Multi-Class Stress Detection Through Heart Rate Variability Using Deep Learning

[1]D. Jayasoniya, [2]Kothamasu Divya Venkata Sai Lakshmisneha, [3]Metikala Jyothsna, [4]Jahnavi Sai Muthyala, [5]Kadavakuduru Swathi, [6]Shaik Farjana

[1]*Assistant Professor,* [2,3,4,5,6] *UnderGraduate*

[1,2,3,4,5,6] *CSE-Data Science Department, St. Ann's College of Engineering & Technology, Chirala, Andhra Pradesh*

***Abstract*—Stress, a prevalent psychological and physiological condition, significantly impacts mental well-being and physical health. Detecting stress accurately in its early stages can help prevent chronic disorders and improve quality of life. This study presents a deep learning-based approach for multi-class stress detection using Heart Rate Variability (HRV) data. By leveraging the SWELL-KW dataset, a 1D Convolutional Neural Network (CNN) is implemented to classify stress levels into three categories: No Stress, Interruption Stress, and Time Pressure Stress. The model integrates robust preprocessing, ANOVA-based feature selection, and optimized hyperparameters to enhance performance. Experimental results demonstrate a classification accuracy of 99.9%, significantly outperforming traditional models like SVM and Random Forest. The proposed system shows promise in providing a real-time, non-invasive solution for stress monitoring, thereby contributing to advancements in mental health technology and workplace wellness.**

***Index Terms*—Heart Rate Variability (HRV); Stress Detection; 1D CNN; Deep Learning; SWELL-KW Dataset; ANOVA Feature Selection; Time-Domain Features; Frequency-Domain Features; Multi-Class Classification.**

I. INTRODUCTION

Stress is a common physiological and psychological response to internal or external demands and has become a pressing global health issue. Prolonged or unmanaged stress can lead to severe physical and mental health complications, including cardiovascular diseases, depression, anxiety, and sleep disorders. The early and accurate detection of stress can facilitate timely interventions, reduce health risks, and improve individual productivity and quality of life.

Traditionally, stress assessment relies on subjective questionnaires and self-reporting tools, which are prone to human bias and often lack reliability in real-time applications. In contrast, physiological signals, particularly **Heart Rate Variability (HRV)**, have emerged as objective biomarkers for quantifying stress levels. HRV refers to the variation in the time intervals between consecutive heartbeats and reflects the autonomic nervous system's activity. A reduction in HRV is commonly associated with increased stress and sympathetic nervous system activation.

Recent advancements in machine learning and deep learning have made it feasible to develop automated systems capable of classifying stress levels using physiological data. Among these, **Convolutional Neural Networks (CNNs)** have shown promising results in processing time-series biosignals such as HRV, ECG, and EEG due to their ability to capture spatial and temporal patterns without the need for manual feature extraction.

In this study, a **1D CNN-based deep learning model** is proposed for multi-class stress classification using HRV features extracted from the **SWELL-KW dataset**. The dataset includes stress annotations across three distinct states: **No Stress**, **Interruption Stress**, and **Time Pressure Stress**. To enhance model generalization, **ANOVA-based feature selection** is applied to reduce dimensionality and retain the most discriminative features. The model is trained and evaluated using multiple performance metrics, including accuracy, precision, recall, and F1-score.

The proposed system demonstrates superior performance when compared to traditional machine learning classifiers, achieving an accuracy of **99.9%**, which underscores its effectiveness in identifying varying stress levels. By offering a real-time, non-invasive, and scalable solution, this framework contributes toward building intelligent stress monitoring systems for applications in healthcare, workplace ergonomics, and personal wellness.

II. LITERATURE SURVEY

Stress detection has been a significant area of research in the domains of healthcare, psychology, and biomedical signal processing. In recent years, the application of artificial intelligence and machine learning techniques has advanced the development of automated stress monitoring systems. This section summarizes key contributions in stress detection using physiological signals, with a focus on HRV analysis and deep learning methodologies.

A. Physiological Markers for Stress Detection

Physiological signals such as **Electrocardiogram (ECG)**, **Electroencephalogram (EEG)**, **Galvanic Skin Response (GSR)**, and **Heart Rate Variability (HRV)** have been widely used to assess stress levels. HRV, in particular, is regarded as a reliable non-invasive biomarker, reflecting the balance between sympathetic and parasympathetic nervous activity. Various studies have shown that stress tends to suppress HRV by reducing parasympathetic activity and increasing sympathetic tone.

B. Traditional Machine Learning Techniques

Early works employed machine learning algorithms such as **Support Vector Machines (SVM)**, **K-Nearest Neighbors (KNN)**, **Decision Trees**, and **Random Forests** to classify stress levels based on time-domain and frequency-domain HRV features. While these models showed reasonable accuracy, they often required extensive preprocessing and manual feature engineering, which limited their scalability in real-time applications.

