Andreas, Andreou, Constandinos X. Mavromoustakis, Houbing Song, and Jordi Mongay Batalla. “Optimisation of CNN through Transferable Online Knowledge for Stress and Sentiment Classification.” IEEE Transactions on Consumer Electronics, 2023, 1–1. https://doi.org/10.1109/TCE.2023.3319111.

https://doi.org/10.1109/TCE.2023.3319111

This article has been accepted for publication in IEEE Transactions on Consumer Electronics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCE.2023.3319111

Optimisation of CNN through Transferable Online Knowledge for Stress and Sentiment Classification

Andreou Andreas, *Graduate Student Member*, *IEEE*, Constandinos X. Mavromoustakis, *Senior Member*, *IEEE*, Houbing Song, *Fellow*, *IEEE*, and Jordi Mongay Batalla, *Member*, *IEEE*

***Abstract*—As we stand on the cusp of an evolution in effective and confidential smart healthcare systems, the disciplines of Psychology and Neuroscience remain a barrier. The obstacle of stress and sentiment classification is an enduring challenge to the field. Therefore, this research identifies mental health by analysing and interpreting acquired biological data. By employing convolutional neural networks in conjunction with transfer learning, the article seeks to leverage physiological signs driven by sensors for health monitoring. More precisely, we elaborate on the correlation between vital signs, arousal, and vigour data to classify a person's sentimental state. A novel algorithmic methodology is proposed in which the source and target domains are leveraged adaptively by homogeneous and heterogeneous transfer learning. A comprehensive analysis of the outcomes from real-world acquired datasets was performed to demonstrate the proposed method's effectiveness compared to state-of-the-art classification techniques in the field.**

***Index Terms*— stress, sentiment, online transfer learning, convolutional neural network**

I. Introduction

STRESS is defined as the feeling of being overwhelmed due to the reaction of the individual immune system to external factors or the inability to cope with mental or sentimental pressure. Therefore, emotional trauma stimulates various sentimental and physical behaviours in stressful situations. The immune system's effectiveness decreases as psychological inflammation occurs in some people. Hence, the immune resistance of people with stress becomes limited against many diseases, such as hypertension, heart failure or diabetes and the standard hormone-related measures are followed. Unobtrusive factors such as variability in respiratory rate, respiratory habits, or skin temperature are functional and practical for quantitative analysis of stress and sentimental interpretation. It is contrary to common approaches that use mainly facial gestures, auditory or sentimental variants, and variations in behavioural habits [1].

A person's discretion cannot realise the onset of stress due to the pressure exerted by his environment. Mainly stress occurs due to feelings of loneliness or deficiency of self-confidence. In addition, acute stressful challenges and negative bias often lead to stress. Therefore, psychological changes affecting emotional and cognitive behaviour are exacerbated by this inflammatory activity [2]. For this reason, the advanced respiratory system regulates hormone levels under normal conditions to maintain human defence and heart performance [3]. In general, non-invasive techniques have been used to predict and calculate hormonal secretion, including evaluating biological signals emanating from different sentiments. Thus, it has led to automated systems classifying stressful emotions [4]. However, its integration remains an unsolved problem that concerns researchers in the field.

Intelligent computing for big data in consumer Internet of Things (IoT) refers to the evolution of advancing machine learning and artificial intelligence techniques. However, processing and analysing massive amounts of data generated by consumer IoT devices leads to system fragments and unclear interpretations, especially in psychology, where the datasets are mostly fuzzy and have essential outliers. Hence, this article elaborates on the correlation between vital signs, arousal, and vigour data to classify a person's sentimental state [5]. Furthermore, a novel algorithmic methodology was proposed in which homogeneous and heterogeneous transfer learning leveraged the source and target domains adaptively.

Predominantly, the automation of recognising the feelings of stress is achieved by measurements taken from the skin and the nerve impulses of the spinal cord. The data are mainly acquired from ElectroEncephaloGraphy (EEG), skin conductance response [6] and heart rate. Using an in-depth learning technique that recognises emotional stress, the researchers compiled the research paper [7], where they identified differences in the harshness of the stress. A dedicated

This research work was funded by the Smart and Healthy Ageing through People Engaging in supporting Systems SHAPES project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857159. The work of the Polish researcher has been supported by the Eureka CELTIC-NEXT project "Intelligent Management Of Next Generation Mobile Networks And Services" (IMMINENCE, C2020/2-2), which has been funded in Poland by Narodowe Centrum Badań i Rozwoju (National Centre for Research and Development). This work was supported in part by the U.S. National Science Foundation under Grant No. 2317117.

A. Andreou is with the Department of Computer Science, University of Nicosia and University of Nicosia Research Foundation, Nicosia, Cyprus, (e-mail: andreou.andreas@unic.ac.cy).

C. X. Mavromoustakis is with the Department of Computer Science, University of Nicosia and University of Nicosia Research Foundation, Nicosia, Cyprus, (e-mail: mavromoustakis.c@unic.ac.cy).

H. Song is with the Department of Information Systems, University of Maryland, Baltimore County (UMBC), Baltimore, MD 21250 USA (e-mail: h.song@ieee.org)

J. M. Batalla is with Warsaw University of Technology, Warsaw, Poland, (e-mail: jordi.mongay.batalla@pw.edu.pl)

transceiver was installed to detect the rate of respiratory fluctuations in response to skin conductance, aiming to detect abnormalities in respiratory function. Also, a set of methods with psychological work from cognitive tests with varying degrees of limitation has been created. Patients experience anxiety that varies in severity when faced with these boundaries. At the baseline and end of the pilot, a cognitive test was performed, and it suggested a deep breathing exercise that should be performed to relieve the person from stress. The CNN method was deployed to train and process the collected data to classify their stress. However, the data processing process required much financial support, an obstacle to further development.

Different strategic approaches to prevent stress require effectively classifying the sentiments that lead to it. Several Deep Transfer Learning (DTL) methods have been developed to classify sentimental stress. DBNTL is one of the primary methods to effectively structure essential generative models to learn domain-specific attributes utilising training data. Nevertheless, the alternation between the distribution of stress and sentiment at different levels has been reduced through this method. Therefore, OTLCNNO has proposed implementing sentiment and stress data classification to learn high-level attributes at top layers. The model aims to improve transitioning along with the domains of stress and sentiment; it determines the Joint Distribution Discrepancy (JDD) of distinct layers and Marginal Distribution Discrepancy (MDD) at similar layers. Although, it consumes too much time for data acquisition from the multi-sensor network to analyse and interpret this data [8]. In addition, it requires a constant and gradually changing input distribution and cannot handle the abrupt shift of the concept in real-time. Hence, through this research work, we aim to perform a novel Online Transfer Learning (OTL) by converting the acquaintance from stress and sentiment domains.

The article elaborates on a novel CNN model via TL to achieve real-time stress and sentiment classification. The novelty of the proposed model lies in evaluating and interpreting biometric data obtained in emotional characteristics and classifying them in the field of stress. Furthermore, the model investigates the scenarios of OTL on homogeneous and heterogenous domains of common and varied attributes, respectively. The OTL with streaming data was considered to resolve the classification by the OTLCNNO classifier in the scenarios mentioned above. In addition, the co-regularisation learning technique for transferable knowledge effectively addresses the learning process in various attribute areas. The learning technique effectively addresses the learning process across multiple attribute areas. Thus, it combines the optimised CNN classifiers co-trained from different points of view of similar training cases for transferable knowledge.

Consequently, we depict our model's performance efficiency according to precision, recall and accuracy metrics. Also, by F_1-score, we obtain a measurement that equilibrium precision and recall in a single value. Hence, we validate the outperformance of the proposed OTLCNNO compared with state-of-the-art classifiers in the field using the SWELL knowledge work (SWELL-KW) and the Wearable Stress and Affect Detection (WESAD) dataset for [38], [39].

