

# CNN and Random Forest based Stress Management System using Physiological Signals

Ms. Isha Amte, Ms. Shravani Ghadi, Ms. Sharvari Jadhav, Dr. Seema M. Hanchate, Ms. Poonam More  
Department of Electronics and Communication  
Usha Mittal Institute of Technology, SNT Women's University, Mumbai, India.

**Abstract** - Stress detection using wearable devices has gained significant importance due to its impact on physical and mental well-being. This paper proposes a hybrid Convolutional Neural Network–Random Forest (CNN–RF) based stress detection system using wrist-based physiological signals including Blood Volume Pulse (BVP), Electrodermal Activity (EDA) and Skin Temperature (TEMP). The approach combines deep learning-based feature extraction with handcrafted physiological features to capture both temporal patterns and domain-specific characteristics. A modified three-class framework is introduced by redefining WESAD labels into Low, Moderate and High stress levels making the problem more practical and challenging. The model achieves an accuracy of 80.25% under this setting using Leave-One-Subject-Out (LOSO) cross-validation. The model is also evaluated under a binary classification setting (No Stress vs Stress), achieving 93.46% accuracy and under the standard three-class configuration (baseline, amusement and stress), achieving 82.71% accuracy. The lightweight CNN–RF architecture ensures low computational complexity making it suitable for real-time wearable applications. Additionally, the system demonstrates application-level integration for stress visualization and user feedback. Overall, the proposed system provides an efficient, robust and practical solution for real-time stress detection and monitoring

**Keywords**—Stress Detection, Multimodal Signals, CNN, Random Forest, WESAD, LOSO, Wearable Devices.

## I. INTRODUCTION

Stress has become a common part of daily life and significantly affects both physical and mental well-being [1]. While moderate levels of stress can be manageable prolonged or excessive stress can negatively impact daily performance decision-making ability and overall health [2]. Early studies have demonstrated the importance of detecting stress in real-world environments using physiological signals. Chronic stress has been associated with serious health conditions such as cardiovascular disorders, depression and weakened immune response [3]. The increasing prevalence of stress in modern lifestyles has led to growing interest in real-time stress detection systems for timely intervention and improved quality of life [4], [5] With the advancement of wearable technology continuous monitoring of physiological signals has become feasible and widely adopted [6], [7]. Signals such as Electrodermal Activity (EDA), Blood Volume Pulse (BVP) and skin temperature (TEMP) are strongly

correlated with the activity of the autonomic nervous system and provide reliable indicators of stress [8]. These signals can be collected in real-world settings using minimally intrusive wearable devices enabling continuous and real-time stress monitoring [9], [10]. The availability of benchmark datasets such as WESAD has further accelerated research in this domain by providing multimodal physiological data recorded under different affective states [11]. Although recent advancements have improved stress detection performance ensuring both high accuracy and computational efficiency remains important for real-time applications as many existing approaches rely on complex models that increase computational cost [12], [13]. Furthermore, while subject-dependent models often achieve higher accuracy they require user-specific training data making them less practical for real-world deployment [13]. Subject-independent models although more scalable are more challenging to optimize and require robust generalization techniques [14]. In recent years several machine learning and deep learning approaches have been proposed for stress detection. Traditional machine learning models such as Random Forest have demonstrated strong performance on physiological data [15], [16] while subject-independent approaches using multimodal signals have shown promising results [17]. Deep learning techniques particularly Convolutional Neural Networks (CNN) have proven effective in automatically extracting meaningful features from physiological signals [18], [19]. Additionally, sensor fusion approaches have been explored to improve prediction accuracy by combining multiple physiological signals [20], [21]. Recent studies further highlight the effectiveness of hybrid machine learning and deep learning models in improving generalization and performance [22], [23]. Building upon these developments this paper proposes a subject-independent stress detection system using multimodal physiological signals such as BVP, EDA and TEMP acquired from wearable sensors with model development and validation performed using the WESAD dataset. The proposed approach combines deep learning and machine learning techniques by employing a 1D Convolutional Neural Network (CNN) for feature extraction followed by the fusion of statistical and Heart Rate Variability (HRV) features [24], [25]. The combined feature representation is then classified using a Random Forest (RF) model to categorize stress levels into three classes: Low, Moderate and High. The proposed system is evaluated using a Leave-One-Subject-Out (LOSO) cross-validation strategy to ensure unbiased and generalizable performance. The system presents the predicted stress levels to the user through an application interface along with simple recommendations for managing higher stress levels. Experimental results demonstrate that the proposed hybrid CNN+RF model achieves competitive performance and benefits significantly from multimodal feature

fusion. Overall, the system provides an efficient, computationally lightweight and practical solution for real-time stress monitoring using wearable devices.

## II. RELATED WORK

Stress detection using physiological signals has been widely studied with continuous improvements in methodology, accuracy and real-time applicability. Early foundational work by Healey and Picard demonstrated the feasibility of detecting stress using physiological signals in real-world scenarios [1]. Subsequent studies established the importance of heart rate variability (HRV) and autonomic nervous system activity as reliable indicators of stress [2], [3]. The increasing adoption of wearable sensors has enabled continuous monitoring of physiological signals making stress detection more practical in real-world environments [6], [10]. The introduction of benchmark datasets such as WESAD by Schmidt et al. further accelerated research by enabling the development and evaluation of multimodal stress detection systems [11]. Early machine learning approaches demonstrated the feasibility of stress detection using physiological data. Bobade and Vani achieved 84.32% accuracy using Random Forest models on WESAD [12], while Ninh et al. proposed a subject-independent approach using BVP, EDA, and TEMP signals, achieving approximately 88% accuracy [13]. Although effective, these approaches rely primarily on handcrafted features and may struggle to capture complex signal patterns. Deep learning methods have shown improved performance by automatically extracting meaningful features from physiological signals. Ghosh et al. achieved high accuracy using an image-encoding-based deep neural network [18] and Benita et al. combined CNN with Random Forest for improved classification performance [22]. However, such approaches often involve higher computational complexity limiting their suitability for real-time applications. Researchers have also explored sensor fusion and multimodal approaches. Rashid et al. proposed a multimodal stress detection framework using physiological signals [17] while Zhu et al. focused on wrist-based electrodermal activity (EDA) signals for stress detection [8]. Abdelfattah et al. further investigated hybrid machine learning and deep learning models highlighting their effectiveness on multimodal datasets [23]. Andrić et al. demonstrated that Random Forest-based models can achieve high accuracy in stress detection tasks [15] although such models may lack deep feature representation. Several review studies have summarized advancements in this domain. Gedam and Paul [4], Giannakakis et al. [5] and Iqbal et al. [10] emphasized the importance of multimodal physiological signals for reliable stress detection while Siirtola highlighted the need for subject-independent and generalizable models [14]. Physiological signal analysis plays a crucial role in stress detection. Shaffer and Ginsberg established HRV as a standard indicator of stress [24] while Castaldo et al. validated the use of short-term HRV features for stress detection [25]. Kreibig analyzed autonomic nervous system activity in emotional states [3] and Greco et al. introduced the cvxEDA method for electrodermal activity signal decomposition [26]. Despite these advancements efficient feature extraction remains a challenge. From a modeling perspective, Winata et al. demonstrated the effectiveness of deep convolutional neural networks for physiological signal analysis [19], while He et al. introduced deep residual learning concepts that influenced modern deep learning architectures [27]. Breiman developed the Random Forest algorithm [16], which remains widely used due to its

robustness and performance on physiological datasets. Additionally, Makowski et al. introduced NeuroKit2, a standardized toolkit for physiological signal processing [28]. Recent works have also explored practical implementations. Tyulepberdinova et al. developed an ESP32-based stress monitoring system [29] while IoT-based CNN models have been proposed for physiological signal analysis [30]. Akmandor and Jha introduced wearable health monitoring systems [7] and Zhu et al. along with Phadke et al. explored wearable IoT-based stress monitoring solutions [31], [32]. Additional implementations using machine learning and embedded systems have also been reported [33] although these systems often increase computational complexity. Methodological improvements such as segmentation strategies have also been investigated. Strinar et al. showed that window size significantly impacts stress detection performance [34] while Eren and Saraç demonstrated effective stress detection using BVP and EDA signals [20]. Supporting studies have further contributed to the field. Gjoreski et al. explored wrist-based stress detection systems [9], [21], while Plarre et al. proposed continuous stress detection in real-life settings [35]. Nath and Thapliyal applied machine learning techniques for anxiety detection using wearable sensors [36] and Koldijk et al. introduced the SWELL dataset for stress analysis [37]. More recently Wu et al. proposed multimodal transformer-based models for stress detection [38] which although powerful introduce higher computational complexity. Overall, while significant progress has been made in stress detection challenges remain in achieving a balance between accuracy, generalization and computational efficiency. These limitations motivate the need for hybrid approaches that combine deep learning and machine learning while maintaining lightweight performance for real-time stress prediction.

## III. PROPOSED SYSTEM MODEL

### A. System Overview

The proposed stress detection system is designed as an integrated framework that combines wearable sensing, signal processing and machine learning into a unified pipeline. The system operates in three main stages: physiological signal acquisition, hybrid model-based processing and application-level visualization enabling continuous monitoring and real-time stress assessment. In the data acquisition stage wrist-based physiological signals including Blood Volume Pulse (BVP), Electrodermal Activity (EDA) and Skin Temperature (TEMP) are collected using wearable sensors. These signals are selected due to their strong correlation with autonomic nervous system activity and stress responses. The sensors are interfaced with an ESP32 microcontroller that facilitates real-time data acquisition, initial signal handling and transmission of signals to the processing unit. In the processing stage, the acquired physiological signals undergo preprocessing, resampling and window segmentation before being provided as input to the proposed hybrid model. A multi-branch 1D Convolutional Neural Network (CNN) is employed to automatically learn and extract meaningful temporal features from the input signals. In addition, handcrafted physiological features are extracted and combined with the deep features to form a fused representation. These fused features are then utilized by a Random Forest (RF) classifier which performs the final classification of stress levels. The system categorizes stress into three distinct classes: Low,

Moderate and High enabling a more informative and realistic assessment.

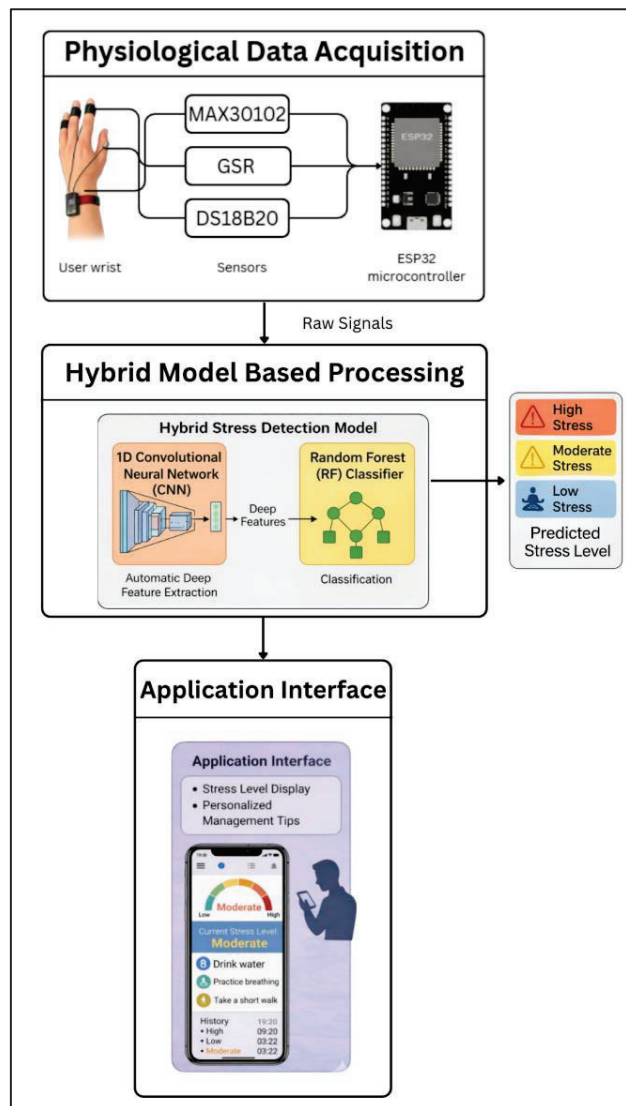


Fig 1: Proposed stress detection system architecture.

In the final stage the predicted stress levels are communicated to a smartphone application. The application serves as a user interface that displays the current stress level, maintains historical records for tracking stress trends over time and provides simple recommendations to help users manage elevated stress levels. This integration of sensing, processing and feedback enables a complete end-to-end system for practical stress monitoring. The overall architecture of the proposed system is illustrated in Fig 1.

### B. Data Acquisition

Physiological signals are widely used for stress detection as they reflect the activity of the autonomic nervous system under different conditions [3]. In this work, three signals are considered: Blood Volume Pulse (BVP), Electrodermal Activity (EDA) and skin temperature (TEMP) as they are reliable indicators of stress responses [5], [8]. The proposed system utilizes wearable sensors for signal acquisition. Specifically, the MAX30102 sensor is used to acquire the BVP signal, the Grove

GSR sensor is used for EDA measurement and the DS18B20 sensor is used for skin temperature. These sensors are interfaced with an ESP32 microcontroller enabling continuous, real-time and non-invasive physiological signal acquisition [29], [32]. For model development and evaluation, the publicly available WESAD dataset is used [11]. This dataset contains multimodal physiological signals recorded under different affective states including baseline, stress, amusement and meditation. In this study only wrist-based signals (BVP, EDA and TEMP) are selected as they closely match the signals obtained from the proposed wearable sensor setup and are more suitable for real-world wearable applications [9], [21]. To formulate the stress detection problem the original WESAD labels are mapped into three classes. The baseline condition is defined as Low stress, the meditation condition as Moderate stress and the stress condition as High stress. This mapping provides a more practical and interpretable representation of stress levels for real-world applications. However, this formulation introduces additional complexity due to overlapping physiological characteristics between moderate and high stress conditions making the classification task more challenging and realistic [14], [38]. The selected signals are used to train and evaluate the proposed model under subject-independent conditions using Leave-One-Subject-Out (LOSO) cross-validation ensuring that the system can generalize effectively across different users [13].

### C. Hardware Setup

The hardware setup used for physiological signal acquisition is shown in Figure 2. The system is developed using an ESP32 microcontroller integrated with wearable sensors for capturing wrist-based physiological signals.

The setup includes:

- MAX30102 sensor for Blood Volume Pulse (BVP).
- GSR sensor for Electrodermal Activity (EDA).
- DS18B20 sensor for skin temperature (TEMP).

The sensors are interfaced with the ESP32 microcontroller enabling continuous and real-time acquisition of physiological signals. The ESP32 platform provides a compact, low-power and cost-effective solution for wearable health monitoring systems [29], [32].

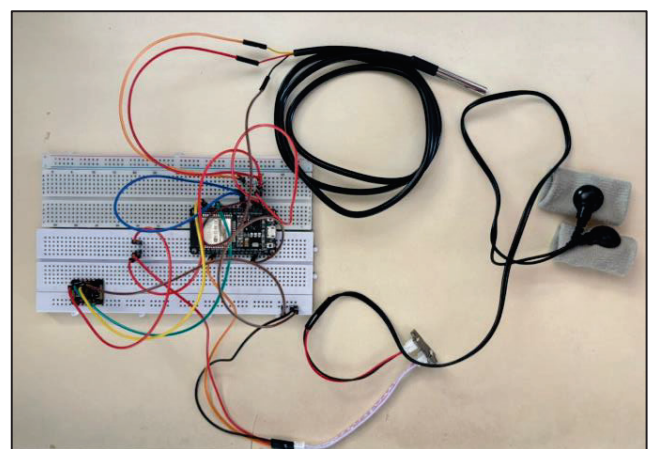


Fig 2: Hardware setup using ESP32 and wearable sensors.

The proposed hardware configuration is designed to be lightweight and suitable for stress monitoring in real-world

environments. Compared to traditional chest-based monitoring systems which often rely on multiple sensors such as ECG and respiration belts the proposed wrist-based setup offers enhanced comfort, portability and usability [6], [7]. This makes it more practical for continuous everyday monitoring while maintaining sufficient physiological information for reliable stress detection [9], [21].

#### D. Signal Preprocessing

Physiological signals acquired from wearable devices are often affected by noise, motion artifacts and inter-subject variability. Therefore, appropriate preprocessing is required to enhance signal quality and ensure reliable feature extraction [4], [10]. In the proposed system signal-specific filtering techniques are applied to remove noise while preserving meaningful physiological components. The Blood Volume Pulse (BVP) signal is processed using a bandpass filter to retain frequency components associated with cardiovascular activity. Electrodermal Activity (EDA) and skin temperature (TEMP) signals are smoothed using low-pass filtering to remove high-frequency noise while preserving relevant trends [26]. These preprocessing steps are essential for improving signal quality and ensuring accurate physiological representation. Since the selected signals are recorded at different sampling rates (BVP at 64 Hz, EDA and TEMP at 4 Hz) a resampling step is performed to align them to a common frequency. Specifically, EDA and TEMP signals are upsampled to 64 Hz to match the sampling rate of the BVP signal ensuring proper temporal synchronization for subsequent processing and model input. To address variations in signal magnitude and the presence of outliers robust scaling is applied to normalize the data. In contrast to conventional approaches the proposed system employs a leakage-free normalization strategy within the Leave-One-Subject-Out (LOSO) framework. For each fold normalization parameters are computed only on the training data and then applied to the corresponding test data preventing information leakage and ensuring unbiased subject-independent evaluation [13]. Additionally, to improve model generalization and handle class imbalance a controlled noise-based data augmentation strategy is applied to the training data. The pre-processed signals are then used for window segmentation and subsequent feature extraction. Overall, the preprocessing pipeline improves signal consistency, reduces noise, prevents data leakage and enhances the robustness of the proposed stress detection system.

#### E. Window Segmentation

After preprocessing the physiological signals are segmented into fixed-length time windows to capture temporal patterns associated with stress responses. Window-based segmentation is a widely used approach in physiological signal analysis as it enables the extraction of meaningful features from short intervals of data while preserving temporal dynamics [34]. In the proposed system a sliding window approach is employed with a window length of 60 seconds and a 75% overlap (15-second stride) between consecutive windows. The choice of a longer window length allows for reliable estimation of physiological features such as heart rate variability (HRV) which require sufficient temporal context for accurate computation [24], [25]. At the same time the use of a smaller stride increases temporal resolution and enables more frequent predictions making the system suitable for real-time monitoring. The use of overlapping windows increases the number of training samples and ensures smoother transitions between consecutive segments. This helps

in capturing gradual changes in physiological signals and improves the robustness of the model by reducing abrupt variations between adjacent windows. Each segmented window is treated as an independent sample and is passed to the subsequent feature extraction stage. This segmentation strategy enables effective learning of temporal patterns while maintaining sufficient contextual information contributing to improved performance of the proposed stress detection system.

#### F. Feature Extraction

Feature extraction plays a crucial role in transforming raw physiological signals into meaningful representations that can be effectively used for stress classification. In the proposed model a hybrid feature extraction approach is adopted combining deep learning-based features with handcrafted physiological features to capture both complex temporal patterns and domain-specific characteristics of the signals.

##### i. Deep Feature Extraction using CNN

The CNN model automatically learns deep feature representations from the segmented physiological signals. Unlike handcrafted features which are explicitly designed based on domain knowledge, deep features are learned directly from data through convolutional operations enabling the model to capture complex temporal patterns in physiological signals [19]. In the proposed multi-branch CNN architecture as shown in Fig 3 each signal modality (BVP, EDA and TEMP) is processed independently through a sequence of convolutional blocks consisting of convolution, batch normalization, ReLU activation and pooling layers. This design allows the model to effectively capture modality-specific temporal characteristics while maintaining robustness to noise and signal variations. Multi-branch architectures have been shown to be effective in multimodal physiological signal analysis as they enable learning both individual signal representations and complementary inter-signal information [19], [38]. Following the convolutional layers a Global Average Pooling (GAP) layer is applied to each branch to reduce feature dimensionality while retaining the most discriminative information and mitigating overfitting [27]. The outputs from the three branches are then concatenated to form a unified deep feature vector.

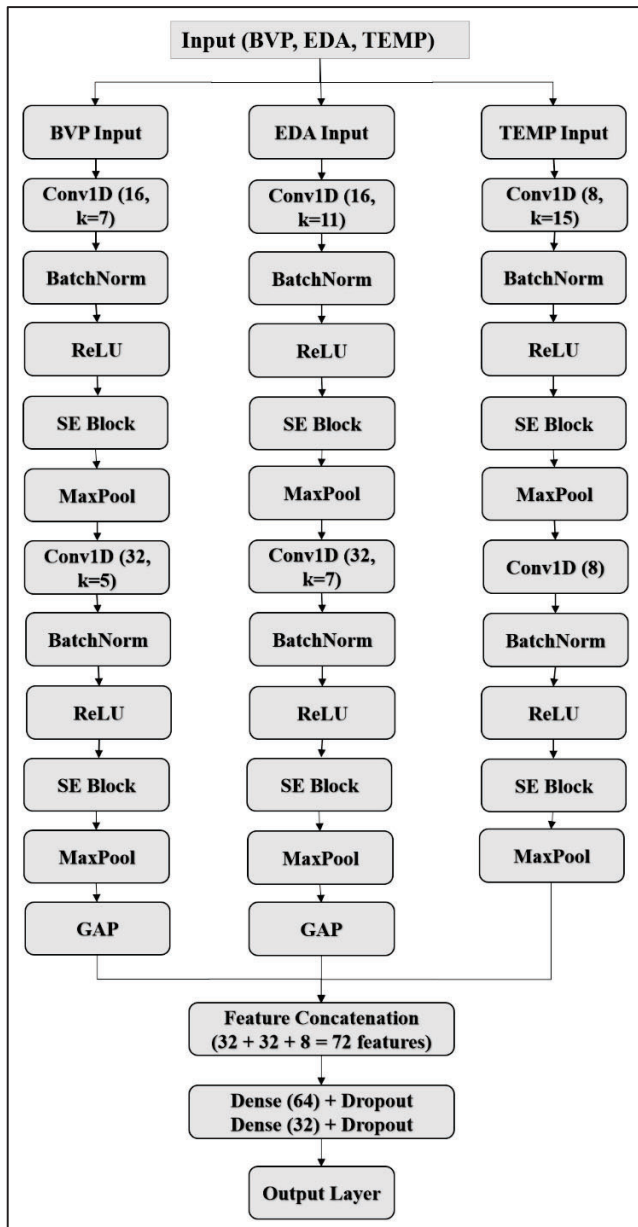


Fig 3: Multi-branch CNN architecture.

The resulting deep feature representation is 72-dimensional comprising:

- 32 features from the BVP branch
- 32 features from the EDA branch
- 8 features from the TEMP branch.

These deep features capture high-level temporal and non-linear characteristics of physiological signals that are difficult to model using handcrafted approaches alone. The extracted deep feature vector is subsequently combined with handcrafted features in the feature fusion stage for final classification.

### ii. Handcrafted Feature Extraction

In addition to deep features handcrafted features are extracted to incorporate domain knowledge related to physiological signal behavior under stress. These features are computed separately

for each signal modality. For the BVP signal, Heart Rate Variability (HRV) features are derived which are widely recognized as reliable indicators of stress and autonomic nervous system activity [2], [24], [25]. A set of statistical and HRV-based features is extracted to capture variations in cardiovascular dynamics. For the EDA signal, statistical features such as mean, standard deviation, minimum and maximum values are computed. Additionally, features related to tonic and phasic components of the EDA signal are considered as they reflect sympathetic nervous system activity [26]. For the TEMP signal statistical features including mean, standard deviation and trend-based features such as slope or rate of change are extracted to capture temperature variations associated with stress responses. In total approximately 32 handcrafted features are obtained by combining features from BVP, EDA and TEMP signals.

Table 1: Handcrafted Features Extracted from Physiological Signals.

Signal	Feature Type	Features
BVP	Statistical	Mean, Standard Deviation, Min, Max
	HRV-based	RMSSD, SDNN, Peak Frequency
EDA	Statistical	Mean, Standard Deviation, Min, Max
	Signal Components	Tonic el (SCL), Phasic Response (SCR)
TEMP	Statistical	Mean, Standard Deviation, Min, Max
	Trend	Slope / Rate of Change

### iii. Feature Fusion

To leverage the strengths of both approaches the deep features extracted using CNN and the handcrafted features are combined to form a unified feature vector. Specifically, the 72-dimensional deep feature vector is concatenated with the 32-dimensional handcrafted feature vector resulting in a final feature representation of 104 dimensions. This fusion strategy enables the model to capture both data-driven patterns and domain-specific characteristics leading improved classification performance. The combined feature set is then used as input to the classification stage. The feature extraction and fusion process is illustrated as part of the overall methodology pipeline in Figure 4.

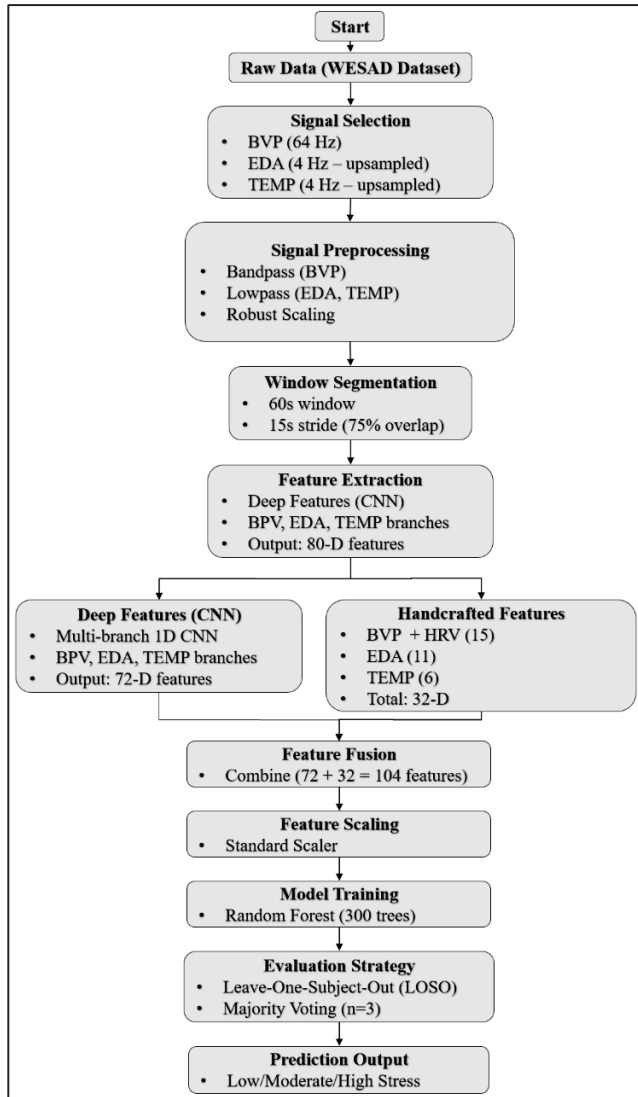


Fig 4: Signal preprocessing pipeline.

### G. Feature Scaling

After feature extraction and fusion, the resulting feature vectors are normalized to ensure consistent scaling across all features. Since the fused feature set consists of both deep learning-based features and handcrafted features with different value ranges normalization is necessary to prevent features with larger magnitudes from dominating the learning process. In this work, StandardScaler is applied to transform the features to zero mean and unit variance defined as:

$$x' = (x - \mu) / \sigma$$

where  $x$  represents the original feature value,  $\mu$  is the mean and  $\sigma$  is the standard deviation. To ensure fair and unbiased evaluation feature scaling is performed in a leakage-free manner within the Leave-One-Subject-Out (LOSO) framework. Specifically, the scaling parameter  $\mu$  and  $\sigma$  are computed only from the training data in each fold and then applied to the corresponding test data. This prevents information leakage and preserves strict subject independence [13]. This scaling step improves the stability, convergence and performance of the

classification model by ensuring that all features contribute equally during training and inference [18].

### H. Class Imbalance Handling

In stress detection tasks class imbalance is a common issue where certain stress categories may have significantly fewer samples compared to others. This imbalance can bias the classification model toward majority classes leading to reduced performance on minority classes. To address this issue in the proposed system a controlled noise-based data augmentation strategy is applied to the training data. This approach generates additional samples by introducing small perturbations to existing signals thereby preserving the underlying physiological characteristics while improving class representation. In addition, class weighting is incorporated during model training to assign higher importance to minority classes particularly the high-stress category. This ensures that the classifier places greater emphasis on correctly identifying underrepresented classes. Both data augmentation and class weighting are applied only to the training data within each Leave-One-Subject-Out (LOSO) fold to maintain strict subject independence and prevent data leakage [13]. This strategy improves the model's ability to generalize across subjects while maintaining a realistic and unbiased evaluation. Overall, the proposed approach effectively mitigates class imbalance without introducing synthetic bias leading to improved classification performance across all stress levels.

### I. Classification Model (Random Forest)

After feature scaling and class balancing the fused feature vectors are used for classification. In this part a Random Forest (RF) classifier is employed to predict stress levels. Random Forest is an ensemble learning method that constructs multiple decision trees and combines their outputs to improve classification performance and robustness [16]. Each decision tree in the forest is trained on a randomly sampled subset of the training data while a random subset of features is considered at each split. This randomness helps in reducing overfitting and improves the generalization capability of the model. The final prediction of the Random Forest model is obtained using majority voting across all decision trees which can be expressed as:

$$\hat{y} = \text{mode}(T_1(x), T_2(x), \dots, T_n(x))$$

where  $T_1, T_2, \dots, T_n$  represents individual decision trees and  $\hat{y}$  is the final predicted class. In the proposed system model the Random Forest classifier is configured with 300 trees to ensure stable and reliable predictions. Additionally, class weighting is incorporated within the classifier to assign higher importance to minority classes particularly the high-stress category thereby improving classification performance under imbalanced conditions. The model takes the fused feature vector as input and classifies each sample into one of three stress levels: Low, Moderate or High. Random Forest is selected due to its robustness to noise, ability to handle high-dimensional feature spaces and relatively low computational complexity compared to deep learning classifiers [15], [16]. These properties make it well-suited for the proposed hybrid framework and real-time stress detection applications.

### J. Evaluation Strategy

To evaluate the performance of the proposed stress detection system a subject-independent evaluation strategy is adopted using Leave-One-Subject-Out (LOSO) cross-validation. LOSO is widely used in physiological signal-based stress detection tasks as it provides a realistic assessment of model generalization across unseen subjects [14]. Evaluation is performed on the redefined three-class stress labels derived from WESAD dataset. In LOSO cross-validation data from one subject is used as the test set while data from all remaining subjects is used for training. This process is repeated for each subject in the dataset ensuring that every subject is used exactly once for testing. The overall performance is then obtained by averaging the results across all folds. This evaluation strategy ensures that the model does not rely on subject-specific patterns and instead learns generalized features that can be applied to unseen users. Such an approach is particularly important for real-world stress monitoring systems where prior data from the end user may not be available. During evaluation standard performance metrics such as Accuracy, Weighted F1-score and Macro F1-score are computed for each fold and the final results are reported as the average across all subjects. This provides a robust and unbiased estimate of the model's performance under subject-independent conditions. Overall, LOSO cross-validation enables a comprehensive and reliable evaluation of the proposed system and demonstrates its effectiveness for real-world stress detection applications.

## IV. RESULT & ANALYSIS

### A. Evaluation Metrics

To evaluate the performance of the proposed stress detection system standard classification metrics including Accuracy, Precision, Recall and F1-score are used. Accuracy represents the overall correctness of the model while Precision and Recall provide insights into class-wise prediction performance. The F1-score defined as the harmonic mean of Precision and Recall provides a balanced evaluation particularly in the presence of class imbalance [27]. For the three classes (Low, Moderate, High) and two classes (Stress, No Stress) these metrics are computed for each class and averaged using macro and weighted strategies to obtain an overall performance measure.

### B. Proposed Model Results

The proposed model is primarily designed for three-class stress classification using redefined WESAD labels:

- Baseline → Low
- Meditation → Moderate
- Stress → High

The model is evaluated using Leave-One-Subject-Out (LOSO) cross-validation. The detailed classification results are shown in Fig. 5. The model achieves an overall accuracy of 80.22% with a weighted F1-score of 80.17% demonstrating strong performance under subject-independent conditions.

CLASSIFICATION REPORT				
	Precision	Recall	F1-score	Support
Low	0.7644	0.8627	0.8106	1151
Moderate	0.7811	0.6975	0.7369	757
High	0.9080	0.8165	0.8599	665
accuracy			0.8022	2573
macro avg	0.8178	0.7923	0.8025	2573
weighted avg	0.8064	0.8022	0.8017	2573

Fig 5: Classification report for three-class.

The proposed model demonstrates strong performance for Low and High stress levels achieving F1-scores of 81.06% and 85.99% respectively. In contrast comparatively lower performance is observed for the Moderate class (F1-score: 73.69%) due to overlapping physiological characteristics between moderate and high stress conditions which is a known challenge in stress detection tasks [14], [38].

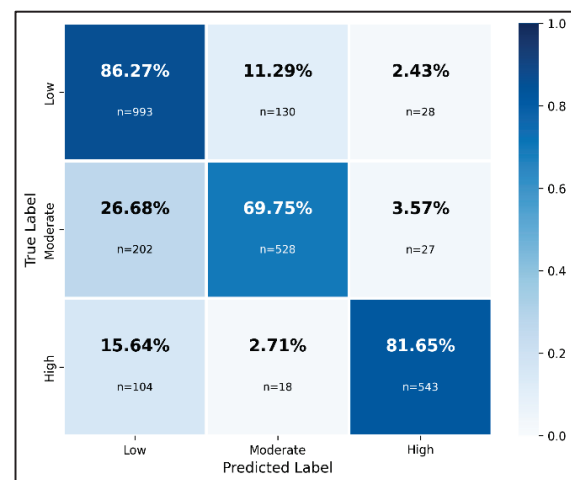


Fig 6: Confusion matrix for three-class classification.

Fig. 6 presents the confusion matrix for the three-class classification. A strong concentration of values along the diagonal indicates correct classification across most samples. Misclassification is primarily observed in the Moderate class, where instances are often confused with Low stress (26.68%) reflecting the subtle transition between relaxed and moderate stress states.

Table 2: Per-subject accuracy (LOSO) for three-class.

Subject	Acc(voted)	Macro F1	CNN val
S2	87.88%	87.97%	95.36%
S3	83.23%	82.02%	93.30%
S4	86.47%	86.77%	94.32%
S5	77.91%	75.45%	91.73%
S6	82.46%	83.79%	94.32%
S7	78.36%	79.22%	93.52%
S8	80.81%	79.52%	91.99%
S9	95.32%	95.13%	95.35%

S10	82.49%	80.50%	93.02%
S11	91.91%	91.24%	90.18%
S13	86.63%	87.21%	91.21%
S14	49.71%	52.31%	86.05%
S15	88.51%	88.49%	93.54%
S16	80.92%	81.26%	91.99%
S17	51.16%	36.43%	93.80%

Mean	80.25%	79.15%
Std	12.59%	14.80%

Per-subject results (Table 2) show variability across individuals which is expected due to inter-subject physiological differences. High performance is observed for subjects such as S9 (95.32%) and S11 (91.91%) while comparatively lower performance is observed for subjects such as S14 (49.71%) and S17 (51.16%). This variation highlights the inherent difficulty of generalizing physiological responses across diverse individuals. Despite this variability the model maintains stable overall performance demonstrating its robustness and generalization capability. The model summary is presented in Fig. 7 highlighting the lightweight architecture (~26K parameters) and the hybrid feature representation consisting of deep and handcrafted features making it suitable for real-time and embedded stress monitoring applications.

Metric	Value
Classification	3-Class: Low / Moderate / High
Label: Low	Baseline (WESAD label 1)
Label: Moderate	Meditation (WESAD label 4)
Label: High	Stress / TSST (WESAD label 2)
<b>Dropped</b>	<b>Amusement (WESAD label 3)</b>
<b>Overall Accuracy</b>	<b>80.22%</b>
<b>Balanced Accuracy</b>	<b>79.23%</b>
<b>Macro F1</b>	<b>80.25%</b>
Weighted F1	80.17%
Low Precision	76.44%
Low Recall	86.27%
Low F1	81.06%
Moderate Precision	78.11%
Moderate Recall	69.75%
Moderate F1	73.69%
High Precision	90.80%
High Recall	81.65%
High F1	85.99%
Evaluation	Full LOSO - 15 subjects
Total Test Windows	2573
Window Size	60s @ 64Hz = 3840 samples
Stride	15s (75% overlap)
Signals	BVP + EDA + TEMP
Feature Vector	72-D CNN + 32-D HC = 104-D -> RF
Dataset	WESAD (Schmidt et al. 2018)

Fig 7: Three-class model summary.

### C. Additional Evaluation Results

Alongside the three-class formulation additional evaluations are also conducted under alternative classification settings to analyze performance across different formulations. In the binary classification setting (No Stress vs Stress) baseline and amusement states are grouped as No Stress while stress is retained as Stress and meditation is excluded. Under this setting the model achieves an overall accuracy of 93.46% with a weighted F1-score of 93.32%. The corresponding classification report and confusion matrix are presented in Fig. 8 and Fig. 9. The results demonstrate strong class separability with high precision and recall for both classes indicating that the model effectively distinguishes between relaxed and stressed conditions under subject-independent evaluation.

- Baseline + Amusement → No Stress
- Stress → Stress
- Meditation → Excluded

	Precision	Recall	F1-score	Support
No Stress	0.9286	0.9821	0.9546	1509
Stress	0.9519	0.8241	0.8834	648
accuracy			0.9346	2157
macro avg	0.9402	0.9031	0.9190	2157
weighted avg	0.9356	0.9346	0.9332	2157

Fig 8: Classification Report for two-class.

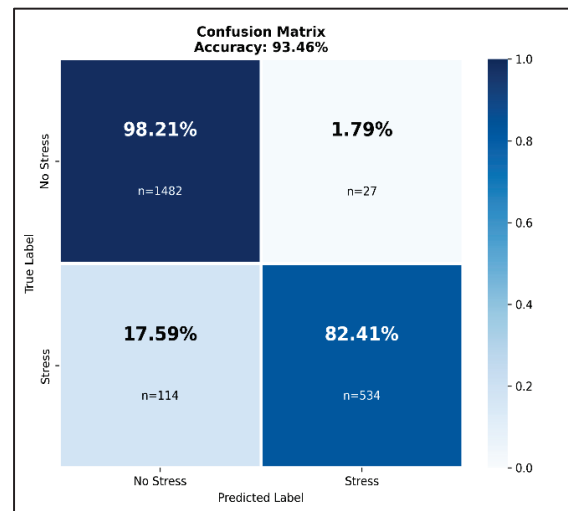


Fig 9: Confusion Matrix for two-class classification.

In addition, the model is evaluated using the original WESAD three-class configuration (baseline, amusement and stress). Under this standard setting the model achieves an accuracy of 82%. The classification report and confusion matrix for this setting are shown in Fig. 10 and Fig.11. The results indicate consistent performance across all classes demonstrating the ability of the model to generalize under conventional label definitions.

CLASSIFICATION REPORT				
	Precision	Recall	F1-score	Support
Baseline	0.8288	0.9296	0.8763	1151
Amusement	0.6059	0.4553	0.5199	358
Stress	0.9213	0.8488	0.8835	648
accuracy			0.8266	2157
macro avg	0.7853	0.7446	0.7599	2157
weighted avg	0.8196	0.8266	0.8193	2157

Fig 10: Classification report for three-class (baseline, amusement, stress)

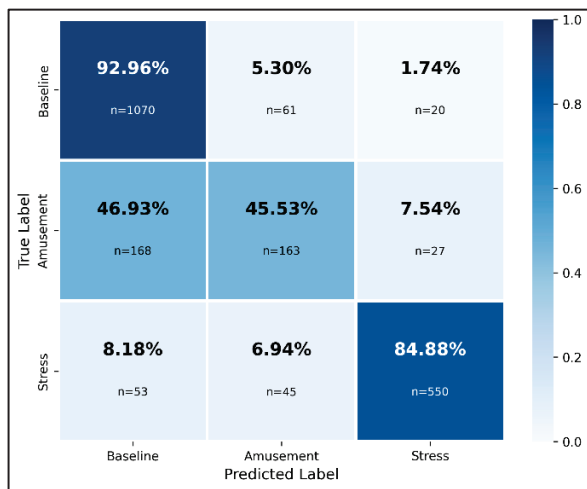


Fig 11: Confusion matrix for three-class (baseline, amusement, stress)

It is observed that the binary classification setting yields higher performance due to reduced class complexity whereas the three-class setting presents a more challenging scenario due to the presence of intermediate or overlapping states. Nevertheless, the model maintains stable and reliable performance across both settings. Overall, these additional evaluations confirm that the proposed hybrid CNN-RF model is robust across multiple classification scenarios and is capable of adapting to both simplified and standard stress detection formulations.

#### D. Application Results

The proposed system also demonstrates an application interface for visualizing predicted stress levels (Low, Moderate and High) along with basic user feedback including relaxation guidance and real-time stress alerts. The application outputs for different stress levels are illustrated in Fig. 12–14.

Fig. 12 shows the interface for Low stress, providing positive reinforcement messages. Fig. 13 presents the Moderate stress condition, where the system suggests basic relaxation activities such as hydration and breathing exercises. Fig. 14 illustrates the High stress condition where stronger intervention recommendations and alerts are provided. The outputs demonstrate consistent and reliable prediction behavior across varying conditions. The application enables intuitive visualization of stress levels and enhances user engagement through actionable feedback facilitating timely stress management interventions.

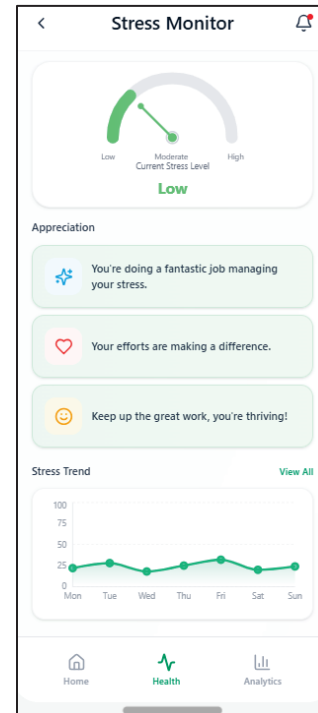


Fig 12: Application interface displaying Low stress level with feedback

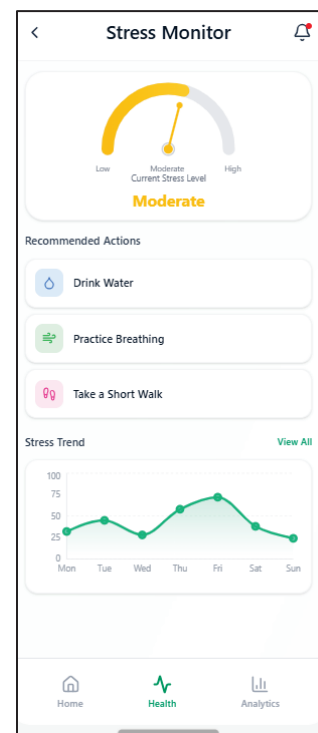
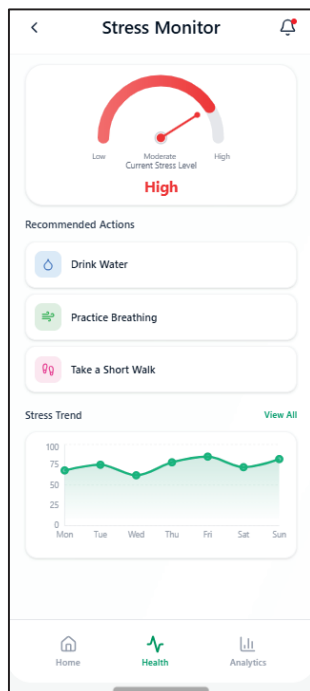


Fig 13: Application interface displaying Moderate stress level with recommended actions



**Fig 14: Application interface displaying High stress level with stress alerts and guidance**

Overall, the system highlights the integration of wearable sensing, machine learning and user interaction into a unified framework for practical and real-time stress monitoring applications.

### E. Comparison with Existing Work

Table 3 presents a comparison of the proposed model with existing studies on the WESAD dataset. For fair comparison only studies utilizing wrist-based physiological signals specifically Blood Volume Pulse (BVP), Electrodermal Activity (EDA) and Skin Temperature (TEMP) are considered. All reported results are based on Leave-One-Subject-Out (LOSO) cross-validation to ensure subject-independent evaluation. The table summarizes key aspects including the classification setting (binary or three-class) the learning approach and the corresponding performance metrics such as accuracy, F1-score and balanced accuracy. For binary classification, the proposed model achieves an accuracy of 93.46% an F1-score of 91.90% and a balanced accuracy of 90.63% outperforming prior wrist-based approaches such as Schmidt et al. and Siirtola & Roning. For the three-class setting two evaluation configurations are considered. Under the original WESAD label configuration (baseline, amusement and stress) the proposed model achieves an accuracy of 82% enabling direct comparison with existing studies. Under the modified label configuration (Low, Moderate and High) the model achieves an accuracy of 80.25% an F1-score of 80.25% and a balanced accuracy of 79.35%. The slightly lower performance in the modified setting is attributed to the increased complexity introduced by the intermediate Moderate class which exhibits overlapping physiological characteristics.

**Table 3: Model performance comparison for two-class.**

Study	Algorithm	Acc.	F1-Score	Bal. Acc
Schmidt et al. [1]	RF	88.33	86.10	-
Siirtola & Rönning [2]	LDA	-	-	87.40
Ninh et al. [4]	NN	90.00	-	92.66
Huynh et al. (StressNAS) [39]	ML	~89.00	-	-
<b>Proposed (2-class)</b>	<b>CNN+RF</b>	<b>93.46</b>	<b>91.90</b>	<b>90.63</b>

**Table 4: Model performance comparison for three-class.**

Study	Algorithm	Acc.	F1-score	Bal. Acc
Schmidt et al. [1]	AdaBoost	73.62	64.24	-
Huynh et al. (StressNAS) [39]	ML	~79	-	-
<b>Proposed (3-class)</b>	<b>CNN+RF</b>	<b>82.66</b>	<b>73.76</b>	<b>75.08</b>

<b>Proposed (3-class modified)</b>	<b>CNN+RF</b>	<b>80.25</b>	<b>80.25</b>	<b>79.35</b>
------------------------------------	---------------	--------------	--------------	--------------

Despite the increased complexity of the task the proposed model demonstrates strong generalization capability under realistic and practical conditions. The performance remains consistent across both the modified and standard three-class settings confirming its robustness under different label formulations. While some studies report higher accuracy they often address simpler classification tasks or rely on less practical sensor configurations. The proposed approach offers a balanced trade-off between accuracy, robustness and real-world usage making it well-suited for wearable stress monitoring systems.

### V. CONCLUSION

The proposed work presents a hybrid Convolutional Neural Network–Random Forest based stress detection system using wrist-based physiological signals including Blood Volume Pulse (BVP), Electrodermal Activity (EDA) and Skin Temperature (TEMP). The approach combines deep learning-based feature extraction with handcrafted physiological features to effectively capture both temporal patterns and domain-specific characteristics. A three-class stress classification framework is introduced by redefining WESAD labels into Low, Moderate and High stress levels providing a more practical yet

challenging formulation compared to conventional approaches. The model is evaluated using Leave-One-Subject-Out (LOSO) cross-validation demonstrating strong generalization capability across unseen subjects. Additional evaluations under binary and standard three-class settings further confirm the robustness and adaptability of the proposed model. The integration of a lightweight CNN architecture with a Random Forest classifier ensures low computational complexity making the system suitable for real-time wearable applications. Furthermore, the system demonstrates application-level integration for visualizing stress levels and providing basic user feedback highlighting its practical usability. Future work can focus on improving the classification of intermediate stress levels by incorporating additional contextual or behavioural information as well as enhancing personalization to better adapt to individual physiological variations. Further optimization and deployment of the model on embedded platforms can enable more robust and scalable real-time stress monitoring systems. Overall, the proposed system offers an efficient, scalable and practical solution for wearable stress detection.

### ACKNOWLEDGEMENT

We extend our sincere gratitude to SNDT Women's University for providing the platform and resources necessary to undertake this research. Our heartfelt thanks go to the Usha Mittal Institute of Technology and its dedicated faculty for their consistent support and encouragement throughout our academic journey. We sincerely thank our respected Principal, Dr. Yogesh Nerkar for his visionary leadership and constant encouragement throughout our research journey.

We are also immensely grateful to our research guide, Dr. Seema M. Hanchate and co-guide, Ms. Poonam More for their dedicated mentorship, thoughtful guidance and unwavering support which contributed to the success of this work. Their expertise, encouragement and commitment played a vital role in shaping and enriching the quality of this work.

### REFERENCES

- [1] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, 2005.
- [2] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review," *Psychiatry Investigation*, vol. 15, no. 3, pp. 235–245, 2018.
- [3] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biol. Psychol.*, vol. 84, no. 3, pp. 394–421, 2010.
- [4] S. Gedam and S. Paul, "A review on mental stress detection using wearable sensors and machine learning techniques," *IEEE Access*, vol. 9, pp. 84045–84066, 2021.
- [5] G. Giannakakis et al., "Review on psychological stress detection using biosignals," *IEEE Trans. Affective Computing*, vol. 13, no. 1, pp. 440–460, 2022.
- [6] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring," *IEEE Trans. Syst., Man, Cybern., Part C*, vol. 40, no. 1, pp. 1–12, 2010.
- [7] A. O. Akmandor and N. K. Jha, "Smart health monitoring using wearable devices," *IEEE Trans. Biomed. Circuits Syst.*, 2017.
- [8] Y. Zhu et al., "Stress detection using wrist-based electrodermal activity signals," *Sensors*, vol. 19, no. 2, 2019.
- [9] M. Gjoreski et al., "Continuous stress detection using a wrist device," *Proc. ACM IMWUT*, vol. 1, no. 4, 2017.
- [10] M. Iqbal et al., "A review of stress detection using wearable physiological sensors," *Sensors*, vol. 22, no. 19, 2022.
- [11] P. Schmidt et al., "Introducing WESAD: A multimodal dataset for wearable stress and affect detection," *Proc. ACM ICMI*, pp. 400–408, 2018.
- [12] P. Bobade and S. Vani, "Stress detection with machine learning and deep learning using multimodal physiological data," *Procedia Computer Science*, vol. 167, pp. 211–218, 2020.
- [13] V.-T. Ninh et al., "An improved subject-independent stress detection model using wearable sensors," *IEEE Access*, vol. 8, pp. 114425–114435, 2020.
- [14] J. Siirtola, "Continuous stress detection using wearable sensors," *IEEE Trans. Affective Computing*, vol. 10, no. 2, pp. 273–283, 2019.
- [15] M. Andrić et al., "Random Forest-based stress detection using physiological signals," *Sensors*, vol. 18, no. 5, 2018.
- [16] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [17] M. Rashid et al., "SELF-CARE: Sensor-based stress detection using multimodal data," *IEEE Access*, vol. 9, 2021.
- [18] R. Ghosh et al., "Image encoding-based deep neural network for stress detection," *IEEE Access*, vol. 7, pp. 156322–156334, 2019.
- [19] G. I. Winata et al., "Deep convolutional neural networks for physiological signal analysis," *IEEE Access*, 2018.
- [20] A. Eren and Z. Saraç, "Stress detection using BVP and EDA signals," *Biomedical Signal Processing*, 2020.
- [21] M. Gjoreski et al., "Wrist-based stress detection using physiological signals," *Sensors*, 2017.
- [22] D. S. Benita et al., "Hybrid CNN and Random Forest model for physiological signal classification," *Proc. Int. Conf. Emerging Systems*, 2021.
- [23] S. Abdelfattah et al., "Hybrid machine learning models for stress detection," *IEEE Access*, vol. 9, 2021.
- [24] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in Public Health*, vol. 5, 2017.
- [25] R. Castaldo et al., "Short-term heart rate variability for stress detection," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 6, pp. 2340–2348, 2019.
- [26] A. Greco et al., "cvxEDA: Convex optimization approach for electrodermal activity processing," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 797–804, 2016.
- [27] K. He et al., "Deep residual learning for image recognition," *Proc. IEEE CVPR*, pp. 770–778, 2016.

- [28] D. Makowski et al., "NeuroKit2: A Python toolbox for neurophysiological signal processing," Behavior Research Methods, 2021.
- [29] D. Tyuleberdinova et al., "ESP32-based stress monitoring system," IEEE Conf. IoT Systems, 2021.
- [30] M. Mouadili et al., "IoT-based CNN models for stress detection," Sensors, 2025.
- [31] Y. Zhu et al., "Wearable IoT-based stress monitoring system," IEEE Internet Things J., 2020.
- [32] S. Phadke et al., "IoT-based wearable stress detection system," IEEE Conf. Healthcare Informatics, 2021.
- [33] R. Rizwan et al., "Machine learning based stress detection using wearable sensors," IEEE Conf. Computing Applications, 2020.
- [34] M. Strinar et al., "Impact of window size on stress detection," IEEE Access, 2020.
- [35] K. Plarre et al., "Continuous stress detection in real-life settings," IEEE Trans. Biomed. Eng., 2011.
- [36] R. K. Nath and H. Thapliyal, "Machine learning based anxiety detection using wearable sensors," IEEE Sensors J., 2021.
- [37] S. Koldijk et al., "The SWELL knowledge work dataset for stress detection," IEEE Trans. Affective Computing, 2014.
- [38] X. Wu et al., "Multimodal transformer-based stress detection," IEEE Trans. Affective Computing, 2023.
- [39] L. Huynh, T. Nguyen, T. Nguyen, S. Pirttikangas, and P. Siirtola, "StressNAS: Affect State and Stress Detection Using Neural Architecture Search," arXiv preprint arXiv:2108.12502, 2021.