ECAI 2024
U. Endriss et al. (Eds.)

doi:10.3233/FAIA241072

Fuzzy Ensemble Learning and Lightweight CNN for Stress Classification

Katarzyna Baran [a,*]

[a]Lublin Univeristy of Technology

Abstract. Everyday life problems, armed conflicts, pandemics, and catastrophes – these are situations that are always accompanied by stress. Its chronic form can lead to so-called stress-related illnesses. Despite the development of health prevention, many people still get sick due to stress. Therefore, it is important to seek detective and classificatory solutions for stress, which may influence its reduction or control in the future. The example of this can be the thermographic stress registration presented in this article, combined with classification using lightweight CNN and Choquet fuzzy ensemble learning. The article proposed new ensemble frameworks for stress classification based on Choquet fuzzy integral, serving as an aggregation function. In the study, three pre-trained lightweight CNN models were used: MobileNetV2, Xception, and EfficientNet. The proposed fuzzy ensemble model achieves a classification accuracy above 90%. This work is of a prospective nature, with the possibility of implementing solutions in biomedical-psychological activities.

1 Introduction

Stress accompanies human daily life. Due to emerging armed conflicts, pandemics, natural disasters, concerns and fears arise among society, leading to stress. In research literature, stress is defined in various ways. The main definition of stress comes from Hans Selye [1]. He regarded stress as the body's reaction to various difficulties and challenges that require adaptation and mobilization of resources [2]. Seyle introduced the concept of the stress syndrome [3] as well as the General Adaptation Syndrome(GAS)[4]. Stevan E Hobfoll [5] presented another definition of stress, referring to the impact of interpersonal relationships on the experience of stress (relational context). Other definitions focus on an individual's reaction to a given experience of stress, as evidenced by changes in physiological or psychological signals [6, 7, 8]. Various types of stress can also be classified. The most commonly identified types of stress include: emotional, cognitive, social, behavioral, and physical [9, 10, 11]. It is also important to distinguish between eustress (positive stress) and distress (negative stress) [12]. Symptoms of stress can manifest in various forms [13], such as physical (feeling cold, body aches) and cognitive (tension, resignation, inability to relax).

Frequent or prolonged exposure to stress can lead to stress-related illnesses [14, 15]. To prevent them, it is important to focus on health prevention and create solutions that allow for monitoring and coping with stress. Due to the specific nature of stress and its individual immeasurability (inability to express in units such as pressure or pulse), detecting and classifying stress is a challenging task that requires an interdisciplinary approach. Given the individuality of people and the differences in their traits, it is important to highlight the most discriminating features when classifying stress. Objective and automated methods, which researchers are working on, can assist in this. Physiological and psychological signals can be helpful in stress detection [16, 17, 18]. Many researchers use popular but expensive devices, such as EEG [19], EKG [20], EMG [21], GSR [22], and wearable electronics [23]. Certainly, thermal stress detection represents a research niche, but there is noticeable interest in this direction. It is worth emphasizing, however, that strictly industrial thermal cameras are most commonly used, while their smartphone counterparts are less frequent. Smartphone cameras such as FLIR ONE Pro [24], Seek Thermal Compact [25], or InfiRay [26] can serve as low-cost research devices, providing accuracy and measurement sensitivity. In case of stress, thermographic recording of the subject's face helps to detect temperature changes and subtle facial expressions. This is particularly important when there are no visible signs of stress on the subject's face, yet the analysis of recorded images indicates the presence of stress and its concealment by the subject. Suppressing or concealing emotions can be detected through changes in temperature distribution and correlated with physiological signals such as blood pressure and pulse.

The development of image processing is connected with the advancement of object classification systems. Psychological stress detection through observations, surveys, and questionnaires is labor-intensive, error-prone, time-consuming, and dependent on the examiner's experience [27]. Computer methods can provide reliable results while ensuring cost-effectiveness and analysis speed. Image processing may include, among others: measuring morphological features, surface textures, object shapes, and color intensities in individual channels [28].

The dynamic development of machine learning contributes to the creation of increasingly accurate classification systems. Neural networks can operate on both small and large datasets, yielding the best results. However, with large datasets, models are often complex and require better computational resources. At the same time, deep networks can extract deeper feature representations, resulting in better object recognition. In the case of image classification, convolutional neural networks [29] play a crucial role. When limited computer resources and the desire to avoid time-consuming processes are present, lightweight convolutional networks [30] are used, such as: MobileNet, ShuffleNet, EfficientNet, and CondenseNet.

Researchers are creating increasingly larger and deeper deep learning models, with a large number of parameters and training samples. However, this is associated with computational burden [31]. There-

* Corresponding Author. Email: k.baran@pollub.pl

fore, it is considered appropriate to search for and create less demanding but equally effective networks In the case of stress classification, it can be observed that many methods rely on a single network. The performance of a single classifier may exhibit high randomness and instability. The solution to this problem is to use an approach based on ensemble learning [32]. By adding Choquet integral to ensemble learning[33], one obtains aggregation of multiple models. Furthermore, Choquet ensemble learning preserves the interpretability of the final model. The Choquet integral is described by the formula:

$$\int_X f(x)d\mu = \sum_{i=1}^{n} \left[f(x_{(i)}) - f(x_{(i-1)}) \right] \mu(A_{(i)}) \tag{1}$$

where: f is a measurable function, X is the domain of integration, and mu is a non-additive measure.

