Metrol. Meas. Syst., Vol. 27 (2020) No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

STUDY ON FACIAL THERMAL REACTIONS FOR PSYCHO-PHYSICAL STIMULI

Jarosław Panasiuk[1)], Piotr Prusaczyk[1)], Artur Grudzień[2)], Marcin Kowalski[2)]

1) *Military University of Technology, Faculty of Mechatronics and Aerospace, gen. Sylwestra Kaliskiego 2, 00-908 Warszawa, Poland* (*jaroslaw.panasiuk@wat.edu.pl, piotr.prusaczyk@wat.edu.pl*)
2) *Military University of Technology, Institute of Optoelectronics, gen. Sylwestra Kaliskiego 2, 00-908 Warszawa, Poland* (*artur.grudzien@wat.edu.pl,* ✉ *marcin.kowalski@wat.edu.pl,* +48 261 839 353)

Abstract

This paper presents a study on the influence of psychophysical stimuli on facial thermal emissions. Two independent groups of stimuli are proposed to investigate facial changes resulting from human stress and physical exhaustion. One pertains to physical effort while the other is linked to stress invoked by solving a simple written test. Subjects' face reactions were measured through collecting and analysing long-wavelength infrared images. A methodology for numerical processing of images is proposed. Results of numerical analysis with respect to different facial regions of interest are provided. An automatic deep learning based algorithm to classify specific thermal face patterns is proposed. The algorithm consists of detection of regions of interests as well as numerical analysis of thermal energy emissions of facial parts. The results of presented experiments allowed the authors to associate emission changes in specific facial regions with psychophysical stimulations of the person being examined. This work proves high usability of thermal imaging to capture changes of heat distribution of face as reactions for external stimuli.

Keywords: facial biometrics, thermal face images, psychophysical stressors.

1. Introduction

Facial thermal heat emission strongly depends on the physical and psychical state of a human. While observing the image of the human face via a thermal infrared camera one can observe the influence of external conditions (cold or hot environment), stress, physical effort as well as changes of thermal emissions in time. A great challenge in the field of human state recognition is to develop a set of representative stimuli which would provide the repeatability of experiments in a larger group of subjects. Since subjects may intentionally simulate some reactions, it is crucial to develop stimuli that will be independent of human intentions. The ability to interpret emotional state related to stress and emotional matters may prove beneficial to diagnoses of various illnesses.

It is widely known in psychology that a stressor is a stress factor. We can distinguish between physical, psychological, sociological and biological stressors [1]. Stressors can be classified by

their cause and duration. When classifying stressors with respect to the source we can distinguish external stressors appearing from the body's external environment, most often resulting in a fight-or-flight response. Internal stressors are stimuli coming from the inside of the body, like traumatic memories, fear of diagnostic tests, *etc*. Virtual stressors (their source being the virtual world – TV/the Internet) present in the scenes of violence and horror affect the human psyche with almost the same strength as in the real world. Then, with regard to the duration of stressors, the so-called one-time events are internal or external stimuli that inevitably lead to anxiety or fear. The remaining, cyclical stressors, correspond to stress reactions appearing periodically, caused by traumatic memories related to a place or a specific time. The latter group are called long-term stressors – these are situations when a stimulus acts on the body constantly, *e.g*. family problems or a chronic illness.

According to [2, 3] psychological stressors cause biological reactions in the body, *e.g*. increasing the heart rate. The human body reaction to strong stressors such as an interrogation, an interview or another mental load task [4, 5] can lead to temperature changes based on the blood veins map all over the body. Each organism has its upper and lower efficiency threshold and reacts with stress to exceeding these thresholds. Stress tolerance also changes as a result of a deterioration of health or a mental condition [1].

Physical stressors can be divided into several groups. Climatic stressors depend on temperature combined with humidity (both of these factors work together). Temperature changes cause thermal stress [6] which can even turn into a heat stroke. Another group comprises of sound stressors (such as noises) that reduce concentration, hinder communication and can also damage hearing. The body cannot adapt to constant noise and this can lead to development of hypertension or a gastric or duodenal ulcer [1]. Visual stressors can be triggered by the brightness of light (induced by the so-called glare) Lack of light also causes stress. Flashing light of a frequency below 10 Hz is also a strong stressor causing seizures in one in every 5 to 10 thousand people. Another group is related to inertial force - a force that acts on the body when it changes its speed and/or direction. This force causes changes in the body's blood and body fluid distribution, leading to hyperaemia and ischemia in different parts of the body.

Long-wavelength infrared imaging (LWIR; 6–15 μm), also referred to as thermal infrared imaging, takes advantage of mature and efficient technology capable of detecting small differences in apparent temperature of objects. Thermal face imaging is sensitive to changes of emotional, physical and health condition of the subject. Moreover, properties of the face depend on the temperature of the body, environmental conditions and occlusions present on the face such as scarfs, hairs, facial hairs, glasses or any disguise accessories that alter the emitted heat pattern. The image of a human face acquired in thermal infrared presents its unique heat-signature presenting thermal energy distribution and, as such, can be used for facial recognition [1]. An analysis of relative temperature distribution on the surface of a face can reveal individual patterns of heat variations and thus can give specific indicators of the subject's current state. The radiation registered is proportionally to relative distribution of apparent temperature of objects. In this paper we present a study on the influence of psychophysical stimulus on facial thermal emissions.

The aim of this work is to report on study of physical and psychological stimuli on facial heat distribution.

The contributions are summarized as follows:

– Long-lasting experiments presenting influence of psychophysical stimuli on facial thermal emissions;
– A methodology for numerical processing of images is proposed;
– Classification of thermal face images based on heat emission analysis;

– Two classification methods based on:
  (a) numerical image analysis;
  (b) a deep learning classifier applied to extracted thermal facial images.

The paper is organized as follows: in section 2, an overview of stressors and related works are discussed. Section 3 presents a methodology of measurement based on classification of facial expressions and thermally significant facial feature points. Description of experiments is provided. Section 4 is dedicated to the description of the results and analysis process. In section 5 the data processing methodology and algorithm are explained. The study is summarized in Section 6.