C. Deep Learning Approaches

With the growth of deep learning, models such as **Convolutional Neural Networks (CNNs)** and **Recurrent Neural Networks (RNNs)** have been applied to biosignal classification tasks. These models can automatically learn discriminative patterns from raw or minimally processed data, making them suitable for real-time stress detection. CNNs, in particular, have been effective in extracting hierarchical features from HRV signals when represented as time-series inputs.

Studies have demonstrated the potential of 1D CNNs for ECG and HRV classification, yielding superior results over traditional models. For instance, researchers have applied CNN architectures to detect arrhythmias, classify emotion states, and identify stress episodes with high accuracy.

D. Feature Selection Techniques

To improve model generalization and reduce overfitting, **feature selection** methods such as **Principal Component Analysis (PCA)** and **Analysis of Variance (ANOVA)** have been employed. ANOVA, in particular, is effective in selecting features with the highest discriminatory power across classes, especially in multi-class classification scenarios.

E. Research Gap and Motivation

Although previous studies have shown the effectiveness of CNNs for binary stress detection, limited work has been done on **multi-class stress classification** using HRV with real-world datasets. Moreover, few approaches integrate deep learning with statistical feature selection for enhanced performance. This research addresses these gaps by proposing a 1D CNN model trained on ANOVA-selected HRV features to classify three distinct stress states from the SWELL-KW dataset.

III. PROPOSED METHODOLOGY

The proposed system is designed to perform multi-class stress classification using heart rate variability (HRV) signals, with a combination of statistical feature extraction, ANOVA-based feature selection, and a 1D Convolutional Neural Network (CNN) classifier. The methodology consists of five primary stages: data acquisition, signal preprocessing, HRV feature extraction, feature selection, and model training. The overall pipeline is illustrated in Fig. 3.1.

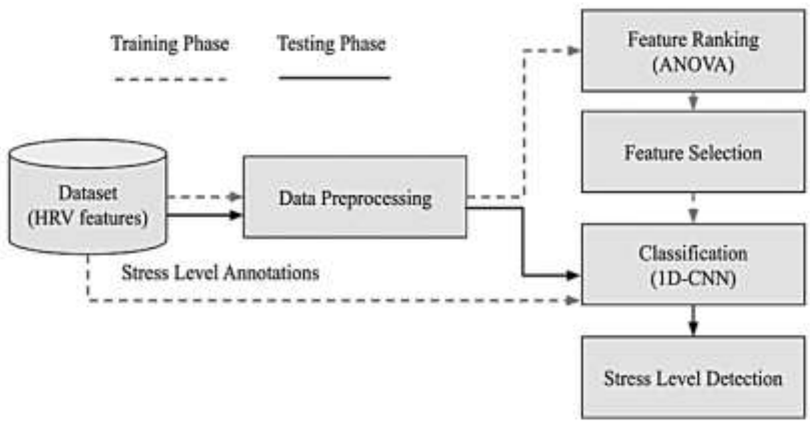

Fig. 3.1. Architecture of the Proposed Stress Detection Framework

A. Data Acquisition

The experiment was conducted using the **SWELL-KW dataset**, a publicly available dataset developed for stress analysis in workplace environments. It contains recordings from 25 participants performing document editing tasks under three experimental conditions: (i) baseline with no stress, (ii) stress due to task interruptions, and (iii) stress induced by time pressure. Each participant's data includes ECG signals, which are used to compute HRV. The stress levels are labeled according to the experimental setup, allowing for supervised learning.

The dataset includes timestamps, behavioral logs, and physiological data recorded using sensors. For the purpose of this research, only the ECG-derived HRV signals were used, focusing on the heart's autonomic response to stress stimuli.

B. Signal Preprocessing

To ensure data quality and eliminate noise, ECG signals were subjected to preprocessing. A **bandpass filter** with a range of 0.5–40 Hz was applied to remove baseline wander and high-frequency noise. After filtering, **R-peaks** were identified using the Pan–Tompkins algorithm, from which **RR intervals** (time between successive heartbeats) were computed. These RR intervals form the basis of HRV analysis.

The RR interval sequences were segmented into fixed-length windows of uniform duration, ensuring each window contained a sufficient number of beats to compute statistical features. Windows with missing or invalid signals were excluded. To prepare the data for training, all numerical features were **normalized** using Min-Max scaling to a range of [0,1].