The Choquet integral as an aggregation operator allows for combining forecasts from multiple base models into one optimal model. In this way, it is possible to achieve better results than individual base models, especially in challenging classification tasks. To determine the importance of each base model and assign them appropriate weights in the aggregation process, the Choquet measure is used [34].

$$\mu(A) = \sup_{E \subseteq A, \mu(E) < \infty} \mu(E) \tag{2}$$

where: A is a subset of X, and mu is a non-additive measure

Ensemble methods can, among other things, increase accuracy, resistance to overfitting, enhance model stability, and improve prediction reliability. Therefore, they are increasingly used in solving complex machine learning problems. It is worth noting that static aggregation functions are often used, where input data are independent of each other for the purpose of simplifying calculations and defining parameters. The fuzzy aggregation function is used when a more general aggregation function is needed, and the researcher intends to pay attention to the relationships between input data during aggregation. Fuzzy aggregation functions allow for combining data in a way that takes into account uncertainty and ambiguity, characteristic of many problems. At the same time, it allows for the consideration of subjective assessments and preferences in the process of data fusion. The application of fuzzy set theory enables a more precise alignment with the data fusion problem, especially when the data is atypical or very similar. Overall, this translates into obtaining more flexible and accurate results.

This article presents a proposal for stress classification using lightweight CNN networks and Choquet ensemble learning based on thermographic imaging. Pre-trained lightweight network models: MobileNet, EfficientNet, and Xception were modified with additional layers to extract the most useful features and functions, as well as to protect against overfitting. The fully connected layer was used for stress classification and result generation for each class. These results were then combined using Choquet integral to obtain the final results.

The main contributions of this work are as follows:

- A technique for stress classification using Choquet fuzzy ensemble learning has been proposed to achieve better results,
- Modified lightweight convolutional networks were enhanced by adding layers: Global Average Pooling, Batch Normalization, Dropout, and Dense, in order to increase classification capabilities and avoid overfitting
- As an aggregation function, the choquet fuzzy integral was used to combine the results of lightweight deep learning models for the final classification of stress.

This article consists of six parts: the first chapter is an introduction, the second chapter indicates the research literature related to the subject matter, chapter 3 presents the research project and the dataset used, and chapter 4 describes the proposed method. Chapter 5 contains the results of the research carried out along with the analysis. The whole thing is finalized by a summary.

2 Related Work

In scientific literature, we can find many stress detection techniques. In the introduction, the use of EEG, EKG, EMG, GSR, and wearable electronics was mentioned most frequently. These are expensive devices primarily used in laboratory conditions. Thermographic stress recording constitutes a developing research niche. Appropriate research work can be divided into those using industrial, expensive cameras and those using smartphone, low-cost thermal cameras. In the study [35], researchers used the FLIR A40 MuSE base to investigate multimodal interactions between stress and affective expressions. The stressors included Question-Answer sessions and emotion-inducing videos. The research group consisted of 28 individuals. Stress classification was conducted using RNN and CAs, achieving an accuracy of 59.9%. Stress, recognized through EKG measurement, NIRS - near-infrared spectroscopy, and NST - nose skin temperature, was presented in the work [36]. Random forest, SVM, and STEP method were also utilized, achieving 76.5%. 10 participants performed arithmetic mental tasks during which ECG and OEG-Sp02 were measured. In [37], the FLIR SC7600 camera imaged hyperspectral oxygen saturation of facial tissues in five ROI areas. Stressor for 20 particpants was a physical stress test - squats. In this article, the HSI system was used, as well as the methods: LD, LR, kNN, DT, EL, SVM, achieving the highest accuracy of 82.11%. The analysis of inflammatory activity in the body and peripheral facial temperature was described in [38]. Researchers utilized FLIR A310 and the TSST test, while measuring interleukin-6 and blood pressure. The work itself consisted of analyzing biotermic biomarkers Researchers [39], using the Microsoft webcam, FLIR SC620, and the ANUStressDB dataset, conducted a study among 45 individuals watching stressful and non-stressful films. They used SVM, LBP-TOP, and HDTP for classification, achieving the highest accuracy of 72%. The entire work boiled down to proposing a method for recording dynamic thermal patterns on histograms. The measurement of the interaction between stress and emails was presented in the work [40]. The reactions of 63 participants were recorded using the Tau 640 thermal camera and the HD Pro Webcam C920, while the proposed technique was the GLM model. Other researchers [41] considered stress reactions during driving on a driving simulator.Ten people were registered by the FLIR Boson 320LW camera and the Intel RealSense D415. SVR with an RBF kernel was used in the analysis, achieving a final accuracy of 77%. Psychological and physical stress among 42 individuals was described in [42]. St02 levels were measured using the FLIR SC7600, PixFly camera, and Specim VNIR as research tools. The tests conducted were TSST and SCWT, with the highest accuracy being 95.56%. The search for specific correlations between temperature patterns and stress markers is discussed in the work [43]. Researchers proposed the DEFP algorithm and used FLIR SC760, Garmin Miroxi HBR heart monitor, and the TSST test for their studies. As seen, the aforementioned works utilized more expensive industrial thermal cameras. However, the use of smartphone thermal cameras in stress detection is still uncommon but promising. In medical areas, this type of camera has been used in challenging tasks, including limb surgery [44], burn diagnosis [45], assessment of diabetic foot [46], early detection of cancer stages [47], sleep apnea diagnostics [48], and groin vein bypass surgery [49]. When it comes to stress detection through smartphone thermal cameras, there is a noticeable decrease in research articles. This may result from the beginnings of this detection technique and its initial development. Considering the trends in mhealth, artificial intelligence, device miniaturization, and improving their parameters, the author observes the necessity of creating an interdisciplinary research foundation for mental and physical health monitoring purposes. When analyzing research works using a smartphone thermal imaging camera, it is noticeable that researchers utilize the readily available but unreliable FLIR ONE camera In their study [50], the researchers tracked breathing patterns and temperature changes around the nostrils of 8 participants. They used FLIR ONE, respiratory variability spectrogram (RVS), the Stroop Color-Word Test (SCWT), and computer-based mental tasks, as well as CNN networks. In binary classification, they achieved 84.59%, while in 3-class classification - 56.52%. In [51], FLIR ONE and PPG were also used. The group consists of 17 people. The stressor was the Stroop test and mathematical tasks. The VAS (Visual Analog Scale) was also used. Parameters were correlated, k-means clustering was applied, and the LOSO method was used. Ultimately, an accuracy of 78.33% was achieved. In the works [52, 53, 54], the FLIR ONE camera was also used. In [52], 10 participants solved mathematical tasks while researchers monitored the thermal variability of the nose and calculated the thermal variability metric. They proposed the Thermal Gradient Flow method. In [53], correlations between thermal changes dynamics and physiological and psychological signals were investigated. The dynamics of changes in a selected color were analyzed using OpenCV. The stressor was a stressful video. In [54], researchers analyzed the nostril area and measured the frequency of breaths using a thermal camera and a breath sensor Depending on the stage of the study, the size of the research group varied. The tasks performed by the participants were various, including walking and reading news. Analyzing the presented work, one can observe that the topic of stress detection and classification has many areas still in development. It would be worthwhile to consider the diversity of research approaches, research conditions, the use of data sets (pre-existing or self-acquired), as well as the accuracy of detection systems in laboratory and real-world conditions. Certainly, in the coming years, due to increasing social concerns, stress will continue to rise, along with the demand for modern stress detection and control systems. As for the correlation of stress detection methods, ensemble learning has recently seen a growth in interest. Several stress detection models were developed at work [55] using ensemble learning. They were based on a video-based plethysmographic application that analyzes a person's face and captures physiological signals. Additionally, information from a research questionnaire was collected. The pilot study lasted for 3 months and involved 28 individuals. Ensemble learning allowed for achieving an accuracy of 86.8% and an F1 score of 87% in binary prediction: stress/no stress. Researchers in [56] focused heavily on anxiety disorders, and thus psychological disorders. They proposed an automatic and intelligent system for identifying anxiety based on physiological signals. Basic machine learning algorithms were tested, and then compared with ensemble learning models for better metric identification. In the work [57], five algorithms were used for stress classification: random forest, SVM, decision tree, logistic regression, and naive Bayes algorithm. Then, the ensemble learning algorithm was applied, achieving an accuracy of 94.25%. It is worth noting that stress was detected through behavior during sleep. In stress research using EEG [58], researchers proposed a CIS-based KNORA-U dynamic ensemble selection (DES) model for human stress identification. They utilized, for instance: discrete wavelet transform (DWT), extreme gradient boosting (XGBoost), and linear discriminant analysis (LDA), extra tree (ET). They achieved 99.14% accuracy. If we analyze works related to thermal stress detection, noticeable is the predominant use of transfer learning [59, 60].

