
Paige Wenbin Tien, Shuangyu Wei, John Kaiser Calautit, Jo Darkwa, Christopher Wood
Department of Architecture and Built Environment, University of Nottingham, Nottingham, NG7 2RD, United Kingdom

Abstract
Occuany and occupant activities within building spaces can significantly impact the energy performance of buildings and the operations of heating, ventilation, and air-conditioning (HVAC) systems. This paper explores the application of a vision-based deep learning approach for occupancy activity detection. 1,200 coloured images of occupancy performing different types of activities were collected and preprocessed before the development of a detection method based on a convolutional neural network, which was deployed to an AI-based camera. Experiments were conducted, data about the occupancy and activities performed were collected and were used to form the deep learning influenced profile (DLIP). Based on the results, overall detection accuracy of 98.65% was achieved. Along with typically ‘scheduled’ occupancy profiles, a building energy simulation (BES) tool was used to further assess the proposed method. Results indicate that using the deep learning approach can provide a more accurate prediction of the occupancy heat gains within a building space while minimising the occurrence of overestimation in occupancy heat gains by up to 64.27%. The application of the occupancy detection approach can also benefit towards the provision of the exact times when the room setpoint temperature can be adjusted to help reduce unnecessary building energy loads, while also maintaining a thermally comfortable environment for occupants.

Key Innovations
- An artificial intelligence-based technique for detecting and recognising occupancy patterns.
- Development of a deep learning detection model, focusing on the typical activities performed within an office-based environment of sitting, standing and walking.
- A vision-based demand-driven solution to assist HVAC operations based on occupancy activities.
- Real-time monitoring of occupancy patterns for better estimation of space heating and cooling requirements and provide a more comfortable environment.

Practical Implications
Occuany behaviour has been identified as an important issue impacting the energy demand of building and HVAC systems. This study proposes a vision-based deep learning approach to capture, detect and recognise in real-time the occupancy patterns and activities within an office space environment. Initial building energy simulation analysis of the application of such an approach within buildings was performed. The proposed approach is envisioned to enable HVAC systems to adapt and make a timely response based on occupancy’s dynamic changes. The results presented here show the practicality of such an approach for various building spaces and environments.

Introduction and Literature Review
Occuany behaviour and patterns within building spaces have been identified as a significant contributing factor towards building energy efficiency (Delzende et al. 2017). Recent studies have analysed ocucany behaviour in buildings and developed novel demand system-based solutions to improve building system operations (Paone and Baacher, 2018). To obtain ocucany data, various technologies were employed, including infrared (Yun and Lee, 2014), Wi-Fi (Simma et al. 2019) and Radio Frequency Identification (RFID) (Li et al. 2011). These solutions provide information about a building space such as ocucany count and location, however, there are several limitations such as the requirement of multiple sensors distributed across the room and not having the ability to detect ocucany activities. Furthermore, indirect methods such as environment-based sensors were used (Yun and Won, 2012), which records the changes within the space when occupants are present. Effectively, the data collected is employed to develop demand response based solutions for more effective system controls (Kathirgamanathan et al. 2021), energy optimisation (Salimi and Hammad, 2020), and also building energy management (Jin et al. 2018). Many of these energy solutions are based on machine learning models that provide advantages in their adaptability and application to different types of buildings (Amasyali and El-Gohary, 2018). Further enhancement towards such strategies includes the prospect of achieving a multi-objective system that enables building energy and comfort management (Shaikh et al. 2018). The above methods that obtain ocucany data for data-driven and forecasting-based solutions that determines the operation of building energy systems were highly dependent on historical data (for example, the conditions obtained from previous days/conditions) (Marinakis, 2020). However, with varying occupancy patterns, it presents a lack of diversity and a potential time delay occurring between the prediction and the provision of the actual building requirements. This indicates the need to develop solutions such as demand-driven controls that adapt to occupancy patterns in real-time and optimise
HVAC operations, while also providing comfortable conditions.

Furthermore, the cooling/heating design setpoint temperature assigned to building spaces is usually based on the indoor space's purpose/function at the early design stages. For instance, (CIBSE, 2015a) suggests operative temperatures for spaces such as offices, libraries and restaurants at 21 - 25°C in the UK. During the design stage and building energy simulation, fixed or scheduled profiles are set for the building HVAC systems. However, the impact of different occupancy patterns and the activities performed are typically not considered, resulting in over or under conditioned building spaces.