2. Related works

There is an increasing number of works showing that robotic devices can be trained to use facial thermography to classify human interaction of subjects' feelings. Recent studies have shown the possibility to use machine learning techniques to classify the valence of the expressed feelings (positivity vs. negativity) [7]. Nhan and Tau [8] discovered that, in addition to distinguishing between positive and negative feelings, infrared thermography can be used to differentiate between high and low self-reported excitement. The proposed method achieved over 80 per cent of classification accuracy of elevated excitement versus the baseline and elevated valence versus the baseline. Success levels were much lower, in the 50–60 per cent range, for the classification of high vs. low excitement and high vs. low valence. To classify discrete feelings, Khan [9] used comparative techniques. However, the work reported difficulty in distinguishing between separate feelings, such as happiness and sorrow, due to participation of comparable groups of muscles.

Although the study mentioned above indicates the potential for using infrared thermography in testing emotional theories, its implementation is limited by the nature of machine-based learning. To overcome this, the authors generated algorithms identifying the temperature readings that provide the most distinctive data and generate equations that use appropriate data to predict emotions. However, predictive equations are not included in the published studies as well as specific patterns of facial temperature that can be used for the classification of specific discrete feelings or emotional dimensions. Nevertheless, the fact that effective classification algorithms using facial temperature measurements were created firmly indicates that emotional reactions appear on the face as heat signatures. Khan [9, 10] and Marzec [11, 12] presented methods for automatic determination of significant points on the face called FTFPs (*Facial Thermal Feature Points*), mainly based on the physiology of the human face. However, future studies must determine how temperature variations in particular areas of the face correspond to emotional aspects and/or discrete feelings in order to facilitate testing sociological theories.

According to Khan [9], humans involuntarily react to emotions. Specific muscular-physiological activities are believed to generate some muscle-thermal cues. As a result, a change of a given affective state can cause a change of the body temperature.

Referring to Jones [13], the normal core body temperature of a healthy person, under normal conditions, ranges between 35.5°C in the morning to 37.7°C in the evening. Humans are capable of maintaining a constant body temperature. An increase (or decrease) of body temperature can cause a body malfunction and even a failure of body organs [14].

The normal temperature of the core body helps to preserve the homeostasis. The skin is a vital organ that receives signals from a temperature-regulating element in brain called a hypothalamus to maintain the body's core temperature through the process called thermoregulation. Physiological thermoregulation in humans comprises changes in heat dissipation (sweating) and heat generation (shivering) in response to various internal and external thermal stimuli [8].

As reported by Tanda [15] and Smith [16], during a graded load exercise, where the load on the body progressively increases together with the blood demand for the working organs, the mean skin temperature of the subjects decreases throughout the exercise. The reduction of skin temperature occurs immediately after starting to run, even during the first steps at moderately low velocities, without the appearance of sweat on the skin and thus it is not related to thermal factors such as evaporation due to skin sweat. In the last stage, during the recovery following the running stage hyperthermal spots were observed probably due to peripheral vasodilation enabling a progressive transfer of warmer blood from the body core to the surface [7]. This can be used to determine the RoI (*Region of Interest*) for specific scenarios.

Cruz-Albaran [17] presented a methodology to obtain a non-invasive smart-thermal system based on an analysis of biomedical infrared images that can diagnose basic emotions *i.e.* joy, disgust, anger, fear and sadness. Such emotions show different thermal facial behaviour due to the blood flow through the vessels when an emotion occurs; therefore, in this study the quantification of the change in the facial temperature was registered at different RoIs in order to propose a biomarker as a response to these emotions. However, subjects have different reactions to the same stimulus. Thus, a stage of self-calibration in the system is proposed in order to obtain a correct diagnosis when one of the five basic emotions occurs [17].

Warmelink [18] stated that people that are involved in terrorism or other illegal activities at the time of their actions appear to be accompanied by alertness, anxiety, and even fear. Since the sympathetic nervous system [19] produces symptoms, it cannot be completely controlled. These symptoms can provide useful hints for security staff of critical infrastructures to identify prospective suspects (*e.g.* first-time offenders). Redistribution of the blood flow in the surface blood vessels creates abrupt changes in the local temperature of the skin. This is easily visible on the human face where the skin layer is very thin.

Pavlidis [19] segmented snapshots from the beginning, middle, and final stages of an activity. The face of the subject was segmented out from the rest of the background into 5 areas, namely, periorbital area, nasal area, cheeks (left and right), chin area and neck area. In [19] image measurements have been performed at 5 designated sub-areas on the face for 2 primary activities: a response to a startle stimulus and a mild aerobic exercise. Significant and measurable facial thermal changes appeared. In response to a startle stimulus, the fear was accompanied by an instantaneous (less than 300 milliseconds) increase in the blood flow around the eyes; it being independent of the movement of the face or eye. Other facial changes, *i.e.* the cooling over the cheeks and the heating over the neck area, occurred simultaneously. Moreover, the average temperature of the nasal area remained more or less the same. All the changes returned within 1 minute to the pre-start resting values [19].

In this study we aim to provide a quantitative analysis of effects of external psychophysical stimuli on subject's state through an analysis of facial thermal infrared images. This paper aims to address the possibility of automatic distinguishing between different subject's states with a deep learning algorithm. During this study, a measurement methodology was proposed together with a methodology for image numerical analysis. The outcomes of image analysis were implemented in the form of a deep learning algorithm.

3. Measurement methodology

In this section, we present the methodology of the measurements. Two main experiments were performed to show the general usability of thermal imaging to capture changes of thermal distribution of the face as reactions to external stimuli. The results are arranged in two main groups.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

Thermal image of the human face presents its unique heat-signature that can reveal individual patterns of intensity variations [20]. Thermal infrared imaging does not need illumination since it relies on heat emitted by all bodies and objects within the field of view of the camera. The radiation registered is proportional to relative distribution of apparent temperature of objects. The projection of objects depends on noise equivalent temperature difference and optics, as well as on temperature difference between objects and their emissivity. The acquired image strongly depends on environmental conditions during the acquisition process. In order to capture the shape and structure of the face, the imager needs to distinguish between very small amounts of energy. The ability to capture and quantify the thermal energy depends on the parameters of the camera, in particular, the *noise equivalent temperature difference* (NETD). This parameter directly defines the ability of a camera to detect the shape of an object. Infrared cameras equipped with uncooled micro-bolometer focal plane arrays offer NETD values between 40 and 130 mK, whereas imagers with cooled detection units have an NETD value below 20 mK.

