, top-hung and pivot (vertical, horizontal).

The reason why only these types of open window images were selected was to demonstrate a different method of labelling. The labelling method assigned bounding boxes to regions where it showed window opening gaps. The change in the types of images for the window dataset and the labelling method resulted in an increase in the number of images used, from a total of 250 images used in [64] to a total of 826 images given in Table 1.

#### 2.2.2. Model Selection, Configuration and Training

Once the images were gathered and pre-processed, a suitable framework platform was selected to configure and train the CNN based model. The TensorFlow Object Detection API was used to develop the occupancy activity and window detector. The TensorFlow API is based on a transfer learning approach. Transfer learning is a machine learning method that utilises a pre-trained model. Therefore, this method allowed the desired model to acquire high detection performance without needing a model to be trained from scratch (which can be very time consuming and the requirement of a larger dataset).

For the present study, a model selected from the TensorFlow detection model zoo was used to assist the pipeline configuration of the model used to train the desired detector. The TensorFlow Detection Model Zoo consisted of a collection of detection models pre-trained on various common image-based datasets. The COCO-trained model of Faster R-CNN (With Inception V2) was selected [54, 55, 64]. Two models were configured and trained separately, with Model A for occupancy activity detection and Model B for window detection. Figure 4 presents the overall architecture and the pipeline configuration of the models used. Once these models were successfully trained, they were combined and deployed in an AI-powered camera.

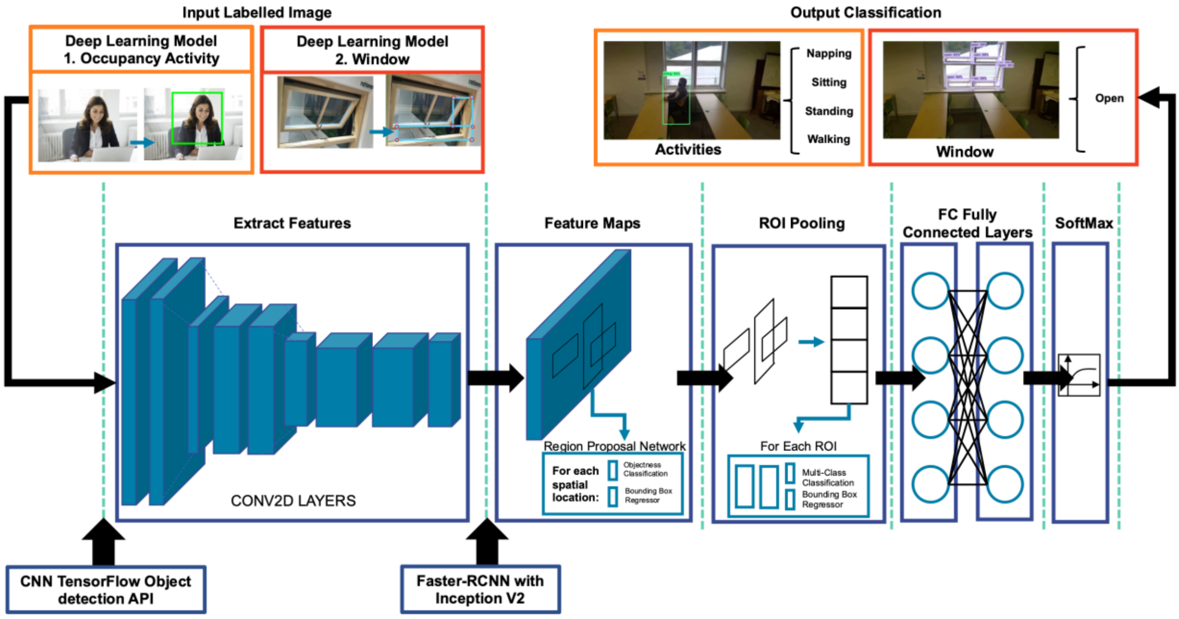



Figure 4. Architecture and configuration of the convolutional neural network (CNN) based model which was used to develop both the occupancy activity and window detector.

### 2.3. Model Application and Real-time Detection

To assess the combined vision-based deep learning detector, a series of tests was conducted. This includes the following: 1. An initial model performance evaluation using common classification evaluation metrics based on the detection performance on still images from the testing dataset (Table 1). 2. An assessment of the detection performance during a real-time experimental test conducted within the selected case study lecture room, and 3. Further assessment of the impact of applying the detection approach on the building energy performance.

#### 2.3.1. Model Performance Evaluation

For the initial test, model performance evaluation was conducted based on the detection and recognition ability using still images from the test dataset. This provided results in the form of a confusion matrix by identifying whether the results provided were classified as either true positive, true negatives, false positives and false negatives. Hence, results presented in the form of a confusion matrix was used to generate results in terms of the common classification-based evaluation metrics of accuracy, precision, recall and the $F_1$ score. Full details about each of these terms can be found in [54].

For the experimental test conducted within the Marmont building lecture room, a full analysis of the detection performance was conducted. This includes evaluating the recognition ability based on the detection accuracy displayed across each of the bounding boxes that appeared at every instance in terms of the generated Intersection over Union (IoU) values. IoU is a standard evaluation metric for convolutional neural network detectors used to evaluate how similar a predicted bounding box is to the ground truth box. For such a case, higher prediction accuracy (near to 100%) would be achieved when there is a direct overlap between the target mask and the prediction output. Further evaluation includes the analysis of each detection in terms of the percentage of the time achieving correct, incorrect and no/missed detections throughout the different segments of the tests.

#### 2.3.2. Real-time Detection and the Formation of the Deep Learning Influenced Profiles (DLIP)

As presented in Figure 1, along with the set-up shown in Figure 2b, a 15-minute experimental test during a typical winter's day afternoon (at 15:00) was performed within the selected room. The experiment is divided into 5 parts. The test started with Part 1 (which lasted for 1 minute), which consisted of all windows being closed and that no occupants were present within the room. Next, Part 2 (15:01 – 15:03) consisted of a person entering the room and performing a series of activities that included sitting, standing, and walking. Within Part 3 (15:03 – 15:08), the person continued to perform the following activities and decided to open all 4 of the windows. Towards the end of Part 3, the person decided to leave the room, and all windows were left open. Hence, in Part 4 (15:08 – 15:11), the windows were all opened, and no occupant was present in the room. Furthermore, Part 5 consisted of the same situation as part 4; however, for this section, the lights were switched off.

During the experimental test, continuous real-time detection provided response output, including sitting, standing, walking, and open window. It was assumed that when none of these responses was made, windows were closed, and no occupancy was present in the room.

Figure 5 presents the formation process of the DLIP using the real-time detection data of both occupancy activities and windows. As shown in the snapshots of the recorded frame indicating the detection and recognition made, along with the percentage of prediction accuracy, a separate profile was generated for each category. Similar to [55], a count-based profile was generated for the occupancy activities, which were then used to estimate the heat emission rates of occupants performing different activities [55].

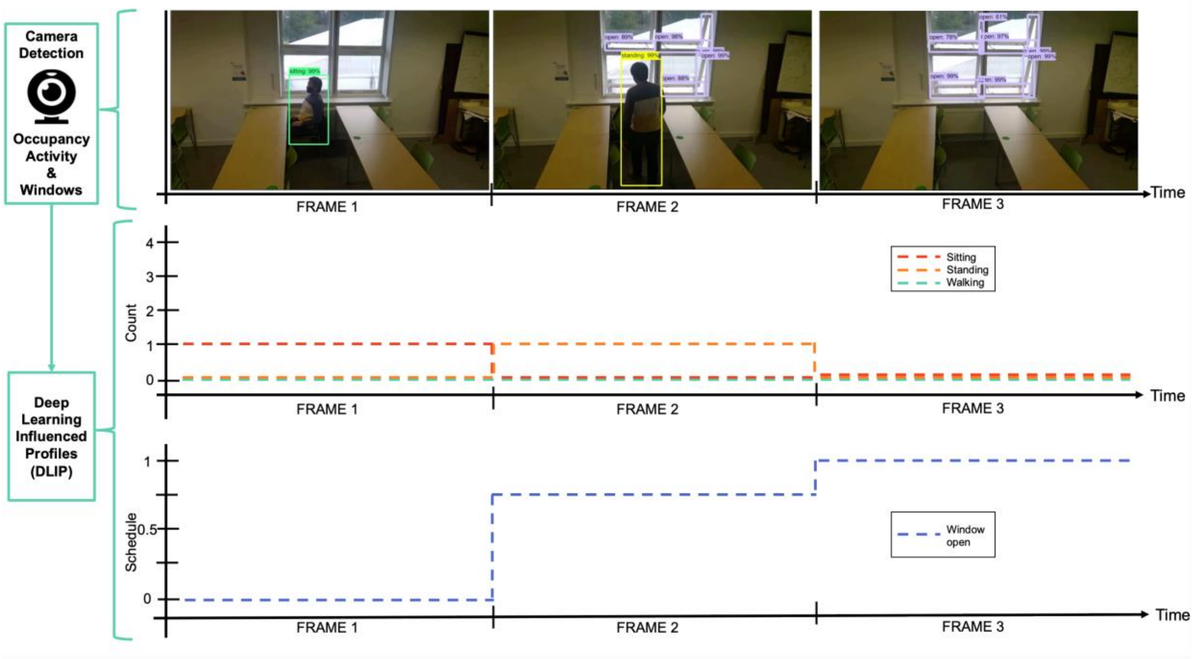



Figure 5. Real-time detection and formation of the deep learning influenced profiles (DLIP) for occupancy activities and windows.

A modulating profile was generated for windows, which correspond to the total number of open windows detected. As shown in Figure 5, due to the method of labelling of the opened windows, it resulted in instances when two overlapping bounding boxes were assigned to one window. Hence, a rule was set to ensure that a single-window opening would only be detected once.

### 2.4. Building Energy Simulation (BES)

As detailed in Figure 1, the case study building described in Section 2.1 was modelled, and BES was performed to evaluate the effect on the building energy demand of the proposed approach. The following section presents the description and setup of scenario-based simulation cases.

#### 2.4.1. Test Scenarios and Simulation Cases

Figure 6 shows the selected lecture room in the Marmont Centre and the activity schedule during a typical four-day period (Friday to Monday) between Friday 10$^{th}$ and Monday 13$^{th}$ January. It was assumed that a timetabled lecture was held on day 1 (Friday between 14:00 – 16:00) with full occupancy (up to 40 people), and another session was held on day 4 (Monday between 10:00 – 1200) with only half the number of occupants present (up to 20 people). The room was unoccupied during the other periods.

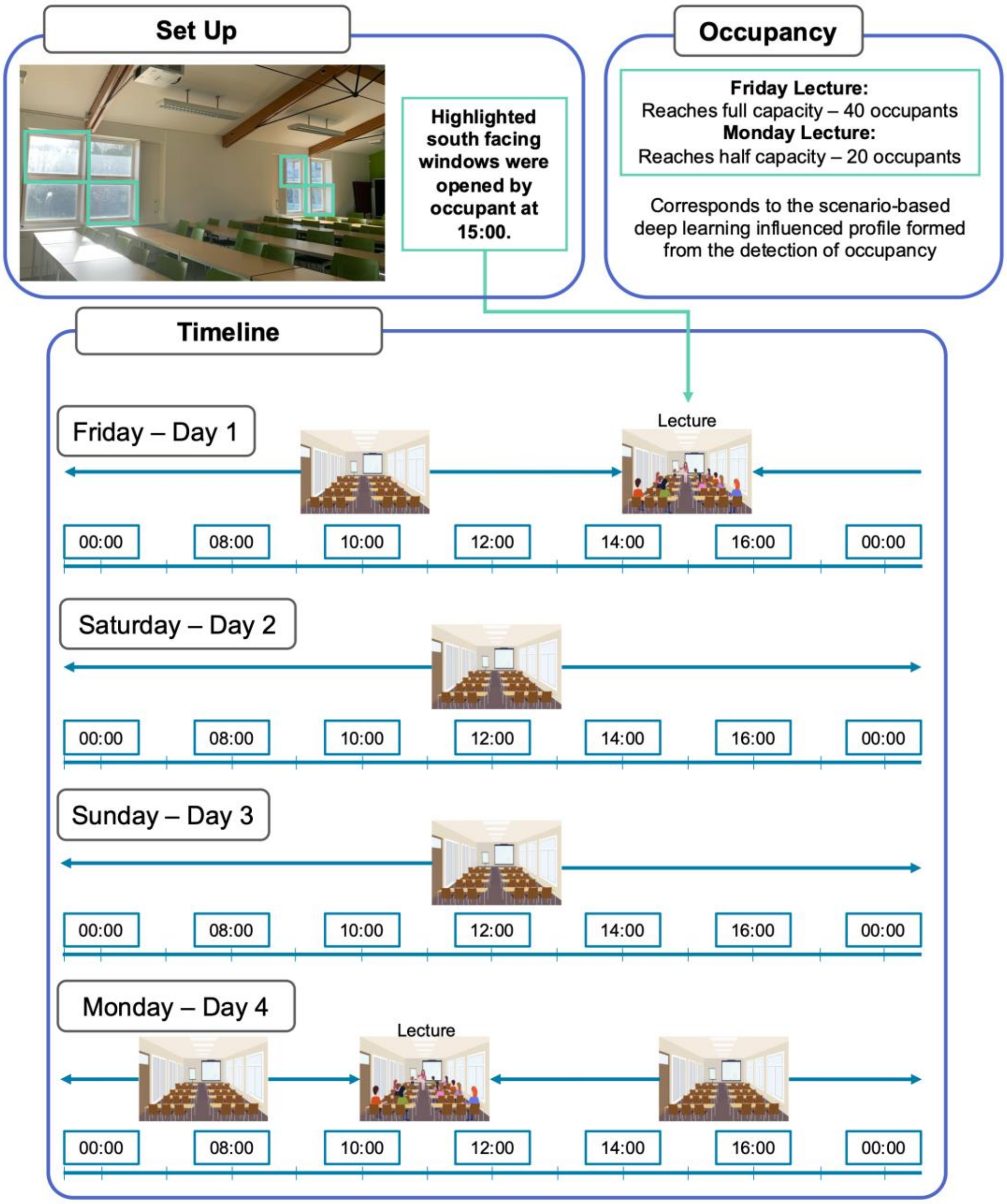



Figure 6. Timeline of activities in the selected lecture room during a typical week.

