IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 February 2024, accepted 10 March 2024, date of publication 13 March 2024, date of current version 28 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3377448

RESEARCH ARTICLE

ISO-Compatible Personal Temperature Measurement Using Visual and Thermal Images With Facial Region of Interest Detection

BARTOSZ PTAK[1], PRZEMYSŁAW ASZKOWSKI[1], JOANNA WEISSENBERG[2], MAREK KRAFT[1], (Member, IEEE), AND MICHAŁ WEISSENBERG[2]

[1]Institute of Robotics and Machine Intelligence, Poznań University of Technology, 60-965 Poznań, Poland
[2]Institute of Communication and Computer Networks, Poznań University of Technology, 60-965 Poznań, Poland

Corresponding author: Marek Kraft (marek.kraft@put.poznan.pl)

This work was supported in part by the European Development Fund under Grant POIR.01.01.01-00-0662/20, in part by squareTec Balicki Kozlowski Weissenberg sp. j. and Milton Essex SA under Grant 0313/PRJG/1309, and in part by the Polish Ministry of Education and Science under Grant 0313/SBAD/1307.

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

ABSTRACT Disease outbreaks and pandemics show us how important it is to limit the spread of diseases. One common indicator of many ailments is body temperature. It's a measurement that can be taken quickly, also using contactless methods. However, it is necessary to ensure the methodological correctness, repeatability and reliability of such measurement. In this manuscript, we introduce a non-intrusive approach for individual body temperature assessment that adheres to the stipulated criteria outlined by ISO/IEC 80601-2-59 standard. The measurements are performed at specific regions of interest (ROIs) of a human face, at the inner canthi of both eyes, which show high robustness to the environment temperature change. The method utilises the fusion of RGB-D (red, green, blue and depth) and thermal cameras. The system detects the ROIs on the RGB image employing deep learning methods and transfers them to the thermal image, from which the temperature can be read. The system was tested on our validation dataset consisting of 210 individuals, achieving ROI's position identification mean error below 3 mm and temperature measurement error below 0.5°C, which is in line with the ISO norm requirements.

INDEX TERMS Thermal imaging, temperature measurement, computer vision, deep learning.

I. INTRODUCTION

Due to the observed increase in infectious diseases in recent years, as well as the ease of their spread [1], as exemplified by the COVID-19 pandemic that began in 2019, it has become necessary to perform quick and screening tests to detect potentially infected people [2]. At the same time, higher body temperature is one of the most common symptoms of viral disease in humans [3].

In hospital conditions, measurement can be performed in many ways, including oral, rectal, and temporal artery thermometry, and by introducing temperature sensors during catheterisation [4]. However, all these methods are relatively time-consuming and often very invasive, which makes it impossible to use them in screening tests in public places such as train stations, airports, etc. At the same time, due to the increasing mobility of people, which leads to an increase in infections [1], conducting tests in these places is crucial to stop the spread of such diseases. Consequently, several

The associate editor coordinating the review of this manuscript and approving it for publication was Yongjie Li.

ideas can be found in the literature which enable real-time temperature determination [5].

This aligns with the research outlined in [6], which delineates the inherent challenges in thermometry regarding the precise interpretation of accuracy and repeatability across various thermometer modalities. The authors undertook a comparative analysis of surface body temperature measurement techniques, notably employing non-contact (infrared) thermometry alongside contact-based thermometers (mercury, mercury-free, electronic). Notably, temperatures recorded via contact thermometers placed under the armpit exhibited elevation compared to non-contact measurements taken at the forehead. Statistical analysis of measurement reliability between contact-based thermometers and non-contact infrared thermometers revealed robust outcomes. However, conventional clinical thermometers typically boast an accuracy margin of approximately 0.1 °C. Despite their efficacy, these methods necessitate patient contact, extended testing duration, and pose challenges in simultaneous screening of numerous individuals to promptly identify those with elevated temperatures. Addressing these challenges necessitates the advancement of a non-intrusive temperature measurement system utilising thermal image analysis, conforming to medical device standards.

One of the most popular solutions used in this case is using infrared thermograms through thermal cameras [2], [7], [8], [9]. A thermal camera uses infrared technology to detect and measure the temperature of objects in its field of view. It works by capturing infrared radiation, or heat, emitted by objects and converting it to a temperature values matrix, which can be interpreted and processed as an image. As research shows, the use of this technology is suitable for detecting people with elevated temperatures in a short time [2], [10], [11], [12]. At the same time, because no human is required to operate the system, it meets all the requirements for screening tests.

The critical aspect in the matter of temperature measurement using thermal cameras is determining the measurement area. In the literature, many articles can be found on the measurement of temperature in different parts of the human face [12], [13], [14]. One has to note, however, that while numerous research and commercial solutions were proposed, there is a shortage of viable, consistently applied standard procedures for deployment and screening [12]. According to the *ISO/IEC 80601-2-59* standard [15], the Region of Interest (ROI) can be distinguished as an elliptical area with the following antipodal points: (*i*) the corner of the eye and the end of the tear duct (forming the main axis of the ellipse) and the upper part of the tear duct (*ii*) the lower boundary of the lower tear duct (forming the minor axis of the ellipse). This area is marked in Fig. 1.

Defining the measurement area in such a specific and small area on the human face is challenging, especially considering the resolution and blur of the thermal imaging and the movement of the subject being examined. The images taken with a thermal camera are, to a large extent, devoid of natural textures and landmarks clearly distinguishable in the visible spectrum RGB images. As a result, it is not easy to identify critical points such as the corners of the eyes, the centre of the nose, the corners of the mouth, and the line of the face oval, which are the basis of building the face grid in many algorithms.

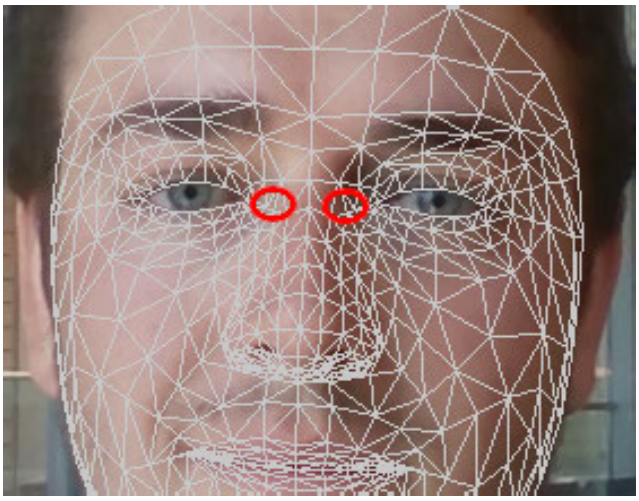


FIGURE 1. The image presents face mesh (grey), obtained with one of the tested algorithms, and temperature measurement areas (red) that are defined in Annex AAA of *EN IEC 80601-2-59* [15], each of ROI-1/ROI-2.

Additionally, in the study of implementing a medical temperature measurement solution, the guidelines presented in the *ISO/TR 13154 Technical Report* [16] must be taken into account. This document provides guidelines for implementing the *ISO/IEC 80601-2-59* standard to prevent the spread of infections. It outlines requirements for device deployment planning and data acquisition, such as performing the measurement in a room free from air currents and with a non-reflective background. The report also includes guidelines for conducting medical examinations, information about the responsibility of the person conducting the measurement, and interpretation of results. However, this part is not relevant to the temperature measurement process, which is the basis for the research carried out within the publication.

In this research, we present the first approach to temperature measurement that meets the requirements of a medical device and the *ISO/IEC 80601-2-59* standard and technical report. The system uses a fusion of information from consumer-grade RGB-D and thermal cameras. Thanks to blackbody and prior system calibration, it identifies the appropriate areas for measurement and ensures the co-location of RGB and thermography data.

As part of the method validation, a custom dataset of 210 persons was collected, containing synchronised RGB-D and thermal images. The demographic composition of our study group revealed a predominant representation of males within the 20-30 age bracket, comprising 187 individuals. Additionally, 17 females within the same age range, along with 6 males aged between 40 and 60, contributed to the dataset. Among the participants, 208 identified as white

persons, whereas 2 identified as black persons. To ensure uniformity and consistency in image acquisition, participants wearing glasses were instructed to remove them during the data collection process. The size of the research sample was determined using statistical methods. The significance threshold was set at 5%, while the confidence level was established at 95%, taking into account an assessment of inherent risk at the mean level. The evaluation of temperature measurements performed using the dataset confirms the correct operation of the presented method and meets the ISO standard requirements. The pipeline steps for performing this task are described in Section II-C of the article. With the implemented solution, it is possible to transfer the detected measurement area to the thermal image. The process of visual-to-thermal transition methodology is described in Section II-D. After the transfer, it is possible to acquire the raw measurement results and then determine the temperature of the person being tested according to the algorithm proposed in Section II-G.

A. RELATED WORK

Thermal imaging is a technology that allows for the detection of heat energy emitted by an object, providing a visual representation of temperature variations within an image. This technology has a wide range of applications, from industrial and scientific uses to security [17] and surveillance [18], [19]. Its ability to detect heat signatures in various conditions such as medical imaging [20], [21], animal behaviour research [22] and social distancing compliance [23] makes it a valuable tool for many industries. In industrial settings, thermal imaging cameras can be used to detect temperature variations in equipment and machinery [24], [25], which can help identify potential issues before they lead to equipment failure. As another example, in the automotive industry, thermal imaging cameras can be used to detect hot spots in engines and electrical systems [26], allowing for early detection and maintenance of potential problems. In the energy industry, thermal imaging cameras can be used to detect areas of heat loss in power plants [27] and other facilities [28], which can help identify ways to improve energy efficiency and reduce costs.