As part of the experimental study, it is proposed that participants will be exposed to two elicitations including physical stimuli and psychological stressors. During the study, the subjects' reactions will be captured with a state-of-the-art thermal infrared camera of the parameters presented in Table 1.

Table 1. Parameters of the FLIR A65 camera.

Parameters	**Values**
Detection unit	Focal plane array (FPA), uncooled VOX microbolometer
Resolution	640 × 512 pixels
Field of View (FOV)	25° × 20° with 25 mm lens
Spectral range	7.5 μm – 13 μm
NETD	50 mK at 300 K
Acquisition frequency	30 Hz

The selection of the stressors in the form of physical and psychological factors resulted from the fact that these factors can simultaneously inform about external stimuli (physical effort or environmental changes) as well as those connected to emotional states, independent of exertion or the environmental factors mentioned. During the study, all the factors were considered following the Cooper experiment. The experiment consisted of two phases, physical and psychological during which the participants were exposed to stimuli including instances of stimulants affecting the sensual domain, *e.g.* loud sound, bright light, flickering light.

This work focused on linking the presence of specific external stimuli with heat emission changes in specific areas of the face. The psychological and physical parts of the experiment had their own criteria based on intensity variations in the function of occurring stressors. In both experiments, images were registered during the presence of the stimuli as well as while the stressor was not influencing the subject directly. During both tests we analysed images to extract differences of pixel intensities in order to provide information for the classification process. The process relied on relative values instead of absolute values. We looked for pixel intensity changes of specific RoIs with respect to the subject being in the neutral state. The measurement was considered successful when a set of images acquired during the experiment was classified correctly.

All the measurements were taken indoors. During the psychological part of the test, a subject sat on a chair in front of the camera at a distance of 1.5 m. The camera was mounted on a tripod at a height of 1.2 m above the floor. After the physical exercises, images were collected of subjects standing still in front of the camera. Ambient temperature during each of the measurement sessions was controlled and recorded using a thermometer and stabilized using an air conditioning system. In order to ensure a uniform background of images, the camera was directed towards the wall covered with cotton fabric. The fabric was used to reduce the non-uniformity of wall as well as eliminate reflections.

3.1. Physical test

The first type of the experiment concerned evaluating the impact of physical factors on facial heat distribution. The physical test, being part of the Cooper test, was performed in the form of a 12-minute run endurance training. The task of each subject was to run the longest distance within 12 minutes. In order to achieve full commitment to the task, the distance was compared to the gender-specific scoreboard.

The Cooper test, designed to determine an athlete's performance, was limited to a run element being the most common factor in everyday life. Conducting the tests in a diversified age group (20, 30, 40) allowed us to check whether the examined face areas recorded on thermographs had common features with the same stimuli.

Thermal photos were taken immediately after the workout and repeated at specified time intervals. Each set of images collected for each subject contained images presenting subjects before the test and during cooling down. Images before the test were acquired for reference.

3.2. Psychological test

The mental stressors considered in the study were evaluated for the same group of subjects performing the same activities comprising solving an arithmetic test, putting words on a sheet of paper in the correct order and finding differences between two images. The duration of this test was limited. During the experiment each subject sat in front of the camera while performing specific actions guided by a person supervising the measurement process.

In order to achieve the stress element in the activities presented, it was decided to use a factor which was simultaneously a stimulus and the consequence of a failure. In the course of designing the experiment various consequences had been considered including physical (pain), mental (rebuke or embarrassment) or financial (financial penalty). A physical penalty is by nature not acceptable for this type of test. Neither is psychological punishment. The application of the financial penalty was accepted as the most acceptable and, at the same time, the most relevant to the binding social order (penalty for traffic offences). Subjects participating in the study agreed to be punished with the financial penalty. All the participants in the study also agreed on the amount of the fine at the level of around 100 USD. The subject to achieve the lowest score in the test was supposed to lose the declared amount in favour of the other participants of the study.

The duration of each task was limited to a time interval between 1 and 3 minutes. The subject was notified by a visual or sound signal to look at the thermal camera in order to acquire a thermal face image. Images were acquired with one minute intervals. As a result, we obtained a dataset of thermal face images. The dataset was then analysed to extract changes of heat distribution of the subjects' faces. For the purpose of comparison, images presenting the subjects in the relaxed state were also acquired.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

4. Data processing and results

During the experiments of both types, thermal face images as well as environment conditions were recorded. The images were then considered as material for processing and analysis. The thermal infrared camera used during the experiment acquired thermal data images with a resolution of 640×512 pixels, each with the depth of 14 bits. Images were analysed with respect to selected RoIs. A sample imaging matrix showing a sample RoI is presented in Fig. 1.

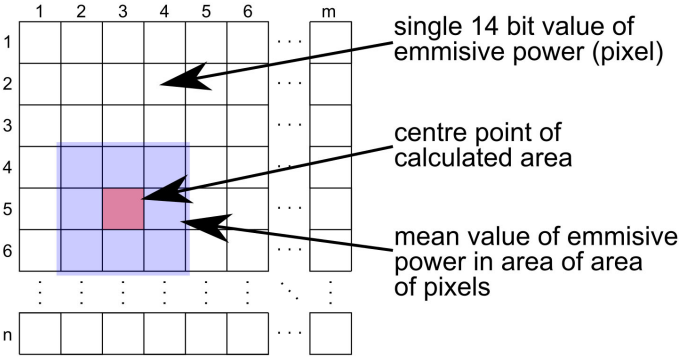

Fig. 1. Array of pixels.

In order to measure the value of emissive power in specific areas of the face it is required to determine the location of the face in the image. Face analysis consists of several stages starting from pre-processing, RoI detection and analysis of heat distribution. Since the RoIs for analysis are located at the entire facial region, the RoI detection algorithm has been trained to detect entire head.

The head detection was performed using the Faster R-CNN [20] algorithm trained with thermal infrared and visible face images. The selection of the Faster R-CNN method is the result of an extensive study on thermal face detection algorithms. We had studied three algorithms including machine learning Viola-Jones [21], and two deep learning methods YOLOv2 [22] and Faster-RCNN. For the purpose of this work, the best performing algorithm was selected for face detection.