The study by (Tien et al. 2020a) presented the use of a vision-based detection approach to identify occupancy activities within a building space and develop real-time occupancy profiles. The generated profiles were compared to fixed or static scheduled profiles. It was highlighted that activities conducted by occupants could affect the building performance due to the internal heat gains directly (Tien et al. 2020b) and also indirectly (Tien et al. 2021) However, no in-depth analysis on the impact of such a detection approach on the building energy performance was carried out. The present study provides insights on how the proposed detection method can enable HVAC systems to adapt and make a timely response to dynamic changes of occupancy, instead of using “static” or fixed occupancy operation schedules.

**Method**

**Approach**

This study aims to use a building energy simulation tool to analyse the performance and feasibility of applying a developed vision-based real-time occupancy activity detection approach. The detection and recognition approach applied during an experimental test within an office space environment were used to assess the effectiveness of the approach and feasibility for integration with a demand-driven control strategy to enhance the performance of building heating, ventilation and air-conditioning (HVAC) systems.

A convolutional neural network-based model configured and trained using an image dataset. It was then deployed to an artificial intelligence-powered camera and was used to perform live detection to form the data-driven real-time Deep Learning Influenced Profiles (DLIP). To evaluate the performance of the trained model an experimental test was performed using a case study office building. Thereafter, the selected case study office space was modelled using a building energy simulation (BES) tool and the proposed method was compared against typical or fixed occupancy operation schedules. This was carried out to assess the potential influence on the building energy demand and the operations of HVAC systems.

**Deep Learning Method**

To establish the deep learning vision-based occupancy detection model, input data in the form of images were collected. As shown in Figure 1, common occupancy activities performed within typical office buildings were selected as the desired model detection responses. This includes the activities of sitting, standing and walking. The images collected were all coloured (RGB) and was used to form the images within both the training and testing dataset as given in Table 1. This table presents the number of images collected and the number of labels assigned to both datasets. It should be noted that both datasets consist of different, but similar images of occupancy activities. As presented in Figure 1, the labels assigned to the images represent each of the bounding boxes drawn manually around each image's specific region interest. For some cases, more than one occupant appears within the image; hence multiple labels were assigned.

![Figure 1: Example images of occupancy with different activities obtained from Google Images.](image)

<table>
<thead>
<tr>
<th>Occupancy Activity</th>
<th>Number of Training Images</th>
<th>Number of Labels</th>
<th>Number of Testing Images</th>
<th>Number of Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting</td>
<td>400</td>
<td>753</td>
<td>100</td>
<td>149</td>
</tr>
<tr>
<td>Standing</td>
<td>400</td>
<td>701</td>
<td>100</td>
<td>134</td>
</tr>
<tr>
<td>Walking</td>
<td>400</td>
<td>1000</td>
<td>100</td>
<td>177</td>
</tr>
<tr>
<td>Total</td>
<td>1200</td>
<td>2454</td>
<td>300</td>
<td>460</td>
</tr>
</tbody>
</table>

To perform the occupancy activity classification task, the CNN TensorFlow Object Detection API (Huang et al. 2016) and the selected pre-trained COCO-trained model of Faster R-CNN (with Inception V2) were used in the form of a transfer learning approach to assist the development of the convolutional neural network (CNN) based model. It provided benefits in terms of the training duration and the overall detection performance compared to training a model from scratch. Figure 2 presents an outline of the overall model architecture and pipeline configuration of the occupancy activity detection model.
To achieve multiple occupancy activity detection methods, high computational power is required in the process of developing (training) the initial model. However, once the model was trained it can be deployed onto a programmable camera to form an AI-based camera for detection within any type of indoor space and low computational power is required.

**Application of the Deep Learning Model**

The developed deep learning model was deployed to a camera to perform real-time detections. To test the model performance, a 15-minute experimental test was conducted within an office space on the 1st floor of the Sustainable Research Building (University of Nottingham, UK). Figure 3a presents the floor plan of the selected area and Figure 3b shows the detection area captured from the wide field of view 1080p resolution camera which was used for the detection. The camera used for detection was near the ceiling of the room. During the experimental test, only three occupants performed a series of activities and for the analysis, the occupants were identified as Person A, B and C. Future work would include a greater number of participants within the experimental tests.