A total of five different cases were presented as described in Figure 7. This enabled the analysis of the different system responses that could assist the HVAC control system provide adequate indoor thermal comfort and air quality while improving the building energy performance. As mentioned previously, many HVAC systems are operated based on predefined or fixed schedules that are usually based on the building operational hours and recommended setpoint temperature [3,4]. Hence, for the present study, heating setpoint temperatures of 21°C and 15°C for occupied and unoccupied hours were set. Additionally, for most buildings, the number of occupants within the building and the window conditions are usually not known. Hence, to represent such situations, Case A was created with 6 different combinations of using constant "static" profiles for occupancy, windows, heating and cooling. For these cases, occupants were either; 1. Assumed to be not present in the room, 2. Present in the room and mainly performing sedentary activities during building operational hours and 3. Present in the room and mainly performing high-intensity activities to represent the maximum occupancy conditions.

Additionally, it was assumed that windows were either constantly opened or closed during occupied hours.

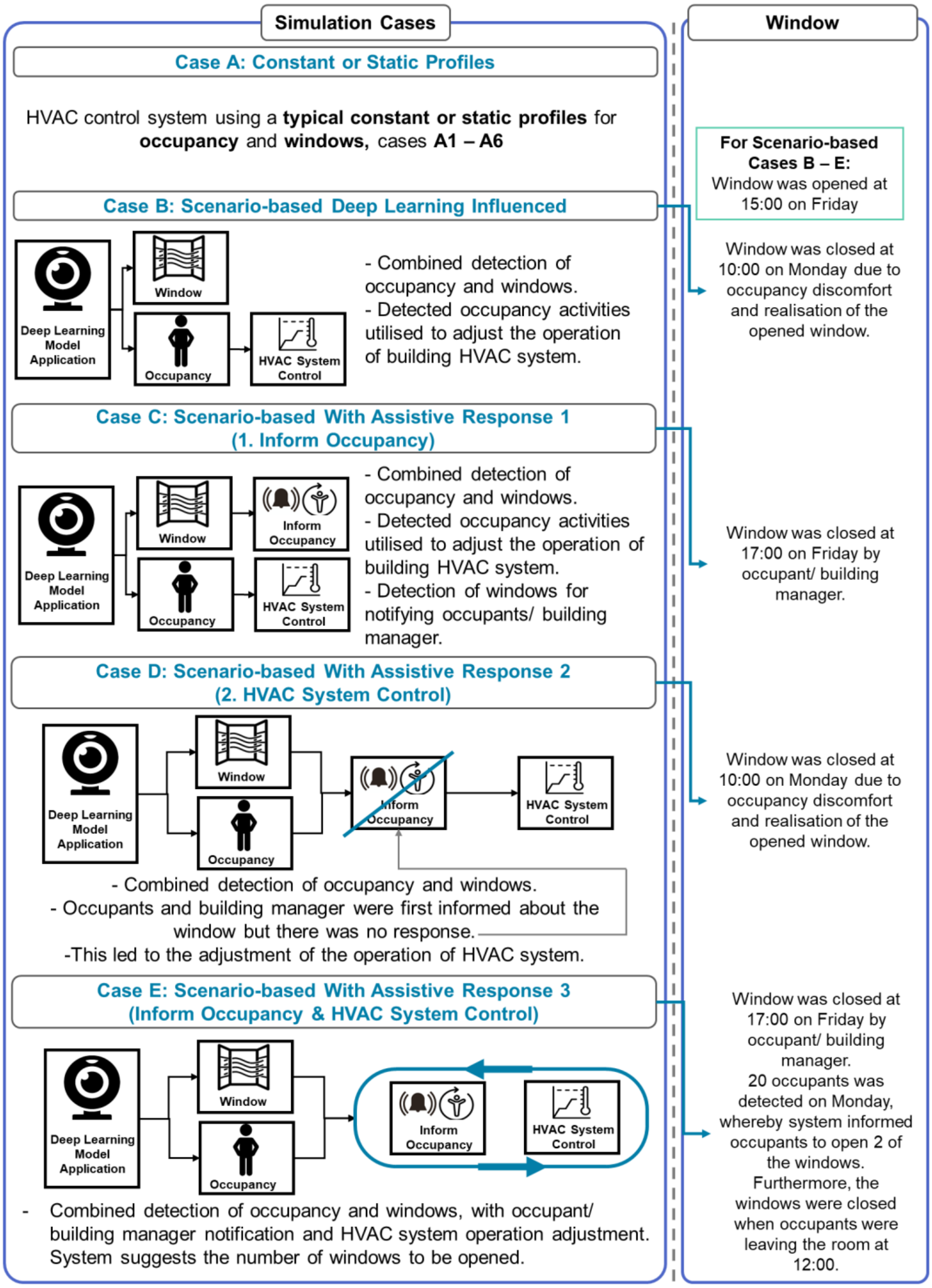



Figure 7. Description of the different simulation cases based on the different system responses.

The vision-based detection approach was implemented in scenario cases B, C, D and E. For the occupancy, the Scenario-based Deep Learning Influenced Occupancy Profile (Figure 8b) was set, and Figure 8c and d present a more detailed version of the profile with the number of occupants detected and the activities performed on both lecture days. Furthermore, only the south-facing windows were assumed to be left open by occupants during the lecture session at 15:00 on Friday (Day 1).

Case B represented the situation when both occupant's activities and window conditions were detected and recognised using the integrated vision-based approach. The detection of occupancy activities aided the adjustments of the operations of the building HVAC. The heating setpoint temperature of 21°C was only set when occupants were detected in the lecture room. The windows were detected to be opened at around 15:00 on Day 1. However, no response-based adjustments were made. Hence, the windows were left open until 10:00 am on Monday, when a person who attended the session decided to close these windows.

Similar to Case B, both occupant's activities and window conditions were detected in Case C, and the building HVAC operation was adjusted based on the occupancy level. In addition, the building users were informed about the window condition. For this, the windows left open after the lecture on Day 1 (Friday) was detected, and the building manager was informed and closed the windows at 17:00 (1 hour after the end of the lecture).

In Case D, the building users or manager did not respond to the notification made, and the building HVAC controls made a direct response by switching off the heating system when no people and opened windows were detected. This is indicated by the heating profile shown in Figure 10c.

In Case E, the system response was further improved. In addition to the adjustment of the HVAC operation and informing building users about window conditions, the approach suggests the number of windows that should be opened depending on the number of occupants within the room and the indoor room temperature.

Hence, for this case, the same scenario-based conditions as for Case C were assumed for occupants, windows, heating, and cooling windows during day 1 (Friday). However, on Monday (day 4), only half of the occupants were present within the room. The system suggested that only a certain number of windows be opened and was later closed just before all the occupants left the room at 12:00.

For all cases, the corresponding profiles are highlighted in Figure 8, Figure 9 and Figure 10. A summary of the scenario cases is detailed in Table 2.

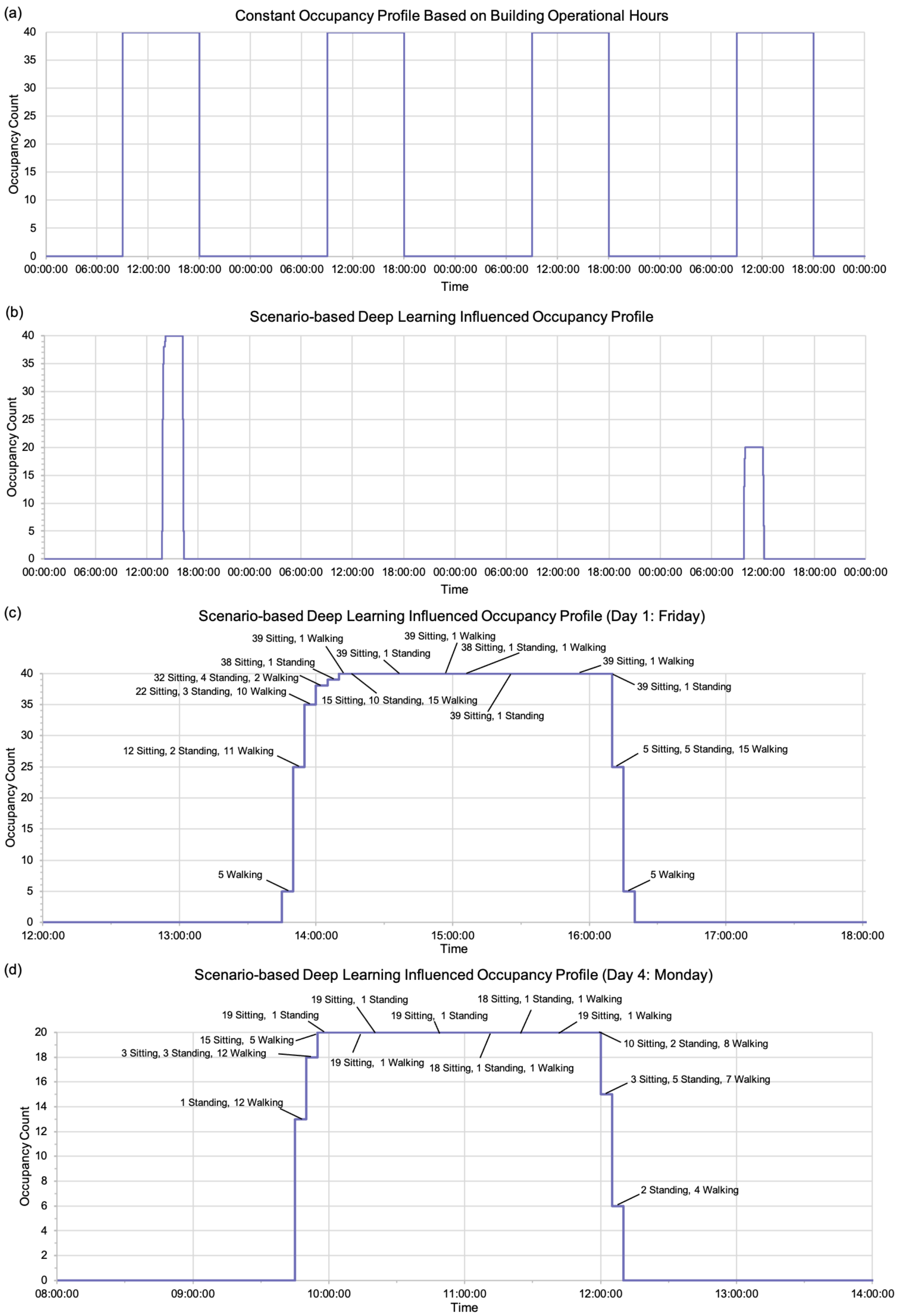



Figure 8. Occupancy profiles used within BES. (a). Typical, constant static occupancy profile that is based on building operational hours. (b) Scenario-based deep learning influenced occupancy profile that corresponds to the timeline given in Figure 7. (c) and (d) A more detailed view of (b), with the description of the occupancy behaviour during day 1 (Friday) and day 4 (Monday).