The Faster R-CNN algorithm is based on the idea of *region proposal network* (RPN). The RPN outputs the objectness score for many proposed boxes which indicates whether the selected part of an image contains a background or a foreground object. All the boxes are examined by a classifier and a regressor to check the occurrence of objects. The Faster R-CNN is composed of two networks, a RPN for generating region proposals and a network using these proposals to detect objects. This algorithm demonstrates an impressive face/head detection performance.

We selected specific face areas for the numerical analysis including forehead, eyes, nose and cheeks as unique markers for determining one's physical and psychological state. The method to determine specific RoI's coordinates is based on [17]. It applies the ratio distribution factor of the elements on the face based on the human anatomical parameters calculated for large group of subjects. According to Fig. 3 the centre of nose RoI can be estimated by using the following formula:

$$\begin{bmatrix} x_{nose} \\ y_{nose} \end{bmatrix} = \begin{bmatrix} x_{face} + 0.5D \\ y_{face} + 0.55d \end{bmatrix}, \tag{1}$$

where D is size of a face along its x axis and d is a face size along its y axis.

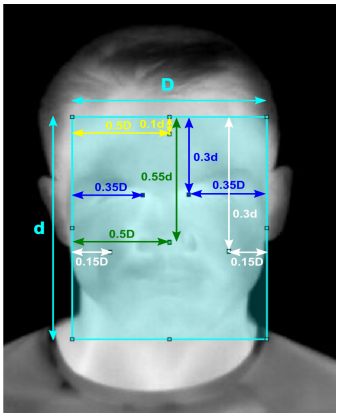

Fig. 2. Estimation method of the centres of specific ROIs.

The RoIs for the calculations of the mean value of emissive power were specified as a particular ratio of corresponding dimensions of the detected face as presented in Fig. 3. These regions were used for calculating the mean value of normalized pixel intensities.

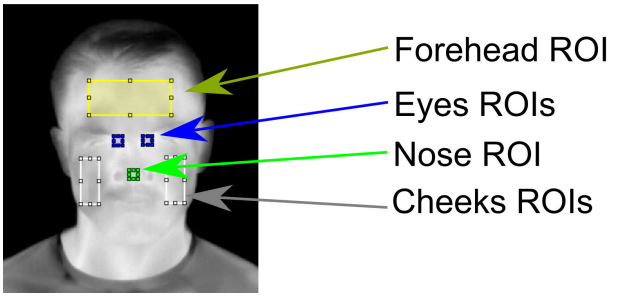

Fig. 3. Regions of interests.

The mean value of normalized intensities for each region of interest were calculated with the following formula:

$$E_{ROI} = \frac{\sum V_{ROI_x ROI_y}}{ROI_x * ROI_y}, \tag{2}$$

where E_{ROI} is mean value of emissivity of a selected RoI, $V_{ROI_x ROI_y}$ are values of emissivity for a specified pixel and ROI_x, ROI_y are width and height of the RoI, respectively. Pixel intensities of each image were normalized before the analysis using min-max normalization.

4.1. Psychological test

The data acquired during the psychological test were divided into 4 groups described in Table 2.

A gallery of sample images presenting a subject during the psychological test is shown in Fig. 4.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/amms.2020.134591

Table 2. Description of subsequent parts of the measurements in the psychological test.

Stage	Task description
Baseline	Data acquired at the beginning of the experiment
Maths task	Data acquired during the mathematical test
Sorting task	Data acquired during the sorting test
Find the differences task	Data acquired during the graphical test of finding the differences between two images

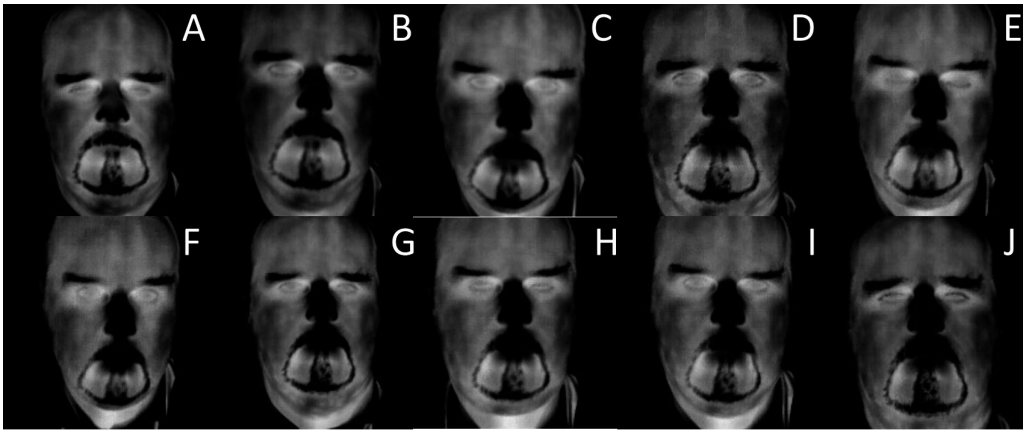

Fig. 4. Gallery of thermal facial images presenting subject during the psychological test; A) before the experiment, B) before the math task, C) during the math task, D) after the math task, E) before the sorting task, F) during the sorting task, G) after the sorting task, H) before the FD task, I) during the FD task, J) after the FD task.

Increasing values of normalized intensities were observed during each test, followed by a decrease after reporting the task as completed. These observations are shown as the effect of stress affecting each participant. Graphs of normalized pixel intensities across the different tasks are presented in Fig. 5.

The analysis of the results showed a certain regularity. After the initial moment of each part of the test (mathematical test, sorting, finding the differences) the normalized intensity rises as shown in Fig. 6.

The change in normalized intensity is directly connected with the change of apparent temperature. However, since we did not use any stabilized radiation source (a blackbody), the measured values of pixel intensities should not be considered real temperatures. Afterwards, the normalized intensities rise to the maximum value. The values of normalized intensity decrease at the end of the task. This trend was repeated for each participant during all the three stages of the experiment. The strongest trend is visible in the case of forehead and eyes and the weakest in the case of nose.

