Building and Environment 252 (2024) 111299

Contents lists available at ScienceDirect

Building and Environment

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

# Performance evaluation of an occupant metabolic rate estimation algorithm using activity classification and object detection models

Ji Young Yun, Eun Ji Choi, Min Hee Chung, Kang Woo Bae, Jin Woo Moon *

*School of Architecture and Building Science, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, Seoul, 06974, Republic of Korea*

ARTICLE INFO

*Keywords:*
Metabolic rate
Object detection model
Occupant-centric control
Thermal comfort
MET estimation algorithm

ABSTRACT

To create a comfortable indoor environment, the metabolic rate (MET), which affects the thermal sensation of occupants, needs to be reflected in real-time. Recently, methods employing computer vision techniques classify activities based on the pose of the body in images. However, these methods face challenges in determining the MET depending on the objects used, even with the same pose. Therefore, the objective of this study is to develop a MET estimation algorithm that can estimate various METs by integrating a pose-based activity classification model and an object detection model. To achieve this, an object detection model capable of detecting and classifying six regularly used objects indoors was developed, and a performance evaluation was conducted. The MET estimation algorithm was assessed through the implementation of a thermal control system, validating its applicability in experimental settings. As a result, the object detection model exhibited a real-time classification accuracy of 89%. Additionally, when evaluating the mode value over 15-s intervals, it demonstrated a classification accuracy of 100%. The algorithm exhibited a real-time estimation accuracy of 83% for the six METs and examining the mode value for 15-s intervals, it demonstrated a classification accuracy of 99%. This study thus confirmed the control capability of the proposed MET estimation algorithm and its potential for the estimation of various METs. The developed method can be used for the real-time estimation of occupant thermal comfort in indoor comfort-based control systems, contributing to the realization of a comfortable environment for occupants that protects their well-being.

## 1. Introduction

### *1.1. Background*

As a consequence of urbanization, many more people worldwide are spending a significant portion of their day indoors [1–3]. Consequently, it has become important to create an appropriate indoor thermal environment that affects the health, productivity, and overall well-being of occupants [4–6]. However, traditional building thermal control methods have typically relied on temperature setpoints based solely on the ambient temperature or fixed assumptions regarding occupant clothing and activity levels [7,8]. This approach fails to reflect the actual information of the occupants, often leading to unnecessary energy consumption and discomfort. There is growing recognition of the importance of incorporating occupant information into building control systems, thus various studies have been conducted to create a comfortable indoor thermal environment [9–12].

In particular, there is growing interest in occupant-centric control (OCC), which aims to enhance occupant thermal comfort while efficiently utilizing energy resources by taking into account occupant information such as their presence, occupancy, clothing insulation (CLO), and metabolic rate (MET) [13–16]. Ouf et al. proposed a simulation framework for optimized OCC based on stochastic occupant behavior and building design. Simulation results illustrated that the consideration of occupants' preferences when configuring the OCC enhances performance. The importance of statistically modeling occupants' data and integrating it effectively into the OCC was emphasized [17]. Pang et al. conducted energy simulations to investigate the energy-saving potential of OCC for office buildings. They found that utilizing occupant presence and occupancy information for building control could reduce energy use by 19–45% compared to the standards outlined in ASHRAE Standard 90.1–2004 [18]. Xie et al. reviewed research on OCC from the perspective of sensing, predicting, and controlling technologies. OCC demonstrated an average energy savings of 22%, with a 29.1% improvement in thermal comfort, highlighting its potential to maximize

* Corresponding author.
*E-mail addresses:* yjyyjy5350@cau.ac.kr (J.Y. Yun), ejjchl77@cau.ac.kr (E.J. Choi), mhloveu@cau.ac.kr (M.H. Chung), rkddn2@cau.ac.kr (K.W. Bae), gilerbert73@cau.ac.kr (J.W. Moon).

https://doi.org/10.1016/j.buildenv.2024.111299
Received 22 December 2023; Received in revised form 6 February 2024; Accepted 12 February 2024
Available online 13 February 2024

**Nomenclature**

| | |
|---|---|
| OCC | Occupant-centric control |
| MET | Metabolic rate [met or $W/m^2$] |
| PMV | Predicted mean vote |
| HPE | Human pose estimation |
| CNN | Convolutional neural network |
| CLO | Clothing insulation [clo or $m^2$•K/W] |
| ASHRAE | American Society of Heating, Refrigerating and Air-Conditioning Engineers |
| ISO | International Organization for Standardization |
| DNN | Deep neural network |
| YOLO | You Only Look Once |

occupant thermal comfort while conserving energy within the building [19].

The thermal sensation experienced by an occupant in OCC is quantitatively evaluated using the predicted mean vote (PMV) [20–22]. The PMV is calculated based on four environmental variables (indoor temperature, relative humidity, mean radical temperature, and air quality) and two individual variables (CLO and MET). Generally, when the PMV is between −0.5 and 0.5, the occupant is considered comfortable [23]. The PMV serves as a representative indicator of comfort, providing valuable information for indoor environmental design, operation, and comfort control. Achieving accurate PMV values during thermal environmental control can enhance the comfort of indoor spaces, thus the variables influencing the PMV should be objectively calculated to ensure their accuracy [24–27].

The four environmental PMV variables can be measured using sensors, but the two personal factors are difficult to measure objectively and directly. Consequently, various approaches have been proposed to measure these factors [28–30], of which the MET represents a particularly important variable influencing thermal comfort. The MET represents the heat production of the human body, which varies based on the activity being performed; for example, a person has a MET of 50 kcal/$m^2h$ when they are seated and at rest [23]. Therefore, in order to create a comfortable indoor thermal environment, various methods have been developed to estimate the MET of occupants [31–34].

### 1.2. Previous research on the measurement of the MET

A variety of methods have been established to measure the MET. For example, ISO 8996 specifies four levels and eight methods for MET measurement (Table 1) [35]. The simplest approach involves using approximate information inferred from the occupant's occupation or making direct observations. Other methods directly measure the occupant's heart rate, oxygen consumption, and/or caloric expenditure to produce a more precise estimate of activity levels. Luo et al. calculated the MET by measuring the oxygen consumption and carbon dioxide production of the human body using a Vmax Encore metabolic cart to experimentally determine the effects of thermal environmental conditions such as the indoor temperature and CLO [36]. Zhai et al. also used a wireless Cosmed K5 wearable metabolic system, which measures the gas exchange in each breath, to determine the MET for sitting, standing, and walking, activities that usually occur in the office [37]. Ji et al. measured carbon dioxide levels by installing transmitters and sensors in an airtight chamber and used the results to evaluate the change in the MET and thermal comfort for a person who is moving [38]. In general, these methods for measuring the MET require direct intervention in the activity of the occupant and/or expensive equipment, making them difficult to measure in an unspecified space in which a large number of people are present.

In order to determine the MET of occupants in a real building, an indirect and objective measurement method is required. For this reason, computer vision approaches that extract information from images have received specific attention, demonstrating suitable accuracy and usability in the field [39–45]. For example, Liu et al. estimated the MET using the random forest machine learning approach for images collected from thermal cameras, demonstrating that the MET of various individuals can be estimated in an uncontrolled multi-room indoor environment [40]. Na et al. also developed an MET calculation model based on the deep learning of images and data using a heart rate sensor and a Kinect camera [41], while Choi et al. established a model that could simultaneously estimate the MET and CLO using deep vision, which combines deep learning and computer vision, and employed this in a comfort-based control experiment. Their control experiments were conducted by classifying activities as either sitting or standing, showing that the occupant's thermal comfort could be effectively improved [42].

The most widely used computer vision method for calculating the MET is human pose estimation (HPE), which was developed as an activity classification model that classifies a person's pose by recognizing the main joint coordinates in an image [43–45]. Choi et al. demonstrated that 10 types of indoor activity can be classified by estimating joint coordinates from images using a deep neural network (DNN) model [44]. Kim et al. also developed a MET estimation model using OpenPose and a DNN and conducted an experiment with multiple occupants to demonstrate its suitability for use in a real building [45]. Because HPE utilizes cameras, making it easy to employ in a real building environment. Though HPE has been found to be effective for objectively calculating the MET while minimizing occupant intervention, similar poses are not easily distinguished, thus different activities can be misclassified as the same activity (Fig. 1). Therefore, in order to calculate various METs, a new model is needed that considers other information in addition to pose information, leading to the accurate identification of the activities of the occupants.

