Delft University of Technology

Enhancing real-time nonintrusive occupancy estimation in buildings via knowledge fusion network
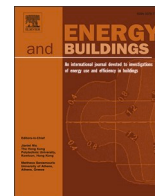
Lu, Chujie

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Enhancing real-time nonintrusive occupancy estimation in buildings via knowledge fusion network

Chujie Lu [a,b]

[a] *Faculty of Architecture and the Built Environment, Delft University of Technology, Julianalaan 134, 2628 BL Delft, The Netherlands*
[b] *School of Computer Science and Technology, Guangdong University of Technology, 510006 Guangzhou, PR China*

ABSTRACT

Real-time nonintrusive occupancy estimation can maximize the use of existing sensors to infer occupant information in buildings with the advantages of fewer privacy concerns and fewer extra device costs. Recently, many deep learning architectures have proven effective in estimating occupancy directly from raw sensor data. However, some handcrafted features manually extracted from statistical and temporal domains might convey additional information for occupancy estimation. In this study, a novel knowledge fusion network for nonintrusive occupancy estimation is proposed to integrate knowledge from two streams, i.e. automatic knowledge stream from a deep learning architecture and handcrafted knowledge stream from manual feature engineering. Moreover, four different fusion modules are investigated to optimize the design of the fusion network. To verify the effectiveness of the proposed network, experiments are conducted in a dataset from the ASHRAE Global Occupant Behavior Database, which is collected from an office space with records of indoor environment parameters, occupant-building interactions, and contextual information. The results demonstrate the superiority of the proposed fusion network, which outperforms five representative algorithms. Furthermore, the ablation study underscores the benefits of knowledge fusion and occupant-building interaction information, showing that the proposed fusion network can enhance the occupancy estimation accuracy by 3.47 % to 9.24 %.

1. Introduction

In 2021, the operation of buildings accounted for 30 % of global final energy consumption and 27 % of total energy sector emissions [1]. To align with the Net Zero Scenario, by 2030, the emissions from building operations need to reduce by more than half. The reduction requires significant efforts to develop efficient and intelligent technologies for all end uses, including leveraging the potential of occupant behavioral changes [2]. One of the main targets of building service is to provide a comfortable and healthy indoor environment for occupants. Real-time occupancy information is an essential input for the control systems of intelligent buildings, including lighting, heating, ventilation, and air conditioning [3,4]. Accurately estimating the occupancy information can be beneficial to predict energy consumption [5]. Furthermore, occupant-centric building operations can effectively improve the energy efficiency [6–9] and space use efficiency [10] of buildings.

The presence and number of occupants are the basic occupancy information that needs to be estimated in intelligent buildings [11–14]. In the past few decades, many methods have been developed for real-time occupancy estimation. Passive infrared (PIR)-based motion sensors are the most common approach to infer occupancy. But motion sensors are limited by the detection distance and also are hard to sense static occupants. Camera-based occupancy estimation has attracted researchers in computer engineering because of its high accuracy [15–17]. However, Camera-based occupancy estimation suffers from privacy issues and difficulty in expensive computation and deployment [12]. Non-intrusive occupancy estimation refers to indirectly inferring occupancy by data mining algorithms from the interactions between occupants and buildings [18–21], such as the indoor environment and the energy use of various electric appliances. In particular, recently, indoor environment quality and plug electricity use have begun to be widely monitored due to the popularity of the concept of the Internet of Things [22]. Non-intrusive occupancy estimation can leverage the existing sensors and achieve reliable accuracy with fewer extra costs and privacy issues. However, due to the indirectness of non-intrusive occupancy estimation, the relationship between occupancy and sensor data can be complicated. Therefore, it is essential to develop data mining algorithms with high accuracy and universality for non-intrusive occupancy estimation.

The data mining algorithms for non-intrusive occupancy estimation include shallow machine learning and deep learning algorithms. The

objective is to build a relationship between the occupancy and the features that can be extracted manually from raw sensor data or automatically learned through deep learning algorithms. For shallow machine learning, because the raw sensor data can be noisy and not representative well for occupancy estimation, data preprocessing and feature extraction are necessary before the inference [23–27]. The important features can be manually extracted from statistical and temporal domains. After carefully obtaining some informative handcrafted features, shallow machine learning algorithms can be employed to estimate the occupancy. With the rapid development of artificial intelligence, deep learning is another powerful choice for non-intrusive occupancy estimation. Instead of manual feature engineering, deep learning algorithms are able to automatically learn representative features and estimate occupancy from raw sensor data. In particular, convolutional neural networks (CNN) can extract robust local features [28–31]. Since the sensor data and occupancy are sequential with temporal dependencies, recurrent neural networks (RNN) are a popular choice [28,29,32–35], especially long short-term memory (LSTM).

However, certain limitations persist when applying data mining algorithms to achieve nonintrusive occupancy estimation. These limitations are succinctly summarized as follows:

- While RNN-based deep learning algorithms have demonstrated their efficacy in nonintrusive occupancy estimation, they may not consistently surpass the performance of shallow machine learning employing carefully designed handcrafted features [32]. This indicates that certain handcrafted features, enriched with domain knowledge, can indeed convey supplementary information for occupancy estimation.
- Numerous investigations into fusion models for occupancy estimation merely explore distinct combinations of sensor features [18,27,36], thus warranting a more comprehensive exploration of fusion strategies at higher feature levels.
- Many studies tend to disregard the influence of occupant-building interactions on non-intrusive occupancy estimation [23,28,37,38]. These interactions encompass factors like the state of windows and doors, as well as the operation of air-conditioning (AC) systems. Notably, these studies often employ simplified closed-room setups, which may not accurately reflect real-world scenarios. The omission of these interactions has the potential to introduce variations in the relationships between occupancy information and the indoor environment. Consequently, it is imperative to incorporate these interactions into the modeling and analysis process to ensure a more accurate and holistic understanding of occupancy patterns.

To solve the limitations above, the contributions of this study are listed as follows.

- A novel knowledge fusion network is proposed for non-intrusive occupancy estimation within buildings. The proposed network integrates knowledge from two distinct streams, i.e. automatic knowledge stream from a deep architecture and handcrafted knowledge stream from a shallow network with informative handcrafted features.
- Diverse fusion modules are explored in this study, thereby refining and optimizing the architecture of the knowledge fusion network tailored specifically for occupancy estimation. This rigorous investigation into fusion strategies aims to extract the most advantageous synergies between different knowledge streams.
- The effectiveness of the proposed network is rigorously validated using a dataset from the ASHRAE Global Occupant Behavior Database, which is collected from an office environment. This dataset encompasses a comprehensive range of indoor environmental parameters, occupant-building interactions, and contextual information, providing a robust foundation for assessing the network's performance.

Table 1

The representative studies on nonintrusive occupancy estimation using data mining algorithms.

Ref.	Input	Feature Extraction	Feature selection	Algorithm
[39]	CO_2, CO, TVOC, Relative humidity PM2.5, Acoustics	First-order difference, Second-order difference, First-order shifted difference, 20-min moving average	Filter based on relative information gain (RIG)	SVM, ANN, HMM
[26]	CO_2, Relative humidity, Temperature, Pressure	First-order shifted difference, Second-order difference, moving average	Wrapper-based ranking, Filter–wrapper hybrid algorithm based on RIG	ELM
[27]	CO_2, Relative humidity, Temperature, Pressure	First-order difference, Second-order difference, First-order shifted difference, 5-min moving average, 10-min moving average.	ELM-based wrapper	ELM, SVM, ANN, K-nearest neighbors (KNN), Classification and regression tree (CART)
[45]	CO_2, Relative humidity, Temperature, Pressure	First-order difference, Second-order difference, First-order shifted difference, 15-min moving average,	Filter-based on RIG	IHMM-MLR
[25]	CO_2, Relative humidity, Temperature, Pressure	First-order difference, Second-order difference, First-order shifted difference, moving average, frequency domain data	Local Receptive Fields with random weights.	ELM
[43]	Temperature, Humidity, Derived humidity ratio, Light, CO_2	Time average, First order differences	–	HMM
[28]	CO_2, Relative humidity, Temperature, Pressure.	–	–	CDBLSTM
[18]	CO_2, TVOC, Temperature, Relative humidity	Moving average, Moving sample variance, First derivative, Second, derivative	Correlation-based filter	Naïve Bayes (NB), KNN, RF, Decision tree, ZeroR, JRip
[33]	Wi-Fi connections	Time-series decomposition	–	RF, ANN, LSTM
[29]	AMI data	–	–	CNNBiLSTM
[36]	Wi-Fi access points, CO_2, PIR motion, Plug and light electricity load	–	–	Linear regression, ANN

Table 1 (*continued*)

Ref.	Input	Feature Extraction	Feature selection	Algorithm
[19]	CO_2, Temperature, Relative humidity, Noise, Illuminance	–	–	C4.5 decision tree, Curve description
[23]	CO_2	Wavelet denoising	–	Multi-grained cascade forests (GcForest)
[42]	Temperature, Relative humidity, CO_2, TVOC, Power, Door state, Window state, AC operation	–	–	Law of Total Probability (LTP), Naïve Bayes classifier (NB), CART
[31]	CO_2, Relative humidity, Temperature	–	–	CNN-XGBoost

- To establish the superiority of the proposed network, a thorough comparison is conducted against five representative models sourced from pertinent studies. Additionally, an ablation study is performed to meticulously dissect and evaluate the individual impacts of various network components on overall model performance.

The rest of the paper is organized as follows. Section 2 reviews and summarizes current studies on data mining approaches for nonintrusive occupancy estimation in buildings. Section 3 introduces the proposed knowledge fusion network for nonintrusive occupancy estimation. Section 4 describes the experiment, including the dataset and model setup. Section 5 presents the results and discussions. Section 6 concludes this paper and provides future research directions.