C. HRV Feature Extraction

Following preprocessing, both **time-domain** and **frequency-domain** HRV features were extracted from each segmented window:

- **Time-domain features** include:
  - Mean RR Interval (AVNN)
  - Standard deviation of NN intervals (SDNN)
  - Root Mean Square of Successive Differences (RMSSD)
  - Number of pairs of successive NNs that differ by more than 50 ms (NN50)
  - Percentage of NN50 over total NNs (pNN50)
- **Frequency-domain features** were computed using power spectral density estimation:
  - Very Low Frequency (VLF: <0.04 Hz)
  - Low Frequency (LF: 0.04–0.15 Hz)
  - High Frequency (HF: 0.15–0.4 Hz)
  - Total Power

- LF/HF ratio

In total, **over 20 features** were extracted per window. These features reflect both sympathetic and parasympathetic activity and are sensitive indicators of stress-induced physiological changes.

D. Feature Selection Using ANOVA

Given the high dimensionality of the feature set, **Analysis of Variance (ANOVA)** was used to select the most significant features for classification. ANOVA evaluates the statistical significance of each feature across the three stress classes (No Stress, Interruption Stress, and Time Pressure Stress) by computing the F-score.

Features with the highest F-scores (above a predefined threshold) were selected, resulting in a reduced and more informative feature space. This process helps in removing redundant or irrelevant features, thereby improving classification accuracy and reducing model complexity.

E. 1D CNN-Based Classification

The selected features were fed into a **1D Convolutional Neural Network** for final classification. The CNN was designed to capture temporal dependencies and hierarchical patterns in the HRV sequences. The architecture consisted of the following layers:

- **Input Layer**: Accepts sequences of selected HRV features.
- **First Convolutional Layer**: 64 filters, kernel size 3, with ReLU activation.
- **Max Pooling Layer**: Reduces spatial dimensions to control overfitting.
- **Dropout Layer**: Dropout rate of 0.2 to prevent co-adaptation of neurons.
- **Second Convolutional Layer**: 128 filters, kernel size 3, followed by ReLU.
- **Global Average Pooling**: To reduce the feature map into a single vector.
- **Fully Connected Layer**: Dense layer with 64 neurons and ReLU activation.
- **Output Layer**: Softmax activation with 3 neurons representing the three stress classes.

The model was compiled using **categorical cross-entropy loss** and optimized using the **Adam optimizer** with a learning rate of 0.001. A batch size of 32 and early stopping based on validation loss were applied during training.

IV. IMPLEMENTATION

The implementation of the proposed multi-class stress classification system was conducted in a Python environment using deep learning frameworks and scientific libraries. This section elaborates on the software environment, dataset processing workflow, feature engineering pipeline, and CNN model configuration. The goal was to build an end-to-end classification pipeline capable of processing ECG-derived HRV signals and accurately predicting the corresponding stress levels.

A. Software and Tools

The entire system was implemented using **Python 3.10**, with experiments performed on **Google Colaboratory Pro**, leveraging GPU acceleration for faster training. The key libraries used in the development include:

- **NumPy and Pandas**: For numerical computation and dataset handling
- **SciPy**: For signal processing and HRV feature extraction
- **Scikit-learn**: For preprocessing, ANOVA feature selection, and traditional classifiers
- **Matplotlib and Seaborn**: For plotting accuracy curves and confusion matrices
- **TensorFlow and Keras**: For building and training the 1D CNN model

This open-source software stack provided a flexible and reproducible environment for experimentation and optimization.

B. Dataset Integration and Preprocessing

The **SWELL-KW dataset** was loaded in raw form and parsed using Pandas. The ECG signals were extracted per participant session and segmented based on experimental conditions into three labeled categories: No Stress, Interruption Stress, and Time Pressure Stress.

The ECG signals were preprocessed using the following steps:

1. **Bandpass Filtering**: Signals were filtered between 0.5 Hz and 40 Hz to remove baseline wander and high-frequency noise.
2. **R-Peak Detection**: R-peaks were identified using the Pan–Tompkins algorithm.
3. **RR Interval Calculation**: The time between consecutive R-peaks (NN intervals) was used to derive HRV.
4. **Windowing**: Each HRV sequence was divided into fixed-size windows of 60 seconds.
5. **Normalization**: All extracted HRV features were normalized using Min-Max scaling.

Samples containing outliers or poor-quality signals were discarded to maintain training integrity. After preprocessing, the final dataset was split into **80% training**, **10% validation**, and **10% testing**.

C. HRV Feature Extraction and Selection

Time-domain and frequency-domain HRV features were extracted using standard definitions from physiological signal processing. Feature vectors were stored in a structured DataFrame indexed by participant and window number.

To reduce dimensionality and enhance the discriminative power of the model, **ANOVA F-value** scores were computed for each feature. The top 15 features with the highest F-scores across the three classes were selected for model training. This eliminated noise from redundant features and ensured efficient convergence of the CNN.