3 Research project and dataset

As part of the research program *Recognition of Stress Using Thermography and Neural Networks*, stress studies were conducted among college students. A total of 100 individuals aged 20-24 participated in the research sessions. Twelve individuals reported visual impairments such as astigmatism, nearsightedness, and farsightedness. However, during thermographic recording, these individuals wore contact lenses instead of glasses. The research was conducted in laboratory conditions. For research purposes, a windowless laboratory room with subdued lighting was adapted. The study was carried out according to the research procedure specified in the research project approved by the Ethics Committee. The examination involved recording the tested participants using an InfiRay smartphone thermal camera during exposure to a stressor (the distance between the participant and the camera was 50 cm). For the purpose of obtaining a diverse database for further multiclassification, the subjects were exposed to various stressors The entire group was divided into 4 teams, each with a different stressor: a stressful film, an arithmetic task, the Stroop test, and a recording of a heated argument. Due to the specific requirements of thermal cameras, the research room was maintained at a constant temperature, free from drafts, reflections, etc. Additionally, during recordings, the blood pressure and pulse of the subjects were measured for control and observation purposes. The acquired recordings were processed to extract images. Then these images were compared with reference images of the subjects, in a relaxed state, in order to divide them appropriately according to the specified stress level. It is worth noting that the smartphone thermal camera featured adjustable focus, thermal sensitivity of $\leq 60\,\mathrm{mK}$, thermal accuracy of $2\% \pm 2\%$, a refresh rate of $25\,\mathrm{Hz}$, and a field of view of $44.9^\circ \times 33.4^\circ$.

An important aspect of the work was the processing of thermographic images. The first step was to scale the images to a size of 256 x 256. It was decided to use the standard value in order to avoid further loss of information. Each photo was scaled and adjusted to the accepted size. This is defined by the mathematical formula:

$$l(x,y) = \sum_{i-0}^{3}\sum_{j-0}^{3}(\boldsymbol{a}_{ij} m^i n^j) \tag{3}$$

where:
n - pixel height,
m - pixel width,
a_{ij} - square area of the images.

