



## Occupant-centric cabin thermal sensation assessment system based on low-cost thermal imaging

Zhenyu Hou<sup>a,c,1</sup>, Junmeng Lyu<sup>b,1</sup>, Dongyuan Wu<sup>a,c</sup>, Jiangping Chen<sup>a,c</sup>, Junye Shi<sup>a,c</sup>, Zhiwei Lian<sup>b,\*</sup>

<sup>a</sup> School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

<sup>b</sup> School of Design, Shanghai Jiao Tong University, Shanghai, China

<sup>c</sup> Shanghai High Efficient Cooling System Research Center, Shanghai, China



### ARTICLE INFO

**Keywords:**

Vehicle cabin  
Thermal sensation  
Thermal imaging  
Facial recognition  
Assessment system  
Low-cost

### ABSTRACT

Assessing cabin occupants' thermal sensation in real-time enables automatic control of car air conditioning, improving driving safety and energy efficiency. Traditional methods, limited by their need for numerous physiological parameters, have restricted practicality. For this purpose, this study employs a low-cost thermal imaging sensor as the hardware core to establish a cost-effective and non-contact real-time assessment system for the thermal comfort of occupants within a cabin. The system comprises a thermal sensation assessment model developed based on subject experiments, along with facial recognition and segmentation algorithms optimized for thermal images. The thermal sensation assessment model developed, which employs cheek temperature, solar radiation intensity, and cabin air temperature as its features, exhibited an  $R^2$  of 0.617 on the test set. Furthermore, the facial recognition algorithm established for thermographic imaging achieved a mean accuracy of 96.5% and a recall rate of 99.0%. The system underwent validation in real-world vehicle environments, proving its ability to accurately detect and measure the cheek temperatures of occupants in the cabin and execute thermal sensation assessments. With a mean absolute error of 0.5 thermal sensation units in its output, its accuracy in practical applications was affirmed. This research provides an effective solution for automatically adjusting cabin air conditioning.

### Abbreviations

$(x_1, y_1)$	Center position of the left eye
$(x_2, y_2)$	Center position of the right eye
$(x_3, y_3)$	Center position of the nose
$(x_4, y_4)$	Left corner of the mouth
$(x_5, y_5)$	Right corner of the mouth
$(x_{box}, y_{box})$	Top left corner of the face frame
$h_{box}$	Frame height
$w_{box}$	Frame width
$(x_{cheek\_lfup}, y_{cheek\_lfup})$	Top left corner of the cheek area
$(x_{cheek\_rtdw}, y_{cheek\_rtdw})$	Bottom right corner of the cheek area
$T_{\max}$	Highest temperature in the 192*256 temper matrix, °C
$T_{\min}$	Lowest temperature in the 192*256 temper matrix, °C
$T_{air}$	Air temperature, °C
$T_0$	Initial cabin air temperature, °C

(continued on next column)

### (continued)

$T_e$	Stabilized cabin air temperature, °C
$S_r$	Solar radiation, W/m <sup>2</sup>
$T_{incheek}$	Skin temperature of the cheek on the non-window side, °C
$T_{outcheek}$	Skin temperature of the cheek on the window side, °C
$T_{forehead}$	Forehead skin temperature, °C
$T_{mouth}$	Perioral skin temperature, °C
$T_{nose}$	Nose skin temperature, °C
MSE	mean squared error
MAE	mean absolute error
ENR	Elastic Net Regression
RFR	Random forest regression
TSV	Thermal sensation vote
A/C	Air conditioner

### 1. Introduction

In current research on automotive indoor environments, the

\* Corresponding author.

E-mail address: [zwlian@sjtu.edu.cn](mailto:zwlian@sjtu.edu.cn) (Z. Lian).

<sup>1</sup> These two authors contributed equally and should be considered as co-first authors.

necessity of studying automatic air conditioning control systems aimed at comfort is increasingly prominent. This approach mitigates the safety issues introduced by drivers manually adjusting air conditioning settings while driving. On the other hand, by identifying passenger comfort to dynamically adjust air conditioning set point, energy utilization can be optimized, reducing energy consumption and enhancing the overall energy efficiency of the vehicle.

Existing researches on incorporating thermal sensation feedback into automotive automatic control systems primarily focuses on two categories: The first involves applying the Fanger's Predicted mean vote (PMV) model in air conditioning systems [1–3]. However, the PMV model assumes that the human body is in a stable state, where skin temperature and sweating rate are influenced solely by metabolic rate [4–6]. Utilizing the PMV model to assess human thermal sensation in non-neutral and transient environments can lead to significant errors. The second category explores the feasibility of automatically adjusting the cabin environment by real-time monitoring of physiological parameters, such as skin temperature [7–11], and heart rate [12]. Among them, skin temperature changes due to the dilation and constriction of blood vessels driven by the body's perception of heat and cold, establishing a correlation between skin temperature and thermal sensation [13–16]. Existing thermal sensation assessment models based on skin temperature developed for cabin environments include the physiological-thermal sensation calculation model by Zhang et al. [9] from UC Berkeley (UCB), the transient thermal sensation calculation model for cabin occupants developed by Zhou et al. [11], and the computational model developed by Li et al. [17], which uses skin temperatures from seven different body parts. Moreover, due to the dynamic and non-uniform characteristics of cabin environments, thermal sensation models developed for dynamic environments are also considered applicable to cabin settings. For example, the dynamic thermal sensation (DTS) model by Fiala et al. [18] and the model proposed by Lai et al. [19], though neither was specifically designed for cabin environments [11] (the former based on an indoor building experiment dataset and the latter on an outdoor experiment dataset). However, these models all require the mean skin temperature of the human body or skin temperatures at multiple parts as inputs, and may also require core body temperature. Despite this, these models provide insights with a high accuracy into the changes in thermal sensation caused by heat exchange between occupants and their environment from a physical perspective. The extensive physiological parameter inputs make these models challenging to be applied in actual vehicles.

Given that non-contact measurement methods can minimize interference with individual activities, in recent years, the use of thermal imaging sensors to measure and extract skin temperature at specific regions for thermal sensation assessment have become a focal point in the field of intelligent thermal environment design [20–28]. Many studies have concentrated on facial temperature due to minimal obstruction on the face [20,22,27,28]. However, most of these studies aim to develop thermal sensation assessment models applicable to indoor building environments. These studies indicate that using facial temperature to assess human thermal sensation is feasible, but few have provided thermal sensation assessment systems that balance sensor costs with practical applicability from an application standpoint. Wu et al. [22] reviewed and summarized the thermal imaging sensors used in thermal comfort research. Among them, the disparity in the cost of sensors reaches a zenith of ninetyfold, with the premium-priced apparatus necessitating an expenditure of \$2700. Conversely, sensors that are economically priced, exemplified by the MLX90640, encounter challenges in facial recognition due to their inferior resolution. For applications based on thermal imaging to assess occupant thermal sensation, a more detailed workflow is required to complete the data acquisition to the output of the calculation results, including not only the calculation of thermal sensation, but also the effective acquisition and integration of the input values, which involves face recognition, temperature region segmentation, and data filtering. That is, the performance of the system

is influenced not only by the accuracy of the thermal sensation assessment model but also by the performance of the sensors and the algorithms used within the workflow. Exclusively pursuing model accuracy while neglecting practical applicability leads to a disconnect between laboratory models and real-world engineering applications.

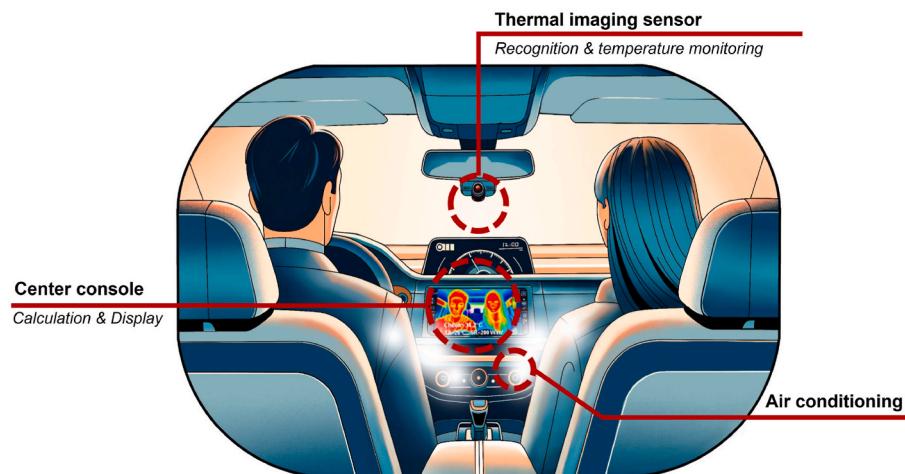
In the domain of visible light image processing, significant advancements have been achieved in the technology of facial parsing. For instance, Google's Mediapipe model can output an estimation of 478 three-dimensional facial landmarks, achieving precise segmentation of the face [29]. However, in the infrared image processing field, the lower resolution of infrared images compared to visible light images somewhat limits their capability to capture facial details. Moreover, facial parsing model development for infrared images is more complex. Due to differences in spectral characteristics, sensitivity to lighting conditions, and the physical information captured between visible light and infrared images, directly applying models designed for visible light images to infrared images often fails to achieve the desired effect. Kuzdeuov et al. [30] conducted an in-depth analysis of 15 major infrared image datasets and concluded that the majority lack annotations for bounding boxes and facial keypoints. Additionally, infrared image collection is often conducted under controlled indoor conditions, limiting their ability to reflect the complexity of real-world scenarios. Therefore, directly obtaining skin temperature data from specific facial regions using infrared images presents difficulties. A current mainstream solution is the fusion of visible and infrared images [31], but this approach involves privacy issues with visible light images. Compared to building indoor environments, cabin environments have unique advantages in addressing facial recognition challenges. The confined space of the cabin reduces the distance between thermal imaging sensors and occupants, allowing for a possible reduction in resolution requirements during the selection phase of thermal imaging sensors. This opens opportunities for high-accuracy but lower-resolution economical sensors. However, the development of a viable solution for facial recognition and region temperature extraction using lower-resolution thermal imaging sensors remains an unresolved issue.

To address existing research gaps, this study aimed to develop a practical, non-contact thermal sensation assessment system for cabin occupants, which is cost-effective (as illustrated in Fig. 1). Due to the cost constraints of thermal imaging sensors, the focus of developing this system was not only on building models but also on creating algorithms for real-time cabin occupant facial thermal imaging region segmentation and feature skin temperature extraction for lower-resolution thermal imaging sensors. This is an essential step in advancing the application of thermal imaging technology in HVAC systems. To develop this system, cabin subject experiments were conducted, and an open-source human portrait dataset was utilized to provide the necessary data foundation. Following its development, the system was deployed in actual vehicles for validation experiments to test its accuracy in real-world applications.