2. Related works

This section provides a literature review of data mining algorithms for nonintrusive occupancy estimation from three main perspectives, including model input, feature engineering, and model selection. The summary of the representative studies is presented in Table 1.

Due to the impact of occupants on the indoor environment, the most common input is indoor environmental sensor data. Dong et al., [39] presented an ambient sensor network that measured many indoor environment parameters to detect the presence of occupants, including indoor carbon dioxide (CO_2), carbon monoxide (CO), total volatile organic compounds (TVOC), PM2.5, temperature, and acoustics. Chen et al., [40] proposed a light-emitting diode (LED) light sensor-based occupancy estimation system, which can infer the occupancy level using the variance of diffusion reflection caused by the presence of occupants. Candanedo et al., [41] detected occupancy in an office using data from light, temperature, humidity, and CO_2 sensors. Also, considering the occupants' activities in the buildings, electricity consumption and Wi-Fi connections also can be effective indicators of indoor occupancy information. Wang et al., [33] inferred occupant counts from the number of Wi-Fi-connected devices for long-term and short-term connection in an office building. Hobson et al., [36] developed an implicit occupancy sensing using Wi-Fi access points, CO_2 sensors, PIR motion detectors, and plug and light electricity load meters in an academic building. Notably, some occupant-building interactions might impact the relationship between the indoor environment and occupancy information, such as opening and closing doors and windows, and operating AC systems. Fajilla et al., [42] improved the performance of occupancy estimation with both indoor environmental parameters and information about occupant-building interactions, such as window and door status, and AC system use. Zhou et al., [19] designed an algorithm

to recognize the occupant presence by indirectly detecting the usage information of light and AC systems. Candanedo et al., [43] also found the performance of occupancy estimation could be affected when occupants opened the windows or doors.

Many data mining algorithms have been established successfully for non-intrusive occupancy estimation in buildings, which can be roughly classified into two categories: shallow machine learning and deep learning algorithms. The popular shallow machine learning algorithms include support vector machine (SVM) [39], artificial neural network (ANN) [36], extreme learning machine (ELM) [24,44], decision tree [41,42], random forest (RF) [23,32,38,41], hidden Markov model (HMM) [39,45], and so on. Notably, the raw sensor data usually has random noise and can be less directly informative for occupancy estimation. Thus, manual feature engineering is necessary and beneficial before inferring occupancy using shallow machine learning algorithms. Candanedo et al., [43] proposed a methodology based on HMM, where the time average was applied to remove noise and spikes in the raw indoor environment sensor data, and then the first-order difference was extracted before modeling. Chen et al., [27] proposed a framework to estimate occupancy based on ELM-based wrapper feature selection, where the important features were extracted from environment parameters, such as the first-order difference, second-order difference, first-order shifted difference, moving average, and shifted difference. These features have proven to be effective representatives and widely used in many studies. Mustafa et al., [24] developed a novel framework with hybrid feature-scaled ELM based on similar extracted features. Zimmermann et al., [18] also applied correlation-based feature selection to identify the important features and performed a comparison of seven supervised machine learning algorithms.

Deep learning algorithms are considered a powerful technique to model the nonlinear relationship between sensor data and occupancy information. In particular, deep learning architecture can automatically learn the significant features from raw sensor data without human intervention. Also, due to the temporal dependency, RNN can effectively capture significant information from occupancy and sensor data. Chen et al., [28] proposed a convolutional deep Bi-LSTM (CDBLSTM) approach for building occupancy estimation that contains an automatic feature learning process for environmental sensor data. Feng et al., [29] developed a convolutional neural network with Bi-LSTM (CNNBiLSTM) to detect occupancy from advanced metering infrastructure (AMI) data. Mohammadabadi et al., [31] proposed CNN-XGBoost to detect occupancy based on environmental data in a residential building, where the high-level features extracted by CNN were fed into XGBoost to classify occupancy. Tekler et al., [46] used five different deep learning architectures for occupancy prediction in an office, library, and lecture room, including deep neural network (DNN), LSTM, Bi-LSTM, Gated Recurrent Unit (GRU), and Bi-directional GRU (Bi-GRU).

However, deep learning algorithms cannot always achieve the best performance for nonintrusive occupancy estimation. Huchuk et al., [32] found that RF outperformed the LSTM network when predicting occupancy using thermostat data. Wang et al., [33] also found RF was better than the LSTM network for inferring occupancy using Wi-Fi connections. Therefore, the knowledge automatically extracted from deep learning is not enough and the handcrafted knowledge from shallow machine learning may contain additional information for occupancy estimation. It is highly motivated to improve the performance of nonintrusive occupancy estimation by integrating the two pieces of knowledge in a reasonable way. Notably, even though some studies have mentioned fusion models for nonintrusive occupancy estimation [18,27,36,47], their focus has predominantly revolved around diverse feature-level combinations, which is insufficient.

3. Methodology

In this section, the overview of the proposed knowledge fusion network for non-intrusive occupancy estimation is presented. Then, the
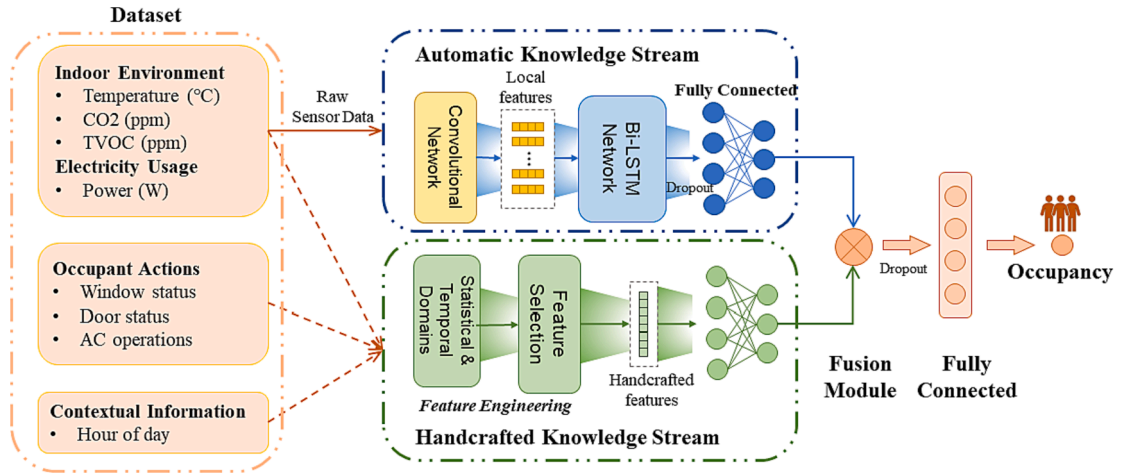
Fig. 1. The Overview of the proposed knowledge fusion network for nonintrusive occupancy estimation.

details of the automatic knowledge stream and the handcrafted knowledge stream are introduced respectively. Finally, the knowledge fusion modules implemented in this work are demonstrated.

3.1. Overview of the proposed knowledge fusion network

For non-intrusive occupancy estimation in buildings, the main purposes are, first, to extract representative and informative features from raw sensor data, and second, to map the complicated and indirect relationship between the extracted features and occupancy. The proposed knowledge fusion network is presented in Fig. 1. The sensor data are from the indoor environment and electricity usage. In the proposed network, the raw sensor data are fed directly into a deep learning architecture, which includes a convolutional network and a Bi-LSTM network. The convolutional network can obtain compressed local features and the Bi-LSTM network can learn the temporal dependencies of the local features. Then, a fully connected network can seek a high-level representation of automatic features for occupancy estimation. Meanwhile, according to expert knowledge, the handcrafted features of the raw sensor data are extracted from the statistical and temporal domains. The information on occupant actions also is added to the handcrafted knowledge stream, which can help modify the impact of occupant-building interactions on non-intrusive occupancy estimation. To remove the redundant information, a feature selection is conducted to determine the most informative handcrafted feature combination. And then, the feature combination is fed into another fully connected network to obtain abstract handcrafted features. Finally, to take full advantage of information from the two knowledge streams, the automatic feature and the handcrafted feature are combined to form a complete feature, which will be fed into a fully connected classifier to estimate the final occupancy.

3.2. Automatic knowledge stream

In the automatic knowledge stream, a widely used deep architecture, the convolutional network with Bi-LSTM (CNN-BiLSTM) [28,29], is utilized to obtain the automatic feature for occupancy estimation. The convolutional network is applied for directly extracting the local features from the raw sensory data, while the Bi-LSTM network is used to encode and learn the temporal dependencies of the local features in two directions, including forward and backward. The details of the convolutional network and the Bi-LSTM network are introduced as follows.

