International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Multi-Class Stress Detection through Heart Rate Variability: A Deep Neural Network based Study

Dr. M.Manivannan[1], G.Rahul[2], S.Shaziya[3], S.A.Rakshith[4], S.Sana[5], N.Tharun[6]

Department of computer science Siddartha institute of science and technology

ABSTRACT

Stress is a natural human reaction to demands or pressure, usually when perceived as harmful or/and toxic. When stress becomes constantly overwhelmed and prolonged, it increases the risk of mental health and physiological uneasiness. Furthermore, chronic stress raises the likelihood of mental health plagues such as anxiety, depression, and sleep disorder. Although measuring stress using physiological parameters such as heart rate variability (HRV) is a common approach, how to achieve ultra-high accuracy based on HRV measurements remains as a challenging task. HRV is not equivalent to heart rate. While heart rate is the average value of heart beats per minute, HRV represents the variation of the time interval between successive heartbeats. The HRV measurements are related to the variance of RR intervals which stand for the time between successive R peaks. In this study, we investigate the role of HRV features as stress detection bio-markers and develop a machine learning-based model for multi-class stress detection. More specifically, a convolution neural network (CNN) based model is developed to detect multi-class stress, namely, *no stress, interruption stress, and time pressure stress*, based on both time- and frequency-domain features of HRV. Validated through a publicly available dataset, SWELL−KW, the achieved accuracy score of our model has reached 99.9% (*Precision=1*, *Recall=1*, *F1-score=1*, and *MCC=0.99*), thus outperforming the existing methods in the literature. In addition, this study demonstrates the effectiveness of essential HRV features for stress detection using a feature extraction technique, i.e., analysis of variance.

INDEX TERMS: Stress detection, heart rate variability, convolution neural network, feature extraction.

INTRODUCTION

Stress occurs when the body experiences physical or mental imbalances due to harmful stimuli, prompting efforts to restore homeostasis. Chronic stress can lead to an overactive sympathetic nervous system, resulting in physiological, psychological, and behavioral disturbances. Traditionally, stress assessment relies on subjective methods based on personal perceptions. However, heart rate variability (HRV) analysis provides a more objective approach, as it reflects the impact of stress on the autonomic nervous system (ANS). Studies indicate that HRV increases during relaxation and decreases under stress, showing an inverse relationship with heart rate. Additionally, HRV fluctuates based on activity levels and occupational stress, making it a reliable physiological marker for stress detection.

The ANS regulates stress through its sympathetic and parasympathetic divisions. The sympathetic system activates the fight-or-flight response, boosting energy to handle external stressors, while the parasympathetic system promotes relaxation and recovery. Since the ANS influences mental stress levels, physiological indicators such as electrocardiogram (ECG), electromyogram (EMG), galvanic skin response (GSR), HRV, blood pressure, respiration rate, and breathing frequency can be used to assess stress levels. With advancements in the Internet of Medical Things (IoMT), wearable and non-wearable devices now enable real-time heart rate monitoring, making stress assessment more accessible. Given the rise of machine learning (ML) and deep learning (DL), various predictive models have been developed to classify stress levels based on ECG and HRV data. However, existing ML and DL models trained on the SWELL-KW dataset have not yet achieved exceptionally high accuracy in multi-class stress classification, highlighting a gap in research.To address this limitation, we propose a one-dimensional convolutional neural network (1D CNN) for multi-class stress classification, outperforming previous models in predictive accuracy. Our research explores stress detection using both traditional ML approaches and multi-layer perceptron (MLP) architectures inspired by fully connected neural networks (FCNN). Unlike MLP, which processes inputs as vectors, CNNs process inputs as tensors, enabling superior spatial feature representation while reducing training parameters. Our 1D CNN model achieves 99.9% accuracy, with high Precision, Recall, and F1-scores (all at 1.0) and a Matthews Correlation Coefficient (MCC) of 99.9%. Additionally, using an Analysis of Variance (ANOVA) F-test, we optimize feature selection, achieving 96.5% accuracy with fewer than half of the original 34 HRV features. This optimization reduces computational complexity while maintaining excellent classification performance, making our model efficient and scalable.