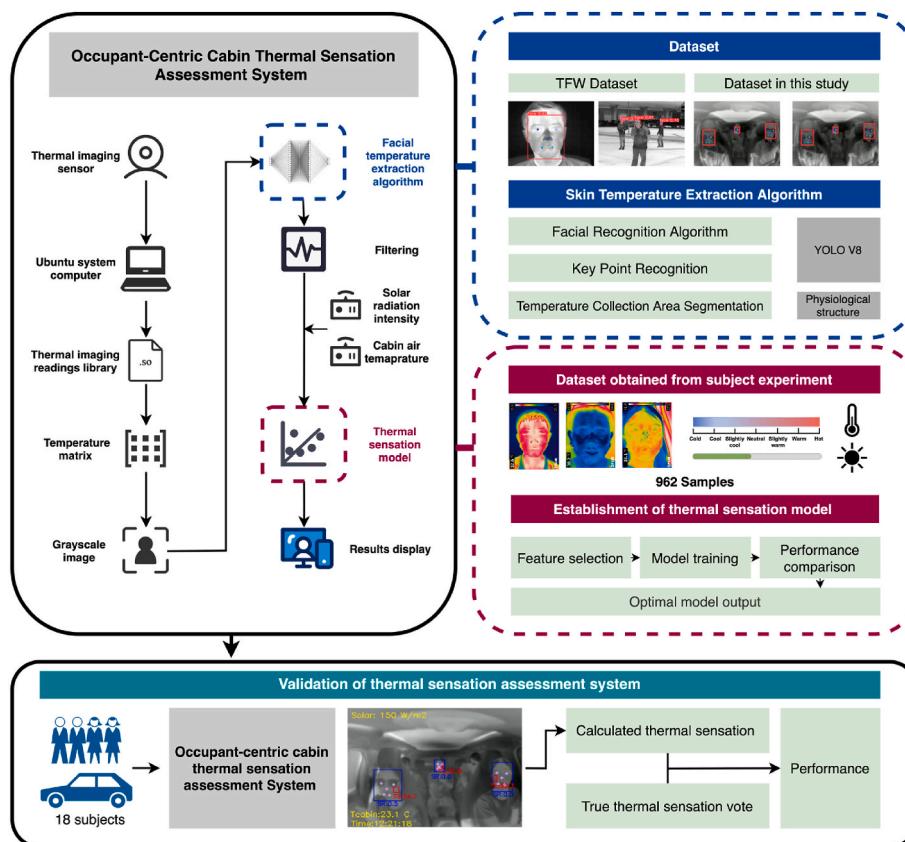
## 2. Methodology

### 2.1. Low-cost cabin thermal sensation assessment system framework

The work flow of the low-cost cabin occupant thermal sensation assessment system established in this study is illustrated in Fig. 2. The hardware core of the thermal sensation monitoring system is a customized low-cost thermal imaging sensor (IRAY Inc, China), with its specifications detailed in Table 1. Non-contact measurement of object temperatures using thermal imaging sensors is influenced by factors such as the materials of the object, imaging distance, and ambient temperature, which often makes high precision and high resolution inevitably associated with high costs. To address this issue, the customized thermal imaging sensor has been adjusted to specialize in measuring human skin temperature. This adjustment potentially diminishes its performance in other temperature measurement applications but enhances its efficacy for specific tasks, achieving a balance



**Fig. 1.** Schematic of practical application of this study.



**Fig. 2.** Framework for establishing a cabin thermal sensation assessment system.

**Table 1**

The specification of the custom low-cost thermal imaging sensor.

Specification	Value
Resolution	256 × 192
Horizontal field of view	120°
Temperature measurement range	25–40 °C
Operating temperature	-45 °C to +85 °C
Temperature measurement accuracy	±0.5 °C
Frame rate	25 Hz
Volume price	\$60

between cost-effectiveness, resolution, and measurement accuracy. The customized thermal imaging sensor has undergone blackbody radiation testing and recalibration under various conditions, including within the set range of skin temperatures, typical distances inside cabins. An example of imaging from this thermal imaging sensor is shown in Appendix [Figure S1](#). Comparison of readings with the high-performance thermal imager FLIR T-460 is shown in Appendix [Figure S2](#). The system also comprises a thermal imaging facial thermal imaging region segmentation and temperature extraction model, along with a thermal sensation assessment model. The thermal imaging sensor is wiredly connected to a computer running the Ubuntu operating system. The system employs the V4L2 framework for video stream reading and

sensor control. After data collection, the sensor's parsing function is invoked to process the video stream, outputting a  $192 \times 256$  temperature matrix, which contains temperature data within the range of  $[T_{min}, T_{max}]$ . These data are linearly mapped to the range of  $[0, 255]$  and converted into integers through rounding, thus forming a  $192 \times 256$  grayscale matrix that encompasses image information. This transformation allows the original temperature data to be presented in a visual format. Further, to meet the input requirements of the YOLO algorithm, grayscale duplication technique is employed to replicate each grayscale value three times, assigning them to the RGB channels respectively. This converts the single-channel black and white image into a three-channel pseudo-RGB image. After resizing, these images are fed into the facial region segmentation and temperature extraction model to obtain temperature data of various facial regions, which are ultimately recorded in CSV format. The system also includes built-in environmental parameter sensors of the vehicle, used to monitor the environmental data, read in real-time via the vehicle's CAN signal. These data are finally integrated into the developed thermal sensation model for real-time calculation of passenger thermal sensation, with the results displayed on a screen. The subsequent sections will detail each component.

## 2.2. Establishment of thermal sensation model

### 2.2.1. Subject experiment

This study conducted subject experiments to obtain the dataset necessary for constructing the thermal sensation assessment model for occupants in a vehicle cabin. The experiments were carried out from July to August 2023 in the real car cabin on the outdoor grounds of Shanghai Jiao Tong University. The sample size was calculated using G\*Power (F-test linear multiple regression). Assuming a regression with 5 independent variables, R<sup>2</sup> was set to 0.5, f<sub>2</sub> was set to 1.0, and  $\alpha$  error was set to 0.05. The minimum sample size calculated was 20 individuals, which gives a statistical power of 0.83. Ultimately, 24 subjects were recruited, meeting the requirement of ISO 14505-1 which stipulates that the number of subjects in automotive thermal sensation experiments should be no less than 8 [32]. The subjects included 23 university students and a 26-year-old female faculty staff, with detailed information provided in Table 2. All subjects possessed valid driving licenses and signed informed consent forms prior to the experiment and were advised to avoid caffeine, alcohol, and vigorous physical activities before the experiment. During the experiment, the clothing thermal resistance of the subjects was approximately 0.5 clo. The experimental scheme of this study was approved by the Science and Technology Ethics Committee of Shanghai Jiao Tong University.

As shown in Fig. 3, each experiment session spanned a minimum of 62 min, encompassing a pre-experimental adaptation phase of at least 20 and no more than 40 min, a 2-min transition period for relocating from the adaptation room to the cabin, and a 40-min formal experimental session inside the cabin. During the adaptation phase, subjects remained in a room characterized by a neutral environment (temperature maintained at 26 °C, air speed less than 0.2 m/s) to mitigate the influence of prior thermal experiences. Following the adaptation period, subjects transitioned from the room located on the third floor to the exterior vehicle, descending the stairs at a slow walking pace. At the beginning of the experiment, the subjects first entered the vehicle cabin to complete an initial overall thermal sensation vote (using the ASHRAE 7-point scale), followed by frontal head thermal imaging at 0.5 m. The air conditioning system was then activated and set to "cooling +

recirculation" mode. Subsequently, the subjects completed overall thermal sensation votes at fixed time intervals and underwent facial imaging. To ensure that the dataset encompasses a wide range of solar radiation intensity, formal experiments during the day were conducted in clear weather at three different times: 9:00–11:00, 12:00–14:00, and 15:00–17:00. The nighttime experiments were designed to simulate conditions without solar radiation, with the formal experiment session taking place from 20:00 to 21:00. During the daytime, each subject experienced three levels of initial cabin air temperatures of 35 °C (measured  $37.1 \pm 1.8$  °C), 45 °C (measured  $44.7 \pm 1.8$  °C), and 55 °C (measured  $53.9 \pm 2.2$  °C), and the nighttime experiments were conducted at an initial temperature of  $30.7 \pm 1.3$  °C. After the cabin was cooled down and stabilized, the cabin air temperatures were maintained at a temperature of between  $25.5 \pm 1$  °C, which is in the neutral temperature range specified by ASHRAE-55 [33]. Considering the varying relative directions of the car to the sun in real situations, the study adjusted the car's orientation during the day to achieve conditions where the sun directly shone on the windshield and laterally entered the cabin. During nighttime experiments, the cabin had a single orientation. For each orientation, each subject took part in one experiment per time period, totaling seven experiments. Two subjects were involved in each experiment, each sitting in a position directly exposed to the sun. Additionally, considering the limited space within the cabin and the proximity of passengers to the air conditioning vents, the direction of the vents was adjusted to 75° upward to minimize the draft sensation of cold air directly on subjects. In the preliminary experiments, the air speed in the human breathing zone was measured, with a mean air speed below 0.3 m/s (measured value:  $0.26 \pm 0.10$  m/s). Throughout the entire experiment, the vehicle's doors and windows remained closed. For safety reasons, subjects remained seated and did not drive during the experiment. Through this experiment, a total of 962 valid sample data sets from 24 subjects was collected. Of these, 46 sets of sample data were removed because overall thermal sensation votes were not completed correctly by the subjects or the thermal image files were not validly output.

In this experiment, a high-performance infrared thermal imager (FLIR T-460, FLIR Systems Inc, USA) with a resolution of  $320 \times 240$  pixels and an accuracy of  $\pm 1$  °C or 1 % of the reading was used to collect accurate skin temperature dataset. The thermal imager was calibrated at the factory to provide accurate measurements. Additionally, the thermal imager automatically performs flat-field calibration when sensor temperature changes, to compensate for the impact of reading drift. The emissivity was set to 0.98, a value suitable for human skin [34–36]. The reflected temperature was set to the cabin air temperature. The temperature collection region is shown in Fig. 4. For verification of instrument accuracy, please refer to our previously published paper [23]. During the experiment, a pyranometer (S-LIB-M003, ONSET Inc, USA) with an accuracy of  $\pm 5$  W/m<sup>2</sup> or  $\pm 5$  % of readings was horizontally positioned atop the instrument cluster to continuously collect solar radiation intensity transmitted through the cabin glass surface. The thermocouples (JTNT-A, Jantytech Inc, China) were used to measure the air temperature inside the cabin. The wires of the thermocouples were attached to the ceiling of the cabin, allowing the probes to hang in front of the occupants' faces. The seat positions were finely adjusted so that the thermocouple probes hung 0.1 m in front of the occupants' faces. The measurement accuracy of the thermocouple is  $\pm 0.5$  °C, which meets the requirements of ISO 7726-2001 for measurements [37]. In the preliminary experiments, the air speed at the height of the human breathing zone inside the cabin was measured using an omnidirectional anemometer (Swema 03+, SWEMA AB, Sweden) with an accuracy of  $\pm 0.03$  m/s.

### 2.2.2. Modeling algorithms

In the field of indoor thermal environment research, machine learning has been widely applied to the development of thermal sensation models [38–42]. Since the focus of this study is not on exploring the

**Table 2**  
Subjects information.

	Number	Age	Height (cm)	Weight (kg)
Male	12	$23.2 \pm 1.6$	$173.5 \pm 5.1$	$64.5 \pm 6.8$
Female	12	$24.4 \pm 2.1$	$163.5 \pm 5.5$	$54.1 \pm 4.8$

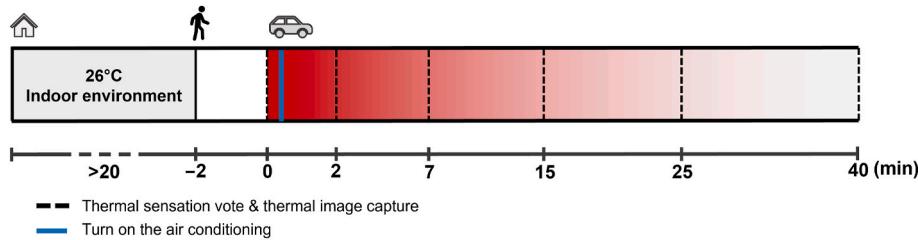


Fig. 3. The procedure of the experiment.