### 1.3. Research purpose

When classifying various activities from images using HPE, information about the objects being used can play a crucial role. Therefore, this study proposes a novel method for simultaneously classifying activities and objects in order to produce accurate estimates of the MET for various activities. To achieve this, a new MET estimation algorithm was developed that considered the pose of the occupant and the object(s) used by the occupant. The proposed algorithm included an object

**Table 1**
Levels for the determination of the metabolic rate (ISO 8996).

| Level | Method | Accuracy |
|---|---|---|
| 1. Screening | 1A: Classification according to occupation<br>1B: Classification according to activity | Rough information Very great risk of error |
| 2. Observation | 2A: Group assessment tables<br>2B: Tables for specific activities | High error risk Accuracy: ±20% |
| 3. Analysis | Heart rate measurement under defined conditions | Medium error risk Accuracy: ±10% |
| 4. Expertise | 4A: Measurement of oxygen consumption<br>4B: Doubly labeled water method<br>4C: Direct calorimetry | Errors within the limits of the accuracy of the measurement or of the time and motion study Accuracy: ±5% |

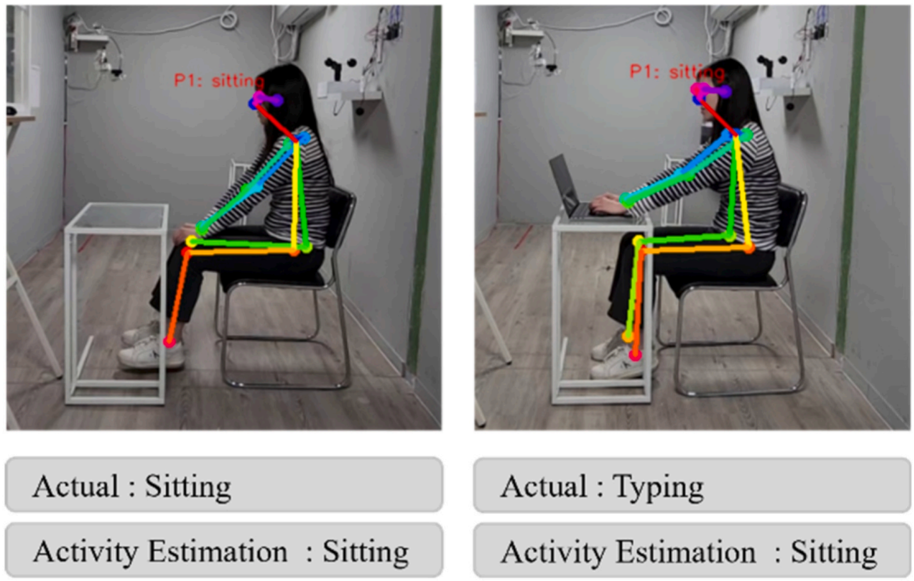



**Fig. 1.** Different activities detected for similar poses.

detection model that detected the object used by an occupant and integrated it with an existing activity classification model. Experimental assessments of the proposed MET estimation algorithm were then conducted to assess its practical utility in real building environments.

## 2. Methods

In the present study, MET estimates for various activities were produced by developing an integrated MET estimation algorithm that combined an activity classification model and an object detection model. As presented in Fig. 2, this process consisted of three stages. The activity classification stage builds on the HPE-based occupant activity classification model developed by Kim et al. [45]. In the object detection stage, objects in use by the occupants detected in the activity classification stage are identified. Finally, in the MET estimation stage, the results from the activity classification and object detection stages are combined to produce a range of MET estimates.

### 2.1. Activity classification stage

The activity classification model employed in the present study to distinguish sitting, standing, and walking was previously developed by Kim et al. [45]. This model used indoor RGB images to classify these three activities based on the pose of the occupant and consisted of two steps: joint recognition and activity classification. In the joint recognition step, HPE technology was employed to recognize the posture of individuals. In particular, the real-time extraction of the joint coordinates of individuals detected within the video images was achieved using the OpenPose library, which is based on a convolutional neural network (CNN). In the activity classification step, a DNN is utilized to train and classify the human activities based on the previously extracted coordinates. Fig. 3 illustrates these steps using example images.

In Kim et al.'s study, a test-bed experiment was conducted to assess the model's performance, focusing on a single occupant [45]. The three activities (sitting, standing, and walking) were alternated every 5 min, and images were captured at 1-s intervals using camera sensors for a total time of 15 min. Real-time activity classification demonstrated an overall average accuracy of 72.7%. To resolve real-time activity classification errors, the activities were predicted based on the mode value for 1 min and 5 min intervals, resulting in an accuracy of 100% across all activities.

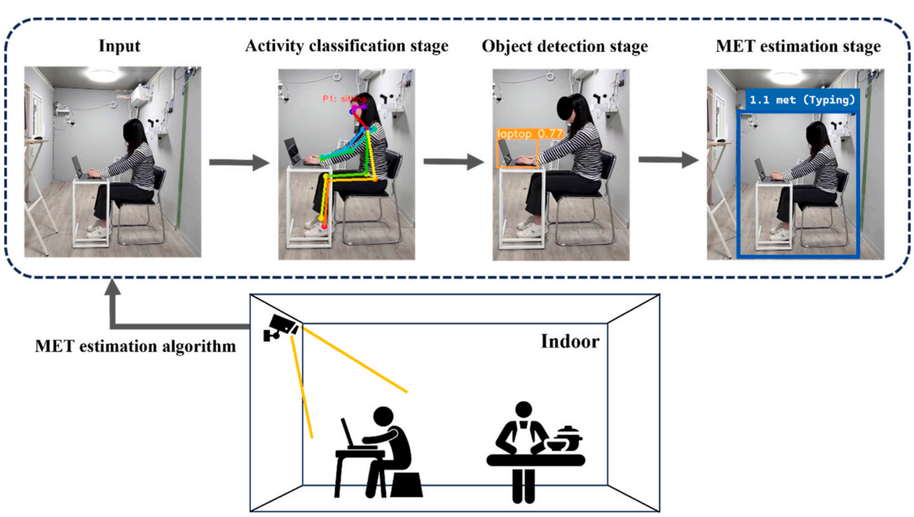



**Fig. 2.** Process used for the MET estimation algorithm.

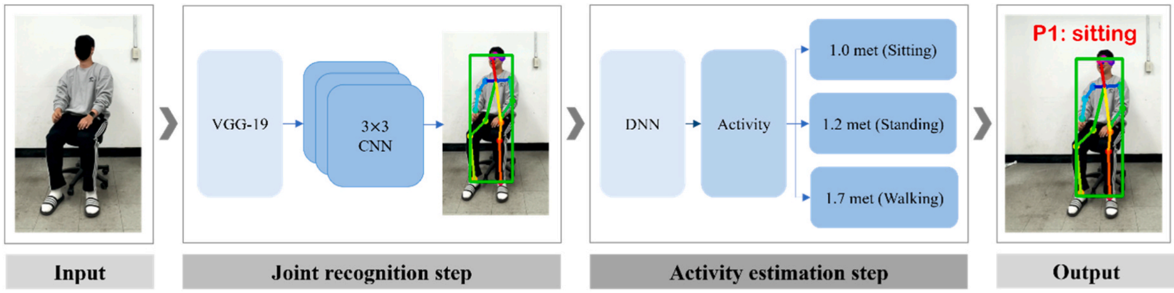



**Fig. 3.** Activity classification process [45].

### 2.2. Object detection stage

The object detection model was trained using You Only Look Once (YOLO) v5, an image-processing deep learning network designed for real-time object detection [46]. YOLOv5 employs a single-stage approach to object detection and classification through the same neural network structure, providing a highly efficient and real-time object recognition technology with fast inference processes [47]. YOLOv5 allows for real-time measurements and does not require direct intervention in the occupant's activity. Therefore, it is widely used in OCC approaches to collect information about the occupants, such as the number of occupants in the building, CLO, and MET [48,49]. Fig. 4 illustrates the object-detection process for the YOLOv5-trained model using indoor images. In these images, a bounding box centered around the occupant recognized in the activity classification stage is employed. This method effectively eliminates unnecessary background information and significantly reduces the likelihood of errors by detecting only the objects used by the occupant.

The object detection model was developed to classify activities that involved the use of an object, such as typing, cooking, and cleaning. The system platform for model training included Ubuntu (Linux-5.11.0-generic), a Quadro RTX 8000 GPU, and Python 3.7.13. The types of objects learned and the number of learning data are shown in Table 2. The image dataset used for training was collected from internet searches and direct photography. The image dataset comprised a total of 2359 images, each configured to include two objects typically detectable during the performance of the respective activity. For typing, the dataset included images of a laptop and a keyboard, cooking images included a knife and a cutting board, and cleaning images featured a broom and a vacuum cleaner. Images of laptops (636), keyboards (452), knives (669), cutting boards (668), brooms (232), and vacuum cleaners (282) were collected for each activity and utilized in the training process. Of these images, 70% were divided randomly for training, 15% for verification, and 15% for testing. For the brooms and vacuum cleaners, which were associated with cleaning, many of the images on the Internet were very similar, so data was collected by photographing these objects directly to reduce the possibility of overfitting. This is why the number of images for these two objects was lower than for the other objects. The performance and stability of the object detection model for all objects were reviewed by conducting accuracy evaluation tests in actual experiments (Section 3.1).

