BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi" din Iaşi
Volumul 69 (73), Numărul 3, 2023
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ
DOI:10.2478/bipie-2023-0013

BUILDING A DATABASE WITH THERMAL IMAGES FOR THE CLASSIFICATION OF EMOTIONAL STATES

BY

MARIUS SORIN PAVEL[*] and DOREL AIORDĂCHIOAIE

„Dunărea de Jos" University of Galați,
Faculty of Automation, Computers, Electrical Engineering
and Electronics, Galați, Romania

Received: June 5, 2024
Accepted for publication: July 31, 2024

Abstract. This research paper aims primarily to highlight the stages regarding the development of a database with facial images in the thermal spectrum, presenting various emotional states (for example, neutral, happy and sad). A series of exploratory results was presented, carried out on five databases with thermal images (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and DBnew-2023_ROI), belonging to the CCETIC research center (CCETIC, 2023). The database development process includes both the thermal image acquisition part and the pre-processing steps (cutting, alignment, selection of areas of interest, removal of annotations/marking elements, etc). And secondly, it aims to present an analysis regarding the general structure of the thermal image processing application, together with a series of image processing methods and techniques, addressed in certain research works. Some experimental results obtained with the help of the five databases are presented, as well as a series of general conclusions regarding the development of a database with facial thermal images.

Keywords: thermal images, database, image processing, pre-processing.

[*]Corresponding author; *e-mail*: sorin.pavel@ugal.ro

1. Introduction

The importance of the subject is given by the fact that the current recognition systems with thermal images present certain shortcomings given the external measurement conditions, such as the temperature of the environment, for example. These systems can benefit from a performance boost by implementing a thermal image processing system, using an infrared thermal camera (e.g., CA 1950), (Arnoux, 2023) and an advanced image processing toolbox (e.g., Matlab), (Nguyen *et al.,* 2013).

Facial expression analysis continues to be an active research topic for scientists. Currently, most research works use images in the visible spectrum and have achieved good results. However, in the absence of illumination, in the dark, rain, fog or in excessive light sources, the result of facial expression analysis in the visible spectrum is not good. Therefore, the use of thermal images helps us to fill in the gaps of visible spectrum images. In addition, skin temperature changes are useful for emotion classification, and facial expression is strictly related to the interlocutor's emotions. We can thus deduce emotions and expressions from skin temperature measurements.

The recognition of the person, the classification or detection of emotional states associated with him, using thermal images has become an increasingly significant part of what it is called "computer vision" (Sebastian, 2015). In practice, human and emotional state detection is a difficult task due to several physical parameters that can affect the final result of image processing. The detection is strictly dependent on the position of the interlocutor, the environment in which the interlocutor is located, the types of camera used, the angle of the camera, the distance from the interlocutor and a possible changing background, and so on. In order to evaluate the effectiveness of facial and emotional recognition systems, an evaluation of the input data, presented in different measuring conditions, is necessary.

Currently, there are quite numerous public databases of facial images acquired in the visible spectrum covering many variations, (Liu *et al*., 2010; Virginia *et al*., 2013; Chen *et al*., 2020). However, the interest in the use of facial images in the thermal spectrum is increasing but relatively low compared to the use of images in the visible spectrum.

The important advantages of thermal imaging compared to visual imaging are:

- Provides a more efficient and safer measurement method (Kölzer *et al*., 1996).
- Obtaining images of interest in low light conditions or even in total darkness (Jędrasiak and Nawrat, 2013).
- Detection of forms of camouflage (Brouant *et al*., 2023).

- Efficiency in avoiding and stopping unwanted incidents (Balaras and Argiriou, 2002).
- Effective detection of intrusions (Kim *et al*., 2017).
- Can provide fast and accurate measurements of hard-to-reach objects such as high-altitude power lines (Jeong *et al*., 2023).
- Can provide sensitive information about the health status of various patients under safe conditions (medical domain) (Rai *et al*., 2017).
- The possibility of detecting faults in electrical/electronic equipment (Jeong *et al*., 2023; Haider *et al*., 2018).
- Can help identify air leaks, document uneven heat dissipation and identify possible insulation irregularities (Baral *et al*., 2018).

The main disadvantages of thermal imaging compared to visual imaging are:

- High initial investment in equipment (thermal imaging cameras) (Byung *et al*., 2015).
- Images are difficult to interpret in certain situations, with objects of interest having irregular temperatures (Vollmer, 2021).
- Accurate temperature measurements are affected by emissivity and reflections, from different surfaces (Zhang *et al*., 2019).
- Impossibility of obtaining thermal images under water (Caimi, 1995).
- Measuring temperatures by thermal imaging is not as accurate as contact methods (Zhang *et al*., 2019).

In Fig. 1 is presented the block structure of the thermal image processing system used in this work. The application blocks addressed in this study are for image acquisition, pre-processing, feature extraction and selection, followed by some classifiers, depending on the imposed objective.

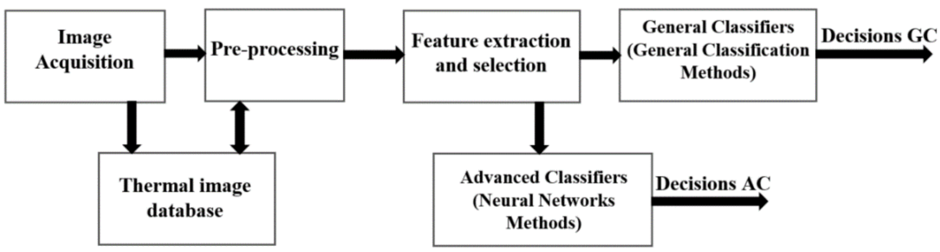

Fig. 1 – Block structure of the application.

The first block “Image Acquisition” from Fig. 1 addresses the thermal image acquisition process for the five databases (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and DBnew-2023_ROI).

The second block “Pre-processing” addresses the pre-processing part of thermal images (cropping, alignment, selection of the area of interest from the thermal images and removing markers and annotations from the thermal images).

The third block “Thermal image database” addresses the five mentioned databases from the research center of CCETIC (CCETIC, 2023).

Block four “Feature extraction and selection” refers the process of selection and feature extraction, for example the selection process of the important features with RF (Random forest) (Pavel *et al*., 2023) and the selection of the regions of interest with DCT (Discrete Cosine Transform) (Aiordachioaie *et al*., 2019) (necessary in the process of detection and diagnosis of people and their associated emotional states).

Block five “General Classifiers (GC)” deals with the general classification methods, for example the K-NN (k-nearest neighbors) algorithm (Moldovanu *et al*., 2023) (classification of people and their associated emotional states).

And block six “Advanced Classifiers (AC)”, deals with neural networks methods, such as feedforward back-propagation neural network (FFBPN) (Pavel *et al*., 2023), and in the future others like CNN (Convolutional Neural Network).

This work is divided into six sections, as follows. In section one has been presented some general information regarding the processing of thermal images for the purposes of detection, diagnosis and recognition of people and their associated emotional states. At the end of the section, some advantages and disadvantages of thermal images were also presented.

In the second section, the acquisition process of facial thermal images for the five mentioned databases from the CCETIC (CCETIC, 2023) research center is presented, and a general description of some databases available with thermal images.

In the third section, the general structure of the facial thermal image processing software application is presented, used in the process of detection, diagnosis and recognition of people and their associated emotional states.

In the fourth section, a series of basic techniques for processing facial thermal images, addressed in certain research works, are presented.

In the fifth section, a series of experimental results from certain research works are presented. The above five mentioned databases (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and BDnew-2023_ROI), (CCETIC, 2023) are used.

The last section includes the conclusions regarding the development of the databases with facial thermal images, for the purposes of event detection, process diagnosis and recognition of people and their associated emotional states.

2. Acquisition and description of the databases

Thermovision and thermography (ThV/ThG) are terms that refer to a modern, high-performance technique that allows the visualization and generation in real time of thermal maps ("thermal images", thermograms) of physical systems (biological or technical) under investigation. To carry out the thermal scanning activity, specialized equipment called thermovision - thermography cameras are used, similar in size and appearance to the well-known photo and video cameras, (Marius and Codreanu, 2019).