Moreover, general precise temperature measurements using thermal imaging are widely applied. Application of this technology is common in the animal husbandry industry. Over a long time, systems were developed to facilitate the control of animal well-being by using non-invasive temperature measurement methods with thermal cameras [29]. In [30], deep learning is employed to detect regions of interest and then measure the surface temperature of the cow's eye, whereas in [31], they focus on measuring cattle temperature with real-time constraints based on a measurement sequence to improve accuracy and reduce animal stress. In contrast, well-being measurement solutions for humans are under development. These systems are grounded in both deep learning and thermal imaging and are designed to predict individual thermal preferences. One can note a characteristic workflow. Firstly, detect the key features of the human body. In [32], the hands, neck, cheeks, and forehead are detected; in [33], the entire face is identified; meanwhile, in [34], facial landmarks selected by the authors are sensed. In the second step, the key points are used to track temperature changes in the fluctuating environment and provide an indication of a personal comfort level. All of these mentioned methods detect features directly on thermal images.

Finally, there are several approaches and algorithms in the literature for determining a person's body temperature based on thermal imaging. In the article [14], the authors presented a solution in which data acquisition was performed through a system consisting of a BeagleBone Black board and a FLIR thermal camera, and the entire solution was based on open-source libraries. The measurement was based on a point determined in the centre of the forehead of the person being examined in relation to the highest and lowest temperature recorded using the Tesseract OCR library. To increase the reliability of the results, the author implemented an averaging mechanism of the result from 10 measurements and discarded values considered false, that is, with increased deviation. However, the entire study does not meet the requirements for a medical device described in the *ISO/IEC 80601-2-59* standard. The authors of the articles [9], [12] conducted clinical studies to evaluate the location of the effect of the measurement point on the temperature measurement results and to compare the calibration methods and determine the best practices for evaluating the performance of infrared thermographs intended to detect elevated body temperatures. In article [9], measurements were taken for 17 points on the face and compared with an oral thermometer. The results confirmed that the ISO-compliant approach of using the inner canthi temperatures gives more accurate results. It was also indicated that similar results were achieved with an approach using the maximum temperature of the entire face. Because the most reliable body temperature measurements are obtained when analysing the area located in the inner canthi, the work [11] proposes a two-phase system to detect measurement points inside the corners of the eyes and nostrils on thermal images. In Phase I, four types of processed images are generated: clipped (CLIP), inverse clipped (ICLIP), gradient magnitude (GRAD), and inverse gradient magnitude (IGRAD). Various image enhancement techniques such as sharpening, image clipping, quantisation, and gradient are used to obtain them. Then, in Phase II, based on the features found in the images in Phase I, face detection and characteristic points detection are performed, and finally, the estimation of the location of the aforementioned measurement areas. Based on the analysis, the authors indicated that for the system tested, the best solution was to use ICLIP images. In the [13], an automatic approach for locating the area of interest (inner canthi) was presented. In the algorithm, the first step was roughly to detect five key facial points (centre of the eyes, tip of the nose, and

ears). In the next step, using the 3D Morphable Face Model (3DMM), a sparse set of corresponding 2D-3D points was calculated, which is used to the projection of the entire 3D face onto the image and then locate the inner corner of the eye along with the improvement of the location of the angle of the eye. The two-step approach is employed, as detecting small regions of interest in thermographic images is problematic due to their mostly textureless and featureless nature. The research showed the possibility of identifying the measurement area on the thermogram. Some researchers focus also on compensation methods for temperature measurement using thermal imaging. In [35], authors propose simple mathematical models for compensating the influence of ambient temperature and measuring distance variations for their system consisting of a combination of a 3-D depth sensor, an electronic temperature sensor, and a reference temperature. A more complex compensation model is proposed in [36]. It uses polynomial regression, allowing for an adaptative approach. However, as mentioned in [37], most thermography-based personal temperature screening systems used in practice ignore the recommendations and guidelines on the measurement protocol. While the details on commercial solutions are in general not available to the public, we found that the vast majority of the solutions described in the scientific literature are also not compliant or only partially compliant with the ISO guidelines.

B. RESEARCH CONTRIBUTION

The main achievements of this article are as follows:

- Creation of a data acquisition system using a RGB and thermal imaging camera which, together with the methods developed, will enable the determination of the temperature of the subject in real time in a non-invasive manner meeting the medical device standard ISO/IEC 80601-2-59
- The development of a method enabling the identification of ROIs (Regions of Interest) in thermal images, thereby facilitating the measurement of human temperature at specific points through the detection of transformations between RGB and thermal images.
- Developing software to detect key points in the RGB image and convert them into thermal image coordinates, while also determining the distance between the camera and the human face using a depth sensor.
- The developed solution complies with the requirements of the medical device standard (ISO/IEC 80601-2-59) which distinguishes it from other known solutions in the literature.
- The solution uses a combination of both deep learning and classical methods to estimate ellipsoidal regions near the canthus. The Mediapipe framework was used for the deep learning part.
- The developed solution achieves an error of 2.53 $\pm$ 1.40 mm in the visual part of the keypoint detection and generates an error of 0.4 $\pm$ 0.39 mm during the visual to thermal conversion. The combined ROI detection error in the RGB image and the transformation error to thermal image is 2.93 $\pm$ 1.79 mm (mean $\pm$ std) and meets the requirements of ISO/IEC 80601-2-59.

II. MATERIALS AND METHODS

A. SENSORS SETUP

This section describes the hardware, data, methods and algorithms used to obtain the results. The numerical results and statistical analyses are enclosed in the next section.

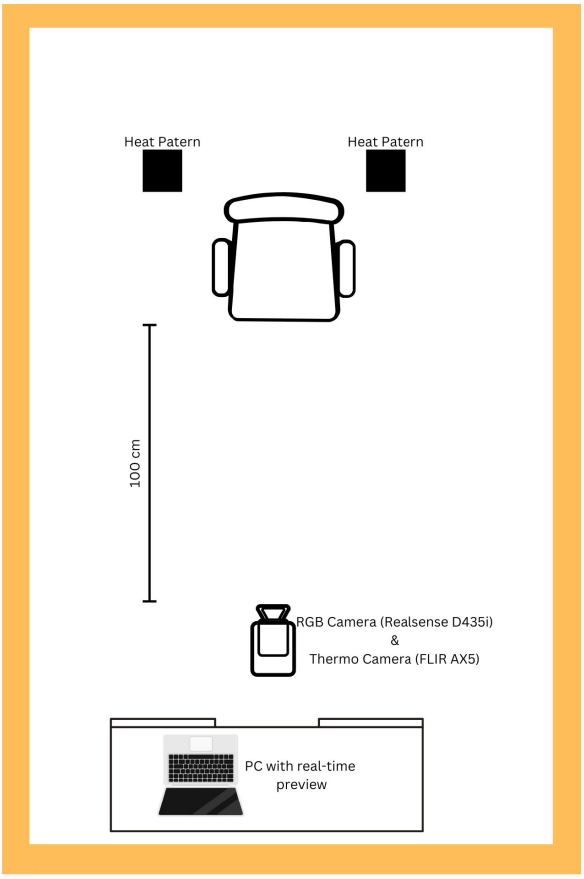


FIGURE 2. Block diagram of configuration elements and their arrangement.

The research platform was constructed with a various of sensors. It allows the collection of a dataset and provides a real-time preview for the user. It consists of the parts described below, a block diagram of the system is presented in Fig. 2, while a view of the built system during data acquisition is shown in Fig. 3.

- **Host computer**. It is a Lenovo Mini PC equipped with Intel i5-8600T CPU and 16 GB DDR4 RAM. It operates on the 64-bit Ubuntu 20.04 and enables communication by Ethernet, WiFi, and many high-speed serial ports.
- **RGB-D camera**. The Intel RealSense D435i is used as a visual (RGB) and depth sensor. It was selected because it is accessible, low-cost, and user-friendly. It provides a resolution of 1920 $\times$ 1080 for colour images and 1280 $\times$ 720 pixels for depth images. The

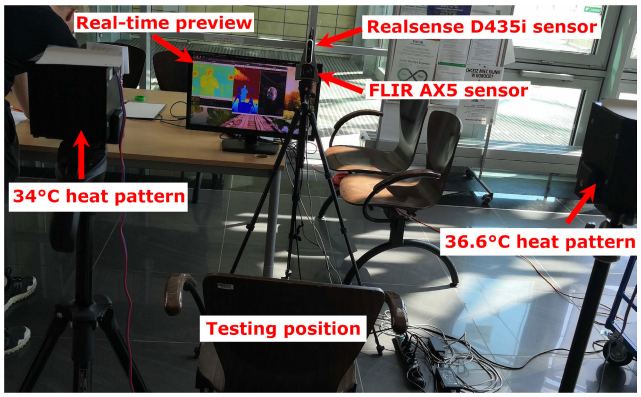


FIGURE 3. An overview of the setup's elements and their arrangement.

communication with the host computer is performed by a USB 3.1 connection.

- **Thermal sensor**. A temperature sensor, the FLIR AX5 camera, was attached on top of the RGB camera so that their optical axes were roughly parallel. The AX5-Series are low-cost infrared cameras. The A65 captures infrared images with a relatively high resolution of 640 × 512 pixels and transports them via Ethernet API.
- **Blackbody calibration sources**. Two temperature patterns are installed in the system for thermal reference. They emit a constant temperature of 34.0 and 36.6 degrees Celsius, respectively. The patterns emit heat at a constant value and, by imitating the behaviour of a blackbody, will not absorb heat from the environment. Due to the high precision of the emitted heat by the units (change every 0.1 degree Celsius), using them as reference results in a low measurement error.