RELATED WORK

The related work in this study examines HRV data quality and various state-of-the-art ML/DL algorithms developed for stress detection. A comprehensive review of HRV data obtained from ECG and IoMT devices such as Elite HRV, H7, Polar, and Motorola Droid reveals minor discrepancies when compared to ECG-based HRV measurements. A review of 23 studies indicated that these small-scale errors are acceptable, as IoMT devices provide a practical, cost-effective alternative to laboratory equipment for HRV data collection. These findings suggest that portable IoMT devices can be effectively utilized for real-time stress detection without requiring clinical setups.

Several recent research efforts have focused on ECG-based stress classification using ML and DL techniques. Existing models have primarily addressed binary classification (stress vs. non-stress) and multi-class stress classification. For instance, in one study, HRV data was classified into stressed and normal states using Naïve Bayes, k-nearest neighbor (KNN), support vector machine (SVM), multi-layer perceptron (MLP), random forest, and gradient boosting, with the best recall score reaching 80%. Another comparative study found that SVM with radial basis function (RBF) achieved 83.33% accuracy using time-domain features and 66.66% accuracy with frequency-domain HRV features. Furthermore, CNN-based binary classification methods have yielded promising results, such as 98.4% accuracy in a recent study. Additionally, StressClick employed a random forest classifier using mouse-click patterns and gaze data to detect stress levels, demonstrating an innovative approach beyond physiological signals.

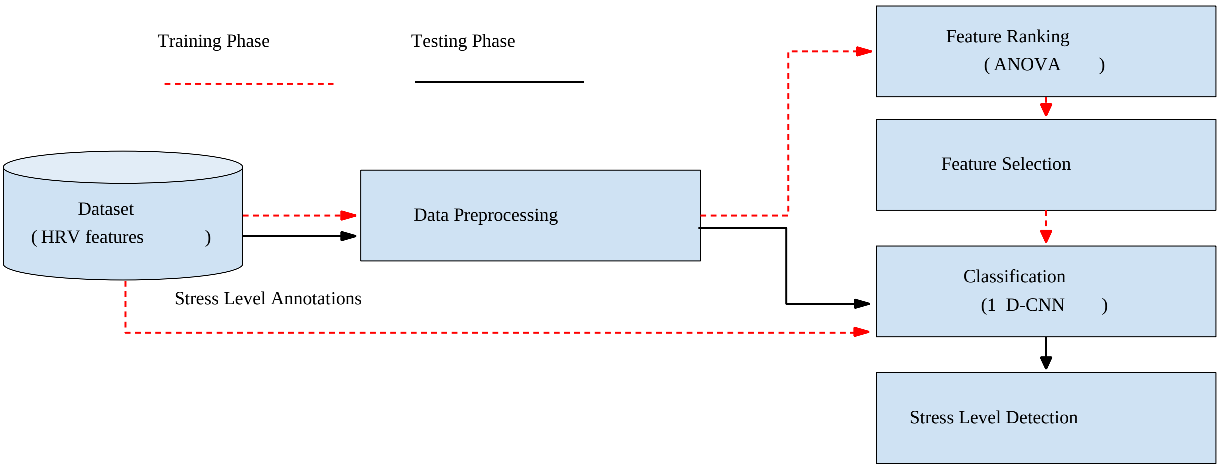


FIGURE: Framework of the proposed stress status classification model: From data collection to stress level classification.

In the domain of multi-class stress classification, SVM-based models trained on the SWELL-KW dataset have achieved 90% accuracy when distinguishing between no stress, interruption stress, and time pressure stress. Another publicly available dataset, WESAD, has been utilized for both binary (stress vs. non-stress) and three-class (amusement vs. baseline vs. stress) classifications, with ML models achieving up to 81.65% accuracy and deep learning models reaching 84.32% accuracy. More recently, novel genetic deep learning convolutional neural networks (GDCNNs) have been explored for two-dimensional stress classification, though adapting them for one-dimensional HRV data requires significant modifications. As summarized in a 2022 study, the best models trained on the SWELL-KW dataset achieved.

DATA COLLECTION AND DATASET

We adopt the SWELL−KW dataset, which was collected in a study report. Various types of data have been recorded, including computer logging, facial expression from camera recordings, body postures from a Kinect 3 dimensional (3D) sensor, heart rate (variability), and skin conductance from body sensors.

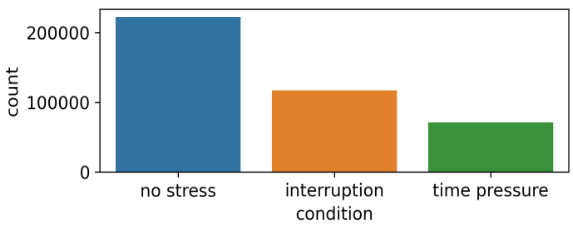
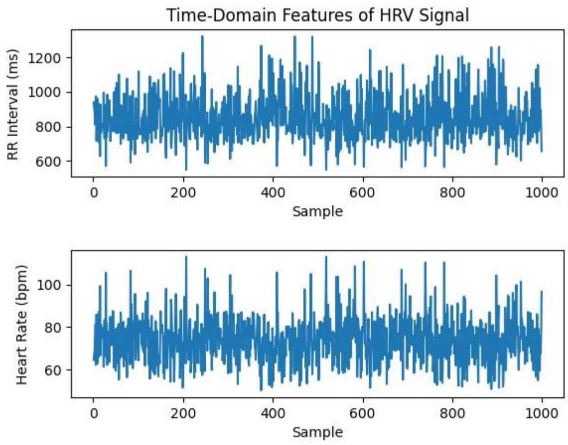


FIGURE : Distribution of Data in SWELL-KW

The SWELL−KW dataset comprises HRV computed for stress and user modeling. The subjective experiences of participants with task load, mental effort, mood, and perceived stress were also recorded. Each participant was exposed to three different working environments and the data are then labeled by medical professionals as follows.

- ***No stress***: The participants are permitted to work on the activities for as long as they need, up to 45 minutes. However, they are unaware of the maximum duration of the task.
- ***Time pressure***: Under time pressure, the time to complete the same job was decreased to 2/3 of its time in the normal condition.
- ***Interruption***: The participants were interrupted when they received 8 emails in the middle of a given activity. Some emails were pertinent to their

tasks, and the participants were asked to take particular actions, whereas others were totally irrelevant to the ongoing tasks.



FEATURE RANKING AND SELECTION FOR SWELL−KW

In this study, we prioritize feature selection by ranking attributes based on their relevance to the classification task using the ANOVA F-test method. ANOVA (Analysis of Variance) is a widely-used statistical tool for determining whether the means of three or more groups come from the same distribution. The F-test computes the ratio between variances—typically comparing between-group and within-group variances—to identify features that significantly influence classification outcomes. This method is particularly useful when comparing a numerical input variable against a categorical output, such as stress classification labels. Initially, all features from the SWELL-KW dataset are used to train the model. Subsequently, less significant features are eliminated based on their ANOVA-derived rankings. This feature reduction step optimizes the model by reducing training time without compromising classification accuracy, ensuring efficient and effective stress detection.