The next steps in image processing were: transforming thermal images to grayscale using a linear function, contrast stretching, and image sharpening. Then, based on the observation of physiological results, expert observation, and analysis of thermal image histograms, the data was divided into three stress levels: low, medium, and high. Additionally, stress absence was taken into account. In specifying stress levels, the works [61, 62, 63] were helpful in determining stress levels. Additionally, after registration of the research participants (during exposure to the stressor), they were asked about the level of stress they felt. The PSS-10 psychological test conducted during the study was also helpful. Also important was control emotions on participants face to additional control predtiction of stress level. Positive emotions corelate with low stress level, negative emotions - with high stress level. After many thermal analyses, psychological consultations, reviewing scientific papers, and a thorough analysis of the feelings of the participants, the following labels were assigned: no stress, low stress, medium stress, high stress. These levels mainly referred to the size of the intensity changes in the designated ROIs: forehead, nasal septum, cheeks (left, right), eye area (left, right), jaw with mouth. The following percentage ranges of pixel intensity changes for stress levels (relative to the value in the relaxed state) were adopted: no stress 0-10, low stress 10-30, medium stress 30-50, high stress 50-100. The designated ROIs are the areas that provide the most information about thermal changes. In these areas, temperature changes due to stress are most often noticeable, visible in the thermal image through a change in color or its intensity.

In the next part of preparations for neural calculations, the size of the datasets was balanced so that each class contained 5000 images. Next, for neural calculations, the data was divided into training, testing, and validation sets using an 80-10-10 split. All coding operations were done in Google Colab: backend - Keras with TensorFlow, disk space - 78GB, GPU RAM - 15GB, Windows 11, Python - version 3.10. Table 1 presents information about the basic parameters used during model training. Figure 1 shows sample thermal images of faces from our own database (before pre-processing).

Table 1. Basic parameters using during model training.

Parameter	Value
Epochs	50
Batch size	32
Learning Rate	0.001
Optimizer	Adam
Loss Function	Categorical Cross Entropy
Weights	Imagenet

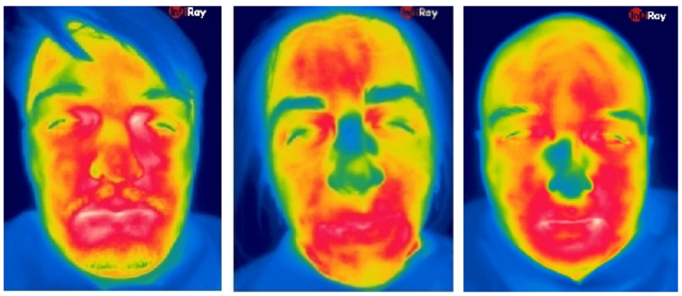

Figure 1. Thermal images of faces - examples (before pre-processing)

4 Proposed Method

The classification based on the implementation of Choquet integral within ensemble methods for integrating the results of lightweight convolutional networks: MobileNet, EfficientNet, Xception, is presented in this work. Ensemble methods involve combining different networks and classifiers, taking into account all uncertainties in order to make a final classification decision. In the case of lightweight CNNs, additional layers were added to their structure: Global Average Pooling, Batch Normalization, Dropout, and Dense. GAP aimed to reduce the spatial dimensions of object maps while preserving their spatial information. This helped decrease the number of parameters and improve the model's generalization ability. The Batch Normalization layer was designed to improve the training of deep neural networks by reducing the internal variation of the co-variables that can occur during training. Batch normalization normalizes the input data for each layer so that they have zero mean and unit variance. Dropout is responsible for regularization and preventing overfitting in a neural network. It is applied to the outputs of the neural network layer and helps reduce co-adaptation of neurons in the network. The Dense layer can learn complex relationships between inputs and outputs In the next stage, fuzzy ensemble learning with Choquet integral is implemented. This allows for data integration, leading to final classification Figure 2 illustrates the concept of the proposed method's architecture.

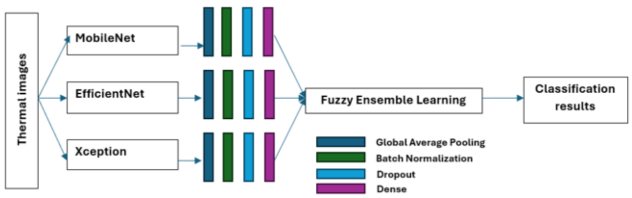

Figure 2. Architecture of proposed model.

5 Results

In Tables 2-5, classification reports have been compiled. As can be observed, the use of fuzzy ensemble learning improved the results of model quality metrics.

Table 2. Classification Report with Fuzzy Ensemble Learning

Category	Precision	Recall	F1-score	Support
No-stress	0.98	0.97	0.98	1000
Low-stress	0.98	0.99	0.99	1000
Medium-stress	0.90	0.92	0.91	1000
High-stress	0.94	0.90	0.92	1000
Accuracy			0.95	4000
Macro avg	0.95	0.95	0.95	4000
Weighted avg	0.95	0.95	0.95	4000

Analyzing the metric values in Tables 2-5, it can be observed that low stress and no stress were classified the best. Another levels of stress obtained lower metric values, which is associated with greater uncertainties in image recognition and assignment to the appropriate group. In the case of the fuzzy ensemble learning model, an accuracy of 95% was achieved, which is a very good result. Classification by single lightweight CNN model is associated with an accuracy level of 89-96%. The difference between the accuracy values in both cases

Table 3. Classification Report for MobileNet

Category	Precision	Recall	F1-score	Support
No-stress	0.96	0.94	0.95	1000
Low-stress	0.94	0.99	0.96	1000
Medium-stress	0.83	0.81	0.82	1000
High-stress	0.85	0.83	0.84	1000
Accuracy			0.89	4000
Macro avg	0.89	0.89	0.89	4000
Weighted avg	0.89	0.89	0.89	4000

Table 4. Classification Report for EfficientNet

Category	Precision	Recall	F1-score	Support
No-stress	0.96	0.96	0.96	1000
Low-stress	0.96	0.99	0.98	1000
Medium-stress	0.83	0.85	0.84	1000
High-stress	0.88	0.82	0.85	1000
Accuracy			0.91	4000
Macro avg	0.91	0.91	0.91	4000
Weighted avg	0.91	0.91	0.91	4000

confirms the validity of the proposed method. However, it is worth emphasizing that it would be beneficial to work on improving the classification of images depicting a stress-free state and those with moderate stress. One can consider using different classifiers or a different aggregation function for this purpose.