For each facial RoI, we calculated a threshold which corresponds to the mean intensity of a RoI for the subject being in the neutral state. In this step, only images presenting subjects in the neutral state were analysed. Based on these thresholds, we compared the RoI values of subjects during the test. Finally, we calculated the total ratio of facial images correctly classified as neutral or during the test which equals 76.40%.

One can also notice here that different parts of the test had a different influence on the stress level of the subject. In the case of the mathematic test, which in terms of content could be the

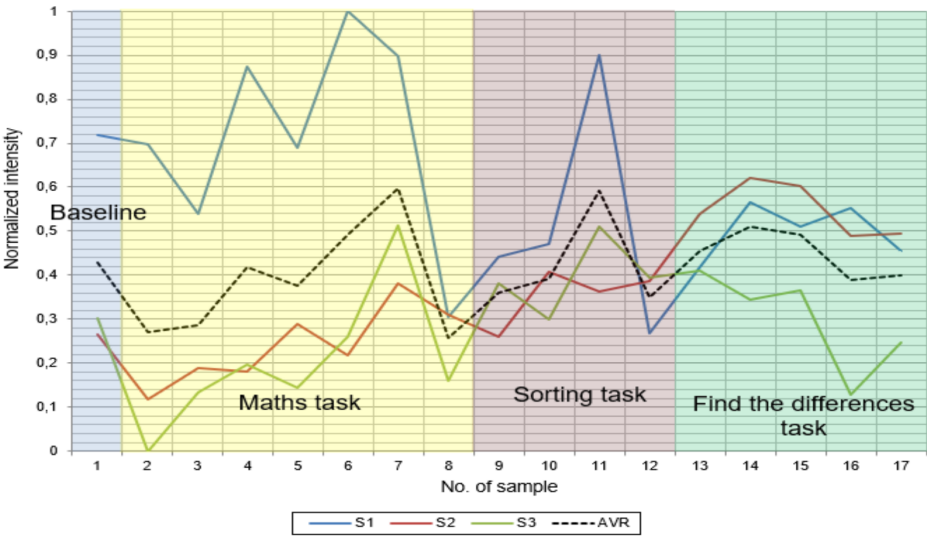

Fig. 5. Graphs of normalized average pixel intensities for the forehead during the psychological test.

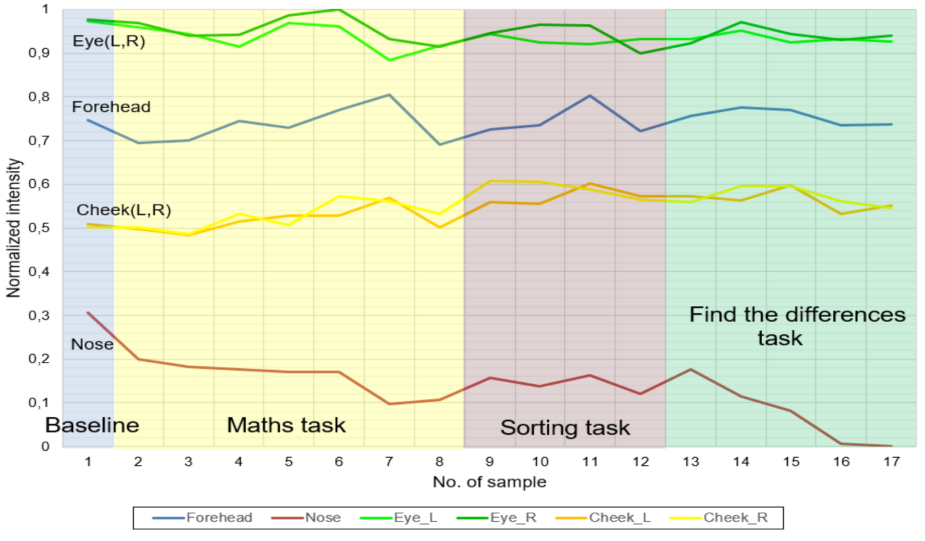

Fig. 6. Graphs of normalized average pixel intensities for facial areas during the psychological test.

most difficult, it is clear that the measurement of normalized pixel intensities at the surface of the forehead implies the highest impact compared with the other parts of the experiment.

4.2. Physical test

The results collected during the second part of the experiment were grouped with regard to the stage of the measurement process. Table 3 provides a description of the measurement process. During each stage 5 images of each subject were collected.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

Table 3. Description of the measurement stages.

Stage number	Part description
Baseline	Beginning of the experiment
Stage 1	Subjects coming in from the cold environment outside
Stage 2	Subjects after a 16-minute run in low temperatures outside
Stage 3	First minute after running
Stage 4	2 minutes after running
Stage 5	3 minutes after running
Stage 6	4 minutes after running
Stage 7	5 minutes after running
Stage 8	10 minutes after running
Stage 9	15 minutes after running

The baseline measurements were acquired before the tests to determine each participants' neutral face heat distribution. At the first stage, the participants were asked to spent 5 minutes outside the building at the temperature of 5°C. After the first stage, subjects were asked to perform the running test according to the Cooper proposition. Each participants' values measured were lower than the baseline (Fig. 5). The running test was performed outside at the temperature of 5°C. The image acquisition started after the physical exercise in the laboratory with air temperature of 22°C and humidity of 35%. A gallery of images acquired before and after the physical test is presented in Fig. 7.

