



Non-invasive human thermal comfort assessment based on multiple angle/distance facial key-region temperatures recognition

Kangji Li, Weiwei Li, Fukang Liu, Wenping Xue*

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, PR China



ARTICLE INFO

Keywords:

Infrared camera
Thermal comfort
Facial region temperatures
Correlation analysis
Data-driven model

ABSTRACT

Accurate monitoring of human thermal comfort is crucial in optimizing HVAC system control scheme and enhancing building energy efficiency. The method based on facial temperature infrared recognition is a non-invasive and real-time thermal comfort predictive strategy, which has been proven great application potential. In response to the limitations of previous fixed-position facial infrared detection in practical applications, this study aims to explore a less restrictive non-intrusive method for assessing human thermal comfort. The temperature features of multiple regions on human face are identified by an infrared camera, and six temperature-sensitive regions are selected through skin thermal sensitivity analysis. Thirty subjects participate in a series of field experiments. To reduce the measurement limitations caused by human postures, the correlation between these six facial regions, indoor air temperature, thermal sensation votes (TSV), and thermal comfort votes (TCV) are analyzed in details to determine two key facial regions for the non-intrusive thermal comfort assessment. By applying the YOLOv5 algorithm, the real-time extraction of facial region temperatures from multiple angles and distances is achieved. On this basis, a data-driven thermal comfort predictive strategy based on facial temperature and optimized SVM model is designed. Results indicate that the method achieves a general accuracy of 85.68% in predicting thermal comfort based on its successful recognition rate of 88.7% for facial key regions.

1. Introduction

The increase in urbanization is promoting the rapid expansion of various public buildings [1]. Relevant data show that in the current energy consumption, the proportion of building energy consumption has exceeded 40% [2]. In this part of energy use, HVAC system accounts for 2/3 [3,4]. Nevertheless, indoor thermal dissatisfaction still exists [5,6]. Previous surveys have shown that only 11% of buildings could meet the comfort requirement of 80% of the occupants [7]. Analyzing the main reasons for low satisfaction of thermal comfort in building environments, on the one hand, indoor environmental variables and human physical/psychological states both affect human thermal comfort [8]. On the other hand, most HVAC systems adopt setting point control strategies, which cannot meet the personalized comfort needs of indoor population [9]. Therefore, accurately identifying the thermal comfort of the occupants in buildings is of great significance for building energy efficiency and human well-being index.

1.1. Related works

It is well known that PMV (Predictive Mean Voting) is a quantitative model widely used for the assessment of human thermal comfort

(ASHRAE 55–2020 [10]). However, it cannot update existing parameters real-time especially from occupants and on-site collection [11]. In recent years, various data-driven models have been introduced into this field [12–15]. For example, Jung et al. [12] placed the temperature sensor on the back of the hand to measure skin temperature, and used the support vector machine model (SVM) to predict the comfort of the human body, achieving an accuracy rate of 87%. Liu et al. [13] and Soo et al. [14] attempted to use wearable devices to measure participants' skin temperature, heart rate, and some environmental parameters. They used machine learning classifiers to calculate human thermal preferences and achieved an accuracy of over 75%.

At present, feedbacking occupant's thermal physiological information and thermal comfort states has been recognized as the key for building environmental optimal control. Factors that reflect the thermal comfort of the human body can be divided into environmental variables (such as indoor air temperature, humidity, air velocity, etc.), physiological information (such as skin temperature, heart rate, gender, age, BMI index, etc.) and other factors (human activity level, health states, etc.). Traditional invasive physiological information collection equipment can collect human physiological data of wrist temperature,

* Corresponding author.

E-mail address: xwpingle@ujs.edu.cn (W. Xue).

Table 1
Related papers on non-invasive thermal comfort prediction by infrared camera.

Literatures	Measurement region	Angle/Distance	Face detection methods	Accuracy
Bouzida et al. (2009) [23]	back of the hand	fixed	none	not reported
Ranjan et al. (2016) [24]	forehead, cheeks, lips, jaw, upper neck, lower neck, palms and back of hands	fixed	none	63–68%
Pavlim et al. (2017) [25]	forehead	fixed	5 key points	not reported
Aryal et al. (2019) [26]	nose, cheek, forehead	fixed	68 key points	85%
Cosma et al. (2018) [27]	hands, elbows, shoulders, chest, forehead	fixed	none	not reported
Metzmacher et al. (2018) [28]	eyes, nose, cheeks, mouth, forehead, chin	fixed	face segmentation	not reported
Li et al. (2018) [29]	nose, cheeks, mouth, ears, neck, forehead	fixed	Haar-cascade	85%
Cosma et al. (2019) [30]	eyes, nose, cheeks, mouth, chin, forehead	fixed	Haar-cascade	76%
Wu et al. (2022) [31]	eyes, nose, cheeks, ears, forehead	fixed	none	79.9%

heart rate [16], mean blood pressure [17], ECG, EEG [18], etc.. But there are some limitations in this invasive manner. Firstly, physical contact and subject cooperation are required during the measurement process. Secondly, the body fat content of the subject and the position of the device worn will affect the measurement accuracy, and this is nearly inevitable. In addition, the measurement equipment for ECG and EEG are usually expensive and unsuitable for on-site measurement for the purpose of thermal comfort assessment.

Compared to traditional invasive methods, semi-contactive measurement requires less physical contact. It is mainly achieved by integrating sensors into wearable accessories [19]. Ghahramani et al. [20] installed infrared sensors on eyeglass frames to monitor the body's thermoregulatory response in different environments, which was further used for thermal comfort assessment. Nkurikiyeyeze et al. [21] estimated pulse rate variability (PRV) by recording photoplethysmography (PPG) signals on the wrist. The heartbeat interval signal was extracted and sent to the mobile client for PRV calculation and thermal comfort modeling. Under different indoor temperature conditions, Choi et al. [22] used wearable sensors to measure skin temperatures of seven selected body regions. The small sample size limited the effectiveness of validation results. The measurement of physiological information collected by wearable devices could establish a good estimation of human thermal sensation, and to certain extent, reduce the subject's foreign body sensation. However, compared with non-invasive manners, wearable accessories still have limitations in terms of physical contact and the resulting subjective influence.

In recent years, non-invasive measurements have been adopted due to their minimal interference with indoor occupants. There are several different research directions, including human skeleton key-point recognition for monitoring thermal comfort/discomfort posture, Euler video magnification (EVM) based human temperature monitoring with an ordinary camera, and infrared camera based human skin temperature monitoring. Yang et al. [32], Na et al. [33] and Meier et al. [34] used Kinect to observe various human postures related to thermal discomfort, and developed a method for non-contact monitoring of human thermal discomfort state. Due to the insufficient sampling of frames in a short time, the human posture recognition is easy to cause misjudgment. The use of Kinect is protected by commercial patents and its development is limited. The skin temperature is the most direct physiological information that reflects the thermal state of the human body. EVM technology can amplify image frames in videos to detect color and subtle motion changes that are invisible to the naked eye [35]. Cheng et al. [36] used EVM to correlate skin temperature with skin color saturation for non-contact measurement of human body temperature. Jung et al. [37] and Jazizadeh et al. [38] also used the EVM technique to infer the state of human thermoregulation for thermal comfort assessment. Compared with RGB cameras which are highly dependent on natural light and lighting, infrared cameras can

measure skin temperature directly without visible light interference. It can also avoid privacy issues related to the use of RGB cameras [39]. Ranjan et al. [24] analyzed temperature features of facial and hand regions and correlated them with thermal sensations for the purpose of HVAC energy efficiency. After skin temperature collection at offline stage, the thermal sensation prediction achieved the accuracy of 63%–68%. Pavlim et al. [25] and Li et al. [29] combined a low-resolution thermal camera with the Haar-Cascade classifier to collect the subject's skin temperature at a given position. The human thermal comfort was assessed with 85% accuracy. Aryal et al. [26] applied facial landmark detection algorithm to extract skin temperature for thermal comfort assessment, and the accuracy rate of 76% in predicting thermal sensation was achieved based on 85% facial recognition rate. However, the performance of their algorithm was unknown when the head is tilted or rotated. Table 1 provides a brief summary of individual thermal comfort assessment using infrared camera.

From the Table 1, it is shown that the skin temperatures of various facial regions are strongly correlated with the thermal comfort state of the human body. On this basis, the human thermal comfort can be predicted accurately with infrared camera. However, it should be noted that the current fixed position based facial infrared detection and thermal comfort assessment have limitations in practice. Firstly, completely non-invasive methods cannot limit the posture of the human body indoor; Secondly, in more general detection for multi person, human postures are inevitably different. Therefore, from the perspective of practical application, facial temperature detection for indoor population needs to consider the influence of multiple angles and distances between the camera and the object, which is generally ignored in current research.

1.2. Main contribution of this work

Starting from the above issues, this paper conducts field experiments and data analysis to investigate the thermal sensitive regions of human face at different distances/angles. On this basis, a non-invasive human thermal comfort assessment method with less limitations is proposed using infrared camera and machine learning method. Major contributions of this paper include:

(1) By using sensitivity analysis method, six facial regions are identified from collected infrared images in typical thermal environments. Based on Pearson correlation coefficients analysis between the six facial temperatures and thermal sensation votes data, two most sensitive regions, i.e., cheeks and nose, are further chosen for multi-angle/distance recognition and thermal comfort prediction.

(2) A multi-angle/distance facial thermal image dataset is constructed by field experiment, based on which a multi-angle facial region recognition model based on the YOLOv5 object detection algorithm is designed.

(3) For non-contact thermal comfort evaluation, an improved SVM pattern recognition method is proposed, in which a modified swarm intelligence algorithm (CLPSO) is applied for SVM's parameter optimization. On this basis, a data-driven online thermal comfort prediction model based on two-region facial recognition is developed. Compared with other non-invasive methods, the proposed framework has less restrictions on subjects and the prediction performance is verified and analyzed.

1.3. Paper organization

The remaining sections of this paper are organized as follows: Section 2 provides a detailed description of the methods used for infrared image analysis. The relevant settings for data collection experiments, and the method for thermal comfort prediction are also provided in this section. In Section 3, the experimental results are analyzed in detail. Section 4 presents the results of thermal comfort prediction. Section 5 provides a brief discussion on research findings and limitations. Section 6 offers the final conclusions and outlines the future work.