The convolutional network usually consists of two main components, a convolutional layer and a pooling layer. The convolutional layer can extract multiple representative output features. Since the sensory data from both the indoor environment and electricity usage are sequences, a one-dimension convolution (1D-CNN) kernel is utilized in the convolutional layer. The pooling layer is used to further compress and down-sample the dimension and parameter quantity of the previous output feature. A widely used max-pooling layer is used in this work. The raw sensory data is assumed to be $X_{r \times d}^{(i)}$, where r is the length of sequence, d is the number of sensors, and i denotes i-th 1D convolution kernel. The 1D convolution operation can be described as follows.

$$C_{l \times d}^{(i)} = \theta\left(W_{l \times d}^{(i)} \oplus X_{r \times d}^{(i)} + B^{(i)}\right) \tag{1}$$
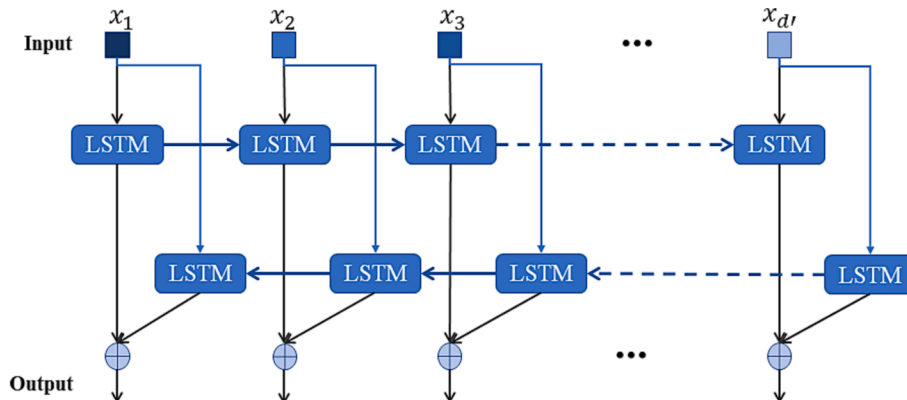

Fig. 2. The illustration of the Bi-LSTM network.

where $\oplus$ represents the concatenation operator; $W_{l\times w}^{(i)}$ represents the weight of the i-th 1D convolution kernel; l is the length of the 1D convolution kernel; $B^{(i)}$ represents the bias of i-th 1D convolution kernel; $C_{\hat{l}\times d}^{(i)}$ is the output feature after 1D convolutional operation; $\hat{l}$ is the length of the output feature; d is the width of the output feature; $\theta(\bullet)$ is the nonlinear activation function that is a rectified linear unit (ReLU) function in this work.

RNN is designed to understand temporal dependencies in time series and has been successfully utilized for sequence modeling, such as non-intrusive occupancy estimation. However, the vanilla RNN suffers from the problem of gradient disappearance and long-term memory deficiency in the processing of long sequence data. To solve this problem, the LSTM network was proposed to be a memory cell. LSTM can capture long-term dependencies because three gates are designed to allow or prevent the passing of information along the sequence. Furthermore, the Bi-LSTM was proposed to take past and future contexts into consideration. Therefore, Bi-LSTM can be more powerful for capturing time dependencies in two directions: directions: forward and backward. The demonstration of the Bi-LSTM is presents in Fig. 2. The computation process of the forward direction can be formulated as follows.

$$\overrightarrow{h}(t) = f\left(\overrightarrow{x}(t), \overrightarrow{h}(t-1), \omega\right)$$

$$= \begin{cases} \overrightarrow{i}(t) = \sigma\left(\overrightarrow{\omega}_{ii}\,\overrightarrow{x}(t) + \overrightarrow{\omega}_{hi}\,\overrightarrow{h}(t-1)\right) \\ \overrightarrow{f}(t) = \sigma\left(\overrightarrow{\omega}_{if}\,\overrightarrow{x}(t) + \overrightarrow{\omega}_{hf}\,\overrightarrow{h}(t-1)\right) \\ \overrightarrow{g}(t) = \tanh\left(\overrightarrow{\omega}_{ig}\,\overrightarrow{x}(t) + \overrightarrow{\omega}_{hg}\,\overrightarrow{h}(t-1)\right) \\ \overrightarrow{o}(t) = \sigma\left(\overrightarrow{\omega}_{io}\,\overrightarrow{x}(t) + \overrightarrow{\omega}_{ho}\,\overrightarrow{h}(t-1)\right) \\ \overrightarrow{c}(t) = \overrightarrow{f}(t) \odot \overrightarrow{c}(t-1) + \overrightarrow{i}(t) \odot \overrightarrow{g}(t) \\ \overrightarrow{h}(t) = \overrightarrow{o}(t) \odot \tanh(\overrightarrow{c}(t)) \end{cases} \tag{2}$$

where $\overrightarrow{i}(t)$, $\overrightarrow{f}(t)$, $\overrightarrow{g}(t)$, and $\overrightarrow{o}(t)$ represent the gates for input, forget, memory cell, and output at the time t, respectively; $\overrightarrow{h}(t)$ and $\overrightarrow{c}(t)$ represent the hidden and cell state at the time t; $\sigma(\bullet)$ represents the *sigmoid* function and $\odot$ denotes the Hadamard product; ω represents the learnable weights.

Similar to the forward direction, the computation process of the backward direction is as follows.

$$\overleftarrow{h}(t) = f\left(\overleftarrow{x}(t), \overleftarrow{h}(t+1), \omega\right)$$

$$= \begin{cases} \overleftarrow{i}(t) = \sigma\left(\overleftarrow{\omega}_{ii}\,\overrightarrow{x}(t) + \overleftarrow{\omega}_{hi}\,\overleftarrow{h}(t+1)\right) \\ \overleftarrow{f}(t) = \sigma\left(\overleftarrow{\omega}_{if}\,\overrightarrow{x}(t) + \overleftarrow{\omega}_{hf}\,\overleftarrow{h}(t+1)\right) \\ \overleftarrow{g}(t) = \tanh\left(\overleftarrow{\omega}_{ig}\,\overrightarrow{x}(t) + \overleftarrow{\omega}_{hg}\,\overleftarrow{h}(t+1)\right) \\ \overleftarrow{o}(t) = \sigma\left(\overleftarrow{\omega}_{io}\,\overrightarrow{x}(t) + \overleftarrow{\omega}_{ho}\,\overleftarrow{h}(t+1)\right) \\ \overleftarrow{c}(t) = \overleftarrow{f}(t) \odot \overleftarrow{c}(t+1) + \overleftarrow{i}(t) \odot \overleftarrow{g}(t) \\ \overleftarrow{h}(t) = \overleftarrow{o}(t) \odot \tanh(\overleftarrow{c}(t)) \end{cases} \tag{3}$$

Finally, two outputs from two directions, $\overrightarrow{o}(t)$ and $\overleftarrow{o}(t)$, will be com-

Table 2
the handcrafted features extracted from the statistical and temporal domains.

Feature	Description
15-min moving average	$ma(t) = \left(\sum_{j=t-14}^{t} raw(j)\right)/15$
First-order difference	$fd(t) = raw(t) - raw(t-1)$
Second-order difference	$fd2(t) = fd(t) - fd(t-1)$
First-order shifted difference	$fds(t) = raw(t) - raw(t-2)$

bined to obtain the complete representation from the Bi-LSTM as follows.

$$o_{final}(t) = \overrightarrow{o}(t) + \overleftarrow{o}(t) \tag{4}$$

3.3. Handcrafted knowledge stream

For non-intrusive occupancy estimation, the raw sensory data may be noisy, which cannot be fed directly into shallow machine learning. Thus, manual feature engineering is necessary before occupancy estimation. Manual feature engineering contains two steps, feature extraction and feature selection. After that, the most informative feature combination will be fed into a fully connected network to obtain more abstract features.

Feature extraction is to obtain more representative and informative features of indoor occupancy, according to expert knowledge. In this work, the features are extracted from the statistical and temporal domains of the raw sensory data, shown in Table 2. The 15-min moving average (ma) is a common statistical feature that takes into account the time delay of the indoor environment parameters. The first-order difference (fd), the second-order difference ($fd2$), and the first-order shifted difference (fds) are also widely used to capture the temporal variances in the environmental data. These four features have proven effective to provide valuable information for occupancy estimation [18,24,27,41,45]. Therefore, we believe the information in these features can be a valuable supplement to the feature learned by the deep learning structure. Furthermore, the occupant-building interactions will significantly impact the relationship between the indoor environment and occupancy [19,41,42]. For example, the state of windows will impact the indoor environment. The model for occupancy estimation in a closed room cannot be simply applied in the situation of a room with an open window. The features representing the interactions can help modify the information and improve the model for occupancy estimation. Thus, three features representing the interactions are also supplied to the handcrafted feature stream, including the state of the window, the state of the door, and the state of the AC systems. after the feature extraction, to remove the redundant information and select the best feature combination, a filter-based feature selection is conducted. The filter-based feature selection quickly assesses the features using common criteria, named Information Gain [39].
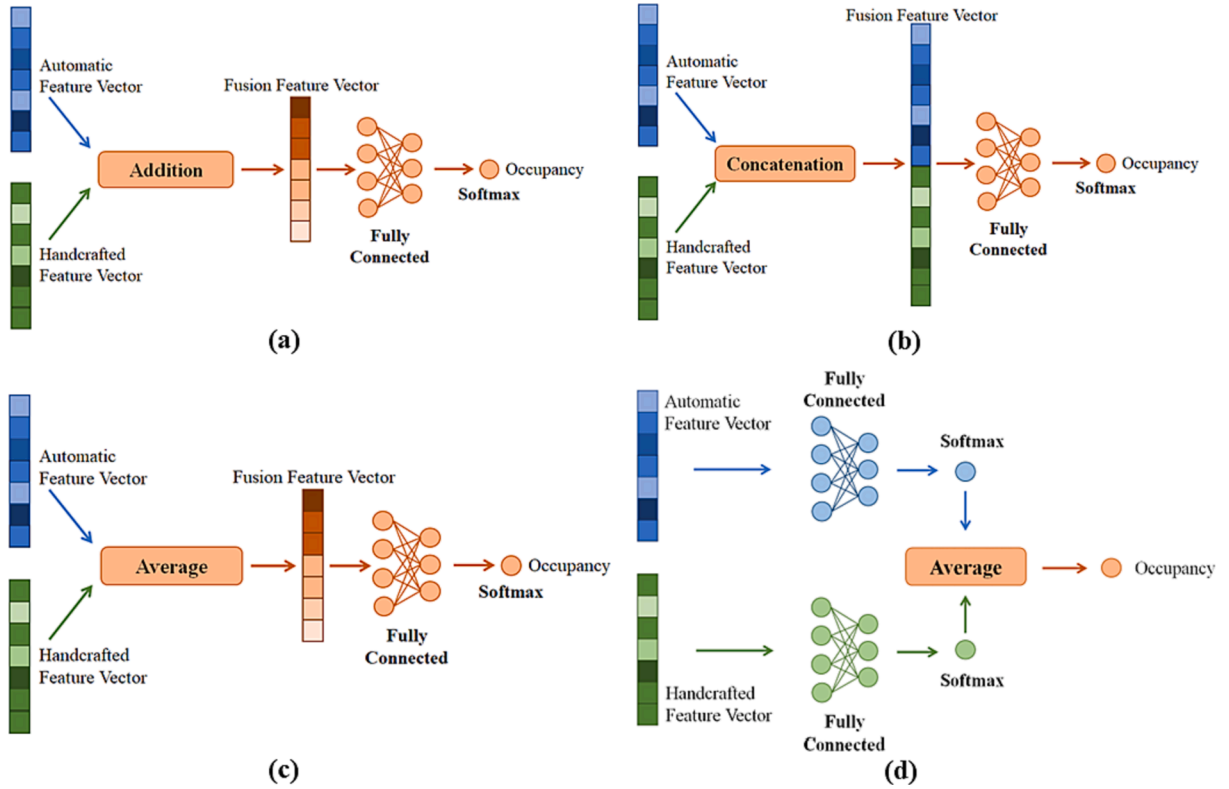
Fig. 3. The illustration of the fusion modules: (a) Addition fusion, (b) Concatenation fusion, (c) Average fusion, and (d) Final average fusion.

