Milestone Transactions on
MEDICAL TECHNOMETRICS
Vol. 03 | Issue. 02 | 2025

RESEARCH ARTICLE OPEN ACCESS

Stress Monitoring With Heart Rate Variability Using Deep Learning

B K Preethika Raj . D Nithish . B Manjunatha . G Praveen . G Sandeep . P Chandra Sekhar

Department of Computer Science Engineering,
Annamacharya Institute of Technology and Sciences (Autonomous)
Kadapa, Andra Pradesh, India.

DOI: **10.5281/zenodo.15266886**
Received: 22 February 2025 / Revised: 27 March 2025 / Accepted: 30 March 2025
©Milestone Research Publications, Part of CLOCKSS archiving

Abstract – Prolonged stress can lead to mental health issues like anxiety and sleep disorders. Heart Rate Variability (HRV) serves as a key physiological marker for stress detection. Unlike heart rate, HRV measures the variation in time intervals between heartbeats (RR intervals). This study proposes a CNN-based model for classifying stress into no stress, interruption stress, and time pressure stress using HRV features. Evaluated on the SWELL-KW dataset, the model achieves 99.9% accuracy, outperforming existing methods. Feature extraction techniques, such as ANOVA, further validate the significance of HRV features in stress detection.

Index Terms – Stress detection, heart rate variability, convolutional neural network, feature extraction.

I. INTRODUCTION

Stress is a response to external stimuli that can disrupt mental and physical well-being. Chronic stress over activates the sympathetic nervous system (SNS), leading to various health issues. Traditional stress assessment relies on subjective evaluations, whereas physiological measures like Heart Rate Variability (HRV) offer objective insights into stress levels. HRV reflects variations in time intervals between heartbeats (RR intervals). Research shows that HRV increases during relaxation and decreases under stress. As the Autonomic Nervous System (ANS) regulates stress responses, physiological markers such as electrocardiograms (ECG), heart rate, blood pressure, and respiration rate help assess mental stress levels. ECG-based HRV extraction requires clinical expertise, but advancements in the Internet of Medical Things (IoMT) have enabled wearable devices for stress monitoring. Various machine learning (ML) and

MILESTONE
RESEARCH.IN
OPEN ACCESS
Electronic ISSN: **2584-072X**

deep learning (DL) algorithms have been explored for stress detection, with the SWELL-KW dataset being a widely used benchmark. However, existing models struggle with ultra-high accuracy in multi-class stress classification.

II. RELATED WORK

Existing research on HRV data quality has analyzed ECG and IoMT devices like Elite HRV, H7, Polar, and Motorola Droid. Studies indicate minor discrepancies between HRV values from IoMT devices and traditional ECG measurements. Despite small errors, IoMT devices remain a cost-effective alternative for stress monitoring. Machine learning and deep learning techniques have been widely used for stress classification. Traditional methods such as Naive Bayes, KNN, SVM, MLP, Random Forest, and Gradient Boosting have achieved up to 80% recall in binary classification. SVM with radial basis function (RBF) reported 83.33% accuracy using HRV features. Deep learning approaches, including CNNs, have demonstrated better accuracy, with one model achieving 98.4% accuracy. Stress Click, a study using Random Forest, classified stress based on mouse-click patterns.

For multi-class stress classification, the SWELL-KW dataset has been used to categorize stress levels as no stress, interruption stress, and time pressure stress. An SVM-based model reached 90% accuracy. The WESAD dataset has also been used for binary and three-class classification, with machine learning models achieving 81.65% accuracy and deep learning models reaching 84.32%. Recent studies have explored advanced deep learning models like Genetic Deep Learning CNNs (GDCNNs), which are effective for 2D data classification but require significant modifications for 1D stress data. A 2022 study on the SWELL-KW dataset achieved 88.64% accuracy, 93.01% precision, and an F1 score of 82.75%. The proposed 1D CNN model in this study outperforms these approaches, achieving 99.9% accuracy in multi-class stress classification.

III. FRAMEWORK OVERVIEW AND DATA PREPROCESSING

Framework Overview

This section outlines the framework for multi-class stress classification. It includes data collection, dataset preparation, and preprocessing, while the CNN model is discussed in the next section.

Data Collection and Dataset

The study uses the SWELL-KW dataset, which includes HRV signals, computer activity logs, facial expressions, body postures, and physiological responses. The dataset was collected from 25 participants engaged in knowledge-based tasks under different stress conditions: normal, time pressure, and interruptions. Participants' stress levels were labeled by medical professionals.

IV. A CNN MODEL FOR STRESS STATUS CLASSIFIVATION

we present the developed deep learning model for stress status classification, the model consists of feature ranking, feature extraction, and tress level classification

Feature Ranking and Extraction

Feature ranking is performed using the ANOVA F-test, which identifies the most relevant features for classification. This statistical method evaluates whether different sample groups share the same distribution. Initially, all features are used, and less significant ones are removed to optimize training time while maintaining accuracy.

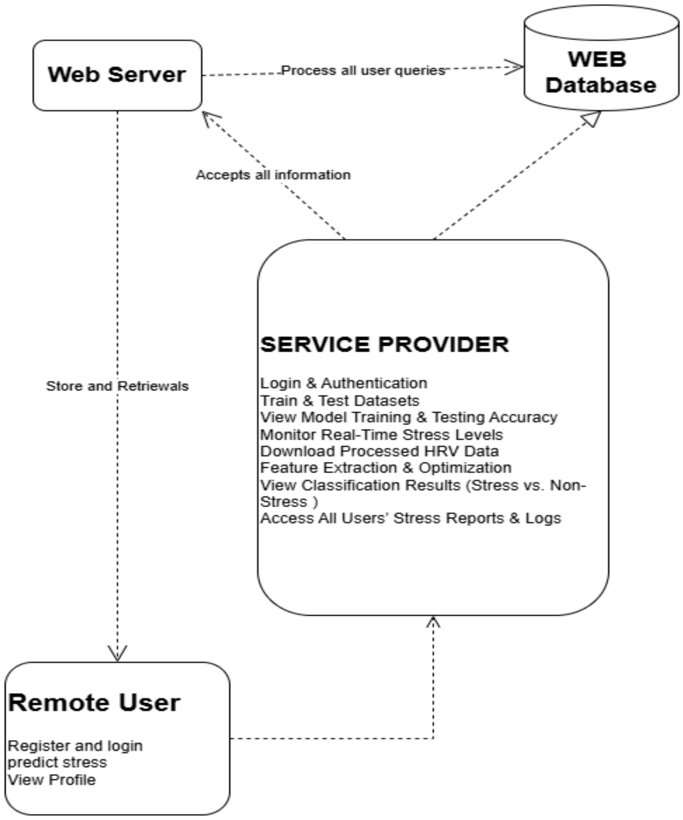

Fig 1: System Architecture

CNN-Based Deep Learning Model

The model is based on a 1D Convolutional Neural Network (CNN), which is effective for learning from sequential data. It includes an input layer, multiple hidden layers, a max-pooling layer, a flattening layer, and an output layer. The input layer applies convolution using 64 filters, a kernel size of 2, and the ReLU activation function to speed up convergence. The max-pooling layer reduces feature dimensions, while the flattening layer converts data into a 1D vector for classification. A softmax function is used in the output layer to categorize stress levels into three classes: no stress, time pressure, and interruption. The model is trained using categorical cross-entropy loss and optimized with the ADAM optimizer for efficient learning.

Data Preprocessing

HRV data is reformatted from a time-series structure to numerical sequences. Noisy, incomplete, or missing data is removed. The dataset undergoes normality testing and is split into training and testing sets. Features are normalized and reshaped for compatibility with the CNN model. The model is trained using Google Colab with batch processing for efficiency

V. PERFORMANCE METRICS

The performance of the 1D CNN model for multi-class stress classification is assessed using discrimination analysis on the SWELL-KW dataset. The key evaluation metrics include Precision, Recall, Accuracy, F1-score, MCC, a classification report, and a confusion matrix. The confusion matrix is a two-dimensional table comparing actual versus predicted values, categorized into True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). TP occurs when the model correctly identifies a positive class, TN when it accurately detects a negative class, FP when it incorrectly predicts a positive class, and FN when it wrongly classifies a negative class. The proposed block chain-based KYC background consists of three main components: customer on boarding, credit allocation, and risk valuation. The background leverages Ethereum smart contracts to automate these processes and ensure data integrity.

VI. CLASSIFICATION OF RESULTS AND DISCUSSION

In this section, we present the experimental results and reveal the importance of ANOVA-based feature selection.