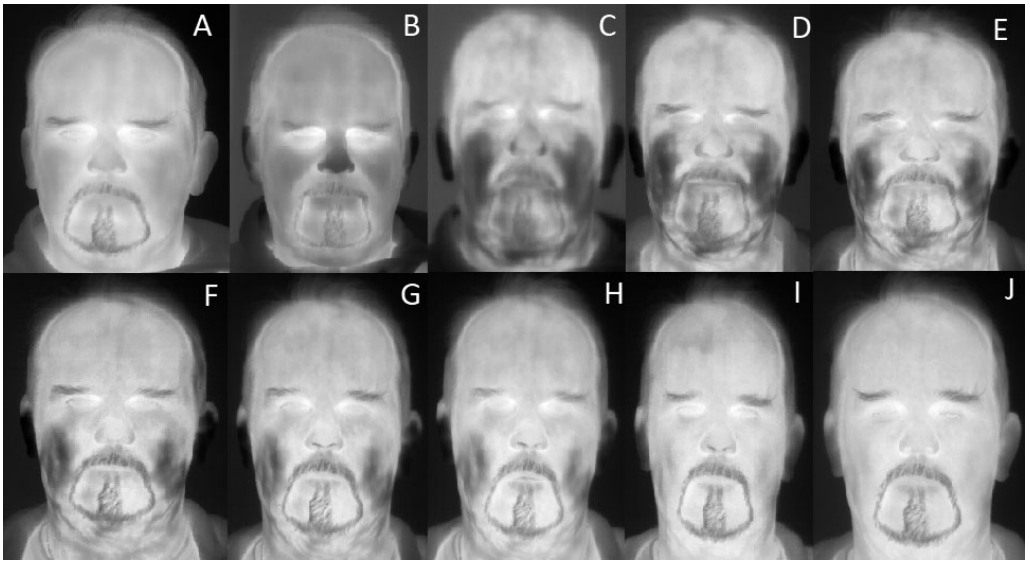

Fig. 7. Gallery of thermal facial images presenting subject during physical experiment; A) before the experiment, B) after 5 minutes in a low temperature environment, C) after 12 minutes of running, D) 1 minute after running, E) 2 minutes after running, F) 3 minutes after running, G) 4 minutes after running, H) 5 minutes after running, I) 10 minutes after running, J) 15 minutes after running.

As presented above, facial thermal emission changed significantly after the exercise. Physical exercise showed to be very strong stimulus since it increased the heart rate and the blood flow. It took the human body several minutes for to regulate the thermal emission and return to the baseline, depending on type of physical stimuli, age and other factors related to the health of a subject. The measurement scheme was applied to collected images to find patterns of heat emission changes as presented in Fig. 8.

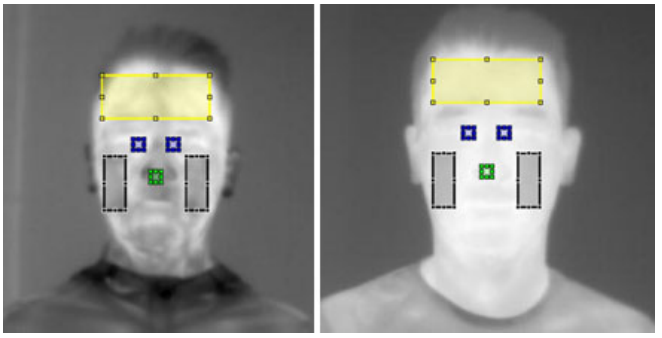

Fig. 8. Sample thermal image, after workout (left), baseline (right).

Graphs of pixel intensities for the forehead and the remaining areas are presented in Fig. 9 and Fig. 10, respectively.

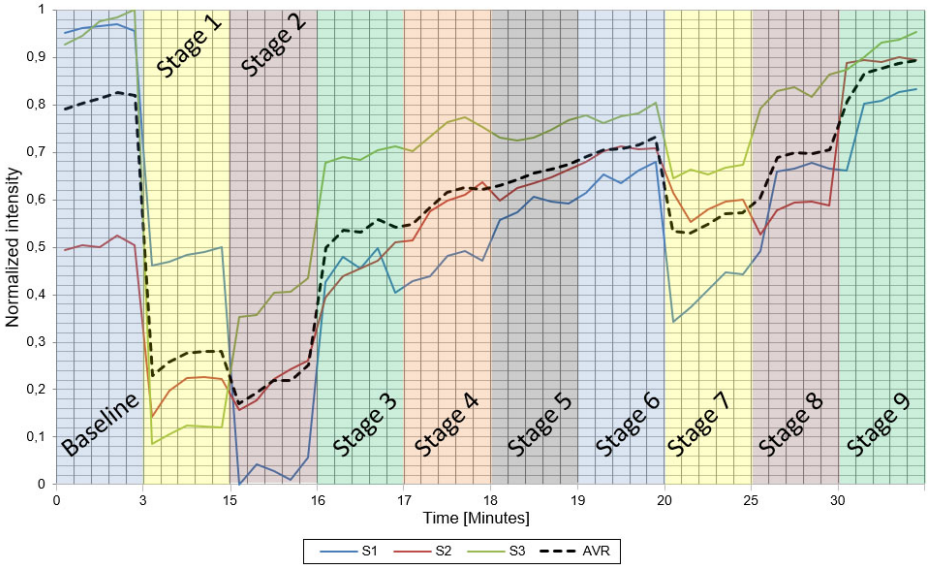

Fig. 9. Graphs of normalized average pixel intensities for the forehead area during the physical test.

An analysis of changes of facial heat distribution registered by the thermal camera after a physical effort brings us to the conclusions below. The coherent nature of the results can be clearly observed in the case of the subjects at different ages (S1 – 32 yr., S2 – 28 yr., S3 – 44 yr.) (Fig. 9). The changes in their characteristics are consistent for all the RoIs as presented in Fig. 10.

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

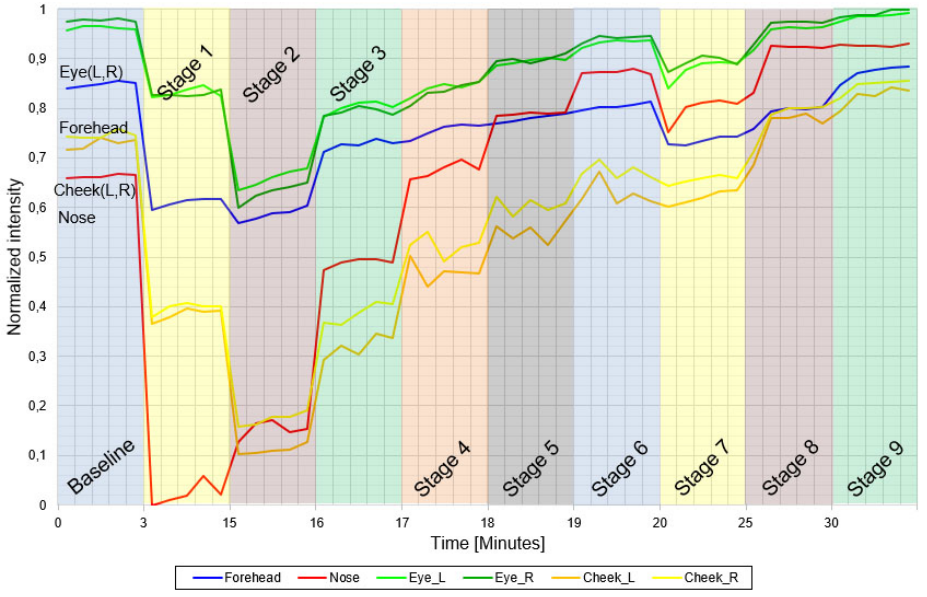

Fig. 10. Graphs of normalized average pixel intensities for facial RoIs during the physical test.