3.4. Knowledge fusion modules

To make full use of the valuable information from both the automatic knowledge stream and the handcrafted knowledge stream, four different fusion modules are implemented in the proposed knowledge fusion network and compared in the experiments, shown in Fig. 3.

a) Addition fusion: an intuitive approach to combining the two streams is to conduct an element-wise addition operation to merge the learned features from the two streams. The dimensions of the two learned feature vectors must be the same and the fusion feature vector will remain the same shape after the addition operation. The process is defined as follows.

$$f_{out} = f_a(X_{raw}) + f_h(X_h)\#$$ (5)

where $f_a(X_{raw})$ indicates the learned feature vector from the raw sensory data through the automatic feature stream; $f_h(X_h)$ represents the learned feature vector from the handcrafted features through the handcrafted feature stream.

b) Concatenation fusion: different from the previous fusion modules, this fusion operation will directly concatenate the learned features together. The dimension of the fusion feature vectors will increase and be equal to the sum of the dimensions of the learned features from two streams.

$$f_{out} = f_a(X_{raw}) \oplus f_h(X_h)$$ (6)

c) Average fusion: another intuitive approach to combining the two streams is to weight sum the fusion feature from the learned features from the two streams. In this work, the same weight is given to the two learned features, so-called average fusion. Similar to the addition fusion, the dimensions of the two learned feature vectors must be the same and the fusion feature vector also remains the same. The process can be formulated as follows.

$$f_{out} = Avg(f_a(X_{raw}), f_h(X_h))$$ (7)

d) Final average fusion: the simplest fusion method to integrate the knowledge from both streams is a weighted sum of the estimated occupancy. Different from the previous three fusion modules, the two knowledge streams are trained separately in the final fusion. When estimating the final occupancy, the fusion network computes the average of the occupancy estimated by the two knowledge streams, which can be formulated as follows.

$$O_{final} = Avg(O_a(f_a(X_{raw})), O_h(f_h(X_h)))$$ (8)

where O_a, O_h, and O_{final} are the occupancy estimated by the automatic knowledge stream, the handcrafted knowledge stream, and the final fusion network respectively. When the final occupancy was not an integer, the rounding operation would be applied.

4. Data preprocessing and model setting

4.1. Data description and preprocessing

An open dataset from the ASHRAE Global Building Occupant Behavior Database [48] is used to verify the effectiveness of the proposed knowledge fusion network for non-intrusive occupancy estimation. The dataset used in this work was collected in an office building of the University of Calabria located in Southern Italy, whose collection details can be found in [42,49]. The climatic condition is Mediterranean hot-summer climate. The dataset contains the number of occupants in the office, as well as the indoor environment (indoor air temperature, indoor air relative humidity, CO_2, and TVOC), the electricity usage, the occupant actions (window state, door state, and AC operations), and the contextual information (Hour of day). The initial resolution of the dataset is one minute.

In this work, due to too many missing values, the indoor air relative humidity was not used. Also, the office was only occupied during working time. To balance the estimation target, i.e., the number of occupants, and avoid too many unoccupied samples, the experiments were
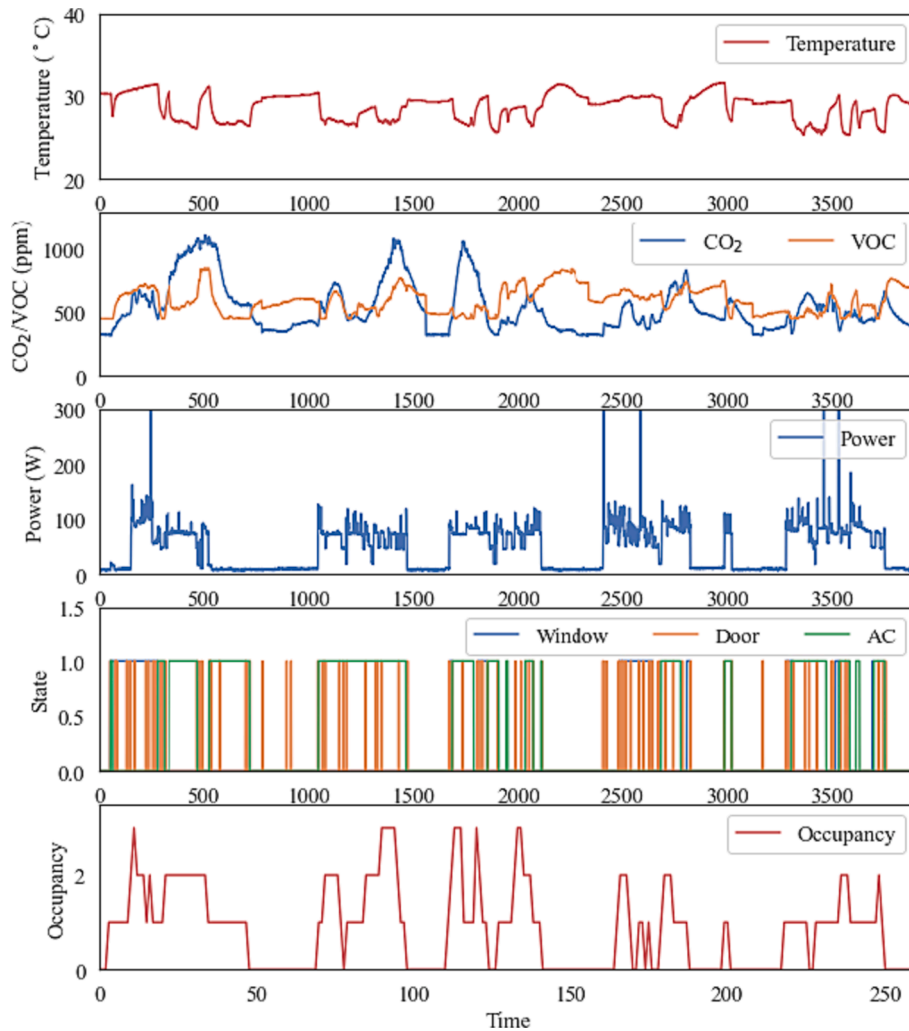
Fig. 4. The illustration of the testing set (five working days) in this work.

based on the data from 8 am to 9 pm on the working days. The data from May 13th to July 3rd, 2016 were selected as the training set, and the data from July 4th to July 8th, 2016 were used as the testing set. During the collection, the hourly average outdoor temperature in the coldest day (May 13th, 2016) was from 16.1 °C to 21.7 °C while in the hottest day (July 8th, 2016) was from 23.3 °C to 29.4 °C, which are provided by "Centro Funzionale Multirischi della Calabria' of the 'Agenzia Regionale per la Protezione dell'Ambiente della Calabria (Arpacal)" (https://www.cfd.calabria.it).The testing set (five working days) is presented in Fig. 4. Due to the delay of the indoor environment and the slow response of the building control systems, 15-minute resolution of the data would be adequate [50]. Thus, the number of occupants was transferred into 15-minute resolution by averaging. When the averaging occupancy was not an integer, the rounding operation would be applied. Thus, as illustrated in Fig. 4, it is important to note that the temporal granularity utilized varies across different data categories. The indoor environment (indoor air temperature, indoor air relative humidity, CO_2, and TVOC), the electricity usage, and the occupant actions (window state, door state, and AC operations) are presented at a minutely time resolution with 3750 samples. The occupancy is aggregated over a 15-minute time interval with 250 samples.

4.2. Model setup and evaluation

Since the occupancy resolution is 15 min and the raw sensory data is recorded every minute, the length of sequence r is 15 for the 1D CNN. In the proposed network, for the automatic feature stream, the filter number, window size, and pooling size of the convolutional network are 32, 3, and 2, respectively. Then, a one-layer Bi-LSTM network with the hidden size of 32 is chosen. The fully connected networks in both two streams are with a size of [100, 10] and a ReLU activation function so that the dimension of the automatic feature vector and the handcrafted feature vector is 10. After the fusion module, the fully connected network with a hidden neuron size of 10 is adopted before the final *softmax* classification for occupancy estimation. Also, the dropout layer is applied for data argumentation and to prevent the overfitting issue in the proposed network, where the dropout rate is set to 0.5. The final and optimal hyperparameters are determined by running several configurations. In the training process, the batch size is 32 and the epoch number is 200. The optimizer is *Adam* and the target training loss is cross-entropy. The experiments are performed in Python and the proposed information fusion network is achieved using Keras. The experiments are conducted in a personal computer with an Intel® Core™ i5-1135G7 central processing unit, 2.4 GHz processor, and a memory size of 32G.