Feature Ranking and Selection

The SWELL-KW dataset includes 34 features, but some are irrelevant or act as outliers. The ANOVA method ranks these features based on their F-values, with higher values indicating greater importance for stress classification. The forward sequential selection method then identifies the most relevant subset. The model achieves over 95% accuracy with fewer than 17 top-ranked feature

Feature Ranking and Selection

The SWELL-KW dataset includes 34 features, but some are irrelevant or act as outliers. The ANOVA method ranks these features based on their F-values, with higher values indicating greater importance for stress classification. The forward sequential selection method then identifies the most relevant subset. The model achieves over 95% accuracy with fewer than 17 top-ranked features.

Performance with All Features

Using all features, the 1D CNN model classifies stress into three categories: no stress, time pressure, and interruption. The model achieves an accuracy of 99.9%, with Precision, Recall, and F1-score all reaching.The confusion matrix shows a classification error of less than 0.01%. Validation tests confirm that the model is not overfitted, as training and validation accuracy remain nearly identical**.**

Performance with Top 15 Features

The model was further evaluated using only the top 15 ANOVA-ranked features. The results showed that Precision, Recall, F1-score, and MCC remained high, averaging 96.5%, 94.6%, 97.0%, and 92.9%, respectively. Using a 70/30 train-test split, the model achieved an accuracy of 96.1%. While using all features yields better results, it comes at the cost of longer training times. Selecting a smaller feature subset offers a balance between accuracy and computational efficiency.

K-Fold Cross-Validation

To ensure robustness, a 5-fold cross-validation was performed. The average scores across the five splits were Precision = 94.4%, Accuracy = 94.5%, Recall = 93.3%, F1-score = 90.8%, and MCC = 90.8%. These results confirm that the model maintains high classification accuracy across different test sets.

Hyperparameter Optimization

Hyper band tuning was applied to optimize the model's hyper parameters using the top 15 features. The best parameters—filters = 160, kernel size = 5, and dense units = 48—yielded a validation accuracy of 99%. However, hyperparameter tuning can be resource-intensive and dataset-specific, making it less practical for all applications.

Comparison with Existing Studies

A comparison with previous studies shows that no existing model using the SWELL-KW dataset outperforms the proposed 1D CNN model in terms of Accuracy, Precision, Recall, F1-score, and MCC. While one study achieved better performance using all available features, it did not apply feature selection techniques. This highlights the effectiveness of the proposed approach in balancing accuracy and computational efficiency.