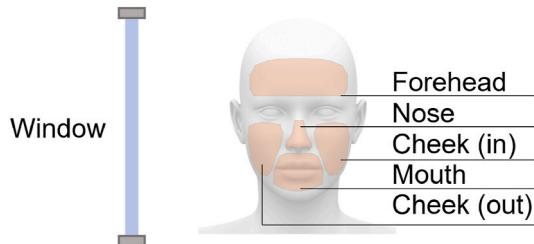


Fig. 4. Facial region segmentation.

best algorithm for modeling cabin occupant thermal sensation, this study has chosen to only select elastic net regression and random forest regression algorithms and compare their performances. Elastic net regression is a linear regression method that combines ridge regression and LASSO regression. It improves traditional linear regression models by incorporating both L1 and L2 regularization terms and has been proven effective in handling datasets with multicollinearity or high feature dimensions. Random forest regression algorithm, a nonlinear regression algorithm based on decision tree, has been proven effective and robust in studies of thermal comfort and thermal sensation prediction. The Scikit-learn library in Python was used to invoke the necessary algorithms [43].

When establishing the thermal sensation assessment model, the "train\_test\_split" function was used to divide the dataset, allocating 80 % as the training set and the remaining 20 % as the test set. The "test\_size" parameter was set to 0.2, and the "random\_state" for random sampling was set to 45. For evaluating the performance of the regression model, the Mean Squared Error (MSE) calculated using Equation (1) served as a metric to measure the model's accuracy. MSE represents the mean magnitude of the differences between the model's predicted values and the actual observed values. Additionally, the  $R^2$  metric, calculated using Equation (2), was introduced.  $R^2$  is typically used to evaluate the goodness of fit of a model, focusing more on the model's ability to explain the variability of the data. In practice, these two metrics are used together to comprehensively assess the performance of a model. Due to the sensitivity of the regression algorithm's performance to hyperparameters, a grid search with 5-fold cross-validation was employed, using  $R^2$  as the metric to find the best combination of hyperparameters. The parameter grid is presented in Table 3.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (1)$$

**Table 3**  
The hyperparameters involved in the optimization.

Algorithm	Hyperparameters	Range
RFR	n_estimators max_depth	1 to 20, with an interval of 1 None to 20, with an interval of 1
ENR	alpha	0.01 to 0.1 with an interval of 0.01 and 0.1 to 1 with an interval of 0.1
	l1_ratio	0.1 to 1, with an interval of 0.1

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n \left( Y_i - \frac{1}{n} \sum_{i=1}^n Y_i \right)^2} \quad (2)$$

Where,  $n$  is the number of observations.  $Y_i$  is the actual value of the observation.  $\hat{Y}_i$  is the predicted value from the model

### 2.3. Establishment of facial temperature extraction algorithm

#### 2.3.1. Dataset

The dataset relied upon to develop the facial temperature extraction algorithm was partially sourced from the TFW open-source dataset [44] and is partly collected by the research team in actual vehicle environments, recording 14 infrared video segments of 30 subjects under various in-cabin conditions. Based on this, 185 images were extracted from the video streams and manually annotated. To mitigate potential issues of portrait dataset imbalance, such as model bias, overfitting, and insufficient generalization capability, a rotation processing technique was employed to augment these 185 images, generating a total of 1110 images. Additionally, 40 images of empty cars under different environmental conditions were incorporated as background images. The final dataset comprises a total of 3950 images, divided into training, testing, and validation sets. The specific distribution and proportions are detailed in Table 4.

#### 2.3.2. Facial recognition and region segmentation algorithms

The facial recognition algorithm is built upon YOLOv8 [45]. The algorithm encompasses two primary tasks: Object Detection and Pose Estimation. Object Detection is utilized to capture the position of faces in the image. Pose Estimation, a regression prediction task, aims to locate keypoint coordinates within the identified facial region. In this algorithm, the five key points identified are the centers of the left eye, right eye, tip of the nose, left mouth corner, and right mouth corner. The model training was conducted on a pre-trained model established by Jocher et al. [45], which has been successfully applied in visible light image facial recognition and keypoint localization. Based on this, transfer training was performed on the training set images, with the parameters and weights of the convolutional and fully connected layers updated through the backpropagation algorithm. This adjustment ensures that the model accurately captures facial features in thermal

**Table 4**  
Overview of portrait dataset distribution.

Dataset	Total number of images	Outdoor images	Indoor images	On-vehicle images	Proportion
Training set	2777	1000	1000	777	70 %
Test set	862	300	300	262	22 %
Validation set	311	100	100	111	8 %
Total dataset	3950	1400	1400	1150	100 %

images.

The YOLO.train function was invoked for training using the ultralytics library in Python. During the model training process, pose loss was used as the primary metric to control iterations. This is based on the distance between the predicted and actual keypoints, while also incorporating elements like keypoint size and loss factors. After the completion of the model training, Precision (as defined in Equation (3)) and Recall (Equation (4)) were used to evaluate the performance of facial recognition. The performance of keypoint prediction was assessed using the Normalized Mean Error (NME) (Equation (5)), quantifying the deviation between the predicted keypoint locations and their actual positions.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{NME}(p, g) = \frac{1}{M} \sum_{i=1}^M \frac{\|p_i - g_i\|_2}{d} \quad (5)$$

Where,  $p$  represents the predicted positions of all keypoints,  $g$  represents the actual (ground truth) positions of all keypoints for each image,  $M$  denotes the number of keypoints, and  $d$  signifies the normalization factor, which, in this task, is taken as the width of the facial bounding box.  $\|p_i - g_i\|_2$  represents the Euclidean distance between the actual and predicted positions of a particular keypoint.

Although the above keypoints are identifiable, thermal images in real environments lack the detail necessary for detailed facial parsing, making it difficult to define boundaries between different regions based on variations in facial textures. The relative spatial structural has been employed to overcome these limitations. The facial recognition algorithm further infers facial subregions based on the structure of the face, guided by the positions of the keypoints. Despite significant variations in human facial structures, the relative positions of various facial regions follow certain patterns. Larrabee [46] proposed that a human face can be divided horizontally into three parts and vertically into five parts. Based on this theory, facial regions are divided and numbered (as shown in Fig. 5) to assist in the facial segmentation task. The nose region can be selected within area B3, determined in conjunction with the midpoint coordinates of the nose tip; the lip region can be chosen within the range from Ca2 to Cb4, determined by the coordinates of the left and right corners of the mouth; and the forehead region can be selected as the rectangular area from A2 to A4. Particularly, given the great variations

in the visible range of the cheeks under different postures, and since existing research indicates that cheek temperature is a key variable in assessing thermal sensation [23,47], a more refined strategy was designed for selecting the cheek area. Firstly, the intersection point of the vertical line through the eyes and the horizontal line through the nose is defined as the midpoint of the cheek boundary (as shown in Fig. 5). Building on this, additional strategies were developed and packaged (see pseudocode in Appendix Figure S3) to adapt to different scenarios of cheek selection.

All captured faces are categorized into two orientations (based on the left and right in the photo): For faces oriented to the right (meeting the condition  $|y_2 - y_3| < |y_1 - y_3|$ ), the left cheek is selected, with the bounding box extending to the left. The width of the box is set to  $\frac{2}{3}|x_1 - x_{box}|$ , and the height is set to  $\frac{1}{2}|y_3 - y_4|$ . For faces oriented to the left (meeting the condition  $|y_2 - y_3| > |y_1 - y_3|$ ), the right cheek is selected, with the bounding box extending to the right. The width of the box is set to  $\frac{2}{3}|x_2 - x_{box} - w_{box}|$ , and the height is set to  $\frac{1}{2}|y_3 - y_5|$ . Once the cheek position is determined, the algorithm matches the cheek region information with the occupant. Based on the height and central coordinates of the facial bounding box, the algorithm automatically archives data: If the center of the face is in the left area of the image (meeting the condition  $y_{box\_center} < 75$ ), data corresponding to a larger facial bounding box (meeting the condition  $25 < h_{box}$ ) is categorized as co-driver information, and the rest as rear co-driver passenger information. If the center of the face is in the middle area of the image (meeting the condition  $75 < y_{box\_center} < 175$ ), it is recorded as information for the middle rear passenger. If the center of the face is in the right area of the image (meeting the condition  $75 < y_{box\_center}$ ), data corresponding to a larger facial bounding box (meeting the condition  $25 < h_{box}$ ) is stored as main driver information, and the rest as rear main driver passenger information.

The above algorithm achieves facial segmentation, providing customized information for monitoring occupants' thermal sensation. To assess the accuracy of the cheek segmentation algorithm, a metric called Cheek Area Recognition Accuracy (CRRA) is proposed (refer to Equation (6)). This metric is defined as the ratio of the area of intersection between the algorithm-identified cheek region and the actual cheek region to the algorithm-identified cheek region. This metric evaluates the effectiveness of the selected area.

$$\text{CRRA} = \frac{S(\text{algorithm - identified cheek region} \cap \text{actual cheek region})}{S(\text{algorithm - identified cheek region})} \quad (6)$$

### 2.3.3. Temperature data filtering

In some scenarios, the selected cheek area may include non-skin elements. For instance, if the detected person is wearing glasses, the area of the glasses might be selected. When the tilt of the head is significant, the selected area may include the background. To address this issue, a data cleaning strategy is set. Firstly, a standard range for skin temperature is established. Existing study states that the human body experiences a tingling sensation when the skin temperature is below 20 °C or above 45 °C for an extended period of time [48]. The extreme outdoor conditions of winter can result in low facial temperatures when occupants enter a vehicle, but very few scenarios in summer result in skin temperatures as high as 45 °C, so the temperature range of 20–40 °C was ultimately chosen. Any temperature data outside this range are preliminarily excluded. Subsequently, a further filtering is applied to the remaining cheek skin temperature data. Assuming that the highest temperature point in the selected area is always the skin temperature, a temperature interval of 2.5 °C is set. If the difference between any temperature point and the highest temperature point exceeds 2.5 °C, that point is determined to be a non-skin temperature point and is excluded from the temperature data. This method further refines the

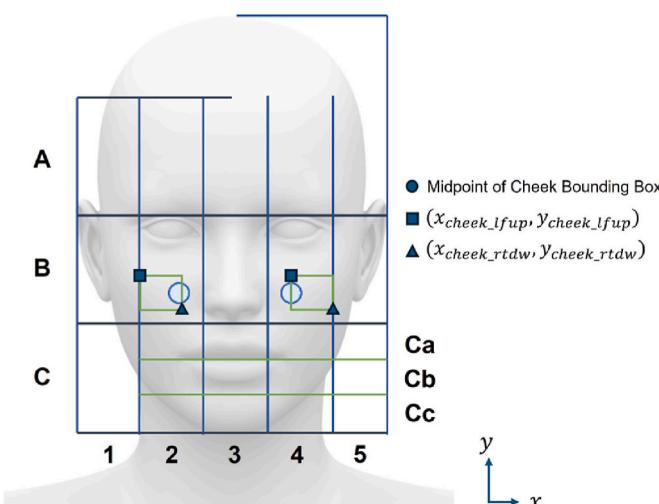


Fig. 5. Horizontal and vertical facial proportions of human face.

cheek area selection strategy, effectively eliminating non-skin temperature points. Particularly in scenarios where occupants have larger facial movements, this approach maintains the stability and reliability of temperature data.

In dynamic environments, thermal imaging sensor readings exhibit significant fluctuations and contain some noise. Therefore, Fourier transform is introduced to filter the time-series data. Firstly, the temperature data are transformed into frequency-based signals using the Fast Fourier Transform (FFT). Secondly, high-frequency noise is eliminated in the frequency domain. Since human skin temperature usually shows a relatively gentle and gradual change pattern, unlikely to produce large fluctuations, a lower cut-off frequency of 0.01 is selected to remove signals above this frequency. Finally, the inverse Fourier transform is applied to obtain the filtered data.