2. Method

2.1. Quantitative analysis of sensitivity

Skin thermal sensitivity refers to the ability of the skin to react to thermal stimulation, which is commonly used to measure skin's heat resistance. By quantifying the differences in facial thermal images under different thermal comfort levels, the thermal comfort sensitivity results of multiple facial regions can be obtained. Specifically, the temperature differences for each pixel in the facial region under different thermal comfort sensations are calculated, i.e., $\Delta T = T_{s1} - T_{s2}$, in which T_{si} represents the temperature value of the thermal image taken in the certain thermal state i . The thermal sensitivity analysis of facial region is then conducted, based on which the key facial regions for thermal comfort prediction are identified. The selection of key facial regions needs to consider several factors: (1) choose facial regions that are susceptible to changes in environmental temperature, as these regions better represent changes in human thermal sensation; (2) choose the facial region close to the core body temperature that is relatively stable and could accurately reflect the overall body temperature; (3) the facial regions that are easily obscured or difficult to measure should not be selected.

Taking the above factors into account, we conduct a thermal comfort vote experiment (see Section 2.3 for details), and perform the thermal sensitivity analysis on facial key regions selection.

2.2. Multi-position facial region detection method

(1) Object detection algorithm

In this study, the model for extracting facial regions is trained by the YOLO (You Only Look Once) version 5 structure [40]. Different from traditional object detection algorithms, YOLOv5 uses a deep neural network model to predict the class and location of objects simultaneously. Its ability of rapidly detecting objects with high accuracy makes it suitable for real-time estimation of skin temperature of indoor occupants. The network structure consists of three main parts: Backbone network, Neck network, and Head network. Concretely, the Backbone network is used to extract feature maps from the input image; The Neck network fuses feature maps from different layers to improve detection accuracy; The target objects are then classified and localized by Head network, which could output the bounding box's class and position information of each detected object. The facial regions detected by the model are displayed in the image as a bounding box containing only the ROI, and the facial region classification results are labeled within the bounding box. Be noted: ROI refers to a region of interest, which is a specific zone of an image that we further extract features from. In addition, the

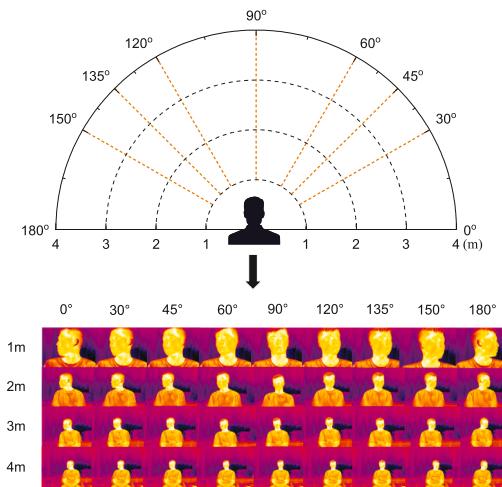


Fig. 1. Diagram of multiple shooting angles/distances.

algorithm of YOLOv5 is directly applied to thermal images for object detection without calibration of the RGB images, which simplifies field experiments' steps. The infrared camera is immunity to lighting variations, which guarantees the performance of object detection across diverse lighting scenarios, and it is also good at privacy protection.

(2) Construction of infrared image dataset with multiple angles and distances

In order to develop a non-invasive thermal comfort prediction strategy with least limitations, fully consider the multiple possible angles and distances between the infrared camera and human faces in an ordinary office environment. A dataset of collected infrared images is then constructed. Fig. 1 is a sample diagram of a typical subject, including 36 infrared images taken from four distances and nine angles. The distances between the subject and infrared camera are within the range of 1–4 m. At each distance, nine equal difference horizontal angles (0° to 180°) between the subject and infrared camera are set to simulate the possible postures in office scenarios. The pitch angle of the human face is not strictly restricted, and angles within 30° up and down are allowed. A total of 898 infrared images from 25 subjects at different distances and angles are collected, and by using the 'LabelImg' tool, all images are labeled and converted to a format that can be adopted by the YOLOv5 networks.

To extract representative temperature of each facial region, the coordinate information of the detected ROI is output by YOLOv5 networks. Through infrared camera's SDK function, the average temperature value of each ROI can be extracted, before which the image should be calibrated to ensure the temperature value's accuracy.

2.3. Thermal comfort prediction and optimization methods

Machine learning methods, including decision trees, random forests, support vector machines, and artificial neural networks, have been frequently reported for thermal comfort prediction. Among them, the SVM [41] algorithm has higher accuracy and better generalization ability [42,43]. Basically, SVM is a kind of binary classification model that separates positive and negative samples by finding a hyperplane. For multi classification problem, the "one-versus-rest" approach is adopted in this study. Three SVMs are constructed for the three thermal comfort categories of TCV, and seven SVMs are constructed for the seven thermal sensation categories of TSV. Each SVM is responsible for distinguishing a certain class of data. Taking the three-class classification as an example, Fig. 2 shows the classification diagram of this method.

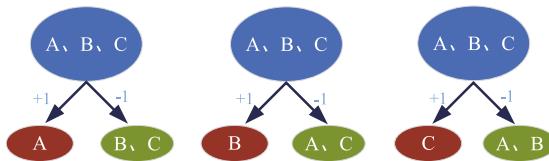


Fig. 2. Three-class SVM example for “one-versus-rest” method.

Based on the basic principles of SVM, the hyperplane partition in the feature space can be described by the following linear equation:

$$y(x) = w^T x + b \quad (1)$$

where w is the normal vector of the hyperplane, x represents the sample data, and b denotes the offset. To maximize the distance between the hyperplane and the nearest sample points, the following optimization problem is derived:

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i \quad (2)$$

$$y_i(w^T \cdot x_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, n \quad (3)$$

$$K(x_i, x) = \exp(-\frac{\|x_i - x\|^2}{2g^2}) \quad (4)$$

where ξ_i represents the slack variable and c is the penalty factor. $K(x_i, x)$ is the kernel function of the SVM model, and g is the kernel function parameter. Choosing an appropriate c and g could help improve the classification accuracy of SVM. For this study, the comprehensive learning particle swarm optimization (CLPSO) algorithm [44] is further applied for these two key parameters’ optimization. In the algorithm, c and g are represented as the position of the particle in the search space. During each iteration, the velocity and position of particles are updated using Formulas 5 and 6:

$$V_{id}^{(k+1)} = w \cdot V_{id}^{(k)} + c_a \cdot r_{id}(pbest_{f(id)}^{(k)} - X_{id}^{(k)}) \quad (5)$$

$$f_i(d) = [f_i(1), f_i(2), \dots, f_i(D)] \quad (6)$$

$$X_{id}^{(k+1)} = X_{id}^{(k)} + V_{id}^{(k)} \quad (7)$$

where V_{id} and X_{id} represent the velocity and position of particle i in d dimension, k is the current iteration number. w is the inertia weight coefficient, c_a is the acceleration factor, and r is a random number in the interval [0,1]. For each particle, the fitness value is calculated based on its current position, and its historical best position $pbest$ is recorded. The $f_i(d)$ defines which $pbest$ of the particles should follow with, which is determined by learning probability P_{ci} :

$$P_{ci} = 0.05 + 0.45 \times \frac{(e^{10(i-1)(pop-1)} - 1)}{e^{10} - 1} \quad (8)$$

where pop represent the population size. Compare the fitness values of all particles and select the best particle position as the global best position $gbest$. Compare with the basic PSO algorithm, the improved CLPSO algorithm uses the comprehensive best position $pbest_{f(id)}$ to update the particles that make it more advantageous in global search capability. The 5-fold cross validation method is used for accuracy verification. The flowchart of the CLPSO-SVM model is shown in Fig. 3.

2.4. Experiment and procedure

(1) Experiment environment and equipment

A typical office room with the size of 7.0 m * 3.6 m (length * width) is used for data collection experiment at Jiangsu University, China. During the experiment, the door and window are both kept

Table 2
Basic physiological features of all subjects.

Gender	Number	Age	Height (cm)	Weight (Kg)	BMI (Kg/m ²)
Male	20	24.4 ± 1.1	176.8 ± 4.3	72.9 ± 5.7	23.3 ± 1.5
Female	10	23.6 ± 1.1	165.2 ± 3.9	56.5 ± 4.9	20.7 ± 1.3

Table 3
Data statistics at different levels of thermal sensation.

Category	Level	Vote number	Percentage (%)
7-class	-3	24	7.7%
	-2	48	15.3%
	-1	46	14.7%
	0	54	17.3%
	1	53	16.9%
	2	50	16.0%
3-class	3	38	12.1%
	1	72	23.0%
	0	153	48.9%
3-class	-1	188	28.1%

closed and the room temperature is regulated within 16–30 °C by a suspended air conditioning system. The indoor air velocity maintains below 0.15 m/s during the experiment. The indoor air temperature (T_{air}) and relative humidity (RH) are monitored by Arduino based sensor (DHT11, DFRobot, China). The sensor is placed on a table about 0.7 m high from the ground, which is following to the ASHRAE standard [10]. An infrared camera (MAGNITY33, Shanghai Magnity Technologies, China, Accuracy: 0.5 °C) is placed at a fixed position in the room. The specific information of all sensors is provided in Appendix (Table A.1). The layout of the experiment site is shown in Fig. 4, and the used experimental equipment is shown in Fig. 5.

(2) Subjects and thermal sensation votes

This study recruits 30 volunteers, consisting of 20 males and 10 females with ages from 23 to 27 and BMI within the normal range. The statistical data of all subjects’ basic physiological features is listed in Table 2. Before the experiment, the volunteers receive physical and psychological health (stress levels) evaluations. They are asked to wear typical summer clothing, including short-sleeved shirts, pants, socks, underwear, and shoes, to maintain a consistent thermal resistance of about 0.57 clo. Since this study involves capturing infrared images of subjects’ faces, the informed consent is required from all participants before the experiment.