The need to create an own database with facial images in the visible and thermal spectrum is due to the fact that, at the moment, there are few public databases with thermal images that meet the technical conditions/requirements, necessary to carry out different researches in the field of person detection and recognition.

When creating the database, it is necessary to define the scenario, to select an appropriate environment that meets the conditions required for the selected research area, taking into account for example lighting and weather conditions. It is also necessary to find volunteers to obtain the desired data set (facial images in the visible and thermal spectrum). Researchers typically create the dataset according to the research area, objectives and tasks they intend to solve. In order to obtain satisfactory results regarding the detection and diagnosis process, it is necessary for the database to contain a sufficiently large number of images (in our case facial thermal images), from at least one hundred candidates.

The acquisition of images from the database should be carried out according to certain rules for each individual emotional state. Each participant/candidate should be positioned so that the background does not influence the desired elements from the image (a background that does not present high values in terms of infrared radiation, and does not present fluctuations).

In order to create a database with thermal images, it is necessary also to take into account a series of parameters regarding the appropriate setting of the thermal imaging camera in order to acquire correct the images. These are related to emissivity, background and transmission. Emissivity is a material property that describes the ability of a material to emit (or radiate) thermal energy in proportion to its surface temperature. It is a value between 0 and 1, or if we consider it as an efficiency factor, between 0 and 100%. When emissivity values are assigned to a material, that value is a comparison to what is known as a "black body" or a perfect thermal emitter. Human skin has an emissivity of 0.98, this is compared to a "black body" at the same temperature and wavelength, (Group, 2020). The background describes/highlights the sources of thermal energy reflected from the surfaces we are trying to inspect and measure. No surface is a perfect emitter, so even highly emissive materials will have some reflections, even if the amount is so small, it is not obvious when viewing the surface via the thermal imaging camera. The background is also what many camera manufacturers call the correction factor, intended to correct for the influence of reflected energy when trying to measure surface temperature. This correction factor matches emissivity in the summation algorithm that thermal cameras use to assign a value to the detected surface radiation, (Group, 2020). We consider setting the transmission only if the thermal imaging camera "looks" through other material, such as an IR window, that is between it and the target (area of interest). There are very few common materials that are transmissive to midwave or longwave IR radiation, so most of the time we're looking at what we call thermally opaque surfaces. We cannot see through these surfaces with IR thermal cameras (Jonathan *et al.*, 2015).

When creating the mentioned databases (DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and DBnew-2023_ROI), the following were considered for the thermal image acquisition process:

- Each participant was positioned in front of a white background at a small distance from the camera;
- The camera was mounted on a tripod and its height was adjusted manually, for a proper centering of the image;
- The distance to the participant was strictly controlled during the acquisition process (approximately one meter);
- A constant lighting condition was maintained, using diffuse lights of the same intensity;
- The resulting images were pre-processed (cutting, alignment, selection of the areas of interest, removal of annotations, etc) to improve the final results regarding the detection and diagnosis process (recognition of emotional states and recognition of persons).

As part of the acquisition process, the thermal camera was set to remove the markers and annotations from the image (cursor and temperature values). These markers/annotations negatively influence the detection process. Another important setting within the thermal camera is given by the correct adjustment of the emissivity value, corresponding to the human body (the value is approximately 0.98).

For the implementation of the five mentioned databases (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and DBnew-2023_ROI) the thermal camera CA 1950 from Chauvin Arnoux Metrix was used, with a resolution of 80 x 80 pixels, thermal sensitivity of 0.08°C at 30°C, 8-14 μm spectral range and automatic focus, (Arnoux, 2023).

Fig. 2 shows the process of acquiring facial thermal images for the five mentioned databases (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise and DBnew-2023_ROI) and the equipment used.

Fig. 2 – The process of acquiring facial thermal images.

Currently, there is a small number of databases with human facial thermal images that are available to the general public. These databases have focused on different aspects of certain studies. Images from the Natural Visible and Infrared Facial Expression Database (NVIE), (Wang *et al.*, 2010) were mainly acquired to investigate the impact of the thermal spectrum on facial expression recognition, so the only variation considered was facial expression. An example from the database is presented in Fig. 3.

Fig. 3 – Images from the NVIE database (Wang *et al.*, 2010).

The Kotani Thermal Facial Emotion database (KTFE), (Hung *et al*., 2014) was proposed for the study of human expression and emotion recognition. This database is among the first to contain natural and spontaneous video images in both the visible and thermal spectrum. This database will allow facial expression and emotion recognition researchers to have more realistic approaches. Several researches have been done with the thermal imaging data available and some positive results have been obtained to support researchers using this database. The results of research in the field offers a promising future in facial thermal image research. An example from the database is presented in Fig. 4.

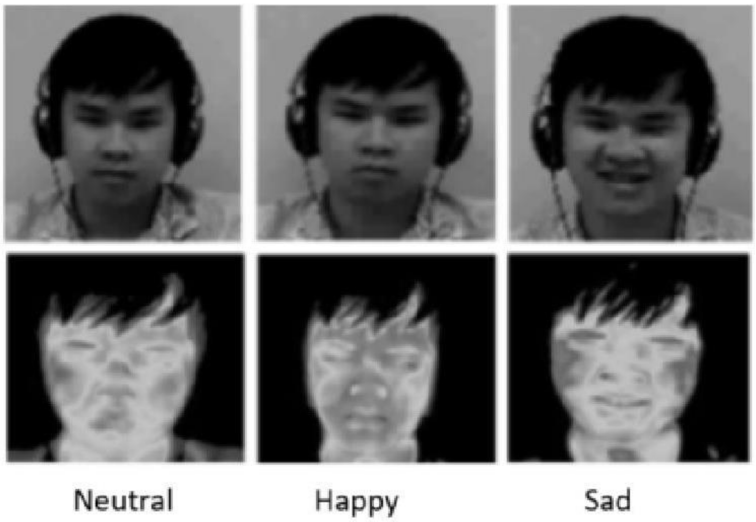

Fig. 4 – Thermal images from the KTFE database, showing three different emotional expressions (Hung *et al*., 2014).

The Tufts database (Karen, 2019) (Tufts University) is very comprehensive and has an extensive dataset of over 10,000 images from more than 15 countries with an age range of 4 to 70 years. The dataset contains six image modes: visible, near-infrared, thermal (contains hand-cropped thermal facial images with various states), computer sketch, recorded video images, and 3D images. An example from the database is presented in Fig. 5.

Interconnected facial recognition modalities are a current topic due to the widespread use of different sensors in everyday applications. The development of facial recognition systems relies heavily on existing databases for evaluation and obtaining training examples for machine learning algorithms. However, there is currently no publicly available facial image database that includes more than two image modes for the same subject.

Fig. 5 – Examples of images in the thermal spectrum from the Tufts database (Karen, 2019).

Five databases with thermal images are available within the Signals and Information Processing Laboratory of CCETIC (Research Center in Electronics, Telecommunications and Information Technology of the Department of Electronics and Telecommunications (ETC), from the University „Dunărea de jos” of Galați), (CCETIC, 2023).

The first database (DBold-2019) contains a relatively small number of images (about 60) in the visible and thermal spectrum with power electronic circuits, to study the thermal behavior under different loading conditions, Fig. 6, and human facial images, Fig. 7, suitable for recognizing/detecting human faces and their associated emotional states (e.g., neutral, happy and sad).

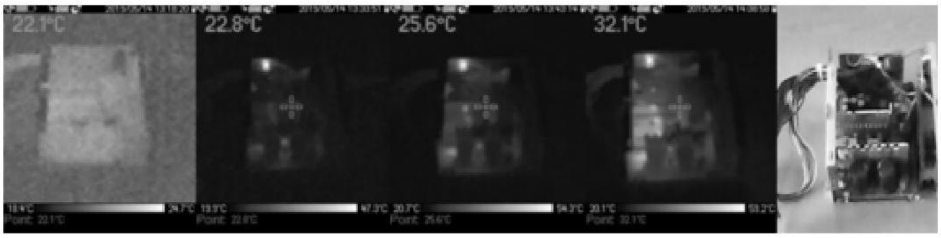

Fig. 6 – Thermal images from the DBold-2019 database from CCETIC, with a power supply under different load conditions, (CCETIC, 2023).

