Smart Techno (Smart Technology, Informatic, and Technopreneurship)
Volume 7, Number 1, 2025 pp. 45-50
E-ISSN: 2541-0679
Open Access: https://lppm.primakara.ac.id/jurnal/index.php/smart-techno/

Convolutional Neural Network-Based Human Stress Detection Using Multivariate Sensor Data and Cross-Validation

Asma' Abu Samah[1*], Nor Fadzilah Abdullah[2]
[1,2]Universiti Kebangsaan, Malaysia

ARTICLE INFO
Article history:
Initial submission
21-04-2025
Received in revised form
24-04-2025
Accepted 24-04-2025
Available online 25-04-2025

Keywords:
Human Stress, Convolutional Neural Networks, K-fold Cross Validation, Multivariate

DOI:
https://doi.org/10.59356/smart-techno.v7i1.155

ABSTRACT

Early identification of human stress levels plays a crucial role in promoting mental well-being and preventing related health issues. However, conventional stress assessment methods often involve multi-step procedures or subjective evaluations, making them inefficient and impractical for continuous monitoring. This study introduces a Convolutional Neural Network (CNN)-based approach to automatically detect human stress using multivariate sensor data, such as heart rate, oxygen saturation, body temperature, and movement signals. Unlike traditional machine learning methods that rely on handcrafted features and shallow classifiers, the proposed deep learning model leverages raw sensor data to learn hierarchical representations of physiological patterns associated with various stress levels. The dataset utilized in this research is the SaYoPillow dataset obtained from Kaggle, which includes labeled physiological signals based on subjective stress assessments. Input features are normalized and reshaped into one-dimensional sequences compatible with the CNN architecture. A stratified 5-fold cross-validation strategy is used to ensure robust and generalizable model performance. The proposed CNN model achieved an outstanding accuracy of 0.999, with a precision of 0.998, recall of 0.991, and F1 score of 0.994, outperforming baseline models such as Decision Tree with accuracy of 0.987 and Random Forest with accuracy of 0.981. These results highlight the CNN model's strong potential for real-time, reliable stress monitoring using wearable sensors, making it a promising solution for digital health and well-being applications.

1. INTRODUCTION

Research on human stress, a multifaceted physiological and psychological response to perceived threats, is increasingly critical due to its link to various health issues and the growing emphasis on mental well-being. While acute stress can be adaptive, chronic stress poses significant risks, underscoring the need for robust detection methods in areas like health monitoring, workplace efficiency, and human-computer interaction (Arsalan et al., 2022).

Traditional stress assessment relies on subjective and retrospective methods like self-report questionnaires and clinical interviews, which are susceptible to bias and lack real-time monitoring capabilities. To overcome these limitations, research has shifted towards physiological signal analysis using wearable biosensors. These sensors capture data such as heart rate, respiration, and body temperature to infer stress. Conventional machine learning techniques like Random Forest and SVM have been applied to this data, often requiring extensive preprocessing and manual feature engineering (Singh & Hooda, 2023). However, these methods can struggle with real-world variability, sensor noise, and the complexity of physiological data, limiting their generalizability and introducing subjectivity.

Recent research explores the potential of deep learning, particularly Convolutional Neural Networks (CNNs), for automated feature extraction and human state recognition. CNNs excel at learning complex patterns from time-series physiological data, potentially overcoming

** Corresponding Author: Asma' Abu Samah: asma@ukm.edu.my*

the limitations of manual feature selection. However, current CNN-based stress detection research often suffers from issues like overfitting and insufficient validation against traditional methods.

Addressing these challenges, this research introduces a 1D CNN framework for stress level detection using multivariate data from the SaYoPillow dataset. The study's key contributions include a CNN model for classifying stress levels, rigorously validated using 5-fold cross-validation to ensure robustness and mitigate overfitting. Furthermore, the study provides a comparative evaluation against established classifiers (SVM and Decision Tree) to demonstrate the efficacy of the proposed deep learning approach. By employing deep learning and thorough validation, this research aims to contribute to a scalable, real-time, and reliable methodology for human stress detection, paving the way for its integration into wearable health technologies and intelligent environments.

2. LITERATURE REVIEW

Traditional methods for stress detection have primarily utilized classical machine learning algorithms such as Support Vector Machines (SVMs), Decision Trees (DTs), and K-Nearest Neighbors (KNN), often in combination with features manually extracted from physiological signals (Guido, Ferrisi, Lofaro, & Conforti, 2024). While these approaches provide a baseline for stress classification, their effectiveness is limited by the variability of human responses and the complexity of biosignal patterns.