1) SETUP REQUIREMENTS

The design of the research and test environment was done based on the specifications and requirements described in the *ISO/IEC 80601-2-59* standard and the *ISO/TR 13154* technical report. The choice of hardware components and system parameters were driven by the constraints of the aforementioned regulatory documents. Satisfying the requirements outlined in the ISO standard and the report enables the developed system to be classified as a medical tool,making it applicable in medical facilities and confirming its accuracy. In order to meet these requirements in our system, a range of specifications is considered, including:

- A skin emissivity of 0.98 is used. This is especially important when measuring body temperature in a medical or clinical setting, as even slight variations in temperature can have significant implications for diagnosis.
- The resolution of the face surface in the measured image has to be at a minimum resolution of 320 × 240 pixels. This is satisfied by using a thermal camera with a resolution of 640 × 512, in which our work area (corresponding to face area) covers 480 × 360 pixels (in this specific thermal imaging camera, this resolution covered at least 56.25% of the total image.
- The external blackbody calibration source is present in the thermal image, and its size is 50 × 50 px, which corresponds to less than 10% of the person's face.
- The system, in accordance with the requirements of the standard, enabled the measurement of temperature in the range of 33°C to 40°C with an expanded uncertainty of 0.3°C and a stability and drift of +- 0.1°C. Moreover, it is possible to measure temperature beyond this range.
- The distance between the thermal camera and human faces is in the range of 0.65 to 0.95 meters, which allows it to meet the above needs while maintaining a safe distance from the subject. This is monitored in real-time using the depth camera.
- For visual and thermal cameras, real-world values of sensor resolution are 0.6 mm and 0.5 mm per pixel, respectively. These values are within the recommended range for evaluating facial keypoint measurements.
- Each ROI-1/ROI-2 area is defined as an elliptical area with the following antipodal points:
  - the corner of the eye and the end of the lacrimal sac (forming the major axis of the ellipse)
  - the upper part of the lacrimal sac and the inferior border of the position of the lower lacrimal duct (forming the minor axis of the ellipse).

The benefit of using the multi-camera pipeline is that it enables the real-time generation of cues for the scanned user if the conditions for performing a reliable measurement are not met. The system checks if the user is not too close or too far and monitors if the user faces the camera – excessive head tilting is also detected. Based on that, the user is asked to correct his/her position, and the measurement is performed only when the face is within the correct range and facing the measurement setup.

It is also critical to pay attention to some camera's technical notes. Because of the different camera communication protocols, only software synchronisation could be performed. Synchronisation is one of the important things while objects are moving, and the lack thereof can cause a decrease in accuracy.

B. THE EVALUATION DATASET

In order to evaluate the developed method, a dataset of face samples has been collected. The dataset contains samples collected from 210 different individuals. Each individual contributed three samples for a total of 630 test samples. The face on each image is centred, with only a slight rotation allowed, which was enforced during recordings. Each sample item contains the following data:

- an RGB image acquired using the Intel D435 depth camera,
- its corresponding depth image acquired using the same camera,
- a thermal image collected using the FLIR AX5 thermal camera.

Two blackbody calibrators (34°C and 36.6°C) were visible in the top corners of each thermal image. Each person who contributed their images to the dataset has signed a consent form. The registration system provided a live preview of the images visible to both cameras. Moreover, based alignment of the subject's head was monitored in real-time to ensure the view was frontal and the head held properly upright. For each of the 630 collected RGB images, hand annotation of the regions of interest is associated with the area described in the regulations.

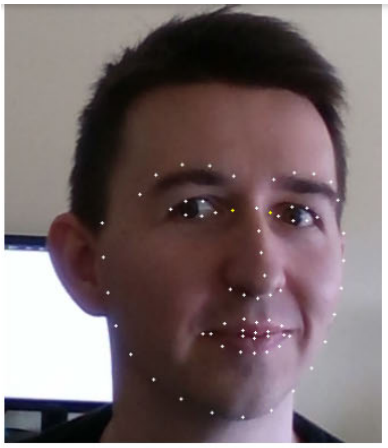
C. FACE KEYPOINT DETECTION ALGORITHMS OVERVIEW

Many approaches to detecting face key points were introduced. Due to this, we decided to compare three of the most popular frameworks in our work and describe them below. Each one of them is a two-stage algorithm: first for face region of interest estimation and second for point coordinate regression inside the region of interest. Dividing the problem into two steps allows us to replace one module with another. The visual comparison of algorithms results is included in Fig. 4.

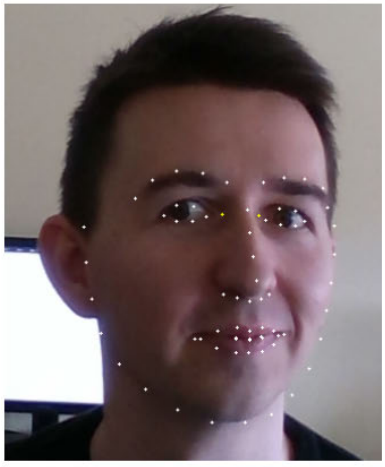
Dlib pipeline [38] is the method that uses Histograms of Oriented Gradient [39] for human face detection. The technique is based on normalised local histograms of image gradient orientations for extracting features and linear SVM to classify them. Next, the regions of interest are forwarded to the second stage [40]. Based on regression trees, it fits prior structure into image data and estimates the face's landmark positions - directly from pixels with real-time performance. This classic pipeline is fast on the CPU, and many implementation additions make it scale- and rotate-invariant.

Menpo pipeline [41] performs the face detection based on Cascade Classifier introduced by Viola and Jones in [42]. It extracts Haar features at multiple scales and classifies the image region as face or non-face using the AdaBoost ensemble classifier [43] based on the extracted features. It covers the human face detection part of the Menpo pipeline. As a facial key point estimator part, the method uses the Patch-based Active Appearance Models (AAM) [44]. AAM is a generative model that consists of a statistical parametric model of the shape and appearance of a human face. Patch-based AAM is a variant that splits and simply samples patches around the image's landmarks. The algorithm is iterative and requires defining a maximum number of iterations over all scales that influences to final result and performance.

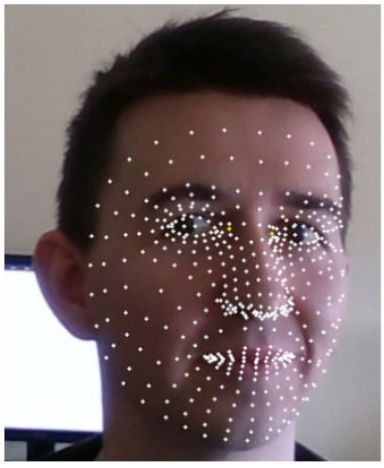
Mediapipe pipeline [45] utilises deep neural networks in both stages. The first stage uses the BlazeFace [46] neural network, which is a very effective architecture for mobile GPUs. Moreover, the network provides not only bounding box regression for the human face but also a few facial key points, which enables the reduction of translation and rotation influence. For face landmarks, the pipeline implements another efficient deep learning model [47]. It allows for inferring an approximate 3D mesh representation of a human face. The mesh contains as many as 468 points as specified in [47]. Moreover, the model can finish a face that is partially



(a) Dlib method.



(b) Menpo method.



(c) Mediapipe method.

FIGURE 4. Results of the comparison of individual facial keypoint detection pipelines.

obscured or crosses the image boundary. However, when the Mediapipe pipeline is used on video input, it uses face tracking mode to make face detector usage redundant. This deep learning method is trained using synthetic renders and fine-tuned using real images. Therefore, the method provides a high capacity for generalisation.

Once the facial keypoints are detected, it is possible to create a quadrangle described by four specific points, which is assumed to be ROI. As for further processing only the centre position of each ROI is required, it is calculated as the centre of gravity of the ROI quadrilateral.

A comparison of ROI centre detection results achieved using the described methods is given in III-A, and visual comparison is shown in Fig. 4.

D. VISUAL-TO-THERMAL TRANSITION METHODOLOGY

Initially, some attempts were made to use face keypoint detection directly on thermal images. However, the methods described in section II-C did not meet the accuracy-related assumptions required for ISO compliance. As mentioned before, this is mostly due to the fact that thermal face images lack distinctive features and texture when compared to RGB images. Moreover, the methods are in general not prepared to operate on thermal images and are easily misguided by features distorting surface temperature such as the arrangement of hair, facial hair etc. In addition, the lack of a dataset with face keypoint markers in thermal imaging made it impossible to fine-tune the algorithms from the RGB domain to the thermal imaging domain. To handle this barrier, it was decided to combine machine learning methods with geometric projection methods. Therefore, after detecting facial key points and calculating the ROI centre points on the RGB image, they need to be translated into the thermal image. Fig. 5 shows an example pair of RGB and thermal images of one face, with keypoints transformation applied.

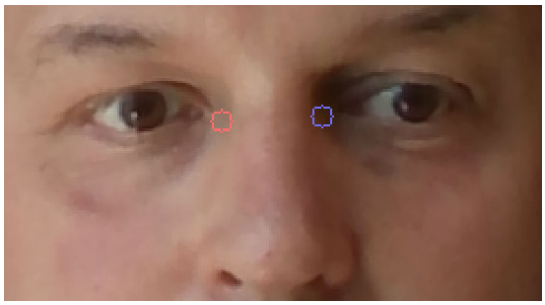
1) IMAGE ALIGNMENT ALGORITHM

The developed method allows the transformation of a single point with fixed transformation parameters computed at calibration time. Both cameras were calibrated individually using Zhang's method [48]. Therefore, both cameras must be physically constrained once the one-time calibration is done. The transformation's input data are as follows:

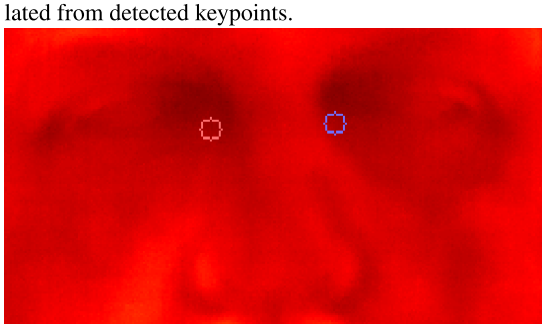
- pixel position on the RGB camera image - coordinates (x_1, y_1),
- pixel value on the depth camera at (x_1, z_1) point - z_1,

which are converted to the global coordinate system position (x_2, y_2, z_2). The global coordinate system is anchored at the RGB camera coordinate system, with overlapping axes, with coordinates expressed in cm units. The result of the transformation is the pixel position on the thermal camera image (x_3, y_3).