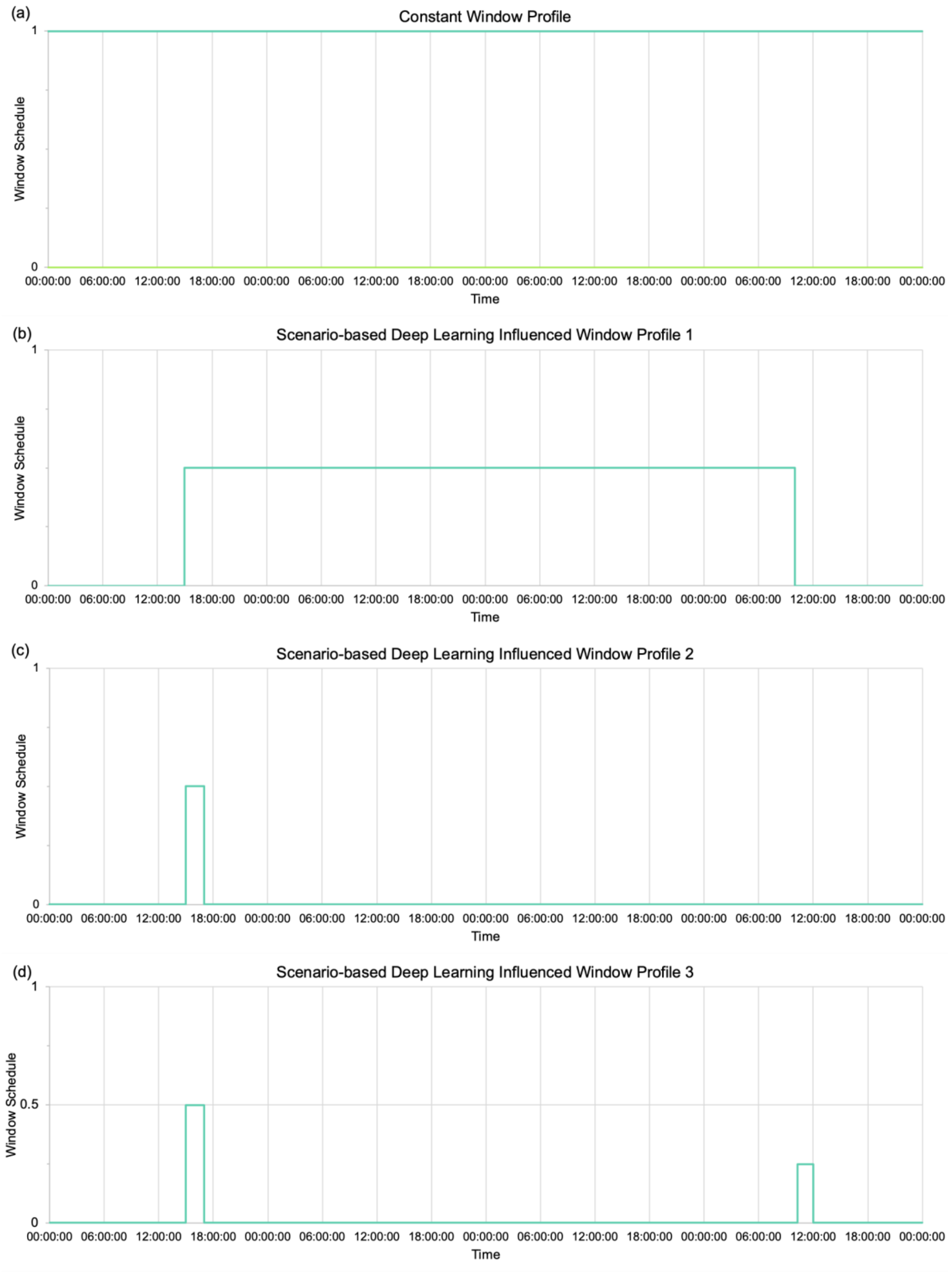



Figure 9. Window profiles used within BES. (a). Typical, constant open and closed window profiles. (b), (c) and (d). Scenario-based deep learning influenced occupancy profile that corresponds to cases highlighted in Figure 8 and Table 2.

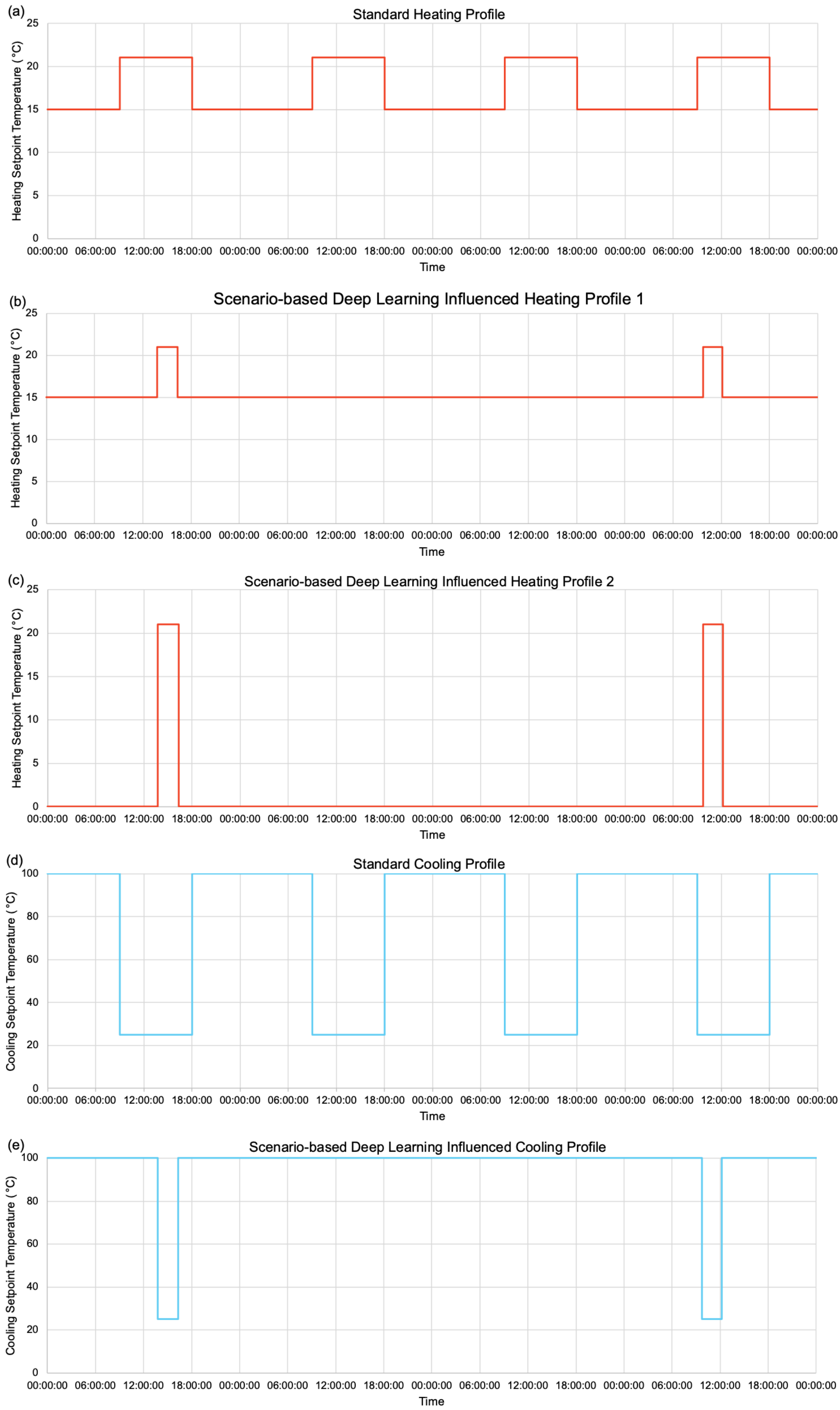



Figure 10. Heating and cooling profiles used in the scenario cases, highlighted in Table 2.

Table 2. Summary of the profiles assigned to the scenario cases.

| Simulation Case | | Assigned Profiles: Occupancy | Window | Heating | Cooling |
|---|---|---|---|---|---|
| A: Typical | A1 | None | Constant closed (Figure 9a) | Standard (Figure 10a) | Standard (Figure 10d) |
| | A2 | | Constant open (Figure 9a) | | |
| | A3 | Constant low activity level during building operational hours (Figure 8a) | Constant closed (Figure 9a) | | |
| | A4 | | Constant open (Figure 9a) | | |
| | A5 | Constant high activity level during building operational hours (Figure 8a) | Constant closed (Figure 9a) | | |
| | A6 | | Constant open (Figure 9a) | | |
| B | | Scenario-based DL Influenced (Figure 8b, c & d) | Scenario-based DL Influenced 1 (Figure 9b) | Scenario-based DL Influenced 1 (Figure 10b) | Scenario-based DL Influenced (Figure 10e) |
| C | | | Scenario-based DL Influenced 2 (Figure 9c) | | |
| D | | | Scenario-based DL Influenced 1 (Figure 9b) | Scenario-based DL Influenced 2 (Figure 10c) | |
| E | | | Scenario-based DL Influenced 3 (Figure 9d) | Scenario-based DL Influenced 1 (Figure 10b) | |

## 3. Results and Discussion

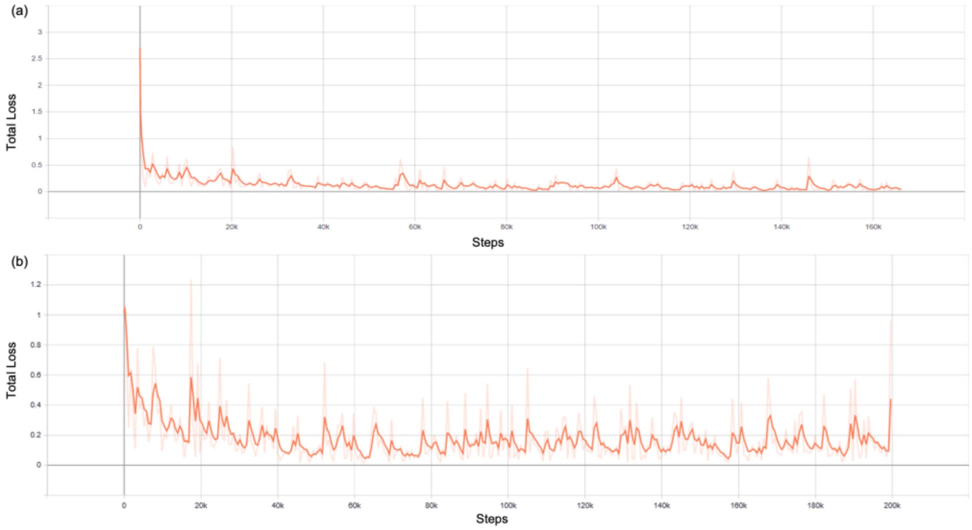
The following sections present the results and discussion of the model detection performance and evaluation of the impact on the building energy performances of the proposed approach.

### 3.1. Model Training Performance and Evaluation

Table 3 presents the training results for the two models, Model A for the detection and recognition of occupancy activity and Model B for window conditions. The converged total loss values across the number of training steps imply that both models were adequately trained.

Table 3. Vision-based occupancy activity and window detector training results.

| Training Conditions and Results | (a) Occupancy Activity Model | (b) Window Model |
|---|---|---|
| Model Used | Faster RCNN with InceptionV2 | |
| Total Steps | 102,194 | 199,630 |
| Training Duration | 10 hours, 29 minutes, 52 seconds | 11 hours, 29 minutes, 46 seconds |

| | | |
|---|---|---|
| Average Loss | 0.13436 | 0.181088 |
| Minimum Loss | 0.005654 | 0.01519 |
| Total loss versus the number of training steps |  | |



Before deploying both models to an AI-based detector, a model performance evaluation was conducted based on the detection and recognition ability on still images from the test dataset (Table 1). The images were similar but different to the ones used for training the models. Table 4 provides the corresponding results in the form of confusion matrix and the results in terms of the common classification evaluation metrics.

Both models suggest the prediction labels were correctly assigned with an average accuracy of 94.04% for sitting, 91.43% for standing, 92.70% for walking, and 87.74% for opened windows. As shown by the confusion matrix, slightly lower accuracy was achieved for standing, which could be due to the similarity in body form and shape of an occupant performing the walking activity. Despite this, the results indicated that both models are adequate for the initial deployment and testing.

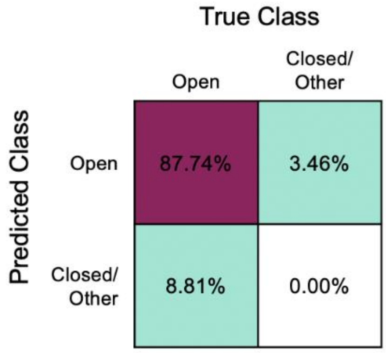
Table 4. Detection performance on still images from the test dataset.

**Confusion Matrix**

**(a) Occupancy Activity**



**(b) Window**





| Classification | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| | | **(a) Occupancy Activity** | | |
| Sitting | 94.04% | 0.9250 | 0.8911 | 0.9077 |
| Standing | 91.43% | 0.9064 | 0.8284 | 0.8657 |
| Walking | 92.70% | 0.8643 | 0.9266 | 0.9047 |
| | | **(b) Window** | | |
| Open | 87.74% | 0.9621 | 0.9088 | 0.9347 |

### 3.2. Vision-based Detection and the Formation of the Deep Learning Influenced Profiles (DLIPs)

Figure 11 and Video 1 presents a preview of the real-time detection and recognition using the integrated vision-based detection approach within the case study lecture room. Figure 11 presented a series of results for each of the key stages of the test, Part 1 – Part 5. The results showed the capabilities of the approach to provide a combined detection of occupancy activities and window conditions.