The learning results were used to evaluate the performance of the classification model, with a confusion matrix constructed to determine the classification accuracy (Fig. 5). When evaluated based on previously classified test data, the classification accuracy for all six objects was higher than 80%. The real-time accuracy for each object is also presented in Table 2.

### 2.3. MET estimation algorithm

The MET estimation algorithm designed to predict various METs using both activity classification and object detection is illustrated in Fig. 6. In the activity classification stage, OpenPose was employed to detect occupants within indoor videos or images and initially identify one of three activities: sitting, standing, or walking. Subsequently, cropped images containing only the detected occupants were created and stored. If no occupants were detected during this process, the algorithm ended without further action. The object detection stage utilizes the object detection model to recognize the six target objects used by the detected occupants in the cropped images obtained from the activity classification stage. Based on this, the final activity was determined and the corresponding MET estimated.

The MET values used in the present study ranged from 1.0 to 2.7 met, with walking restricted to speeds below 2.5 km/h, corresponding to 1.7 met, following ASHRAE Standard 55. Additionally, while typing, cooking, and cleaning may involve a range of poses in reality, assumptions were made based on the typical posture associated with each activity, i.e., sitting for typing, standing for cooking, and walking for cleaning. ISO 7730 [50] and ASHRAE Standard 55 [23] provide specific ranges for certain activity categories, suggesting a range of 1.6–2.0 met for cooking and 2.0–3.4 met for cleaning. However, they do not distinguish between detailed activity types or the use of objects. In this study, based on ISO 7730 [50] and ASHRAE Standard 55 [23], cooking was assumed to involve typical cooking objects in a standing posture with an estimated

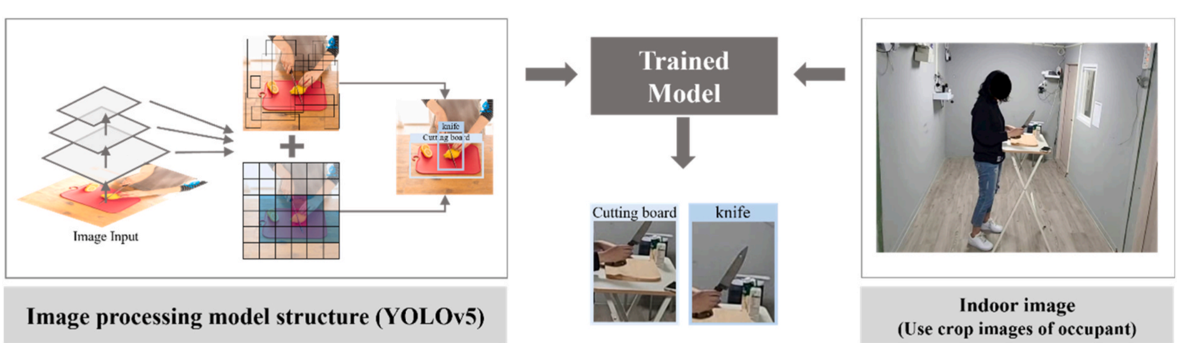



**Fig. 4.** Object detection process for indoor images using YOLOv5.

**Table 2**
Number of images and real-time accuracy for each object.

| Object type | Laptop | Keyboard | Knife | Cutting board | Broom | Vacuum cleaner |
|---|---|---|---|---|---|---|
| Activity | Typing | Typing | Cooking | Cooking | Cleaning | Cleaning |
| Number of images | 636 | 452 | 669 | 668 | 232 | 282 |
| Real-time accuracy | 94% | 82% | 75% | 73% | 81% | 82% |

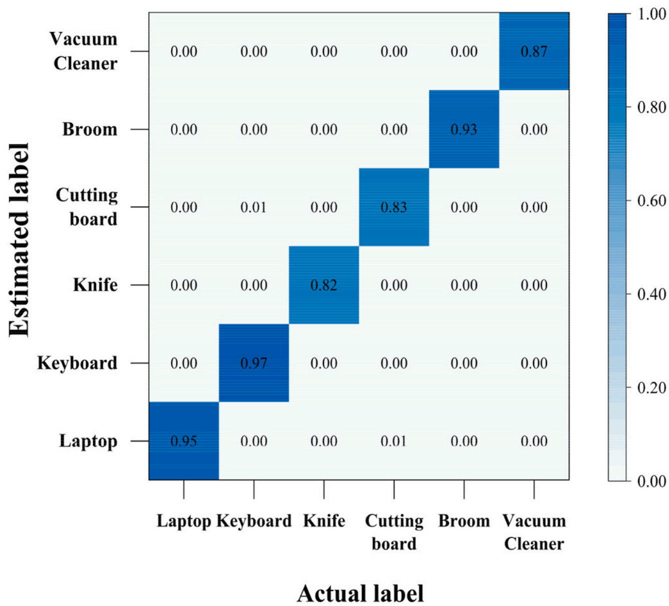



**Fig. 5.** Confusion matrix for the object detection model.

MET of 1.8 met, and cleaning was assumed to involve common cleaning objects in a walking posture with an estimated 2.7 met. Nevertheless, for an accurate reflection of activity levels in real buildings, it is necessary to refine the activity classification system through prior measurements using equipment such as thermal manikins to obtain more precise values.

In the MET estimation stage of the algorithm, when a laptop or keyboard was detected in a sitting posture, the algorithm estimated 1.1 met, corresponding to typing. If no object was detected, 1.0 met was estimated, corresponding to sitting. In a standing posture, if a knife or cutting board was detected, the algorithm estimated 1.8 met, corresponding to cooking. If no object was detected, it returned 1.2 met, corresponding to standing. In a walking posture, if a broom or vacuum cleaner was detected, the algorithm estimated 2.7 met, corresponding to cleaning. If no object was detected, it estimated 1.7 met, corresponding to walking. If it was not possible to calculate a value based on the combinations in Fig. 6, the process concluded without any output.

## 2.4. Experimental process

### 2.4.1. Test-bed information

The developed MET estimation algorithm was evaluated for its utility in real environments in an experiment. The performance of the object detection model embedded in the MET estimation algorithm was also evaluated. The experiments were conducted in a space measuring 8.4 m × 4.2 m located at University C in Seoul. The experiments involved the use of a new space and objects not included in the training dataset. Fig. 7 schematically presents the experimental setup. This study, conducted after obtaining Institutional Review Board (IRB) approval, secured the participants' consent before the experiment began. The experiment was conducted with a single occupant positioned in the center of the experimental space, utilizing tripods installed at a height of 1.7 m on the left and right sides, each equipped with cameras.

### 2.4.2. Experimental setting for the MET estimation algorithm

Prior to the evaluation of the MET estimation algorithm, the on-site performance of the trained object detection model was verified using participants engaged in experiments using six objects related to typing, cooking, and cleaning. The experiment focused on a single subject, employing two objects for each activity, as outlined in Table 2. The subject performed each activity with the designated object for 5 min, amounting to six repetitions. Image data were captured at 1-s intervals, resulting in 600 data points per object and a total of 3600 data points. The object detection model evaluation conducted real-time classification performance at 1-s intervals and performance based on mode data over 15-s.

To evaluate the performance of the MET estimation algorithm, two experiments were conducted. The first experiment involved the algorithm facing each of the six METs consistently for an equal duration (the case-based experiment). The second experiment assessed the utility of the algorithm when the MET changed by freely alternating activities for 30 min after dividing the six METs into two categories (the scenario-based experiment).

The accuracy of the proposed algorithm was evaluated by engaging in the same specific activity for a specific duration. However, in real buildings, occupants are not always involved in a single activity continuously. Therefore, to assess the utility of our proposed approach for buildings, different scenarios were established in which the activity (and thus the MET) of the occupants randomly changed. Scenarios were defined by randomly categorizing the previous six activities into two groups, and participants were instructed to freely engage in the activities for each scenario over a period of 30 min. The activities used for each scenario are illustrated in Fig. 8. Because the mode value for intervals of at least 15 s needed to be recorded, each activity was performed for a minimum of 2 min. The experimental conditions were consistent with those of the object detection model. Fig. 8 presents the actual activities of the participants during the experiment. In Scenario 2, the transition from typing to cleaning or from cleaning to typing, where the participant sits down to organize a laptop and desk, is classified as sitting. The performance assessment of the MET estimation algorithm for each scenario entailed analyzing the real-time accuracy of MET estimation at 1-s intervals as well as the accuracy based on mode values at 15-s and 30-s intervals.