To verify the effectiveness of the filtering algorithm, a 160-s video was recorded indoors. During this period, the subject was instructed to slowly rotate their head and perform actions of putting on and taking off glasses. Throughout this process, the subject's facial temperature was assumed to be a constant 33 °C. Using the MSE introduced in [Equation \(1\)](#) as the evaluation standard, the errors between the observed values and true values were quantified. The performances of three data processing strategies (no processing, single cleaning, and cleaning combined with filtering) in terms of temperature data accuracy were compared, thereby validating the effectiveness of the algorithm.

### 3. Results

#### 3.1. Thermal sensation model

The dataset obtained from subject experiments, encompassing air temperature, facial skin temperature, solar radiation, and thermal sensation voting, is presented in a kernel density plot ([Fig. 6](#)). As the thermal sensation and environmental data were collected from a dynamic cooling process, ranging from a high-temperature cabin with the air conditioning off to 40 min after air conditioning activation, the range of air temperature in the dataset varied from 24.2 to 57.8 °C. This wide temperature range covers various cabin temperature levels that occupants might experience during summer. The thermal sensation votes primarily clustered around neutral and warm, with cold votes mostly originating from nighttime conditions and being less frequent. Votes on the warm side were relatively evenly distributed across different levels. The distribution of thermal sensation voting represents the various levels of thermal sensation that cabin occupants might experience in summer, without significant imbalance. Solar radiation intensity ranged from 0 to 487 W/m<sup>2</sup>, with a relatively wide range for daytime samples. The facial skin temperature data of occupants exhibited a normal distribution. The mean skin temperatures of the window-side cheek and non-window-side cheek were 35.7 °C and 35.5 °C, respectively, while the forehead was slightly higher at 35.9 °C, and the nose was the highest at 36.1 °C, possibly due to its higher exposure to direct sunlight. The skin temperature around the mouth was similar to that of the window-side cheek, with a mean of 35.7 °C, but it had the smallest variation range, with a standard deviation of 1.0 °C. Overall, the dataset had a wide and relatively balanced sample distribution, adequately covering a variety of thermal environmental conditions in summer cabins and the characteristics of thermal sensation and facial skin temperatures experienced by occupants.

Furthermore, [Fig. 6](#) illustrates the linear relationship between temperatures in different facial regions and thermal sensation votes. The slope of the regression line represents the sensitivity of temperature in that region to thermal sensation. Overall, the sensitivities of various facial regions to thermal sensation are similar. The forehead region shows the lowest sensitivity to thermal sensation, with a sensitivity of 0.48 TSV unit/°C. The region around the mouth demonstrates the highest sensitivity, with a sensitivity of 0.61 TSV unit/°C. The skin temperatures of the cheek and nose show comparable sensitivities to

thermal sensation, ranging between 0.54 and 0.56 TSV unit/°C.

The Spearman correlation coefficients between facial skin temperatures, cabin air temperatures, solar radiation intensity, and thermal sensation were calculated, as shown in [Fig. 7](#). Except for the forehead skin temperature, strong correlations ( $r > 0.7$ ) were observed among temperatures in different facial regions. Notably, there are strong correlations between cabin air temperature, cheek skin temperature, nose skin temperature, and thermal sensation ( $r > 0.7$ ). The total solar radiation inside the vehicle cabin shows a moderate correlation with thermal sensation ( $r > 0.7$ ).

It is important to note that having many highly correlated input variables does not necessarily enhance the information content of a model. On the contrary, in practical applications, increasing the number of skin temperature measurement regions can make the assessment more susceptible to noise interference [49]. To establish a thermal sensation assessment model for cabin occupants based on low-cost infrared thermal imaging sensor, while maintaining a balance between model stability and accuracy, this study progressively introduced independent variables based on their correlation strengths, systematically observing the impact of various variable combinations on the model. The performance of the models on the test set with different combinations of features is shown in [Table 5](#).

The performance of models obtained using Elastic Net Regression and Random Forest Regression algorithms is approximately similar. When assessing thermal sensation using only facial skin temperature data obtained from the thermal imaging sensor or relying solely on air temperature, the performance of the model is not ideal. By integrating the thermal imaging sensor, solar radiation intensity sensor, and air temperature sensor, and using cheek temperature, solar radiation intensity in the cabin, and cabin air temperature as inputs, the Random Forest Regression model achieved an R2 of 0.632 and an MSE of 0.519 on the test set. The Elastic Net Regression algorithm achieved an R2 of 0.617 and an MSE of 0.539 on the test set. When expanding the facial skin temperature collection region, the incremental improvement in model fitting is minimal. In practical applications, the forehead may be obscured by hair, leading to ineffective extraction of forehead skin temperature or significantly inaccurate readings, whereas the cheek is usually less covered. Therefore, using a thermal imaging sensor to capture cheek skin temperature and incorporating it as a physiological parameter in the model not only ensures the stability of model outputs in practical applications but also maintains a certain level of model accuracy. The trained model is used for the system's thermal sensation calculations.

#### 3.2. Performance of facial temperature extraction model

##### 3.2.1. Facial recognition and key point estimation

In the training process shown in [Fig. 8](#), the initial rapid decline of pose loss indicates that the model swiftly mastered the key features of infrared images in the early stages. After undergoing 250 training epochs, the loss curve began to stabilize, signifying a gradual decrease in the acquisition of new knowledge as the model's learning process progressed. To prevent overfitting and maintain the model's generalization ability on unseen data, training was terminated at this stage. Finally, the parameters obtained at the 263rd training epoch were selected as the final model parameters, ensuring good performance on new data. The performance of the trained face recognition model across various test sets is depicted in [Fig. 9](#). With a confidence threshold set at 0.5, the model achieved an accuracy of 96.48 % and a recall rate of 98.95 % on the entire test set. The training stride and performance of the model are similar to those reported in other literature for models performing similar recognition tasks [44,50,51].

[Fig. 10](#) indicate that the model performs best on the building indoor test set, with a mean NME of 0.025. In contrast, the mean NMEs on the outdoor and on-vehicle test sets increased to 0.055 and 0.054, respectively, indicating a decline in model performance. This reduction in

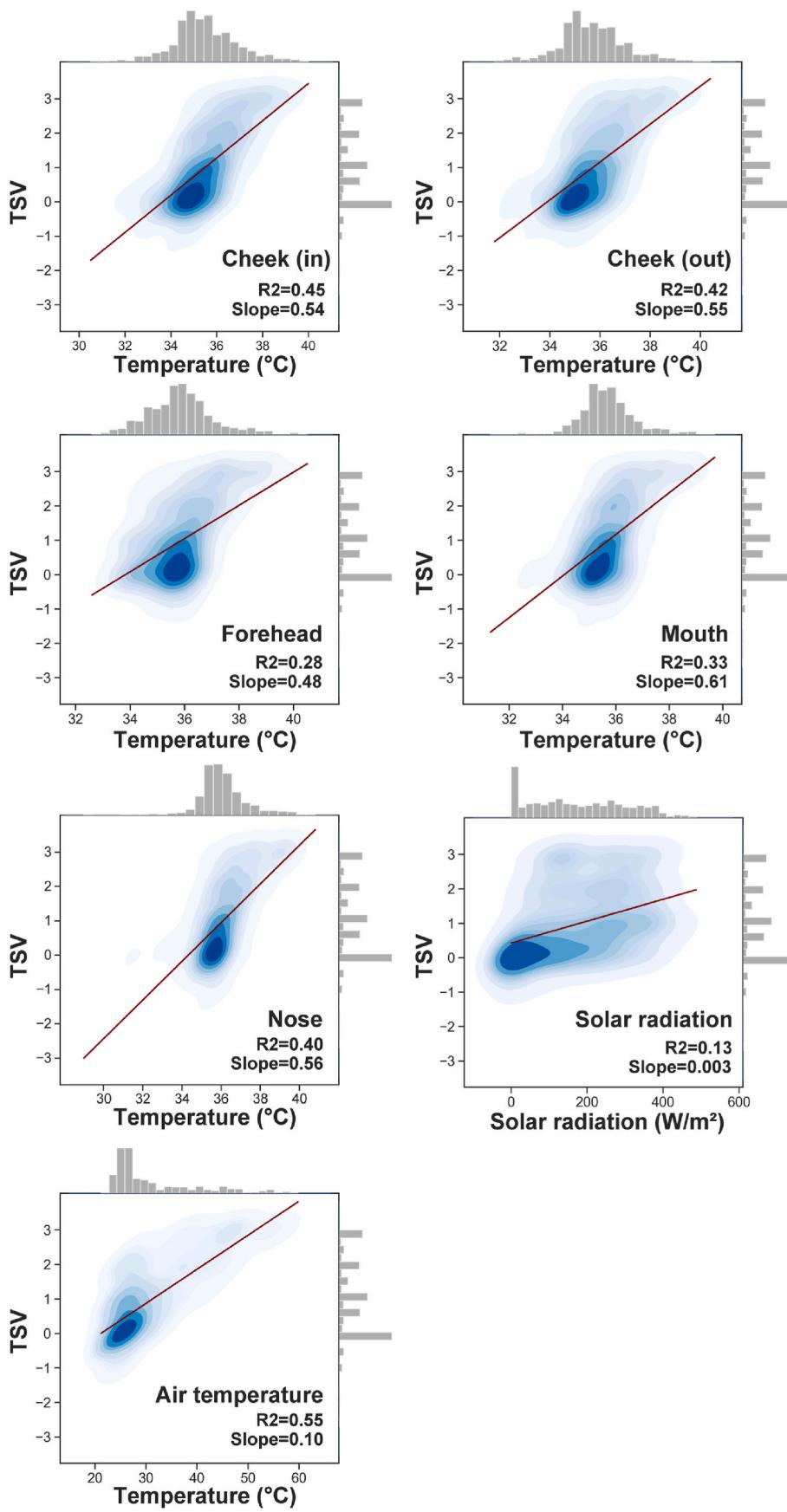
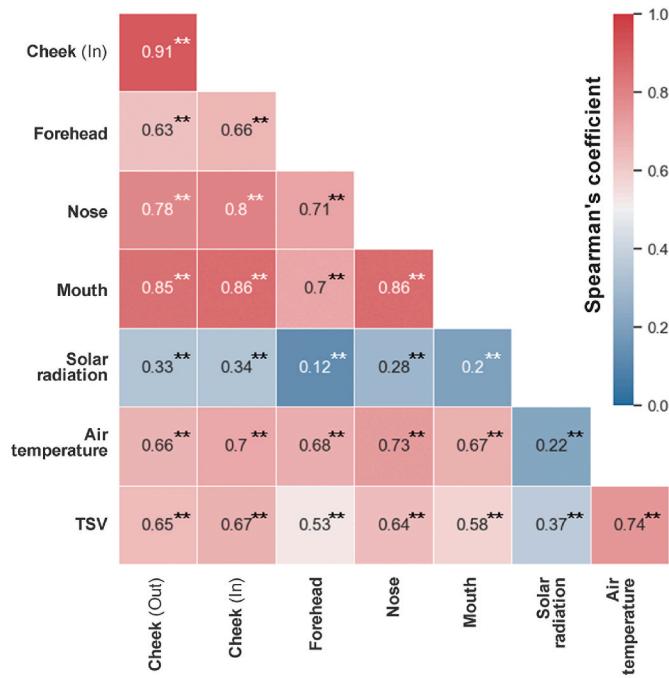


Fig. 6. Sample distribution of the dataset obtained from the experiment.