During the real-time detection, for each of the instances, bounding boxes were assigned across the desired object that was recognised, and the detection accuracy based on the IoU value was presented. Since this is a vision-based approach that requires a camera to perform the following detections, limitations in terms of obstruction can ultimately affect the performance of such an approach. Furthermore, the labelling of the images of the gaps of opened windows in the training dataset resulted in instances of windows achieving two overlapping bounding boxes assigned to one window. This was shown in Figure 11, with overlapping horizontal and vertical bounding boxes.


Figure 11. Key stages of the occupancy activity and window detection during the experimental test.

In practice, the images or videos of the detection (for example, Figure 11 and Video 1) are not stored and are only shown here to demonstrate how it works. Figure 12 presents the generated DLIP for occupancy activities and window conditions during the experimental test. The profile in Figure 12a details the activities performed by the detected occupant, which was utilised to predict the internal heat gains from occupancy and form a heat emission profile, as shown in Figure 12b.

Figure 12b compares the DLIP against other profiles, including two static scheduled profiles assuming fixed occupancy rates and the Actual Observation profile, representing the ground truth or actual activities performed by the occupant during the test. Figure 12c presents the comparison between the predicted and actual conditions (open or close) of the windows. The results show that there were some errors when comparing the DLIP against the actual conditions. The error was 8.20% for occupancy activity detection and 20.98% for window detection. This suggests that further development is necessary to improve the detection performance.

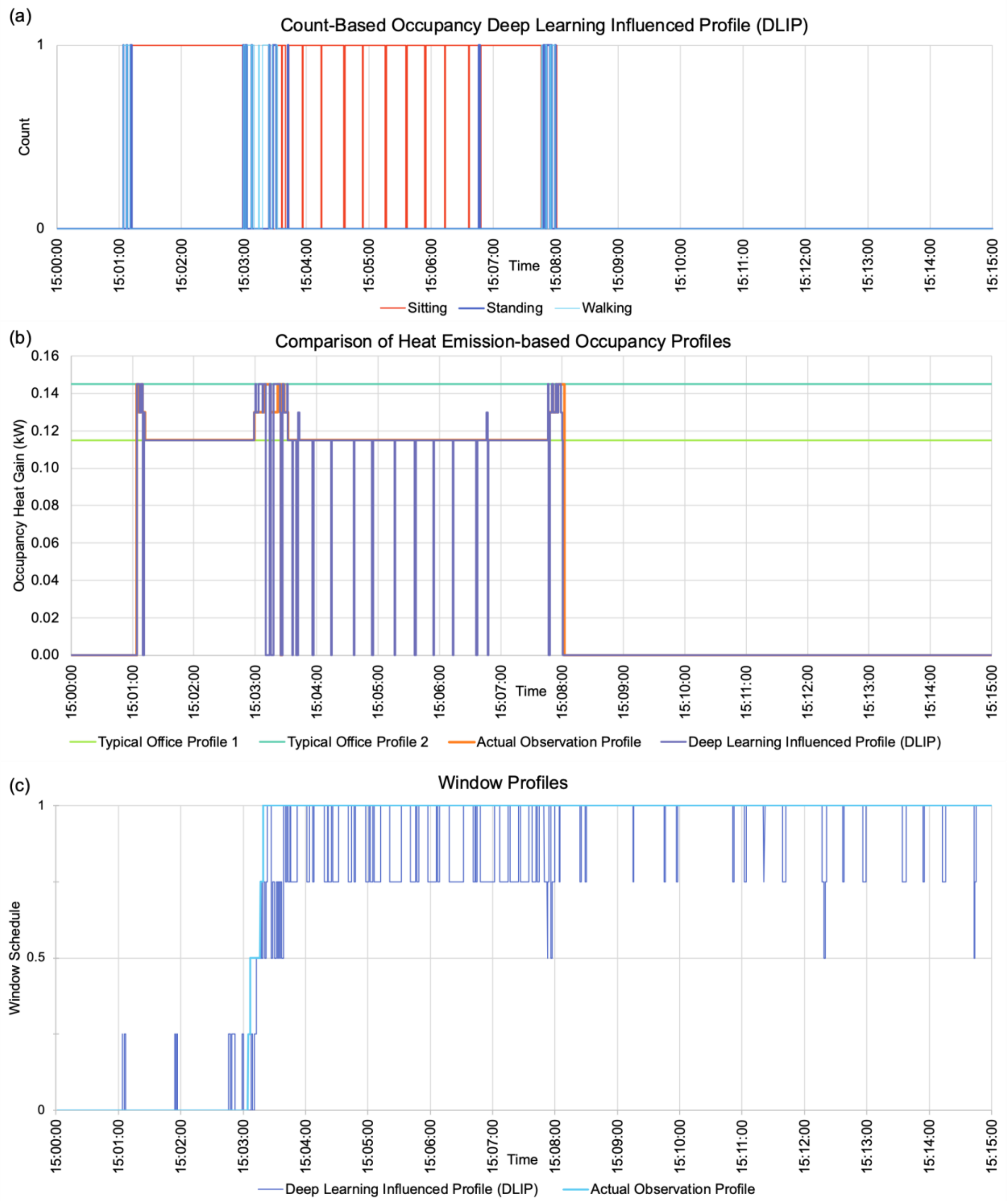



Figure 12. Generated (a) count-based occupancy deep learning influenced profile (DLIP) during the experimental test. (b) Comparison between heat emissions DLIP and the static scheduled and the actual observation profiles. (c) Generated DLIP for windows during the experimental test plotted against the Actual Observation Profile.

### 3.3. Experimental Test: Detection Performance Analysis

Based on the three model performance evaluation methods detailed in Section 2.3.1, the following sections presents the analysis of the detection performance of the vision-based approach during the experimental test. Section 3.3.1 provides the performance based on the whole, full duration of the experimental test. Since only one person was present during the experimental test, the following analysis will be based on the detection of each activity. Furthermore, since the south-facing windows 1 consisted of 4 windows, the following analysis denoted the window from the top left, top right, bottom left and bottom right as Windows A – D.

### 3.3.1. Experimental Test Detection Results

Figure 13 presents the average IoU values obtained from the detections made for each window and activity performed by the occupant over time. Overall, a high IoU was achieved for the occupancy activity detection, with the activity of sitting achieving 97.67%, 96.17% for standing and 97.00% for walking, giving an overall IoU of 96.95% for all activities.

Although the experimental test consisted of only one occupant, the person performed all the activities required to be detected. Hence, the results can verify the model's capability to recognise the differences between the corresponding human poses for each specific activity. Due to the variation in the images of occupants within the training dataset, the performance should not be impacted by the appearance of the occupants and specific features such as the clothing that the person was wearing.

The IoU for the detection of windows A, B, C and D were 84.53%, 95.67%, 94.56% and 98.92%, which gave an overall IoU value of 93.42%. Many factors could have influenced Window A to achieve a lower accuracy as compared to the other windows. The angle of the window opening gap in relation to the camera, the influence of the indoor-outdoor lighting conditions, and the outside view could have impacted the detection of the window opening gap.

Furthermore, the results achieved for Window C suggests that the detection approach still managed to recognise the open windows when an occupant is obstructing the window. Hence, this further suggests that the method of labelling the window opening gaps was effective. Overall, the results suggest that the integrated model could detect multiple and different objects in the indoor space.

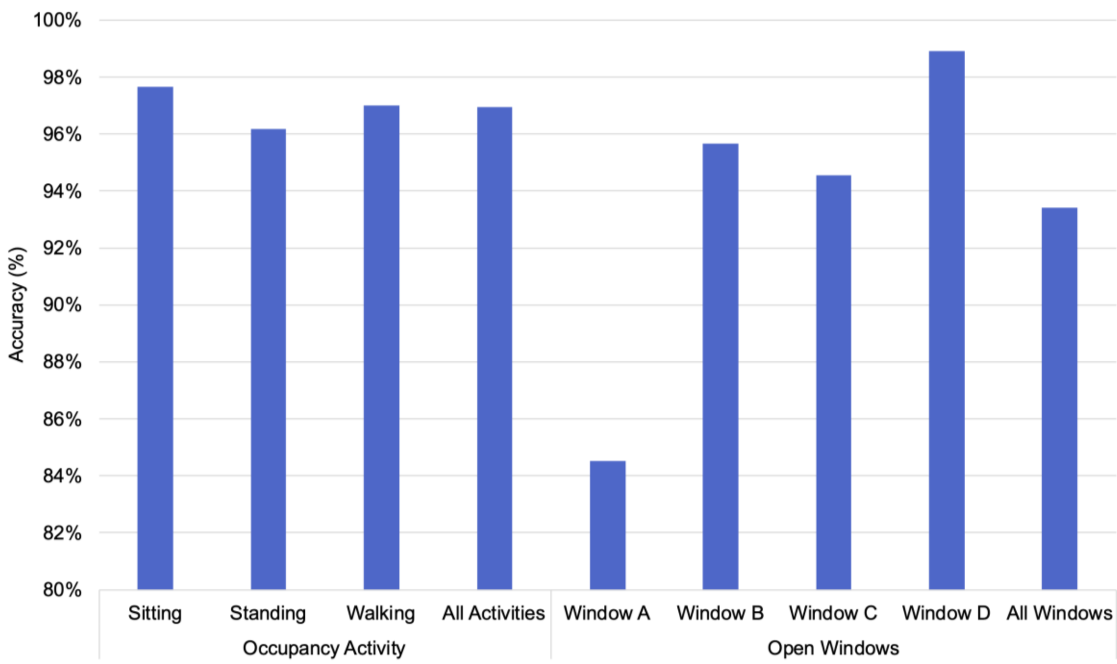



Figure 13. Average IoU for the detection of occupancy activity and open windows during the experimental test.

Figure 14 presents the detection performance in terms of the model providing correct, incorrect, and no/missed detections during the test. Correct detection was achieved when the activity performed by the person was correctly identified and when detection was not made when that activity was not performed. Similarly, this was also achieved when the open windows were correctly identified as open and for the times when detection was not made as the windows were closed.

Based on the detection for all three occupancy activities (Figure 14a), the best results were achieved for sitting as this action was different to the standing and walking activities. Hence, sitting achieved the highest correct detections with up to 91.99% compared to 46.88% for standing and 48.00% for walking. Due to the similarities in the human pose when standing and walking, the results showed a higher number of no or missed detections, 40.63% and 20.00%, and incorrect detections of 12.50% and 32.00%.

Figure 14b shows the results of the window detection. The presence of the occupant near the window resulted in missed detections (up to 16.54% of the time), with a lower amount of correct detections (83.35% of the time) compared to the other windows (Windows A, B and D) achieving correct detections for up to 93.67%, 99.22% and 99.11% of the time.