Fig. 7 – Examples of facial images in the thermal spectrum from the DBold-2019 database from CCETIC, showing three states (neutral - N, happy - V and sad - T), (CCETIC, 2023).

The second database (DBnew-2023) contains exclusively human facial images (full uncropped images), Fig. 8, suitable for the recognition/detection of a person's face and emotional state, it was recently created and contains approximately four hundred and eighty images in the visible and thermal spectrum (approximately one hundred and sixty student volunteers participated in the creation of the database, from the faculties of ACIEE (Faculty of Automation, Computers, Electrical Engineering and Electronics) and Letters from the „Dunărea de Jos" University of Galați).

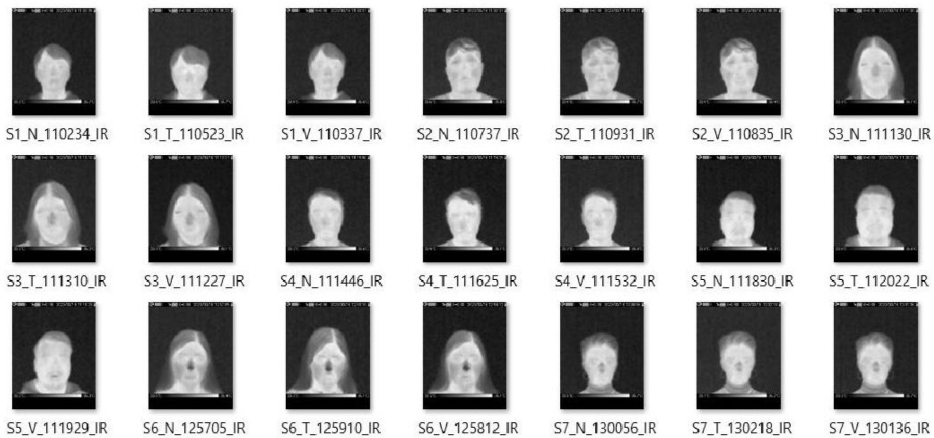

Fig 8 – Examples of facial images in the thermal spectrum from the DBnew-2023 database of CCETIC, showing three states (neutral - N, happy - V and sad - T), (CCETIC, 2023).

The third database (DBnew-2023_Crop) contains exclusively cropped images from the DBnew-2023 database (only the area of interest was cropped from the images belonging to database DBnew-2023). An example from the database is presented in Fig. 9.

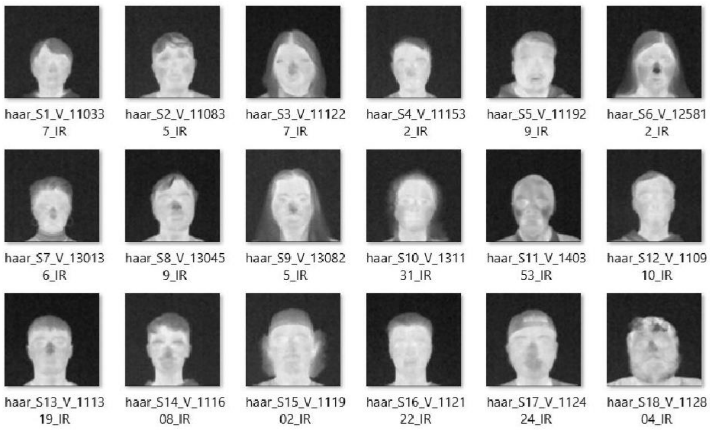

Fig. 9 – Examples of cropped facial images in the thermal spectrum from the DBnew-2023_Crop database of CCETIC, showing happy states – V, (CCETIC, 2023).

The fourth database (DBnew-2023_Noise) contains exclusively cropped thermal facial images in gray tones from the database DBnew-2023 with different noise levels (with PSNR (Peak Signal-to-Noise Ratio) from inf (infinite) to 19 dB). The extended database includes approximately forty-eight thousand facial thermal images. An example from the database is presented in Fig. 10.

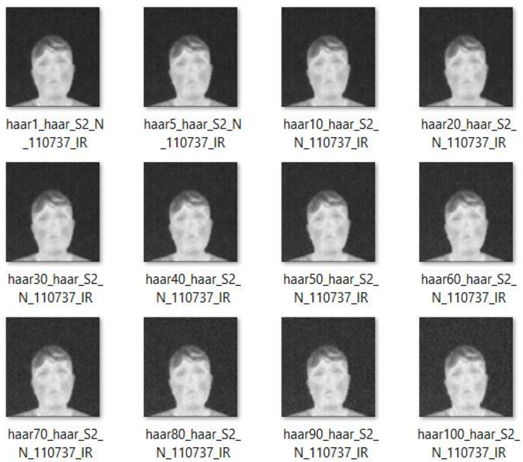

Fig. 10 – Thermal images with grayscale conversions with different levels of noise from the DBnew-2023_Noise database of CCETIC (Gaussian noise). Showing neutral state – N, (CCETIC, 2023).

The fifth database (DBnew-2023_ROI) consists of ROIs (regions of interest) composed of cropped RGB and grayscale images with the mouth area and the eyes area (the main regions of interest in the face area). The development of this database (image category) was considered because in the case of human facial images when the emotional state changes, it is considered that only certain facial areas show noticeable changes in structure and temperature. Some examples of thermal images (ROIs) in grayscale from the DBnew-2023_ROI database are shown in Fig. 11.

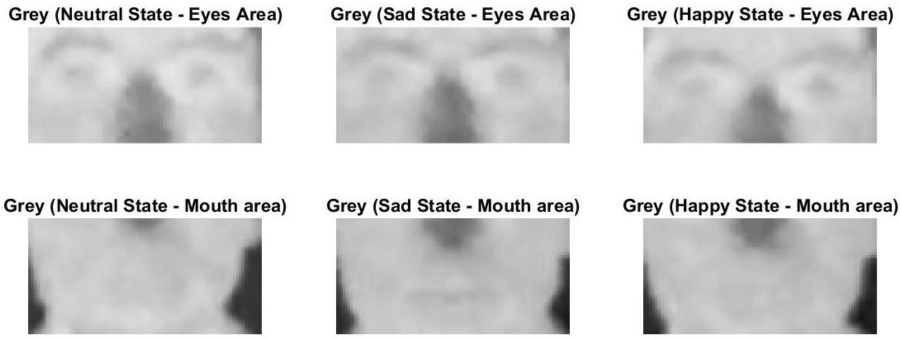

Fig. 11 – Examples of grayscale images (ROIs) from DBnew-2023_ROI database, showing three emotional states (neutral, happy and sad), (CCETIC, 2023).

Due to the fact that the DBnew-2023 database has a relatively small to medium size, it was considered necessary for certain research objectives and techniques to increase the number of images in the database. This was achieved by implementing a database expansion technique by applying different levels of noise to each thermal image, the type of noise chosen was Gaussian. The maximum number of resulting images was selected based on a qualitative assessment (if the resulting images became blurry, the process was stopped).

Fig. 12 shows an example of a thermal image after conversion to gray levels (image 2) and adding two types of noise, salt and pepper (image 3) and Gaussian (image 4).

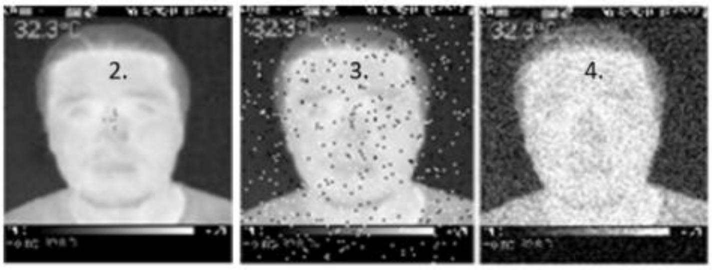

Fig. 12 – An example of a facial thermal image with different types of noise.

Fig. 13 shows the results obtained based on the three sets of images in gray tones from Fig. 12 (images 2, 3 and 4) after applying average filtering (images A1 - gray level image, A2 - salt and pepper noise image and A3 - Gaussian noise image), Gaussian filtering (images B1 - gray level image, B2 - salt and pepper noise image and B3 - Gaussian noise image) and median filtering (images C1- gray level image, C2 - salt and pepper noise image and C3 - Gaussian noise image).