**Fig. 7.** Spearman correlation coefficients between facial regional skin temperatures, cabin air temperature, solar total radiation, and thermal sensation.

**Table 5**  
Performance of models on the test set with different combinations of features.

Features	ENR		RFR	
	R <sup>2</sup>	MSE	R <sup>2</sup>	MSE
$T_{incheek}$	0.515	0.684	0.499	0.705
$T_{air}$	0.505	0.697	0.457	0.765
$T_{incheek}, T_{air}$	0.584	0.586	0.586	0.584
$T_{incheek}, T_{air}, S_r$	0.617	0.539	0.632	0.519
$T_{incheek}, T_{forehead}, T_{air}, S_r$	0.615	0.546	0.618	0.539
$T_{incheek}, T_{forehead}, T_{nose}, T_{air}, S_r$	0.615	0.543	0.617	0.540
$T_{incheek}, T_{forehead}, T_{nose}, T_{mouth}, T_{air}, S_r$	0.621	0.534	0.651	0.492

performance is primarily attributed to the generally smaller facial dimensions in the outdoor image set and the complexity and challenges of the on-vehicle test environment. In the on-vehicle context, on one hand, the image resolution from low-cost sensors is lower, leading to blurrier

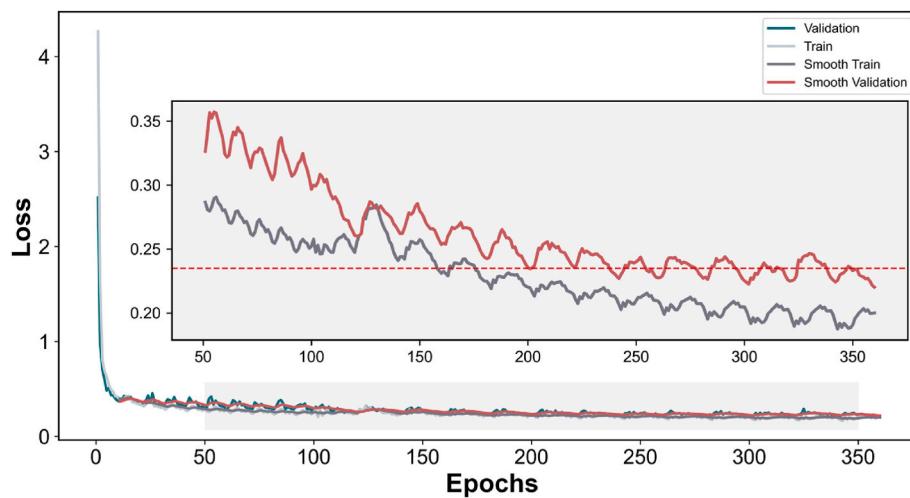
images. On the other hand, the vehicle indoor environment is subject to varying and complex conditions due to weather, solar radiation, and other factors, all of which significantly increase the difficulty of the key point estimation task. Despite these challenges, particularly in the on-vehicle environment, the model still demonstrates good applicability, as specifically illustrated in Fig. 11.

### 3.2.2. Cheek region segmentation

For evaluating facial region segmentation performance, 142 indoor images (containing 142 faces) and 10 in-vehicle images (containing 23 faces) were selected from the test set. The building indoor image set includes thermal imaging along with corresponding visible light imaging. Therefore, the Mediapipe tool [29] was utilized to segment the visible light images to obtain the cheek region, which was then compared to the cheek area identified by the developed model. Due to the absence of visible light images in the on-vehicle image set, the cheek regions in these images were determined through manual annotation. The building indoor facial images cover three viewing angles: looking down, straight ahead, and looking up, while the majority of the on-vehicle images are from a downward viewing angle (see Fig. 12). The cheek segmentation achieved a mean CRRA of 82.11 % in these images (see Fig. 12(c)). The results demonstrate that the model is more adept at handling images from a downward viewing angle, performing best on building indoor downward-looking images and on-vehicle images, with mean CRRA of 88.92 % and 91.85 %, respectively. These results indicate that the developed model is capable of effectively segmenting the cheek region.

### 3.2.3. Performance of temperature data filtering

Since the frames are rectangular, in some cases, the cheek recognition boxes provided by the model may contain portions of non-cheek areas. These invalid data segments are further processed through a filtering algorithm. Fig. 13 illustrates the changes in cheek skin temperature data under different filtering strategies. By eliminating non-skin temperature data, the MSE was reduced from 0.91 to 0.33, effectively excluding significant outliers. On this basis, a filtering algorithm was further introduced to mitigate the variability in sensor readings, reducing the MSE to 0.17. It can be observed that the temperature data before filtering exhibited considerable fluctuations, while the post-filtering data retaining the original trend, was comparatively smoother and more stable. These results demonstrate that a filtering algorithm combined with a cleansing strategy can effectively reduce data disturbances caused by background temperature values and noise.



**Fig. 8.** Convergence of pose loss during training epochs for YOLO model.

		Building indoor test set		Outdoor test set		On-vehicle test set	
		Face	Background	Face	Background	Face	Background
True	Face	300	1	559	14	839	1
	Background	0	0	47	0	13	37
	Predicted			Predicted		Predicted	

Fig. 9. Confusion matrices across different test sets.

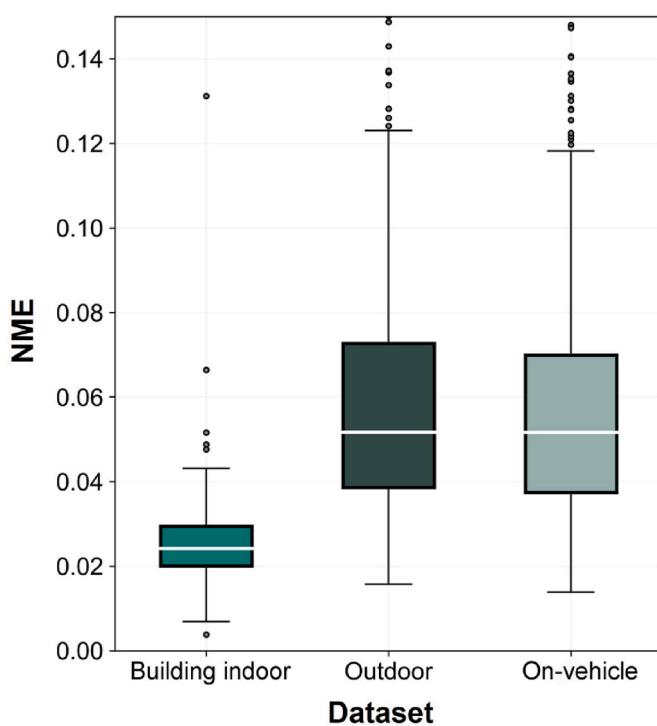


Fig. 10. NME of key point estimation on different datasets.

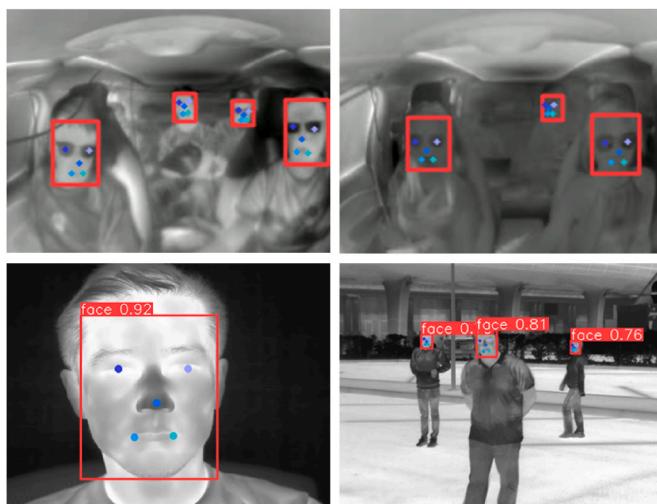


Fig. 11. Output of face recognition and key point estimation model.

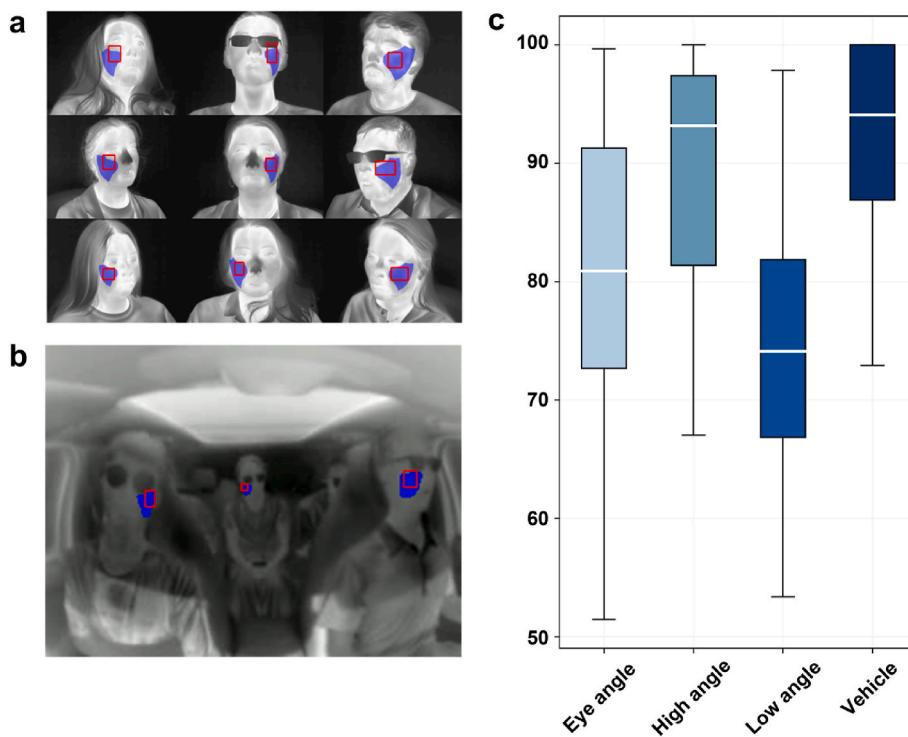
#### 4. Validation of thermal sensation assessment system

To further validate the performance of the developed low-cost cabin thermal sensation assessment system, an independent validation experiment was conducted. The validation experiment took place in October 2023 at the outdoor space of Shanghai Jiao Tong University, inside a car cabin. A total of 18 subjects were recruited to participate in the experiment to validate the system's performance. During the experiment, all subjects wore standard summer clothing. They entered the cabin from outside, undergoing a process of cooling and temperature stabilization, with thermal sensation assessments conducted every 2 min. Subjects also voted for true overall thermal sensation every 2 min. A low-cost thermal imaging sensor, installed under the rearview mirror, monitored facial temperatures in real-time, while the vehicle's built-in sensors recorded solar radiation and cabin air temperature. All data were transmitted in real-time to an Ubuntu computer for real-time analysis of thermal sensation. The experiment did not restrict car orientation and initial temperature, to test the system's performance under real conditions. The air conditioning vents were kept in a horizontal direction (air speed is  $0.61 \pm 0.14$  m/s in the breathing zone of occupants). Each experimental session included three subjects, with subject number 5 withdrawing midway due to personal reasons. Detailed information of the 18 subjects is available in Appendix Table S1. The air temperature changes experienced by each subject are shown in Appendix Figure S4.