A. Motivation

Neural Networks, or NNs are divided into Biological Neural Networks (BNNs) and Artificial Neural Networks (ANNs). BNNs are part of the central nervous system of biological systems, for example, humans, and consist of biological tissue, chemicals and electrical signals that separate them from ANNs that try to mimic the former through a set of electronic and mechanical systems accompanied by intelligent algorithms.

Artificial Neural Networks essentially try to mimic the function of the human brain by adopting the architecture that characterises it. At the same time, their education is done as in the BNN, i.e., through a set of examples. Thus, the standard features of the two categories of intelligence are the following:

- They are trained through a set of criteria, helping them to learn their environment
- The synapse points of the BNNs, depending on the strength of their bonds and the synaptic weights in the ANNs, are used to store the acquired knowledge.

B. Novelty & Contribution

This research aims to develop a CNN model that can accurately classify human stress and emotion by optimising the model through transferable online knowledge. It is a novel and promising area of research that has the potential to impact the field of mental health significantly. CNN is a deep learning architecture widely used in computer vision, natural language processing, and speech recognition. However, applying CNNs to psychology is relatively new, and the concept of transferable online knowledge is particularly innovative.

The transferable online knowledge approach involves using pre-trained models on a large dataset to improve the accuracy and speed of training the CNN model for stress and emotion classification. By transferring the knowledge gained from pre-trained models, the CNN model can learn to identify patterns and features in data sets that would have otherwise required a significantly more considerable amount of training data. The novelty and contribution of this research lie in applying CNNs in psychology for stress and emotion classification and using transferable online knowledge to optimise the CNN model. Using CNNs in psychology can improve mental health diagnosis and treatment accuracy and speed. By accurately classifying stress and emotions, healthcare professionals can identify individuals at risk of developing mental health issues and provide timely interventions.

II. Literature Review

Research work [11] presented "SensorN" and a scalable and low-power integrated Deep Convolutional Neural Network (DCNN). Its purpose was to classify multimodal time-series signals produced by different sensors with different sampling rates initially converted into images. However, in contradiction to this research work, accuracy and performance were major barriers to classifying stress and sentiment. DCNN then used them to adopt their standard features from the images and perform the classification. Using CNN and generative adversarial networks, the paper [12] attempted to accurately train the system to predict when a person experience negative sentiments. The authors also addressed the problems of identifying stress and negative sentiments through deep learning and proposed strategies that can be applied to reduce them. Compared with the proposed method, the computation effort of the techniques we mentioned above was high, which led to the need to evaluate more EEG features.

Considering the sentiment factor, the paper [13] presents a system for extensive data application in sentiment-aware healthcare to improve healthcare services. Societies need to develop guidelines to help provide health care and well-being to older adults by addressing aspects that promote age-friendliness in intelligent urban environments, primarily in the post-Covid era.

Confronting these challenges, [14] proposed a framework to predict the outbreak's progress. Continuing their previous research work, the authors focused more on dealing with the pandemic that had established the new reality in our modern society. In particular, by modifying the Levenberg – Marquardt algorithm, they achieved a better approach to data analysis and therefore offered a timely information solution to predict the increase in cases [15]. Based on the same orientation, the researchers who conducted research [16] attempted to classify the characteristics of COVID-19 and estimate its quantity indoors. By optimising data mining and machine learning, they successfully attempted to detect the virus indoors effectively.

Notwithstanding, the required time to acquire the data from the sensors and interpret them to procure the knowledge leads to time complexities. Thus, the efficiency of the procedures we quote above is limited. Throughout the proposed research contribution in the field we aim to deal with these issues by handling the abrupt drift during the online transfer assignment process.

The CNNs are categorised into two types convolutional layers and subsampling layers. We increase the characteristics depending on the depth of the layers in the convolutional networks. Improving the performance of DNNs occurs through six modifications, increasing the depth or width, modifying the convolution operation [17], the pooling operation [18], or activation function [19], and reducing the number of parameters [20]. By studying changes in blood glucose levels through CNN, paper [21], researchers concluded that they could lead to stimulation of the autonomic nervous system, which would cause a change in the results of the Electrocardiogram. Article [22] introduced an innovative framework based on a productive, competitive network PulseG to create realistic photoplethysmography biometric signals, an unobtrusive technique to monitor cardiac signals from facial videos through attenuation colouring signals [23]. The above-mentioned non-intrusive techniques evaluate vital information generated regarding emotions, while the classification accuracy is less than the proposed one.

Research paper [24] which processed in-depth data to classify emotions, applied the classification process LIBLINEAR in conjunction with the Bert model [25], resulting in higher accuracy for the multi-classification problem. Developing the same CNN model, his research [26] studied the classification of cognitive stress by creating five different sessions. During some of them, they included playing relaxing music. While the rest, of the participants evaluated mathematical expressions. This model, however, had high demands on computing power and storage. Two years later, the researchers gave the golden ratio and the authors of the research work [27], reducing the computational memory requirements and the learning time but without lowering the model's efficiency in terms of accuracy. It was achieved by using the channel-based pruning of a structured model. Current stress and sentiment classification techniques start once the knowledge from the set of stress to the collection of sentiments has been transferred. However, this causes delays in the interpretation, and the information is not acquired in real-time. Hence, through this article, we aim to overcome this barrier and obtain performance in a real-time environment.

Consumer electronics is a rapidly growing field that focuses on the design, development, and manufacturing of electronic devices, software, and services that are intended for use by consumers [28]. It mainly focuses on the engineering and research aspects of the theory, design, construction, manufacture, or end use of mass-market electronics [29]. Through this research, we aim to develop a CNN model for consumer electronics that can accurately classify human stress and emotion by optimising the model through transferable online knowledge [30].

The main goal is to create user-friendly, intuitive, and reliable products while being affordable and accessible to the public [31]. Therefore, it requires a deep understanding of consumers' needs, preferences, and the latest technological developments and innovations [32]. The engineering and research aspects of consumer electronics involve developing new technologies, optimising existing technologies, and integrating these technologies into products and services that meet the needs of consumers [33].

Overall, the field of consumer electronics is a dynamic and exciting area of research and development focused on creating innovative and user-friendly products and services that enhance consumers' lives [34]. The engineering and research aspects of consumer electronics are crucial to the success of this field, and they play a vital role in driving innovation and advancing technological progress [35].

III. PROPOSED MODEL

This section elaborates on the homogenous and heterogeneous scenarios concerning the stress sentiments classification. The process workflow consists of three steps: classification, attribute extraction, and physical sampling signals. The implementation of the classification leads to three pillars stress, indifference, and merriment. TABLE I lists the mathematical notations to enhance the readability.

The architecture of the CNN model used for stress and sentiment classification are illustrated in Fig. 1. As can be seen, the CNN methodology is used for mental visualisation, while taking the input and differentiating the output value.

It is the normalised version of a multi-layer perceptron which is one layer of the cell connected to the next layer. A CNN consists of combined rank input and output layer and multiple hidden layers. The hidden layers of the CNN contain a series of convolutional layers convoluted with multiplication or alternative actual number.

This article has been accepted for publication in IEEE Transactions on Consumer Electronics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCE.2023.3319111

TABLE I. GLOSSARY OF NOTATIONS

Notation	Description
s	Stress
e	Sentiment
$\mathcal{B}_s$	Stress attributes domain
$\{x_i^{n_s} \times y_i^{n_s}\}$ $\{x_i^{n_e} \times y_i^{n_e}\}$	Attribute domains of the i^{th} stress and sentiment domain, respectively
$\varkappa_1$	Convolution function of a kernel filter
g	Activation function
f	Ensemble activation function
$w_{i,j}$	Weighting factors
sgn	Signum function
$\sigma(x)$	Standardisation factor
$*$	Operation of convolution
ℓ	Loss factor
$\varrho_1,\ \varrho_2,\ T$	Parameters
D_i	Window size parameter

The convolutional layer within the neural network has convolutional kernels described by dimension and height. Also, the variety of input channels, output channels and the depth of the convolutional filter are sufficient for the number of channels of the input feature map.

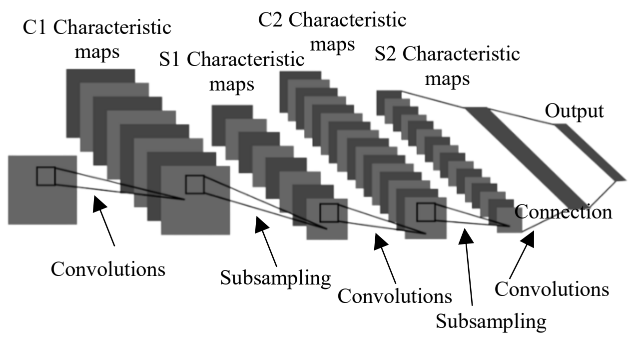

Fig. 1. Proposed CNN architecture

We assume that the stress attributes domain $\mathcal{B}_s$ and sentiment attributes domain are formed by ordered pairs (1) and (2), respectively. Where $x_i^{n_s} \times y_i^{n_s}$ and $x_i^{n_e} \times y_i^{n_e}$ are the attribute domains of the i^{th} stress and sentiment domain respectively. Let also, $x_i^{n_s} = \mathbb{R}^m$, $x_i^{n_e} = \mathbb{R}^m$ and $y_i^{n_s} = \{-1,1\}$, $y_i^{n_e} = \{-1,1\}$, where $n_s \gg n_e$.

$$\mathcal{B}_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s} \tag{1}$$

$$\mathcal{B}_e = \{(x_i^e, y_i^e)\}_{i=1}^{n_e} \tag{2}$$

We denote the convolution function of a kernel filter by $\varkappa_1: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. The model OTLCNNO is based on (3) and (4) for stress and sentiment, respectively.

$$f(x^s) = \sum_{i=1}^{n_s} y_i^s\, \varkappa_1(x_i^s, x) \tag{3}$$

$$f(x^e) = \sum_{i=1}^{n_e} y_i^e\, \varkappa_1(x_i^e, x) \tag{4}$$

A prerequisite for enabling the OTL is to procure the knowledge of the activation functions $g \in \mathcal{R}_{\varkappa_1}$ on a sentiment domain of cases (5) in the attribute domain $x_i^{n_e} \times y_i^{n_e}$.

$$\left\{\left(x_{i,j}^e, y_{i,j}^e\right) \middle|\ j = 1, \dots, J\right\}_{i=1}^{n_e} \tag{5}$$

Throughout the implementation of the procedure, we assume that the learner assigns a case $x_{i,j}^e$ at the j^{th} iteration. We use the online learning process to optimise the activation function so that the sorted class label $g_j\left(x_{i,j}^e\right)$ associate with the original label $y_{i,j}^e$. Foremost, we need to explore the solution of effectively transferring knowledge from stress to the domain of sentiments to increase online learning effectiveness.

A. Homogeneous Domain Adaptation

Assume that the domains of stress and sentiment are equal $x_i^{n_e} = x_i^{n_s}$ and $y_i^{n_e} = y_i^{n_s}$. The problem, known as concept shift, complicates the learning of a model because the variable that denotes the sentiment varies according to the time. Therefore, transferring knowledge from the domain of stress to sentiments becomes more difficult.

We target the solution by a fundamental method called ensemble learning. This method is a meta-approach to Machine Learning that seeks to optimise prediction and mitigation by combining multiple modelling [36]. Initially, a new activation function g driven from the sentiment domain towards an ensemble activation function f. We must also find a way to effectively combine the two activation functions to solve the concept shift problem.

The fundamental notion for the solution is based on the combination of the activation functions $f(x)$ and $g_j(x)$ at j^{th} iteration of the online learning process. Therefore, we take into consideration the weighting factors $w_{1,j}$ and $w_{2,j}$ where the signum function (6) classifies the case $x_{i,j}^e$ at the j^{th} iteration according to their sign.

$$\hat{y}_{i,j}^e = sgn\left[w_{1,j} \prod_{i=1}^{n_e}\left(f\left(x_{i,j}^e\right)\right) + w_{2,j} \prod_{i=1}^{n_e}\left(g_j\left(x_{i,j}^e\right)\right) - \frac{1}{2}\right] \tag{6}$$

Correspondingly, at j^{th} iteration, the ensemble function (7) classifies the case $x_{i,j}^s$.

$$\hat{y}_{i,j}^s = sgn\left[w_{1,j} \prod_{i=1}^{n_s}\left(f\left(x_{i,j}^s\right)\right) + w_{2,j} \prod_{i=1}^{n_s}\left(g_j\left(x_{i,j}^s\right)\right) - \frac{1}{2}\right] \tag{7}$$

Where $\sigma(x)$ indicates the standardisation factor determined by (8).

$$\sigma(x) = max\left(0, min\left(1, \frac{x+1}{2}\right)\right) \tag{8}$$

Initially, we substitute the weighting factors by the value of a half $w_{1,1} = w_{2,1} = 0.5$. We process these two weighting factors by the dynamic fine-tuning that trails the commonly used method in CNN, the backpropagation method. Also, with the development of the Passive-Aggressive Classifier (PAC), which is a classification algorithm that uses online learning in machine learning, we effectively transfer, for the next iteration of the online transfer learning process, along with updating the function $g_{j+1}(x)$. The formulation of the weighting factors is presented in iterative formulas (9) and (10).

$$w_{1,j+1} = \frac{w_{1,j} * \lambda_j(f)}{w_{1,j} * \lambda_j(f) + w_{2,j} * \lambda_j(g_j)} \tag{9}$$

$$w_{2,j+1} = \frac{w_{2,j} * \lambda_j(g_j)}{w_{1,j} * \lambda_j(f) + w_{2,j} * \lambda_j(g_j)} \tag{10}$$

Where ℓ is the loss factor and $\lambda_j(h^e)$ and $\lambda_j(h^s), \forall h \in \mathcal{R}_{\varkappa_1}$ are defined in (11) and (12), respectively. From the operation of convolution denoted by $*$ a new function emerges. The obtained function is determined by integrating the product of the two functions after one is inverted and shifted [37].

<u>*Algorithm. 1*</u>

Input: OTLCNNO classifier $f(x^s), f(x^e)$, initial exchange T and weights $w_{1,1} = w_{2,1} = 0.5$

Set $g_1 = 0$;
$\boldsymbol{for}(j = 1, \dots, J)$
  Let $x_{i,j}^e \in \mathcal{X}_i^{n_e}$ and $x_{i,j}^s \in \mathcal{X}_i^{n_s}$;
  Classify $\hat{y}_{i,j}^e$ and $\hat{y}_{i,j}^s$ by (6) and (7);
  Assign label: $y_{i,j}^e \in \{-1,1\}$ and $y_{i,j}^s \in \{-1,1\}$;
  Calculate $w_{1,j+1}$ and $w_{2,j+1}$ by (9) and (10);
  Calculate the loss:
  $1 - \ell_j^e = y_{i,j}^e g_j(x_{i,j}^e)$ and $1 - \ell_j^s = y_{i,j}^s g_j(x_{i,j}^s)$;
  $\boldsymbol{if}(\ell_j^e > 0)$
    $\varepsilon_j = \min\left\{L, \frac{\ell_j^e}{\varkappa_2 \|x_{i,j}^e\|^2}\right\}$;
    $g_{j+1} = g_j + \varepsilon_j y_{i,j}^e \varkappa_2(x_{i,j}^e)$;
  $\boldsymbol{end\ if}$
  $\boldsymbol{if}(\ell_j^s > 0)$
    $\varepsilon_j = \min\left\{T, \frac{\ell_j^s}{\varkappa_2 \|x_{i,j}^s\|^2}\right\}$;
    $g_{j+1} = g_j + \varepsilon_j y_{i,j}^s \varkappa_2(x_{i,j}^s)$;
  $\boldsymbol{end\ if}$
$\boldsymbol{end\ for}$

$$-\frac{1}{\eta\ell}\ln[\lambda_j(h^e)] = \left(\prod_{i=1}^{n_e}\left(h(x_{i,j}^e)\right), \prod_{i=1}^{n_e}\left(y_{i,j}^e\right)\right) \tag{11}$$

$$-\frac{1}{\eta\ell}\ln[\lambda_j(h^s)] = \left(\prod_{i=1}^{n_s}\left(h(x_{i,j}^s)\right), \prod_{i=1}^{n_s}\left(y_{i,j}^s\right)\right) \tag{12}$$

B. Heterogeneous Domain Adaptation

The problem of adaptation to a heterogeneous field through the online transfer learning process due to differences in stress and sentiments must be solved. Therefore, we approach the solution using a sentiment domain subset. However, due to the difference between the two attribute domains, the algorithmic process cannot be implemented unswervingly, and thus we use a Multiview methodology.

We take into consideration the initial m dimensions of $x_i^{n_e}$ and we denote the old attribute domain $x_i^{n_s}$. According to the Multiview methodology, every data case $x_{i,j}^e$ is distributed into two cases $x_{i,j}^{e(1)} \in x_i^{n_s}$ and $x_{i,j}^{e(2)} \in \mathcal{X}_i^{n_e}/\mathcal{X}_i^{n_s}$ and we determine the convolutional function of the kernel filter as $\varkappa_2: \mathbb{R}^{n-m} \times \mathbb{R}^{n-m} \to \mathbb{R}$ for the next case. Heterogeneous OTL aims to implement a co-regulation of online learning for the two classifiers $g_j^{(1)}$ and $g_j^{(2)}$ Simultaneously from both perspectives, classify unknown data on the field of sentiments and stress respectively from (14) and (14).

$$\hat{y}_{i,j}^e = sgn\left(\frac{g_j^{(1)}(x_{i,j}^{e(1)}) + g_j^{(2)}(x_{i,j}^{e(2)})}{2}\right) \tag{13}$$

$$\hat{y}_{i,j}^s = sgn\left(\frac{g_j^{(1)}(x_{i,j}^{s(1)}) + g_j^{(2)}(x_{i,j}^{s(2)})}{2}\right) \tag{14}$$

The first step of the algorithmic technique for the OTLCNNO classifier will be the substitution of $g_j^{(1)} = f$ and $g_j^{(2)} = 0$. By using the co-regulation (15) for new insertions of data, we update the functions $g_{j+1}^{(1)}$ and $g_{j+1}^{(2)}$ for $g^{(1)} \in \mathcal{R}_{\varkappa_1}, g^{(2)} \in \mathcal{R}_{\varkappa_2}$.

$$\left(g_{j+1}^{(1)}, g_{j+1}^{(2)}\right) = ArgMin\frac{1}{2}\Big(\varrho_1\left\|g^{(1)} - g_j^{(1)}\right\|_{\mathcal{R}_{\varkappa_1}}^2 + \varrho_2\left\|g^{(2)} - g_j^{(2)}\right\|_{\mathcal{R}_{\varkappa_2}}^2\Big) + T\ell_j \tag{15}$$

Where, $\varrho_1 > 0$, $\varrho_2 > 0$ and $T > 0$ are parameters. Also, the computed loss factor ℓ_j is determined in (16) and (17) as well as the $ArgMin$ is used to determine minimum values from a set of attributes considering some specific constraints. The mathematical notations of (16) and (17) explained in TABLE I.

$$\ell_j^e = 1 - y_{i,j}^e/2\left(g_j^{(1)}(x_{i,j}^{e(1)}) + g_j^{(2)}(x_{i,j}^{e(2)})\right) \tag{16}$$

$$\ell_j^s = 1 - y_{i,j}^s/2\left(g_j^{(1)}(x_{i,j}^{s(1)}) + g_j^{(2)}(x_{i,j}^{s(2)})\right) \tag{17}$$

The insertion of new order pairs $(x_{i,j}^e, y_{i,j}^e)$ and $(x_{i,j}^s, y_{i,j}^s)$ will be classified by using this algorithmic technique that constructs the update of the ensemble classifier. Also, the use of this approach forces two-sided classifiers without contradiction from the previous classifiers $(g_j^{(1)}, g_j^{(2)})$ using the initial two adjustment terms.

The activation function update will be implemented as soon as the learner's loss is calculated based on some criteria related to the updated data. The learning process has the ultimate goal of reducing data loss. However, the OTL process will not be adequately implemented due to the drift concept scenario since the distribution will be frequently modified during each period. Nevertheless, effectively managing the concept drift with frequently metadata distributions requires a combination of the proposed adaptive algorithms, monitoring strategies, and thoughtful design considerations.

Algorithm. 2

Input: OTLCNNO classifier $f(x^s), f(x^e), \varrho_1, \varrho_2$ and T
Set $g_j^{(1)} = f$ and $g_j^{(2)} = 0$;
for $(j = 1, \ldots, J)$
Let $x_{i,j}^e \in x_i^{n_e}$ and $x_{i,j}^s \in x_i^{n_s}$;
Classify $\hat{y}_{i,j}^e$ and $\hat{y}_{i,j}^s$ by (13) and (14);
Assign label: $y_{i,j}^e \in \{-1,1\}$ and $y_{i,j}^s \in \{-1,1\}$;
Calculate the loss ℓ_j^e and ℓ_j^s using (16) and (17);
if $(\ell_j^e > 0)$
$\varepsilon_j = \min\left\{T, \frac{4\varrho_1\varrho_2\ell_j^e}{\varkappa_{1,j}\varrho_2+\varkappa_{2,j}\varrho_1}\right\}$;
$g_{j+1}^{(1)} = g_j^{(1)} + \frac{\varepsilon_j y_{i,j}^e \varkappa_1}{2\varrho_1}\left(x_{i,j}^{e(1)}\right)$;
$g_{j+1}^{(2)} = g_j^{(2)} + \frac{\varepsilon_j y_{i,j}^e \varkappa_2}{2\varrho_2}\left(x_{i,j}^{e(2)}\right)$;
end if
if $(\ell_j^s > 0)$
$\varepsilon_j = \min\left\{T, \frac{4\varrho_1\varrho_2\ell_j^s}{\varkappa_{1,j}\varrho_2+\varkappa_{2,j}\varrho_1}\right\}$;
$g_{j+1}^{(1)} = g_j^{(1)} + \frac{\varepsilon_j y_{i,j}^s \varkappa_1}{2\varrho_1}\left(x_{i,j}^{s(1)}\right)$;
$g_{j+1}^{(2)} = g_j^{(2)} + \frac{\varepsilon_j y_{i,j}^s \varkappa_2}{2\varrho_2}\left(x_{i,j}^{s(2)}\right)$;
end if
end for

C. Machine Learning Concept Drift

The conjecture that a procedure of binary classification in a concept drift case where we can access the learner over time with a set of data. Hence, at time j, the cases of the algorithm $x_j = \{x_{i,j}^e, x_{i,j}^s\} \in \mathbb{R}^m$ and will classify and assign the label as $\hat{y}_j = \{\hat{y}_{i,j}^e, \hat{y}_{i,j}^s\} = sgn(g_j x_j) \in \{-1,1\}$ where g_j denotes the activation function.

The original label $\hat{y}_j$ will be exposed after the implementation of the classification. Thus, the learner will lose (18), which is the discrepancy between his classification and the original label.

$$\ell_j = \{\ell_j^e, \ell_j^s\} = \ell\left((x_j, y_j), g_j\right) \tag{18}$$

Therefore, we will divide the basic plan into multiple repetitions during online learning to solve the problem. The assimilated knowledge in each rediscovery is transferred from the OTLCNNO by initialising it as a zero vector. Hence, the novel classifier is deployed; alternatively, the OTLCNNO is used. The formulation of the concept-drifting OTL is applied through the window size parameter denoted as D_i and which is the quantity of cases established in the i^{th} iteration. Also, the functions of the classifiers remain unchanged in j^{th} the iteration, using the function (21), the class label of x_j is classified.

$$w_{1,j+1} = \frac{w_{1,j} * \lambda_j(f_j)}{w_{1,j} * \lambda_j(f_j) + w_{2,j} * \lambda_j(g_j)} \tag{19}$$

$$w_{2,j+1} = \frac{w_{2,j} * \lambda_j(g_j)}{w_{1,j} * \lambda_j(f_j) + w_{2,j} * \lambda_j(g_j)} \tag{20}$$

$$mod(j, D_i) \neq 0$$

Algorithm. 3

Set $f_1 = 0, g_1 = 0, w_{1,1} = 0$ and $w_{2,1} = 1$ and $i = 1$
for $(j = 1, \ldots, J)$
Let $x_j \in \mathcal{X}$
Classify $\hat{y}_j$ using (21);
Assign label: $y_j \in \{-1,1\}$;
Calculate the loss $\ell_j = \max\{0, 1 - y_j g_j x_j\}$;
if $(\ell_j > 0)$
$\varepsilon_j = \min\left\{T, \frac{\ell_j}{\varkappa_2\|x_j\|^2}\right\}$;
$g_{j+1} = g_j + \varepsilon_j y_j x_j$;
end if
$f_{j+1} = f_j$;
$w_{1,j+1} = \frac{w_{1,j}*\lambda_j(f_j)}{w_{1,j}*\lambda_j(f_j)+w_{2,j}*\lambda_j(g_j)}$
$w_{2,j+1} = 1 - w_{1,j+1}$;
if $(mod(j, D_i) = 0)$
$f_{j+1} = \begin{cases} f_{j+1}, & if\ w_{1,j+1} \geq w_{2,j+1} \\ g_{j+1}, & Alternatively \end{cases}$
$g_{j+1} = 0$ and $w_{1,j+1} = w_{2,j+1} = 0.5$ and $i = i + 1$;
end if
end for

$$\hat{y}_j = sgn\frac{1}{2}\left[2w_{1,j}\prod_{j=1}^{J}\left(f_j(x_j)\right) + 2w_{2,j}\prod_{j=1}^{J}\left(g_j(x_j)\right) - 1\right] \tag{21}$$

However, the main challenge remains effective weight management. The OTLCNNO classifier is zero in the original entry, so the activation function is weighted with zero, while the new method is weighted with one. The fundamental weighting equations (20) and (21) are used for the dynamic fine-tuning process for the remaining iteration.

Algorithm. 4

Input: Small window size D and exchange T
Set $k_{t,1} = 0, N_t = 0$ where $t = 1,2$ and $D_1 = D, i = 1$;
for $(j = 1, \ldots, J)$
Let $x_j \in \mathcal{X}$
Classify $\hat{y}_{t,j} = sgn(k_{t,j}, x_j)$;
Assign label: $y_j \in \{-1,1\}$;
Calculate $N_t = N_t + \mathbb{I}_{(\hat{y}_{t,j} \neq y_j)}$;
Calculate the loss $\ell_{t,j} = \max\{0, 1 - y_j k_{i,j} x_j\}$;
$k_{t,j+1} = k_{t,j} + \varepsilon_{t,j} y_j x_j$ where $\varepsilon_{t,j} = \min\left\{T, \frac{\ell_{t,j}}{\varkappa_2\|x_j\|^2}\right\}$;
if $(mod(j, D) = 0)$
if $(N_1 > N_2)$
$i = i + 1, D_i = D$;
else
$D_i = D_i + D$;
end if
$k_{1,j+1} = k_{2,j+1}, k_{2,j+1} = 0, N_t = 0, t = 1,2$;
end if
end for

On the other hand, D_i data at different iterations depending on the fundamental frequency, intensity, and signal changes that can be input and parsed simultaneously affect performance. To overcome this issue, we could replace all the parameters D_i with the constant D which is necessary to be consistent with the concept drift iteration.

However, it is impossible to apply the above method because, we cannot be precise in determining the exact parameter D for the substitution to achieve the desired result. Secondly, we must remember that the shift of concepts will not be fixed so the input will not be repaired, leading to erroneous results. Therefore, as an alternative, we propose Algorithm. 4, which implements the automatic selection of different D_i parameters similarly to an online window tuning algorithm.

IV. Experimental Results

In this section, we present the outcomes of our experimental investigations, detailing the statistical findings. To rigorously assess the significance of the proposed model, we conducted a comprehensive comparison between OTLCNNO and state-of-the-art algorithmic technics in the field.

The SWELL-KW and the WESAD dataset were utilised for the validation. Upon preprocessing, the TL technique and fine-tuning on pre-trained CNNs deployed to adopt samples from the datasets.

The WESAD open-source dataset comprised various physical biometric signals, i.e. ECG, blood volume pulse, electromyogram, electrodermal activity, temperature and heart rate composed with a sampling of triaxial acceleration signals at 32Hz. The data are essential for the interpretation of stress and sentiment-domain attributes.

While the SWELL-KW dataset includes results from 25 individuals regarding their working conditions with multiple stressors, it contains body sensors' facial expression data, body postures, heart rate, and skin conductance. Metadata and extracted features are also included, except for the raw data.

To handle the great quantity of diverse data, we select a sample for processing. More precisely, from approximately 250-500 thousand raw data from each of the 17 subjects participating in generating data for the WESAD dataset, we select a sample of 1230. While from the data in SWELL-KW obtained from 25 participants, 1018 values for each category were selected.

The comparison of the performance has been implemented between the proposed OTLCNNO and the following state-of-the-art classifiers after utilising TL:

Library for Large Linear (LIBLINEAR): This method is widely used for training linear classifiers, particularly in machine learning and data mining. It is designed to be highly scalable and efficiently handle large datasets. It uses a linear SVM formulation that efficiently addresses large-scale data with a relatively small memory footprint. Also, it is a highly efficient algorithm that can train linear classifiers quickly and accurately. Advanced optimisation techniques achieve this speed, such as stochastic gradient descent and coordinate descent [9].

Deep Belief Network (DBN): This model, combined with a transfer learning (TL) strategy (DBNTL), is a multi-layer neural network that can learn and extract hierarchical features from input data. Thus, it is highly effective in learning complex patterns and structures in data. Furthermore, when combined with a TL strategy, the model can leverage the knowledge learned from pre-trained models to improve learning efficiency and accuracy [10].

LinearSVC (Linear Support Vector Classification): LinearSVC is a classifier from the scikit-learn library [39], that provides an interface like other scikit-learn classifiers and is widely used for linear classification tasks like sentiment analysis [40].