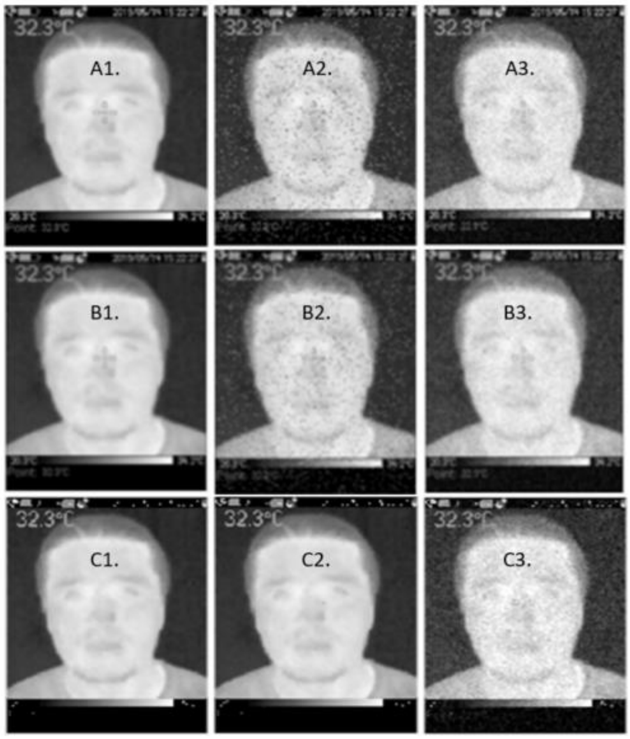

Fig. 13 – The results of the filtering process of a thermal image in gray levels.

This experiment tries to highlight the results of the filtering process of a thermal image in gray with different types of noise (salt and pepper and Gaussian). The purpose of the filtering process is to eliminate unwanted noise from the images and to make the resulting images (filtered images) look as close as possible to the originals (images without noise), without losing the original details.

In the case of the first set of images in gray tones (A1, B1 and C1), without added noise, regardless of the filtering method used, the results are comparable, almost identical from a visual point of view.

In the case of the second set of images (A2, B2 and C2), with salt and pepper type noise, the best result was obtained with median filtering (C2), the other two results are comparable (A2 and B2).

In the case of the third set of images (A3, B3 and C3), with Gaussian noise, the best result was obtained for the filtering mode specific to this type of noise, i.e. Gaussian filtering (B3), followed by average filtering (A3) and the worst result was obtained with median filtering (C3).

The databases from the CCETIC research center (CCETIC, 2023) highlighted in this section present a number of minor imperfections in terms of execution and design. These imperfections refer to the equalization of the size and the orientation of the face in the case of some thermal images from the databases, problems that can be corrected or resolved by applying specific methods and techniques. These imperfections could negatively affect the process of extracting features and regions of interest (ROIs) from target images and can also affect the accuracy of the final classification process (results).

The original DBnew-2023 database (thermal images only), the corresponding database DBnew-2023_Crop with cropped images, the database DBnew-2023_Noise (thermal images only in gray levels with noise) and the database DBnew-2023_ROI (composed of cropped RGB and grayscale images with the mouth area and the eyes area) are available for download from the web page of the CCETIC research center, at the following internet address (http://www.etc.ugal.ro/ccetic/eng/database2.html; CCETIC, 2023).

3. The general structure of the software application for thermal images

The structure of the thermal image acquisition and processing system (Aiordachioaie *et al*., 2019) is shown in Fig. 14. The images were provided by a dedicated thermal camera, and come from the database DBold-2019 and DBnew-2023 from CCETIC (CCETIC, 2023). The processing functions are organized in toolboxes. The first toolbox, which is called „Image processing toolbox" contains classic processing functions. The second toolbox, which is called „Machine Learning Toolbox (for advanced processing functions)" is planned to be addressed in the next stages of the research (Aiordachioaie *et al*., 2019).

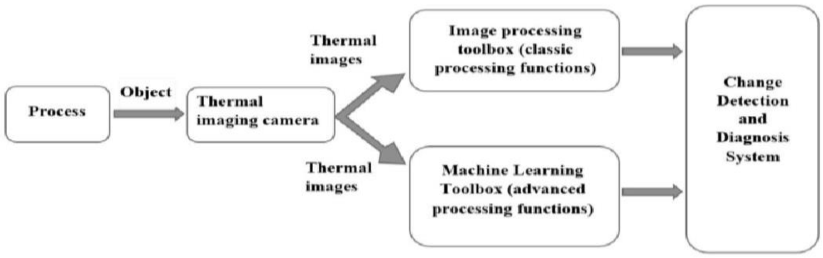

Fig. 14 – The structure of the system (Aiordachioaie *et al*., 2019).

The toolbox is organized in packages. At the primary level, pre-processing and auxiliary functions are considered to be from (1) to (5). At the second level, advanced processing functions are considered to be from (6) to (8). The main functions implemented are classified into (Aiordachioaie *et al*., 2019):

(1) Image format conversions (code IFC):
- Image conversion formats, e.g. from RGB to Gray;
- Data structures conversions, i.e. from 2D and 3D to 1D and vice versa;

(2) Image pre-processing (code IPP):
  - Image filtering (median filtering);
  - Image resizing;

(3) Automated image processing (code AIP):
- Image registration;
- Image segmentation;
- Image fusion;
- Image enhancement;

(4) User-based image processing (code UIP):
- Extraction of the region of interest;
- Temperature visualization at the pixel level, i.e. specifying the coordinates of the pixels;

(5) Numerical transforms (code NT):
- Temperature computation from raw images;
- Temperature computation when the range of the temperatures is changing (rescaling);
- Histogram computation and visualization;
- Peaks temperature detection;

(6) Change detection (code CD):
- Object detection;
- Change detection based on statistical signal processing;
- Change detection based on information processing.

(7) Pattern Recognition (code PR):
- Feature extraction;
- Compression;
- Object recognition;
- Classification.

(8) Machine learning (ML):
- Supervised;
- Unsupervised.

The above declarative neutral functions are graphically represented in Fig. 15, where the structure of the toolbox is presented.

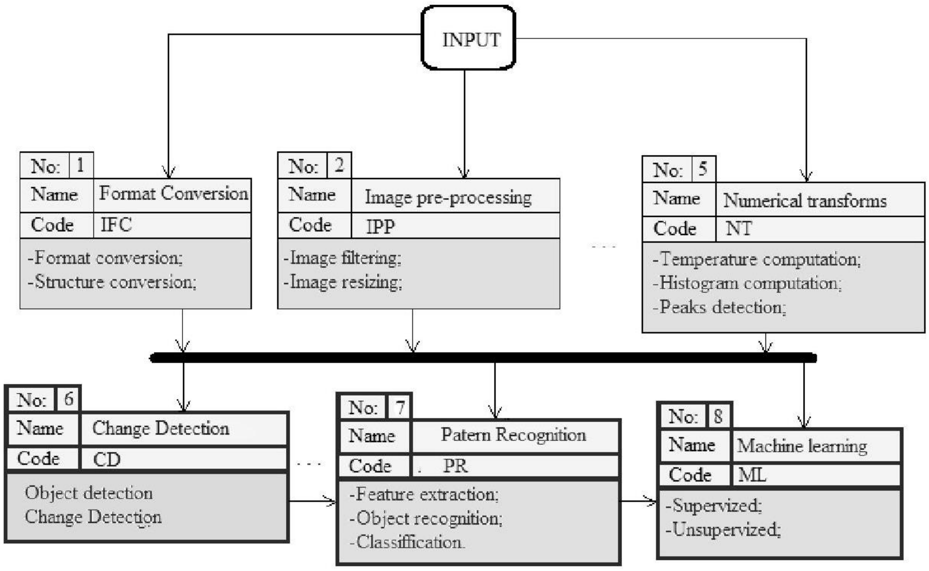

Fig. 15 – The structure of the toolbox, the main components (Aiordachioaie *et al*., 2019).

Object recognition involves identifying and locating specific objects within an image. In the context of facial thermal image processing, object recognition would be used to accurately detect and locate human faces within the thermal images.