As the transformation is performed only on the ROI centre points, the working area can be constrained to a rectangular cuboid in the space where these points occur, based on the distribution of the collected data. The distribution of the pixel x position (horizontal coordinate) in the collection is presented in Fig. 6.



(a) RGB image with two ROI areas marked, both calculated from detected keypoints.



(b) Thermal image with two ROI areas marked, both calculated as points transformation from the RGB image. Images shown zoomed into the area of interest.

FIGURE 5. Example of ROI detection on the RGB image, and transformation to the thermal image.

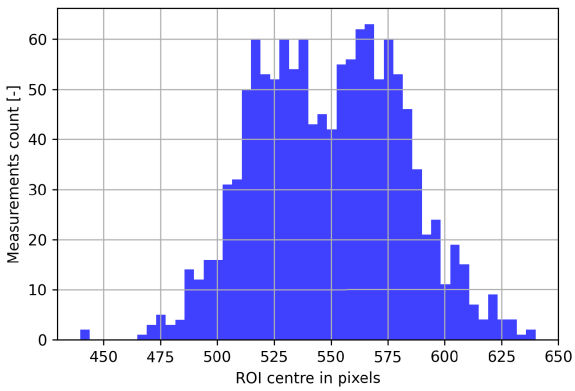


FIGURE 6. Distribution of ROI centre points x coordinate in the dataset. Peaks for left and right eye are clearly visible.

The points are transformed according to the following algorithm:

1) Take position x_1, y_1 of the pixel on the RGB image.
2) Find depth value on the depth image at position x_1, y_1. The median value from a patch 5×5 around the pixel is calculated to remove noise. The patch size of 3×3 mm is small enough, in the scale of a face, to be of a similar depth, but big enough to remove outliers. The depth is already expressed in cm, therefore $z_2 = z_1$,

since the RGB-D camera is internally calibrated. For the detected ROI, the patch is created around the ROI's centre point.

3) Calculate x_2, y_2 position in global coordinates, assuming pinhole camera model, as:

$$\begin{cases} x_2 = (x_1 - x_{1centre}) \cdot z_2/f_x \\ y_2 = (y_1 - y_{1centre}) \cdot z_2/f_y \end{cases} \quad (1)$$

where f_x, f_y are the focal lengths of the camera in x and y axes and $(x_{1centre}, y_{1centre})$ is the pixel position of the camera optical axis, known from the camera calibration matrix.

4) Calculate x_3, y_3 with the following equation:

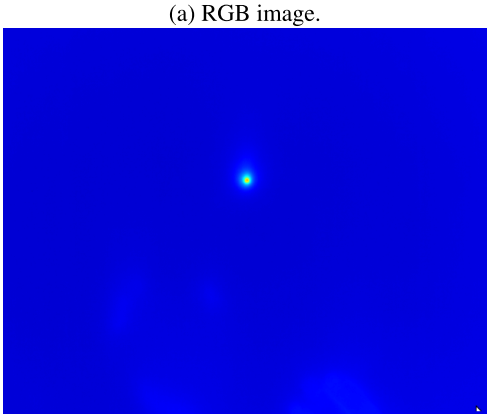
$$[x_3, y_3, 1] = M_{thermal} \cdot [M_r|M_t] \cdot [x_2, y_2, z_2, 1] \quad (2)$$

It needs to be noted, that due to the similar resolution of both cameras (in cm per pixel) and the nature of this transformation, the calculated x_3 and y_3 errors are proportional to the error of the z_1 depth measurements.

Cameras matrices have been calculated during single camera calibration with a chessboard pattern.

$M_{thermal}$ is the calibration matrix of the thermal camera, which can be calculated with a chessboard pattern using Zhang's method [48], or filled without distortion parameters if the image distortion is not significant. Rotation matrix M_r and translation matrix M_t have been calculated during the transformation calibration procedure which we'll now describe.

(a) RGB image.



(b) Thermal image.

FIGURE 7. Calibration board with a single point (powered resistor).

2) MULTISENSOR CALIBRATION PROCEDURE

In order to perform the multisensor calibration, we need to establish correspondence between a set of points on the RGB and depth images and the thermal images. To collect such data a special marker is required, which can be easily identified in both RGB and thermal modalities and manually annotated. This condition means high contrast on both images - in terms of colour on the RGB and temperature on the thermal image. To meet these requirements, a small powered resistor on white cardboard was used, as presented in Fig. 7.

Figure 8 presents a zoomed and annotated area of the resistor target on both cameras.

During data collection, two calibration collections were gathered - one for the calibration process and one for validation. During each experiment, the resistor target was moved to cover the entire working area described by a rectangular cuboid. Then the collected image pairs were manually annotated, as presented in Fig. 8.

First, the annotated points on the RGB and depth image were translated to the global coordinate, as already described in subsection II-D1. Then, having a set of x, y, z points in the global coordinate system and pixel position on the thermal camera, the solvePnP (RANSAC) [49] method from the OpenCV library [50] [51] was used to calculate the matrices M_r and M_t.

3) ALTERNATIVE MULTISENSOR CALIBRATION METHODS

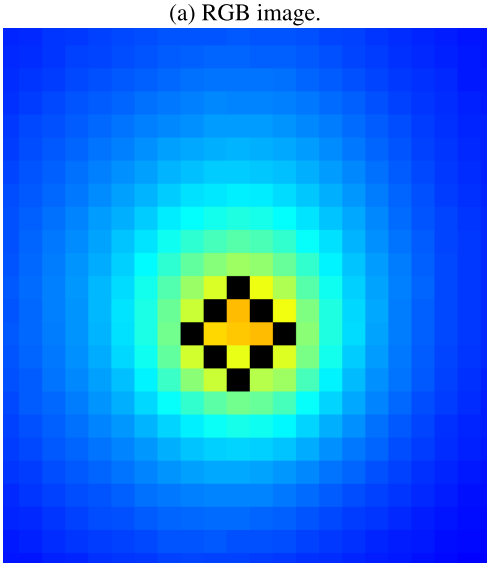
In [52] multi-modal geometric calibration target was developed, based on steel planes with an active heating system. Similarly, in [53] aluminium composite plates were prepared for both circular and chessboard targets.

Unfortunately, these approaches require a huge amount of precise tools to be built. Therefore, we opted for using a method that was described in [54] where a classical chessboard was used for the thermal camera after preheating it with a 250W heat lamp. We adopted the approach and collected images for a 7×4 chessboard pattern and for a 9×3 circle pattern heated by sunlight. The visual effect was illustrated in Fig. 9. Using the OpenCV library, the points of interest were extracted and a calibration process was performed. However, obtained re-projection errors were very high, because of blurred marker edges. Due to the problem with low-quality calibration results, we compute the calibration matrix using the method described in II-D2.

E. KEYPOINTS DETECTION DIRECTLY ON THERMAL IMAGE

An alternative approach to detecting keypoints on the RGB image and transforming them to the thermal image coordinates is detecting the keypoints directly on the thermal image. Doing it would allow simplifying the setup and methodology.

(a) RGB image.



(b) Thermal image.

FIGURE 8. Zoomed calibration board with a single point (powered resistor).

Unfortunately, the thermal spectrum does not carry as rich information as visible light in terms of surface colours and texture. Therefore, detecting a face and its keypoints by relying solely on the thermal image is less precise, which was confirmed, as the method was also validated. The thermal image was converted directly to a grayscale image, with temperature values clipped between 20 and 45 degrees Celsius and scaled up to the full pixel range. As the next step, the grayscale image was provided as input to the Mediapipe model. An example image with detected keypoints is presented in Fig. 10. One can notice, that although the keypoints mesh is misaligned at the top and bottom of the face. Moreover, on our test dataset, for 7 thermal images, the Mediapipe library failed to detect a face on it. As the dataset contains annotations only on the RGB images, and

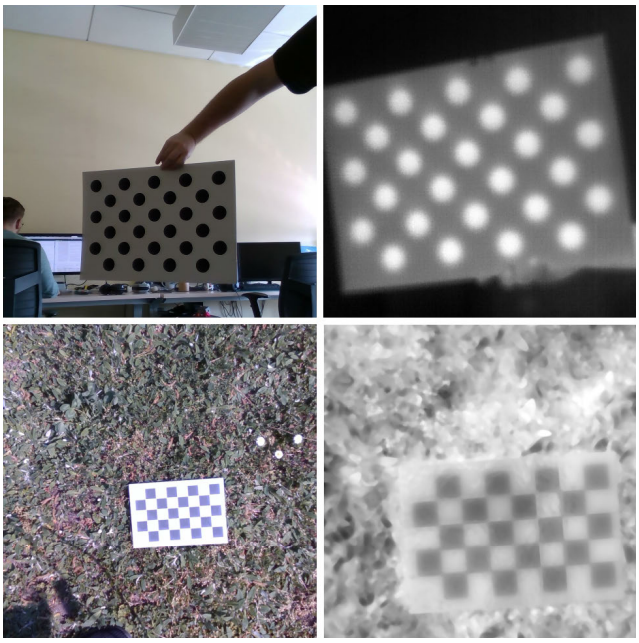


FIGURE 9. Classical calibration targets heated by the sun. Comparison of circle pattern (upper) and chessboard (lower) captured by the RGB camera (left) and a thermal camera (right).

not on the thermal images, ROIs detection accuracy cannot be performed directly. Nevertheless, the position of ROIs detected directly on the thermal image was compared with the position obtained with the transformation of annotations. The average distance between the aforementioned ROIs is 6.41 pixels, with a standard deviation of 7.16 pixels. Inspection of the results showed a significant number of outliers, which is also suggested by a high standard deviation value, compared to the mean. For 10 results the ROI centre error is higher than 50 pixels.