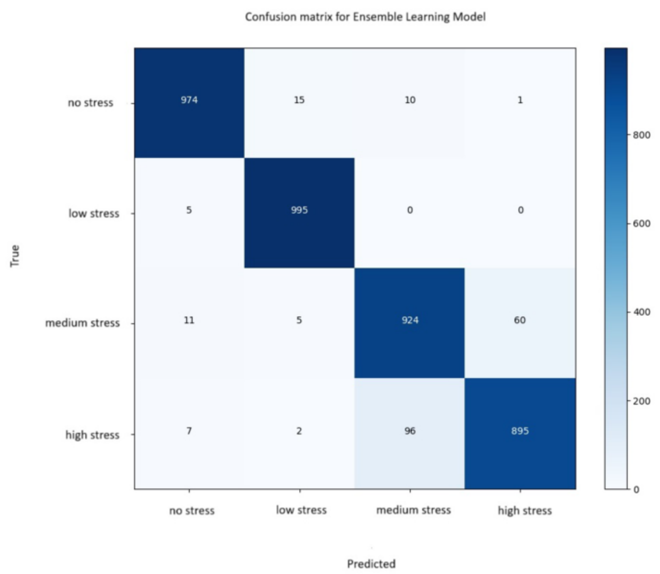

Figure 3. Confusion matrix

On Figure 3, graphical confusion matrix for Ensemble Learning model was presented. The best prediction was for "low stres" class. Also, "no stress" class has high true prediction. Some bad predictions are for "medium stress" and "high stress" - the reason may be the greater level of diversity of photos in these classes, which is related to the variety of thermal patterns on the human face (individual human features). Thermal patterns in photos of faces without stress coincide between people. In the case of low stress, a change in thermal patterns is noticed, especially around the eyes, mouth and nose.

Table 5. Classification Report for Xception

Category	Precision	Recall	F1-score	Support
No-stress	0.98	0.98	0.98	1000
Low-stress	0.99	1.00	0.99	1000
Medium-stress	0.92	0.93	0.93	1000
High-stress	0.95	0.92	0.93	1000
Accuracy			0.96	4000
Macro avg	0.96	0.96	0.96	4000
Weighted avg	0.96	0.96	0.96	4000

In the case of medium or high stress, thermal changes are more dynamic and may differ from person to person, for example: in one person, due to high stress, a thermal change will be noticed in the forehead, and in another person - in the jaw. So, report classification results and confusion matrix confirm this statement.
In order to check the model, it was decided to test it also on other databases. Due to the lack of public availability of thermal images databases of stressed people taken with a smartphone thermal camera, it was decided to use databases with images taken with industrial versions of thermal cameras. The following databases were used for comparative studies: DBnew [64], Tufts Face Database [65], SF-TL54 [66]. DBnew contains images of faces of people in various emotional states: neutral, happy, sad. Tufts Face Databases includes 4 facial expressions: neutral, smiling, closed eyes, shocked. SF-TL54 contains thermal images with marked facial landmarks and pairing with a visual image. Before testing the model, images from individual databases were subjected to pre-processing and labeling described in Chapter 3. The same number of input images was also taken care of using data augmentation. The division of images by emotion facilitated correlation with the stress level. It is worth noting that individual databases were created based on different procedures and guidelines, e.g.: different distance of the camera from the subject, registration of the head and not just the face, different angle of the face. Due to the fact that only the face was taken into account in the author's database, when testing on other databases, the photos were modified so that they also showed only the face. Table 6 summarizes the accuracy results of the proposed model obtained for individual databases.

Table 6. Model accuracy for datasets - comparison

Dataset	Accuracy
DBnew	0.78
Tufts Face Database	0.91
SF-TL54	0.90
Author dataset	0.95

Comparing the obtained accuracy results from Table 6, it can be stated that the model provides high classification accuracy also on databases other than the author database. However, it is worth emphasizing that the precise and thorough pre-processing of thermal images plays an important role, which translates into the enhancement of details and ultimately into the recognition and classification of the image. Thermal registration is also important - the distance of the examined person from the thermal imaging camera, the type and technical parameters of the thermal imaging camera, as well as the area of body being registered together with the positioning of the camera in relation to the body. Thermal imaging stress detection has many advantages, including: non-invasiveness, ease of use, comfort of examination, safety, mobility - the ability to record in various conditions (laboratory, home, outdoor, indoor - with appropriate setting of recording parameters. However, there are also disadvantages of this method, especially when using smartphone cameras, for example: the need for regular calibration to ensure the accuracy of measurements, the need for complex data interpretation if the examined person has a disease factor or physical exertion (the need to use advanced algorithms to distinguish the cause of a given thermal change - stress or illness), limited accuracy depending indirectly on the distance between the recorded object and the camera (optimally 50 cm), as well as on the correctness of the recording settings (emission, palette, distance, etc.). However, with the currently still developing machine learning algorithms, the disadvantages of this method can be eliminated. This requires increasingly complex computer analysis and a thorough evaluation of the obtained results.