Classification involves categorizing a detected object into predefined classes or categories. In the case of detecting and classifying emotional states in humans from thermal images, classification would be used to assign emotions such as happy, sad, angry, etc., to the detected facial features.

Machine learning involves training a model to perform a specific task without being explicitly programmed for that task. In the context of facial thermal image processing for emotional state detection, machine learning algorithms can be employed to analyze and learn features from the thermal images to predict emotional states accurately.

In supervised learning, the model is trained on labeled data, where each input example is paired with the corresponding output. For the facial thermal image processing scenario, training a supervised learning model would involve providing thermal images with labeled emotional states (happy, sad, angry, etc.) to teach the system to predict emotional states accurately from new images.

Unsupervised learning involves training a model on unlabeled data, where the system identifies patterns and relationships within the data without predefined labels. In the context of facial thermal image processing, unsupervised learning could be used to uncover hidden patterns or groupings in the thermal images that could potentially reveal insights about emotional states without explicit labels.

4. Basic techniques and methods for thermal image processing

Image processing consists of several methods and techniques that involves the manipulation and analysis of images and it refers to certain operations which are performed in order to obtain an improved image or to extract some useful information from the target image. Currently, image processing is among the rapidly developing technologies and forms a core research area within engineering and computer science disciplines, (Anbarjafari, 2023).

Image processing includes the following steps (Anbarjafari, 2023):

- Image import through image acquisition tools;
- Image analysis and manipulation;
- Image results (obtaining the result of a modified image or a report based on the analysis of that image).

There are two types/categories of methods used for image processing, namely, analog and digital image processing. Digital processing techniques help manipulate digital images through the use of computers. The four general phases that all types of data must go through during the application of digital techniques are pre-processing, enhancement and display, and information extraction.

There are many image processing methods, which depend on the specific objectives, e.g., such as thresholding (e.g., Otsu method) (Suryani *et al*., 2020), geometric transformations (Wolberg, 2011), image filtering (Chandel and Gupta, 2013), segmentation (Saini and Arora, 2014), contour detection (Arbeláez *et al.*, 2010), alignment (Shima *et al*., 2010), feature extraction (Tian, 2013), adding different types of noise (impulsive, salt and pepper and Gaussian) (Ilango and Marudhachalam, 2011), and so on.

In paper (Moldovanu *et al*., 2023), is implemented a database expansion technique by applying different levels of noise to each thermal image, the type of noise chosen in this work was Gaussian. The maximum number of resulting images was selected based on a qualitative assessment (if the resulting images became unclear, the process was stopped). This was done for two reasons. First of all, due to the low number of thermal images in the database. In order to check and validate the detection and classification algorithms, a database with a larger number of images is necessary. Secondly, we wanted to test the effect of different levels of noise applied to the thermal images, in order to test the performance of the processing algorithms. Taking into account the fact that thermal images have a lower informational content, compared to images in the visible spectrum.

In paper (Pavel *et al*., 2023), is presented a different way of processing facial thermal images with the aim of detecting the emotional states associated with each person. Three emotional states were considered: normal, sad and happy. The thermal images are processed so that we can select and extract the important features by using the random forest (RF) classifier. All features and the selected set are used as input to a feedforward back-propagation neural network (FFBPN) with binary outputs. The input images are organized into pairs (normal with sad and normal with happy). This paper gives a perspective on Haralick's texture features, RF and FFBPN techniques.

In paper (Aiordachioaie *et al*., 2019), is presented a method for processing facial thermal images (human facial images), for classification purposes. The main task is to find correct representations and descriptors/features for high recognition rates. Various data transformations are considered for image processing, accompanied by information-based transformations, e.g. transformed Renyi. Classification is considered and based on pattern similarity. Image compression is widely used in many applications, mainly for storage and transmission. This is based on a well-known theory with numerous standards and examples available, (Sikora, 1997) and (Watson, 1994). For the issue/situation corresponding to change detection, only the selection and extraction of the main coefficients are mandatory. Solutions addressed in image compression are also used here, i.e. a zigzag trajectory in terms of reading the DCT (Discrete Cosine Transform) coefficients and extracting a limited number based on their values, (Aiordachioaie *et al*., 2019).

In paper (Pavel and Aiordachioaie, 2021), from various available methods and discriminant functions, is taken into account/consideration the analysis of the informational content through two informational measures/quantities, the Renyi entropy and the Renyi divergence. The probability distributions involved are estimated in two ways, with interest in implementing the simpler one in custom programmable hardware such as FPGA. The chain/links of the main pre-processing blocks involved in thermal image processing, from acquisition to database generation, is shown in Fig. 16. The first block performs a conversion from RGB format to gray levels. The next processing block adapts/adjusts the images by temperature scaling (e.g., in a dynamic range between 20 and 35ºC). The last two blocks perform the image registration operations, respectively their size scaling (useful for automatic selection of important/interesting areas). In the processing of thermal images, a specialized software toolbox could contribute to all these steps, as described in (Aiordachioaie *et al*., 2019).

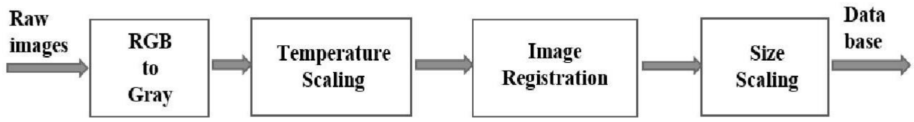

Fig. 16 – The main pre-processing blocks.

Fig. 17 shows the blocks developed in this work, in order to apply the change detection process. The first block is represented by the selection of interest areas (automatically or manually), the second block addresses source identification (i.e. source distribution estimation), the third block is represented by information measures/quantities (basically addresses entropy and divergence calculation), and the last block is intended for the classification process.

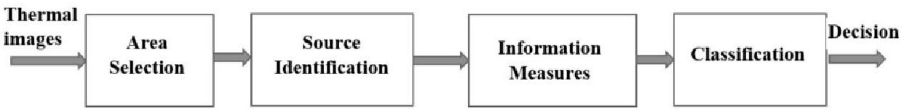

Fig. 17 – The blocks involved in the change detection process.

The classifier uses three classes, corresponding to the three emotional states (normal/neutral, cheerful/happy and angry/sad) and measures the Manhattan distance between the input model and the prototype of each class.

In the paper (Pavel *et al*., 2024), the study utilized Run-Length Features, which are statistical measures that characterize the distribution of consecutive homogeneous regions within an image. These features were then used as input to a TPOT AutoML, a Python-based machine learning library that utilizes/employs low-code methods to automate the generation and handling of machine learning workflows. Subsequently, a classifier was selected/chosen and the outcomes were validated using the bagging technique. The detection and classification process involved three image categories: the face area, the mouth area, and the eyes area.

5. Experimental results with thermal facial images

The objective of this section is to present some general aplications of thermal images and the results we obtained in the field of change detection and emotion recognition in human faces, based on thermal images (classifications of facial thermal images).

There are many applications related to the automatic analysis of thermal images. As image quality and resolution increase and prices decrease, more and more thermal imaging applications will emerge (Berg, 2016). The main application areas for thermal imaging are in the following fields/categories of activities: agricultural, (Roselyne *et al*., 2014), automotive safety, (Berg *et al*., 2015), building inspection, (Preda *et al*., 2019), fire monitoring, (Paugam *et al*., 2013), industry, (Berg *et al*., 2016), medical, (Ahlberg *et al*., 2015), military, (Siegel, 2002), personal use, (Jones and Plassmann, 2002), search and rescue, (Pere *et al*., 2012) and security, (Berg *et al*., 2015a, b).

Thus, in paper (Pavel *et al*., 2023) the selection process of the important features with RF (Random forest) and experimental findings were made with different FFBPN (Feedforward Backpropagation Neural Network) architectures approaches, which are based on the observations generated in this study. The dataset with a total of 231 samples (thermal facial images) was divided into 60% training data, 20% validation, and 20% testing data. The experimental environment (hardware and software) consists of Python 3.9 programming language for implementing RF classifiers and Matlab R2020a for implementing FFBPNs. Superior FFBPN test results were found when the architecture contains 400 neurons in the hidden layer, for the specified architecture and MSE (mean squared error) of 0.052, the classification accuracy exceeded 80%. In this work, thermal images from the database of the Research Center for Electronics, Communications and Information Technology (CCETIC), Signal and Information Processing Laboratory were used (CCETIC, 2023).