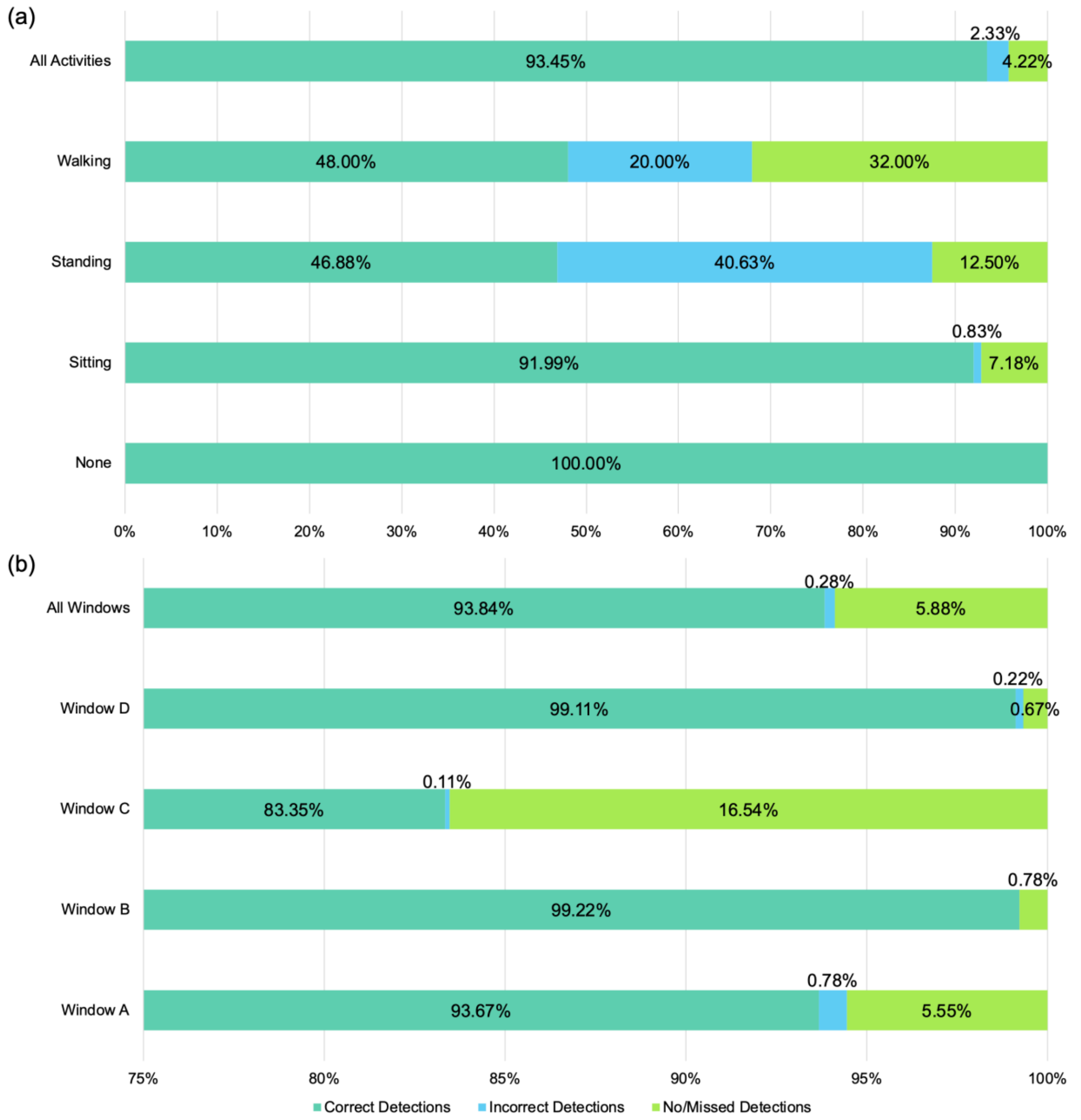


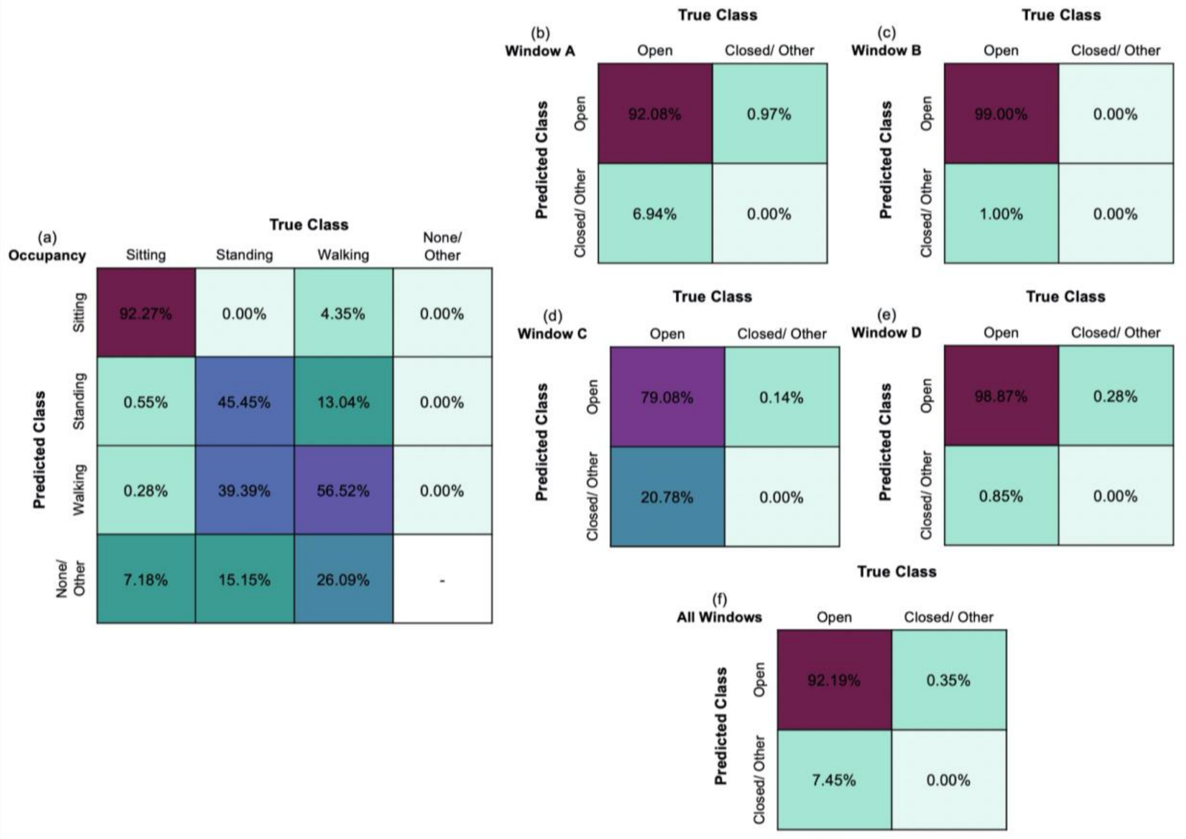
Figure 14. Detection performance in terms of the correct, incorrect, and no/missed detections during the experimental test.

Further evaluation of the detection performance during the experimental test is presented in Table 5. The results are presented in the form of a confusion matrix. Similarly, the confusion matrix for occupancy activity detection showed that the best performance was achieved for the sitting activity, and lower accuracy was achieved for standing and walking activities. The results suggest that the activity of standing can be confused by the model as walking. However, it was less likely for the activity of walking to be identified as standing. Additionally, the standing activity achieved the lowest recall value and $F_1$ scores.

For the window detection,Window C detection performance was slightly affected by the occupant's presence. Overall, good performance was achieved, with the percentage of true positives (opened windows correctly ) reaching an average of 92.19%, and only occasional instances when the opened windows were not identified, with no labels assigned. Also, times when open labels were assigned to windows that were actually closed.

Table 5. Detection performance results presented in a confusion matrix and based on common classification evaluation metrics.

**Confusion Matrix**




| Class | Activity | Accuracy | Precision | Recall | $F_1$ Score |
|---|---|---|---|---|---|
| **Occupancy Activity** | | | | | |
| 1 | Sitting | 95.97% | 0.9550 | 0.9226 | 0.9385 |
| 2 | Standing | 77.38% | 0.7734 | 0.4545 | 0.5725 |
| 3 | Walking | 72.28% | 0.5875 | 0.5652 | 0.5761 |
| Average for all activities | | 85.63% | 81.88% | 0.7720 | 0.6474 |
| **Windows** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| | Window A | 92.09% | 0.9896 | 0.9299 | 0.9588 |
| | Window B | 99.00% | 1.0000 | 0.9900 | 0.9950 |
| Open | Window C | 79.08% | 0.9982 | 0.7919 | 0.8832 |
| | Window D | 98.87% | 0.9972 | 0.9915 | 0.9943 |
| | All windows | 92.20% | 0.9962 | 0.9252 | 0.9594 |

The detection performance achieved during the experimental test showed that training the model separately and combining and deploying it to a single detector was a feasible approach. Furthermore, The results also suggested that the detection of one object can subsequently impact the detection of another object. This was observed when detecting the occupant in front of the opened window.

## 3.4. Building Performance Analysis

The following section presents the BES results for the scenario cases detailed in Table 2.

### 3.4.1. Occupancy Heat Gains

Based on the scenario cases detailed in Figure 6 and the generated DLIP as shown in Figure 8b, Figure presents the predicted occupancy sensible and latent heat gains. Results for Typical Office 1 and 2 provided benchmark values to represent static occupancy profiles employed in conventional control systems. Typical Office 1 assumed occupants to be carrying out sedentary activities, while Typical Office 2 assumed that the occupants were more active. For both cases, overall heat gains of 165.6kWh and 208.8kWh were predicted. Figure 5a presents the distribution of heat gains across the simulated period. With the lecture room unoccupied for most of the time and only a small number of occupants were present for a few hours, hence a lower total occupancy heat gain (16.6kWh) was predicted using the DLIP. This shows that if the HVAC was assumed to be operated using static occupancy profiles, it could lead to a significant overestimation of the indoor heat gains. This shows the importance of employing strategies which can  recognise whether a room is occupied or unoccupied, along with the knowledge of the type of activities performed by occupants at a given time.

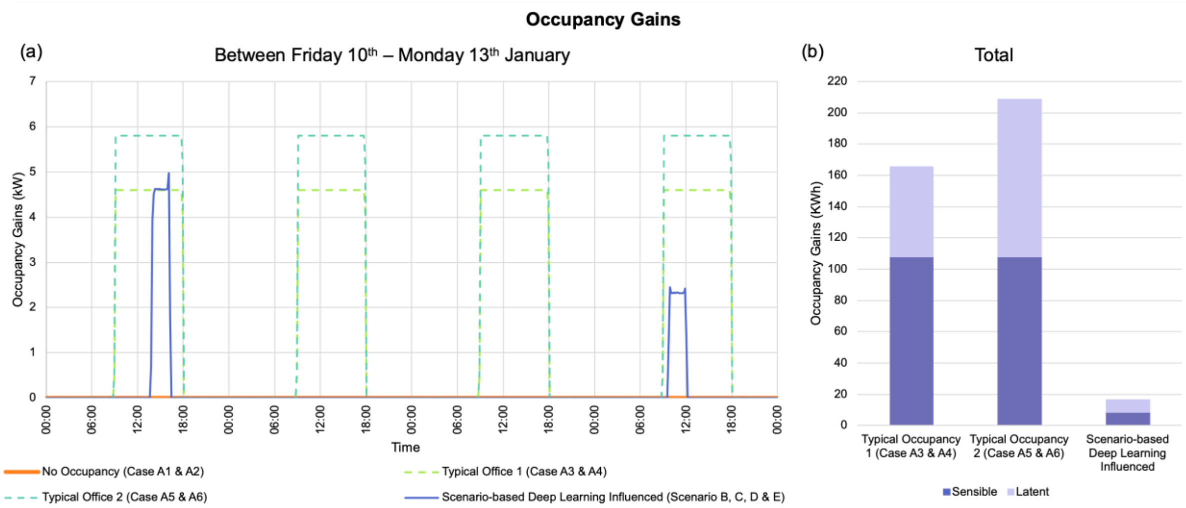



Figure 15. Comparison of the (a) occupancy heat gains across time and (b) the predicted total sensible and latent occupancy heat gains based on the scheduled profiles and scenario-based DLIP.

### 3.4.2. Impact on Ventilation Heat Loss

Figure a presents the total ventilation heat loss, and Figure b shows the distribution of the ventilation heat loss for all cases during the 4-day period. The ventilation heat loss was influenced by the indoor-outdoor conditions and the number of opened windows. The constant open and closed results show the maximum and minimum heat loss. The results for Case B, C, D and E, were directly influenced by the window profiles given in **Error! Reference source not found.**9, which shows the advantage of knowing whether windows are either opened or closed, as it can significantly affect the conditions within an indoor environment.

Case B and D had higher ventilation heat losses (90kWh and 77.9kWh) as the windows were left open after the lecture on Friday (day 1) afternoon. Using the proposed approach, the open windows were detected, and the building manager was alerted and managed to close the windows, which led to the lowest ventilation heat loss (6.8kWh) during the 4-day period. A slightly higher ventilation heat loss (10.4kWh) was predicted for Case E as the windows were suggested to be opened by the system during the lecture on Day 4 to improve the indoor air quality. Although the windows were left open in Case D, unnecessary energy demand can be minimised by adjusting the setpoints or turning off the heating system after the system detected no response.