Recent studies have explored the use of deep learning to overcome these challenges. For instance, (Lima, Osório, & Gamboa, 2021) demonstrated the application of deep CNNs for stress level prediction using heart rate variability and electrodermal activity. Their approach outperformed SVM classifiers by automatically identifying time-series patterns that correlate with stress conditions. Similarly, Airlangga, Bata, Adi Nugroho, and Lim (2025) investigated LSTM and CNN-LSTM hybrid models, showing the benefits of capturing both spatial and temporal dependencies in multivariate sensor inputs.

Several studies also emphasized the importance of evaluating model generalizability through cross-validation. (White & Power, 2023) applied a k-fold cross-validation approach in their emotion recognition system, highlighting improved consistency in performance metrics compared to single-split evaluations. Moreover, the work of Alzubaidi et al. (2021) revealed that CNNs trained on normalized sensor data outperformed other architectures in both accuracy and inference speed.

Despite these advancements, few studies have focused on a structured comparison of CNN-based models under robust cross-validation settings using publicly available stress datasets. This study fills that gap by implementing a CNN architecture tailored for 1D physiological signal input and evaluating it comprehensively using stratified 5-fold cross-validation. The results are also benchmarked against traditional classifiers, providing insight into the model's applicability for real-world stress monitoring systems.

3. METHOD

In this research, we present a system designed for the classification of human stress levels. The methodology involves the application of a 1D Convolutional Neural Network (CNN) to analyze multivariate physiological data. The initial steps of the research include data acquisition and preprocessing, encompassing cleaning, normalization, and label encoding of the raw dataset. Subsequently, a 1D CNN architecture is developed to automatically learn relevant spatial features from the multi-sensor input. To ensure the reliability and generalizability of the model, a k-fold cross-validation strategy is employed for both training and evaluation. The performance of the proposed system is quantitatively assessed using standard classification metrics, specifically accuracy, precision, recall, and F1-score.

3.2 Dataset Collection and Preprocessing

This study utilizes the SaYoPillow dataset, which includes physiological readings such as heart rate, temperature, blood oxygen level, snoring range, respiration rate, limb movement, and eye movement (Rachakonda, Bapatla, Mohanty, & Kougianos, 2021). These features correspond to labeled stress levels such as low, medium low, medium high, and high. Low converted into integer label as 1, followed by medium low as 2, then medium high and high transformed into 3 and 4, respectively. The dataset provided by Rachakonda et al. (2021) detailed in comma separated column (.csv) with total number of instances of 630 data. The number of data is limited, to ensure the originality of the dataset characteristic, data augmentation process is not conducted. The description of the SaYoPillow dataset is detailed in Table 1.

Table 1. Dataset Description of the SaYoPillow dataset.

No	Feature Name	Symbol	Min	Max	Std
1	Snoring Rate	sr	45.0	100.0	19.372
2	Respiration Rate	rr	16.0	30.0	3.966
3	Body Temperature	t	85.0	99.0	3.529
4	Limb Movement	lm	4.0	19.0	4.299
5	Blood Oxygen	bo	82.0	97.0	3.902
6	Eye Movement	rem	60.0	105.0	11.893
7	Sleeping Hours	sr.1	0.0	9.0	3.054
8	Heart Rate	hr	50.0	85.0	9.915

In order to train the model proposed in this work, SaYoPillow dataset were processed using several steps for supervised learning approach, such as dividing data features and label, label encoding to categorize the stress label and convert it to integer values, normalization to scale the data to a 0-1 range using Min-Max scaling to ensure uniform contrbution, and reshaping to match the 1D CNN model input format. The dataset is then split into training and testing sets while ensuring class distribution is preserved using stratification.

3.3 CNN Model Architecture

The core of the proposed system in this work is a 1D CNN model designed to extract feature patterns across multivariate time-series data. The architecture consists of several 1D Convolutional layers with MaxPooling1D, Dropout and Dense layer. The overview of the proposed architecture shown in Figure 1.

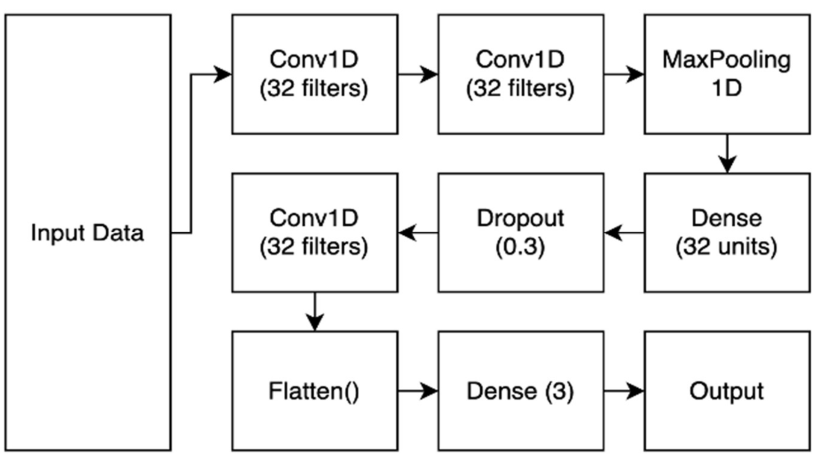

Figure 1. The general overview of the proposed Layered CNN-based MRI Classification.