There is a significant difference in heat emissions reflected with different pixel intensities between a person who entered a laboratory from a 5°C environment and a person who came from a similar environment conditions after 12 minutes of physical effort. Based on the baseline pixel intensities and environmental conditions, it is possible to determine whether the examined person came from a low temperature environment or the so-called room temperature. Due to the ability of human body to adapt to different conditions, this information disappears after several minutes. The thermal balancing process depends on a variety of factors, both personal and environmental.

The acquired thermal data of measured RoIs can be interpreted in the context of environmental conditions as well as their changes during the experiment. The observed heat distribution trends allow us to detect the changes that results from stress, entering a room from a low-temperature area or a physical effort. The results of the psychological test reveal that stress caused a heat distribution change related to a local increase in temperature for all the measured RoIs except for the nose. At the end of the psychological test we recorded a noticeable intensity decrease which is particularly visible at the forehead and eyes. This situation repeats in the case of each person examined at each stage of the test.

Based on the physical test, it is possible to see a significant deviation from the initial face heat map (the baseline). The physical exercise caused a strong change in thermal heat emission of the face for each of the subjects. Restoring the baseline heat distribution takes several minutes.

Similarly to the psychological test, we calculated a classification ratio for the physical test. For each facial image of a subject in the neutral state, we calculated a threshold which corresponds to the mean intensity of a RoI. Based on these thresholds, we compared the RoI values of subjects after the physical exercise. As a final step, we calculated the total ratio of facial images correctly classified as neutral or after the exercise. The classification accuracy is 78.1%. We noted that most of the images incorrectly classified were registered at the end of the experiment (after 10 minutes). Therefore, in order to achieve high accuracy, the measurement should be performed immediately after the physical task, no later than 10 minutes.

5. Classification algorithm

The starting point for the algorithm registering the environmental conditions under which the test is performed and used as a reference for results obtained from the specific individual measurement areas.

A block diagram of the classification algorithm is presented in Fig. 11. The algorithm aims to recognize subject's psychophysical state based on registered images as well as external, environmental conditions. Since the state of a subject depends on numerous factors, the facial heat distribution analysis has to be linked with other parameters. The images are categorized with specific metadata including at least air temperature and air humidity. The data registered can be supplemented with the pulse ratio.

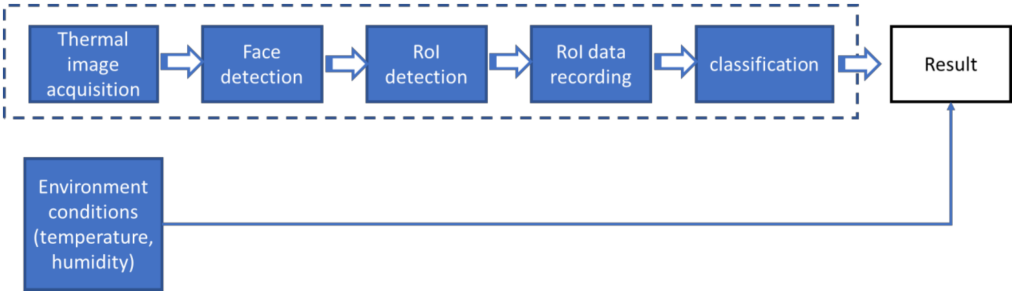

Fig. 11. Block diagram of the classification algorithm.

The algorithm starts with acquisition of metadata and thermal infrared images of a subject's face. The facial images need to be pre-processed before the numerical analysis. Head detection was performed with the Faster R-CNN algorithm achieving a high detection rate of over 95% with zero false detection rate. In order to train the head detection algorithm, 1148 thermal images, each containing a single face were manually annotated. The training dataset is a composition of images brought from two databases (in-house [23] and PROTECT [24]).

Currently, deep learning classification algorithms offer high performance for a variety of tasks. Numerous methods for image classification are known including AlexNet [25], VGG networks [26], residual neural networks, inception networks [27], *etc*.

For automatic classification, a *residual neural network* (ResNet-50) was considered. ResNets is a family of networks proposed in many variants [28]. ResNet-18, ResNet-50 and ResNet-101 were proposed as a solution for training very deep networks with a limited amount of data. These CNNs use the so-called identity shortcut connection or residual connections that skip some of the connections to jump over some layers. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. The network constructs pyramidal cells in the cerebral cortex during data processing. ResNet-50 consists of 50 hidden layers. Since the collected dataset of samples is relatively small, we chose this network as it had been proven less prone to overfitting.

The network analyses the thermal facial images in order to classify the state of a subject as neutral, after a physical effort or stressed. The algorithm aims to provide a good separation rate between different classes of a subject's state. For development of the classification algorithm, we selected 500 images presenting subjects in three states – neutral, after a physical effort and during the psychological test. Extracted face images were grouped into three classes corresponding to two investigated subject's states and the neutral state. In order to classify the subject's state,

Metrol. Meas. Syst., Vol. 27 (2020), No. 3, pp. 399–415
DOI: 10.24425/mms.2020.134591

thermal images need to be acquired at certain time intervals during the experiment. The process should start with acquisition of images of a subject in the neutral state, followed by images captured during the experiment and after it was completed. During our study, all the predictions were based on 6 images including two images in the neutral state, two images taken during the experiment and two images acquired afterwards.

A simple transfer learning scheme was applied training of the classifier. The network model was first trained on the ImageNet dataset in order to provide general representation of various objects and retrained on visible and thermal infrared images. The entire dataset was divided into training, testing and validation parts with the split ratio of 70%, 20%, and 10%, respectively. Our study has revealed that it is possible to determine whether a person performed physical exercise, entered a room of different temperature, or is under stress. Classification performance was calculated in a 5-fold cross-validation scheme with the mean classification performance of 88.21% ± 2.1%.