Apart from the low details quality of the thermal image, it has to be noted that the algorithms from the Mediapipe library are not aimed to be used on thermal images, and fine-tuning it on such data would probably increase the quality of the model output. This could be, to some extent, alleviated by performing training using thermal images, but doing so would require a sizeable annotated training set. Nevertheless, thermal images are largely textureless and precise localisation of the ROI for temperature measurement is much easier and more accurate using RGB images.

F. CORRECTION OF TEMPERATURE WITH BLACKBODY CALIBRATION SOURCES

Thermal sensors are characterised by measurement drift, which is a well-known effect explained in the literature [55], [56]. Aware of this phenomenon, we used reference temperature sources during data collection. In our case, we utilised two blackbodies emitting a temperature of 34°C and 36.6°C. These are illustrated in the workstation overview 3. Fig. 11 illustrates the changes that occurred while the camera was capturing these patterns' values. The high variability of the

FIGURE 10. Detection of keypoints directly on a thermal image (interpreted as a grayscale image) with the Mediapipe algorithm.

raw thermal camera measurements emphasises the need for using reference temperature sources. It is also clearly visible, that fluctuations in the measured values for both blackbody calibrators follow the same pattern.

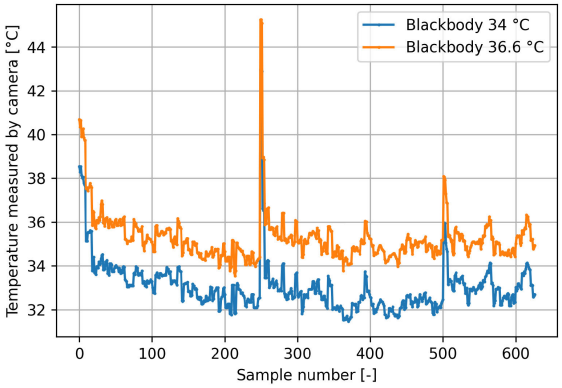


FIGURE 11. The temperature comparison of two blackbody calibration sources (34°C – orange and 36.6°C – blue) measured by the thermal camera. Samples span over 3 days of data collection.

Based on the specific requirements of our case at hand, we selected only one of them - 36.6°C. This particular pattern is chosen because it is closer to the average human body temperature, which makes it a suitable reference point for various measurements and observations. Additionally, using the 36.6°C pattern as the baseline can help to correct any potential errors or discrepancies. The correction of temperature readings is performed by measuring blackbody value and difference calculation. Then, the difference is applied to the temperature of sensed facial regions.

G. TEMPERATURE CALCULATION METHOD

Once the ROI centres areas are transferred to the thermal image, the following steps are applied to determine a single temperature measurement for a person:

TABLE 1. ROI centre position detection error for different algorithms on the RGB images. The error is calculated as a difference between the manual annotation point and algorithm detection.

Algorithm	ROI detection error [mm]: *mean* $\pm$ *std*
Dlib	6.57 ± 26.76
Menpo	13.72 ± 27.65
Mediapipe	2.53 ± 1.40

- Take the temperature value for each pixel in the circle with a two-pixel radius around the ROI centre, yielding 13 values.
- Calculate the median value of these 13 pixels for each ROI, yielding two values for the image (both left and right canthi)
- Take the average value of the median computed from both ROIs

Taking the median temperature value within the ROI allows for the elimination of outliers and measuring errors, and was determined to be superior to taking the average value. This method is commonly used in medical imaging to measure the temperature of a specific area of interest [57], [58].

III. RESULTS

A. ROI DETECTION ON RGB IMAGES

Some of the keypoint detection algorithms described in section II-C have a limited number of keypoints. Moreover, the detected keypoints may not be aligned with the required ROI. Therefore, for each algorithm, the ROI centre position was calculated as a weighted average of the four closest neighbouring keypoints it is capable of detecting. Moreover, the calculation ensures that selected ROIs do not overlap the eye area. ROI centre points obtained from each algorithm have been compared with the manual annotations from the dataset, with results presented in Table 1. The errors were calculated in pixels and converted to millimetres, using depth measurements and intrinsic camera parameters. The results show that the Mediapipe method generates 2.6 times and 5.4 times smaller detection errors than other evaluated methods.

B. POINTS TRANSFORMATION TO THERMAL IMAGE

Matrices M_r and M_t calculated on the training dataset were tested on an independent validation dataset, consisting of 24 point pairs. The points from the RGB image were transferred to the thermal image and compared with the manual annotations by calculating the absolute position error. The average error was calculated as 0.85 pixels, which corresponds to 0.4 mm. The standard deviation of the error is 0.84 pixels (0.39 mm).

C. TEMPERATURE DISTRIBUTION IN THE DATASETS

Fig. 12 presents the distribution of the measured temperature across the entire dataset of 210 persons, with each value being a median temperature of each ROI.

As there are three images of each person, and 2 ROIs on each image, it yields six ROIs for each person, with

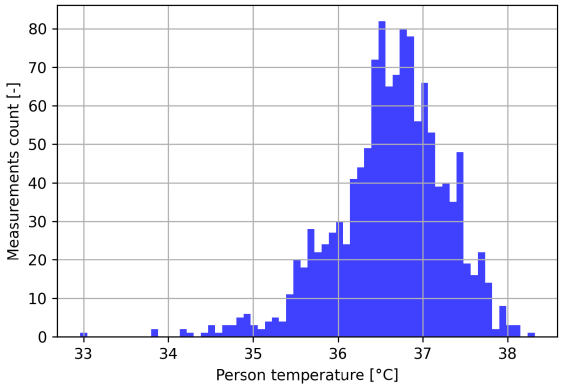


FIGURE 12. Distribution of median ROI temperatures across the entire dataset. Temperature values were corrected with the blackbody calibration source.

six median temperature values. For these 6 values, each person's standard deviation was calculated. Fig. 13 presents the distribution of this standard deviation among the dataset.

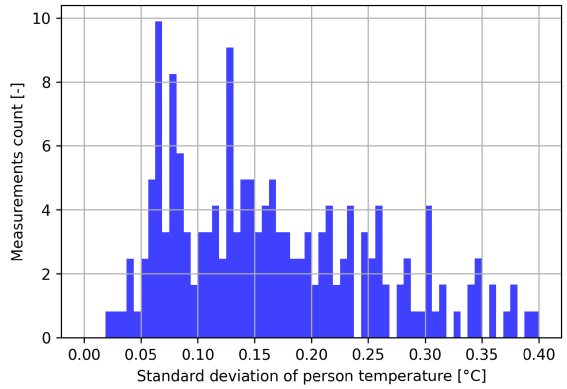


FIGURE 13. Distribution of standard deviation of six median ROIs temperatures for each person, across the entire dataset. Temperature values corrected with the blackbody calibration source.

The distribution shows no clear outliers, confirming the correct operation of ROI selection, transfer, filtering and final measurement generation.

D. BASIC DESCRIPTIVE STATISTIC OF DATASET

The basic descriptive statistics for the data set are presented in table 2. Statistics were performed for the temperatures of the ROI1 and ROI2 areas. The temperatures for ROI1 and ROI2 areas were determined based on the median of the pixel values obtained in the area from the thermographic images and were denoted by the parameters *TCL* and *TCR*, for the left and right eye canthus, respectively. The final temperature result (*TC*) is equal to the median of the results obtained for ROI1 and ROI2. The most common temperature value regardless of the area was a temperature of 36.75°*C* and this occurred for 26, 17, and 8 records for *TCL*, *TCR* and *TC*, respectively. Exactly a quarter of the records were characterised by a temperature of 36.2°*C* or less, and exactly a quarter of the records were characterised by a temperature

TABLE 2. Descriptive statistic of dataset. TCL - temperature of ROI1; TCR - temperature of ROI2; TC - median temperature. Temperature in degrees Celsius.

	TCL (left ROI)	TCR (right ROI)	TC (median)
Mode	36.76	36.76	36.76
Crd. no. of mode	26	17	8
1st quartile	36.24	36.28	36.26
3rd quartile	37	37.12	37.00
Variance	0.46	0.41	0.39
Variability Coeff	1.85	1.75	1.70

TABLE 3. Difference of temperature measurements of ROI1 and ROI2 areas ($|TCL - TCR|$).

EV	0.31
Mode	0.00
Crd. no. of mode	41
1st quartile	0.08
3rd quartile	0.39
Variance	0.09
Variability Coeff	99.76

greater than 37 °*C*. The scatter of the results in each case is centred around the mean values for each region, which is also confirmed by the low coefficients of variation values.

According to the standard [15], the difference between the temperature measurements of ROI1 and ROI2 areas ($|TCL-TCR|$, a measure of measurement uncertainty) should not exceed 0.5°C. The statistics of temperature differences are presented in Table 3. The most common difference is 0 degrees and occurs in 41 records. Exactly one-quarter of the temperature results in the left and right ROI1 and ROI2 areas do not exceed 0.08°C, and exactly three-quarters of the results have values not greater than 0.39°C. The results are highly variable, but they are centred around the expected value of 0.31°C. In light of the obtained results, it can be stated that the *TC* model meets the standard assumption.