To assess the performance of the proposed model, three widely-used metrics were applied to compare the estimation results [23,28,45], including estimation accuracy (EA), mean absolute error (MAE), and detection accuracy (DA). EA represents the accuracy for estimating the number of occupants. MAE represents the error between the estimated and actual number of occupants. DA represents the accuracy for detecting the presence and absence of occupants, which is useful to the

Table 3
The handcrafted features extracted by manual feature engineering.

Raw data	Handcrafted features
Temperature	**Temp_ma**, **Temp_fd**, Temp_fd2, **Temp_fds**
CO_2	**CO₂_ma**, CO₂_fd, CO₂_fd2, CO₂_fds
TVOC	**TVOC_ma**, TVOC_fd, TVOC_fd2, TVOC_fds
Power	Power_ma, **Power_fd**, Power_fd2, Power_fds
Window status	**Window status**
Door status	Door status
AC operation	**AC operation**
Hour of days	**Hour of days**

light control systems. The calculation of metrics is as follows.

$$EA = \frac{\sum_{k=1}^{n} \chi(o_k - \widehat{o_k})}{n} \qquad (7)$$

$$MAE = \frac{\sum_{k=1}^{n} |o_k - \widehat{o_k}|}{n} \qquad (8)$$

$$DA = \frac{n_{(p,p)} + n_{(a,a)}}{n} \qquad (9)$$

where o_k is the ground truth of the number of occupants in the office at time k; $\widehat{o_k}$ is the estimated number of occupants; n is the number of samples; $\chi(o_k - \widehat{o_k}) = \begin{cases} 1, & if o_k - \widehat{o_k} = 0 \\ 0, & otherwise \end{cases}$; $n_{(p,p)}$ is the number of samples that both the estimated occupancy and the ground truth are "presence"; $n_{(a,a)}$ is the number of samples that both the estimated occupancy and the ground truth are "absence".

5. Results and discussion

5.1. Results of handcrafted feature engineering

Before training the information fusion network, manual feature engineering is conducted to select the most informative feature combination as inputs in the handcrafted information stream. First, the extracted handcrafted features are presented in Table 3. Here, 'Temp_ma', 'Temp_fd', 'Temp_fd2', and 'Temp_fds' refer to the moving average feature, the first-order difference feature, the second-order difference feature, and the first-order shifted difference feature of temperature respectively, and so forth. Furthermore, the information gain of each handcrafted feature to occupancy is calculated to select important

features, presented in Fig. 5. The extracted features are sorted by the absolute values of the information gain. Among indoor environmental parameters, 'CO_2_ma' is the most important feature for occupancy estimation, followed by 'Temp_ma'. Also, some features representing the occupant-building interactions are important, such as 'Window state' and 'AC operation', which affect the indoor environmental parameters. But 'Door state' shows its low importance, probably because the opening time of the door is very short. Due to the regularity of occupant activities in the office, the contextual information (Hour of Day) also is a useful feature. Finally, the top nine handcrafted features are selected and they are in bold in Table 3. The selected handcrafted features will be fed into the handcrafted knowledge stream in the proposed network to estimate real-time occupancy.
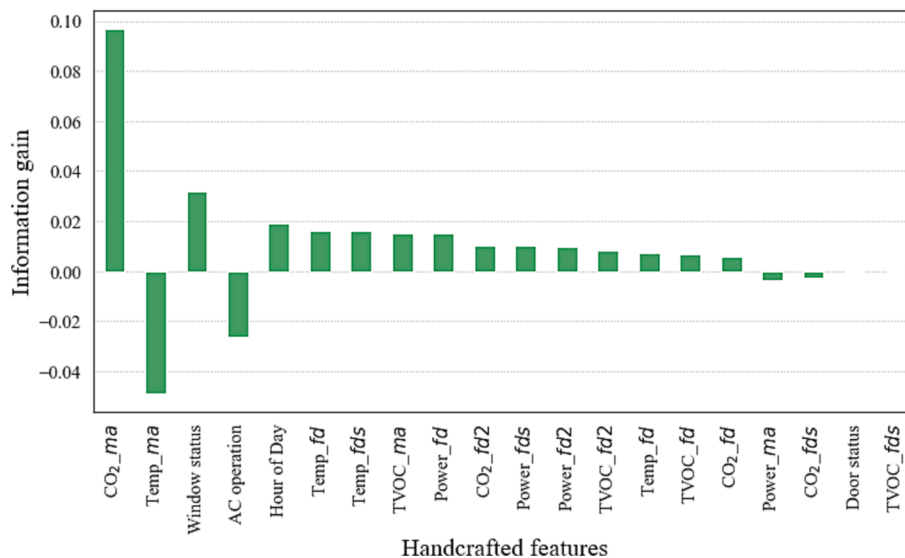
5.2. Results of different fusion modules

To effectively integrate the two knowledge streams, four fusion modules are implemented and the results of the proposed network with different fusion modules are presented in Table 4. Overall, in the training set, the performance of the four fusion modules is similar, where the average fusion obtains slightly better results on EA and DA. In the testing set, the proposed network with the addition fusion obtains the highest DA, which means that it can accurately detect the presence and absence of occupants in the buildings. The final fusion is significantly worse than other fusion modules because it is essentially the simplest bagging ensemble learning with equal weights to the results of the two knowledge streams, which cannot effectively integrate the information. The proposed network with the average fusion achieves the best performances generally. Thus, the proposed network with the average fusion will be used for performance comparison with other models.

Table 4
The results of the proposed network with different fusion modules.

	Training set			Testing set		
	EA (%)	MAE	DA (%)	EA (%)	MAE	DA (%)
Addition Fusion	76.75	0.278	97.10	68.08	0.358	**91.54**
Average Fusion	**77.68**	0.270	**97.87**	**74.62**	**0.280**	90.00
Concatenation Fusion	77.34	0.273	97.61	69.62	0.338	91.19
Final Fusion	75.55	**0.265**	90.89	64.23	0.408	85.77

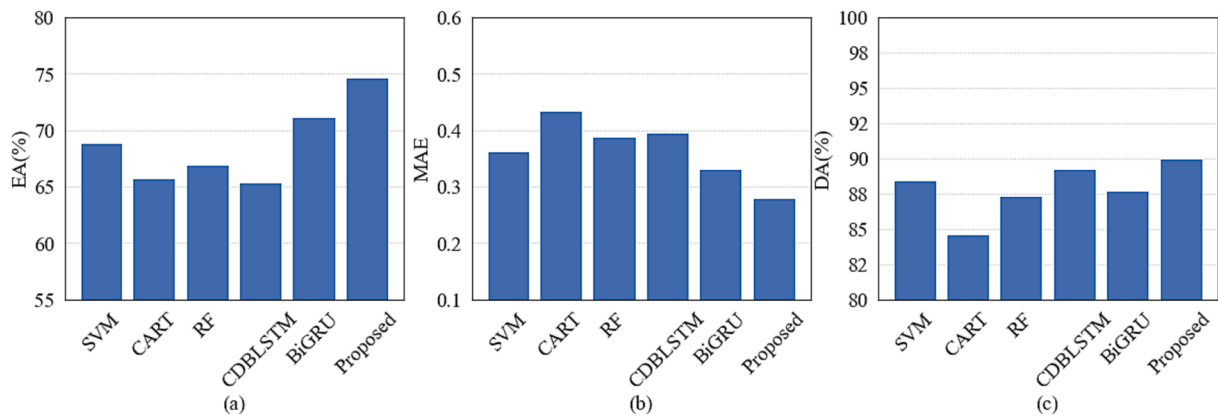**Fig. 5.** The Information Gain ranking of the handcrafted features.