Based on given architecture, the first Convolution layers detect local signal patterns, followed by the next Convolution to conduct additional feature extraction. The MaxPooling1D is used to reduce the spatial dimension of the model learning space, whereas the Dropout prevents overfitting during training process. Lastly, Flatten used to convert learning weights into 1D vector before it passed to softmax activation to give final prediction. Additionally, the model uses the Adam optimizer with a learning rate of 0.001, categorical cross-entropy loss, and ReLU activation in hidden layers. Training is performed for 100 epochs with a batch size of 8.

3.4 Cross-Validation and Evaluation Metrics

To assess the generalizability of the proposed model, a k-fold stratified cross-validation strategy is implemented (Gorriz et al., 2024). The fold adopted in this work is 5. This is due to the limitation of the data instances provided in the datasets. In each fold, the dataset is split that 90% is used for training and 10% for testing, while preserving class balance. The final results are averaged across all folds. For each fold, the proposed CNN model is re-compiled to avoid continous learning from the previous learning process.

To further investigate the model performance, various performance metrics adopted in this work, such as Accuracy to show the ratio of correct predictions to total predictions, Precision to indicate the proportion of true positive predictions among all predicted positives, recall assessing the proportion of true positives among all actual positives and F1 Score which represent the harmonic mean of precision and recall. These metrics offer a comprehensive performance assessment, especially in multi-class classification settings. The accuracy calculated using Equation 1, followed by Precision in Equation 2, Recall and F1 Score in Equation 3 and 4, respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{2\ x\ Precision\ x\ Recall}{Precision+Recall} \quad (4)$$

where TP represents the model correctly predict the positive class, followed by TN when the model successfully predict the negative class. False positive denotes as FP indicates the model failed to predict the positive class and False negative refers as FN means that the model incorrectly predicts the negative class.

4. RESULT AND DISCUSSION

The evaluation in this work conducted using simulation-based approach using Kaggle Notebook as the development environment. Main library used in this work are Keras, Tensorflow, Numpy, Pandas and Scikit-learn. In context of evaluationg the proposed CNN-based classifier performance, several machine learning-based models were adopted, such as Random Forest and Decision Tree. The Random Forest model considered in this work configured with 100 estimators, with gini criteria, and sqrt as the max features. In terms of Decision Tree model, the criterion settings are set to gini with best splitter and 2 minimal sample split and without any maximum depth of the tree.

The first evaluation conducted without the 5-fold cross-validation technique to show the performance of the model in single data spliting approach. The proposed CNN-based model compared with RF and DT. Figure 2 shows the accuacy performance of the DT and RF model along with the proposed CNN model.

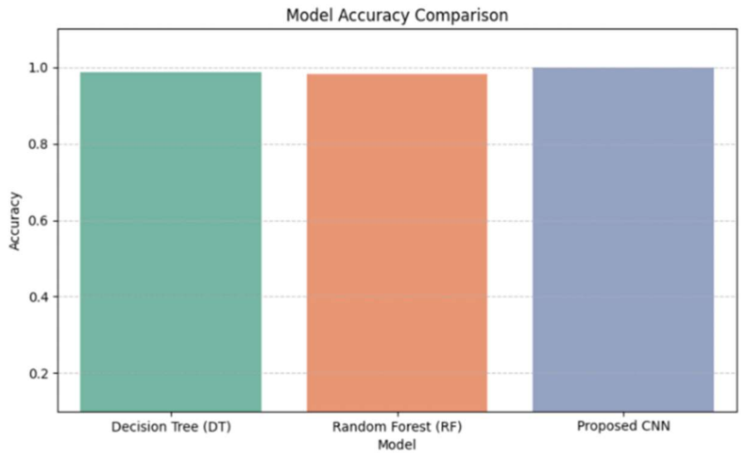

Figure 2. The accuracy of various models compared in this work.