# 3. Results and discussion

## 3.1. Performance of the object detection model

The classification accuracy of the object detection model using YOLOv5 was evaluated to confirm that it could be used in actual control systems. The classification accuracy for each object is summarized in Table 3, and Fig. 9 illustrates an example of the objects monitored in the experiment. The average classification accuracy for the six objects was 89%. The keyboard and laptop, which were associated with typing, had accuracies of 92% and 95%, respectively, because both objects have well-defined shapes. The knife and cutting board had accuracies of 85% and 90%, respectively, with errors occurring with the knife due to its dynamic usage at various angles. Additionally, the shape of the cutting board occasionally led to confusion with the table. The broom and vacuum cleaner, used for cleaning, exhibited accuracies of 82% and 92%, respectively, with errors mainly arising from dynamic poses during the use of these objects, leading to occlusions, which was in contrast to

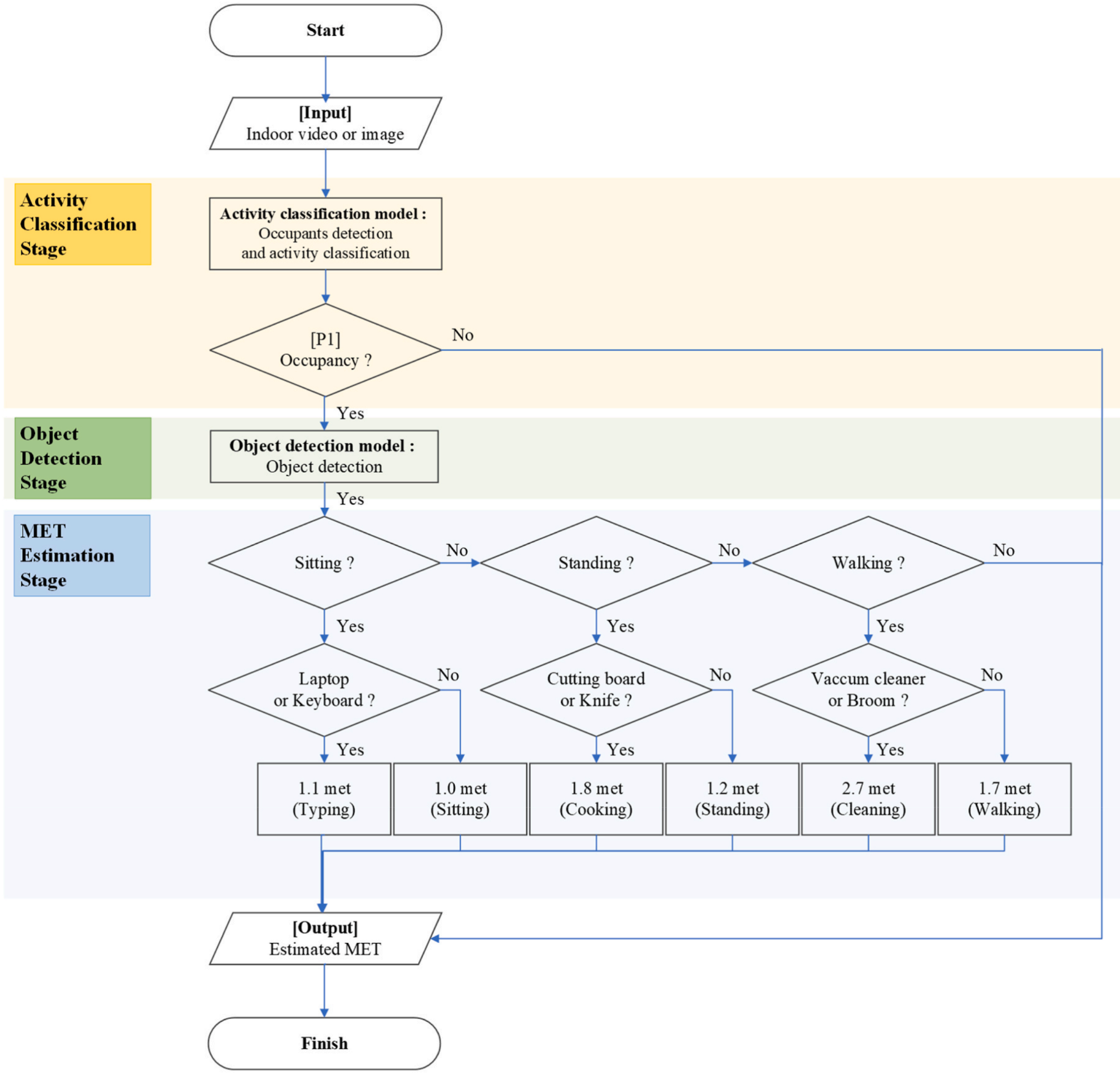



**Fig. 6.** MET estimation algorithm.

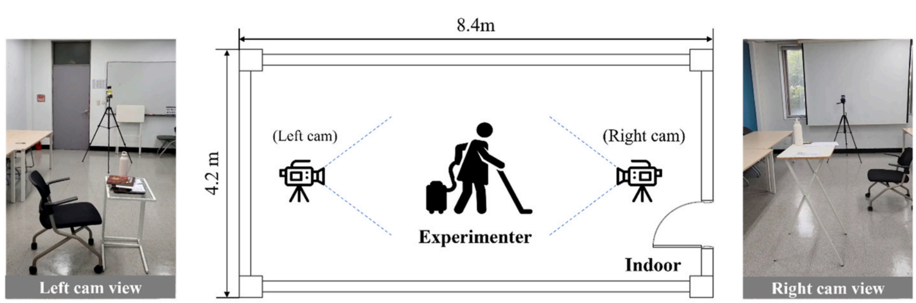



**Fig. 7.** Experimental setup.

typing or cooking, where the objects were typically static. On average, the accuracy for typing was 93.5%, for cooking was 87.5%, and for cleaning was 85.0%. Focusing on training images with diverse angles and scenarios where people are actively using the objects, especially for objects associated with dynamic METs, can reduce errors.

In this case, to improve the usability of the learning model in buildings, it is necessary to calculate a single value representing a certain period of time, rather than reflecting the estimated MET in real-time control. Therefore, in the present study, the mode of accumulated values over a designated period was employed to mitigate instantaneous errors that may occur during real-time object detection. This approach can also be employed to determine the activity level and PMV that represent the control cycle in actual building control systems. To determine the feasibility of this, the mode value for 15-s intervals was examined in the experiment, demonstrating a 100% classification accuracy across all objects.

<table>
<tr><th colspan="20">15 second mode value of Scenario 1</th></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td></tr>
<tr><td colspan="14">1.2 met (Standing)</td><td colspan="6"></td></tr>
<tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td></tr>
<tr><td colspan="15">1.8 met (Cooking)</td><td colspan="5"></td></tr>
<tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>47</td><td>48</td><td>49</td><td>50</td><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>57</td><td>58</td><td>59</td><td>60</td></tr>
<tr><td colspan="20">1.0 met (Sitting)</td></tr>
<tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>67</td><td>68</td><td>69</td><td>70</td><td>71</td><td>72</td><td>73</td><td>74</td><td>75</td><td>76</td><td>77</td><td>78</td><td>79</td><td>80</td></tr>
<tr><td colspan="13">1.8 met (Cooking)</td><td colspan="7">1.2 met (Standing)</td></tr>
<tr><td>81</td><td>82</td><td>83</td><td>84</td><td>85</td><td>86</td><td>87</td><td>88</td><td>89</td><td>90</td><td>91</td><td>92</td><td>93</td><td>94</td><td>95</td><td>96</td><td>97</td><td>98</td><td>99</td><td>100</td></tr>
<tr><td colspan="5"></td><td colspan="12">1.8 met (Cooking)</td><td colspan="3"></td></tr>
<tr><td>101</td><td>102</td><td>103</td><td>104</td><td>105</td><td>106</td><td>107</td><td>108</td><td>109</td><td>110</td><td>111</td><td>112</td><td>113</td><td>114</td><td>115</td><td>116</td><td>117</td><td>118</td><td>119</td><td>120</td></tr>
<tr><td colspan="12">1.0 met (Sitting)</td><td colspan="8">1.2 met (Standing)</td></tr>
<tr><td colspan="4">MET Type</td><td colspan="16">1.0 met (sitting), 1.2 met (standing), 1.8 met (cooking)</td></tr>
<tr><td colspan="4">Times</td><td colspan="16">1.0 met (Sitting) : 40 / 1.2 met (Standing) : 34 / 1.8 met (Cooking) : 46</td></tr>
</table>