To analyse the potential influence of the application of such a deep learning-based occupancy activity detection model on the building energy demand, building energy simulation using the simulation tool, IESVE was employed. The simulation included several scenarios based on 4 days (Wednesday – Saturday) during the UK heating season (winter). Since current buildings and building energy simulation, mechanical systems are often designed to assume constant internal heat gains or static occupancy during building operational hours (CIBSE, 2015c), the following profile in Figure 3a was generated for comparison. For this selected room, the typical number of occupants in the room would be 8. Two conventional fixed scheduled-based profiles of Typical Office Profiles 1 and 2 were established and applied. With ‘sitting’ being the most common activity performed within an office space environment, Typical Office 1 represents this during the building operational hours. In addition, Typical Office 2 represents the maximum condition (walking). However, in reality, variations occur...
in occupancy levels and activities. As an example, a scenario-based occupancy DLIP is generated and simulated in BES, as shown in Figure 5b. This was compared with the use of the fixed or scheduled profile given in Figure 5a.

Figure 5: Occupancy profiles used within the building energy modelling representing part of a typical office week (Wednesday – Saturday, a). Typical static occupancy profiles and b). Scenario-based deep learning influenced occupancy profile (DLIPs).

Results and Discussion

Deep Learning Model Training Performance

The occupancy activity detection model was trained for 102,194 steps, and a minimum loss of 0.005654 was reached.

Using common classification evaluation metrics to assess the detection performed on the images from the test dataset in terms of true positives, true negatives, false positives and false negatives, suggesting that the activity of sitting achieved the highest performance with a detection accuracy of 94.04%. However, occupants performing activities such as standing and walking can have similar occupancy body form and shape, which could create difficulties in identifying the true activity. It, therefore, resulted in a lower accuracy value of 91.43% for standing and 92.70% for walking. Despite this, the results (Table 3) indicated that the majority of the prediction labels were correctly assigned to the desired occupancy activity which also suggested the capability of the approach to provide sufficiently accurate detections of occupants within indoor spaces.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting</td>
<td>94.04%</td>
<td>0.9250</td>
<td>0.8911</td>
<td>0.8911</td>
</tr>
<tr>
<td>Standing</td>
<td>91.43%</td>
<td>0.9064</td>
<td>0.8284</td>
<td>0.8284</td>
</tr>
<tr>
<td>Walking</td>
<td>92.70%</td>
<td>0.8643</td>
<td>0.9266</td>
<td>0.9266</td>
</tr>
</tbody>
</table>

Detection Performance Evaluation and DLIP

To provide further analysis of the detection performance, an experimental test was performed within the selected office space. Figure 6 presents a timeline indicating the key stages of the test with example detection images from various stages. The approach provides multiple occupancy recognition within the space. Above each bounding box, it also shows the detection accuracy. Furthermore, both the appearance and the shape and sizes of the bounding boxes varied between each of the detection intervals. It should be noted that in practice, the device won’t be storing or outputting images. It will only output real-time information based on the number of occupants performing each of the activities to generate and establish the deep learning influenced heat emission profiles, DLIPs.

Figure 6: Example of images from various stages during the experimental detection test.

Table 3: Occupancy activity detection model performance based on the application of images from the testing dataset.
All three occupants performed the activity of sitting, and the detection accuracy for Person A, B and C were 97.88%, 98.91% and 98.38%, respectively. Since Person B achieved the highest value and Person A achieved the lowest, it suggests that distance had a minimal effect on detecting and recognising people with a sitting position, based on the tests carried out. However, the camera angle used for detection may impact the result. The standing activity was only performed by Person A, and an accuracy of 98.81% was achieved. Additionally, the performance of walking achieved an average accuracy of 98.74%. These initial results highlight the model’s capabilities in recognising the differences between the different human poses within an office space environment.

During the experimental test, time-stamped data were generated by the occupancy detection method. Data were used to form the deep learning influenced profiles (DLIP) given in Figure 8a, and the heat emission-based deep learning influenced profile (DLIP) indicated in Figure 8b. The results achieved by the DLIP were compared with static profiles and with the Actual Observation Profile corresponding to the ‘true’ activities performed by the occupants.

BES results indicate that using the DLIP can result in up to 29.09% and 10.60% decrease in occupancy gains if occupants were not assumed to be following the Typical Profiles 1 and 2. However, the results also suggest the DLIP still alternates between different responses and led to a 4.14% error when compared to the actual observations, indicating the need for further improvement towards the model detection performance to provide more accurate DLIPs.