Based on the results, the accuracy of DT model is 0.987, followed by RF with 0.981 accuracy value. The proposed CNN model able to produce better performance, with accuracy value of 0.99 (99%). To further show the superiority of the proposed CNN model, extensive performance evaluation was conducted. The performance evaluation of three different models is detailed in Table 2.

Table 2. Extensive performance evaluation of three models investigated in this work.

No	Model Name	Accuracy	Precision	Recall	F1 Score
1	Decision Tree	0.987	0.976	0.952	0.963
2	Random Forest	0.981	0.984	0.960	0.972
3	Proposed CNN	0.999	0.998	0.991	0.994

Table 2 presents the performance comparison of three machine learning models, DT, RF, and the proposed CNN in the context of human stress level detection using multivariate sensor data. The evaluation metrics include accuracy, precision, recall, and F1 score. The Decision Tree model achieved an accuracy of 0.987, with a precision of 0.976 and recall of 0.952, resulting in an F1 score of 0.963. While it shows strong performance overall, its relatively lower recall indicates a higher tendency to miss stress instances compared to the other models. The Random Forest model, being an ensemble of decision trees, demonstrated an improved balance between precision (0.984) and recall (0.960), producing a higher F1 score of 0.972. Despite having a slightly lower accuracy (0.981) than the Decision Tree, the Random Forest exhibited a more consistent and generalized classification performance.

Among the three, the proposed CNN model outperformed the traditional approaches across all metrics, achieving an accuracy of 0.999, precision of 0.998, recall of 0.991, and F1 score of 0.994. This remarkable result indicates the CNN’s superior ability to learn complex patterns from multivariate sensor inputs, enabling highly accurate stress level predictions. The minimal gap between precision and recall further suggests the model's robustness and reliability in classifying both stressed and non-stressed states without significant bias. Overall, these findings affirm the effectiveness of deep learning, particularly CNN architectures, in enhancing stress detection systems and overcoming limitations associated with conventional machine learning models.

5. CONCLUSION

This study presents a CNN-based model for detecting human stress levels using multivariate physiological sensor data, addressing key limitations in conventional stress assessment techniques. By leveraging raw input signals such as heart rate, blood oxygen levels, and body temperature from the SaYoPillow dataset, the proposed 1D CNN architecture successfully learns hierarchical representations that distinguish between low, medium, and high stress states. The use of stratified 5-fold cross-validation ensures the robustness and generalizability of the model, mitigating overfitting and yielding highly reliable results.

The CNN model outperforms traditional machine learning classifiers like Decision Tree and Random Forest across all performance metrics, achieving a remarkable accuracy of 0.999, precision of 0.998, recall of 0.991, and F1 score of 0.994. These findings underscore the capability of deep learning in capturing the complex, nonlinear patterns of physiological stress responses, providing a promising direction for real-time stress monitoring through wearable technologies.

6. REFERENCES

Airlangga, G., Bata, J., Adi Nugroho, O. I., & Lim, B. H. P. (2025). Hybrid CNN-LSTM Model with Custom Activation and Loss Functions for Predicting Fan Actuator States in Smart Greenhouses. AgriEngineering, 7(4), 118. https://doi.org/10.3390/agriengineering7040118

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Arsalan, A., Majid, M., Nizami, I. F., Manzoor, W., Anwar, S. M., & Ryu, J. (2022). Human stress assessment: A comprehensive review of methods using wearable sensors and non-wearable techniques. arXiv preprint arXiv:2202.03033.

Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. Information, 15(4), 235. https://doi.org/10.3390/info15040235

Gorriz, J. M., Segovia, F., Ramirez, J., Ortiz, A., & Suckling, J. (2024). Is K-fold cross validation the best model selection method for Machine Learning?. arXiv preprint arXiv:2401.16407.

Jegan, R., Mathuranjani, S., & Sherly, P. (2022). Mental Stress Detection and Classification using SVM Classifier: A Pilot Study. In 2022 6th International Conference on Devices, Circuits and Systems (ICDCS) (pp. 139-143).

Rachakonda, L., Bapatla, A., Mohanty, S., & Kougianos, E. (2021). SaYoPillow: Blockchain-Integrated Privacy-Assured IoMT Framework for Stress Management Considering Sleeping Habits. IEEE Transactions on Consumer Electronics, 67(1), 20-29.

Singh, S., & Hooda, S. (2023). A Study of Challenges and Limitations to Applying Machine Learning to Highly Unstructured Data. In 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA) (pp. 1-6).

White, J., & Power, S. D. (2023). k-Fold Cross-Validation Can Significantly Over-Estimate True Classification Accuracy in Common EEG-Based Passive BCI Experimental Designs: An Empirical Investigation. Sensors, 23(13), 6077. https://doi.org/10.3390/s23136077