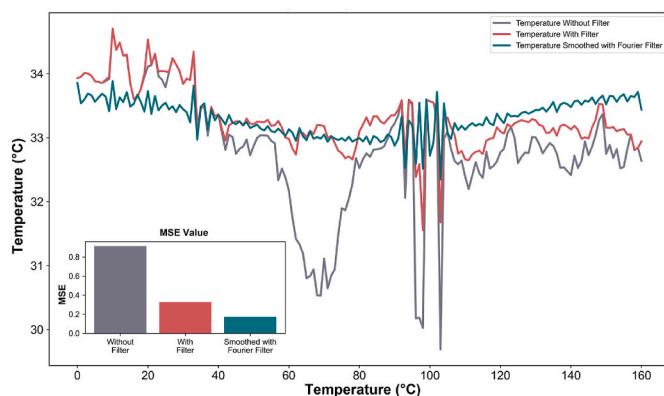
During the validation experiment, the system's user interface, as operational, is depicted in Fig. 14. A comparison of the system's output results, using both the embedded Elastic Net regression model and the Random Forest regression model, with actual thermal sensation votes is shown in Fig. 15. The results demonstrate that the system's output is consistent with the actual thermal sensation votes. The system output based on the Elastic Net regression model shows a mean absolute error (MAE) of 0.49 thermal sensation units. It is noteworthy that even though the model trained using Random Forest regression demonstrated slightly higher performance on the test set compared to the Elastic Net regression during model training, the former exhibited a higher output error (MAE = 0.52) in the validation experiment, with fluctuations observed in certain subjects (such as subjects 3, 11, and 16). This is attributed to the more stable and accurate readings from the thermal imaging devices used in the modeling experiment, whereas the actual application system utilized low-cost sensors with less stable readings. For the Random Forest regression model, it typically does not extrapolate beyond the range of training data. Coupled with the greater temperature variability of low-cost sensors, this leads to more fluctuations in the model output. On the other hand, the Elastic Net regression model, being based on linear regression, produces more robust outputs with less fluctuation.

#### 5. Discussion

This study has developed a low-cost, non-contact thermal sensation

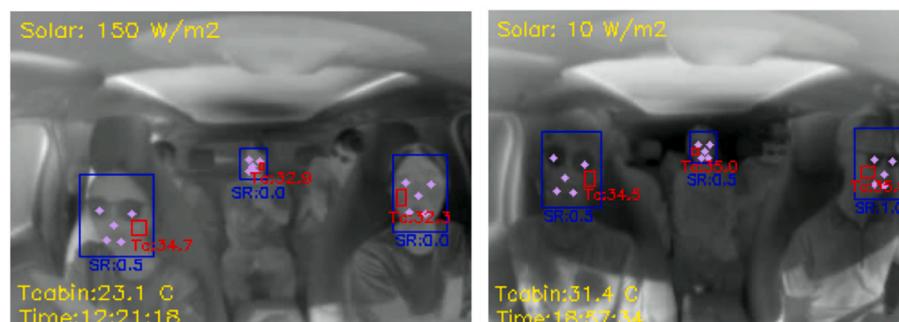


**Fig. 12.** Performance of the region segmentation (the areas selected by the model are indicated by red boxes). (a) Performance on the building indoor test set, with the actual cheek region delineated by Mediapipe marked in blue. (b) Performance on the on-vehicle test set, with the actual cheek area manually demarcated shown in blue. (c) CRRA under different views. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

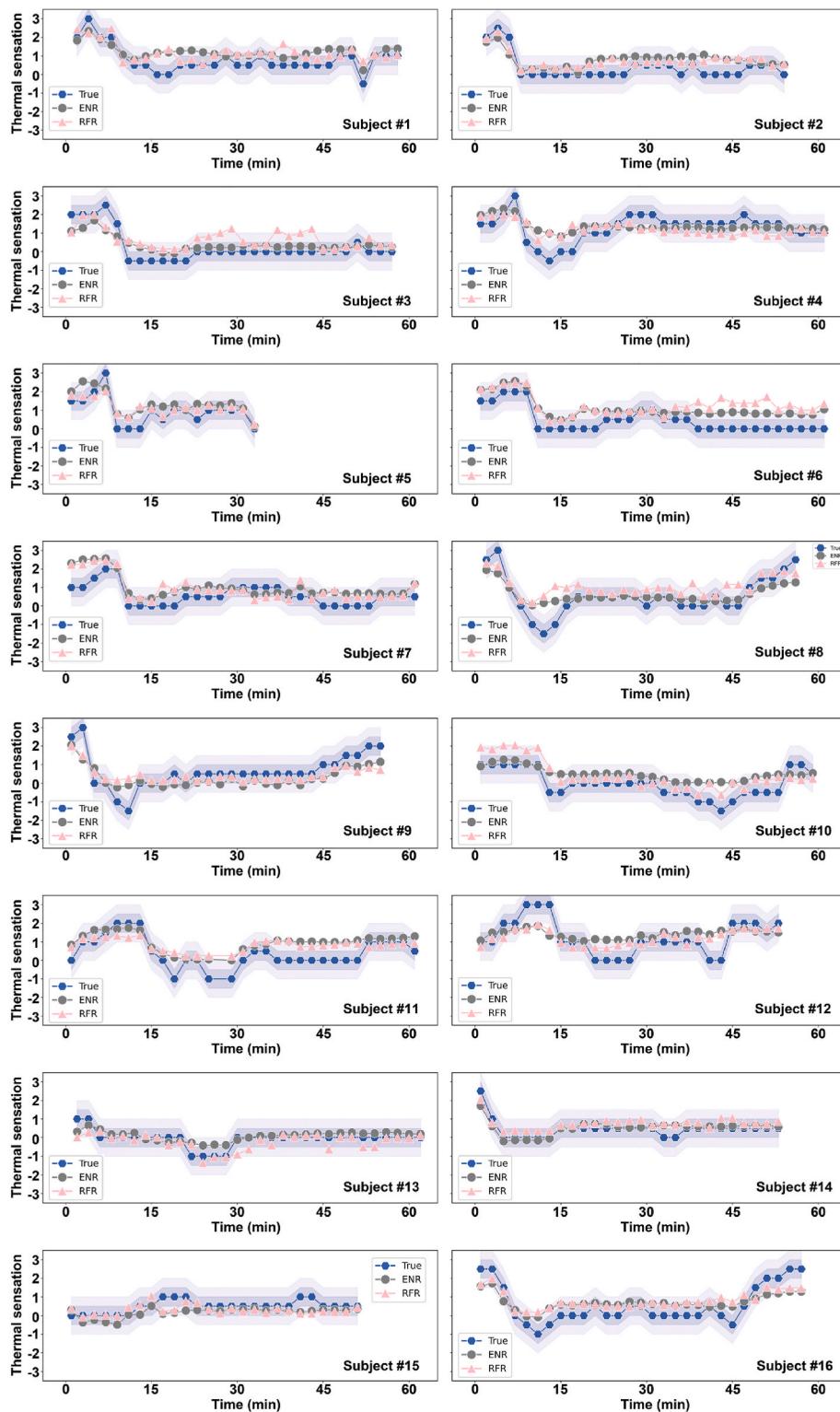


**Fig. 13.** Comparison of cheek skin temperature data before and after application of the temperature filtering.

assessment system for cabin occupants. The system has been validated and demonstrates high accuracy. This system supports the air conditioning system in performing feedback automatic control. This integration can be done through a vehicle's central processing unit or a dedicated microcontroller. A simple and feasible strategy is to preset an initial air conditioning setpoint. Once an occupant enters the vehicle, the goal is to achieve a thermal sensation output value of +0.5 through periodic iterative adjustments to the setpoint based on the developed thermal sensation assessment system. At each calculation point, the thermal sensation of the occupant is calculated, and depending on the difference between the current thermal sensation calculation and +0.5, the air conditioning setpoint is either lowered or raised. For example, for each unit difference in thermal sensation value, the adjustment to the temperature setpoint is 2 °C. For situations involving multiple people, the set points are adjusted by zone. For different automobile manufacturers, there is a specific corresponding relationship between the air conditioning set points and the supply air temperature, allowing the adjustment of setpoints to be translated into adjustments of supply air temperature. Furthermore, based on such fundamental control strategy, optimization algorithm can be introduced to explore the optimal control



**Fig. 14.** Thermal imaging-based cabin occupant thermal sensation assessment system operation interface.



**Fig. 15.** Validation results of the low-cost cabin occupant thermal sensation assessment system output.

strategies that quickly improve the thermal sensation of cabin occupants while balancing energy consumption [52]. With the gradual reduction of physical buttons in electric vehicles, this automatic control can effectively reduce user intervention in air conditioning adjustments during driving, thereby enhancing driving safety. Another potential application is that the developed system can capture the different thermal sensations of occupants in the vehicle, potentially meeting all occupants' personalized thermal needs through local thermal comfort management such

as seat heating or cooling. These types of personal thermal comfort systems are considered to have great potential to replace traditional air conditioning systems, balancing comfort with energy savings.

Table 6 summarizes the models and their performance related to the evaluation of thermal sensation of occupants inside the cabin, as presented in relevant studies. Currently, there are relatively few studies focused on the development of models for assessing the thermal sensation of occupants within cabin environments. Moreover, research that

**Table 6**

Comparison of existing cabin occupant thermal sensation models and system performance.

Model	Air temperature range (°C)	Input <sup>a</sup>	Performance		Actual evaluation system established?
			Self-reported	Peer validation in cabin environment <sup>b</sup>	
DTS [18]	10–48	$\Delta T_{sk,m}$ , $\frac{dT^{(-)}_{sk,m}}{dt}$ , $\frac{dT^{(+)}_{sk,m}}{d_{tmax}}$ , $\Delta T_{hy}$	No specific metric values are given (comparison result figures can be found in Ref. [18])	RMSE = 1.25	No
Lai's model [19]	0–35 (outdoor)	$T_L$ , $T_{sk,m}$ , $\frac{dT_{sk,m}}{dt}$	Model fitting: $R^2 = 0.811$ ; Model validation: $R^2 = 0.766\text{--}0.938$ .	RMSE = 1.2	No
UCB model [53–55]	20–32	$\Delta T_{sk,m}$ , $\Delta T_{sk,i}$ , $\frac{dT_{sk,i}}{dt}$ , $\frac{dT_{cr}}{dt}$	Model fitting: $R^2 = 0.94$ (more performance details can be found in Ref. [55])	RMSE = 1.19	No
Zhou's model [11]	20–50	$T_{Lg}$ (face), $T_{Ls}$ (face), $T_{sk,m}$ , $T_{out}$	Model fitting: $R^2 = 0.91$ ; Model validation: RMSE = 0.56	–	No
Li's model [17]	16–32	$T_{head}$ , $T_{trunk}$ , $T_{back}$ , $T_{arm}$ , $T_{hand}$ , $T_{leg}$ , $T_{foot}$	Model fitting: $R^2 = 0.91$	–	No
He's model [1]	25–40	$T_{out}$ , $T_{air\_return}$ , RH, SR, $T_{air\_supply}$ , $V_{air}$	Model validation: MAE = 0.764–0.863	–	No
Zhang's model [56]	Not given	$T_{air}$	Model accuracy >91.7 %	–	Yes

<sup>a</sup>  $\Delta T_{sk,m}$  = difference in mean skin temperature under neutral thermal sensation and current status,  $dT^{(-)}_{sk,m}/dt$  = negative change rate of the mean skin temperature,  $dT^{(+)}_{sk,m}/d_{tmax}$  = maximum positive change rate of skin temperature,  $\Delta T_{hy}$  = difference in head core temperature under neutral thermal sensation and current status,  $T_{sk,m}$  = mean skin temperature,  $dT_{sk,m}/dt$  = change rate of the mean skin temperature,  $\Delta T_{sk,i}$  = Local skin temperature,  $dT_{sk,i}/dt$  = change rate of the local skin temperature,  $dT_{cr}/dt$  = change rate of the core temperature,  $T_{Lg}$  (face) = thermal load of face changes gradually,  $T_{Ls}$  (face) = thermal load of face changes suddenly,  $T_{out}$  = outdoor air temperature,  $T_{head}$  = head skin temperature,  $T_{trunk}$  = trunk skin temperature,  $T_{back}$  = back skin temperature,  $T_{arm}$  = arm skin temperature,  $T_{hand}$  = hand skin temperature,  $T_{leg}$  = leg skin temperature,  $T_{foot}$  = foot skin temperature,  $T_{air\_return}$  = air return temperature, RH = relative humidity, SR = solar radiation,  $T_{air\_supply}$  = air supply temperature,  $V_{air}$  = air speed,  $T_{air}$  = cabin air temperature,  $V_{air}$  = air speed.