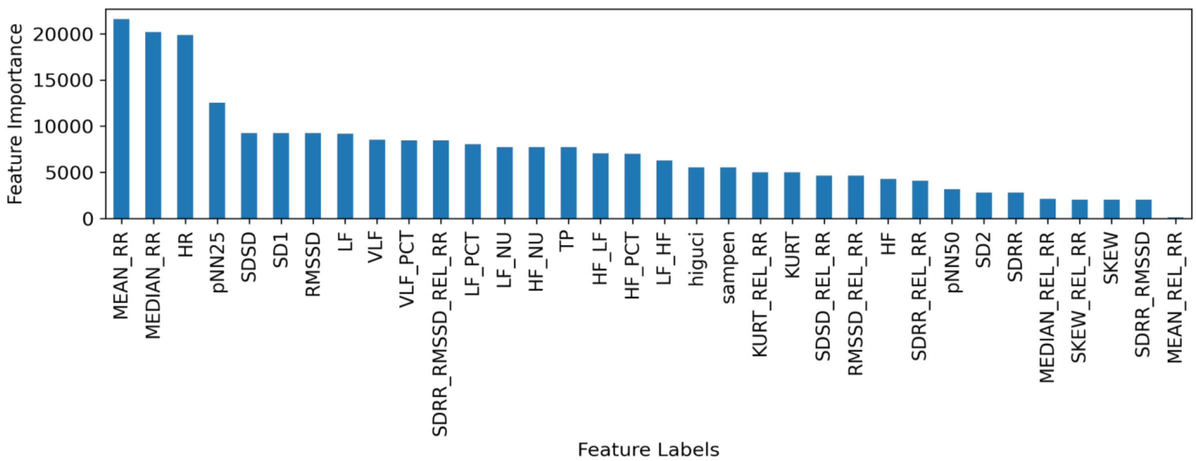
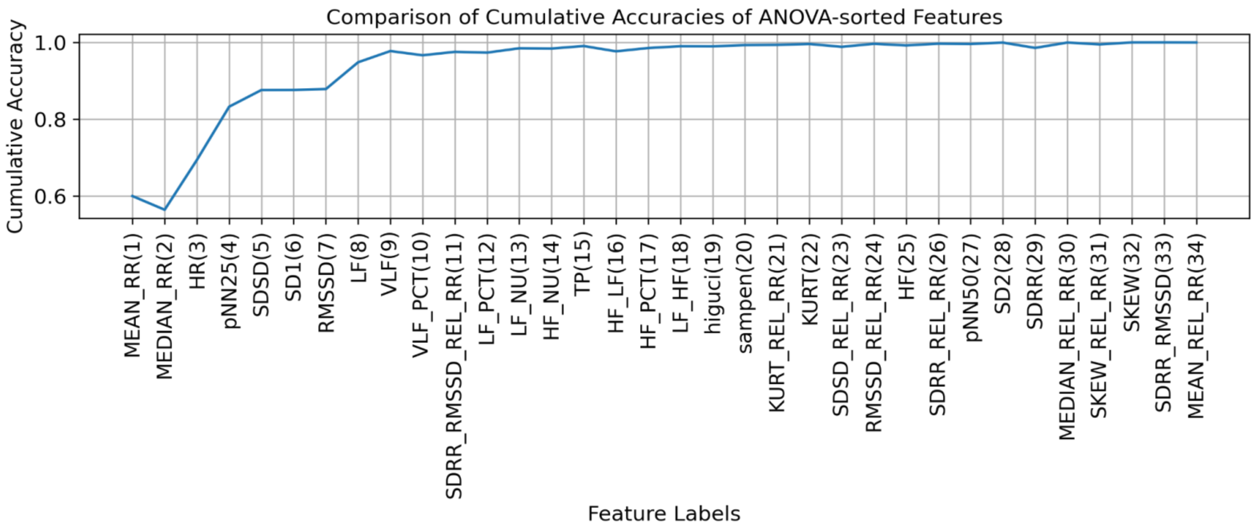


FIGURE: Feature ranking of the 34 features using ANOVA



FRAMEWORK OVERVIEW

Figure illustrates the schematic diagram of the proposed stress level classification framework. Briefly, the framework constitutes the following procedures.

- Data collection and datasets. HRV signals are collected and separated into a training dataset and a testing dataset. They will use to define the model's
- Architecture and to assess the proposed model's effectiveness.
- Data preprocessing and feature extraction. Data are preprocessed to fit into the feature ranking algorithm. In this study, ANOVA F-tests [28] and forward sequential feature selection are employed for feature ranking and selection respectively.
- Classification and validation. The designed DL-based multi-class classifier is trained, tested, and validated with significant features and annotations (e.g., *no stress, interruption condition*, and *time pressure*) labeled by medical professionals.
- Testing. In the testing phase, distinctive features are considered from the new test samples, and the class label is resolved using all classification parameters estimated in training. Different numbers of features are extracted and tested.
- Performance assessment. The performance of the classifier is measured against discrimination analysis metrics, such as *Accuracy, Precision, Recall, F1-score, and*
- *MCC.*