**Energy Performance Analysis**

The selected office space was modelled and simulated with the various occupancy profiles detailed in Figure 4. Figure 9 presents the total and distribution of occupancy heat gains achieved across the selected four days. The monitoring of occupancy activities through the detection approach provided an overall heat gain of 17.4kWh. However, Typical Office 1 and 2 achieved an overall heat gain of 38.6kWh and 48.7kWh. Overall, it indicates that the use of Typical Profiles 1 and 2 leads to an overestimation up to 54.92% and 64.27% compared to the use of the deep learning influenced profile whereby the time and the type of activities performed by occupants were captured.
Based on the current standards and guidelines, (ASHRAE 90.1, 2019) and (ASHRAE 55, 2017) it suggests a generalised set point range and schedule for room heating and cooling during occupied and unoccupied hours. For example, during occupied hours, it suggests 22 – 27°C for cooling and 17 – 22°C for heating, while during unoccupied hours it suggests 27 – 30°C for cooling and 14 – 17°C for heating. Furthermore, (CIBSE, 2015a) suggest office buildings maintain an operative room temperature of 21 – 23°C during the winter and 22 – 25°C during summer. Simulations were performed with the heating setpoint temperature of 21°C and 22°C during the building operational hours of 06:00 – 18:00, and 15°C was set during the unoccupied hours. In addition, the cooling setpoint temperature was assumed to be 25°C.

Figure 10 presents the predicted heating load. When the office space heating setpoint was modified from 21°C to 22°C, the heating load would increase by an average of 11kWh for all the cases. Therefore, changing the room temperature requirements by 1°C can ultimately change the total heating demand by an average of 15.45%.

Understanding occupancy patterns and activities with the use of the proposed method can help improve building energy predictions and system operations. Given the results in Figure 10, where the Deep Learning Scenario-based results showed lower occupancy heat gains and higher heating demand as compared to Typical Office 1 and 2. Furthermore, Figure 11 presents the distribution of heating load across time. It suggests that both the Typical Office results follow the same pattern, as variation in heating was only dependent on the outdoor and indoor condition. However, the Deep Learning Scenario-based results indicated that a demand-driven approach could help provide the actual requirements of the space while minimising the unnecessary energy demand.

Figure 9: Comparison of the occupancy heat gains achieved using the deep learning approach compared to the different typical occupancy scheduled profiles (a) Total gains and (b) Variation across time.

**Conclusion and Future Works**

This study presents the analysis of the application of a vision-based deep learning approach for occupancy activity detection conducted within an open-plan office space. The detection model was first developed by using a transfer learning-based method to establish and train a convolutional neural network for the classification of occupancy activities. The model was then deployed towards a camera which enabled the evaluation of the performance of real-time detections. The initial performance of the model was evaluated based on a 15-minute experimental detection test. Average detection accuracy of 98.65% was achieved across all activities. During the real-time detection of a selected space, constant data about the number of occupants performing each of the selected activities was generated from the deep activity setpoint conditions.

Figure 10: Comparison of the total heating load achieved using the deep learning approach compared to the different typical occupancy profiles, with room heating setpoint at 21°C and 22°C.

The proposed approach can help provide the actual requirements of the space while minimising the unnecessary energy demand.
learning influenced profile (DLIP). The generated DLIP was compared to typical static profiles indicating the ability to provide more valuable data about occupants within a building space regarding heat emission for more effective HVAC system operations.

Building energy simulation were conducted within the selected office building. A scenario-based occupancy profile generated from the application of the deep learning detection approach during 4 days was created. It consisted of a variation in occupancy activities. Through building energy simulation, this was compared with the use of static occupancy profiles. Results indicate that the deep learning approach can minimise the occurrence of overestimation in occupancy heat gains by up to 48.7kWh. For such a scenario-based situation, an increase of heating by up to 8.92% (6.1kWh) in heating was required to maintain a comfortable indoor environment.

The application of the occupancy detection approach can help reduce the unnecessary building energy loads while also assist in maintaining a thermally comfortable environment for occupants.

Limitations and Future Works

Instead of using conventional sensors, the proposed approach can be used. An AI-powered camera devices can be used to provide accurate and real-time monitoring of the occupancy count, location and activity levels within a space of a building space which could assist the operations of a building HVAC system. The camera is used for detection purposes only. Therefore, privacy issues can be avoided as this approach does not require the collection or storing of data in the form of images or videos. Instead, data will be used to form the real-time DLIP (with only text and numerical data in the form of profiles).

Improvements to the current approach are required before the integration with controls of building HVAC systems. Future works include the improvement of the detection model through enhancing the input image dataset. A series of experimental tests to be conducted to verify the approach's feasibility in a diverse range of indoor environments. Furthermore, a streamlined framework-based solution must be developed to define the exact HVAC control system conditions based on the given real-time detection data responses.

Acknowledgement

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


CIBSE Chartered Institution of Building Services Engineer (2015a). Table 1.5 Recommended comfort criteria for specific applications. Environmental design: CIBSE Guide A. In: London: CIBSE; 2015.