In order to collect the subjective evaluation of indoor thermal comfort in the experiment, the thermal sensation votes and thermal comfort votes are both used. This study adopts the 7-level thermal sensation index of the ASHRAE-55 standard [10]: hot (+3), warm (+2), slightly warm (+1), neutral (0), slightly cool (-1), cool (-2), and cold (-3), and 3-point thermal comfort state standard categories: warm discomfort (+1), comfort (0), and cool discomfort (-1) [45]. Fig. 6 shows the details. In both fixed-position and multi-angle/distance experiments, a total of 313 (163+150) valid questionnaires were collected. The data statistics at different levels of thermal sensation are listed in Table 3.

(3) Experimental procedure and content

The experiment was conducted in the summer, June to August, 2022. The whole experimental procedure involves three phases. In preparation phase A, the natural ventilation of the room is conducted for 20 min, after which both door and window are closed to form a manually regulated thermal environment. At the end of phase A, the subject enters the room and phase B begins. To obtain a low metabolic rate and make his (her) body fully adapt to the test environment, the subject is required to sit quietly (as daily office status) in the room for 20 min. At the same time, the subject would be informed with the experimental procedure, thermal sensation and thermal comfort votes and other test information. Then in the next 5 min, the infrared images

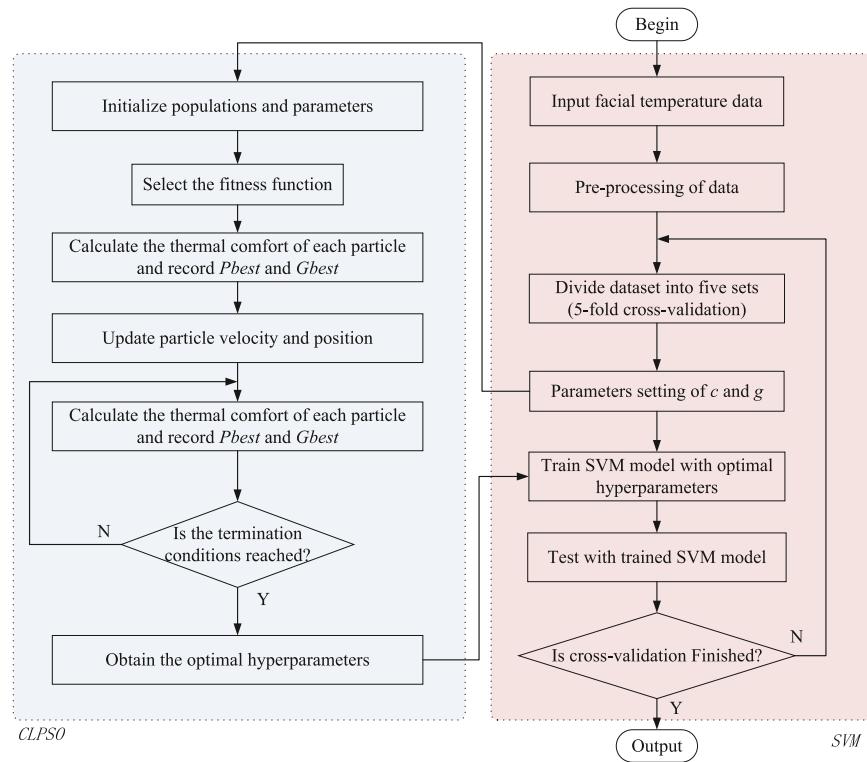


Fig. 3. Flow chart of CLPSO-SVM model.

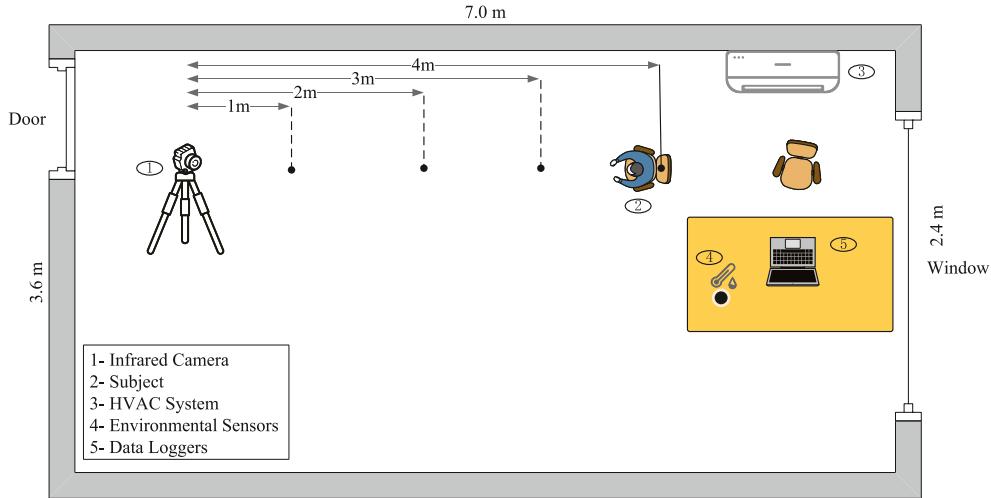


Fig. 4. The layout of the laboratory.

of the subject are taken, and votes/environmental data are collected. After that, the repetitive phase C begins. The temperature set point of air conditioning system would be decreased by step of 2 degrees for five times. The entire experiment dropped the room temperature from around 30 °C to around 20 °C. In order to avoid the influence of negative temperature jump on human thermal sensation, there are twenty minutes for thermal environment re-stabilization and physical adaptation, as indicated in Fig. 7. Be noted: during the whole experiment process, less than half an hour is for subjects' data sampling. The other time is for temperature adjusting and stabilization. During this period, subjects may do some reading, chatting and other relaxation to prevent physiological discomfort caused by mood swings.

The following three experimental contents are carried out following to the above experimental procedure. Firstly, we investigate the facial temperature changes of typical subjects under different thermal

sensations to determine key regions of facial sensitivity. Secondly, we conduct a fixed-position data collection experiment based on the facial key regions. In this experimental part, an infrared camera is placed at a fixed distance of 1.5 m from the subjects. The subjects' faces are required to be facing the camera (90°, Fig. 1) during data collection. Manually select the ROI regions and extract temperatures from the infrared images captured in the experiment. Through this method, the representative regions for thermal comfort prediction are selected for further correlation analysis.

In the third experimental part, the distances between the subject and the camera are extended to 1–4 m (Fig. 4), and for each distance, nine different horizontal angles of human face (0° to 180°, Fig. 1) are set to simulate the possible postures of sitting occupant in real office environment. Based on the results of correlation analysis, only two facial regions with the highest correlation with human thermal

Table 4
Absolute values of facial temperature difference.

TSV	Nose	R-cheek	L-cheek	Mouth	Chin	R-eye	L-eye	Forehead
Cold-Neutral	5.4	2.4	2.4	2.8	2	1.4	1.5	1.4
Neutral-Hot	2.6	2.8	2.8	2.5	2	2.1	2.1	1.6
Cold-Hot	7.8	5.2	5.2	5.1	4.4	3.7	3.6	2

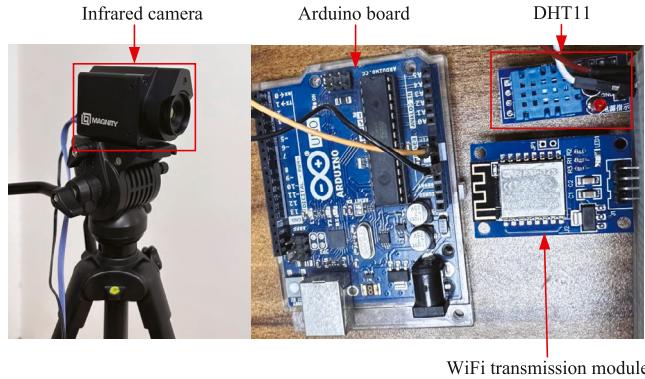


Fig. 5. Infrared camera MAGNITY33, DHT11, Arduino board and WiFi transmission module.

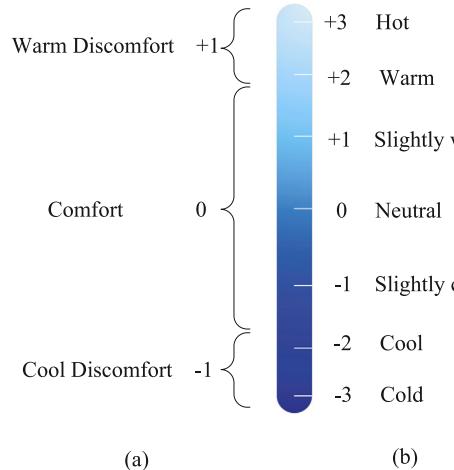


Fig. 6. The scales of thermal comfort state and thermal sensation: (a) Three-class thermal comfort state, (b) Seven-class thermal sensation.

sensation are detected. Different from the second part, the YOLOv5 object detection algorithm is used to achieve online extraction of facial regions' temperatures. In this way, the proposed predictive model for thermal comfort could less limit subject's face angle to the camera, and the steps of artificial selection of ROI are omitted.

2.5. The overall framework

This study aims to propose a non-invasive human thermal comfort assessment method based on facial multi-region temperature recognition. Firstly, by facial sensitivity analysis and quantification, the facial regions for thermal sensation are determined. Then by fixed-position data collection experiment and correlation analysis, the representative regions for thermal comfort prediction are further selected. A YOLOv5-based object detection model is trained to detect ROI using facial images from multi-angle/distance. On this basis, statistical analysis is performed on the extracted multi region temperature data, and a thermal sensation predictive model based on CLPSO-SVM is developed. The framework is shown in Fig. 8.

3. Experimental results and analysis

3.1. Quantitative results of facial thermal sensitivity

We select infrared images under Cold, Neutral, and Hot conditions based on thermal sensation vote, and perform temperature matrix subtraction (as stated in Section 2.1). By using pseudo-color display technology, we convert the temperature difference matrix of all pixels into an image, as shown in Fig. 9. Furthermore, the absolute values of temperature differences in each facial region are presented in Table 4. It is observed that the most significant difference mostly appears in nose region, where the apparent temperature change occurs, with some differences up to 7.8 °C. Next are the left and right cheeks, mouth, chin, eyes, and forehead. Moreover, it is noticed that the temperature of forehead region remains relatively stable (with a maximum difference of 2 °C). Considering that some subjects wear glasses that can affect the infrared image, the eye region is excluded from subsequent analysis. The remained six facial regions are selected as key regions for feature extraction in the fixed position experiment. They are nose, R-cheek, L-cheek, mouth, chin, and forehead.