<sup>b</sup> Details of the validation comparison can be found in the study by Zhou et al. [11].

provides practical and applicable assessment systems from an application perspective is even less common. Zhou et al. [11] used data sets obtained from cabin experiments to validate the performance of the DTS model, the UCB model, and Lai's model in their research. They found that the UCB model, which was developed based on cabin experiments, performed similarly to the other two models in assessing the thermal sensation of cabin occupants. This demonstrates that the DTS model and Lai's model are feasible for assessing thermal sensation in cabin environments, achieving at least comparable validation performance to the UCB model. However, from the practical standpoint of real-time assessment of thermal sensation to guide the automatic adjustment of air conditioning systems, these models are difficult to apply. Similarly, while Li's model [17] has achieved good fit, the challenge of effectively obtaining skin temperatures from seven body parts without disturbing the individual in real-world scenarios is inevitable. This is complicated by the fact that at least four of these temperature measurement locations are typically covered by clothing. The model proposed by He et al. [1] fundamentally relies on the PMV model for calculations. While the PMV model is suitable for neutral, steady-state environments, its accuracy may be reduced in non-steady-state cabin environments, leading to a relatively lower accuracy of He's model. The system developed in our study enables non-contact assessment of thermal sensation. Validation experiments have shown that the system's output performance is comparable to Zhou's model but offers greater practical application performance. However, it must be noted that there is currently a lack of similar datasets available for peer validation of this study, and relying solely on the dataset from our own study for validation could lead to an overestimation of the system's generalizability.

The two experiments conducted in this study used the same sport utility vehicle (SUV) (model: NIO ES6). For different types of vehicles, variations in the insulating performance of glass and enclosure structures may lead to significant differences in the interior surface temperatures within the same outdoor environmental conditions. Since interior surface temperatures are difficult to measure in practical applications, this study did not include them as inputs for the system. However, this omission may result in the system losing some adaptability when assessing the thermal sensation of occupants in different automotive

cabins. Furthermore, the black globe temperature is considered an effective parameter that reflects the radiant heat exchange between the environment and the human. In the preliminary experiments, a black globe thermometer was used to measure the radiant temperature to capture the effects of radiant heat from different directions. However, tests showed that its stabilization time is relatively long, consistent with the 5–30 min noted in existing research [57]. This results in the black globe temperature potentially being unable to immediately and accurately reflect rapid temperature changes caused by the activation or deactivation of air conditioner. Additionally, there are practical limitations associated with the use of black globe temperature sensors. Therefore, the system did not include black globe temperature as an input parameter. Furthermore, air speed is a crucial parameter for thermal comfort of occupants. Due to the confined space inside vehicle cabins and the close proximity of occupants to the air vents, combined with the low supply air temperature, there is a significant risk of draft sensation when the air vents directly face the occupants during cooling operations. Therefore, in the experiments to collect data for this study, the air vents were fixed at an upward angle of 75° to reduce the air speed in the occupant zone to below 0.3 m/s. This consideration aligns with the optimal settings for cooling the cabin air conditioning in summer, utilizing the natural descent of cold air to lower the cabin temperature while avoiding direct drafts on occupants, which can cause discomfort. Consequently, air speed was not included as an input parameter in the thermal sensation assessment system. However, it must be noted that this approach potentially undermines the generalizability of the model. First, occupants may use the ventilation mode of the air conditioning to improve thermal sensation during transitional seasons. Existing research suggests that when the supply air temperature is higher, even if the air speed is somewhat elevated, the risk of a draft sensation may still be within acceptable limits [58,59]. However, in this usage scenario, the system may have inaccuracies in assessing the thermal sensation of occupants. On the other hand, due to the design of air conditioner vents in some vehicles, it is difficult to prevent cold airflow from blowing directly on occupants, which may result in underestimations by the system. In the system performance validation experiments, to test the accuracy of the system in assessing thermal sensation under conditions of airflow,

the vents were set to deliver horizontal airflow, with the supply air speed set to "medium". The validation results showed that at certain times, the system output was higher than the actual votes, indicating that the cooling effect of the air speed might not have been effectively captured.

Privacy concerns are always a focal point in the field of visual technology, particularly in the realm of visible light imaging. However, thermal imaging sensors collect temperature data rather than visual images. This means that the technology does not capture details that can directly identify occupants, such as detailed facial features, skin color, or clothing, nor does it have the capability to precisely identify items within the cabin. This significantly reduces the potential for privacy invasion of cabin occupants. For this system, on one hand, the characteristic of offline local operation reduces the possibility of temperature data leakage. On the other hand, the system only converts the skin temperature of specific regions into data formats, without involving the storage of video streams. These characteristics collectively help address cabin occupants' concerns about privacy.

## 6. Limitation

This study developed system based on the conditions of summer cooling; therefore, its applicability under winter heating conditions necessitates further examined. Additionally, this study did not include air speed and air supply direction as input variables, and all subjects were aged 20–30 years. Future experiments could involve a broader age range of subjects and incorporate variables such as air speed and direction to enhance the generalizability of the thermal sensation assessment system. Lastly, regarding the facial temperature extraction model, the angle of imaging might affect the model's recognition capabilities. For non-frontal faces, it might be challenging to identify facial features, leading to difficulties in recognizing cheek areas. While the head position of drivers tends to be relatively fixed, occupants in a vehicle may turn their heads or look down, which could impact the effectiveness of cheek temperature extraction.

## 7. Conclusion

This research utilizes low-cost thermal imaging sensors, in-vehicle air temperature sensors, solar radiation sensors, and an onboard computer to construct a system for assessing the thermal sensation of cabin occupants. This system is capable of real-time face recognition, cheek skin temperature extraction, thermal sensation calculation, and real-time display of results. After development, the system was deployed in a vehicle cabin and tested with subjects to validate its performance further. The main conclusions drawn are as follows.

1. A thermal sensation assessment system was established utilizing a low-cost thermal imaging sensor. This system was deployed in vehicles, and its performance was validated through independent validation experiments. The output demonstrated an MAE of 0.5 thermal sensation units, indicating that the system performs effectively overall.
2. A thermal imaging facial cheek temperature extraction algorithm, utilizing YOLO v8 and facial structure characteristics, was developed. The model developed based on this algorithm demonstrated high efficiency in thermal image-based facial recognition, achieving a mean accuracy of 96.5 % and recall of 99.0 %. In the task of cheek segmentation, the model achieved a mean CRRA of 88.9 %. A filtering algorithm further minimized noise-induced data disturbances, yielding a cheek temperature extraction MSE of 0.17.
3. Considering both stability and accuracy of the thermal sensation assessment model, the Elastic Net regression trained using cheek skin temperature, solar radiation intensity, and cabin air temperature, was found to be the best. It achieved an  $R^2$  of 0.62 and an MSE of 0.54 on the test set.

## CRediT authorship contribution statement

**Zhenyu Hou:** Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Junmeng Lyu:** Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Dongyuan Wu:** Writing – original draft, Methodology, Formal analysis. **Jiangping Chen:** Writing – review & editing, Supervision, Funding acquisition. **Junye Shi:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Zhiwei Lian:** Writing – review & editing, Supervision.

## Declaration of competing interest

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

## Data availability

Data will be made available on request.

## Acknowledgements

This work was supported by National Key R&D Program of China (2020YFA0711500), National Natural Science Foundation of China (52306104). This study extends its gratitude to the NIO Co., Ltd. for their support.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.buildenv.2024.111692>.

## References

- [1] X. He, X. Zhang, R. Zhang, J. Liu, X. Huang, J. Pei, J. Cai, F. Guo, Y. Wang, More intelligent and efficient thermal environment management: a hybrid model for occupant-centric thermal comfort monitoring in vehicle cabins, *Build. Environ.* 228 (2023) 109866, <https://doi.org/10.1016/j.buildenv.2022.109866>.
- [2] ISO/TS 14505-3, Ergonomics of the Thermal Environment — Evaluation of Thermal Environments in Vehicles - Part 3: Evaluation of Thermal Comfort Using Human Subjects, 2006.
- [3] ISO/TS 14505-1, Ergonomics of the Thermal Environment—Evaluation of Thermal Environments in Vehicles—Part 1: Principles and Methods for Assessment of Thermal Stress., (n.d.).
- [4] H. Du, Z. Lian, D. Lai, L. Duanmu, Y. Zhai, B. Cao, Y. Zhang, X. Zhou, Z. Wang, X. Zhang, Z. Hou, Evaluation of the accuracy of PMV and its several revised models using the Chinese thermal comfort Database, *Energy Build.* 271 (2022) 112334, <https://doi.org/10.1016/j.enbuild.2022.112334>.
- [5] M.A. Humphreys, J. Fergus Nicol, The validity of ISO-PMV for predicting comfort votes in every-day thermal environments, *Energy Build.* 34 (2002) 667–684, [https://doi.org/10.1016/S0378-7788\(02\)00018-X](https://doi.org/10.1016/S0378-7788(02)00018-X).
- [6] K. Humphreys, A. Winzelberg, E. Klaw, Psychologists' ethical responsibilities in Internet-based groups: issues, strategies, and a call for dialogue, *Prof. Psychol. Res. Pract.* 31 (2000) 493–496, <https://doi.org/10.1037/0735-7028.31.5.493>.
- [7] Y. Don Guan, M.H. Hosni, B.W. Jones, T.P. Giedla, Investigation of human thermal comfort under highly transient conditions for automotive applications-Part 1: experimental design and human subject testing implementation, *Build. Eng.* 109 (2003) 885–897. <https://www.proquest.com/scholarly-journals/investigation-human-thermal-comfort-under-highly/docview/192539219/se-2?accountid=13818>.
- [8] Y. Don Guan, M.H. Hosni, B.W. Jones, T.P. Giedla, Investigation of human thermal comfort under highly transient conditions for automotive applications-Part 2: thermal sensation modeling, *Build. Eng.* 109 (2003) 898–907. <https://www.proquest.com/scholarly-journals/investigation-human-thermal-comfort-under-highly/docview/192538953/se-2?accountid=13818>.
- [9] H. Zhang, *Human Thermal Sensation and Comfort in Transient and Non-uniform Thermal Environments*, University of California, Berkeley, 2003.
- [10] X. Zhou, D. Lai, Q. Chen, Experimental investigation of thermal comfort in a passenger car under driving conditions, *Build. Environ.* 149 (2019) 109–119, <https://doi.org/10.1016/j.buildenv.2018.12.022>.
- [11] X. Zhou, D. Lai, Q. Chen, Thermal sensation model for driver in a passenger car with changing solar radiation, *Build. Environ.* 183 (2020) 107219, <https://doi.org/10.1016/j.buildenv.2020.107219>.