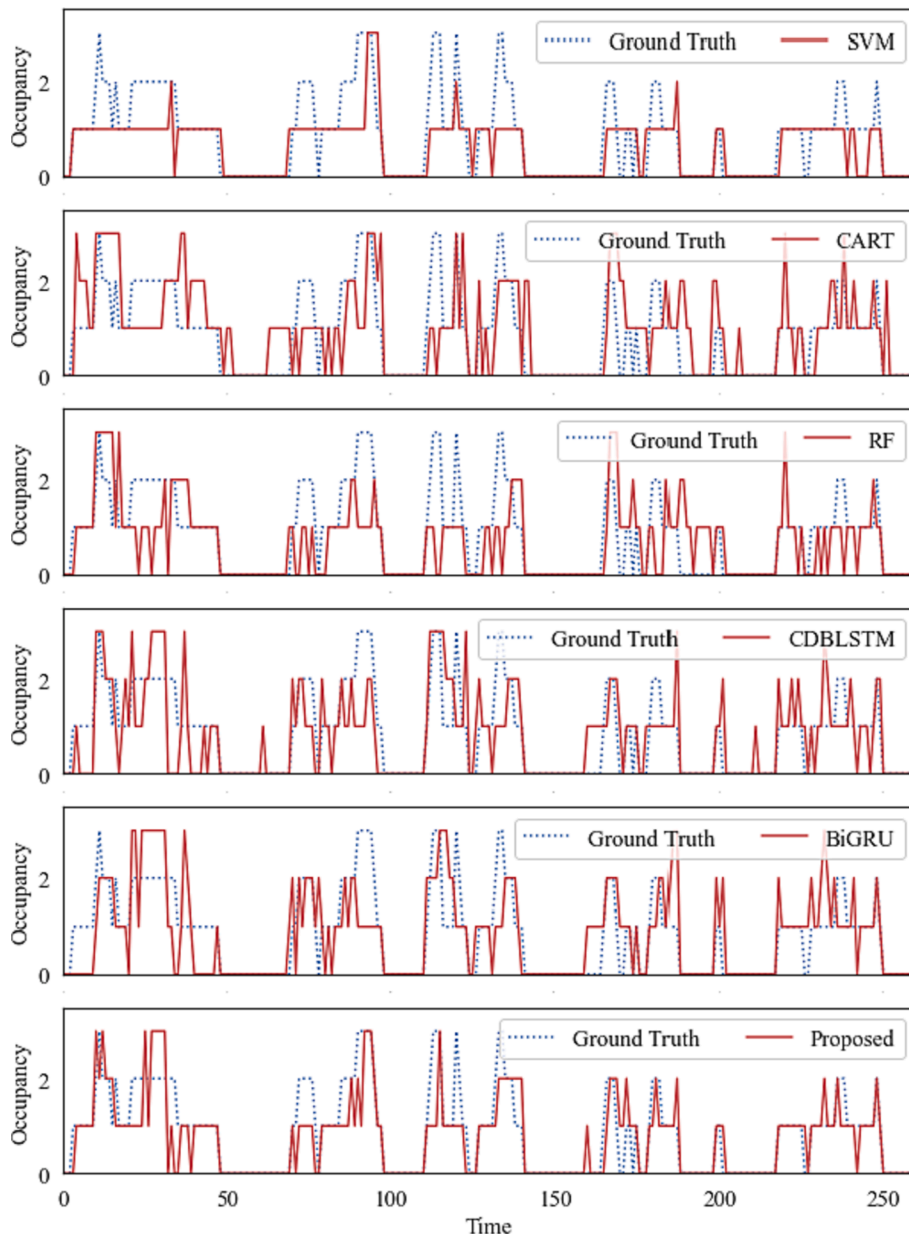**Fig. 6.** The performance of different models for nonintrusive occupancy estimation, (a) EA, (b) MAE, and (c) DA.

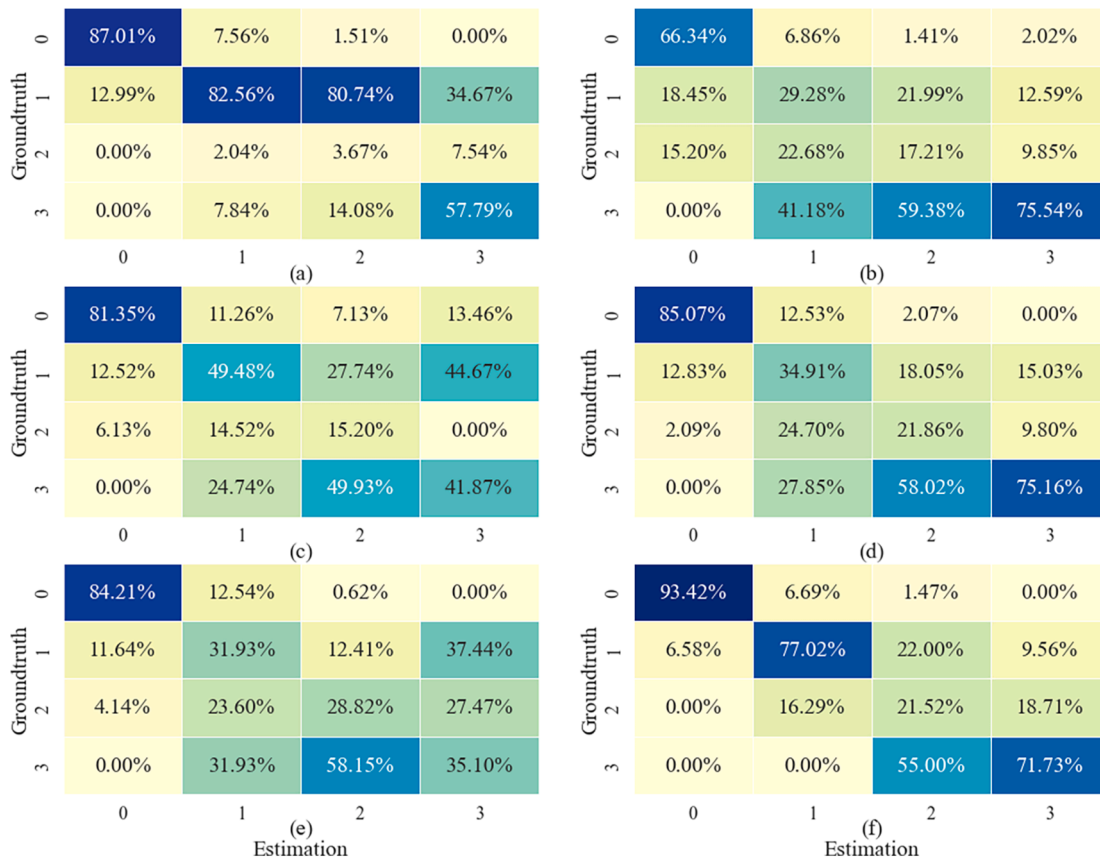**Fig. 7.** The occupancy estimation results of different models.

Fig. 8. Confusion matrix of different models for occupancy estimation, (a) SVM, (b) CART, (c) RF, (d) CDBLSTM, (e) Bi-GRU, and (d) Proposed.

5.3. Performance comparison with other models

In this section, to assess the performance for nonintrusive occupancy estimation, the proposed knowledge fusion network is compared with other representative models in previous studies, including SVM [39], CART [41], RF [32,33], CDBLSTM [28,30], and Bi-GRU [46]. Since SVM, CART, and RF are shallow machine learning models and could not automatically extract features, the selected handcrafted features would as inputs for these three models. CDBLSTM and Bi-GRU are state-of-the-art deep learning algorithms that would directly process the raw sensor data.

The performance for nonintrusive occupancy estimation of these models is presented in Fig. 6. Overall, the proposed fusion network achieves the best performance with EA of 74.62 % and MAE of 0.28. Meanwhile, SVM obtains the highest DA and performs better than the CDBLSTM and Bi-GRU on all the metrics. We notice that CDBLSTM and Bi-GRU contain three layers of RNN, which makes the over-complexity of these two models that causes the overfitting problem. This also means that even though deep learning algorithms have powerful automatic feature learning abilities, they cannot replace shallow machine learning and manual feature learning, indicating that it is worth integrating the knowledge from both deep learning and shallow machine learning. Also, compared to CART, RF obtains better performance in occupancy estimation because it is an ensemble learning algorithm made up of multiple decision trees in an effective way.

The occupancy estimation results of different models are presented in Fig. 7. SVM obtains a stable estimation while ignoring many details of occupancy changes. CART, RF, CDBLSTM, and Bi-GRU generate many fluctuations. Furthermore, the shallow machine learning algorithms miss the peak occupancy on the third day, which deep learning algorithms can estimate accurately. The discrepancy may arise due to the challenge in accurately capturing the dynamics of indoor environment

Table 5
The description of models for the ablation study.

Model	Description
M1	Only automatic knowledge stream
M2	Only handcrafted knowledge stream
M3	Automatic knowledge stream (without the convolutional network) with handcrafted knowledge stream
M4	Automatic knowledge stream with handcrafted knowledge stream (without the occupant actions features)
M5	Automatic knowledge stream with handcrafted knowledge stream (without the manual feature selection)

parameters, despite the extraction of crucial handcrafted features of raw sensor data in statistical and temporal domains through manual feature engineering. This difficulty often leads to an incomplete representation of the intricate nuances within the indoor environment, consequently overlooking vital occupancy-related information. Also, CDBLSRM and Bi-GRU cannot accurately the first arrival time on the first day. This might be misled by the low electricity use and the unusual changes in indoor environment parameters, such as temperature, CO_2, and TVOC, which are obviously affected by the opening of windows in the office. These two deep learning algorithms ignore these parameters representing the occupant-building interactions. The confusion matrix of different models for occupancy estimation is presented in Fig. 8. The proposed fusion model achieves a significant improvement in the precision rates, which are the number on the diagonal. Overall, the proposed fusion network performs the best occupancy estimation through effectively integrating the knowledge from shallow machine learning and deep learning algorithms.
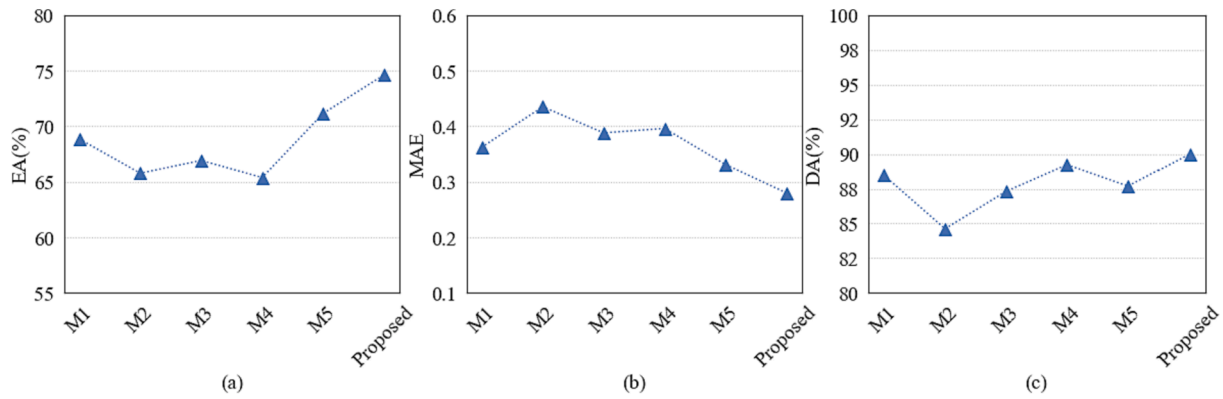
Fig. 9. The performance of the models of the ablation study for nonintrusive occupancy estimation.