(a) Scenario 1

<table>
<tr><th colspan="20">15 second mode value of Scenario 2</th></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td></tr>
<tr><td colspan="15">(1.1 met) Typing</td><td colspan="5"></td></tr>
<tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td></tr>
<tr><td colspan="8">1.7 met (Walking)</td><td colspan="9">2.7 met (Cleaning)</td><td colspan="3"></td></tr>
<tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>47</td><td>48</td><td>49</td><td>50</td><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>57</td><td>58</td><td>59</td><td>60</td></tr>
<tr><td colspan="10">1.7 met (Walking)</td><td colspan="10">(1.1 met) Typing</td></tr>
<tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>67</td><td>68</td><td>69</td><td>70</td><td>71</td><td>72</td><td>73</td><td>74</td><td>75</td><td>76</td><td>77</td><td>78</td><td>79</td><td>80</td></tr>
<tr><td colspan="2"></td><td colspan="16">2.7 met (Cleaning)</td><td>1.0 met</td><td></td></tr>
<tr><td>81</td><td>82</td><td>83</td><td>84</td><td>85</td><td>86</td><td>87</td><td>88</td><td>89</td><td>90</td><td>91</td><td>92</td><td>93</td><td>94</td><td>95</td><td>96</td><td>97</td><td>98</td><td>99</td><td>100</td></tr>
<tr><td colspan="9">(1.1 met) Typing</td><td>1.0 met</td><td colspan="10">2.7 met (Cleaning)</td></tr>
<tr><td>101</td><td>102</td><td>103</td><td>104</td><td>105</td><td>106</td><td>107</td><td>108</td><td>109</td><td>110</td><td>111</td><td>112</td><td>113</td><td>114</td><td>115</td><td>116</td><td>117</td><td>118</td><td>119</td><td>120</td></tr>
<tr><td colspan="20">1.7 met (Walking)</td></tr>
<tr><td colspan="4">MET Type</td><td colspan="16">1.0 met (sitting), 1.2 met (standing), 1.8 met (cooking)</td></tr>
<tr><td colspan="4">Times</td><td colspan="16">1.0 met (Sitting) : 2 / 1.1 met (Typing) : 37 / 1.7 met (Walking) : 46 / 2.7 met (Cleaning) : 35</td></tr>
</table>

(b) Scenario 2

**Fig. 8.** Experimental activities in Scenarios 1 and 2.

**Table 3**
Accuracy of the proposed object detection model.

| Object | Real-time | 15-s mode value |
|---|---|---|
| Keyboard | 92% | 100% |
| Laptop | 95% | 100% |
| Knife | 85% | 100% |
| Cutting board | 90% | 100% |
| Broom | 82% | 100% |
| Vacuum cleaner | 92% | 100% |

### *3.2. Performance of the MET estimation algorithm*

#### *3.2.1. Case-based experiment*

The performance of the MET estimation algorithm was evaluated through experimentation in a test environment identical to that used to test the performance of the object detection model. In the first experiment, activities corresponding to six target METs were performed individually. Each activity was conducted for 10 min, with the two objects associated with each of the typing, cooking, and cleaning activities utilized for 5 min each to ensure uniform object exposure. The accuracy of the algorithm was analyzed in terms of both the real-time classification accuracy and the mode of the cumulative data at 15-s and 30-s intervals. By calculating representative values, errors arising from short-term phenomena such as image shaking and focus issues can be minimized.

Table 4 illustrates the accuracy for the real-time, 15-s, and 30-s modes for the six METs. The real-time accuracy for sitting, standing, and walking, which did not involve object detection, averaged 86%. In contrast, the real-time accuracy for typing, cooking, and cleaning, which incorporate object detection, averaged 80%. This lower accuracy was attributed to the simultaneous estimation of both the activity and object.

In the activity classification stage, the classification of walking was challenging because it had the potential to be misclassified as standing when the legs overlapped or as sitting when the leg joints bent momentarily. As a result, the accuracy for walking was lower than that for sitting or standing. This lower accuracy for walking classification thus affected the accuracy of cleaning, an activity that involves walking while using a broom or vacuum cleaner. This problem was exacerbated further by the errors in broom and vacuum cleaner detection when these

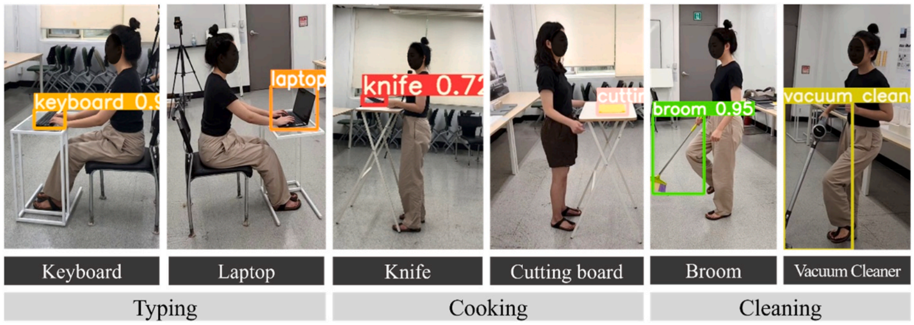



**Fig. 9.** Object detection examples.

**Table 4**
Accuracy of the MET estimation algorithm.

| METs | Object | Real-time | 15-s mode | 30-s mode |
|---|---|---|---|---|
| 1.0 met (sitting) | None | 94% | 100% | 100% |
| 1.2 met (standing) | None | 84% | 100% | 100% |
| 1.7 met (walking) | None | 80% | 100% | 100% |
| 1.1 met (typing) | Laptop | 89% | 100% | 100% |
| | Keyboard | 87% | 100% | 100% |
| 1.8 met (cooking) | Knife | 76% | 100% | 100% |
| | Cutting board | 82% | 100% | 100% |
| 2.7 met (cleaning) | Broom | 71% | 93% | 100% |
| | Vacuum cleaner | 73% | 95% | 100% |

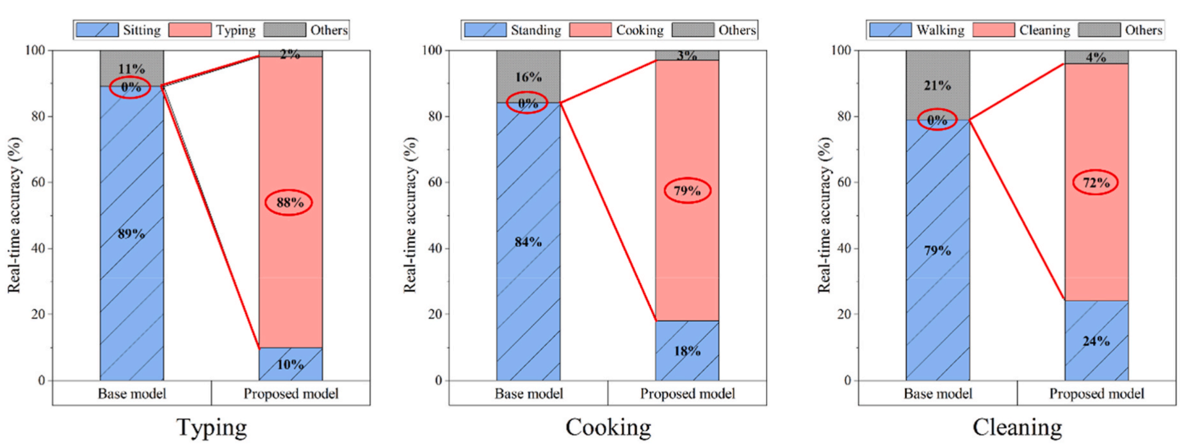
objects were obscured by the pose of the occupant, reducing the accuracy of cleaning classification. Additionally, classification errors occurred when the blade of the knife was thin and the object was not clearly visible due to reflections, reducing the classification accuracy for cooking.

To minimize real-time errors and achieve accurate MET classification, the mode values for data accumulated over regular intervals were analyzed. Using the 15-s mode values, all activities except for cleaning exhibited 100% accuracy while, with the 30-s mode values, all six METs exhibited 100% accuracy. By utilizing the mode values of the accumulated data for periods exceeding 30-s, it was confirmed that errors occurring in real-time data could be reduced, leading to an improvement in the MET estimation accuracy.

The accuracy of the proposed MET estimation algorithm for typing, cooking, and cleaning was also compared to that of an activity classification model that does not include object detection [45] (Fig. 10). The conventional activity classification model classified typing as sitting with a probability of 89%. However, the proposed MET estimation algorithm reduced the rate for classification as sitting to 10%, with a classification accuracy for typing of 88%. The activity classification model classified cooking as standing in 89% of cases, while the MET estimation algorithm correctly classified it as cooking in 79% of cases. Similarly, the activity classification model classifies cleaning as walking in 79% of cases, whereas the MET estimation algorithm correctly identified cleaning 72% of the time. Unlike the activity classification model, which simply classifies a few activities based on their poses, the MET estimation algorithm demonstrated the ability to produce more various MET estimates for the same poses depending on the object(s) that were in use.

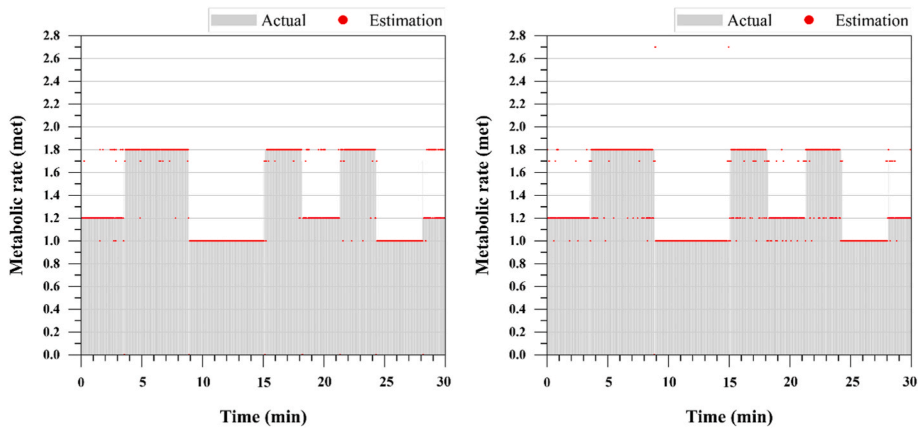
#### 3.2.2. Scenario-based experiment

Figs. 11 and 12 compare the actual MET values and the values predicted by the algorithm for Scenarios 1 and 2, respectively, based on video image data collected from cameras installed on the left and right of the room. The real-time MET exhibited by the participant depending on the activity is indicated by the vertical bars, while the algorithm's predicted results are denoted by the red dots. Looking at the average real-time accuracy for each scenario, the left and right cameras in Scenario 1 produced accuracies of 91% and 87%, respectively. The main errors occurred for the standing posture, with the algorithm mistakenly recognizing a phone as a knife, leading to the prediction of cooking. In Scenario 2, the real-time accuracy for the left and right cameras was 80% and 82%, respectively. The main errors occurred during cleaning, where objects were obscured by the participant's body while walking, leading to detection failure and misclassification as walking. Furthermore, some other activities such as picking up and lifting a broom from



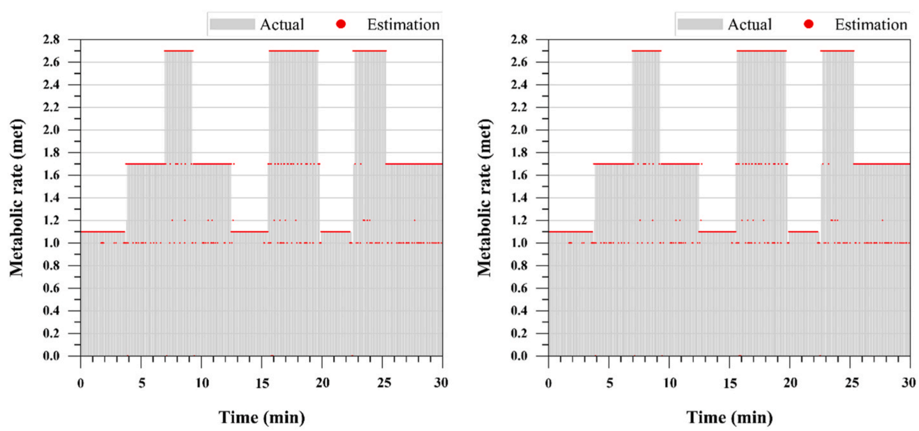

- Base model : Activity classification model [44]
- Proposed model : MET estimation algorithm

**Fig. 10.** Comparison of the activity classification model and MET estimation algorithm.




(a) Left camera (b) Right camera

**Fig. 11.** Actual and estimated MET values for Scenario 1.

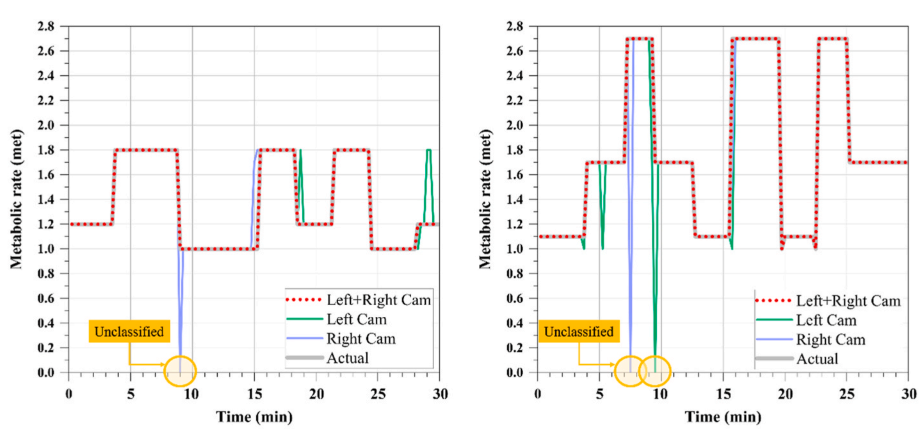



(a) Left camera (b) Right camera

**Fig. 12.** Actual and estimated MET values for Scenario 2.

the floor or getting up from a chair were mistakenly classified as sitting instead of being unclassified, leading to errors.

Fig. 13 presents the classification results using the 15-s mode value for each scenario based on images from the left camera (Fig. 13a) and right camera (Fig. 13b). The orange highlights in Fig. 13 represent cases in which the activity was not classified, which was denoted as 0.0 met




(a) Scenario 1 (b) Scenario 2

**Fig. 13.** Actual and estimated MET using the 15-s mode value for Scenarios 1 and 2.

for visualization purposes. The accuracy of the 15-s mode values for the left and right cameras in Scenario 1 was 96% and 98%, respectively. Similarly, in Scenario 2, the accuracy for the left and right cameras was 96% and 98%, respectively. Even when performing the same activity, the classification of activities or objects differed based on the direction or angle observed by the cameras. Therefore, to enhance the accuracy of the proposed algorithm, continuous training with additional objects prone to misrecognition, such as mobile phones, is necessary. Moreover, training the algorithm to recognize activities such as sitting, standing, and walking based on the joint positions when using objects can help to improve the accuracy and reduce errors.

When an error occurred in estimating the results for either the left or right camera, combining the estimated values from both camera directions resulted in 100% accuracy when using the 15-s mode value. Though it is possible to calculate the MET for building control using a single camera, the error decreases as the number of cameras increases.

Based on the analysis of Fig. 13, a detailed examination of the major errors for each scenario is presented in Fig. 14. In Scenario 1, the most common error was the misclassification of standing as cooking, attributed to the algorithm mistakenly detecting a phone as a knife. In Scenario 2, the predominant error was the misclassification of cleaning as walking, primarily occurring when the algorithm failed to detect the presence of a broom or vacuum cleaner. Errors may arise when the objects are swapped, causing individuals to bend their waist or when some joints leave the field of view of the camera, resulting in a failure to detect any activity. To mitigate these errors, using images captured from multiple directions or slightly longer mode values can prevent momentary errors. Although this study evaluated the algorithm based on very short cumulative data periods of 15-s and 30-s, the developed algorithm has been designed for use in actual HVAC (Heating, Ventilation, and Air Conditioning) control. Thus, in practical environments, the algorithm can be adapted to calculate representative activities based on data accumulated over 5-min or 10-min intervals for more accurate estimations.

## 4. Discussion and conclusion

This study introduces a novel MET estimation algorithm that incorporates both an activity classification model and an object detection model. This approach enables a more diverse range of MET estimates to be produced than is possible with conventional methods by concurrently considering both the activity and objects involved. The key results of this study are as follows:

1) A novel MET estimation algorithm that integrated pose-based activity classification with object detection was developed. The object detection model was trained on six objects: laptops, keyboards, knives, cutting boards, brooms, and vacuum cleaners. The average classification accuracy of the model for these six objects was 89%, with an average accuracy for typing-associated objects of 93.5%, for cooking-associated objects of 87.5%, and for cleaning-associated objects of 85%. When using 15-s mode values, 100% detection accuracy was observed for all objects.
2) The MET estimation algorithm was used to estimate six METs, with two performance-assessment experiments conducted in an actual indoor environment. First, each of the six activities was performed individually for a duration of 10 min, and the accuracy was evaluated. The average classification accuracy for all activities was 83%. When using the 15-s mode values, 100% accuracy was achieved for all activities except cleaning; with the 30-s mode values, 100% accuracy was observed for all activities.
3) The second experiment focused on the accuracy of scenarios involving random changes in the activities. In Scenario 1, which included sitting, standing, and cooking, the real-time accuracy was 89%, with a 15-s mode value accuracy of 97%. In Scenario 2, which included typing, walking, and cleaning, the real-time accuracy was 80%, and the 15-s mode value accuracy was 97%. Considering images from both the left and right cameras simultaneously in both scenarios resulted in 100% accuracy.

In summary, the performance of the proposed MET estimation algorithm, which combined both activity classification and object detection, was assessed in terms of its MET estimates and practical applicability. It was found that using the mode value of cumulative real-time data for a certain period can lead to excellent MET estimation performance. The intervals employed in the present study (15-s and 30-s) were conservative; based on the accuracy of these intervals, a single representative value for a control period of 5 min or 10 min can be determined and utilized in practical applications.

Thus, this technology can be used to effectively monitor occupant behavior and improve comfort, which are important tasks when seeking to establish a well-functioning building environment, as outlined in the Annex 79 project [51]. The proposed methodology provides information that can be used to ensure the thermal comfort of occupants by combining the building's HVAC system, CCTV, the Internet of Things, and the building energy management system. Additionally, technology that can detect human activities or objects can be used not only in buildings but also in city planning. When planning a city, safety and health are important factors to consider in order to improve the convenience and quality of life for the residents [52]. The proposed methodology can contribute to creating a safe and healthy city by detecting dangerous objects or situations and monitoring the behavior of vulnerable groups, such as the elderly, to prevent accidents or enable a rapid response to dangerous situations. In addition, with the recent increase in environmental, economic, and climate uncertainty, global challenges such as achieving carbon neutrality and Sustainable Development Goals (SDGs) have recently gained greater attention. There is a general aim to promote good health and well-being by decreasing energy consumption and greenhouse gas emissions. The methodology developed in this study can thus play a crucial role in supporting these goals by creating a comfortable environment while reducing unnecessary energy consumption.

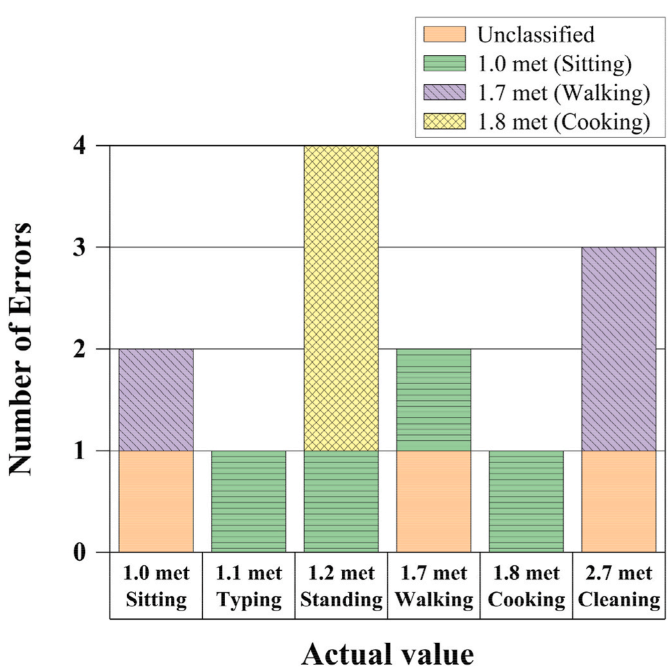



**Fig. 14.** Actual values for incorrectly estimated MET values in the scenario-based experiment.

To apply this technology diversely, the following follow-up studies are necessary. The applicability of the proposed approach in real thermal environment control was confirmed via performance verification based on actual experiments. By building upon the proposed method, it would be possible to expand the MET estimates by increasing the number of classified activities and types of objects. In addition, to enhance the accuracy of the MET estimation algorithm, further training of both the activity classification and object detection models is necessary. Utilizing this technology, improvements in thermal comfort and building operation, including the reduction in unnecessary energy use, can be achieved in indoor environments where the MET of the occupants is variable. However, because the performance evaluation of this study was conducted targeting a single occupant, further research is necessary to validate its suitability for environments with multiple occupants.

Furthermore, image-based technologies that collect data using cameras can generate privacy concerns. Therefore, it is essential to introduce privacy protection technology to address these concerns. To prevent data from being accessed by unauthorized individuals, data encryption and secure transmission technology should be employed in conjunction with the proposed system [53]. In addition, by combining the developed algorithms with an HVAC system, the edge computing technology within the HVAC system itself can be harnessed, simplifying the data transmission process and enhancing security [54].

Moreover, subsequent studies are also needed to assess the effectiveness of indoor thermal comfort control by incorporating real-time MET values obtained from the MET estimation algorithm into PMV-based control. In this context, the MET estimation method for determining individual thermal sensations can be applied not only to PMV but also to the development of novel thermal comfort models. It is essential to consider CLO in the control strategy to comprehensively address individual variables [34,47]. Through this research, this technology has the potential to contribute as a key element in the control of indoor thermal comfort and energy efficiency.

## CRediT authorship contribution statement

**Ji Young Yun:** Writing – original draft, Software, Conceptualization. **Eun Ji Choi:** Writing – review & editing, Software, Methodology. **Min Hee Chung:** Writing – review & editing, Validation. **Kang Woo Bae:** Investigation, Formal analysis. **Jin Woo Moon:** Supervision, Project administration, Conceptualization.

## 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.

## Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2019R1A2C1084145 and RS-2023-00217322) and Chung-Ang University research grant in 2023.