The results obtained with the K-NN (k-nearest neighbors) algorithm for features classification are presented in paper (Moldovanu *et al*., 2023), where two ways of approach were proposed for the data set: the first way of approach involves extracting features from raw images, and in order to increase the number of facial thermal images (amount of data) used in the classification stage, the data (thermal images) were augmented with the SMOTE method and the second way of approach involved extracting features from noise-corrupted images. In both cases, K-NN (k-nearest neighbors) algorithm classified the features and the obtained confusion matrices are presented in the paper (Moldovanu *et al*., 2023) along with the accuracy values, highlighting some promising results (for raw images the accuracy was about 92% and for the images corrupted by noise about 70%). In this work, thermal images from the database of the Research Center for Electronics, Communications and Information Technology (CCETIC), Signal and Information Processing Laboratory were used (CCETIC, 2023). Fig. 18 presents a set of six images with different level of noise (with PSNR from inf (infinite) to 19 dB).

The results obtained with DCT (Discrete Cosine Transform) technique for the selection and analysis of the areas of interest, combined with different distances to determine the percentage of detection accuracy, such as Euclidean (E), Manhattan (M), Canberra (C) and Pearson (P), were presented in the paper (Aiordachioaie *et al*., 2019). For each person there are three images, one for each emotional state. Fig. 19 shows the basic transformations applied to the raw image generated by the thermal imaging camera, resulting in an image at a resolution of 200x200 pixels in RGB and gray format. The final image is a thermal image, taking into account the boundaries used for each person (IR image). In Fig. 20 the process of analysis and selection of the region with DCT (Discrete Cosine Transform) applied to a facial thermal image in gray levels is presented. In this work, thermal images from the database of CCETIC (CCETIC, 2023) were used.

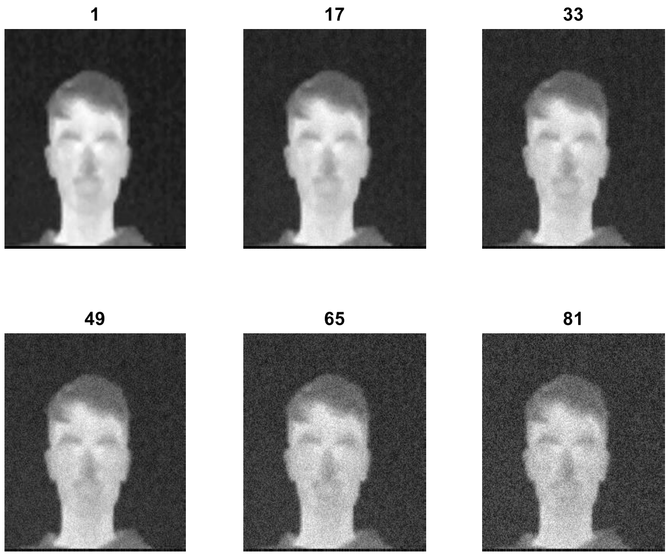

Fig. 18 – Thermal images with grayscale conversions and with different levels of noise (Gaussian noise), (Moldovanu *et al*., 2023).

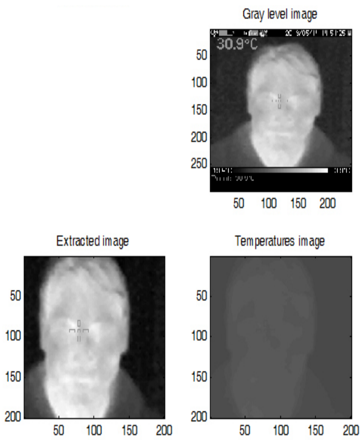

Fig. 19 – Basic transformations of thermal images generated by the IR camera (Aiordachioaie *et al*., 2019).

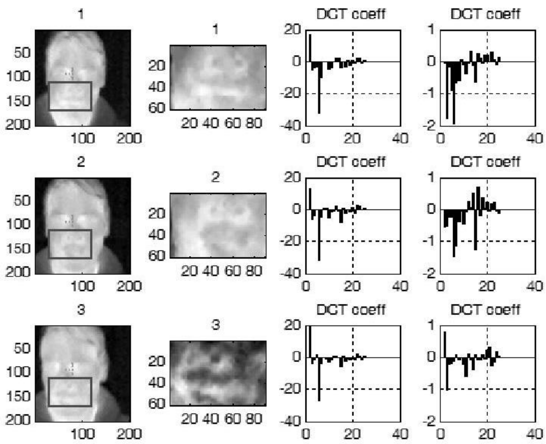

Fig. 20 – Analysis and region selection with DCT (discrete cosine transform) (Aiordachioaie *et al*., 2019).

The results obtained in the process of classifying emotional states (eg neutral, happy and sad) using different distances such as Euclidean (E), Manhattan (M), Canberra (C) and Pearson (P) are promising. Only three classes with the same number of images are considered. There are seven images for each class. The best distance is Canberra (maximum accuracy of 100%) and the minimum number of coefficients is 400 = 20 x 20, (Aiordachioaie *et al*., 2019).

The results obtained with Renyi divergence (RD) and Renyi entropy (RH) for determining the percentage of accuracy, concerning the detection of people and their associated emotional states, are presented in paper (Pavel *et al*., 2021). In this case were applied different image processing steps to remove the annotations and to select the area of interest from the facial thermal image, as in Fig. 21. A downscaled image with a size of 200x200 pixels is selected and later converted into a temperature image (ie, each pixel has a value equal to the body temperature, indicated by the thermal camera, in the range from 20 to 35^{o}C). In this work, thermal images from the database of CCETIC (CCETIC, 2023) were used.

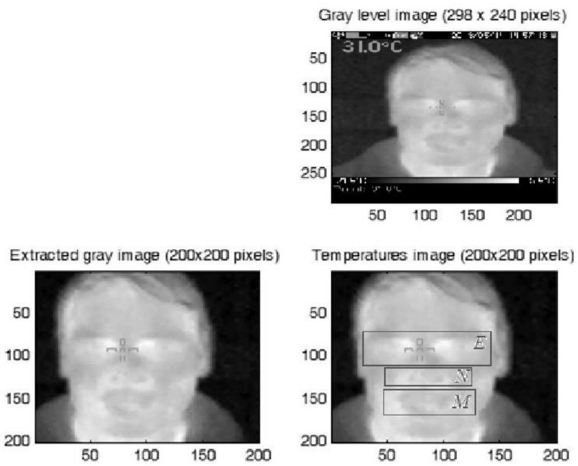

Fig. 21 – Thermal images transformation from RGB model to gray levels and selecting the areas of interest (Pavel *et al*., 2021).

The numerical results for the cases corresponding to the four areas of interest (M, M+N, E, M+N+E) are promising, each area includes four cases. The most relevant/sensitive areas of the face for emotional state recognition are considered to be the mouth (M) and eye (E) areas (maximum accuracy of approx 71%). Experiments with more complex areas such as mouth area plus nose area (M+N) (maximum accuracy of approx 76%) and mouth area plus nose area plus eye area (M+N+E) (maximum accuracy of approx 72%) were also developed. The best results were obtained for the mouth area plus the nose area (M+N). For each region of interest, the Renyi divergence (RD) indicates the best results, that is, the highest recognition rate, and the statistical distribution is estimated by normalizing the image. The results obtained with the Renyi entropy (RH) are not acceptable, regardless of the estimation method (pdf – probability density function), being below 50%.

In the paper (Pavel *et al*., 2024), Run-Length Features were used/employed to extract statistical features/measures, while TPOT AutoML was utilized for classification and assessing the accuracy rate. The initial findings demonstrate promising classification accuracy, with the highest accuracy achieved for the entire/complete face area ranging between approximately 70% and 77%. In contrast, the accuracy rates for the mouth and eyes areas were lower, falling between approximately 60% and 67%. The area with the lowest accuracy was the mouth region, with an accuracy rate of about 60%. Fig. 22 shows the three categories of images with the three emotional states, used in this work.