E. TEMPERATURE MEASUREMENT ACCURACY AND CORRELATION

Temperatures for ROI1 (*TCL*) and ROI2 (*TCR*) canthi regions, and the final temperature measurement (*TC*) were compared to a blackbody reference temperature (*TR*). Pearson's correlation coefficients r for all measurement locations ranged from 0.86 to 0.95, which implies a strong correlation between the measurements used and the reference temperature. In the study, temperature *TC* showed the highest correlation and agreement with *TR*, followed by *TCR*. As *TC* had the highest value of r, it can be assumed that the *TC* metric will best estimate *TR*.

Correlations with inner canthi temperature were further analysed using scatter plots of all internal canthal regions

TABLE 4. Pearson correlation coefficients (*r* values) between canthi temperatures and reference temperature.

	TCL (left ROI)	*TCR* (right ROI)	*TC* (median)
series 0	0.86	0.93	0.95
series 1	0.87	0.89	0.96
series 2	0.86	0.94	0.95
total	0.86	0.94	0.95

against *TR* (Fig. 14-16). This included *TCL*, as well as *TCR* and *TC*, which showed the best r values. The results were carried out for the entire dataset. As expected, the scatter plots show a large group of *TR* values near $36-38°C$, representing a high proportion of subjects. *TCL* results showed the greatest variability, especially in the lower temperature range. *TCR* and *TC* showed a better correlation with TR in this range, and TC also at higher values of *TR*. *TC* measurements were less variable than other measurements and correlated well with TR, approaching a linear relationship.

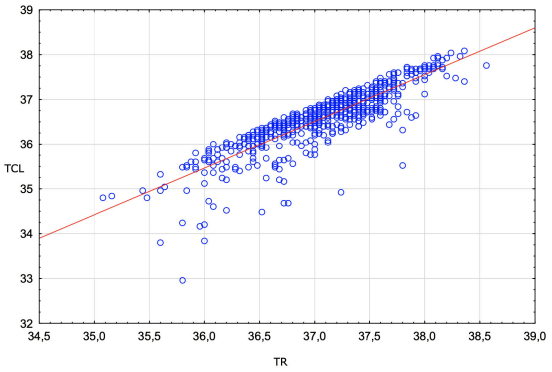


FIGURE 14. Scatter plots of *TCL* versus *TR* (expressed in °*C*), with linear fits.

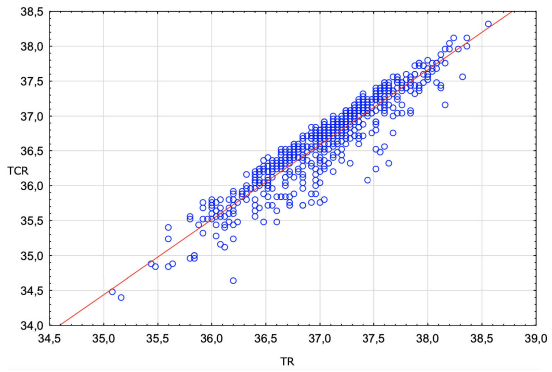


FIGURE 15. Scatter plots of *TCR* versus *TR* (expressed in °*C*), with linear fits.

F. EFFICIENCY OF FEVER DETECTION

ROC curve analysis was used to assess the fever detection performance based on the model proposed in this article.

The ROC curves in Fig. 17 illustrate the performance of *TCL*, *TCR*, and *TC* in discriminating between positive and negative cases. The ROC curve for each measurement shows how sensitivity (true positive rate) and specificity

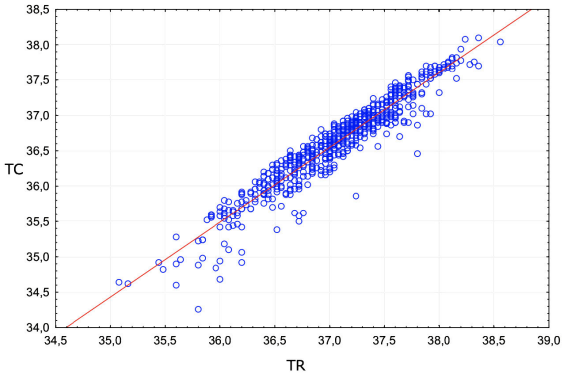


FIGURE 16. Scatter plots of *TC* versus *TR* (expressed in °*C*), with linear fits.

TABLE 5. AUC values from the ROC curves for *TCL*, *TCR* and *TC*.

TCL (left ROI)	*TCR* (right ROI)	*TC* (median)
0.989	0.772	0.958

(true negative rate) change as the discrimination threshold is varied. The line $y = x$ represents random discrimination, and the closer the ROC curve is to the upper left corner of the plot, the higher the discrimination efficiency.

The AUC (Area Under the Curve) values in Table 5 indicate the overall performance of each measurement. An AUC value of 1 represents perfect discrimination, while a value of 0.5 represents random discrimination. The AUC values for *TCR* and *TCR* were both $>$ 0.95, indicating high discrimination efficiency. In contrast, the AUC value for *TCL* was $<$ 0.78, indicating a significantly lower performance. The ROC curve for *TCL* is characterised by a slow approach to high sensitivity as specificity decreases. The difference between the measurements between individual eyes results from the relative location of the camera and the subject, with the right inner canthus being more exposed and directly facing the camera axis, while the left inner canthus' local plane is more parallel to the camera's axis. This emphasises the need for aggregation of performing the measurement in both regions and aggregating the result, as in practice it may be difficult to arrange for all subjects to be facing the camera perfectly.

The correlation coefficients presented in Table 4 indicate the strength and direction of the linear relationship between the measurements. The good fit of the discrimination performance with these correlation coefficients suggests that the discrimination ability of each measurement is consistent with their intercorrelations.

IV. DISCUSSION

The system's characteristics conform to the requirements set by the *ISO/IEC 80601-2-59* norm. This is ensured by the proper choice and spatial placement of the system's components (camera resolution, accuracy, component distances) and the choice of image processing and measurement methods (dense facial keypoint mesh generation, RGB to thermal calibration) and settings (e.g., emissivity). Statistical analysis

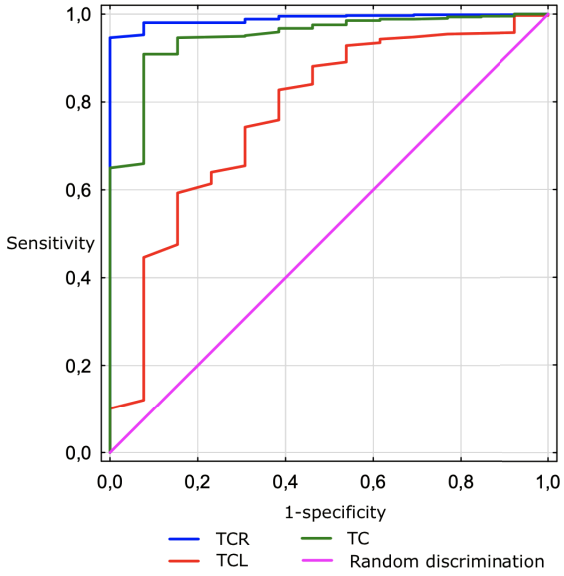


FIGURE 17. ROC curves for temperature detection using temperature of *TCL*, *TCR*, and *TC*.

of the results also confirms that all critical computed and measured parameters are within the prescribed ranges. Worst case canthi location error resulting from both the face keypoint detection and RGB to thermal transfer does not exceed 5 mm. Considering the ROI's size, the detected regions of interest cover the correct anatomical region. Further statistical processing of resulting single measurements (individual pixel ROI temperature values on the thermal image) and using both canthi regions provides an additional degree of error rejection. It was demonstrated that the calibration blackbody is sufficient to correct the thermal camera's drift and provide a robust reference point for accurate temperature measurement. The statistical tests performed with a sizeable dataset further confirm the correct sensor fusion and operation of the system – the measured value distributions follow the expected distributions, correlations indicate the capability for drift correction, and the system is efficient at detecting fever.

The calibration is only necessary to perform once, so it's not complicated to integrate the system's sensors into a standalone arrangement or device. The capability of providing the user with head placement and orientation hints in real-time thanks to the knowledge of the head's 3D pose is a very valuable feature of a field-deployed system such as this.

Taking a single measurement requires less than 200 milliseconds. It could be improved significantly by using a PC with a faster processor and dedicated graphics processing unit delegated to the portion of processing involving neural network inference. The system is built using off-the-shelf components and provides freedom to use other hardware, as long as their specification is at least on par with the ones presented. The system can be easily integrated with other modules – adding the capability for performing face recognition would enable tracking of personal temperature changes over time.

As emphasised in [37], most thermography-based personal temperature screening systems, both commercial and scientific, are also not compliant or only partially compliant with the ISO guidelines. For example, the solution described in [59] performs only triage-like classification into febrile and non-febrile subjects based on complete face thermal image analysis performed using a neural network. The system presented in [11] performs the temperature measurements in ISO-compliant regions, but since the canthi region detection is performed using feature-deficient thermal images, the location is not established as accurately as in our system (few millimetres vs sub-millimetre). The system presented in [9] is also built to adhere to the ISO requirements. The method uses RGB images for landmark detection. Since the landmarks are transferred to the thermal image using 2D to 2D registration techniques (deformable registration), the landmark transfer procedure is not as accurate as the one presented in this paper. This results in lower Pearson correlation between the reference and the measured temperature and lower fever detection accuracy as measured by the AUC. The method presented in [60] matches landmark (photometric ORB features) between the thermal and RGB images to facilitate canthi region detection and registration between the two modalities, but does not implement any other measurement steps and the registration accuracy is not analysed. We believe that the addition of the depth perception functionality significantly enhances the accuracy of RGB-thermal registration. The system achieves performance gains from detecting the landmarks in RGB modality and treating the face as a 3D object during registration. While it comes at the cost of extending the measurement setup, consumer-grade RGB-D cameras such as the one included in the presented system are available as inexpensive, integrated devices, so the benefits outweigh the additional costs.