6. Conclusions

This paper presents a study of influence of psychophysical stimuli on facial thermal emissions. The study consists in applying two groups of stimuli including physical effort and stress invoked by solving a simple written test. The paper presents the methodology used as well as results of numerical analysis of thermal infrared images collected during the experiments. An automatic algorithm to classify specific thermal face patterns is proposed.

This work proves high usability of thermal imaging to capture changes in heat distribution of the face as reactions for external stimuli. The analysis of collected images revealed that in the case of psychological stimuli the dynamics of pixel intensity changes is much higher as compared to the physical effort experiment. The presented methodology can be used to develop an automatic method for classification of state of a subject based on thermal infrared images. However, it should be emphasized that individual reaction depends on a variety of factors including personal and environmental ones. Interpretation of the thermal image of the human face can serve as an additional factor in recognizing a subject's psychophysical conditions.

The conducted numerical analysis, based on specific facial RoIs, achieved the classification ratio of 76.40% and 78.10% for psychological and physical tests, respectively. It has been also revealed that in order to achieve high classification accuracy the measurement should be made immediately after the physical task, no later than 10 minutes after it was completed. The proposed deep-learning based algorithm achieves mean classification performance of 88.21% ± 2.1% in a 5-fold cross-validation scheme.

In order to improve the reliability of the algorithm, a study based on a larger group of subjects would be beneficial. Auto-calibration of the thermal infrared imager as well as reference data of subject being in relaxed state should also improve the method's performance. Nevertheless, this paper shows that both measurement methodology and algorithm are valid and can be used in further investigations. Such extensive research could be used to build a fully functional system allowing to read information about the human psychophysical state from facial thermal infrared images.

References

[1] Terelak, J. (2001). *Psychologia stresu*. Bydgoszcz: BRANTA.

[2] Adelson, R. (2004). Detecting deception. *APA Monitor on Psychology*, 35(7), 70–73.

[3] Gołaszewski, M., Zając, P., Widacki, J. (2015). Thermal Vision as a Method of Detection of Deception: A Review of experiences. *European Polygraph*, 9(1), 5–24.

[4] Powar, N.U., Schneider, T.R., Skipper, J.A., Petkie, D.T., Asari, V.K., Riffle, R.R., Sherwood, M.S. (2017). Thermal Facial Signatures for State Assessment during Deception. *Proc. of the IS&T International Symposium on Electronic Imaging Science and Technology*, USA, 95–104.

[5] Or, C.K.L., Duffy, V.G. (2007). Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement. *Occupational Ergonomics*, 7(2), 83–94.

[6] Ghahramani, A., Castro, G., Becerik-Gerber, B., Yu, X. (2016). Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort. *Building and Environment*, 109, 1–11.

[7] Khan, M.M., Ingleby, M., Ward, R.D. (2006). Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variation. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 1(1), 91–113.

[8] Nhan, B.R., Chau, T. (2010). Classifying affective states using thermal infrared imaging of the human face. *IEEE Transactions on Biomedical Engineering*, 57(4), 979–987.

[9] Khan, M.M. (2008). *Cluster-analytic classification of facial expressions using infrared measurements facial thermal features*. [Doctoral Dissertation, University of Huddersfield]. http://eprints.hud.ac.uk/id/eprint/732/ (accessed on Aug. 2020).

[10] Oz, I.A., Khan, M.M. (2012). Efficacy of biophysiological measurements at FTFPs for facial expression classification: A validation. *Proc. of the IEEE-EMBS International Conference on Biomedical and Health Informatics: Global Grand Challenge of Health Informatics*, Hong Kong, China, 108–111.

[11] Marzec, M., Koprowski, R., Wróbel, Z., Kleszcz, A., Wilczński, S. (2015). Automatic method for detection of characteristic areas in thermal face images. *Multimedia Tools and Applications*, 74(12), 4351–4368.

[12] Marzec, M., Koprowski, R., Wróbel, Z. (2009). Determination of the characteristic face regions on thermograms. *Biomedical Engineering*, 15, 2007–2010.

[13] Jones, B.F. (1998). A re-appraisal of the use of infrared thermal image analysis in medicine. *IEEE Trans. Medical Imaging*. 17(6), 1019–1027.

[14] Jones, B.F., Plassmann, P. (2002). Digital infrared thermal imaging of human skin. *IEEE Engineering in Medicine and Biology Magazine*, 21(6), 41–48.

[15] Tanda, G. (2016). Skin temperature measurements by infrared thermography during running exercise. *Experimental Thermal and Fluid Science*, 71, 103–113.

[16] Smith, C.J., Havenith, G. (2011). Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia. *European Journal of Applied Physiology*, 111(7), 1391–1404.

[17] Cruz-Albarran, I.A., Benitez-Rangel, J.P., Osornio-Rios, R.A., Morales-Hernandez, L.A. (2017). Human emotions detection based on a smart-thermal system of thermographic images. *Infrared Physics & Technology*, 81, 250–261.

[18] Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., Fisher, R.P. (2001). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior*, 35(1), 40–48.

[19] Pavlidis, I., Levine, J., Baukol, P. (2001). Thermal Imaging for Anxiety Detection. *Proc. of the IEEE International Conference on Image Processing (ICIP)*, Thessaloniki, Greece, 2, 315–318.

[20] Ren S., He K., Girshick R., Sun J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149.

[21] Viola, P., Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2), 137–154.

[22] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2006). You Only Look Once: Unified, Real-Time Object Detection. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 779–788.

[23] Kowalski, M., Grudzień, A. (2018). High-resolution thermal face dataset for face and expression recognition. *Metrology and Measurement Systems*, 25(2), 403–415.

[24] Sequeira A.F., *et al.* (2018). PROTECT Multimodal DB: a multimodal biometrics dataset envisaging Border Control. *Prof. of the International Conference of the Biometrics Special Interest Group (BIOSIG)*, Germany.

[25] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*, 25.

[26] Simonyan, K., Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *Proc. of the 3rd International Conference on Learning Representations*, San Diego, USA.

[27] Szegedy, Ch., *et al*. (2015). Going Deeper with Convolutions. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA.

[28] He, K., Zhang, X, Ren, S., Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA.