Evaluating Mental Well-being through Wearable Sensors Utilizing Machine Learning

Aarti[1,†], Swathi Gowroju[2,*,†], Raju Pal[3,†] and Pushpendra Kumar Rajput[4,†]

[1]Lovely Professional University, Punjab, INDIA
[2]Sreyas Institute of Engineering and Technology, Hyderabad, Telangana, INDIA
[3]Jaypee Institute of Information Technology,Noida, Uttar Pradesh, INDIA
[4]Sharda University, Gr. Noida, INDIA

Abstract

Stress is known to play a significant role in the development of serious medical illnesses such diabetes, hypertension, and cardiovascular disease. Numerous studies on the viability of using several physiological markers to detect stress have been conducted in light of the increased emphasis on wearable health monitoring. The goal of this study is to classify individuals based on physiological information, using the easily accessible WESAD (Wearable Stress and Affect Detection) dataset. The main objective is to use this dataset to create algorithms that can forecast stress based on physiological markers. Using the Synthetic Minority Oversampling Technique (SMOTE), a model is developed in this research to improve the precision of stress level detection. SMOTE is designed to balance out dataset imbalances by oversampling the minority class. This study used the SMOTE approach to efficiently balance the dataset groups due to the uneven nature of the data.

Keywords

Artificial intelligence, machine learning, stress detection, physiological signal, mental health, WESAD

1. Introduction

Real-time physiological data gathering for stress evaluation has become easier in recent years because of the development of wearable sensors and physiological monitoring technology. Physiological signals are useful markers of stress levels because they shed light on how the body reacts to stress. These signals provide a clear indicator of physiological arousal and reactivity since they are a direct reflection of the autonomic nervous system's activation and the release of stress hormones. Numerous physiological signals, including the electroencephalogram (EEG), electrocardiogram (ECG), heart rate variability (HRV), skin conductance (SC), electromyography (EMG), and respiration rate, can be continuously monitored thanks to these technologies. Using these physiological signs, there has been an increase in interest in the subject of automatic mental stress detection. This interest results from the requirement to create approaches and tools that can precisely extract features from diverse physiological data, process those signals using machine learning methods, and then analyze those signals. In order to identify stress, a

ACI'23: Workshop on Advances in Computational Intelligence at ICAIDS 2023, December 29-30, 2023, Hyderabad, India
*Corresponding author.
†These authors contributed equally.
aarti.1208@gmail.com (Aarti); swathigowroju@sreyas.ac.in (S. Gowroju); raju3131.pal@gmail.com (R. Pal); pushpendrarajputs@gmail.com (P. K. Rajput)

lot of people now use the WESAD (Wearable Sensor and Affect Detection) dataset. WESAD dataset was utilized by numerous studies to examine stress detection [1, 2, 3, 4, 5].The WESAD dataset includes numerous physiological parameters that were taken from 15 participants under varied circumstances. An Empatica E4 bracelet and a RespiBAN device worn on the chest were used to record these measurements. Blood volume pulse (BVP), electrodermal activity (EDA), skin temperature, and accelerometer data were all recorded using the Empatica E4 wristband. The BVP signal, which is produced from a photoplethysmography (PPG) sensor, was used to extract information on the interval between beats and heart rate. The RespiBAN gadget, on the other hand, captured data from the accelerometer, ECG signal, EDA, electromyography (EMG), respiration rate, temperature, and other sensors. In the WESAD dataset, this combination of sensors and measurements enables a thorough evaluation of physiological reactions during various activities and situations.
The following main reasons are the motivations behind the use of the WESAD dataset for research:

- Multimodal data collection: The WESAD dataset contains information from a number of different modalities, including electrocardiography (ECG), electrodermal activity (EDA), electromyography (EMG), respiration, and temperature.
- Comparability and benchmarking: The availability of a standardized dataset like WESAD makes it possible to benchmark and compare against existing techniques.
- Open accessibility: Researchers can freely access the WESAD dataset.

2. Related Work

The performance of K-nearest neighbors (KNN) models used to classify the WESAD dataset is evaluated in [1]. The K-fold cross-validation parameter and the total number of nearest neighbors taken into consideration are two crucial factors that are changed to conduct the evaluation. The researchers evaluate the effects of various cross-validation procedures on the effectiveness of the models by altering this parameter. The task is addressed by F. D. Martino et al. [2] proposed solution, which makes use of ensemble learners and recurrent neural networks (RNNs). The research uses the Leave-One-Subject-Out (LOSO) cross-validation scheme to evaluate each model's ability to generalize in predicting individual stress scores.

Based on their significance to the classifier and their association with other features, Hsieh et al. [4] stressed the selection of dominating features. For classification, the extreme gradient boosting technique (XGBoost) was used. To manage variations in sensing settings, often known as the noise context, N. Rashid et al. [5] suggested a system dubbed SELF-CARE (Selective Sensor Fusion for Stress Detection). The study measures SELF-CARE performance using wearable sensors that are worn on the wrist and chest. SELF-CARE obtains an accuracy of 86.34% using wrist-based sensors and 86.19% using chest-based sensors for the 3-class stress classification problem. Similar to this, SELF-CARE achieves an accuracy of 94.12% (wrist-based) and 93.68% (chest-based) for the 2-class stress categorization problem. The WESAD dataset was used in [6] to examine the effectiveness of six classifiers. Of the classifiers studied, the random forest (RF) classifier demonstrated the best performance. Overall, the performance of the wrist-worn sensors was worse to that of the sensors positioned on the chest. The wrist-worn sensors performed at their peak level 95.54%.

The performance of the chest-worn sensors, however, was higher at 97.15%. A unique technique for identifying mental stress is presented by S. Ghosh et al. [7], employing the ECG and GSR, two widely utilized physiological signals. The adaptive synthetic minority oversampling (ADASYN) technique was used to alleviate dataset imbalance and guarantee uniform classifier training. The detection of mental stress was then carried out using a multi-class random forest (RF) classifier. Using the WESAD dataset, the performance of the suggested method was assessed. The outcomes show the usefulness of the suggested approach, with a 97.08% overall accuracy. In [8], PPG and EDA data were acquired from the Empatica E4 bracelet and binary categorization was done using Random Forest, Support Vector Machine (SVM), and Logistic Regression.

Certain machine learning approaches were expermented in Ayurveda based physical constituent balancing detection and psychological impact og COVID-19 on hman being [9, 10, 11]. The study's findings showed that the Random Forest model had the maximum stability and, when including all the features, had an accuracy of 76.5%. A novel multimodal artificial intelligence (AI)-based technique for stress detection and categorization as well as the recognition of stress patterns across time is presented by Rahee W. et al. in their paper published in Science Translational Medicine [12]. The proposed method achieves accuracy of 96.07% using the ANN and SWELL-KW dataset.

3. Proposed Methodology:

The major goal of the study paper is to suggest a novel and promising method for recognizing stress using CNN and encoding the multiple-variant time series dataset to GAF pictures following accurate pre-processing, required transformation, and normalization of the dataset. When it comes to the WESAD dataset, each subject's chest data that has been captured was taken, removed, and transformed to data frames with the chest sensor keys serving as the columns. The labels were taken separately from the data frames. The labels list the stress level, which ranges from '0' to '3'. The proposed methodology is shown in Fig. 1.

In our method, we classify each person's state every 0.1 seconds. PPG value (ppg), PPG autocorrelation value (ppgau), HRV value (hrv), and EDA value (eda) are the four measurements that make up each state. Therefore, we considered for each x_t =ppgt, ppgaut, hrvt, edat as a sample at time t as the feature vector, where t equals sampled at intervals of 0.1 seconds. Each sample xt was tagged with the current condition of the subject. i.e., under stress or not. Consequently, our dataset consisted of the measurements. With their matching label as D =, the results that were achieved at each time interval. (Xt, Lt), where L = "stressed, not stressed" is the formula. The person's condition was the activity that individual was engaged in at time t.

3.1. Dataset Description:

The WESAD dataset includes information gathered from people who were exposed to emotional and stressful stimuli. This data collection was conducted in a controlled laboratory setting with 15 volunteers, 3 of whom were female. Each participant underwent three basic effect

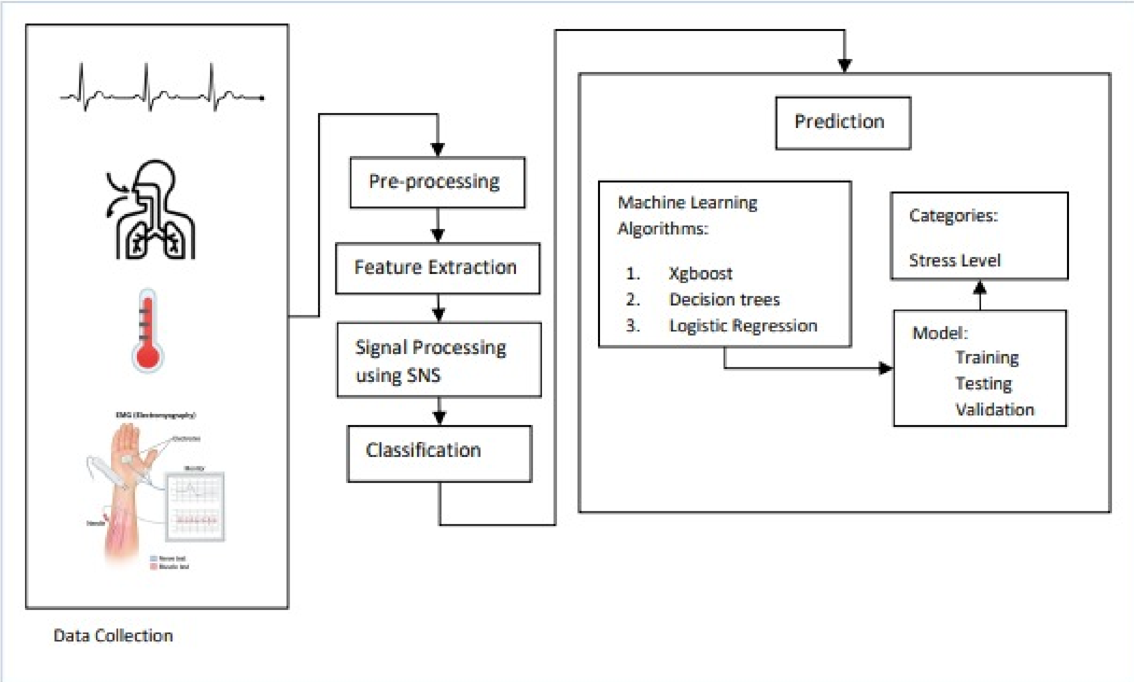

Figure 1: Proposed Methodology

conditions: "baseline" (consisting of an impartial reading task), "amusement" (where subjects watched comedic video clips), and "stress" (using the Trier Social Stress Test - TSST).This dataset includes physiological and movement data that was obtained from a wrist- and chest-worn device. Electrodermal activity (EDA), electrocardiogram (ECG), blood volume pulse (BVP), electromyogram (EMG), respiration (RESP), skin temperature (TEMP), and an accelerometer with three axes (ACC) are some of the sensor modalities. For the purposes of comparison with the stress class in a binary classification task, we combine the baseline and amusement states into a non-stress category in our particular inquiry.

3.2. Pre-processing and Feature Extraction:

With the exception of EMG-features, all physiological signal-based characteristics were calculated across a window size of 60 seconds.To recognize individual heartbeats, peak detection techniques were applied to the raw ECG/BVP signals. By monitoring the time interval between successive peaks, the heart rate (HR) was determined using these identified peaks as reference points. The mean and standard deviation were also calculated from these heart rate readings in addition to the HR calculation.

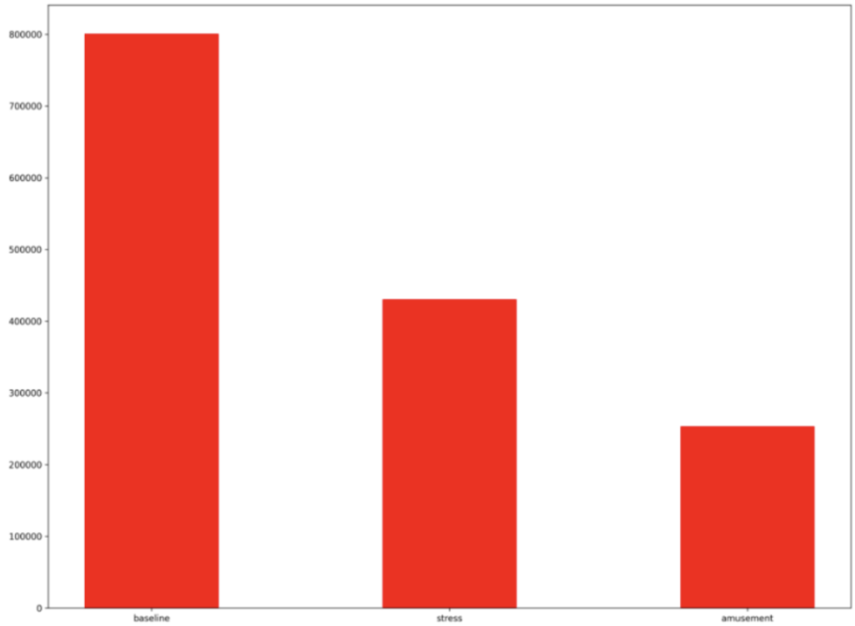

Figure 2: Total no of Samples in each Label

3.3. The sympathetic nervous system (SNS):

SNS has an impact on the EDA signal, making it more sensitive to states of increased arousal. The raw EDA signal was initially passed through a 5 Hz low pass filter. A tonic segment (known as skin conductance level, or SCL) and a phasic segment (known as skin conductance response, or SCR), are also included in the unprocessed EDA signal. The SCR denotes a brief reaction to a stimulus, whereas the SCL represents a steady change in baseline conductivity [13, 14].

The electromyogram (EMG) signal was processed using two different processing steps. A highpass filter was used to remove the DC component in the original sequence. The filtered signal was then divided into 5-second windows by segmentation. Both statistical characteristics and frequency-domain features (such peak frequency) were calculated from these windows. Additionally, seven frequency bands evenly spaced from 0 to 350 Hz were used to compute the power spectral density (PSD) of the EMG signal.The raw EMG signal was subjected to a lowpass filter application in the second processing chain, especially at 50 Hz. This signal after it had been processed was then divided into 60-second segments for study. The respiratory (RESP) signal was first subjected to a bandpass filter with cutoff frequencies of 0.1 and 0.35 Hz. Within the desired frequency range, this filtering procedure assists in isolating pertinent respiratory data. The minima and maxima sites in the filtered RESP signal were then found using a peak detector method. To reduce variances, the data were standardized using the min-max normalization method. The number of samples in each group is shown in Fig. 2. The Synthetic Minority Over-sampling Technique (SMOTE) was used to alleviate dataset imbalance [15]. SMOTE is a well-liked method for addressing class imbalance in datasets for machine learning. Traditional machine learning methods may fail to perform well when faced with imbalanced datasets when one class contains much less samples than the other because they may be biased towards the dominant class.

3.4. Classification:

Linear Discriminate Analysis (LDA), Decision Tree, XGBoost, and Logistic Regression were the four machine learning techniques used in the study. The main goal was binary classification, with the intention of classifying people as either stressed or unstressed. The performance of various classifiers on the WESAD dataset is shown in Table 1. Entropy, a measure of information gain, was used as the criterion for evaluating splits in the instance of Decision Tree classifiers, which had a maximum depth of 5. For XGBoost, the number of estimators (trees) was set to 50, the maximum depth of trees was set at 5, and the learning rate was set at 0.1. In order to maximize the model's convergence in Logistic Regression, the Newton-cg solver was used, with a maximum iteration count set at 1000. The Singular Value Decomposition (svd) solver was chosen for linear discriminant analysis. A 10-fold cross-validation strategy was used for all of these models, enabling thorough performance analysis.

Table 1
Performance Comparison of different Classifiers

Techniques	Precision	Recall	F1-Score	Accuracy
Decision Tress	85.76%	96.62%	90.34%	92.27%
Xgboost	90.83%	97.14%	93.16%	94.22%
Logistic Regression	90.21%	90.36%	90.19%	90.36%
LDA	71.48%	84.20%	74.55%	71.60%

4. Results and Discussion

The following experiment was carried out on a DELL Inspiron 15 5518 laptop with an 11th generation Intel Core processor, 16 GB of memory, and 8 GB of random access memory (RAM). The whole source code for this experiment was produced using a Jupyter notebook on an Ubuntu 22.04 1 LTS 64-bit computer. We also take note of the four evaluation metrics, accuracy, precision, recall, and F1 score, used in the current work before calculating the training and testing accuracies reached by the current image-encoding-based deep neural network. From table 1 it is observed that XGBoost outperforms all the other classifiers with maximum accuracy of 94.22%.

The experiment was run in the WESAD dataset to forecast a person's degree of stress, which ranged from 0 (Baseline) to 3 (Amusement). A promising training accuracy of 99.48% and a testing accuracy of 94.77% were attained after training for roughly 100 epochs. The proposed image-encoding-based deep neural network model for the WESAD dataset yielded the confusion matrix, which is displayed in Fig. 3 with the predicted labels on the X-axis and the actual labels

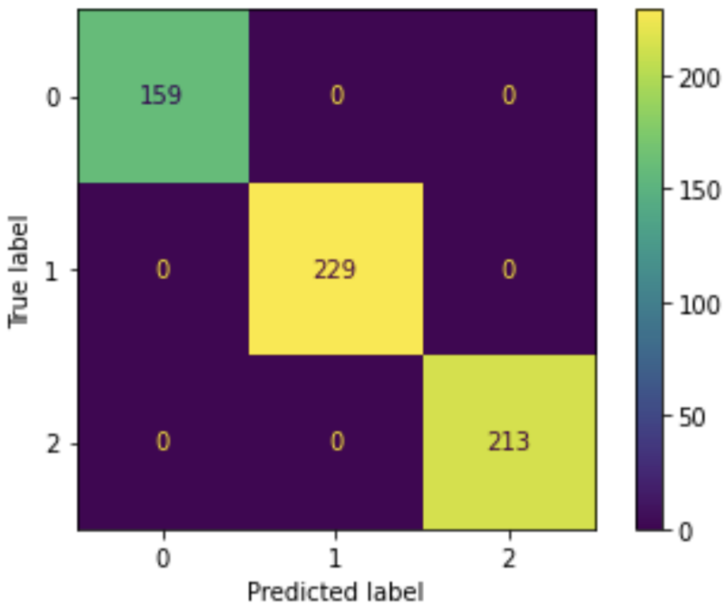

Figure 3: Confusion matrix for the labeled data

of the data on the Y-axis. Table 2 presents the accuracy, precision, recall, and F1 score for each stress identification label as well as the average of all the individual stress-wise performances for the proposed model for the WESAD dataset. Fig. 4 show that as the model is trained for a larger number of epochs, the classification accuracy increases while the graphs presented for loss function substantially reduce. The results also show that, in comparison to other relevant works without the use of encoding time series images, the accuracy is increased when a multivariate time series dataset is encoded to its corresponding image.

We ran the study on the two benchmark datasets, calculated each class's accuracy as well as their F1 score, recall, and precision, and averaged them all to produce the results shown in Table 4. Table 4 shows that the proposed image-encoding-based deep neural network achieves classification accuracy for the WESAD and SWELL datasets of 94.77% and 99.39%, respectively. The SWELL dataset's data length is shorter than the WESAD dataset's, and it required a lot fewer epochs to train the model. Therefore, the plots of accuracy versus epoch size and loss function versus epoch size are not required in that case because the difference in the number of epochs is so small.

5. Future Work

By addressing these issues and enhancing the WESAD dataset, researchers would be able to investigate more intricate research questions, create more precise models, and learn novel affective computing and activity detection insights.

The WESAD dataset can be connected with a number of new trends and technologies to

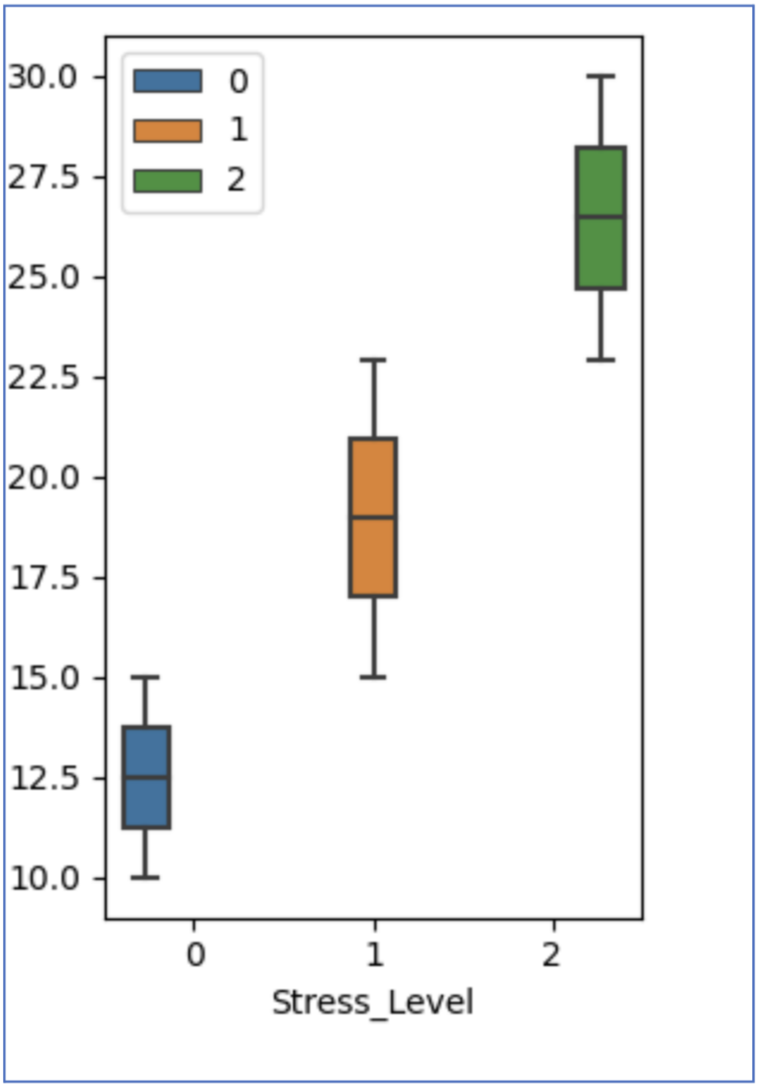

Figure 4: Predicted stress graph

expand its capabilities and open up new research opportunities, including:Integrating the WESAD dataset with wearable devices equipped with advanced sensors can provide richer and more comprehensive physiological data. For instance, incorporating devices that capture brain activity (e.g., electroencephalography, EEG) or eye-tracking technology would enable researchers to study cognitive processes and visual attention alongside physiological responses.

The WESAD dataset can be used in conjunction with mobile sensing technology, such as smartphones or smartwatches, to gather a wider variety of data and offer contextual information. In addition to physiological signals, mobile sensing enables researchers to gather information on location, activity, and social interactions. This provides a more comprehensive understanding of affecting states and activity recognition in real-world environments.

Multi-modal analysis can be facilitated by combining the WESAD dataset with information from other sources, such as social media, sentiment analysis, or natural language processing. The accuracy and applicability of models would be increased by combining physiological data with textual and contextual information to provide a more thorough understanding of affecting states and activity patterns.

6. Conclusion

Different human physiological indicators have been studied in recent study to evaluate and monitor degrees of physical and mental stress. Many of these signs have been used in wearable sensor-driven devices on their own. However, the focus of this study is on utilizing physiological cues for stress detection. This research has carefully assessed the effectiveness of several classifiers through a meticulous comparative analysis. SMOTE was used to the problem of imbalanced datasets in order to improve model accuracy, producing better outcomes. The results particularly emphasize the XGBoost classifier's outstanding performance, which led to an impressive accuracy of 94.22%.

References

[1] D. Bajpai, L. He, Evaluating KNN performance on WESAD dataset, in: 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, 2020.

[2] Praveenkumar., T. Karthick, Automatic stress recognition system with deep learning using multimodal psychological data, in: 2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC), IEEE, 2022.

[3] F. D. Martino, F. Delmastro, High-resolution physiological stress prediction models based on ensemble learning and recurrent neural networks, in: 2020 IEEE Symposium on Computers and Communications (ISCC), IEEE, 2020.

[4] C.-P. Hsieh, Y.-T. Chen, W.-K. Beh, A.-Y. A. Wu, Feature selection framework for XGBoost based on electrodermal activity in stress detection, in: 2019 IEEE International Workshop on Signal Processing Systems (SiPS), IEEE, 2019.

[5] N. Rashid, T. Mortlock, M. A. A. Faruque, Stress detection using context-aware sensor fusion from wearable devices (2023). arXiv:2303.08215.

[6] R. Gupta, A. Bhongade, T. K. Gandhi, Multimodal wearable sensors-based stress and affective states prediction model, in: 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2023.

[7] S. Ghosh, S. Mukhopadhyay, R. Gupta, A new physiology-based objective mental stress detection technique with reduced feature set and class imbalanced dataset management, in: 2021 IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES), IEEE, 2021.

[8] S. Campanella, A. Altaleb, A. Belli, P. Pierleoni, L. Palma, A method for stress detection

using empatica E4 bracelet and machine-learning techniques, Sensors (Basel) 23 (2023) 3565.
[9] C. Gupta, D. Gaur, P. Agrawal, D. Virmani, Huda_covid human disposition analysis during covid-19 using machine learning, International Journal of E-Health and Medical Communications (IJEHMC) 13 (2021) 1–15.
[10] R. Kaur, V. Madaan, P. Agrawal, Fuzzy expert system to calculate the strength/immunity of a human body, Indian Journal of Science and Technology (2016).
[11] K. Ranjit, M. Vishu, A. Prateek, S. S. Kumar, K. Amandeep, Fuzzy expert system for identifying the physical constituents of a human body, Indian Journal of Science and Technology (2016).
[12] R. Walambe, P. Nayak, A. Bhardwaj, K. Kotecha, Employing multimodal machine learning for stress detection, J. Healthc. Eng. 2021 (2021) 1–12.
[13] S. Kim, W. Rhee, D. Choi, Y. J. Jang, Y. Yoon, Characterizing driver stress using physiological and operational data from real-world electric vehicle driving experiment, Int. J. Automot. Technol. 19 (2018) 895–906.
[14] W. Du, Application of improved SMOTE and XGBoost algorithm in the analysis of psychological stress test for college students, J. Electr. Comput. Eng. 2022 (2022) 1–8.
[15] S. Gowroju, S. Kumar, Aarti, A. Ghimire, Deep neural network for accurate age group prediction through pupil using the optimized UNet model, Math. Probl. Eng. 2022 (2022) 1–24.
[16] A. Ragav, N. H. Krishna, N. Narayanan, K. Thelly, V. Vijayaraghavan, Scalable deep learning for stress and affect detection on resource-constrained devices, in: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), IEEE, 2019.