3.2. Correlation between ambient temperature and facial regions' temperatures

The infrared images of 30 subjects under different thermal environmental conditions are observed and analyzed through the fixed-position experiment (Experiment content 2). The facial infrared images exhibit different change patterns with ambient temperature's variation. To show this variation, three typical subjects are picked, and their infrared images, temperature histograms, and reported thermal sensations under three different temperatures are shown in Fig. 10.

From the facial infrared images of subjects, it is observed that the facial temperature changes with the variation of the ambient temperature. When the ambient temperature drops from about 30 °C to about 20 °C, the values of the subjects' facial region temperatures decrease significantly. Among them, the temperature drops of cheek and nose are more pronounced than other regions. In addition, for the same row of infrared images in Fig. 10 (approximately the same T_{air}), the individual differences between the temperatures of subjects' facial regions could be obviously observed.

To show the skin temperature distribution of the entire human face, the histogram is made to count the distribution of pixel temperature values within the oval frame. From the pixel temperatures' histograms, the temperature distributions in the facial regions of the three subjects are not the same. In the thermal environment reported as "hot" by TSV (top row), the facial temperature distributions of subject (#2) and (#3) are more concentrated on the high temperature side (35 °C–38 °C) of the graphs. Relatively, subject (#1) exhibits a different facial temperature distribution, with facial temperature concentrated at 33 °C–36 °C. This phenomenon possibly dues to sweating on the face. Subject (#2) is more prone to sweating, and the evaporation of sweat regulates facial temperature, resulting in differences of temperature distribution. In addition, we find that in the second and third row of images, even though the indoor air temperature is similar, individual differences cause differences in subjects' reported thermal sensations.

Further, the correlations between the human face's sensitive regions and the ambient temperatures are analyzed. According to the six key facial regions determined by sensitivity analysis, ThermoScope software is used to mark each ROI in the collected images (as shown in Fig. 11),

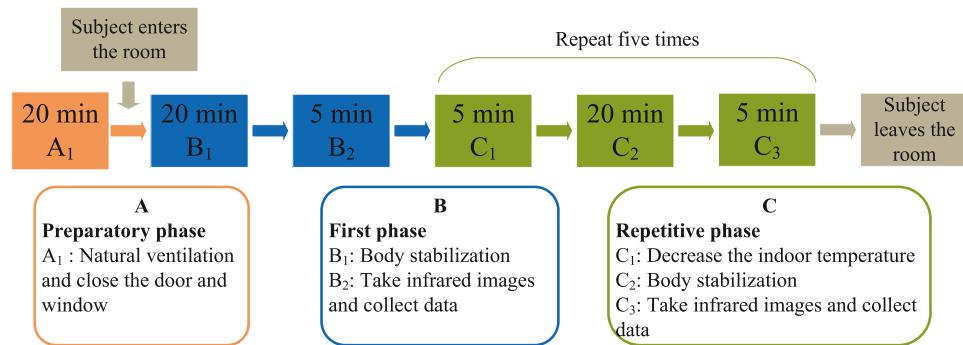


Fig. 7. The basic experiment procedure.

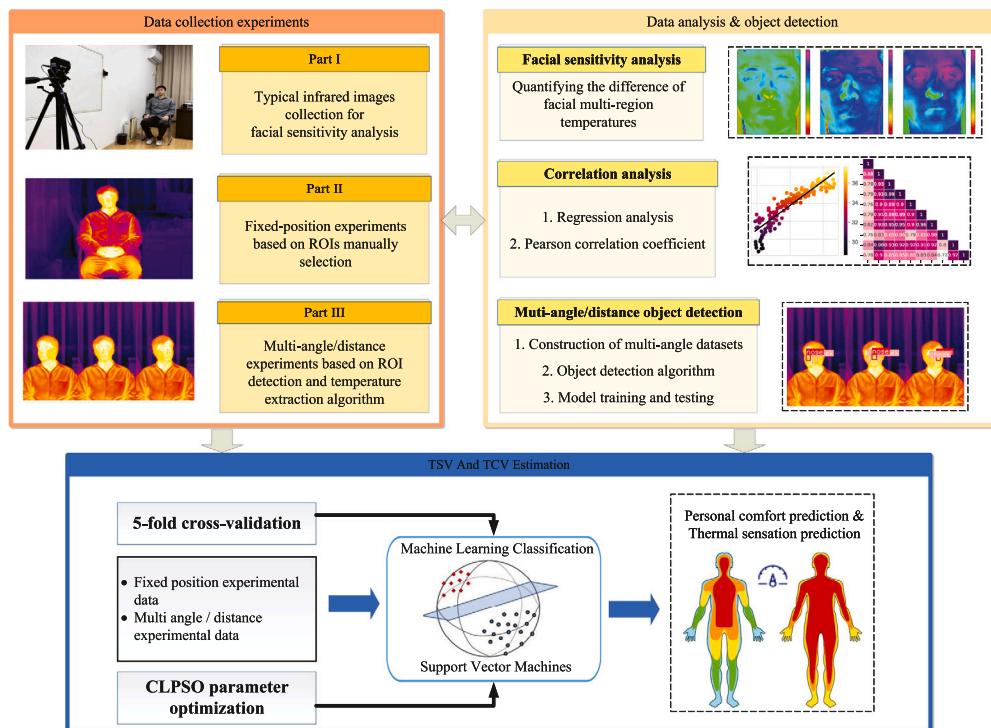


Fig. 8. The overall research framework.

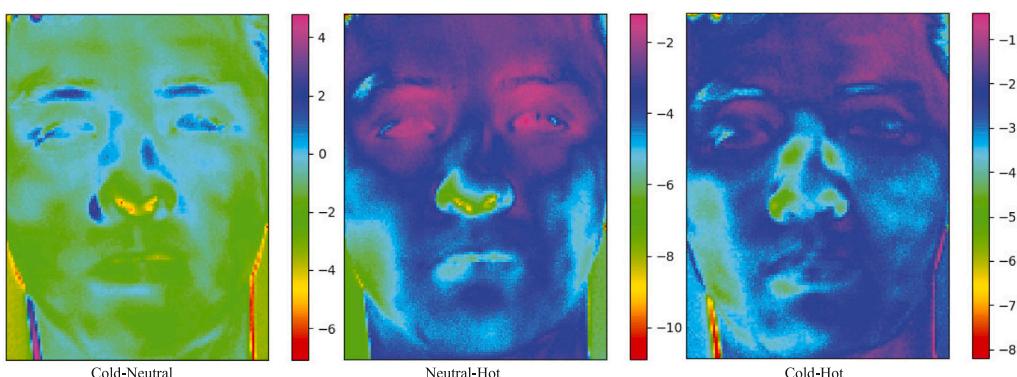


Fig. 9. Temperature differences between thermal images: Cold-Neutral (left), Neutral-Hot (Median), Cold-Hot(right).

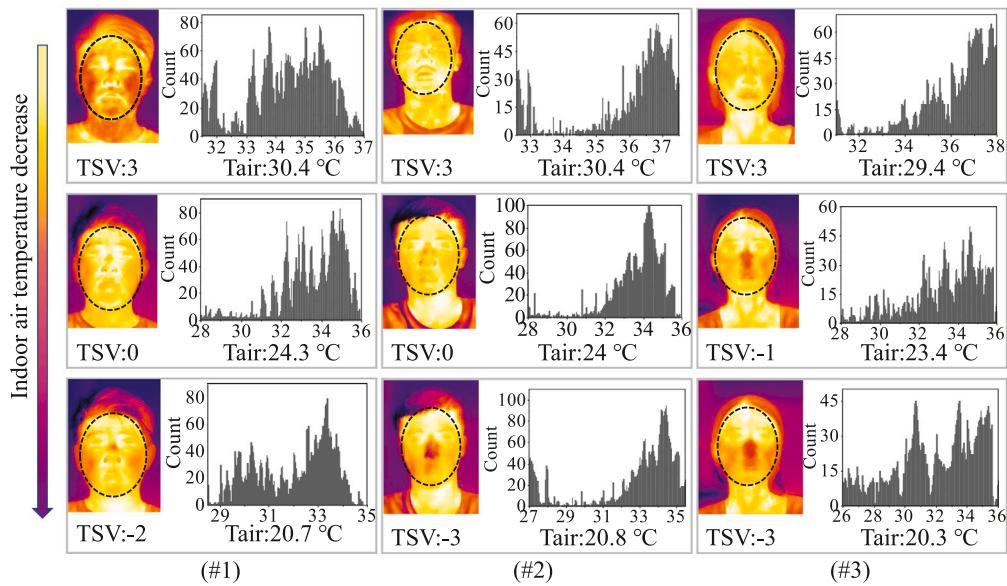


Fig. 10. Infrared images and temperature histograms at different air temperatures.

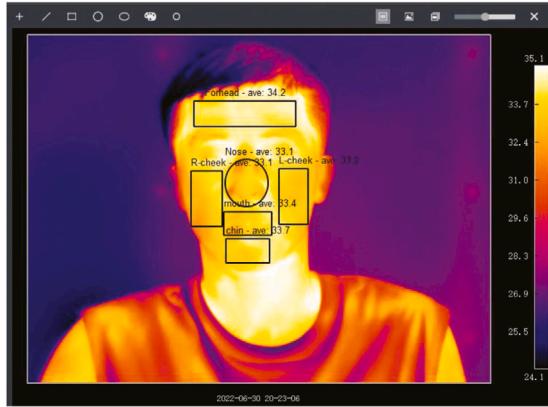


Fig. 11. Infrared image taken by MAGNITY33 (boxes represent the manually labeled ROI).

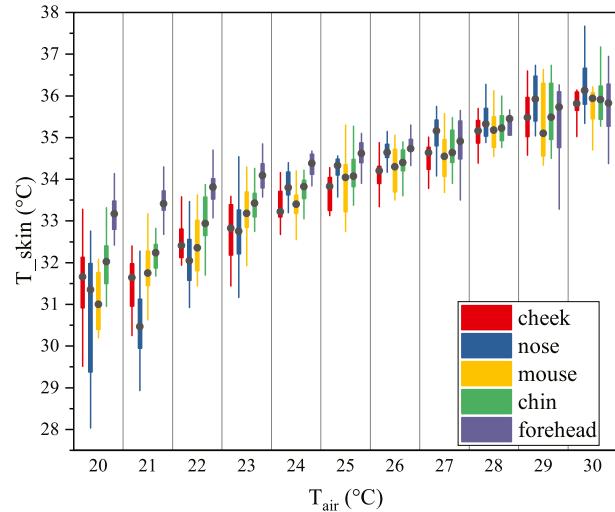


Fig. 12. Box line diagram of facial temperature change.

based on which, the correlation analysis of ROI's average temperature is conducted. Be noted: Thermoscope is a thermal camera supporting software developed by MAGNITY. By using this tool, the average temperature values of ROIs could be extracted directly.

Results indicate a linear correlation between the temperature of all facial regions and indoor air temperature. But the skin temperature of different facial regions is affected by indoor air temperature differently. Fig. 12 shows the boxplot of temperature changes for each facial region under different indoor air temperatures.

Generally, when the indoor ambient temperature rises from 20 °C to 30 °C, the facial temperature rises from 31 °C to 37 °C around. For different ambient temperature intervals, the lower the indoor air temperature, the more significant the temperature difference between various facial regions.

For higher ambient temperature, due to the limitation of the human core temperature, the skin temperature changes tend to flatten out and stabilize around 36 °C. Among the six facial regions, the temperature of forehead exhibits the smallest fluctuation difference, and the temperatures of mouth and chin are very similar, which may be due to their close proximity. But when indoor air temperature is below 24 °C, the temperature of the mouth is significantly lower than that of the chin. When indoor air temperature is low, the nose temperature is noticeably

lower than other regions. This phenomenon of rapid heat loss in nose region may be linked to the protruding structure of nose [46].

In addition, the rising trends of various facial regions are different. To observe these differences more clearly, the linear fitting method is used to plot the relationship between facial temperatures and ambient temperature, as shown in Fig. 13. In the figure, the obtained R^2 values are as follows: 0.82 for nose, 0.67 for forehead, 0.81 for mouth, 0.85 for chin, and 0.86 for cheeks. The corresponding fitting functions are also provided in the figure. By comparing the slopes of all fitting lines, it is indicated that the linear trend is most apparent in nose and least in forehead, which implies that nose is the most sensitive to indoor ambient temperature changes. This finding is consistent with the sensitivity quantification results in Section 3.1.

3.3. Relationship between skin temperature and TSV, TCV

Pearson correlation coefficients analysis is performed between the facial temperatures and thermal sensation votes data, as shown in Fig. 14. From the relationship between each facial region and indoor

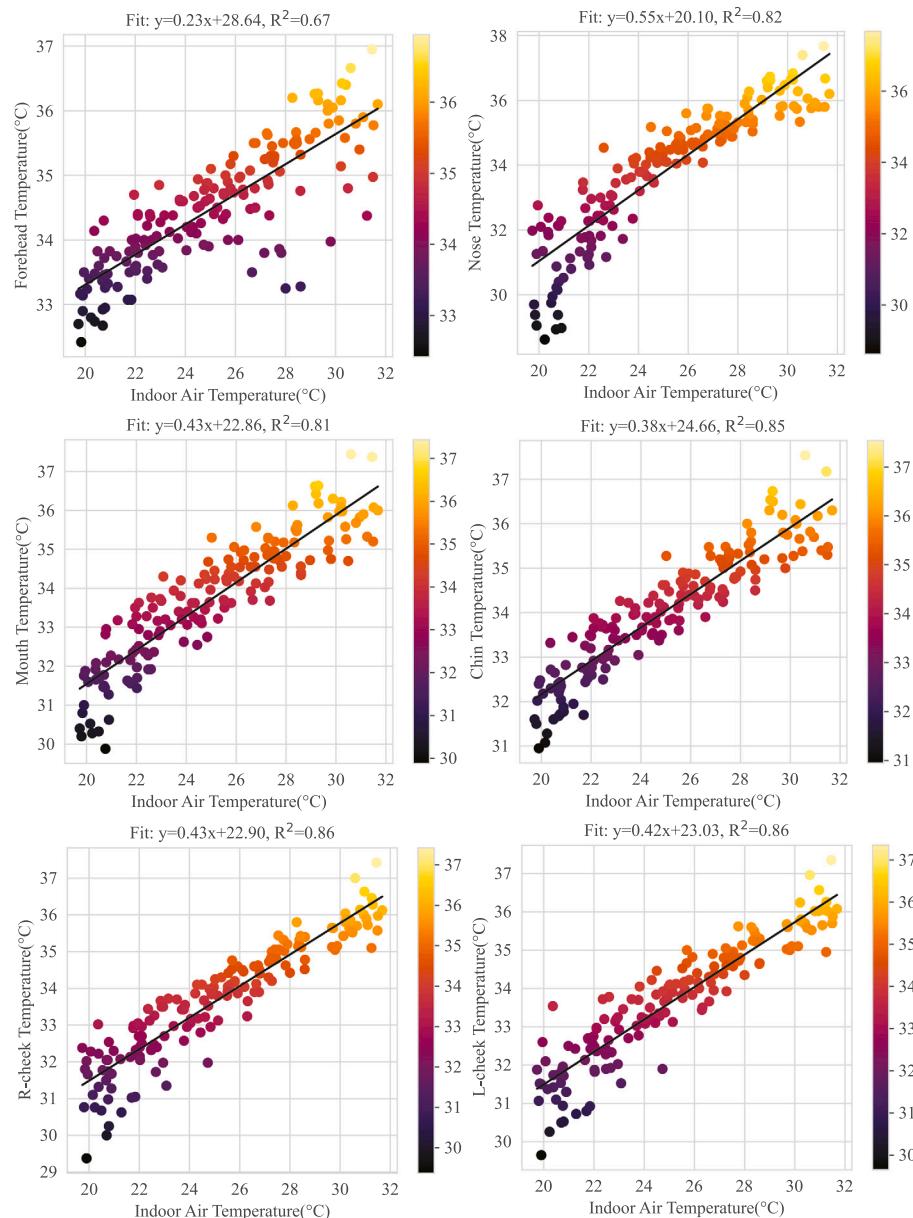


Fig. 13. Relationship between facial skin temperature and indoor air temperature.

air temperature, the forehead has the lowest correlation with indoor air temperature. The temperature of cheeks has the highest correlation with TSV/TCV followed by nose and chin. The correlation between forehead and TSV/TCV is relatively low. These findings are consistent with the results of previous research [29]. We find that the study of Wu [31] and Ashrafi [47] also identified the cheeks and nose regions as the best indicators of human thermal comfort.

Based on the above analysis, the nose and cheek regions are chosen as the key facial regions for multi-angle/distance infrared objection detection. Compared to identifying the temperatures of six regions (nose, left/right cheeks, forehead, mouth, chin), two regions (nose and cheeks) are more easily captured by cameras when the human posture is not restricted.

3.4. Results of multi-angle/distance facial region recognition

(1) Parameter setting and recognition results

Based on the correlation analysis results mentioned above, a multi-angle/distance facial region recognition framework for non-invasive

Table 5
System platform and parameter setting.

System platform	Parameter setting	
OS	size	640*640
GPU	batch size	16
CUDA	epoch	300
Programming language	learning rate	0.01
Deep learning framework	weight decay	0.005

thermal comfort evaluation is designed using the YOLOv5 algorithm. The system platform and parameter settings are listed in Table 5. The infrared image dataset (898 images) constructed in Section 2.2 (2) is randomly divided into training set/test set/validation set in a ratio of 7:2:1 for model construction.

In order to validate the effectiveness of the object detection model, this study conducts a preliminary evaluation of the YOLOv5 algorithm on the test set by the indexes of precision rate, recall rate, and detection

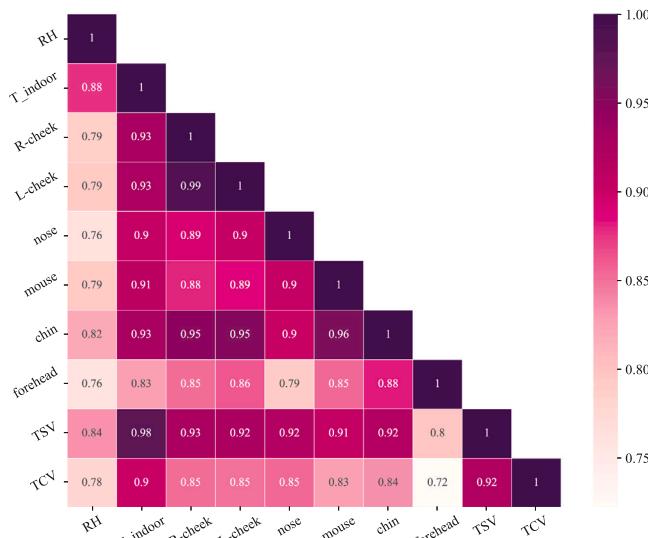


Fig. 14. Correlation heat map between parameters.

Table 6
Results of YOLOv5 model evaluation.

Model	Precision	Recall	FPS
YOLOv5	98.3	96.7	78.74

speed, which are formulated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{FPS} = \frac{1000ms}{\text{sum}(t)} \quad (11)$$

where TP represents the number of correctly detected samples, FP represents the number of samples detected as “nose and cheek” but actually belonging to the “non-nose and non-cheek” region, and FN represents the number of samples belonging to the “nose and cheek” region but incorrectly detected as “non-nose and non-cheek”. The closer the precision is to 1, the more likely the object detection model is to correctly recognize the “nose and cheek” regions. Recall represents the proportion of samples correctly detected as “nose and cheek” in the total number of classifications in the test set. FPS (frames per second) is the number of frames processed per second, where $\text{sum}(t)$ is the sum of model pre-process time, inference time, and NMS (non-max Suppression) processing time. The evaluation results of YOLOv5 on the test set are shown in Table 6.

Fig. 15 shows examples of infrared detection for a typical subject at four distances and nine different horizontal angles. It can be observed that this method could accurately recognize and classify the key facial regions of “nose and cheeks”.

(2) Performance investigation in practical applications

Further, to investigate the performance of the trained YOLOv5 model in practical scenarios, an extra experiment is conducted on eight new participants who did not attend earlier experiment. The participant with the highest detection performance and the participant with the lowest detection performance are excluded from the subsequent analysis. To capture facial images within a range of 0° – 180° , we recorded a video of the participants gradually turning their heads horizontally from right to left by 180° . Camera distance could be within a range of 1–4 m. For convenience, only four fixed distances of 1, 2, 3, and 4 meters are recorded, and by this method, a total of 4439 video frames

are collected for analysis. Then, the trained YOLOv5 model is applied for object detecting of “nose and cheek” regions to further evaluate its accuracy.

The YOLOv5 model successfully detects the nose and cheeks in 3937 of 4439 video frames (*Precision* : nearly 88.7%). The number of successful detections is higher at distances of 1 m and 2 m, and as the distance increases, the success rate of detection gradually decreases. This is due to limitations in infrared camera resolution and imaging texture. The success rate of nose and cheek detection varies among participants, but overall shows a higher success rate for nose detection than cheek. This discrepancy is primarily influenced by the subject’s facial features, as cheeks are usually flatter and less distinctive in appearance. The details of success rates of detection for six subjects are described in Fig. 16.

Additionally, we also count the recognition rates of all images from different angles, as shown in Fig. 17. When the infrared camera is roughly facing the face (80° – 100°), the detection effects of the nose and cheek are optimal, with accuracy rates of 98.3% and 96.9% respectively. As the face rotates to both sides, the detection rates gradually decrease. When the angle reaches 0° or 180° , the detection rates of the nose and cheeks are the lowest. Moreover, the recognition rates of the nose are all higher than that of the cheeks at each angle. Be noted: considering that it is impossible to detect both cheeks at all angles, the aforementioned results are based on the single cheek’s detection.

Based on the above analysis, it is inferred that the trained YOLOv5 object detection model can accurately recognize the nose and cheek regions in infrared images and provide their location information. Subsequently, by combining this object detection model with the infrared camera, the key skin temperatures of multiple subjects can be simultaneously extracted within the available ranges (1–4 m, 0° – 180°). As an example, Fig. 18 captures three subjects facing the camera at different angles. By combining with the object detection model, it is seen that the two facial region temperatures of all subjects can be recognized simultaneously.

4. Thermal comfort prediction

On the basis of the above object detection results, the CLPSO-SVM model described in Section 2 is further used to predict personal thermal comfort. In this section, the data collected from the fixed-position and multi-angle/distance experiments are separately used. For the fixed-position experiment, a total of 164 sets of sample data including thermal comfort questionnaire are collected from 30 subjects. For the multi-angle/distance experiment, a total of 150 sets of sample data including thermal comfort questionnaire are collected.

Firstly, a basic SVM model (with $c = 2$ and $g = 1$) is modeled for thermal comfort prediction. To compare the prediction performances of different facial regions, two different input datasets are used. One data set includes six facial regions’ temperatures, and the other one only contains two facial regions’ temperatures (nose, cheek). In addition, two environmental variables (air temperature and humidity) are both included as the inputs. The prediction accuracy of the model is evaluated with 5-fold cross-validation. The prediction targets are classified into 3-class of thermal comfort state and 7-class of thermal sensation (corresponding to Fig. 6). The results are listed in Table 7, from which, it is seen that compared to the six facial regions, using the nose and cheeks as detection targets can achieve higher thermal comfort prediction accuracy.

The 7-class prediction accuracy based on multi-angle/distance detection is better than that based on fixed-position detection ($84.07\% > 80.98\% > 74.23\%$). For 3-classification prediction, multi-angle/distance detection based thermal comfort prediction can also achieve nearly 90% accuracy (88.64%), which is similar with that of fixed-position detection’s results. For multi-angle/distance prediction, the general classification accuracy rate (General Acc) is the product of SVM classification accuracy rate (Acc) and YOLOv5 recognition rate (88.7%). In the

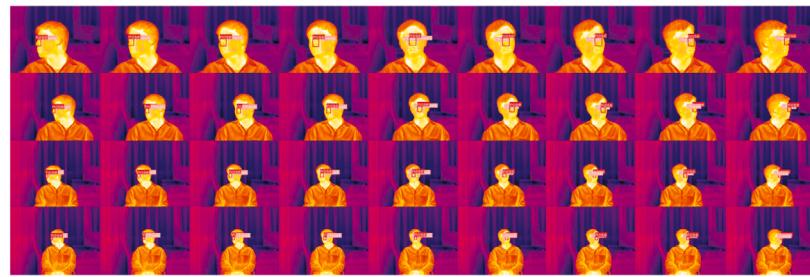


Fig. 15. Result of object detection.

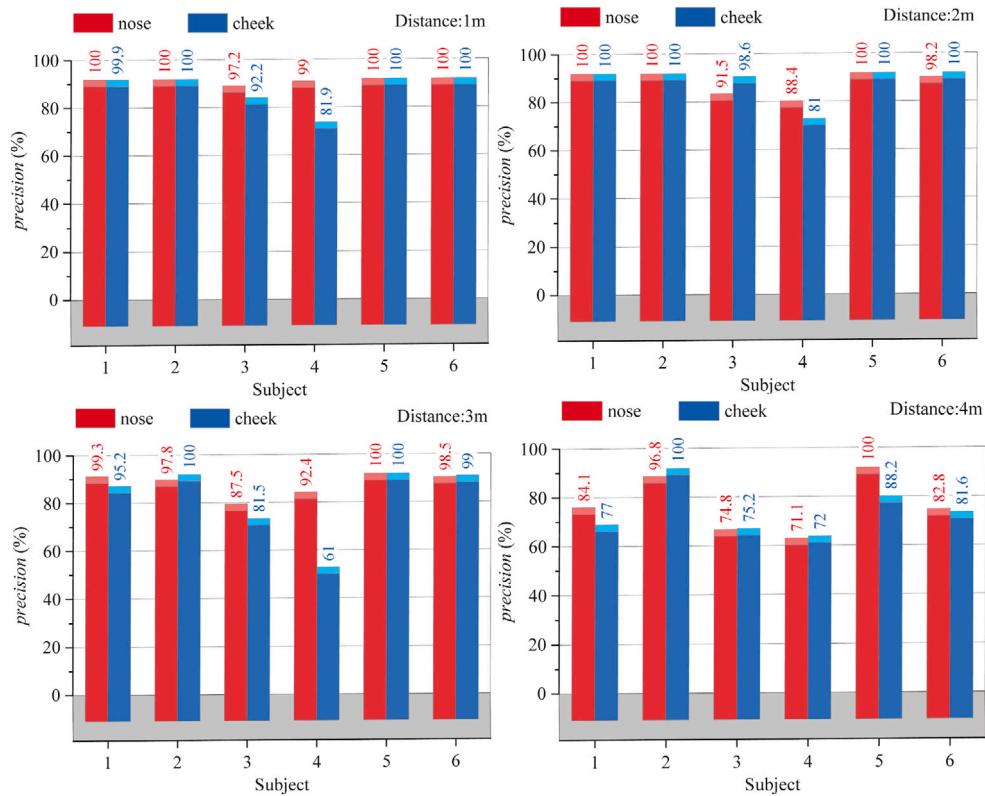


Fig. 16. Percentage of successful detection of ROI frames at different distances.

Table 7
Classification results of SVM model.

Experiment	Input parameters	Target	Acc	General Acc
Fixed-position	T_{air} , RH	3-class	88.96%	/
		7-class	68.71%	/
	T_{air} , RH , T_{skin} (all)	3-class	90.08%	/
		7-class	74.23%	/
	T_{air} , RH , T_{skin} (nose, cheek)	3-class	91.41%	/
		7-class	80.98%	/
Multi-angle/distance	T_{air} , RH	3-class	86.00%	/
		7-class	70.00%	/
	T_{air} , RH , T_{skin} (nose, cheek)	3-class	88.64%	78.62%
		7-class	84.07%	74.57%

Table 8
Classification results of CLPSO-SVM model.

Experiment	Input parameters	Target	Acc	General Acc
Fixed-position	T_{air} , RH	3-class	89.00%	/
		7-class	74.20%	/
	T_{air} , RH , T_{skin} (all)	3-class	92.64%	/
		7-class	74.85%	/
	T_{air} , RH , T_{skin} (nose, cheek)	3-class	93.87%	/
		7-class	80.98%	/
Multi-angle/distance	T_{air} , RH	3-class	94.70%	/
		7-class	91.30%	/
	T_{air} , RH , T_{skin} (nose, cheek)	3-class	96.59%	85.68%
		7-class	92.05%	81.65%

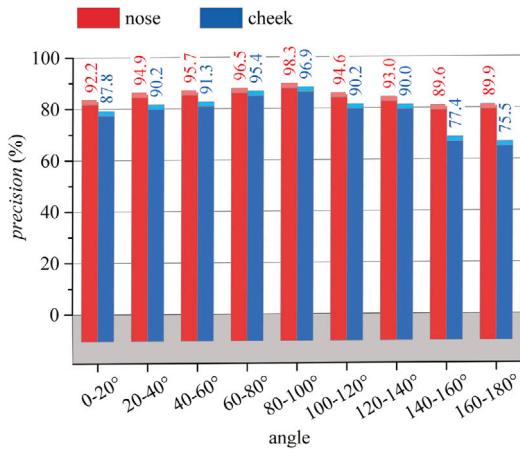


Fig. 17. Percentage of successful detection of ROI frames at different angles.