6 Conclusion

Continuous development of machine learning and improvement of research devices allows for the advancement of stress detection and classification techniques. Using single neural network models for stress recognition is problematic due to the potential for model generalization and lower classification efficiency. Basic CNNs have many parameters and hardware requirements, which results in time-consuming and computationally intensive calculations. Therefore, to minimize this, in this article, a fuzzy ensemble method based on lightweight CNNs was applied. Lightweight CNNs such as MobileNet, EfficientNet, and Xception were used as base models, while Choquet fuzzy integral was used to integrate the results for the final classification of human stress from thermographic images. The results indicate that the proposed method achieves 95% accuracy.

It is worth mentioning that thermographic registration of the subjects was associated with strict adherence to the research procedure and ensuring almost ideal measurement conditions. Additionally, due to its specific nature, the smartphone thermographic camera required frequent calibration and test recordings. Equally important was familiarizing the participants with the research setting and avoiding any additional stress. Therefore, before the studies, the participants underwent relaxation.

In the context of the conducted calculations, it has been noticed that lightweight CNN networks perform well with not very large datasets and complex tasks. For greater complexity and large datasets, the consideration remains to use larger and deeper neural networks. Furthermore, it would be worthwhile to perform data augmentation on the dataset to ensure greater diversity and a larger quantity of images for neural computations.

The use of Choquet integral and fuzzy ensemble learning is considered effective in the context of stress classification. However, it would be worth considering the use of other classifiers, taking into account the test accuracy. The research space also includes searching for optimal weights or voting schemes to combine the results of individual models. The ensemble performance may be sensitive to the choice of weights or voting scheme. Therefore, it is worth creating increasingly advanced neural networks, focusing on more accurate data processing, and developing the idea of fuzzy ensemble learning in further research. In the future, the author plans to conduct advanced stress studies, focused on further correlating thermography with neural networks. However, it should move towards cost-effectiveness, adaptation to average computer resources, integration with mHealth solutions, and comprehensibility for the average person.

References

[1] Seyle, H. (1984). The Stress of Life. New York: McGraw-Hill Company, 1984.

[2] Rochette, L., Dogon, G., and Vergely, C. (2023). Stress: Eight Decades after Its Definition by Hans Selye:"Stress Is the Spice of Life". Brain Sciences, 13(2), 310.

[3] Selye, H. (1975). Confusion and controversy in the stress field. Journal of human stress, 1(2), 37-44.

[4] Selye, H. (1951). The general-adaptation-syndrome. Annual review of medicine, 2(1), 327-342.

[5] Hobfoll S.E. (1989). Conservation of resources: A new attempt at conceptualizing stress. American psychologist, 44(3), 513-524.

[6] Castro-Garcia, J. A., Molina-Cantero, A. J., Gomez-Gonzalez, I. M., Lafuente-Arroyo, S., and Merino-Monge, M. (2022). Towards human stress and activity recognition: A review and a first approach based on low-cost wearables. Electronics, 11(1), 155.

[7] Lu, S., Wei, F., and Li, G. (2021). The evolution of the concept of stress and the framework of the stress system. Cell stress, 5(6), 76.

[8] Sharma, S., Singh, G., and Sharma, M. (2021). A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans. Computers in Biology and Medicine, 134, 104450.

[9] Swanson, G. (2023). The Anxiety and Panic Attacks Toolbox of Cures: The ani-anxiety healers guide book and first aid kit for extreme anxiety relief, cbt, diet, plus a survival quide toolkit for teens, adults and kids. Z Media.

[10] Attia, M., Ibrahim, F. A., Elsady, M. A. E., Khorkhash, M. K., Rizk, M. A., Shah, J., and Amer, S. A. (2022). Cognitive, emotional, physical, and behavioral stress-related symptoms and coping strategies among university students during the third wave of COVID-19 pandemic. Frontiers in psychiatry, 13, 933981.

[11] Romain, T., and Verdick, E. (2018) Stress can really get on your news. Free Spirit Publishing.

[12] Bienertova-Vasku, J., Lenart, P., and Scheringer, M. (2020). Eustress and distress: neither good nor bad, but rather the same?. BioEssays, 42(7), 1900238.

[13] Schulz, A., Schultchen, D., and Vögele, C. (2021). Interoception, stress, and physical symptoms in stress-associated diseases. European Journal of Health Psychology.

[14] Notaras, M., and van den Buuse, M. (2020). Neurobiology of BDNF in fear memory, sensitivity to stress, and stress-related disorders. Molecular psychiatry, 25(10), 2251-2274.

[15] Han, X., Shen, Q., Hou, C., Yang, H., Chen, W., Zeng, Y., et al. (2024). Disease clusters subsequent to anxiety and stress-related disorders and their genetic determinants. Nature Communications, 15(1), 1-13.

[16] Verma, H., Kumar, N., Sharma, Y. K., and Vyas, P. (2024). StressDetect: ML for Mental Stress Prediction. Optimized Predictive Models in Health Care Using Machine Learning, 341.

[17] Sun, M., and Cao, X. (2024). A Mental Stress Classification Method Based on Feature Fusion Using Physiological Signals. Journal of Circuits, Systems and Computers, 33(01), 2450016.

[18] Fontes, L., Machado, P., Vinkemeier, D., Yahaya, S., Bird, J. J., and Ihianle, I. K. (2024). Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques. Sensors, 24(4), 1096.