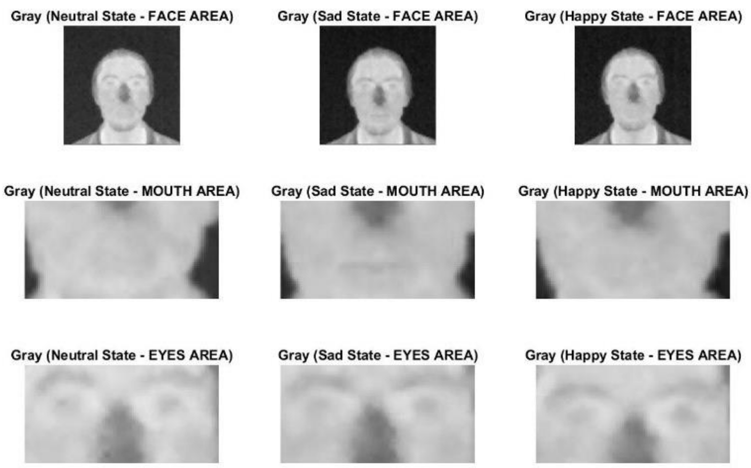

Fig. 22 – The three categories of images with the three emotional states in gray levels (Pavel *et al*., 2024).

The experimental results presented above, obtained through techniques, methods and algorithms such as K-NN (k-nearest neighbors), highlight certain encouraging results regarding the processing of facial thermal images for the purposes of detection and diagnosis of people and their associated emotional states. Regardless of the classification method used, the accuracy results (ACC) are comparable, varying between approximately 70% and 90%, with few exceptions.

It seems that the research field of thermal imaging is not sufficiently explored, and in future research, it is proposed that facial expressions to be explored with features based on models and convolutional neural networks (CNN).

6. Conclusions

The general objective of the paper is to introduce the main basic characteristics regarding the design and development of a facial thermal image database, and also to present the structure of the thermal image databases from the CCETIC (CCETIC, 2023) research center.

A secondary objective was the presentation of the general structure of the thermal image processing software application and the overview of the block structure of the application.

Another objective was to present a series of thermal image processing methods and techniques, together with certain research results, carried out in the field of thermal image processing, with the help of thermal images from the databases of the CCETIC research center.

Taking into account everything that has been presented and discussed so far regarding the process of developing a thermal image database, we can conclude the following:

- In order to obtain satisfactory results regarding the event detection and diagnosis process, it is necessary for the database to contain a sufficiently large number of images (in our case thermal facial images), from at least one hundred candidates;
- The acquisition of images from the database should be done according to certain rules for each individual emotional state;
- The need to select an appropriate environment that meets the conditions required for the selected research objectives;
- Creating the dataset according to the research area, objectives and tasks they intend to solve;
- Setting some parameters regarding the appropriate adjustment of the thermal imaging camera in order to acquire correct the data set (for example, the appropriate setting of the emissivity corresponding to the human body, which is 0.98);
- Each participant should be positioned so that the background does not influence the desired elements from the image (a background that does not present high values in terms of infrared radiation, and does not present fluctuations);
- The camera should be mounted on a tripod and its height could be adjusted manually, for a proper centering of the image;
- A constant lighting condition and the distance to the participant should be strictly controlled during the acquisition process;
- The thermal camera should be set to remove the markers and annotations from the image (cursor and temperature values). That information/text negatively influence the detection process;
- The resulting images should be pre-processed (cutting, alignment, selection of the areas of interest, removal of annotations, etc) to

improve the final results regarding the detection and diagnosis process (recognition of emotional states and recognition of persons).

The next research steps involves the following:

- Use of other transforms in the analysis of images and/or image blocks;
- Define an adaptive and selective system regarding the size and content of the processed image for facial recognition and emotion detection;
- Using other naturally inspired systems (neural networks) in thermal image processing;
- Develop a new identification procedure for thermal transfer modeling, useful for estimating the temperature of points, where this information is not available or when the point of interest is outside the normal working range;
- Develop a feature integration system for pattern recognition, including feature selection and extraction (regions of interest), for change detection and diagnosis.

Acknowledgments. The authors would like to thank all reviewers. Their effort and recommendations are very high appreciated and were very useful in preparing of the final version of the manuscript.

REFERENCES

Ahlberg J., Berg A., Markuš N., *Multi-person fever screening using a thermal and a visual camera*, Linkoping: Swedish Symposium on Image Analysis (SSBA), (2015).

Aiordachioaie D., Culea-Florescu A., Pavel M.S., *On Human Faces Thermal Image Processing for Classification Purposes*, 6th International Symposium on Electrical and Electronics Engineering (ISEEE), pp. 1-6, (2019), Galați, Romania, doi: 10.1109/ISEEE48094.2019.9136103.

Anbarjafari G., University of Tartu, *Digital Image Processing. Introduction to image processing*, Retrieved 1 21, 2023, https://sisu.ut.ee/imageprocessing/book/1.

Arbeláez P., Maire M., Fowlkes C., Malik J., *Contour Detection and Hierarchical Image Segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, (2010), pp. 898-916, 10.1109/TPAMI.2010.161.

Arnoux C. (n.d.). *CA 1950 Thermal camera*, https://www.arc.ro/chauvin-arnoux-ca-1950-diacam2#tab-2, Accessed April 2023.

Balaras C., Argiriou A. (n.d.), *Infrared thermography for building diagnostics*, *Energy and Buildings*, Volume 34, Issue 2, February 2002, pp. 171-183, https://doi.org/10.1016/S0378-7788(01)00105-0.

Berg A., Ahlberg J., Felsberg M., *Enhanced analysis of thermographic images for monitoring of district heat pipe networks*, Linköping: Submitted to Pattern Recognition Letters (PRL), (2016).

Byung M., Dong G.J., Soon J.B., *Fabrication of Silicon Window for Low-price Thermal Imaging System*, Journal of Sensor Science and Technology, Volume 24 Issue 4, pp. 264-269, (2015).

Baral S., Miandashti R.A., Richardson H., *Near-field thermal imaging of optically excited gold nanostructures: scaling principles for collective heating with heat dissipation into the surrounding medium*, Nanoscale, (2018).

Berg A., *Detection and Tracking in Thermal Infrared Imagery*, Linköping: Linköping University, (2016), Linköping studies in science and technology.

Berg A., Ahlberg J., Felsberg M., *A thermal object tracking benchmark*, Linkoping: Advanced Video and Signal Based Surveillance (AVSS), (2015a), 12th IEEE International Conference.

Berg A., Öfjäll K., Ahlberg J., Felsberg M., *Detecting rails. Lecture Notes in Computer Science*, Springer International Publishing, pp. 492-503, Copenhagen, (2015b).

Brouant L., Pham H., Foucher É., Picard A., *VIS-NIR multispectral for camouflage detectio*, Proceedings Volume 12740, Emerging Imaging and Sensing Technologies for Security and Defence VIII, 1274004, Amsterdam (2023), https://doi.org/10.1117/12.2678661.

Caimi F.M., *Technical challenges and recent developments in underwater imaging*, Society of Photo-Optical Instrumentation Engineers (SPIE), (1995), Micro-Optics/Micromechanics and Laser Scanning and Shaping.

CCETIC, Laboratorului de Semnale și Prelucrarea Informațiilor, Centru de cercetare în electronică, Galați: ETTI (2023), http://www.etc.ugal.ro/ccetic/eng/database2.html.

Chandel R., Gupta G., *Image Filtering Algorithms and Techniques: A Review*, International Journal of Advanced Research in Computer Science and Software Engineering, Computer Science Department, (2013).

Chen J., Norman J., Nam Y., (2020, 03 15), *American Multiracial Face Database*, (CC-By Attribution 4.0 International) Retrieved 2023, from https://osf.io/qsdrp/.

Group T.S., *Measurement Parameters*, (2020), https://www.thesnellgroup.com/featured-tips/ir-camera-measurement-tools-and-parameters-what-do-they-mean.

Haider M., Doegar A., Verma K.R., *Fault Identification in Electrical Equipment using Thermal Image Processing*, 2018 International Conference on Computing, Power and Communication Technologies (GUCON), (2018), Greater Noida, India.