V. CONCLUSION

In this paper, a method of efficient identification of ROIs (regions of interest) on a thermal image was presented, allowing for a measure of the human temperature at specific points. Different approaches to keypoints' detection and transformation between the RGB and thermal images were explored, enabling sensor fusion for the task. The best result was achieved by detecting the keypoints on the RGB image and transforming them to the thermal image coordinates, while the distance between the camera and a human face was measured by the depth sensor. Moreover, our approach is simple, requiring just an easy setup of off-the-shelf components. Moreover, it requires only one-time calibration, enabling the construction of easy to use, ISO-compliant, testing station. To validate the research, a dataset of 210 persons was created. It contains 630 instances – for each one: visual, depth and thermal image.

A distinctive aspect of the presented solution, in comparison to previous work, is complying with the requirements

of the standard of a medical device (*ISO/IEC 80601-2-59*). According to it, we employ the mix of both deep learning and classical methods to estimate ellipsoid regions near canthi. Moreover, in the deep learning part, we employ ready-to-use the Mediapipe framework, which does not require additional finetuning to fulfil our requirements. Our algorithm achieves an error of 2.53 ± 1.40 mm in the visual keypoints detection part and generates an error of 0.4 ± 0.39 mm during visual-to-thermal conversions. The summed error of ROIs detection on the RGB image and error of transformation to the thermal image is 2.93 ± 1.79 mm (mean $\pm$ std). The most contributing part of this error is the detection of ROIs, as they were compared with manual annotations. Nevertheless, the ROI location accuracy satisfies the requirements enclosed in the norm.

Additionally, we performed a statistical analysis of the dataset and measurements. The descriptive statistic of the dataset showed that the results are distributed around the mean value and the temperature difference between the left and right ROIs has an excepted value of 0.31°C. Furthermore, we observed a strong correlation between measurements and the reference temperature. We consider, using a more precise thermal camera to achieve even higher accuracy of the system. Nevertheless, we showed, it is technically possible to develop a non-invasive temperature measurement system following the standard for medical devices. Overall, we hope the results will pave the way for fast, reliable, non-contact personal temperature measurement and control applications.

FUTURE WORK

Our current research lays a strong foundation, but further exploration is necessary to solidify the system's capabilities. Fine-tuning pre-trained models on our specific dataset holds promise for enhancing accuracy in keypoint detection and ROI estimation. Additionally, incorporating verification techniques in more places within the system and expanding the study group for more population diversity validation will be considered. Furthermore, collecting user comments during use will provide invaluable insights for improving the user experience (UX) as we move towards real-world applications.

ETHICS APPROVAL

All individuals whose images extracted from the system in the form of thermal and RGB images were used to develop the algorithm, learn it and test it, have signed a consent for their image to be used for research purposes. This document has been prepared in accordance with GDPR regulations, and each person has the right to withdraw their image from the dataset created. The full dataset is not publicly available. The images used in the article represent the faces of the authors of the article, who fully consent to their publication.

REFERENCES

[1] R. E. Baker, A. S. Mahmud, I. F. Miller, M. Rajeev, F. Rasambainarivo, B. L. Rice, S. Takahashi, A. J. Tatem, C. E. Wagner, L.-F. Wang, A. Wesolowski, and C. J. E. Metcalf, "Infectious disease in an era of global change," *Nature Rev. Microbiol.*, vol. 20, pp. 193–205, Apr. 2022.

[2] G. B. Dell'Isola, E. Cosentini, L. Canale, G. Ficco, and M. Dell'Isola, "Noncontact body temperature measurement: Uncertainty evaluation and screening decision rule to prevent the spread of COVID-19," *Sensors*, vol. 21, no. 2, p. 346, Jan. 2021.

[3] A. S. El-Radhi, "Fever in common infectious diseases," *Clin. Manual Fever Children*, pp. 85–140, Jan. 2019.

[4] S. Adams, T. Bucknall, and A. Kouzani, "An initial study on the agreement of body temperatures measured by infrared cameras and oral thermometry," *Sci. Rep.*, vol. 11, no. 1, p. 11901, Jun. 2021.

[5] T. Falcone, F. Cordella, V. Molinaro, L. Zollo, and S. Del Ferraro, "Real-time human core temperature estimation methods and their application in the occupational field: A systematic review," *Measurement*, vol. 183, Oct. 2021, Art. no. 109776.

[6] P. Dolibog, B. Pietrzyk, K. Kierszniok, and K. Pawlicki, "Comparative analysis of human body temperatures measured with noncontact and contact thermometers," *Healthcare*, vol. 10, no. 2, p. 331, Feb. 2022.

[7] P. Chan, G. Wong, T. Dinh Nguyen, T. Nguyen, J. McNeil, and I. Hopper, "Estimation of respiratory rate using infrared video in an inpatient population: An observational study," *J. Clin. Monitor. Comput.*, vol. 34, no. 6, pp. 1275–1284, Dec. 2020.

[8] R. Vardasca, C. Magalhaes, D. Marques, J. Moreira, R. Frade, A. Seixas, J. Mendes, and F. Ring, "Bilateral assessment of body core temperature through axillar, tympanic and inner canthi thermometers in a young population," *Physiological Meas.*, vol. 40, no. 9, Sep. 2019, Art. no. 094001.

[9] Y. Zhou, P. Ghassemi, M. Chen, D. McBride, J. P. Casamento, T. J. Pfefer, and Q. Wang, "Clinical evaluation of fever-screening thermography: Impact of consensus guidelines and facial measurement location," *J. Biomed. Opt.*, vol. 25, Sep. 2020, Art. no. 097002.

[10] J. Foster, A. B. Lloyd, and G. Havenith, "Non-contact infrared assessment of human body temperature: The journal temperature toolbox," *Temperature*, vol. 8, no. 4, pp. 306–319, Oct. 2021.

[11] Z. M. Lazri, Q. Zhu, M. Chen, M. Wu, and Q. Wang, "Detecting essential landmarks directly in thermal images for remote body temperature and respiratory rate measurement with a two-phase system," *IEEE Access*, vol. 10, pp. 39080–39094, 2022.

[12] Q. Wang, Y. Zhou, P. Ghassemi, D. McBride, J. P. Casamento, and T. J. Pfefer, "Infrared thermography for measuring elevated body temperature: Clinical accuracy, calibration, and evaluation," *Sensors*, vol. 22, no. 1, p. 215, Dec. 2021.

[13] C. Ferrari, L. Berlincioni, M. Bertini, and A. Del Bimbo, "Inner eye canthus localization for human body temperature screening," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 8833–8840.

[14] J. Güttler, C. Georgoulas, and T. Bock, "Contactless fever measurement based on thermal imagery analysis," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Apr. 2016, pp. 1–6.

[15] *International Electrotechnical Commission*, Standard 80601-2-59:2017, IEC, 2017.

[16] *Medical Electrical Equipment—Deployment, Implementation and Operational Guidelines for Identifying Febrile Humans Using a Screening Thermograph*, Standard ISO/TR 13154:2017, IEC, 2017.

[17] M. K. Bhowmik, K. Saha, S. Majumder, G. Majumder, A. Saha, A. N. Sarma, D. Bhattacharjee, D. K. Basu, and M. Nasipuri, "Thermal infrared face recognition—A biometric identification technique for robust security system," *Rev. Refinements New Ideas Face Recognit.*, vol. 7, pp. 113–138, Jul. 2011.

[18] W. K. Wong, P. N. Tan, C. K. Loo, and W. S. Lim, "An effective surveillance system using thermal camera," in *Proc. Int. Conf. Signal Acquisition Process.*, Apr. 2009, pp. 13–17.

[19] A. Gautam and S. Singh, "Neural style transfer combined with EfficientDet for thermal surveillance," *Vis. Comput.*, vol. 38, no. 12, pp. 4111–4127, Dec. 2022.

[20] R. Balcerak, M. Diakides, N. Diakides, J. Lupo, and J. Paul, "Advances in medical infrared imaging," in *Medical Infrared Imaging*. Boca Raton, FL, USA: CRC Press, 2007, pp. 19–32.

[21] R. Roslidar, K. Saddami, F. Arnia, M. Syukri, and K. Munadi, "A study of fine-tuning CNN models based on thermal imaging for breast cancer classification," in *Proc. IEEE Int. Conf. Cybern. Comput. Intell. (CyberneticsCom)*, Aug. 2019, pp. 77–81.

[22] D. Karp, "Detecting small and cryptic animals by combining thermography and a wildlife detection dog," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, Mar. 2020.