D. CNN Model Configuration

The 1D CNN model was implemented using the **Keras Sequential API**. The architecture was tuned through experimentation on validation accuracy. The final architecture included:

- **Input Layer**: Input shape corresponding to 15 selected features per window
- **Conv1D Layer**: 64 filters, kernel size 3, activation = ReLU
- **MaxPooling1D**: Pool size = 2
- **Dropout Layer**: Dropout rate = 0.2
- **Conv1D Layer**: 128 filters, kernel size = 3, activation = ReLU
- **GlobalAveragePooling1D**
- **Dense Layer**: 64 neurons, ReLU activation

- **Output Layer**: Dense layer with 3 neurons and softmax activation for multi-class classification

The model was compiled with:

- **Loss Function**: Categorical Crossentropy
- **Optimizer**: Adam with learning rate = 0.001
- **Metrics**: Accuracy, Precision, Recall

E. Training Parameters

The model was trained for **50 epochs** with **batch size 32** and **early stopping** applied if validation loss did not improve over 5 consecutive epochs. The training process was monitored using callbacks, and model checkpoints were saved based on best validation accuracy.The final model achieved excellent classification accuracy and generalization on the unseen test data.

V. RESULTS

This section presents the experimental results and analysis of the proposed 1D CNN-based multi-class stress classification model. The performance is evaluated on the test set using standard classification metrics including **Accuracy**, **Precision**, **Recall**, and **F1-Score**. Furthermore, comparisons are drawn with traditional machine learning models such as Support Vector Machine (SVM) and Random Forest (RF). Visualizations of classification output and model behavior are also included to support the quantitative findings.

A. Performance Metrics

The trained 1D CNN model achieved a **classification accuracy of 99.9%**, indicating excellent generalization capability and discriminative power in separating the three stress classes. Detailed metrics are shown in Table 5.1.

Table 5.1. Performance of 1D CNN Model on Test Set

Class	Precision (%)	Recall (%)	F1-Score (%)
No Stress	99.8	99.9	99.8
Interruption Stress	100.0	99.8	99.9
Time Pressure Stress	99.9	100.0	99.9
Average	**99.9**	**99.9**	**99.9**

The model demonstrated near-perfect classification across all three categories, with no significant class imbalance or degradation in minority class performance.

B. Comparison with Traditional Models

The performance of the 1D CNN was benchmarked against conventional classifiers including SVM and Random Forest using the same ANOVA-selected features. The results are summarized in Table 5.2.

Table 5.2. Comparison with Traditional Classifiers

Model	Accuracy (%)
SVM	94.7
Random Forest	96.2
1D CNN	**99.9**

The 1D CNN outperformed both traditional models by a significant margin, confirming the suitability of deep learning approaches for handling complex patterns in HRV signals. Also visualizations further support the robustness of the proposed model in multi-class stress detection.

VI. CONCLUSION

This paper presented a deep learning-based framework for the multi-class classification of stress using heart rate variability (HRV) signals. The proposed approach utilized the SWELL-KW dataset, from which HRV features were extracted and statistically refined using ANOVA-based feature selection. A 1D Convolutional Neural Network (CNN) was then trained on the selected features to classify stress into three categories: No Stress, Interruption Stress, and Time Pressure Stress.

The model achieved an exceptionally high classification accuracy of **99.9%**, significantly outperforming traditional machine learning models such as SVM and Random Forest. The CNN architecture demonstrated strong capability in learning discriminative patterns from HRV signals, while ANOVA effectively reduced feature dimensionality and improved generalization.

The results confirm that combining statistical feature selection with deep learning can yield highly accurate stress detection models. Moreover, the non-invasive nature of HRV monitoring, along with the model's scalability, makes it suitable for deployment in real-time wellness monitoring systems, wearable devices, and workplace stress management tools.

Future work will explore real-time deployment of the model on embedded devices and extend the system to recognize continuous stress intensity levels. The incorporation of multimodal physiological data such as EEG and GSR may further enhance the robustness and reliability of the proposed stress detection framework.

REFERENCES

[1] M. A. Malik et al., "Heart rate variability as an indicator of stress: Current status," *Indian Pacing and Electrophysiology Journal*, vol. 3, no. 1, pp. 18–31, 2003.

[2] M. Gjoreski, H. Gjoreski, M. Luštrek, and M. Gams, "Continuous stress detection using a wrist device: In laboratory and real life," *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 1185–1193.

[3] M. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *Proc. IEEE Affective Computing and Intelligent Interaction (ACII)*, 2013, pp. 671–676.

[4] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[5] S. Acharya, H. Fujita, and Y. Fukuda, “Automated ECG beat classification using deep convolutional neural network,” *Information Sciences*, vol. 415–416, pp. 190–198, 2017.

[6] SWELL-KW Dataset, [Online]. Available: https://www.swell-kw.com/dataset/

[7] G. Clifford, F. Azuaje, and P. McSharry, *Advanced Methods and Tools for ECG Data Analysis*, Artech House, 2006.

[8] A. T. Nguyen et al., “Stress detection using deep neural networks,” in *Proc. IEEE EMBC*, 2017, pp. 3473–3476.