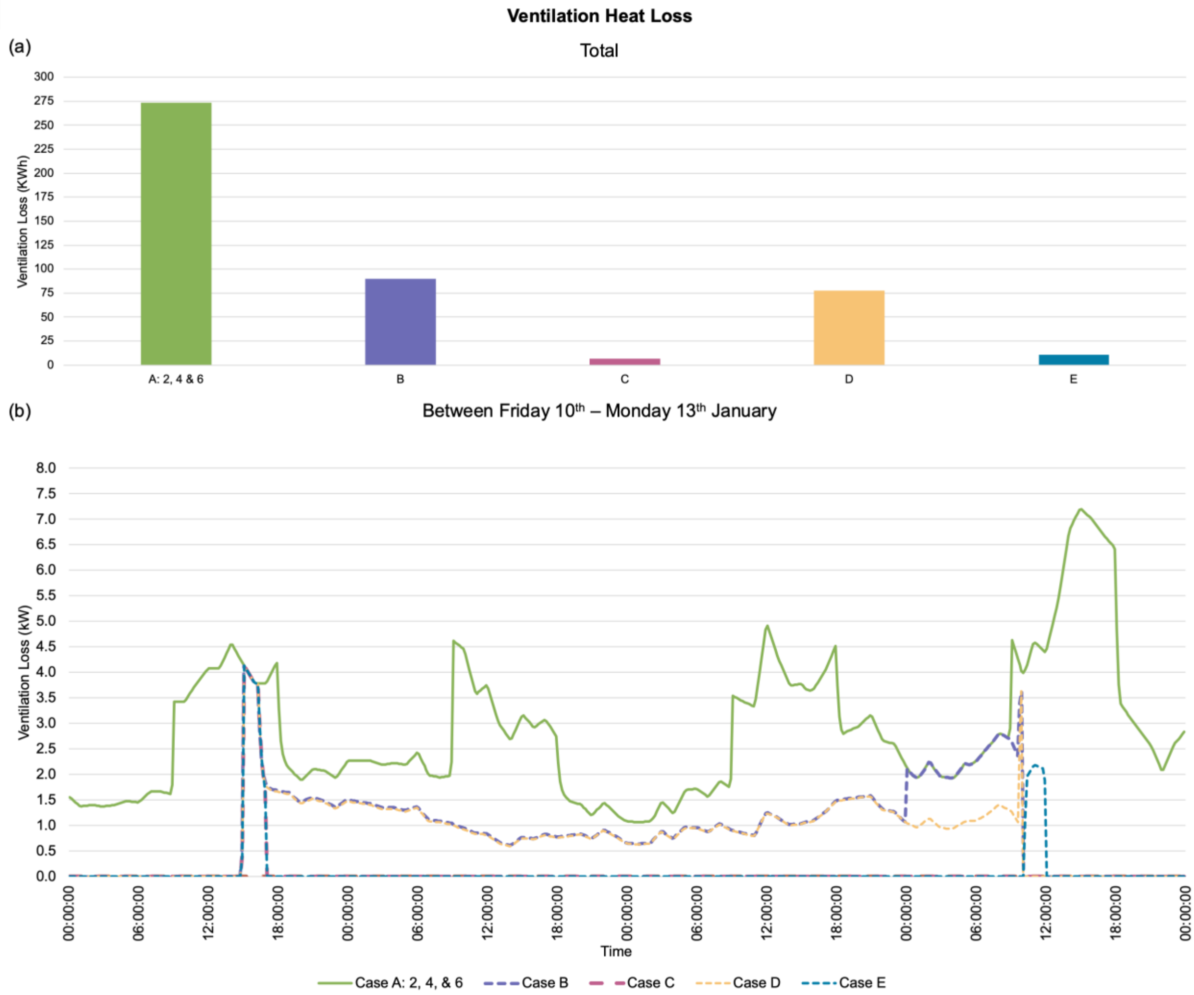



Figure 16. Total building ventilation heat loss prediction for all simulated cases (Case A, B, C, D and E), with (a) and (c) presenting total heat losses for all cases under the 4-day scenario. (b) and (d) indicates the variation in heat losses across time.

#### 3.4.1. Impact on Heating Energy Demand

The results of the heating energy demand based on the different scenario cases are presented in **Error! Reference source not found.**. **Error! Reference source not found.**a and b present the predicted results for Cases A1 – A6, which employed fixed scheduled profiles for occupancy and windows. For all cases, the building HVAC system was operated based on an indoor setpoint temperature of 21°C during the operational hours.

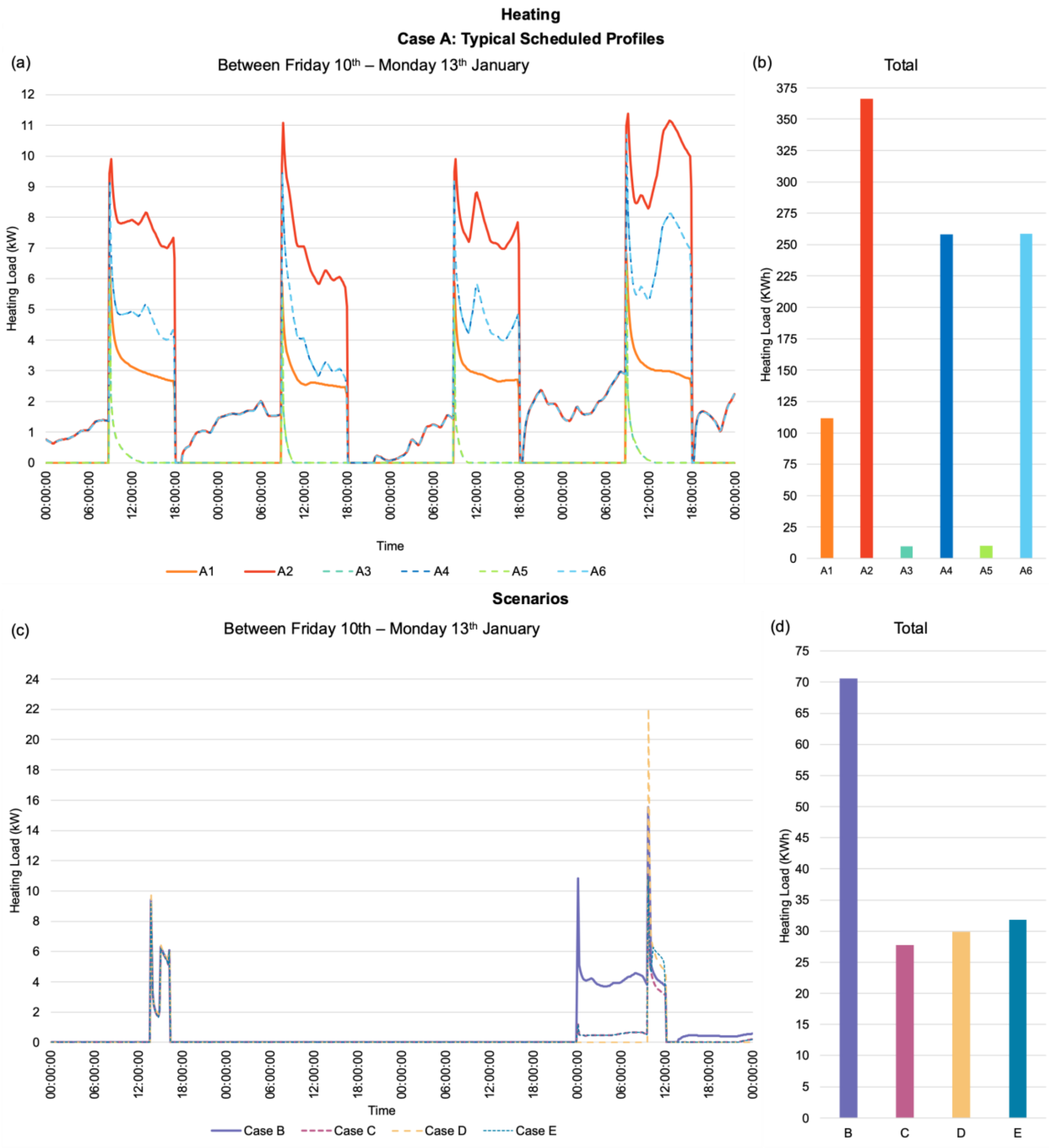



Figure 17. Total building heating load prediction for Case A and the different scenario-based cases (Case B, C, D and E), with (a) and (c) presenting the variation in loads across time. (b) and (d) shows the total heating load for all cases under the 4-day scenario.

From the constant open and constant closed results, it generally shows the maximum and minimum possible heating demand. When the windows were constantly closed, the number of occupants present within the room influenced the internal occupancy heat gains, affecting the heating requirement.

However, the results show that its impact is not as significant as the ventilation heat loss from the windows. Figure 17c and d present the results for Cases B, C, D, and E.

For Case B, the opened windows at 15:00 on day 1 were detected, but no adjustments to the building HVAC operations were made. As a result, on Monday (day 4), the room temperature reached below 15°C (the set temperature during unoccupied hours); hence constant heating occurred during Monday morning, and more heating occurred when the occupants were detected within the room on Monday (day 4) at 09:45. Case C employed the detection approach, which provided notifications or alerts to the building users or manager. For this case, heating was not required until Monday (day 4), when occupants were present in the room for the lecture. Hence a total heating load of 27.8kWh was predicted.

In Case D, no changes were made even after the system provided notifications, and the windows were left open from Friday night to Monday morning. When the system detected that the windows were left open for some time, the heating was turned off until Monday morning, when occupants were detected in the space.

In Case E, a balance between energy reduction and good indoor air quality was aimed to be provided; it informs the occupants and makes adjustments to the HVAC system operations to minimise unnecessary heating demand and maintain the indoor air quality. The heating load achieved across were identical to Case C. However, on Monday, when there is half occupancy in the room, it suggested that the occupant open two of the windows to ensure that good IAQ is maintained (see Figure 18), which then led to the increased heating energy demand of up to 31.8kWh.

#### 3.4.2. Impact on the Indoor Air Quality

The indoor air quality can be assessed in terms of the room $CO_2$ concentration levels. Generally, $CO_2$ levels in rooms below 1,000ppm were assumed to be fairly adequate, and anything above this level would indicate the room is highly polluted [65]. This can affect occupancy productivity and human health [66, 67]. **Error! Reference source not found.**b presents a comparison of the distribution of the $CO_2$ concentration across the 4-day scenario for the three selected cases. Although a lower ventilation heat loss was achieved for Case C during the lecture period in Day 4, it also led to very high $CO_2$ concentrations levels peaking at 2,288ppm when the occupants did not open the windows. In Case E, both the number of occupants and windows open were considered in the decision-making process of the control system. It assumed that 2 windows were opened during the lecture period as per the recommendation by the system. This resulted in the $CO_2$ concentration reducing from 2,288ppm to approximately 1,000ppm. Although the air quality is still poor, it can be further reduced by suggesting to the users to have more windows to be opened.

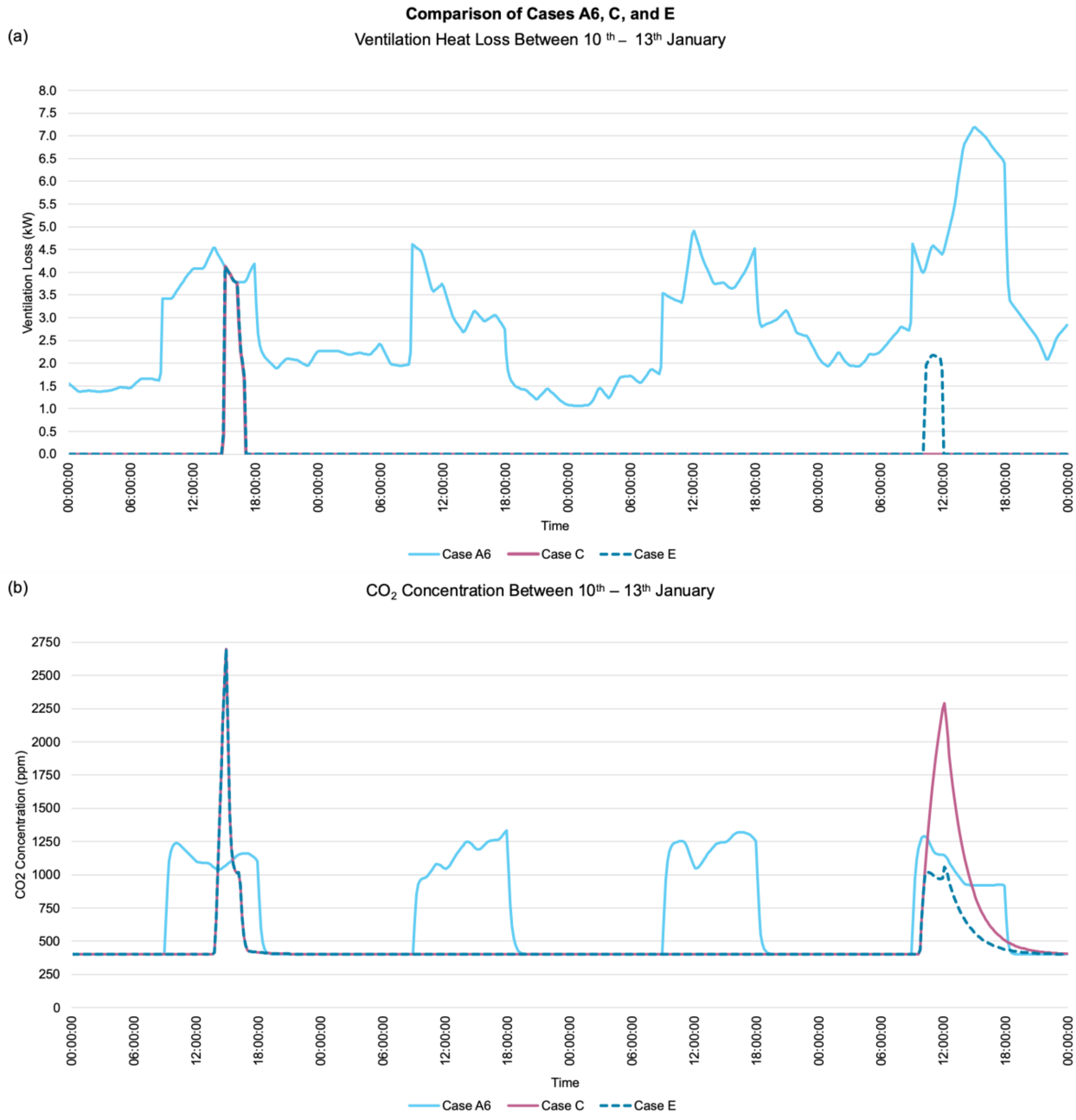



Figure 18. Variation in (a) ventilation heat losses and (b) $CO_2$ concentration across time during the 4-day period comparing Cases A6, C and E.