[19] Sharma, R., and Chopra, K. (2020). EEG signal analysis and detection of stress using classification techniques. Journal of Information and Optimization Sciences, 41(1), 229-238.

[20] Pourmohammadi, S., and Maleki, A. (2020). Stress detection using ECG and EMG signals: A comprehensive study. Computer methods and programs in biomedicine, 193, 105482.

[21] Al-Jumaily, A. A., Matin, N., and Hoshyar, A. N. (2021). Machine learning based biosignals mental stress detection. In Soft Computing in Data Science: 6th International Conference, SCDS 2021, Virtual Event, November 2–3, 2021, Proceedings 6 (pp. 28-41). Springer Singapore.

[22] Khowaja, S. A., Prabono, A. G., Setiawan, F., Yahya, B. N., and Lee, S. L. (2021). Toward soft real-time stress detection using wrist-worn devices for human workspaces. Soft Computing, 25, 2793-2820.

[23] Chen, J., Abbod, M., and Shieh, J. S. (2021). Pain and stress detection using wearable sensors and devices—A review. Sensors, 21(4), 1030.

[24] Alexander, Q. G., and Lunderman, C. V. (2021). Thermal camera reliability study: FLIR One Pro.

[25] Baran, K. (2024). Application of thermal imaging cameras for smartphone: Seek Thermal Compact Pro and Flir One Pro for human stress detection - comparison and study. Applied Computer Science, 20(1), 122-138.

[26] Li, F., Wang, M., Wang, T., Wang, X., Ma, X., He, H., et al. (2023).