MLP (Multi-Layer Perceptron): This supervised learning algorithm learns a function used for binary classification. Also, it is often utilised as a building block for more complex neural network architectures [41], [42].

SGDClassifier (Stochastic Gradient Descent Classifier): This algorithm is based on stochastic gradient descent optimisation. It is efficient and suitable for large-scale learning tasks [43], [44].

Ridge Classifier: This linear classifier uses ridge regression for binary classification tasks, and it is useful when dealing with multicollinearity in the feature set [45], [46].

Recursive Auto-Encoders (RAE): RAE is a neural network model used in natural language processing (NLP) to learn hierarchical representations of sentences or other sequential data. RAEs are particularly effective for capturing the syntactic structure of sentences and have applications in tasks such as parsing, sentiment analysis, and machine translation [47].

Left-Right Bidirectional Long Short-Term Memory (LR-Bi-LSTM): LR-Bi-LSTM model is a type of neural network architecture commonly used in natural language processing (NLP) tasks, such as text classification, sentiment analysis, and named entity recognition. It combines the strengths of bidirectional Long Short-Term Memory (Bi-LSTM) networks with left-to-right and right-to-left processing [48].

TABLE II includes the confusion matrix generated by the statistical computing program R-4.1.2. From these results, the metrics for the statistical performance comparison were derived and the corresponding Confidence Intervals (CIs).

The underpinning of the statistical analysis rests upon the iterative execution of experiments employing the SWELL-KW and the WESAD datasets. This orchestrated approach culminates in the derivation of outcomes and the subsequent computation of CIs across various and distinct temporal time points, achieved through repeated iterations. Leveraging techniques adapted from the binomial distribution, we calculate the respective metrics along with their Margin of Error (MoE), indicating the CI for each metric. We chose a confidence level of 95% to align with standard practice and provide intervals encompassing a reasonable spectrum of potential outcomes. As delineated in TABLE II, the MoE exhibits values below 0.05, denoting a confidence level of 95%. This outcome substantiates the precision and heightened confidence of all methodologies under consideration. Nonetheless, it is noteworthy to underscore that the most diminutive MoE, signifying the most constricted CI, is evident in the proposed OTLCNNO method. This observation accentuates the method's exceptional consistency and steadfastness.

TABLE II. COMPARISON OF STATISTICAL PERFORMANCE METRICS

Method	Confusion Matrix		WESAD CI Precision	WESAD CI Recall	WESAD CI F_1-Score	WESAD CI Accuracy	Confusion Matrix		SWELL-KW CI Precision	SWELL-KW CI Recall	SWELL-KW CI F_1-Score	SWELL-KW CI Accuracy
OTLCNNO	386	21	0.948± 0.012	0.937± 0.014	0.943± 0.013	0.962± 0.011	289	7	0.976± 0.010	0.900± 0.018	0.937± 0.015	0.962± 0.012
	26	797					32	690				
LibLinear	361	30	0.923± 0.015	0.926± 0.015	0.924± 0.015	0.952± 0.012	256	38	0.870± 0.021	0.850± 0.022	0.861± 0.021	0.918± 0.017
	29	810					45	679				
DBNTL	379	25	0.938± 0.013	0.894± 0.017	0.915± 0.016	0.943± 0.013	247	145	0.630± 0.030	0.876± 0.020	0.733± 0.027	0.823± 0.023
	45	781					35	591				
LinearSVC	375	27	0.933± 0.014	0.906± 0.016	0.919± 0.015	0.946± 0.013	291	29	0.909± 0.018	0.882± 0.020	0.895± 0.019	0.933± 0.015
	39	789					39	659				
MLP	381	31	0.925± 0.015	0.929± 0.014	0.927± 0.015	0.951± 0.012	258	49	0.840± 0.022	0.863± 0.021	0.851± 0.022	0.912± 0.017
	29	789					41	670				
SGD	383	43	0.899± 0.017	0.943± 0.013	0.920± 0.015	0.946± 0.013	289	48	0.858± 0.021	0.873± 0.020	0.865± 0.021	0.912± 0.017
	23	781					42	639				
Ridge	335	43	0.886± 0.018	0.758± 0.024	0.817± 0.021	0.878± 0.018	293	30	0.907± 0.018	0.862± 0.021	0.884± 0.020	0.924± 0.016
	107	745					47	648				
RAE	342	39	0.898± 0.017	0.781± 0.023	0.835± 0.021	0.890± 0.017	283	25	0.919± 0.017	0.898± 0.019	0.909± 0.018	0.944± 0.014
	96	753					32	678				
LR-Bi-LSTM	362	37	0.907± 0.016	0.923± 0.015	0.915± 0.016	0.946± 0.013	269	58	0.823± 0.023	0.879± 0.020	0.850± 0.022	0.907± 0.018
	30	801					37	654				

When applying any classification procedure, as in our case, the accuracy for a class is the number of data correctly selected as positive and the total number of positive data, whether true or not; therefore, if we denote the classified stress and sentiment cases as TP and FP, the precision formula will be the (22).

$$Precision = \frac{TP}{TP + FP} \tag{22}$$

The confusion matrix in TABLE III represents amounts from predicted and actual values. True Negative (TN) indicates the number of negative examples classified accurately. Similarly, True Positive (TP) presents the number of positive models categorised accurately. False Positive (FP) values are the number of actual negative examples classified as positive, and respectively, False Negative (FN) shows a value which is the number of true positive models classified as negative.

TABLE III. CONFUSION MATRIX

		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

According to the sentimental orientation, the actual positive case depicts the number of precisely classified stress sentiment cases, and the correspondingly false positive is the number of imprecisely classified stress sentiment cases. Therefore, rendering to TABLE II, the proposed model is more precise than the other models.

With an analogous performance metric called recall, we aim to evaluate the sensitivity of our model by the other two. Considering the correct classified stress sentiment cases TP and the FN, we compute each method's recall using (23).

$$Recall = \frac{TP}{TP + FN} \tag{23}$$

TP depicts the relative classified cases of stress sentiments recovered successfully, and FN the number of irrelevant classified cases retrieved unsuccessfully. According to TABLE II, the proposed model is more sensitive than the similar latest models from the research field.

Using the values of precision and recall from the previous paragraphs will be beneficial for statistical analysis reasons to include F_1-score, which measures a test's accuracy and is given by (24). As shown in TABLE II, the F_1-score representing the harmonic mean of the precision and recall is higher for the proposed model, indicating better precision and recall than the other two.

$$F_1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{24}$$

If the model is systematically erroneous, then by incrementing the quantity of data we have taken as the sample, we will increase precision but not improve the increment of accuracy. Therefore, we evaluate its accuracy as well. Accuracy given by (25) is the percentage of appropriate classification of stress sentiment cases over the total quantity of tests performed.

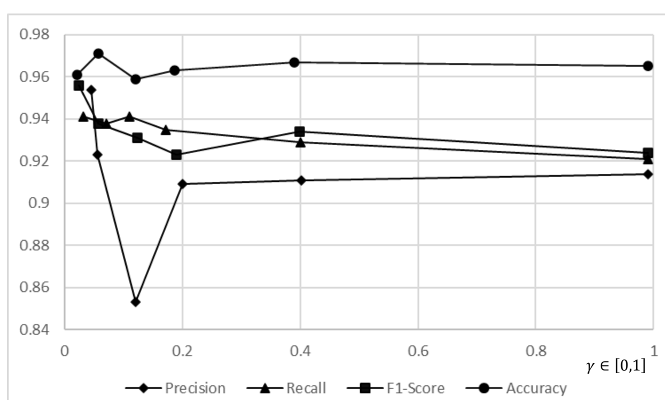

Fig. 2. Performance of OTLCNNO concerning the $\gamma \in [0,1]$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (25)$$

Fig. 2 presents precision, recall, accuracy and F_1-score in respect to $\gamma \in [0,1]$, the correlation between ordinal variables. The proposed method results from the highest accuracy when $\gamma \rightarrow 0$. As γ increases, the decision limit tents towards the majority class. Hence, recall is increasing, and precision is decreasing. Besides, $\gamma = 0.1$ generates the lowest precision (0.8502). Overall, based on the graphs, it is also derived that the method performs best when $\gamma = 0.085$.