## References

[1] M. Frontczak, P. Wargocki, Literature survey on how different factors influence human comfort in indoor environments, Build. Environ. 46 (4) (2011) 922–937.
[2] C. Schweizer, R.D. Edwards, L. Bayer-Oglesby, W.J. Gauderman, V. Ilacqua, M. Juhani Jantunen, H.K. Lai, M. Nieuwenhuijsen, N. Künzli, Indoor time–microenvironment–activity patterns in seven regions of Europe, J. Expo. Sci. Environ. Epidemiol. 17 (2) (2007) 170–181.
[3] J.A. Leech, W.C. Nelson, R.T. Burnett, S. Aaron, M.E. Raizenne, It's about time: a comparison of Canadian and American time–activity patterns, J. Expo. Sci. Environ. Epidemiol. 12 (6) (2002) 427–432.
[4] E.J. Choi, B.R. Park, N.H. Kim, J.W. Moon, Evaluation of thermal comfort by PMV-based control applying dynamic clothing insulation, KIEAE J. 22 (1) (2022) 53–60.
[5] G. Brager, H. Zhang, E. Arens, Evolving opportunities for providing thermal comfort, Build. Res. Inf. 43 (3) (2015) 274–287.
[6] J.G. Allen, P. MacNaughton, J.G.C. Laurent, S.S. Flanigan, E.S. Eitland, J. D. Spengler, Green buildings and health, Curr. Environ.l Health Rep. 2 (2015) 250–258.
[7] W.-T. Sung, S.-J. Hsiao, The application of thermal comfort control based on Smart House System of IoT, Measurement 149 (2020) 106997.
[8] J. Wu, X. Li, Y. Lin, Y. Yan, J. Tu, A PMV-based HVAC control strategy for office rooms subjected to solar radiation, Build. Environ. 177 (2020) 106863.
[9] G. Gao, J. Li, Y. Wen, Energy-efficient Thermal Comfort Control in Smart Buildings via Deep Reinforcement Learning, 2019 arXiv preprint arXiv:1901.04693.
[10] Y. Peng, Z. Nagy, A. Schlüter, Temperature-preference learning with neural networks for occupant-centric building indoor climate controls, Build. Environ. 154 (2019) 296–308.
[11] N. Haidar, N. Tamani, Y. Ghamri-Doudane, A. Bouju, Towards a new graph-based occupant behavior modeling in smart building, in: 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), IEEE, 2019, pp. 1809–1814.
[12] C. Zhong, J.H. Choi, Development of a data-driven approach for human-based environmental control, Procedia Eng. 205 (2017) 1665–1671.
[13] T. Yang, A. Bandyopadhyay, Z. O'Neill, J. Wen, B. Dong, From Occupants to Occupants: A Review of the Occupant Information Understanding for Building HVAC Occupant-Centric Control, Building Simulation, Springer, 2022, pp. 913–932.
[14] S. Naylor, M. Gillott, T. Lau, A review of occupant-centric building control strategies to reduce building energy use, Renew. Sustain. Energy Rev. 96 (2018) 1–10.
[15] H.N. Choi, C.Y. Um, K.M. Kang, H.K. Kim, T.Y. Kim, Review of vision-based occupant information sensing systems for occupant-centric control, Build. Environ. 203 (2021) 108064.
[16] J. Kim, S. Schiavon, G. Brager, Personal comfort models–A new paradigm in thermal comfort for occupant-centric environmental control, Build. Environ. 132 (2018) 114–124.
[17] M.M. Ouf, Optimizing occupant-centric building controls given stochastic occupant behaviour, in: Journal of Physics: Conference Series, IOP Publishing, 2021 012140.
[18] Z. Pang, J. Zhang, Y. Chen, H. Cheng, Z. O'Neill, B. Dong, Nationwide energy saving analysis for office buildings with occupant centric building controls, Build. Eng. 126 (2) (2020).
[19] J. Xie, H. Li, C. Li, J. Zhang, M. Luo, Review on occupant-centric thermal comfort sensing, predicting, and controlling, Energy Build. 226 (2020) 110392.
[20] P.O. Fanger, Thermal comfort. Analysis and applications in environmental engineering, Therm. Comf. Anal. Appl. Environ. Eng. (1970).
[21] S.H. Hong, J.M. Lee, J.W. Moon, K.H. Lee, Thermal comfort, energy and cost impacts of PMV control considering individual metabolic rate variations in residential building, Energies 11 (7) (2018) 1767.
[22] H. Choi, H. Na, T. Kim, T. Kim, Vision-based estimation of clothing insulation for building control: a case study of residential buildings, Build. Environ. 202 (2021) 108036.
[23] A. Standard, Thermal environmental conditions for human occupancy, ANSI/ ASHRAE 55 (2017).
[24] R. Kosonen, F. Tan, Assessment of productivity loss in air-conditioned buildings using PMV index, Energy Build. 36 (10) (2004) 987–993.
[25] W. Huo, Y. Cheng, Y. Jia, C. Guo, Research on the thermal comfort of passenger compartment based on the PMV/PPD, Int. J. Therm. Sci. 184 (2023) 107876.
[26] J.S. Park, H.N. Choi, D.H. Kim, T.Y. Kim, Development of novel PMV-based HVAC control strategies using a mean radiant temperature prediction model by machine learning in Kuwaiti climate, Build. Environ. 206 (2021) 108357.
[27] D. Esteves, J. Silva, N. Rodrigues, L. Martins, J. Teixeira, S. Teixeira, Simulation of PMV and PPD thermal comfort using energyplus, in: International Conference on Computational Science and its Applications, Springer, 2019, pp. 52–65.
[28] J. Miura, M. Demura, K. Nishi, S. Oishi, Thermal comfort measurement using thermal-depth images for robotic monitoring, Pattern Recogn. Lett. 137 (2020) 108–113.
[29] E.J. Choi, B.R. Park, N.H. Kim, J.W. Moon, Effects of thermal comfort-driven control based on real-time clothing insulation estimated using an image-processing model, Build. Environ. 223 (2022) 109438.
[30] Y. Tang, H. Yu, H. Ye, K. Zhang, F. Wang, H. Mao, Z. Wang, Estimating local thermal insulation of clothing garments: modelling and application, Build. Environ. 243 (2023) 110558.
[31] S.I.-u.-H. Gilani, M.H. Khan, M. Ali, Revisiting Fanger's thermal comfort model using mean blood pressure as a bio-marker: an experimental investigation, Appl. Therm. Eng. 109 (2016) 35–43.
[32] M.H. Hasan, F. Alsaleem, M. Rafaie, Sensitivity study for the PMV thermal comfort model and the use of wearable devices biometric data for metabolic rate estimation, Build. Environ. 110 (2016) 173–183.
[33] Y. Zhang, X. Zhou, Z. Zheng, M.O. Oladokun, Z. Fang, Experimental investigation into the effects of different metabolic rates of body movement on thermal comfort, Build. Environ. 168 (2020) 106489.
[34] T. Akimoto, S.-i. Tanabe, T. Yanai, M. Sasaki, Thermal comfort and productivity-Evaluation of workplace environment in a task conditioned office, Build. Environ. 45 (1) (2010) 45–50.
[35] ISO 8996, 2004–Ergonomics of the Thermal Environment–Determination of Metabolic Rate, 2004.

[36] M. Luo, X. Zhou, Y. Zhu, J. Sundell, Revisiting an overlooked parameter in thermal comfort studies, the metabolic rate, Energy Build. 118 (2016) 152–159.

[37] Y. Zhai, M. Li, S. Gao, L. Yang, H. Zhang, E. Arens, Y. Gao, Indirect calorimetry on the metabolic rate of sitting, standing and walking office activities, Build. Environ. 145 (2018) 77–84.

[38] W. Ji, M. Luo, B. Cao, Y. Zhu, Y. Geng, B. Lin, A new method to study human metabolic rate changes and thermal comfort in physical exercise by CO2 measurement in an airtight chamber, Energy Build. 177 (2018) 402–412.

[39] J. Liu, I.W. Foged, T.B. Moeslund, Automatic estimation of clothing insulation rate and metabolic rate for dynamic thermal comfort assessment, Pattern Anal. Appl. 25 (3) (2022) 619–634.

[40] J. Liu, I.W. Foged, T.B. Moeslund, Clothing insulation rate and metabolic rate estimation for individual thermal comfort assessment in real life, Sensors 22 (2) (2022) 619.

[41] H.S. Na, H.N. Choi, T.Y. Kim, Metabolic Rate Estimation Method Using Image Deep Learning, Building Simulation, Springer, 2020, pp. 1077–1093.

[42] H. Choi, B. Jeong, J. Lee, H. Na, K. Kang, T. Kim, Deep-vision-based metabolic rate and clothing insulation estimation for occupant-centric control, Build. Environ. 221 (2022) 109345.

[43] O. Mata, J.I. Méndez, P. Ponce, T. Peffer, A. Meier, A. Molina, Energy savings in buildings based on image depth sensors for human activity recognition, Energies 16 (3) (2023) 1078.

[44] E.J. Choi, J.W. Moon, J.H. Han, Y.S. Yoo, Development of a deep neural network model for estimating joint location of occupant indoor activities for providing thermal comfort, Energies 14 (3) (2021) 696.

[45] N.H. Kim, E.J. Choi, D.H. Park, J.W. Moon, Performance evaluation of the multiple occupants real-time MET estimation model for thermal comfort control of building, KIEAE J. 23 (1) (2023) 69–74.

[46] K.N.G. Jocher, T. Mineeva, R. Vilarino, YOLOv5, vol. 5, 2020.

[47] J.W. Park, Y.J. Kim, A Study on Deep Learning Performance Improvement Based on YOLOv5, Proceedings of Symposium of the Korean Institute of communications and Information Sciences, 2022, pp. 1592–1593.

[48] Y. Liu, B. Yang, Z. Lin, A pilot study of occupant centric control stratum ventilation based on computer vision, in: E3S Web of Conferences, EDP Sciences, 2022 01029.

[49] E.J. Choi, Y.J. Choi, N.H. Kim, J.W. Moon, Seasonal effects of thermal comfort control considering real-time clothing insulation with vision-based model, Build. Environ. 235 (2023) 110255.

[50] ISO 7730, 2005-Ergonomics of the Thermal Environment: Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria, ISO, 2005.

[51] IEA EBC, Occupant Behaviour-Centric Building Design and Operation EBC Annex 79, IEA, 2018.

[52] S. Koley, Challenges in sustainable development of smart cities in India, Sustain. J. Rec. 13 (4) (2020) 155–160.

[53] C. Tiken, R. Samli, A comprehensive review about image encryption methods, Harran Üniversitesi Mühendislik Dergisi 7 (1) (2022) 27–49.

[54] S. Márquez-Sánchez, J. Calvo-Gallego, A. Erbad, M. Ibrar, J.H. Fernandez, M. Houchati, J.M. Corchado, Enhancing building energy management: adaptive edge computing for optimized efficiency and inhabitant comfort, Electronics 12 (19) (2023) 4179.