TABLE . Explanation of the HRV features in the SWELL–KW dataset

No.	Feature	Meaning
1	MEAN_RR	Mean of RR intervals
2	MEDIAN_RR	Median of RR intervals
3	SDRR	SD of RR intervals
4	RMSSD	Root mean square of successive RR interval differences
5	SDSD	SD of successive RR interval differences
6	SDRR_RMSSD	Ratio of SDRR over RMSSD
7	HR	Heart rate
8	pNN25	Percentage of successive RR intervals that differ more than 25 ms
9	pNN50	Percentage of successive RR intervals that differ more than 50 ms
10	SD1	Measures short-term HRV in ms and correlates with baroreflex sensitivity (BRS)
11	SD2	Measures of long-term HRV in ms and correlates with BRS
12	KURT	Kurtosis of RR intervals
13	SKEW	Skewness of RR intervals
14	MEAN_REL_RR	RR Mean of relative RR intervals
15	MEDIAN_REL_RR	Median of relative RR intervals
16	SDRR_REL_RR	SD of relative RR intervals
17	RMSSD_REL_RR	Square root of the mean of the sum of the squares of the difference between adjacent relative RR intervals
18	SDSD_REL_RR	SD of interval of differences between adjacent relative RR intervals
19	SDRR_RMSSD_REL_RR	Ratio of SDRR_REL over RMSSD_REL
20	KURT_REL_RR	Kurtosis of relative RR intervals
21	SKEW_REL_RR	Skewness of relative RR intervals
22-23	VLF; VLF_PCT	Very low (0.003 Hz - 0.04Hz) frequency activity of the HRV spectrum
24-26	LF; LF_PCT; LF_NU	Low frequency activity in the 0.04 - 0.15 Hz range
27-29	HF; HF_PCT; HF_NU	High-frequency activity in the 0.15 - 0.40 Hz range
30	TP	Total HRV power spectrum
31	LF_HF	Ratio of low to high frequency
32	HF_LF	Ratio of high to low frequency
33	sampen	Sample entropy of the RR sign
34	higuci	Higuchi Fractal Dimension

CONCLUDING REMARKS

In this study, we have developed novel a 1D CNN model for stress level classification using HRV signals and validated the proposed model based on a publicly available dataset, SWELL−KW. In our model, we also applied an ANOVA feature selection technique for dimension reduction. Through extensive training and validation, we demonstrate that our model outperforms the state-of-the-art models in terms of major performance metrics, i.e., *Accuracy, Precision, Recall, F1-score*, and *MCC* when all features are employed. Furthermore, our approach with ANOVA feature reduction also achieves excellent performance. For future work, we plan to further investigate the feasibility of optimizing the model to fit it into edge devices so that real-time stress detection can become a reality.

REFERENCES

1. 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 of the literature," *Psychiatry Investigation*, vol. 15, no. 3, pp. 235–245, Feb. 2018.
2. D. Muhajir, F. Mahananto, and N. A. Sani, “Stress level measurements using heart rate variability analysis on Android based application," *Procedia Comput. Sci.*, vol. 197, no. 8, pp. 189–197, Jan. 2022.
3. J. Held, A. Vîsla, C. Wolfer, N. Messerli-Bürgy, and C. Flückiger, “Heart˘ rate variability change during a stressful cognitive task in individuals with anxiety and control participants," *BMC Psychology*, vol. 9, no. 1, pp. 1–8, Mar. 2021.
4. K. M. Dalmeida and G. L. Masala, “HRV features as viable physiological markers for stress detection using wearable devices," *Sensors*, vol. 21, no. 8, art. no. 2873, pp. 1–18, Mar. 2021.
5. M. C. J. Abdon, A. M. Khomami, S. Nicu, and P. Ioannis, “Amigos: A dataset for affect, personality and mood research on individuals and groups," *IEEE J. Biomed. Health Informat.*, vol. 12, no. 2, pp. 479–493, Apr.–Jun. 2021.
6. E. Won, and Y.-K. Kim, “Stress, the autonomic nervous system, and the immune-kynurenine pathway in the etiology of depression," *Current Neuropharmacology.*, vol. 14, no. 7, pp. 665–673, Oct. 2016.
7. B. Olshansky, H. Sabbah, N.Hani, J.P. Hauptman, and W.S. Colucci, “Parasympathetic nervous system and heart failure: Pathophysiology and potential implications for therapy," *Circulation.*, vol. 118, no. 8, pp. 863– 871, Aug. 2008.
8. S. Goel, P. Tomar, and G. Kaur, “ECG feature extraction for stress recognition in automobile drivers," *Electron. J. Biol.*, vol. 12, no. 2, pp. 156–165, Mar. 2016.
9. V. N. Hegde, R. Deekshit, and P. S. Satyanarayana, “A review on ECG signal processing and HRV analysis," *J. Medical Imaging Health Informat.*, vol. 3, no. 2, pp. 270–279, Jun. 2013.
10. M. Vollmer, “A Robust, simple and reliable measure of heart rate variability using relative RR intervals," in *Proc. Comput. Cardiology Conf.*, 2015.