5.4. Ablation study

To further demonstrate the effectiveness of the proposed fusion network and the contribution from each component, an ablation study is conducted in this section. The ablation study is a crucial method to understand the influence of each component and investigate the causation of these parts in the complicated network.

The models proposed for the ablation study are presented in Table 5 and their performances for nonintrusive occupancy estimation are shown in Fig. 9. Generally, the automatic knowledge stream outperforms the handcrafted knowledge stream. Removing the convolutional network would lose the ability to automatically extract the robust local features and the EA decreases to 66.92 %. Removing the manual feature selection would make the handcrafted features redundant and the EA slightly decreases to 71.15 %. Also, the network without the occupant actions features would significantly affect occupancy estimation, whose EA would decrease to 65.38 % and MAE would increase to 0.396. In summary, integrating the knowledge in a reasonable way can effectively improve occupancy estimation performance. Through assessing the contribution from each component, the improvement of the EA is by 3.47 % to 9.24 %, the reduction of the MAE is by 0.051 to 0.155, and the improvement of the DA is by 0.77 % to 5.38 % in the proposed fusion network.

5.5. Contribution, limitation, and future direction

The core goal of nonintrusive occupancy estimation: obtaining valuable occupancy insights through the utilization of the existing sensors without imposing additional privacy concerns or substantial costs. Indeed, over the past decade, numerous studies have substantiated the correlation between the existing sensor data and occupancy, establishing a foundation for advancing the field of nonintrusive occupancy estimation. As we shift focus toward maximizing the utilization of this sensor-derived information, algorithm development emerges as a central research emphasis. The primary goal is to effectively transform raw sensor data into accurate and meaningful occupancy estimates.

The main contribution of this study is to underscore the efficacy of high-level knowledge fusion in enhancing the accuracy of occupancy estimation. Notably, the impact of occupant-building interactions on estimation accuracy, particularly concerning actions such as window openings and AC operations, has been clearly demonstrated. These insights emphasize the need to account for these interactions in pursuit of improved accuracy.

However, it is crucial to recognize that the current study's conclusions are rooted in data collected exclusively from an office setting during the summertime in Italy. Given the dynamic nature of outdoor factors such as temperature, CO_2 levels, and TVOC concentrations throughout the day and across seasons, the influence of window openings can exhibit varying and even contradictory effects. Additionally,

the modes and operations of AC systems can impart distinct impacts on the intricate interplay between occupancy and the indoor environment, varying across different seasons.

In the future, advancing nonintrusive occupancy estimation involves two key directions: firstly, an extended and diversified data collection and testing approach encompassing various building types, settings, seasons, and occupant interactions is crucial to ensure the robustness and generalizability of findings. This comprehensive dataset will illuminate the intricate relationships between occupancy and the indoor environment. Secondly, the development of advanced algorithms, such as data fusion and transfer learning, is pivotal for more accurate and adaptable occupancy estimation in real-world scenarios. These advanced algorithms will drive practical solutions for occupant-centric control strategies that enhance efficiency and comfort in the built environment.

6. Conclusion

In this study, we propose a novel knowledge fusion network for nonintrusive occupancy estimation in buildings. There are two knowledge streams in the proposed fusion network, so-called automatic knowledge stream and handcrafted knowledge stream. The automatic knowledge stream can directly process the raw sensor data by extracting the robust local features using the convolutional network. Then, the Bi-LSTM network is applied to obtain the temporal dependency in the local features and occupancy information. The handcrafted knowledge stream utilizes manual feature engineering to obtain handcrafted features from statistical and temporal domains of the raw sensor data. Then, these handcrafted features are fed into a fully-connected network after the feature selection. To investigate the most efficient method to integrate the knowledge from the two streams, four different fusion modules are discussed, including addition fusion, concatenation fusion, average fusion, and final fusion.

The experiments are conducted on an open dataset collected from an office space, including indoor environment parameters, occupant-building interactions, and contextual information. The results show that the proposed fusion network with the average fusion module achieves the best performance and outperforms five representative models in previous studies, including SVM, CART, RF, CDBLSTM, and Bi-GRU. The results also show that even though deep learning algorithms have powerful automatic feature learning abilities from raw sensor data, they cannot replace shallow machine learning and manual feature learning, indicating that it is worth integrating the knowledge from both deep learning and shallow machine learning for nonintrusive occupancy estimation. Moreover, the ablation study assesses the influence of each component, and the features representing the occupant-building interactions have the most significant impact on the performance of occupancy estimation. The proposed fusion network achieves the improvement of the EA by 3.47 % to 9.24 %, the reduction of the MAE by

0.051 to 0.155, and the improvement of the DA by 0.77 % to 5.38 %.

CRediT authorship contribution statement

Chujie Lu: Conceptualization, Methodology, Software, Investigation, Visualization, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

References

[1] International Energy Agency, Buildings, 2022. https://www.iea.org/reports/buildings (accessed December 25, 2022).

[2] C. Lu, J. Gu, W. Lu, An improved attention-based deep learning approach for robust cooling load prediction: public building cases under diverse occupancy schedules, Sustain Cities Soc. 96 (2023), https://doi.org/10.1016/j.scs.2023.104679.

[3] B. Gunay, W. Shen, Connected and distributed sensing in buildings: improving operation and maintenance, IEEE Syst. Man. Cybern. Mag. 3 (2017) 27–34, https://doi.org/10.1109/msmc.2017.2702386.

[4] R. Jia, B. Jin, M. Jin, Y. Zhou, I.C. Konstantakopoulos, H. Zou, J. Kim, D. Li, W. Gu, R. Arghandeh, P. Nuzzo, S. Schiavon, A.L. Sangiovanni-Vincentelli, C.J. Spanos, Design automation for smart building systems, Proc. IEEE 106 (2018) 1680–1699, https://doi.org/10.1109/JPROC.2018.2856932.

[5] C. Lu, S. Li, Z. Lu, Building energy prediction using artificial neural networks: a literature survey, Energy Build. 262 (2022), https://doi.org/10.1016/j.enbuild.2021.111718.

[6] W. O'Brien, A. Wagner, M. Schweiker, A. Mahdavi, J. Day, M.B. Kjærgaard, S. Carlucci, B. Dong, F. Tahmasebi, D. Yan, T. Hong, H.B. Gunay, Z. Nagy, C. Miller, C. Berger, Introducing IEA EBC annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation, Build Environ. 178 (2020), https://doi.org/10.1016/j.buildenv.2020.106738.

[7] W. Zhang, Y. Wu, J.K. Calautit, A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment, Renew. Sustain. Energ. Rev. 167 (2022), https://doi.org/10.1016/j.rser.2022.112704.

[8] D. Yan, T. Hong, B. Dong, A. Mahdavi, S. D'Oca, I. Gaetani, X. Feng, IEA EBC Annex 66: definition and simulation of occupant behavior in buildings, Energy Build. 156 (2017) 258–270, https://doi.org/10.1016/j.enbuild.2017.09.084.

[9] C. Lu, S. Li, J. Gu, W. Lu, T. Olofsson, J. Ma, A hybrid ensemble learning framework for zero-energy potential prediction of photovoltaic direct-driven air conditioners, J. Build. Eng. 64 (2023), https://doi.org/10.1016/j.jobe.2022.105602.

[10] S. Azizi, G. Nair, R. Rabiee, T. Olofsson, Application of Internet of Things in academic buildings for space use efficiency using occupancy and booking data, Build Environ. 186 (2020), https://doi.org/10.1016/j.buildenv.2020.107355.

[11] W. Jung, F. Jazizadeh, Human-in-the-loop HVAC operations: a quantitative review on occupancy, comfort, and energy-efficiency dimensions, Appl. Energy. 239 (2019) 1471–1508, https://doi.org/10.1016/j.apenergy.2019.01.070.

[12] Z. Chen, C. Jiang, L. Xie, Building occupancy estimation and detection: a review, Energy Build. 169 (2018) 260–270, https://doi.org/10.1016/j.enbuild.2018.03.084.

[13] J. Yang, M. Santamouris, S.E. Lee, Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings, Energy Build. 121 (2016) 344–349, https://doi.org/10.1016/j.enbuild.2015.12.019.

[14] Z. Nagy, B. Gunay, C. Miller, J. Hahn, M. Ouf, S. Lee, B.W. Hobson, T. Abuimara, K. Bandurski, M. André, C.-L. Lorenz, S. Crosby, B. Dong, Z. Jiang, Y. Peng, M. Favero, J.Y. Park, K. Nweye, P. Nojedehi, H. Stopps, L. Sarran, C. Brackley, K. Bassett, K. Govertsen, N. Koczorek, O. Abele, E. Casavant, M. Kane, Z. O'Neill, T. Yang, J. Day, B. Huchuk, R.T. Hellwig, M. Vellei, Ten questions concerning occupant-centric control and operations, Build Environ. (2023), 110518, https://doi.org/10.1016/j.buildenv.2023.110518.

[15] K. Sun, Q. Zhao, J. Zou, A review of building occupancy measurement systems, Energy Build. 216 (2020), https://doi.org/10.1016/j.enbuild.2020.109965.

[16] N. Alishahi, M.M. Ouf, M. Nik-Bakht, Using WiFi connection counts and camera-based occupancy counts to estimate and predict building occupancy, Energy Build. 257 (2022), https://doi.org/10.1016/j.enbuild.2021.111759.

[17] Y. Yang, Y. Yuan, T. Pan, X. Zang, G. Liu, A framework for occupancy prediction based on image information fusion and machine learning, Build Environ. 207 (2022), https://doi.org/10.1016/j.buildenv.2021.108524.