Hung N., Kazunori K., Fan C., Bac L., *A Thermal Facial Emotion Database and Its Analysis*, Ho Chi Minh: Japan Advanced Institute of Science and Technology, (2014).

Ilango G., Marudhachalam R., *New Hybrid Filtering Techniques for Removal of Gaussian Noise from Medical Images*, ARPN Journal of Engineering and Applied Sciences (2011).

Jeong S., Kim M.G., Oh K.Y., *Thermal monitoring of live-line power transmission lines with an infrared camera mounted on an unmanned aerial vehicle*, SageJournals, Structural Health Monitoring, Volume 22, Issue 6, (2023), https://doi.org/10.1177/14759217231156359.

Jędrasiak K., Nawrat A., *The Comparison of Capabilities of Low Light Camera, Thermal Imaging Camera and Depth Map Camera for Night Time Surveillance Applications*, Advanced Technologies for Intelligent Systems of National

Border Security, pp. 117-128, *Part of the Studies in Computational Intelligence book series (SCI,volume 440)*, (2013).

Jonathan T., Xiaopeng H., Svetlana B., James L., Yanfei X., Gang C., *Infrared-Transparent Visible-Opaque Fabrics for Wearable Personal Thermal Management*, ACS Photonics, American Chemical Society, pp. 769-778, (2015), https://doi.org/10.1021/acsphotonics.5b00140.

Jones B., Plassmann P., *Digital infrared thermal imaging of human skin*, IEEE Engineering in Medicine and Biology Magazine, pp. 41-48 (2002).

Karen P., *Tufts Face Database Thermal Cropped. Panetta's Vision thermal face images*, (2019), http://tdface.ece.tufts.edu/.

Kölzer J., Oesterschulze E., Deboy G., *Thermal imaging and measurement techniques for electronic materials and devices,* Microelectronic Engineering, Volume 31, Issues 1–4, pp 251-270, February 1996.

Kim H.S., Lim C.S., Kim Y.D., *Intelligent intrusion detection system featuring a virtual fence, active intruder detection, classification, tracking, and action recognition*, (2017), Annals of Nuclear Energy Volume 112, February 2018, pp. 845-855.

Liu Z., Wang S., Wu G., Peng P., *A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference*, (2010), IEEE Transactiond on Multimedia, pp. 682-691.

Marius P., Codreanu N.D., *Termoviziune şi termografie în industrie*, (2019), https://www.electronica-azi.ro/print.php?id=1247.

Moldovanu S., Pavel M.S., Aiordachioaie D., *First Order Features Extraction from Thermal Images for Human State Recognition, The 8th International Symposium on Electrical and Electronics Engineering (ISEEE)*, (2023), Galați.

Nguyen H., Kazunori K., Fan C., Bac L., *A Thermal Facial Emotion Database and its Analysis*, (2013), Pacific-Rim Symposium on image and video Technology, pp. 397-408, Guanajuato.

Paugam R., Wooster M., Roberts G., *Use of handheld thermal imager data*, Geoscience and Remote Sensing, IEEE Transactions on, (2013), 51(6):3385-3399, IEEE.

Pavel M.S., Aiordachioaie D., *On Change Detection in the State of the Human Faces by Thermal Images and Renyi Information Measures*, 2021 International Symposium on Signals, Circuits and Systems (ISSCS), (2021), Iași, Romania, DOI: 10.1109/ISSCS52333.2021.9497373.

Pavel M.S., Moldovanu S., Aiordachioaie D., *On the use of the Ensemble Learning Method for the Classification of the Human Emotions*, (2024), the 28th International Conference on System Theory, Control and Computing (ICSTCC 2024), Sinaia.

Pavel M.S., Moldovanu S., Aiordachioaie D., *Emotion Recognition in Human Thermal Images with Artificial Intelligence Technology*, (2023), IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia.

Pere M., Ismael C., Teo V., *Drones to the Rescue! Unmanned Aerial Search Missions Based on Thermal Imaging and Reliable Navigation*, (2012), InsideGNSS.

Preda A., Scurtu I., *Thermal image building inspection for heat loss diagnosis*, (2019), 5th International Scientific Conference SEA-CONF 2019.

Rai M., Maity T., Yadav R.K., *Thermal imaging system and its real time applications: a survey*, American Society for Engineering Education, Journal of Engineering Technology (ISSN: 0747-9964), Volume 6, Issue 2, July, pp. 290-303 (2017).

Roselyne I., Faruk A., Khaled A., *Applications of Thermal Imaging in Agriculture - A Review*, Scientific Research Publishing Inc. Johannesburg (2014).

Saini S., Arora K., *A Study Analysis on the Different Image Segmentation Techniques*, International Journal of Information & Computation Technology, ISSN 0974-2239, Volume 4, Number 14, pp. 1445-1452 (2014).

Sebastian B., *Human detection in thermal images using low-level features*, https://www.researchgate.net/publication/282328996_Human_detection_in_thermal_images_using_low-level_features, (2015), Gliwice.

Shima T., Sugimoto S., Okutomi M., *Comparison of image alignment on hexagonal and square lattices*, 2010 IEEE International Conference on Image Processing, (2010), 10.1109/ICIP.2010.5654351.

Siegel R., *Land mine detection*, Instrumentation Measurement Magazine, IEEE, (2002), 5(4):22-28.

Sikora T., *MPEG Digital Video-Coding Standards*, (1997), IEEE SP Magazine, pp. 82-100.

Suryani E., Asmari E.I., Harjito B., *Image Segmentation of Acute Myeloid Leukemia Using Multi Otsu Thresholding*, Journal of Physics: Conference Series, Volume 1803, 2nd International Conference on Enhanced research and Industrial Application 2020, 8-9 October 2020, Yogyakarta, Indonesia.

Tian D.P., *A Review on Image Feature Extraction and Representation Techniques*, (2013), International Journal of Multimedia and Ubiquitous Engineering.

Virginia E.D., Marcos F.Z., Jiří M., *Carl Database*, (2013), Signal Processing Laboratory.

Vollmer M., *Infrared Thermal Imaging*, (2021), Computer Vision pp. 666-670.

Wang S., Liu Z., Lv S., Lv Y., Wu G., Peng P., *A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference*, IEEE transactiond on Multimedia, (2010), pp. 682-691.

Watson A., *Image Compression Using the Discrete Cosine Transform*, Math. Journal, (1994), 4(1).

Wolberg G., *Geometric Transformation Techniques for Digital Images: A Survey,* Columbia University Computer Science Technical Reports, Department of Computer Science, (2011), Columbia University, https://doi.org/10.7916/D8TH8VRW.

Zhang Y., Wang Z., Fu X., *An experimental method for improving temperature measurement accuracy of infrared thermal imager*, Infrared Physics & Technology, Volume 102, November 2019, 103020, https://doi.org/10.1016/j.infrared.2019.103020.

CONSTRUIREA UNEI BAZE DE DATE CU IMAGINI TERMICE PENTRU CLASIFICAREA STĂRILOR EMOȚIONALE

(Rezumat)

Această lucrare de cercetare își propune în primul rând să evidențieze etapele privind dezvoltarea unei baze de date cu imagini faciale în spectru termic, evidențiind diverse stări emoționale (de exemplu, neutru, fericit și trist). Au fost prezentate și o serie de cercetări exploratorii, efectuate pe cinci baze de date cu imagini termice (DBold-2019, DBnew-2023, DBnew-2023_Crop, DBnew-2023_Noise și DBnew-2023_ROI), aparținând centrului de cercetare CCETIC (CCETIC, 2023). Procesul de dezvoltarea a bazei de date incluzând atât partea de achiziție a imaginilor termice cât și partea de pre-procesare a acestora (decupare, aliniere, selecția zonelor de interes, eliminarea adnotațiilor/elementelor de marcaj, etc). Și în al doilea rând, vizează prezentarea unei analize în ceea ce privește structura generală a aplicației software de prelucrare a imaginilor termice, împreună cu o serie de metode și tehnici de prelucrarea a imaginilor, abordate în anumite lucrări de cercetare. Sunt prezentate și câteva rezultate experimentale obținute cu ajutorul celor cinci baze de date, și o serie de concluzii generale în ceea ce privește dezvoltarea unei baze de date cu imagini termice faciale.