- [12] Y. Kim, M. Lee, Y. Shin, H. Cho, Investigation of changes in Driver's biosignals and thermal comfort according to the heating method in winter, *Case Stud. Therm. Eng.* 42 (2023) 102749, <https://doi.org/10.1016/j.csite.2023.102749>.
- [13] J.D. Cotter, N.A.S. Taylor, The distribution of cutaneous sudomotor and alliesthesia thermosensitivity in mildly heat-stressed humans: an open-loop approach, *J Physiol* 565 (2005) 335–345, <https://doi.org/10.1113/jphysiol.2004.081562>.
- [14] M. Cabanac, B. Massonnet, R. Belaiche, Preferred skin temperature as a function of internal and mean skin temperature, *J. Appl. Physiol.* 33 (1972) 699–703, <https://doi.org/10.1152/jappl.1972.33.6.699>.
- [15] J.-H. Choi, V. Loftness, Investigation of human body skin temperatures as a bio-signal to indicate overall thermal sensations, *Build. Environ.* 58 (2012) 258–269, <https://doi.org/10.1016/j.buildenv.2012.07.003>.
- [16] 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, <https://doi.org/10.1016/j.enbuild.2017.09.032>.
- [17] W. Li, J. Chen, F. Lan, H. Xie, Human thermal sensation and its algorithmic modelization under dynamic environmental thermal characteristics of vehicle cabin, *Indoor Air* 32 (2022) e13168, <https://doi.org/10.1111/ina.13168>.
- [18] D. Fiala, K.J. Lomas, M. Stohrer, First principles modeling of thermal sensation responses in steady-state and transient conditions, *Build. Eng.* 109 (2003) 179. <https://www.proquest.com/scholarly-journals/first-principles-modeling-thermal-sensation/docview/192512186/se-2?accountid=13818>.
- [19] D. Lai, X. Zhou, Q. Chen, Modelling dynamic thermal sensation of human subjects in outdoor environments, *Energy Build.* 149 (2017) 16–25, <https://doi.org/10.1016/j.enbuild.2017.05.028>.
- [20] Y. Wu, B. Cao, Recognition and prediction of individual thermal comfort requirement based on local skin temperature, *J. Build. Eng.* 49 (2022) 104025, <https://doi.org/10.1016/j.jobe.2022.104025>.
- [21] Y. Wu, B. Cao, Y. Zhu, Development of an automatic personal comfort system (APCS) based on real-time thermal sensation prediction, *Build. Environ.* 246 (2023) 110958, <https://doi.org/10.1016/j.buildenv.2023.110958>.
- [22] Y. Wu, J. Zhao, B. Cao, A systematic review of research on personal thermal comfort using infrared technology, *Energy Build.* 301 (2023) 113666, <https://doi.org/10.1016/j.enbuild.2023.113666>.
- [23] J. Lyu, H. Du, Z. Zhao, Y. Shi, B. Wang, Z. Lian, Where should the thermal image sensor of a smart A/C look?–Occupant thermal sensation model based on thermal imaging data, *Build. Environ.* 239 (2023) 110405, <https://doi.org/10.1016/j.buildenv.2023.110405>.
- [24] D. Li, C.C. Menassa, V.R. Kamat, Robust non-intrusive interpretation of occupant thermal comfort in built environments with low-cost networked thermal cameras, *Appl. Energy* 251 (2019) 113336, <https://doi.org/10.1016/j.apenergy.2019.113336>.
- [25] 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, <https://doi.org/10.1016/j.enbuild.2018.07.025>.
- [26] Y. He, H. Zhang, E. Arens, A. Merritt, C. Huizenga, R. Levinson, A. Wang, A. Ghahramani, A. Alvarez-Suarez, Smart detection of indoor occupant thermal state via infrared thermography, computer vision, and machine learning, *Build. Environ.* 228 (2023) 109811, <https://doi.org/10.1016/j.buildenv.2022.109811>.
- [27] 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, <https://doi.org/10.1016/j.enbuild.2020.110261>.
- [28] Y. Wu, B. Cao, M. Hu, G. Lv, J. Meng, H. Zhang, Development of personal comfort model and its use in the control of air conditioner, *Energy Build.* 285 (2023) 112900, <https://doi.org/10.1016/j.enbuild.2023.112900>.
- [29] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Ubweja, M. Hays, F. Zhang, C.-L. Chang, M.G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, M. Grundmann, MediaPipe: a framework for building perception pipelines, *ArXiv abs/1906 (2019) 08172*.
- [30] K. Mallat, J.-L. Dugelay, A benchmark database of visible and thermal paired face images across multiple variations, in: International Conference of the Biometrics Special Interest Group, BIOSIG 2018, GI/IEEE, Darmstadt, Germany, 2018, pp. 199–206.
- [31] J. Ma, Y. Ma, C. Li, Infrared and visible image fusion methods and applications: a survey, *Inf. Fusion* 45 (2019) 153–178, <https://doi.org/10.1016/j.inffus.2018.02.004>.
- [32] ISO, *Ergonomics of the Thermal Environment - Evaluation of Thermal Environments in Vehicles - Part 1: Principles and Methods for Assessment of Thermal Stress*, 14505–1, ISO, 2006.
- [33] L. Lan, J. Tang, P. Wargocki, D.P. Wyon, Z. Lian, Cognitive performance was reduced by higher air temperature even when thermal comfort was maintained over the 24–28°C range, *Indoor Air* 32 (2022) e12916, <https://doi.org/10.1111/ina.12916>.
- [34] R.M.S.F. Almeida, E. Barreira, M.L. Simões, T.S.F. Sousa, Infrared thermography to evaluate thermal comfort under controlled ambient conditions, *Appl. Sci.* 12 (2022), <https://doi.org/10.3390/app122312105>.
- [35] 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 (2022) 619–634, <https://doi.org/10.1007/s10044-021-00961-5>.
- [36] A. Speak, L. Montagnani, C. Wellstein, S. Zerbe, Forehead temperatures as an indicator of outdoor thermal comfort and the influence of tree shade, *Urban Clim.* 39 (2021) 100965, <https://doi.org/10.1016/j.ulclim.2021.100965>.
- [37] ISO, *Ergonomics of the Thermal Environment - Instruments for Measuring Physical Quantities*, ISO, 2001, p. 7726.
- [38] X. Shan, E.-H. Yang, Supervised machine learning of thermal comfort under different indoor temperatures using EEG measurements, *Energy Build.* 225 (2020) 110305, <https://doi.org/10.1016/j.enbuild.2020.110305>.
- [39] Z. Wang, J. Wang, Y. He, Y. Liu, B. Lin, T. Hong, Dimension analysis of subjective thermal comfort metrics based on ASHRAE Global Thermal Comfort Database using machine learning, *J. Build. Eng.* 29 (2020) 101120, <https://doi.org/10.1016/j.jobe.2019.101120>.
- [40] Z. Wu, N. Li, J. Peng, H. Cui, P. Liu, H. Li, X. Li, Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings, *Energy Build.* 173 (2018) 117–127, <https://doi.org/10.1016/j.enbuild.2018.05.031>.
- [41] 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, <https://doi.org/10.1016/j.enbuild.2021.111771>.
- [42] Q. Zhao, Z. Lian, D. Lai, Thermal comfort models and their developments: a review, *Energy and Built Environment* 2 (2021) 21–33, <https://doi.org/10.1016/j.enbenv.2020.05.007>.
- [43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: machine learning in Python, *J. Mach. Learn. Res.* 12 (2011) 2825–2830.
- [44] A. Kuzdeuov, D. Aubakirova, D. Koishigarina, H.A. Varol, TFW: annotated thermal faces in the wild dataset, *IEEE Trans. Inf. Forensics Secur.* 17 (2022) 2084–2094, <https://doi.org/10.1109/TIFS.2022.3177949>.
- [45] G. Jocher, A. Chaurasia, J. Qiu, YOLO by Ultralytics, 2023, Version 8.0.0. <https://github.com/ultralytics/ultralytics>.
- [46] W.F. Larrabee, H.K. Makielski, B.M. Zide, Surgical anatomy of the face, *Plast. Reconstr. Surg.* 94 (1994) (1963) 402, <https://doi.org/10.1097/00006534-199408000-00034>.
- [47] 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, <https://doi.org/10.1016/j.enbenv.2016.09.005>.
- [48] H.A. Braun, J. Bligh, K. Bruck, K. Voigt, G. Heldmaier, *Thermoreception and Temperature Regulation*, Springer, 2012.
- [49] S. Khalid, T. Khalil, S. Nasreen, A survey of feature selection and feature extraction techniques in machine learning, in: 2014 Science and Information Conference, 2014, pp. 372–378, <https://doi.org/10.1109/SAI.2014.6918213>.
- [50] C. Jiang, H. Ren, X. Ye, J. Zhu, H. Zeng, Y. Nan, M. Sun, X. Ren, H. Huo, Object detection from UAV thermal infrared images and videos using YOLO models, *Int. J. Appl. Earth Obs. Geoinf.* 112 (2022), <https://doi.org/10.1016/j.jag.2022.102912>.
- [51] L. Jiang, S. Hesham, K.P. Lim, M. Manoj Krishnadhas, Razi, Z. Zishuo, M. Chenxin, B. Jiang, H. Saeedipour, YOLO based thermal screening using artificial intelligence (AI) for instinctive human facial detection, in: W. Xie, S. Gao, X. He, X. Zhu, J. Huang, W. Chen, L. Ma, H. Shu, W. Cao, L. Jiang, Z. Shu (Eds.), 2022 IEEE 17TH CONFERENCE ON INDUSTRIAL ELECTRONICS AND APPLICATIONS(ICIEA), 2022, pp. 1063–1068, <https://doi.org/10.1109/ICIEA54703.2022.10005908>.
- [52] B. Yang, Y. Liu, P. Liu, F. Wang, X. Cheng, Z. Lv, A novel occupant-centric stratum ventilation system using computer vision: Occupant detection, thermal comfort, air quality, and energy savings, *Build. Environ.* 237 (2023) 110332, <https://doi.org/10.1016/j.enbenv.2023.110332>.
- [53] H. Zhang, E. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments: Part I: local sensation of individual body parts, *Build. Environ.* 45 (2010) 380–388, <https://doi.org/10.1016/j.enbenv.2009.06.018>.
- [54] H. Zhang, E. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments, part III: whole-body sensation and comfort, *Build. Environ.* 45 (2010) 399–410, <https://doi.org/10.1016/j.enbenv.2009.06.020>.
- [55] H. Zhang, E. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments, part III: whole-body sensation and comfort, *Build. Environ.* 45 (2010) 399–410, <https://doi.org/10.1016/j.enbenv.2009.06.020>.
- [56] C. Zhang, B. Hou, M. Li, Z. Chen, R. Zhang, S. Li, Z. Han, X. Li, A novel online prediction method for vehicle cabin temperature and passenger thermal sensation, *Appl. Therm. Eng.* 245 (2024) 122853, <https://doi.org/10.1016/j.applthermaleng.2024.122853>.
- [57] F.R. d'Ambrosio Alfano, M. Dell'isola, G. Ficco, B.I. Palella, G. Riccio, On the measurement of the mean radiant temperature by means of globes: an experimental investigation under black enclosure conditions, *Build. Environ.* 193 (2021) 107655, <https://doi.org/10.1016/j.buildenv.2021.107655>.
- [58] B. Yang, P. Liu, Y. Liu, F. Wang, Performance evaluation of ductless personalized ventilation combined with impinging jet ventilation, *Appl. Therm. Eng.* 222 (2023) 119915, <https://doi.org/10.1016/j.applthermaleng.2022.119915>.
- [59] J. Lyu, Y. Shi, H. Du, Z. Lian, Sex-based thermal comfort zones and energy savings in spaces with joint operation of air conditioner and fan, *Build. Environ.* (2023) 111002, <https://doi.org/10.1016/j.enbenv.2023.111002>.