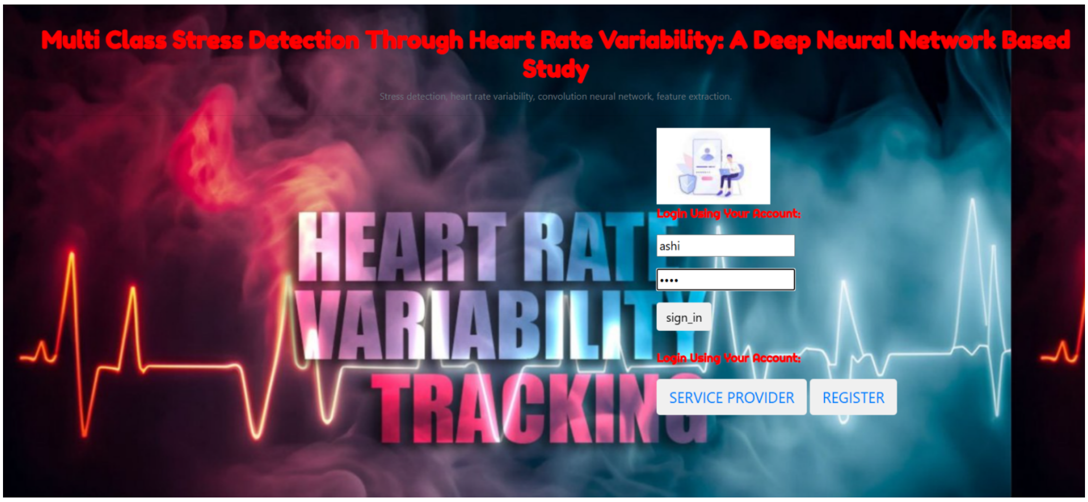

Fig 2: User Interface For Login Page

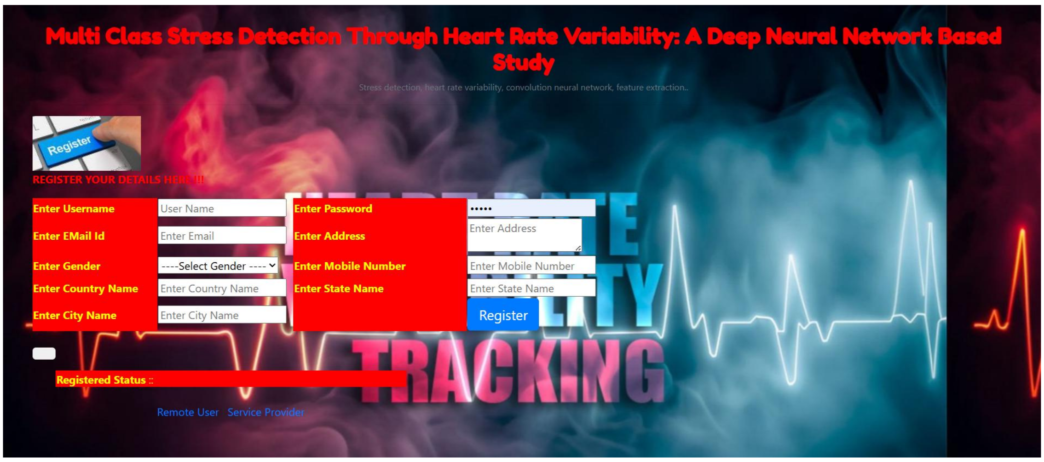

Fig 3: User Interface For Registration page

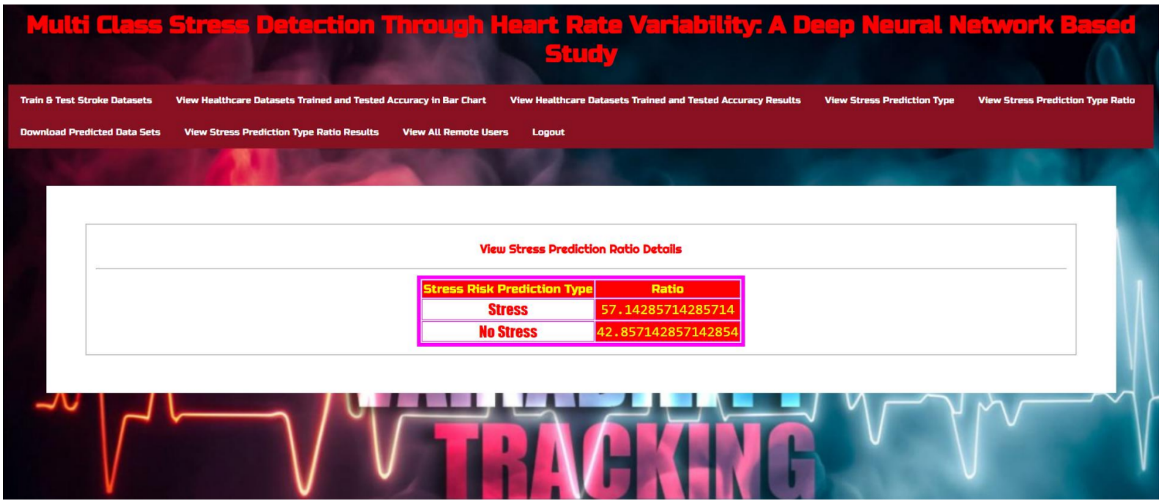

Fig 4: -Stress Prediction Ratio Details

Fig 5: Prediction of Stress

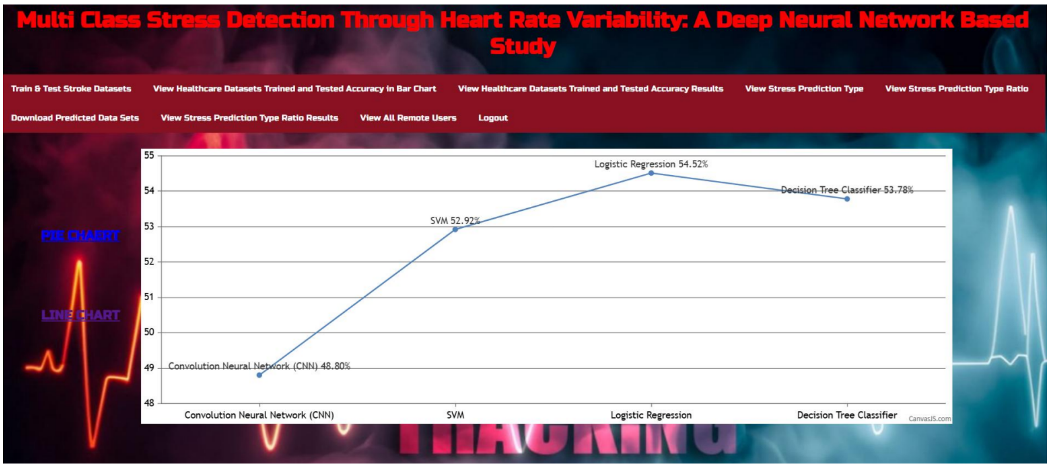

Fig 6: Accuracy Results

VII. FURTHER DISCUSSIONS

The execution time difference between the full-feature and top-15 feature models is minimal due to the dataset's moderate size (410,322 records, 34 features) and the use of Google Colab's high-performance CPUs and GPUs. However, for larger datasets or limited computing resources, feature reduction significantly improves efficiency, especially during validation. Although this model is based on the SWELL-KW dataset, it can be adapted to other mental health datasets with proper tuning. As part of an ongoing study, real-life physiological data is being collected from a Norwegian hospital using non-wearable IoT devices. Future work will evaluate the model's performance on this dataset, though those results are beyond this paper's scope.

VIII. CONCLUSION

This study presents a novel 1D CNN model for stress classification using HRV signals, validated with the SWELL-KW dataset. ANOVA-based feature selection was applied for dimensionality reduction. Extensive training and validation show that the model outperforms existing methods in Accuracy, Precision, Recall, F1-score, and MCC when all features are used. Even with ANOVA-based feature reduction, performance remains high. Future work will focus on optimizing the model for edge devices to enable real-time stress detection.

REFERENCES

1. Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018, March). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation, 15*(3), 235–245.
2. Muhajir, D., Mahananto, F., & Sani, N. A. (2022, January). Stress level measurements using heart rate variability analysis on Android based application. *Procedia Computer Science, 197*, 189–197.
3. Held, J., Vîslă, A., Wolfer, C., Messerli-Bürgy, N., & Flückiger, C. (2021, March). Heart rate variability change during a stressful cognitive task in individuals with anxiety and control participants. *BMC Psychology, 9*(1), Article 44.
4. Dalmeida, K. M., & Masala, G. L. (2021, April). HRV features as viable physiological markers for stress detection using wearable devices. *Sensors, 21*(8), 2873.
5. Miranda-Correa, J. A., Abadi, M. K., Sebe, N., & Patras, I. (2021). AMIGOS: A dataset for affect, personality and mood research on individuals and groups. *IEEE Transactions on Affective Computing, 12*(2), 479–493.
6. Won, E., & Kim, Y.-K. (2016, August). Stress, the autonomic nervous system, and the immune-kynurenine pathway in the etiology of depression. *Current Neuropharmacology, 14*(7), 665–673.
7. Olshansky, B., Sabbah, H. N., Hauptman, P. J., & Colucci, W. S. (2008, August). Parasympathetic nervous system and heart failure: Pathophysiology and potential implications for therapy. *Circulation, 118*(8), 863–871.
8. Goel, S., Tomar, P., & Kaur, G. (2016, March). ECG feature extraction for stress recognition in automobile drivers. *Electronic Journal of Biology, 12*(2), 156–165.
9. Hegde, V. N., Deekshit, R., & Satyanarayana, P. S. (2013, June). A review on ECG signal processing and HRV analysis. *Journal of Medical Imaging and Health Informatics, 3*(2), 270–279.
10. Vollmer, M. (2015, September). A robust, simple and reliable measure of heart rate variability using relative RR intervals. In *Proceedings of Computing in Cardiology Conference (CinC)* (pp. 609–612).
11. Madapuri, R. K., & Mahesh, P. C. S. (2017, December). HBS-CRA: Scaling impact of change request towards fault proneness: Defining a heuristic and biases scale (HBS) of change request artifacts (CRA). *Cluster Computing, 22*(S5), 11591–11599. https://doi.org/10.1007/s10586-017-1424-0
12. Ahmed, S. T., & Syed Mohamed, E. (2020, February). Phonocardiography (PCG) signal optimization and compression for low line transmission in telemedicine. In *International Conference on Automation, Signal Processing, Instrumentation and Control* (pp. 1127-1137). Singapore: Springer Nature Singapore.