A framework and software infrastructure [68] should be developed for the proposed approach to be fully utilised within buildings. This system will connect the real-time vision-based detector and setpoint optimiser with the demand-driven controls for the HVAC system. This will ensure that the HVAC is operated according to the real-time building or space requirements. The potential of a co-simulation approach using BES [69, 70] should also be explored in future works. Using Case E as an example to demonstrate such a decision support system, this approach will combine the two vision-based approaches for detecting occupancy activity and windows to assist the alert and control system in determining the window and HVAC setpoint adjustments required to ensure adequate indoor air quality is achieved and minimise the unnecessary ventilation heat loss. (Figure 19) shows a simple example of the decision making process based on the application during the heating season. Depending on the selected building space, integrated with the vision-based occupancy and window detector, a set number of occupants will be defined to decide if the windows should be opened or closed and/or if adjustment to the HVAC setpoint is required. $N_{occupancy}$ represents the number of occupants detected using the vision-based approach, and $N_{set}$ represents a set number of occupants. It should be noted that the control flow process shown here is simplified and does not take into account the number of windows detected, which can be included in the decision-making process to maximise natural ventilation while minimising heat loss. Additional steps can be added to the control flow process, informing the building users about the optimum number of windows to open or close; this will be developed in future works. Furthermore, more scenario-based analysis is required before establishing the decision support system that could comply with different indoor spaces and buildings.

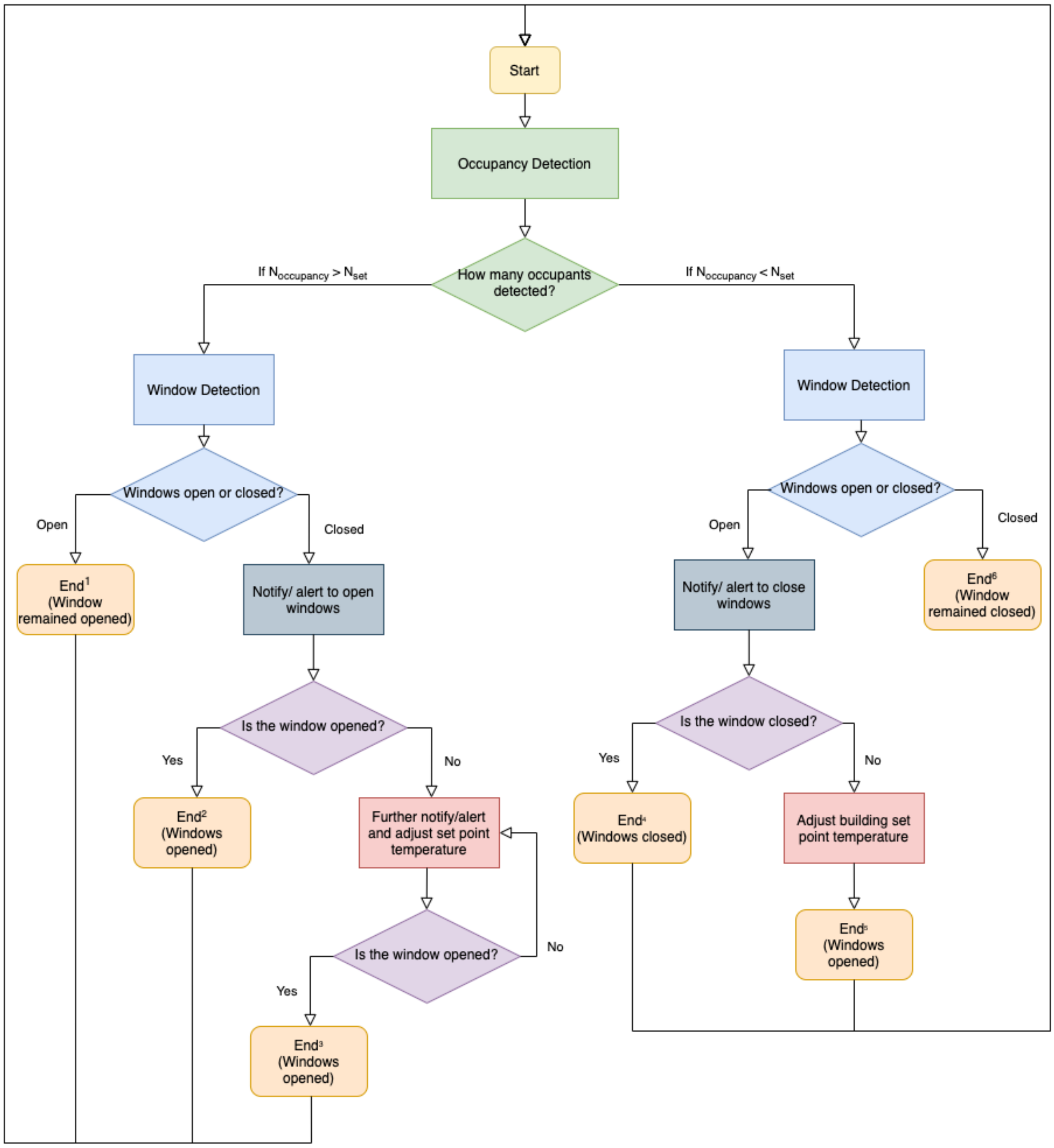



Figure 19. An example flow chart demonstrating the decision-making process of the proposed integrated control system.

## 4. Conclusions and Future Work

This present study proposes a data-driven deep learning framework for the detection and recognition of occupancy activity and windows. The data generated can be used to make real-time adjustments to the HVAC system operations and provide notifications to building users and managers to minimise unnecessary energy usage and effectively manage indoor air quality.

To enable the detection and recognition of occupancy behaviour, a Faster R-CNN model was employed and trained using an image-based dataset. The models were deployed to an AI-based camera. The proposed approach was evaluated based on the experimental test conducted within a lecture room at the University of Nottingham. Average IoU detection accuracy of 96.95% for occupancy activities and

93.42% for opened windows were achieved. During the detection, real-time data about the number of occupants performing each of the selected activities and windows open were generated and used to form the deep learning influenced profile (DLIP).

BES was conducted simulating various scenario-based cases to assess the impact of the proposed approach and to provide insights into how the proposed detection method can enable HVAC systems to adapt and respond to occupancy's dynamic changes. The case study building was modelled, and 5 different scenario-based cases were considered. The cases focused on the application of different response-based solutions. Results indicate that the proposed approach can reduce the under or overestimation of occupancy heat gains.

The combined vision-based deep learning approach enabled the real-time monitoring of the number of occupants, activity performed by the occupants and the number of opened/ closed windows. This led to the prediction of the room internal heat gains and the levels of the room $CO_2$ concentration. This enables the system to inform occupants to open/ close a specific number of windows and/ or to enable the demand-controlled heating system to provide the requirements when required.

The results presented here showed the potential of the proposed approach and could be used as a basis for its further development. To ensure that this framework can be fully implemented and integrated with building energy management systems, further work is necessary. This includes exploring other types of model configurations and evaluating the influence of the training data on its detection performance. A setpoint optimiser will be integrated into the framework to automatically adjust the HVAC operation according to the detection data and requirements of the space. Further testing is required in different types of spaces and scenarios, for example, in spaces with high occupancy and movement.

**Acknowledgements**

This work was supported by the Department of Architecture and Built Environment, University of Nottingham and the PhD studentship from EPSRC, Project References: 2100822 (EP/R513283/1).

**References**

1. Haniff, M.F., Selamat, H., Yusof, R., Buyamin, S., Ismail, F.S., Review of HVAC scheduling techniques for buildings towards energy-efficient and cost-effective operations. Renewable and Sustainable Energy Reviews, 27 (2013), 94-103.
2. ASHRAE, Standard 55 – Thermal environmental conditions for human occupancy, 2017.
3. ASHRAE, ANSI/ASHRAE/IES Standard 90.1-2019 -- Energy Standard for Buildings Except Low-Rise Residential Buildings, 2019.
4. CIBSE Chartered Institution of Building Services Engineers. Table 1.5 Recommended comfort criteria for specific applications. Environmental design: CIBSE Guide A. In. London: CIBSE; 2015.
5. Energy.gov. Quadrennial Technology Review 2015. An assessment of energy technologies and research opportunities. Chapter 5: Increasing Efficiency of Building Systems and Technologies. September 2015.
6. Li, W., Koo, C., Hong, T., Oh, J., Cha, S.H., Wang, S. A novel operation approach for the energy efficiency improvement of the HVAC system in office spaces through real-time big data analytics. Renewable and Sustainable Energy Reviews, 127 (2020), 109885.

7. US Energy information administration (EIA). Residential energy consumption survey (RECS). 2015 RECS Survey Data. https://www.eia.gov/consumption/residential/data/2015/ (Accessed May 2021).
8. Kontes, G., Giannakis, G.I., Sánchez, V., De Agustin-Camacho, P., Romero-Amorrortu, A., Panagiotidou, N, Rovas, D.V., Steiger, S., Mutschler, C., Gruen, G. Simulation-Based Evaluation and Optimisation of Control Strategies in Buildings. Energies, 11 (2018), 3376.
9. Paone, A., Bacher, J-P., The Impact of Building Occupant Behavior on Energy Efficiency and Methods to Influence It: A Review of the State of the Art. Energies, 11 (2018), 953.
10. Chen, S., Zhang, G., Xia, X., Chen, Y., Setunge, S., Shi, L. The impacts of occupant behavior on building energy consumption: A review. Sustainable Energy Technologies and Assessments, 45 (2021), 101212.
11. Esrafilian-Najafabadi, M., Haghighat, F. Occupancy-based HVAC control systems in buildings: A state-of-the-art review. Building and Environment, 197 (2021), 107810.
12. D'Oca, S., Pisello, A.L., De Simone, M., Barthelmes, V.M., Hong, T., Corgnati, S.P. Human-building interaction at work: Findings from an interdisciplinary cross-country survey in Italy. Building and Environment, 132 (2018), 147-159.
13. CIBSE Chartered Institution of Building Services Engineers. Top Tips – Ventilation in Buildings (2015) (webpage). CIBSE Top Tips Information Sheet 1: Ventilation in Buildings. 2015. https://www.cibse.org/knowledge/knowledge-items/detail?id=a0q20000006oamlAAA (Accessed May 2021).
14. Brager, G., Alspach, P., Nall, D. H. Natural vs. mechanical ventilation and cooling. UC Berkeley: Center for the Built Environment. 2011. https://escholarship.org/uc/item/0tp7v717 (Accessed May 2021).
15. Shen, W., Chuanqi, X., Song, P., Jiale, S., Yunmo, W., Xiaoyan, L., Tarek, H., Steven, F., Farid, F., Rory J., D.W. Pieter, Analysis of factors influencing the modelling of occupant window opening behaviour in an office building in Beijing, China, 2016, Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association, Hyderabad, India, Dec. 7-9, 2015.
16. Fabi, V., Andersen, R.V., Corgnati, S., Olesen, B.W., Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models. Building and Environment 58 (2012), 188-198.
17. Andersen, R., Fabi, V., Toftum, J., Corgnati, S.P., Olesen, B.W., Window opening behaviour modelled from measurements in Danish dwellings. Building and Environment, 69 (2013), 101-113.
18. Pan, S., Xiong, Y., Han, Y., Zhang, X., Xia, L., Wei, S., Wu, J., Han, M., A study on influential factors of occupant window-opening behavior in an office building in China. Building and Environment, 133 (2018), 41-50.
19. Oropeza-Perez, I., Østergaard, P.A. Potential of natural ventilation in temperate countries – A case study of Denmark. Applied Energy, 114 (2014), 520-530.
20. de Rubeis, T., Falasca, S., Curci, G., Paoletti, D., Ambrosini, D. Sensitivity of heating performance of an energy self-sufficient building to climate zone, climate change and HVAC system solutions. Sustainable Cities and Society, 61 (2020), 102300.
21. Kallio, J., Tervonen, J., Räsänen, P., Mäkynen, R., Koivusaari, J., Peltola, J. Forecasting office indoor CO2 concentration using machine learning with a one-year dataset. Building and Environment, 187 (2021), 107409.
22. Yun, J. and S. Lee. Human movement detection and identification using pyroelectric infrared sensors. Sensors, 14 (2014), 8057-8081.