[23] S. Savazzi, V. Rampa, L. Costa, S. Kianoush, and D. Tolochenko, "Processing of body-induced thermal signatures for physical distancing and temperature screening," *IEEE Sensors J.*, vol. 21, no. 13, pp. 14168–14179, Jul. 2021.
[24] G. R. Hernández, M. A. Navarro, N. Ortega-Sánchez, D. Oliva, and M. Pérez-Cisneros, "Failure detection on electronic systems using thermal images and metaheuristic algorithms," *IEEE Latin Amer. Trans.*, vol. 18, no. 8, pp. 1371–1380, Aug. 2020.
[25] M. Piechocki, T. Pajchrowski, M. Kraft, M. Wolkiewicz, and P. Ewert, "Unraveling induction motor state through thermal imaging and edge processing: A step towards explainable fault diagnosis," *Eksploatacja I Niezawodność—Maintenance Rel.*, vol. 25, no. 3, pp. 1–16, Jul. 2023.
[26] A. Glowacz and Z. Glowacz, "Diagnosis of the three-phase induction motor using thermal imaging," *Infr. Phys. Technol.*, vol. 81, pp. 7–16, Mar. 2017.
[27] Å. F. Skomedal, B. L. Aarseth, H. Haug, J. Selj, and E. S. Marstein, "How much power is lost in a hot-spot? A case study quantifying the effect of thermal anomalies in two utility scale PV power plants," *Sol. Energy*, vol. 211, pp. 1255–1262, Nov. 2020.
[28] S. Lee, K. E. An, B. D. Jeon, K. Y. Cho, S. J. Lee, and D. Seo, "Detecting faulty solar panels based on thermal image processing," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2018, pp. 1–2.
[29] Z. Cai, J. Cui, H. Yuan, and M. Cheng, "Application and research progress of infrared thermography in temperature measurement of livestock and poultry animals: A review," *Comput. Electron. Agricult.*, vol. 205, Feb. 2023, Art. no. 107586.
[30] Y. Wang, X. Kang, M. Chu, and G. Liu, "Deep learning-based automatic dairy cow ocular surface temperature detection from thermal images," *Comput. Electron. Agricult.*, vol. 202, Nov. 2022, Art. no. 107429.
[31] H. Cuthbertson, G. Tarr, and L. A. González, "Methodology for data processing and analysis techniques of infrared video thermography used to measure cattle temperature in real time," *Comput. Electron. Agricult.*, vol. 167, Dec. 2019, Art. no. 105019.
[32] J. Ranjan and J. Scott, "ThermalSense: Determining dynamic thermal comfort preferences using thermographic imaging," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Sep. 2016, pp. 1212–1222.
[33] A. C. Cosma and R. Simha, "Using the contrast within a single face heat map to assess personal thermal comfort," *Building Environ.*, vol. 160, Aug. 2019, Art. no. 106163.
[34] A. Aryal and B. Becerik-Gerber, "A comparative study of predicting individual thermal sensation and satisfaction using wrist-worn temperature sensor, thermal camera and ambient temperature sensor," *Building Environ.*, vol. 160, Aug. 2019, Art. no. 106223.
[35] S. Rayanasukha, A. Somboonkaew, S. Sumriddetchkajorn, K. Chaitavon, S. Chanhorm, B. Saekow, and S. Porntheeraphat, "Self-compensation for the influence of working distance and ambient temperature on thermal imaging-based temperature measurement," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–6, 2021.
[36] Y. Li, D. Pan, Z. Jiang, H. Yu, and W. Gui, "Adaptive compensation method for the infrared temperature measurement error based on 3-D thermal imaging," *IEEE Sensors J.*, vol. 23, no. 10, pp. 10525–10537, May 2023.
[37] K. J. Howell, J. B. Mercer, and R. E. Smith, "Infrared thermography for mass fever screening: Repeating the mistakes of the past," *J. Med. Virol*, vol. 30, pp. 5–6, Jan. 2020.
[38] D. E. King, "Dlib-ml: A machine learning toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Jul. 2009.
[39] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.
[40] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1867–1874.
[41] J. Alabort-i-Medina, E. Antonakos, J. Booth, P. Snape, and S. Zafeiriou, "Menpo: A comprehensive platform for parametric image alignment and visual deformable models," in *Proc. 22nd ACM Int. Conf. Multimedia*, NY, NY, USA, Nov. 2014, pp. 679–682.
[42] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
[43] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.
[44] M. Wilms, H. Handels, and J. Ehrhardt, "Representative patch-based active appearance models generated from small training populations," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017*. Cham, Switzerland: Springer, 2017, pp. 152–160.
[45] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. Guang Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, "MediaPipe: A framework for building perception pipelines," 2019, *arXiv:1906.08172*.
[46] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "BlazeFace: Sub-millisecond neural face detection on mobile GPUs," 2019, *arXiv:1907.05047*.
[47] Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann, "Real-time facial surface geometry from monocular video on mobile GPUs," 2019, *arXiv:1907.06724*.
[48] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
[49] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[50] E. Marchand, H. Uchiyama, and F. Spindler, "Pose estimation for augmented reality: A hands-on survey," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 12, pp. 2633–2651, Dec. 2016.
[51] G. Bradski, "The OpenCV library," *Dr. Dobb's J. Softw. Tools J. Softw. Tools Prof. Programmer*, vol. 25, no. 11, pp. 120–123, 2000.
[52] J. Strohmayer and M. Kampel, "A compact tri-modal camera unit for RGBDT vision," in *Proc. 5th Int. Conf. Mach. Vis. Appl. (ICMVA)*, NY, NY, USA, Feb. 2022, pp. 34–42.
[53] R. Usamentiaga, D. F. Garcia, C. Ibarra-Castanedo, and X. Maldague, "Highly accurate geometric calibration for infrared cameras using inexpensive calibration targets," *Measurement*, vol. 112, pp. 105–116, Dec. 2017.
[54] P. Saponaro, S. Sorensen, S. Rhein, and C. Kambhamettu, "Improving calibration of thermal stereo cameras using heated calibration board," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4718–4722.
[55] O. Riou, S. Berrebi, and P. Bremond, "Nonuniformity correction and thermal drift compensation of thermal infrared camera," in *SPIE Proc.*, Apr. 2004, pp. 294–302.
[56] R. Olbrycht and B. Wiecek, "New approach to thermal drift correction in microbolometer thermal cameras," *Quant. Infr. Thermography J.*, vol. 12, no. 2, pp. 184–195, Jul. 2015.
[57] J. Law, D. E. Morris, C. Izzi-Engbeaya, V. Salem, C. Coello, L. Robinson, M. Jayasinghe, R. Scott, R. Gunn, E. Rabiner, T. Tan, W. S. Dhillo, S. Bloom, H. Budge, and M. E. Symonds, "Thermal imaging is a noninvasive alternative to PET/CT for measurement of Brown adipose tissue activity in humans," *J. Nucl. Med.*, vol. 59, no. 3, pp. 516–522, Mar. 2018.
[58] M. Soroko, K. Howell, K. Dudek, R. Henklewski, and P. Zielińska, "The influence of breed, age, gender, training level and ambient temperature on forelimb and back temperature in racehorses," *Animal Sci. J.*, vol. 88, no. 2, pp. 347–355, Feb. 2017.
[59] M. L. Brioschi, C. D. Neto, M. de Toledo, M. A. G. M. Moreira, N. Civiero, E. B. Neves, J. V. C. Vargas, and M. J. Teixeira, "Non-fever COVID-19 detection by infrared imaging," in *Proc. Artif. Intell. Over Infr. Images Med. Appl. Med. Image Assist. Biomarker Discovery, First*, 2022, pp. 57–72.
[60] J. Mathew and S. N. George, "Cross spectral image registration for fever detection," in *Proc. IEEE Int. Conf. Signal Process., Informat., Commun. Energy Syst. (SPICES)*, vol. 1, Mar. 2022, pp. 118–122.

BARTOSZ PTAK received the Bachelor of Engineering degree in computer science and the master's degree (Hons.) in the automatic control and robotics field, in July 2021. He is currently pursuing the Ph.D. degree in the automation, electronic, electrical engineering and space technologies discipline with Poznań University of Technology (PUT).

He is also a Computer Vision Engineer with the Institute of Robotics and Machine Intelligence. His research interests include computer vision processing, embedded AI devices, and machine learning aspects for the real-world application of artificial intelligence in everyday life.

PRZEMYSŁAW ASZKOWSKI received the master's degree in automation and robotics from Poznań University of Technology (PUT), in 2018, where he is currently pursuing the Ph.D. degree in the automation, electronic and electrical engineering discipline. He has more than four years of professional experience with commercial research and development projects, including digital signal processing and laboratory equipment control. Since 2021, he has been a part of the research team with the Institute of Robotics and Machine Intelligence Vision Laboratory. His research interests include computer vision with machine learning, with an emphasis on multispectral imaging in environmental and agricultural applications.

JOANNA WEISSENBERG received the M.Sc. degree in mathematics from Casimir the Great University, Poland, in 2007, and the Ph.D. degree in telecommunication from Poznań University of Technology, Poland, in 2015. The main area of her professional activity is Markovian analysis of multiservice networks. Her research interests include the application of stochastic processes theory in telecommunication systems and the analysis of queueing delay models in nervous systems.

MAREK KRAFT (Member, IEEE) received the degree in the field of automation and robotics, in 2013.

He is currently an Assistant Professor and a Researcher with the Faculty of Control, Robotics and Electrical Engineering, Poznań University of Technology. For over ten years, he has been working in the field of image processing, machine learning, and robotics, and in engineering practice, he has also been involved in designing embedded and measurement systems. He is the author of over 40 scientific publications and a member of the international organizations ACM. His area of competence covers the processing of images and video streams, with particular emphasis on machine learning methods: recognition, detection, and tracking of objects in images and video, binary and semantic segmentation, and medical image processing. In addition, his proficiency includes implementation of the above-mentioned methods in embedded architectures and their practical use in vision-based navigation, medical diagnostics, autonomous vehicles, safety, and monitoring applications.

MICHAŁ WEISSENBERG received the M.Sc. degree in electronics and telecommunications from the Faculty of Electronics and Telecommunications, Poznań University of Technology, Poland, in 2017, where he is currently pursuing the Ph.D. degree with the Faculty of Computing and Telecommunications, Institute of Communications and Computer Networks. His research interests include the analytical modeling of ICT systems and cybersecurity.