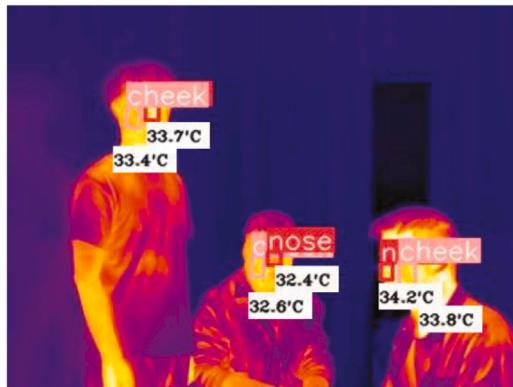


Fig. 18. An example for ROI detection and temperature extraction.

table, the predicted results using only two environmental variables are also provided for comparison. Since all subjects in the experiment were required to have similar clothing insulation and metabolic rates, the thermal sensation differences caused by personal factors are reduced and the correlation of thermal comfort with environmental variables is significant.

Furthermore, the CLPSO algorithm is applied to optimize the c and g of SVM for the purpose of accuracy improvement. The classification result of the CLPSO-SVM model is shown in Table 8. For both two object detection methods, the CLPSO-SVM model exhibits better accuracy than the basic SVM model, with an improvement of 0.62% to 7.98%. For both 3-class and 7-class predictions, the CLPSO-SVM model plus multi-angle/distance detection method achieve the best accuracies (96.59%, 92.05%). Fig. 19 shows the corresponding confusion matrixes. After multiplying by the recognition rate of YOLOv5 object detection,

Table 9

Performances of PMV and different data-driven models (multi-angle/distance, two facial region).

Model	Target	ACC
PMV	7-class	33.70%
RF	3-class	87.50%
	7-class	79.55%
ELM	3-class	88.64%
	7-class	72.73%
CLPSO-SVM	3-class	96.59%
	7-class	92.05%

the general accuracies of thermal comfort prediction are 85.68% and 81.65% respectively. Although the accuracies may be slightly reduced, the proposed method with less restrictions would have better engineering application value compared to fixed-position facial acquisition methods. It is noted that the parameter setting of CLPSO-SVM model is as follows: the population size of particles (pop) is set to 40, the acceleration factor (c_a) is set to 1.49, the initial inertia weight (w) is set to 0.9, the final inertia weight (w) is set to 0.4, and the termination condition is set to exceed 40 iterations.

In addition, the PMV and other two data-driven models, i.e., RF (Random Forest) and ELM (Extreme Learning Machines), are also realized for performance comparison. The relevant PMV parameters are measured values, such as air temperature, air flow velocity, relative humidity and clothing insulation. The metabolic rate is set to 1.0 met. From the brief results in Table 9, PMV cannot predict 7-Class human thermal sensation with high performance. For the multi-angle/distance thermal sensation assessment, the CLPSO-SVM model also achieves the best accuracies among the three data-driven models both in 3-Class and 7-Class predictions.

5. Discussion

Based on the results of above experimentation and modeling work, a brief discussion on this study is provided as follows:

(1) From the facial temperature data collected in the field experiment, it reveals a clear linear relationship between skin temperature of various facial regions and indoor air temperature. During the thermal responding procedure, the forehead is the least affected by indoor air temperature. The nose and cheeks are most affected by environmental temperature and have strong correlations with TSV/TCV. From the results of Section 4, only using two environmental variables (T_{air} , RH) also achieved high prediction accuracy. Be noted that to validate the performance of proposed method, the subjects in the experiment have similar clothing insulation and metabolic rates, which enhances the linear correlation of thermal comfort with environmental variables. And when more personal factors are considered, non-invasive predictive method that use facial temperatures as physiological indicators would have more performance advantages.

(2) YOLOv5 can be used for object detection in infrared images. The recognition accuracy is better within 2 m than at 3 m or 4 m.

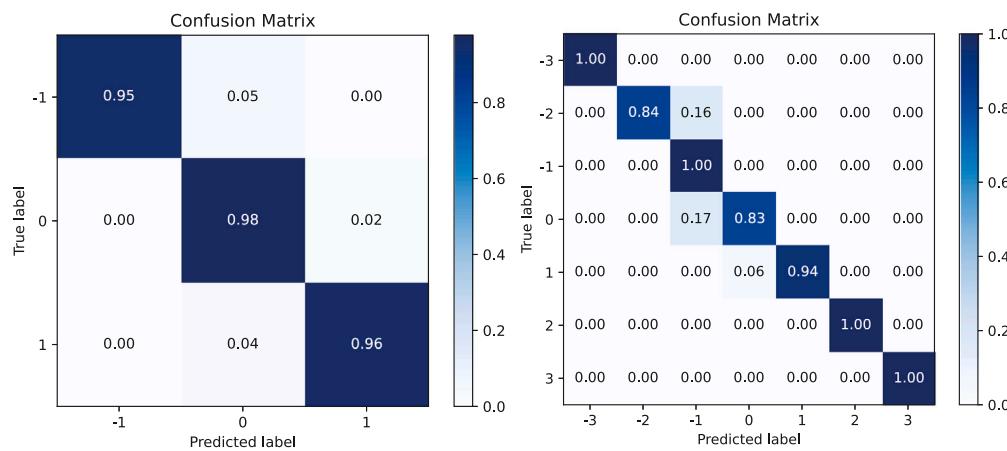


Fig. 19. The confusion matrixes of 3-class and 7-class thermal comfort predictions using CLPSO-SVM model.

The highest recognition rate occurs when the face is directly facing the camera (90°), and as the face rotates to both sides, the recognition rate gradually decreases. In addition, the recognition rate for nose region is better than that for cheek, which may be due to the flatness of the cheek without obvious features.

(3) From the classification results of the thermal comfort model, it is shown that the CLPSO algorithm can effectively improve SVM's classification accuracy by optimizing two parameters of c and g . Moreover, the use of more facial regions' temperature as inputs to the model does not necessarily result in higher classification accuracy. The nose and cheeks are not only the most sensitive to changes in indoor air temperature, but the temperatures obtained from these two regions also have been proven to be the most important facial factors for thermal comfort assessment. Based on our field experiment data, the overall accuracies of the proposed framework are 85.68% (3-class) and 81.65% (7-class) considering the recognition rate of object detection. Despite some of previous studies could achieve similar accuracies (85% in [29], 85% in [48], and 82.8% in [49], all 3-class prediction), the facial recognition rate was not taken into account and the restriction of face angle issue was not considered in their studies. Comparatively, the proposed multi-angle/distance modeling method would have wider application value than previous reported method.

This study also has some limitations. Firstly, some cases of facial obstruction are not considered (masks, glasses, etc.) in this paper. Secondly, our object detection model is only effective within the range of 0° to 180° . When the subject turns their back to the camera, the data collection will fail. Thirdly, our experiment is conducted only in summer. The correlation analysis results of the experiment and the performance of the thermal comfort prediction model may not apply to the heating season directly. In addition, since our data set is collected from a young population with normal BMIs, other types of population (such as the elderly) are not guaranteed to achieve the same predictive performance.

6. Conclusion

Accurate assessment of human thermal comfort was important for HVAC system control and building energy efficiency. This study conducts a series of infrared facial temperature collection experiments. Considering the recognizability of facial regions in practice, the nose and cheeks were selected as key facial regions for thermal sensation assessment. The YOLOv5 model was applied in conjunction with an infrared camera to collect ROI's temperature values. The model achieved a precision of 98.3% in identifying the nose and cheeks

regions using a dataset of infrared facial images (a total of 898 images). For eight new participants in practice, the recognition rate was 88.7%. Then the CLPSO-SVM model was used to forecast human thermal sensation through the collected facial temperature data. After incorporating the recognition rate of YOLOv5 object detection, the general accuracies of thermal comfort prediction were 85.68% (3-class) and 81.65% (7-class), respectively. Comparing to fixed-position facial acquisition methods, the proposed method of this paper was with less restrictions of human postures, and thus would have better engineering application values. In the future work, an extensive data set covering more individual differences would be collected, and more experiments in different seasons would be conducted for model validation.

CRediT authorship contribution statement

Kangji Li: Methodology, Investigation, Funding acquisition. **Weili Li:** Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Fukang Liu:** Visualization, Validation, Software, Data curation. **Wenping Xue:** Validation, Supervision, Resources, 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 National Natural Science Foundation of China (Grant No. 61873114) and “Six Talents Peak” High-level Talents Program of Jiangsu Province (Grant No. JZ-053).

Appendix

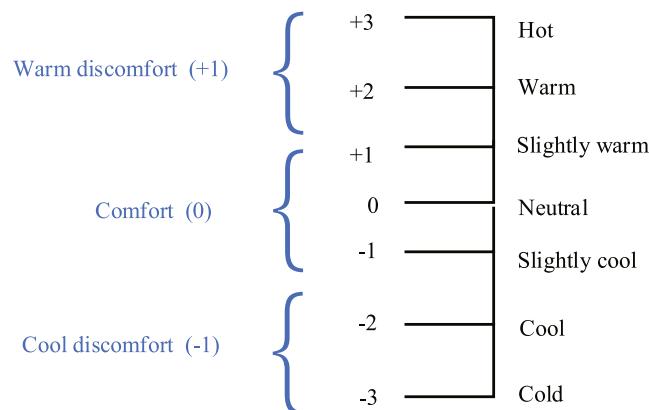
See [Table A.1](#) and [Fig. A.1](#).

Table A.1
The specific information of all sensors.

Target	Sensor type	Manufacturer	Measurement range	Accuracy	Resolution
Ambient air temperature	DHT11	YIMENGWEI, China	-20 °C~60 °C	± 2 °C	-
Relative humidity	DHT11	YIMENGWEI, China	20%~95%	±5%	-
Infrared camera	MAGNITY33	MAGNITY, China	20 °C~45 °C	±0.4 °C	384(h) * 288(v)

Thermal Sensation Questionnaire

Name:	Gender:
Weight:	Contact:
Age:	Height:



Test No.	Thermal state		Thermal sensation	
	Air temperature(°C)	Air humidity(%)	Thermal state	Thermal sensation
No.1				
No.2				
No.3				
No.4				
No.5				

Signature:

Fig. A.1. The example of thermal sensation questionnaire.