23. Simma, K.C.J, Mammoli, A., Bogus, S.M. Real-time occupancy estimation using Wi-Fi network to optimise HVAC operation. Procedia Computer Science, 155 (2019), 495-502.
24. Li, S., Li N., Becerik-Gerber, B., Calis, G. RFID-based occupancy detection solution for optimising HVAC energy consumption. 2011 Proceedings of the 28th ISARC, Seoul, Korea, 587-592.
25. Khalilnejad, A., French, R.H., Abramson, A.R., Data-driven evaluation of HVAC operation and savings in commercial buildings. Applied Energy, 278 (2020), 115505.
26. Kim, D., Kim, D.D., Kim, T. Energy performance assessment of HVAC commissioning using long-term monitoring data: A case study of the newly built office building in South Korea. Energy and Buildings, 204 (2019), 109465.
27. Yun, J., Won, K-H. Building environment analysis based on temperature and humidity for smart energy systems. Sensors, 12 (2012), 12458-13470.
28. Kathirgamanathan, A., de Rosa, M., Mangina, E., Finn, D.P. Data-driven predictive control for unlocking building energy flexibility: A review. Renewable and Sustainable Energy Reviews, 135 (2021), 110120.
29. Salimi, S., Hammad, A.Optimizing energy consumption and occupants comfort in open-plan offices using local control based on occupancy dynamic data. Building and Environment, 176(2020), 106818.
30. Jin, X., Wang, G., Song, Y., Sun, C. Smart building energy management based on network occupancy sensing. Journal of International Council on Electrical Engineering, 8 (2018), 30-36.
31. Hobson, B.W., Huchuk, B., Gunay, H.B., O'Brien, W. A workflow for evaluating occupant-centric controls using building simulation. Journal of Building Performance Simulation, (2021), 1-19.
32. Wang, J., Tse, N.C.F., Poon, T.Y.Chan, J.Y.C. A Practical Multi-Sensor Cooling Demand Estimation Approach Based on Visual, Indoor and Outdoor Information Sensing. Sensors, 18 (2018), 3581.
33. Girshick, R., Donahue, J., Darrell, T., Malik, J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. EEE Transactions on Pattern Analysis and Machine Intelligence, 38 (2016), 142-158.
34. Ren, X., Du, S., Zheng, Y. Parallel RCNN: A deep learning method for people detection using RGB-D images. 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017, 1-6.
35. Tien, P.W., 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. Proceedings of BSO-V 2020 5th Building Simulation and Optimization Virtual Conference 21-22nd September 2020 Loughborough, UK.
36. Perez, H., Tah, J.H.M., Mosavi, A. Deep Learning for Detecting Building Defects Using Convolutional Neural Networks. Sensors 19 (2019), 3556.
37. Chahyati, D., Fanany, M.I., Arymurthy, A.M.Tracking people by detection using CNN features. Procedia Computer Science, 124 (2017), 167-172.
38. Ahmad, M., Ahmed, I., Khan, F.A., Qayum, F., Ajuaid, H. Convolutional neural network-based person tracking using overhead views. International Journal of Distributed Sensor Networks, 16, 2020.
39. Kajabad, E.N., Ivanov, S.V. People detection and finding attractive areas by the use of movement detection analysis and deep learning approach. Procedia Computer Science, 156(2019), 3270337.

40. Wu, X., Sahoo, D., Hoi, S.C.H., Recent advances in deep learning for object detection, Neurocomputing, 396 (2020), 39-64.
41. Singaravel, S., Suykens, J., Geyer, P. Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction, Advanced Engineering Informatics, 38 (2018), 81-90.
42. Zhang, Z., Chong, A., Pan, Y., Zhang, C., Lam, K.P. Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. Energy and Buildings, 199 (2019) 472-490.
43. Konstantakopoulos, I.C., Barkan, A.R., He, S., Veeravalli, T., Liu, H., Spanos, C., A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure. Applied Energy, 237 (2019), 810-821.
44. Salimi, S. and A, Hammad. Optimising energy consumption and occupants comfort in open-plan offices using local control based on occupancy dynamic data. Building and Environment, 176 (2020), 106818.
45. Jin, X., Wang, G., Song, Y., Sun, C. Smart building energy management based on network occupancy sensing. Journal of International Council on Electrical Engineering, 8 (2018), 30-36.
46. Peng, Y., Rysanek, A., Nagy, Z., Schlüter, A. Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. Applied Energy. 211 (2018), 1343-58.
47. Dong, B., Lam, K.P. A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting. Building Simulation. 7 (2014), 89-106.
48. Amasyali, K., El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. Renewable and Sustainable Energy Reviews, 81 (2018), Part 1, 1192-1205.
49. Wang, X., Zhao, Y., Pourpanah, F. Recent advances in deep learning. International Journal of Machine Learning and Cybernetics, 11 (2020), 747-750.
50. Shaikh, P.H., Nor, N.B.M., Nallogownden, P., Elamvazuthi, I. Intelligent multi-objective optimisation for building energy and comfort management. Journal of King Saud University-Engineering Sciences, 30,2 (2018), 195-204.
51. Zheng, H., Li, F., Cai, H., Zhang, K. Non-intrusive measurement method for the window opening behavior, Energy and Building, 197 (2019), 171-176.
52. Tien, P.W., Wei, S., Calautit, J.K., Darkwa, J, Wood, C. A vision-based deep learning approach for the detection and prediction of occupancy heat emissions for demand-driven control solutions, Energy and Buildings, 226 (2020), 110386.
53. Wei, S., Tien, P.W., Calautit, J.K., Wu, Y., Boukhanouf, R. Vision-based detection and prediction of equipment heat gains in commercial office buildings using a deep learning method, Applied Energy, 227 (2020), 115506.
54. Tien, P.W., Wei, S., Calautit, J.K. A computer vision-based occupancy and equipment usage detection approach for reducing building energy demand. Energies, 14(2020), 156.
55. Tien, P.W., Wei, S., Calautit, J.K., Darkwa, J., Wood, C. Occupancy heat gain detection and prediction using deep learning approach for reducing building energy demand, J. sustain. dev. energy water environ. syst., 2020, 1080378.
56. Wei, S., Calautit, J. Development of deep learning-based equipment heat load detection for energy demand estimation and investigation of the impact of illumination. International Journal of Energy Research, 45 (2021), 7204-7221.
57. IESVE Integrated Environmental Solutions, 2021. https://www.iesve.com/. (Accessed May 2021).

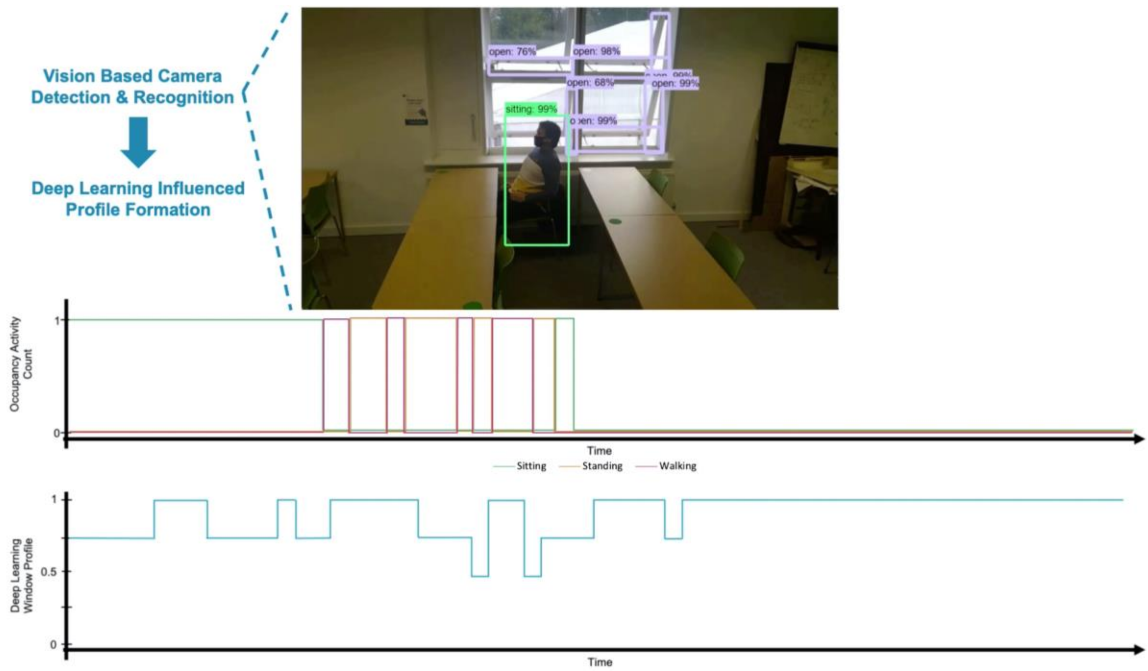
58. Dino, I.G., Sari, A.E., Iseri, K., Akin, S., Kalfaoglu, E., Erdogan, B., Kalkan, S., Alatan, A.A. Image-based construction of building energy models using computer vision. Automation in Construction, 116 (2020), 103231.
59. Zhao, Z., Zheng, P., Xu, S., Wu, X. Object Detection with Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems, 30, 11 (2019) 3212-3232.
60. Wang, W., Yang, J., Xiao, J., Li, S., Zhou, D. Face Recognition Based on Deep Learning. In: Zu Q., Hu B., Gu N., Seng S. (eds) Human Centered Computing. HCC 2014. Lecture Notes in Computer Science, vol 8944. Springer, Cham.
61. Maqueda, A.I., Loquercio, A., Gallego, G., García, N., Scaramuzza, D. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 5419-5427
62. Khan, S., Rahmani, H., Shah, S.A.A., Bennamoun, M. A Guide to Convolutional Neural Networks for Computer Vision. Synthesis lectures on Computer Vision. (2018), 207.
63. Sladojevic, S., Arsenovic, M, Anderla, A., Culibrk, D., Stefanovic, D. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Computational Intelligence and Neuroscience, (2016), 3289801.
64. Tien, P.W., Wei, S., Calautit, J.K., Darkwa, J, Wood C. A computer vision deep learning method for the detection and recognition of manual window openings for effective operations of HVAC systems in buildings. Imaginable Futures: Design Thinking, and the Scientific Method. 54th International Conference of the Architectural Science Association 2020, Ali Ghaffarianhoseini, et al (eds), pp. 21–30. Published by the Architectural Science Association (ANZAScA), 2020.
65. ASHRAE. Ventilation for Acceptable Indoor Air Quality. ANSI/ASHRAE Addendum p to ANSI/ASHRAE Standard 62.1-2013. https://www.ashrae.org/File%20Library/Technical%20Resources/Standards%20and%20Guidelines/Standards%20Addenda/62_1_2013_p_20150707.pdf (Accessed May 2021).
66. Kane. What are safe levels of CO and CO2 in rooms? https://www.kane.co.uk/knowledge-centre/what-are-safe-levels-of-co-and-co2-in-rooms (Accessed May 2021).
67. VCP. Ventilationcontrolproducts.net. Room CO2 detector RC24 and room CO2+RH+T detector RCHT 24. http://www.ventilationcontrolproducts.net/co2-detector-rc24 (Accessed May 2021).
68. Pallonetto, F., Mangina, E., Milano, F., Finn, D.P. SimApi, a smartgrid co-simulation software platform for benchmarking building control algorithms. SoftwareX, 9, (2019), 271-281.
69. Wetter, M. Co-simulation of building energy and control systems with the Building Controls Virtual Test Bed. Journal of Building Performance Simulation, 4, 3(2011).
70. Schiera, D.S., Barbierato, L., Lanzini, A., Borchiellii, R., Pons, E., Bompard, E.F., Patti, E., Machii, E., Bottaccioli, L. A distributed platform for multi-modelling co-simulations of smart building energy behaviour. 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), 2020, pp. 1-6.

**Nomenclature and Abbreviations**

| | |
|---|---|
| AI | Artificial intelligence |
| BEMS | Building energy management systems |
| BES | Building energy simulation |
| CIBSE | Chartered Institution of Building Services Engineers |
| CNN | Convolutional Neural Network |
| $CO_2$ | Carbon Dioxide |

| | |
|---|---|
| DL | Deep learning |
| DLIP | Deep learning influenced profile |
| HVAC | Heating, ventilation and air-conditioning |
| IAQ | Indoor air quality |
| IESVE | Integrated Environmental Solutions Virtual Environment |
| IoU | Intersection over Union |
| R-CNN | Region-based Convolutional Neural Network |
| RFID | Radio frequency identification |
| U | U-value ($W/m^2K$) |
| UK | United Kingdom |

**Appendix**




Video 1. A preview of the real-time detection and DLIP formation using the integrated vision-based detection approach conducted within the selected experimental test within the case study lecture room.