Smartphone-based infrared thermography to assess progress in thoracic surgical incision healing: a preliminary study. International Wound Journal, 20(6), 2000-2009.
[27] Li, F., Wang, M., Wang, T., Wang, X., Ma, X., He, H., et al. (2023). Smartphone-based infrared thermography to assess progress in thoracic surgical incision healing: a preliminary study. International Wound Journal, 20(6), 2000-2009.
[28] Dougherty, E. R. (2020). Digital image processing methods. CRC Press.
[29] Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33(12), 6999-7019.
[30] Liu, Y., Xue, J., Li, D., Zhang, W., Chiew, T. K., and Xu, Z. (2024). Image recognition based on lightweight convolutional neural network: Recent advances. Image and Vision Computing, 105037.
[31] Gawande, N. A., Daily, J. A., Siegel, C., Tallent, N. R., and Vishnu, A. (2020). Scaling deep learning workloads: Nvidia dgx-1/pascal and intel knights landing. Future Generation Computer Systems, 108, 1162-1172.
[32] Mohammed, A., and Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences, 35(2), 757-774.
[33] Zhang, X., Wang, J., Zhan, J., and Dai, J. (2021). Fuzzy measures and Choquet integrals based on fuzzy covering rough sets. IEEE Transactions on Fuzzy Systems, 30(7), 2360-2374.
[34] Grabisch, M. (2000). A graphical interpretation of the Choquet integral. IEEE Transactions on Fuzzy Systems, 8(5), 627-631.
[35] Bara, C. P., Papakostas, M., and Mihalcea, R. (2022). A Deep Learning Approach Towards Multimodal Stress Detection. AffCon@ AAI 67-81.
[36] Kaga, S., and Kato, S. (2019, March). Extraction of useful features for stress detection using various biosignals doing mental arithmetic. In 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech) (pp. 153-154). IEEE.
[37] Liu, X., Xiao, X., Cao, R., and Chen, T. (2020, April). Evolution of facial tissue oxygen saturation and detection of human physical stress. In 2020 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) (pp. 144-147). IEEE.
[38] Cruz-Albarran, I. A., Rodriguez-Medina, D. A., Leija-Alva, G., Dominguez-Trejo, B., Osornio-Rios, R. A., amd Morales-Hernandez, L. A. (2020). Physiological stressor impact on peripheral facial temperature, Il-6 and mean arterial pressure, in young people. Journal of Thermal Biology, 91, 102616.
[39] Sharma, N., Dhall, A., Gedeon, T., and Goecke, R. (2013, September). Modeling stress using thermal facial patterns: A spatio-temporal approach. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (pp. 387-392). IEEE.
[40] Akbar, F., Bayraktaroglu, A. E., Buddharaju, P., Da Cunha Silva, D. R., Gao, G., Grover, T., et al. (2019, May). Email makes you sweat: Examining email interruptions and stress using thermal imaging. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-14).
[41] Cardone, D., Perpetuini, D., Filippini, C., Spadolini, E., Mancini, L., Chiarelli, A. M., and Merla, A. (2020). Driver stress state evaluation by means of thermal imaging: A supervised machine learning approach based on ECG signal. Applied Sciences, 10(16), 5673.
[42] Liu, X., Shan, Y., Peng, M., Chen, H., and Chen, T. (2020). Human stress and StO2: Database, features, and classification of emotional and physical stress. Entropy, 22(9), 962.
[43] Hong, K., and Hong, S. (2016). Real-time stress assessment using thermal imaging. The Visual Computer, 32(11), 1369-1377.
[44] Pereira, N., and Hallock, G. G. (2021). Smartphone thermography for lower extremity local flap perforator mapping. Journal of reconstructive microsurgery, 37(01), 059-066.
[45] Goel, J., Nizamoglu, M., Tan, A., Gerrish, H., Cranmer, K., El-Muttardi, N., et al. (2020). A prospective study comparing the FLIR ONE with laser Doppler imaging in the assessment of burn depth by a tertiary burns unit in the United Kingdom. Scars, Burns and Healing, 6, 2059513120974261.
[46] Van Doremalen, R. F. M., Van Netten, J. J., Van Baal, J. G., Vollenbroek-Hutten, M. M. R., and van der Heijden, F. (2019). Validation of low-cost smartphone-based thermal camera for diabetic foot assessment. Diabetes research and clinical practice, 149, 132-139.
[47] Al Husaini, M. A. S., Hadi Habaebi, M., Gunawan, T. S., and Islam, M. R. (2021). Self-detection of early breast cancer application with infrared camera and deep learning. electronics, 10(20), 2538.
[48] Chung, H., Lee, J. H., Chung, J. H., Hong, J. H., Kang, S. G., Lee, K. S., and Kim, Y. M. (2023). Camera and IR Sensor based Sleep Apnea Diagnosis Device. IEEE Sensors Journal.
[49] Kevin, C., and Theuma, F. (2019). Smartphone-Mounted Infrared Thermography Camera to Assess Success of Open Infrainguinal Revascularisation. European Journal of Vascular and Endovascular Surgery, 58(6), e164-e165.
[50] Cho, Y., Bianchi-Berthouze, N., and Julier, S. J. (2017, October). DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In 2017 Seventh international conference on affective computing and intelligent interaction (acii) (pp. 456-463). IEEE.
[51] Cho, Y., Julier, S. J., and Bianchi-Berthouze, N. (2019). Instant stress: detection of perceived mental stress through smartphone photoplethysmography and thermal imaging. JMIR mental health, 6(4), e10140.
[52] Cho, Y., Bianchi-Berthouze, N., Oliveira, M., Holloway, C., and Julier, S. (2019, September). Nose heat: Exploring stress-induced nasal thermal variability through mobile thermal imaging. In 2019 8th international conference on affective computing and intelligent interaction (ACII) (pp. 566-572). IEEE.
[53] Baran, K. (2021). Stress detection and monitoring based on low-cost mobile thermography. Procedia Computer Science,Proceedings KES2021, 192, 1102-1110.
[54] Cho, Y., Julier, S. J., Marquardt, N., and Bianchi-Berthouze, N. (2017). Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging. Biomedical optics express, 8(10), 4480-4503.
[55] Rodrigues, F., and Correia, H. (2024). Semi-supervised and ensemble learning to predict work-related stress. Journal of Intelligent Information Systems, 62(1), 77-90.
[56] Khullar, V., Tiwari, R. G., Agarwal, A. K., and Dutta, S. (2022). Physiological signals based anxiety detection using ensemble machine learning. In Cyber Intelligence and Information Retrieval: Proceedings of CIIR 2021 (pp. 597-608). Springer Singapore.
[57] Jayawickrama, J. G., and Rupasingha, R. A. H. M. (2022, December). Ensemble Learning Approach to Human Stress Detection Based on Behaviours During the Sleep. In 2022 4th International Conference on Advancements in Computing (ICAC) (pp. 132-137). IEEE.
[58] Malviya, L., and Mal, S. (2023). CIS feature selection based dynamic ensemble selection model for human stress detection from EEG signals. Cluster Computing, 26(4), 2367-2381.
[59] He, Y., Deng, B., Wang, H., Cheng, L., Zhou, K., Cai, S., and Ciampa, F. (2021). Infrared machine vision and infrared thermography with deep learning: A review. Infrared physics and technology, 116, 103754.
[60] Arasu, D. B. L., AzlanMohamed, A. S., Ruhaiyem, N. I. R., Annamalai, N., Lutfi, S. L., and Al Qudah, M. M. (2022). Human Stress Recognition from Facial Thermal-Based Signature: A Literature Survey. CMES-Computer Modeling in Engineering and Sciences, 130(2).
[61] Ribeiro, G., Postolache, O., and Martin, F. F. (2024). A New Intelligent Approach for Automatic Stress Level Assessment Based on Multiple Physiological Parameters Monitoring. IEEE Transactions on Instrumentation andd Measurement, 73, 1-14.
[62] Resendiz-Ochoa, E., Cruz-Albarran, I. A., Garduno-Ramon, M. A., Rodriguez-Medina, D. A., Osornio-Rios, R. A., and Morales-Hernandez, L. A. (2021). Novel expert system to study human stress based on thermographic images. Expert systems with applications, 178, 115024.
[63] Saha, S., and Ghosh, L. (2024). Predicting stress levels for smartphone users using transfer learning induced residual net. Entertainment Computing, 48, 100609.
[64] Pavel, S., Moldovanu, S., and Aiordachioaie, D. (2023). Emotion Recognition in Human Thermal Images with Artificial Intelligence Technology. 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (EFTA), Sinaia, Romania, 1-5.
[65] Panetta, K., Wan, Q., Agaian, S., Rajeev, S., Kamath, S., Rajendran, R., Rao, S. P., Kaszowska, A., Taylor, H. A., Samani, A., and Yuan, X. (2020). A Comprehensive Database for Benchmarking Imaging Systems. IEEE Transactions on Pattern Analysis and Machine Inteligence, 42, 3, 509-520.
[66] Kuzdeuov, A., Koishigarina, d., Aubakirova, D., Abushakimova, S., and Varol, H. A. (2022). SF-TL54: A Thermal Facial Landmark Dataset with Visual Pairs, 2022 IEEE/SICE International Symposium on System Integration (SII), Narvik, Norway,748-753.

Acknowledgements

The research project titled: *"Recognition of human stress using thermographic imaging and neural networks"* was approved by the Commission for Research Ethics, No. 4/2024 dated 19 February 2024.