V. Conclusion, Discussion & Future Orientation

Our contribution through this research work, captured with the help of this article, lies in the mathematical processing of the CNN model with OTL for further improvement. The evaluation results using various statistical performance metrics show that the proposed method's effectiveness outperforms state-of-the-art classification techniques in the field. In addition, TABLE II yields insightful findings regarding the CIs across compared methodologies. A salient point of distinction emerges with the proposed OTLCNNO method, where the lowest MoE value manifests. This characteristic delineates OTLCNNO's propensity for generating the narrowest CIs among the examined methodologies. The coherence in results across multiple instances further highlights the method's intrinsic consistency, fortifying its potential as a reliable approach in sentiment classification techniques. Our study not only elucidates the precision of estimations but also accentuates the methodological steadiness of the OTLCNNO technique.

These findings collectively contribute to a deeper understanding of sentiment classification techniques and provide valuable insights for future research and application. The ultimate goal was to process sentiments and stress as accurately as possible by processing biological evidence. Also, the development process used prior knowledge gained from training data on stress and sentiments. It also elaborates on the scenarios of online learning transfer in homogeneous and heterogeneous areas of common and diverse characteristics. The solution derives from OTL with concept drift challenges when performing online CNN classification via streaming data in the cases mentioned earlier.

Advancing state-of-the-art technological improvements in healthcare strategic planning are shaping the future of smart and healthy quality of life. However, as we are on the cusp of an evolution in efficient and secure intelligent healthcare systems, the disciplines of Psychology and Neuroscience remain an impediment to amelioration. Therefore, researchers investigate mental health identification by analysing and interpreting acquired data from devices and sensors. Since it is an enduring challenge to the healthcare community, through convolutional neural networks, we seek to leverage physiological data, optimising classification performance by targeting a domain of datasets received in an online manner called online transfer learning. More precisely, we aim to contribute by implementing deep learning technology in the medical area, as it would be vital to mitigate the burden on mental health.

Acknowledgement

This research work was funded by the Smart and Healthy Ageing through People Engaging in supporting Systems SHAPES project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857159. The work of the Polish researchers has been supported by the Eureka CELTIC-NEXT project "Intelligent Management Of Next Generation Mobile Networks And Services" (IMMINENCE, C2020/2-2), which has been funded in Poland by Narodowe Centrum Badań i Rozwoju (National Centre for Research and Development).

References

[1] B. S.Oken, I. Chamine and W. Wakeland, "A systems approach to stress, stressors and resilience in humans," *Behavioural Brain Research,* vol. 282, pp. 144-154, 1 Apr. 2015.

[2] A. P. Allen, P. J. Kennedy, J. F. Cryan, T. G. Dinan and G. Clarke, "Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test," *Neuroscience & Biobehavioral Reviews,* vol. 38, pp. 94-124, Jan. 2014.

[3] E. T. Attar, V. Balasubramanian, E. Subasi and M. Kaya, "Stress Analysis Based on Simultaneous Heart Rate Variability and EEG Monitoring," *IEEE Journal of Translational Engineering in Health and Medicine,* vol. 9, pp. 1-7, 23 Aug. 2021.

[4] B. He, L. Yang, C. Wilke and H. Yuan, "Electrophysiological Imaging of Brain Activity and Connectivity—Challenges and Opportunities," *IEEE Transactions on Biomedical Engineering,* vol. 58, no. 7, pp. 1918-1931, Jul. 2011.

[5] A. Lalos, A. Antonopoulos, E. Kartsakli, M. D. Renzo, S. Tennina, L. Alonso and C. Verikoukis, "RLNC-Aided Cooperative Compressed Sensing for Energy Efficient Vital Signal Telemonitoring," *IEEE Transactions on Wireless Communications,* vol. 14, no. 7, pp. 3685 - 3699, 09 Mar. 2015.

[6] L. Crameri, I. Hettiarachchi and S. Hanoun, "Feasibility Study of Skin Conductance Response for Quantifying Individual Dynamic Resilience," in *2020 IEEE International Conference on Systems, Man, and Cybernetics*, 2020.

[7] K. Masood and M. A. Alghamdi, "Modeling Mental Stress Using a Deep Learning Framework," *IEEE Access,* vol. 7, pp. 68446-68454, 2019.

[8] E. Ibarra, A. Antonopoulos, E. Kartsakli, J. Rodrigues and C. Verikoukis, "QoS-Aware Energy Management in Body Sensor Nodes Powered by Human Energy Harvesting," *IEEE Sensors Journal,* vol. 16, no. 2, 28 Sep. 2015.

[9] R. E. Fan, K. W. Chang, C. J. Hsieh, X. R. Wang and C. J. Lin, "LIBLINEAR: a library for large linear classification," *Journal of Machine Learning Research,* vol. 9, no. 9, pp. 1871-1874, Aug. 2008.

[10] D. Banerjee, K. Islam, G. Mei, L. Xiao, G. Zhang, R. Xu, S. Ji and J. Li, "A Deep Transfer Learning Approach for Improved Post-Traumatic Stress Disorder Diagnosis," in *2017 IEEE International Conference on Data Mining*, New Orleans, LA, USA, 2017.

[11] A. Jafari, A. Ganesan, C. S. K. Thalisetty, V. Sivasubramanian, T. Oates and T. Mohsenin, "SensorNet: A Scalable and Low-Power Deep Convolutional Neural Network for Multimodal Data Classification," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 66, no. 1, pp. 274-287, Jan. 2019.

[12] H. Burrows, J. Zarrin, L. B. Saheer and M. M. D. Oghaz, "Realtime Emotional Reflective User Interface Based on Deep Convolutional Neural Networks and Generative Adversarial Networks," *Electronics,* vol. 11, no. 1, 2022.

[13] K. Lin, F. Xia, W. Wang, D. Tian and J. Song, "System Design for Big Data Application in Emotion-Aware Healthcare," *IEEE Access,* vol. 4, pp. 6901-6909, 2016.

[14] A. Andreas, C. X. Mavromoustakis, G. Mastorakis, S. Mumtaz, J. M. Batalla and E. Pallis, "Modified Machine Learning Techique for Curve

Fitting on Regression Models for COVID-19 projections," in *IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks*, 2020.

[15] A. Andreou, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla and E. Pallis, "Evaluation of the COVID-19 Era by Using Machine Learning and Interpretation of Confidential Dataset," *Electronics,* vol. 10, no. 23, 24 Nov. 2021.

[16] A. Andreas, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, J. N. Sahalos, E. Pallis and E. Markakis, "Enhancement of COVID-19 Detection by Unravelling its Structure and Selecting the Optimal Attributes," in *IEEE Global Communications Conference*, 2022.

[17] H. Shao, M. Xia, G. Han, Y. Zhang and J. Wan, "Intelligent Fault Diagnosis of Rotor-Bearing System Under Varying Working Conditions With Modified Transfer Convolutional Neural Network and Thermal Images," *IEEE Transactions on Industrial Informatics,* vol. 17, no. 5, pp. 3488-3496, May 2021.