[18] L. Zimmermann, R. Weigel, G. Fischer, Fusion of nonintrusive environmental sensors for occupancy detection in smart homes, IEEE Internet Things J. 5 (2018) 2343–2352, https://doi.org/10.1109/JIOT.2017.2752134.

[19] H. Zhou, J. Yu, Y. Zhao, C. Chang, J. Li, B. Lin, Recognizing occupant presence status in residential buildings from environment sensing data by data mining approach, Energy Build. 252 (2021), https://doi.org/10.1016/j.enbuild.2021.111432.

[20] K. Kornbluth, L. Slaughter, S. Gul, S.R. Pamireddy, A. Meier, Identifying energy savings opportunities in vacant commercial buildings using a semi-supervised sensor fusion model, Energy Build. 265 (2022), https://doi.org/10.1016/j.enbuild.2022.112084.

[21] H.P. Díaz-Hernández, M.N. Sánchez, R. Olmedo, M.M. Villar-Ramos, E.V. Macias-Melo, K.M. Aguilar-Castro, M.J. Jiménez, Performance assessment of different measured variables from onboard monitoring system to obtain the occupancy patterns of rooms in an office building, J. Build. Eng. 40 (2021), https://doi.org/10.1016/j.jobe.2021.102676.

[22] J. Ma, Internet-of-things: technology evolution and challenges, in, IEEE MTT-S International Microwave Symposium (IMS2014) IEEE 2014 (2014) 1–4, https://doi.org/10.1109/MWSYM.2014.6848429.

[23] Y. Zhou, J. Chen, Z.J. Yu, J. Li, G. Huang, F. Haghighat, G. Zhang, A novel model based on multi-grained cascade forests with wavelet denoising for indoor occupancy estimation, Build Environ. 167 (2020), https://doi.org/10.1016/j.buildenv.2019.106461.

[24] M.K. Masood, C. Jiang, Y.C. Soh, A novel feature selection framework with hybrid feature-scaled extreme learning machine (HFS-ELM) for indoor occupancy estimation, Energy Build. 158 (2018) 1139–1151, https://doi.org/10.1016/j.enbuild.2017.08.087.

[25] Q. Zhu, Z. Chen, M.K. Masood, Y.C. Soh, Occupancy estimation with environmental sensing via non-iterative LRF feature learning in time and frequency domains, Energy Build. 141 (2017) 125–133, https://doi.org/10.1016/j.enbuild.2017.01.057.

[26] M.K. Masood, Y.C. Soh, C. Jiang, Occupancy estimation from environmental parameters using wrapper and hybrid feature selection, Appl. Soft Comput. J. 60 (2017) 482–494, https://doi.org/10.1016/j.asoc.2017.07.003.

[27] Z. Chen, M.K. Masood, Y.C. Soh, A fusion framework for occupancy estimation in office buildings based on environmental sensor data, Energy Build. 133 (2016) 790–798, https://doi.org/10.1016/j.enbuild.2016.10.030.

[28] Z. Chen, R. Zhao, Q. Zhu, M.K. Masood, Y.C. Soh, K. Mao, Building occupancy estimation with environmental sensors via CDBLSTM, IEEE Trans. Ind. Electron. 64 (2017) 9549–9559, https://doi.org/10.1109/TIE.2017.2711530.

[29] C. Feng, A. Mehmani, J. Zhang, Deep learning-based real-time building occupancy detection using AMI data, IEEE Trans Smart Grid. 11 (2020) 4490–4501, https://doi.org/10.1109/TSG.2020.2982351.

[30] D. Stjelja, J. Jokisalo, R. Kosonen, Scalable room occupancy prediction with deep transfer learning using indoor climate sensor, Energies (basel). 15 (2022), https://doi.org/10.3390/en15062078.

[31] A. Mohammadabadi, S. Rahnama, A. Afshari, Indoor occupancy detection based on environmental data using CNN-XGboost model: experimental validation in a residential building, Sustainability. 14 (2022) 14644, https://doi.org/10.3390/su142114644.

[32] B. Huchuk, S. Sanner, W. O'Brien, Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data, Build Environ. 160 (2019), https://doi.org/10.1016/j.buildenv.2019.106177.

[33] Z. Wang, T. Hong, M.A. Piette, M. Pritoni, Inferring occupant counts from Wi-Fi data in buildings through machine learning, Build Environ. 158 (2019) 281–294, https://doi.org/10.1016/j.buildenv.2019.05.015.

[34] B. Howard, S. Acha, N. Shah, J. Polak, Implicit sensing of building occupancy count with information and communication technology data sets, Build Environ. 157 (2019) 297–308, https://doi.org/10.1016/j.buildenv.2019.04.015.

[35] S. Kim, S. Kang, K.R. Ryu, G. Song, Real-time occupancy prediction in a large exhibition hall using deep learning approach, Energy Build. 199 (2019) 216–222, https://doi.org/10.1016/j.enbuild.2019.06.043.

[36] B.W. Hobson, D. Lowcay, H.B. Gunay, A. Ashouri, G.R. Newsham, Opportunistic occupancy-count estimation using sensor fusion: a case study, Build Environ. 159 (2019), https://doi.org/10.1016/j.buildenv.2019.05.032.

[37] A.G. Alam, H. Rahman, J.K. Kim, H. Han, Uncertainties in neural network model based on carbon dioxide concentration for occupancy estimation, J. Mech. Sci. Technol. 31 (2017) 2573–2580, https://doi.org/10.1007/s12206-017-0455-z.

[38] M. Parzinger, L. Hanfstaengl, F. Sigg, U. Spindler, U. Wellisch, M. Wirnsberger, Comparison of different training data sets from simulation and experimental measurement with artificial users for occupancy detection — using machine learning methods Random Forest and LASSO, Build Environ. 223 (2022), https://doi.org/10.1016/j.buildenv.2022.109313.

[39] B. Dong, B. Andrews, K.P. Lam, M. Höynck, R. Zhang, Y.S. Chiou, D. Benitez, An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network, Energy Build. 42 (2010) 1038–1046, https://doi.org/10.1016/j.enbuild.2010.01.016.

[40] Z. Chen, Y. Yang, C. Jiang, J. Hao, L. Zhang, Light sensor based occupancy estimation via bayes filter with neural networks, IEEE Trans. Ind. Electron. 67 (2020) 5787–5797, https://doi.org/10.1109/TIE.2019.2934028.

[41] L.M. Candanedo, V. Feldheim, Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models, Energy Build. 112 (2016) 28–39, https://doi.org/10.1016/j. enbuild.2015.11.071.

[42] G. Fajilla, M. Chen Austin, D. Mora, M. de Simone, Assessment of probabilistic models to estimate the occupancy state in office buildings using indoor parameters and user-related variables, Energy Build. 246 (2021). https://doi.org/10.1016/j. enbuild.2021.111105.

[43] L.M. Candanedo, V. Feldheim, D. Deramaix, A methodology based on Hidden Markov Models for occupancy detection and a case study in a low energy residential building, Energy Build. 148 (2017) 327–341, https://doi.org/10.1016/ j.enbuild.2017.05.031.

[44] C. Jiang, Z. Chen, R. Su, M.K. Masood, Y.C. Soh, Bayesian filtering for building occupancy estimation from carbon dioxide concentration, Energy Build. 206 (2020), https://doi.org/10.1016/j.enbuild.2019.109566.

[45] Z. Chen, Q. Zhu, M.K. Masood, Y.C. Soh, Environmental sensors-based occupancy estimation in buildings via IHMM-MLR, IEEE Trans Industr Inform. 13 (2017) 2184–2193, https://doi.org/10.1109/TII.2017.2668444.

[46] Z.D. Tekler, A. Chong, Occupancy prediction using deep learning approaches across multiple space types: a minimum sensing strategy, Build Environ. 226 (2022), https://doi.org/10.1016/j.buildenv.2022.109689.

[47] W. Wang, J. Chen, T. Hong, Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings, Autom Constr. 94 (2018) 233–243, https://doi.org/10.1016/j.autcon.2018.07.007.

[48] B. Dong, Y. Liu, W. Mu, Z. Jiang, P. Pandey, T. Hong, B. Olesen, T. Lawrence, Z. O'Neil, C. Andrews, E. Azar, K. Bandurski, R. Bardhan, M. Bavaresco, C. Berger, J. Burry, S. Carlucci, K. Chvatal, M. de Simone, S. Erba, N. Gao, L.T. Graham, C. Grassi, R. Jain, S. Kumar, M. Kjærgaard, S. Korsavi, J. Langevin, Z. Li, A. Lipczynska, A. Mahdavi, J. Malik, M. Marschall, Z. Nagy, L. Neves, W. O'Brien, S. Pan, J.Y. Park, I. Pigliautile, C. Piselli, A.L. Pisello, H.N. Rafsanjani, R.F. Rupp, F. Salim, S. Schiavon, J. Schwee, A. Sonta, M. Touchie, A. Wagner, S. Walsh, Z. Wang, D.M. Webber, D. Yan, P. Zangheri, J. Zhang, X. Zhou, X. Zhou, A global building occupant behavior database, Sci Data. 9 (2022), https://doi.org/10.1038/ s41597-022-01475-3.

[49] D. Mora, G. Fajilla, M.C. Austin, M. de Simone, Occupancy patterns obtained by heuristic approaches: cluster analysis and logical flowcharts. A case study in a university office, Energy Build. 186 (2019) 147–168, https://doi.org/10.1016/j. enbuild.2019.01.023.

[50] Z. Yu, L. Jia, M.C. Murphy-Hoye, A. Pratt, L. Tong, Modeling and stochastic control for home energy management, IEEE Trans Smart Grid. 4 (2013) 2244–2255, https://doi.org/10.1109/TSG.2013.2279171.