References

- [1] Y. Xiao, H. Huang, X.M. Qian, L.Y. Zhang, B.-W. An, Can new-type urbanization reduce urban building carbon emissions? New evidence from China, *Sustainable Cities Soc.* 90 (2023) 104410:1–15.
- [2] A. Ahmed, T. Ge, J. Peng, W.C. Yan, B.T. Tee, S. You, Assessment of the renewable energy generation towards net-zero energy buildings: A review, *Energy Build.* 256 (2022) 111755:1–20.
- [3] S.J. Ra, J.H. Kim, C.S. Park, Real-time model predictive cooling control for an HVAC system in a factory building, *Energy Build.* 285 (2023) 112860:1–19.
- [4] F. Zhang, N. Saeed, P. Sadeghian, Deep learning in fault detection and diagnosis of building HVAC systems: A systematic review with meta analysis, *Energy AI* 12 (2023) 100235:1–28.
- [5] T. Breideband, P.T. Sukumar, G. Mark, M. Caruso, S. D'Mello, A.D. Striegel, Home-life and work rhythm diversity in distributed teamwork: A study with information workers during the COVID-19 pandemic, in: Proceedings of the ACM on Human-Computer Interaction, Vol. 6, 2022, pp. 1–23.
- [6] T. Cheung, S. Schiavon, L.T. Graham, K.W. Tham, Occupant satisfaction with the indoor environment in seven commercial buildings in Singapore, *Build. Environ.* 188 (2021) 107443:1–12.
- [7] C. Huizinga, S. Abbaszadeh, L. Zagreus, E. Arens, Air quality and thermal comfort in office buildings: Results of a large indoor environmental quality survey, in: Proceedings of Healthy Buildings, Vol. 3, 2006, pp. 393–397.
- [8] G. Brager, L. Baker, Occupant satisfaction in mixed-mode buildings, *Build. Res. Inf.* 37 (4) (2009) 369–380.
- [9] G.A. Ganesh, S.L. Sinha, T.N. Verma, S.K. Dewangan, Investigation of indoor environment quality and factors affecting human comfort: A critical review, *Build. Environ.* 204 (2021) 108146:1–16.
- [10] ANSI/ASHRAE, ANSI/ASHRAE Standard 55-2020: thermal environmental conditions for human occupancy, ASHRAE (2020).
- [11] 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.
- [12] W. Jung, F. Jazizadeh, Comparative assessment of HVAC control strategies using personal thermal comfort and sensitivity models, *Build. Environ.* 158 (2019) 104–119.
- [13] S. Liu, S. Schiavon, H.P. Das, M. Jin, C.J. Spanos, Personal thermal comfort models with wearable sensors, *Build. Environ.* 162 (2019) 106281:1–17.
- [14] S.Y. Sim, M.J. Koh, K.M. Joo, S. Noh, S. Park, Y.H. Kim, K.S. Park, Estimation of thermal sensation based on wrist skin temperatures, *Sensors* 16 (4:1–11) (2016) 420:1–11.
- [15] J. Kim, Y. Zhou, S. Schiavon, P. Raftery, G. Brager, Personal comfort models: Predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning, *Build. Environ.* 129 (2018) 96–106.
- [16] W. Li, J. Zhang, T. Zhao, R. Liang, Experimental research of online monitoring and evaluation method of human thermal sensation in different active states based on wristband device, *Energy Build.* 173 (2018) 613–622.
- [17] S.I. ul Haq 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.
- [18] Y. Yao, Z. Lian, W. Liu, C. Jiang, Y. Liu, H. Lu, Heart rate variation and electroencephalograph – the potential physiological factors for thermal comfort study, *Indoor Air* 19 (2) (2009) 93–101.
- [19] B. Yang, X. Li, Y. Hou, A. Meier, X. Cheng, J.H. Choi, F. Wang, H. Wang, A. Wagner, D. Yan, A. Li, T. Olofsson, H. Li, Non-invasive (non-contact) measurements of human thermal physiology signals and thermal comfort/discomfort poses -a review, *Energy Build.* 224 (2020) 110261:1–10.
- [20] A. Ghahramani, G. Castro, B. Becerik-Gerber, X. Yu, Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort, *Build. Environ.* 109 (2016) 1–11.
- [21] K. Nkurikiyezu, G. Lopez, Toward a real-time and physiologically controlled thermal comfort provision in office buildings, in: The 14th International Conference on Intelligent Environments, 2018, pp. 168–177.
- [22] J.H. Choi, D. Yeom, Study of data-driven thermal sensation prediction model as a function of local body skin temperatures in a built environment, *Build. Environ.* 121 (2017) 130–147.
- [23] N. Bouzida, A. Bendada, X.P. Maldague, Visualization of body thermoregulation by infrared imaging, *J. Therm. Biol.* 34 (3) (2009) 120–126.
- [24] J. Ranjan, J. Scott, ThermalSense: Determining dynamic thermal comfort preferences using thermographic imaging, in: UbiComp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ISBN: 978-1-4503-4461-6, 2016, pp. 1212–1222.
- [25] B. Pavlin, G. Pernigotto, F. Cappelletti, P. Bison, R. Vidoni, A. Gasparella, Real-time monitoring of occupants' thermal comfort through infrared imaging: A preliminary study, *Buildings* 7 (1) (2017) 1–11.
- [26] A. Aryal, B. Becerik-Gerber, Skin temperature extraction using facial landmark detection and thermal imaging for comfort assessment, in: BuildSys '19: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, ISBN: 978-1-4503-7005-9, 2019, pp. 71–80.
- [27] A.C. Cosma, R. Simha, Thermal comfort modeling in transient conditions using real-time local body temperature extraction with a thermographic camera, *Build. Environ.* 143 (2018) 36–47.
- [28] H. Metzmacher, D. Wölki, C. Schmidt, J. Frisch, C. van Treeck, Real-time human skin temperature analysis using thermal image recognition for thermal comfort assessment, *Energy Build.* 158 (2018) 1063–1078.
- [29] D. Li, C.C. Menassa, V.R. Kamat, Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography, *Energy Build.* 176 (2018) 246–261.
- [30] A.C. Cosma, R. Simha, Using the contrast within a single face heat map to assess personal thermal comfort, *Build. Environ.* 160 (2019) 106163:1–12.
- [31] Y. Wu, B. Cao, Recognition and prediction of individual thermal comfort requirement based on local skin temperature, *Build. Eng.* 49 (2022) 104025:1–18.
- [32] B. Yang, X. Cheng, D. Dai, T. Olofsson, H. Li, A. Meier, Real-time and contactless measurements of thermal discomfort based on human poses for energy efficient control of buildings, *Build. Environ.* 162 (2019) 106284:1–10.
- [33] H. Na, J.H. Choi, H. Kim, T. Kim, Development of a human metabolic rate prediction model based on the use of Kinect-camera generated visual data-driven approaches, *Build. Environ.* 160 (2019) 106216:1–14.
- [34] A. Meier, W. Dyer, C. Graham, Using human gestures to control a building's heating and cooling system, in: In Proceedings of the 9th International Conference on Energy Efficiency in Domestic Appliances and Lighting, EEDAL, 2017, pp. 627–635.
- [35] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. Freeman, Eulerian video magnification for revealing subtle changes in the world, *ACM Trans. Graph.* 31 (4) (2012) 65:1–8.
- [36] X. Cheng, B. Yang, T. Olofsson, G. Liu, H. Li, A pilot study of online non-invasive measuring technology based on video magnification to determine skin temperature, *Build. Environ.* 121 (2017) 1–10.
- [37] W. Jung, F. Jazizadeh, Vision-based thermal comfort quantification for HVAC control, *Build. Environ.* 142 (2018) 513–523.
- [38] F. Jazizadeh, W. Jung, Personalized thermal comfort inference using RGB video images for distributed HVAC control, *Appl. Energy* 220 (2018) 829–841.
- [39] P. Li, P. Dai, D. Cao, B. Liu, Y. Lu, Non-intrusive comfort sensing: Detecting age and gender from infrared images for personal thermal comfort, *Build. Environ.* 219 (2022) 109256:1–13.
- [40] G. Jocher, K. Nishimura, T. Mineeva, R. Vilariño, Yolov5, 2020, Code repository <https://github.com/ultralytics/yolov5>.
- [41] C. Chih-Chung, L. Chih-Jen, LIBSVM: A library for support vector machines, *ACM Trans. Intell. Syst. Technol.* 2 (3) (2011) 1–27, <http://dx.doi.org/10.1145/1961189.1961199>.
- [42] B. Yang, X. Li, Y. Liu, L. Chen, R. Guo, F. Wang, K. Yan, Comparison of models for predicting winter individual thermal comfort based on machine learning algorithms, *Build. Environ.* 215 (2022) 108970:1–17.
- [43] Z. Qavidel Fard, Z.S. Zomorodian, S.S. Korsavi, Application of machine learning in thermal comfort studies: A review of methods, performance and challenges, *Energy Build.* 256 (2022) 111771:1–20.
- [44] J. Liang, A. Qin, P. Suganthan, S. Baskar, Comprehensive learning particle swarm optimizer for global optimization of multimodal functions, *IEEE Trans. Evol. Comput.* 10 (3) (2006) 281–295.
- [45] K. Li, R. Yu, Y. Liu, J. Wang, W. Xue, Correlation analysis and modeling of human thermal sensation with multiple physiological markers: An experimental study, *Energy Build.* 278 (2023) 112643:1–15.
- [46] T.K. Bergersen, A search for arteriovenous anastomoses in human skin using ultrasound Doppler, *Acta Physiol. Scand.* 147 (2) (1993) 195–201.
- [47] R. Ashrafi, M. Azarbayjani, H. Tabkhi, Charlotte-ThermalFace: A fully annotated thermal infrared face dataset with various environmental conditions and distances, *Infrared Phys. Technol.* 124 (2022) 104209:1–16.
- [48] B.B.-G. Ashrant Aryal, A comparative study of predicting individual thermal sensation and satisfaction using wrist-worn temperature sensor, thermal camera and ambient temperature sensor, *Build. Environ.* 160 (2019) 106223:1–11.
- [49] A. Ghahramani, G. Castro, S.A. Karvigh, B. Becerik-Gerber, Towards unsupervised learning of thermal comfort using infrared thermography, *Appl. Energy* 211 (2018) 41–49.