[18] Z. Ma, D. Chang, J. Xie, Y. Ding, S. Wen, X. Li, Z. Si and J. Guo, "Fine-Grained Vehicle Classification With Channel Max Pooling Modified CNNs," *IEEE Transactions on Vehicular Technology,* vol. 68, no. 4, pp. 3224-3233, Apr. 2019.

[19] Z. Liu, Z. Shen, M. Savvides and K. T. Cheng, ReActNet: Towards Precise Binary Neural Network with Generalised Activation Functions, vol. 12359, Springer, 2020.

[20] L. Chen, J. Fu, Y. Wu, H. Li and B. Zheng, "Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals," *Sensors,* vol. 20, no. 3, 26 Jan. 2020.

[21] J. Li, I. Tobore, Y. Liu, A. Kandwal, L. Wang and Z. Nie, "Non-invasive Monitoring of Three Glucose Ranges Based On ECG By Using DBSCAN-CNN," *IEEE Journal of Biomedical and Health Informatics,* vol. 25, no. 9, pp. 3340-3350, Sept. 2021.

[22] R. Song, H. Chen, J. Cheng, C. Li, Y. Liu and X. Chen, "PulseGAN: Learning to Generate Realistic Pulse Waveforms in Remote Photoplethysmography," *IEEE Journal of Biomedical and Health Informatics,* vol. 25, no. 5, pp. 1373-1384, May 2021.

[23] A. Andreas, C. X. Mavromoustakis, H. Song and J. M. Batalla, "CNN-Based Emotional Stress Classification using Smart Learning Dataset," in *Smart Data*, Espoo, Finland, 2022.

[24] Y. Xue, X. Wang and Z. Gao, "Multi-classfication Sentiment Analysis Based on the Fused Model," in *IEEE 31st International Conference on Tools with Artificial Intelligence*, 2019.

[25] L. Cai, Y. Song, T. Liu and K. Zhang, "A Hybrid BERT Model That Incorporates Label Semantics via Adjustive Attention for Multi-Label Text Classification," *IEEE Access,* vol. 8, pp. 152183-152192, 2020.

[26] J. He, K. Li, X. Liao, P. Zhang and N. Jiang, "Real-Time Detection of Acute Cognitive Stress Using a Convolutional Neural Network From Electrocardiographic Signal," *IEEE Access,* vol. 7, pp. 42710-42717, 2019.

[27] K. Chen, K. Franko and R. Sang, "Structured Model Pruning of Convolutional Networks on Tensor Processing Units," 9 Jul. 2021.

[28] D. Ayata, Y. Yaslan and M. Kamasak, "Emotion Based Music Recommendation System Using Wearable Physiological Sensors," *IEEE Transactions on Consumer Electronics,* vol. 64, no. 2, pp. 196 - 203, May 2018.

[29] M. Wazid, A. K. Das and S. Shetty, "BSFR-SH: Blockchain-Enabled Security Framework Against Ransomware Attacks for Smart Healthcare," *IEEE Transactions on Consumer Electronics,* vol. 69, no. 1, pp. 18 - 28, Feb. 2022.

[30] Y. Li, G. Han, S. Huang, C. Liu, M. Zhang and F. Wu, "Exploiting Metadata to Estimate Read Reference Voltage for 3-D nand Flash Memory," *IEEE Transactions on Consumer Electronics,* vol. 69, no. 1, pp. 9 - 17, Feb. 2023.

[31] M. K. Yi, W. K. Lee and S. O. Hwang, "A Human Activity Recognition Method Based on Lightweight Feature Extraction Combined with Pruned and Quantized CNN for Wearable Device," *IEEE Transactions on Consumer Electronics,* 2023.

[32] S. Lim, M. Shin and J. Paik, "Point Cloud Generation Using Deep Adversarial Local Features for Augmented and Mixed Reality Contents," *IEEE Transactions on Consumer Electronics,* vol. 68, no. 1, pp. 69 - 76, Feb. 2022.

[33] T.-M. Chen, Y.-H. Tsai, H.-H. Tseng, K.-C. Liu, J.-Y. Chen, C.-H. Huang, G.-Y. Li, C.-Y. Shen and Y. Tsao, "SRECG: ECG Signal Super-resolution Framework for Portable/Wearable Devices in Cardiac Arrhythmias Classification," *IEEE Transactions on Consumer Electronics,* Jan. 2023.

[34] L. Rachakonda, A. K. Bapatla, S. P. Mohanty and E. Kougianos, "SaYoPillow: Blockchain-Integrated Privacy-Assured IoMT Framework for Stress Management Considering Sleeping Habits," *IEEE Transactions on Consumer Electronics,* vol. 67, no. 1, pp. 20 - 29, Feb. 2021.

[35] J. Bai, S. Lian, Z. Liu, K. Wang and D. Liu, "Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People," *IEEE Transactions on Consumer Electronics,* vol. 64, no. 1, pp. 136 - 143, Feb. 2018.

[36] L. Sanabria-Russo, J. Serra, D. Pubill and C. Verikoukis, "CURATE: On-Demand Orchestration of Services for Health Emergencies Prediction and Mitigation," *IEEE Journal on Selected Areas in Communications,* vol. 39, no. 2, pp. 438 - 445, Feb 2021.

[37] Y. Pang, M. Sun, X. Jiang and X. Li, "Convolution in Convolution for Network in Network," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 29, no. 5, pp. 1587-1597, May 2018.

[38] S. Koldijk, M. Sappelli, S. Verberne, M. Neerincx and W. Kraaij, "The SWELL Knowledge Work Dataset for Stress and User Modeling Research," in *16th ACM International Conference on Multimodal Interaction*, Istanbul, Turkey, 2014.

[39] "sklearn.svm.LinearSVC," [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html. [Accessed 01 Aug. 2023].

[40] H. M. Ahmed, M. J. Awan, N. S. Khan, A. Yasin and H. M. F. Shehzad, "Sentiment Analysis of Online Food Reviews using Big Data Analytics," *Elementary Education Online,* vol. 20, no. 2, pp. 827-836, 2021.

[41] "Neural network models (supervised)," [Online]. Available: https://scikit-learn.org/stable/modules/neural_networks_supervised.html. [Accessed 01 Aug. 2023].

[42] N. S. Jaddu, S. S. R. S and S. A., "Voice Emotion Detection: Acoustic Features Extraction Using Multi-layer Perceptron Classifier Algorithm," in *International Conference on Innovative Computing and Communications*, Singapore, 2023.

[43] "sklearn.linear_model.SGDClassifier," [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html. [Accessed 01 Aug. 2023].

[44] C. Yang, X. Lai, Z. Hu, Y. Liu and P. Shen, "Depression Tendency Screening Use Text Based Emotional Analysis Technique," *Journal of Physics: Conference Series,* vol. 1237, no. 3, 2019.

[45] "sklearn.linear_model.RidgeClassifier," [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.RidgeClassifier.html. [Accessed 01 Aug. 2023].

[46] G. G. Jayasurya, S. Kumar, B. K. Singh and V. Kumar, "Analysis of Public Sentiment on COVID-19 Vaccination Using Twitter," *IEEE Transactions on Computational Social Systems,* vol. 9, no. 4, pp. 1101 - 1111, 08 Nov. 2021.

[47] J. Sun and M. Zhao, "Attention-Based Recursive Autoencoder For Sentence-Level Sentiment Classification," in *International Conference on Pattern Recognition, Machine Vision and Intelligent Algorithms*, 2023.

[48] W. Li, F. Qi, M. Tang and Z. Yu, "Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification," *Neurocomputing,* vol. 387, pp. 63-77, Apr. 2020.