International Journal of Advances in Applied Sciences (IJAAS)
Vol. 14, No. 4, December 2025, pp. 1137~1145
ISSN: 2252-8814, DOI: 10.11591/ijaas.v14.i4.pp1137-1145

Analysis of feature reduction for identifying stress levels electroencephalogram signal based

Setyorini[1,2], Ilham Ari Elbaith Zaeni[1], Hakkun Elmunsyah[1]
[1]Department of Electrical Engineering and Informatics, Faculty of Engineering, Universitas Negeri Malang, Malang, Indonesia
[2]Department of Informatics Engineering, Faculty of Technology and Design, Institut Teknologi dan Bisnis Asia, Malang, Indonesia

Article Info

Article history:

Received Jun 17, 2025
Revised Oct 20, 2025
Accepted Nov 4, 2025

Keywords:

Ensemble
Feature extraction
Feature reduction
Independent component analysis
Naive bayes
Principal component analysis
Support vector machine

ABSTRACT

Stress identification based on electroencephalogram (EEG) signals has become a rapidly growing research topic, with the main approaches utilizing features from the frequency domain and time-frequency domain. This research aims to combine principal component analysis (PCA) and independent component analysis (ICA) for feature extraction to improve the accuracy of stress identification. Additionally, PCA+ICA features are reduced from 64 to 32 columns to optimize computational efficiency without losing important information from the EEG signal. The stress identification models used in this research include Ensemble, naive Bayes, and support vector machine (SVM). The data used are from the SAM-40 task Stroop color trials 1, 2, and 3. Experimental results indicate that the combination of PCA+ICA features improves accuracy only in the ensemble method. Reducing PCA+ICA features from 64 to 32 columns led to an improvement in accuracy only for Stroop trial 2 data with the naive Bayes method.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Setyorini
Department of Electrical Engineering and Informatics, Faculty of Engineering, Universitas Negeri Malang
Malang, Indonesia
Email: setyorini.2205349@students.um.ac.id

1. INTRODUCTION

The electrical activity of the brain is commonly observed, measured, and recorded as voltage using electroencephalogram (EEG) signals [1], [2]. Stress has emerged as a major health concern in many developed countries, exerting a significant impact on both physical and mental well-being. It affects individuals across diverse professions and can contribute to or exacerbate disorders related to the brain, immune system, and endocrine system. Negative stress, in particular, triggers the release of triglycerides and cholesterol into the bloodstream, increasing blood flow and heart rate. Early-life stress has been linked to migraines and reduced physical activity, while long-term stressors are associated with cardiac dysfunction. Because stress is a key determinant of both physical and mental health, accurately identifying stress levels is critical in fields such as healthcare, psychology, and occupational performance [3]–[5]. EEG has proven to be an effective method for detecting stress [6]–[8]. By recording the brain’s electrical activity, EEG provides valuable insights into an individual’s physiological condition. However, EEG signals are inherently high-dimensional, often containing redundant information, which increases computational complexity and may reduce classification performance [9]–[11]. To address this challenge, feature reduction plays a vital role in improving both the efficiency and accuracy of EEG-based stress analysis.

Principal component analysis (PCA) and independent component analysis (ICA) are two widely used methods for EEG feature reduction and extraction [12]. PCA reduces data dimensionality by retaining components that account for the greatest variance, thereby minimizing redundancy and noise. ICA, on the other hand, extracts statistically independent components, which can improve data quality by isolating meaningful signal sources [12], [13]. Both methods have shown effectiveness in enhancing EEG analysis for stress classification, yet their combined potential remains underexplored. Previous research has employed a variety of EEG feature extraction methods, including statistical measures from the time domain, wavelet-based features from the time-frequency domain [14], frequency domain features extracted from five-channel EEG signals, and hybrid deep learning approaches utilizing the discrete wavelet transform (DWT) [15]–[17]. EEG has also been extensively applied in related domains such as epilepsy detection, sleep stage classification, emotion recognition, and stroke diagnosis [18]–[20]. These works highlight the versatility of EEG in biomedical and computational applications [21]–[24]. Despite these advances, EEG data processing continues to face major challenges, including signal noise, inter-individual variability, and the complexity of brain wave patterns [25]–[29]. To mitigate these issues, researchers have applied preprocessing techniques to remove noise and artifacts [30]–[33]. While many researchers have relied on frequency and time-frequency domain features to identify stress, these methods often struggle with high-dimensional complexity, limiting their accuracy and efficiency [34], [35].

This research addresses these limitations by proposing a hybrid feature extraction framework that integrates PCA and ICA for EEG-based stress classification. PCA reduces dimensionality by capturing global variance patterns, while ICA complements this by isolating independent signal sources. The integration of both methods is expected to produce a more compact, informative, and robust representation of EEG signals, preserving critical features while minimizing redundancy. This, in turn, can improve the performance of machine learning classifiers, particularly algorithms such as naive Bayes, which are sensitive to high-dimensional data.

While both PCA and ICA have been extensively researched in isolation for EEG-based stress analysis, there is a notable lack of systematic research exploring their combined use for multi-level stress classification. Addressing this gap, the present research introduces a PCA+ICA hybrid approach to simultaneously enhance classification accuracy and computational efficiency. Prior work has shown that combining PCA and ICA can yield richer feature representations and improved classification performance in EEG-related tasks such as emotion recognition [34], [35]. However, their application in multi-level stress classification remains largely unexamined.

Therefore, this research contributes to the field by systematically evaluating a hybrid PCA+ICA framework for feature reduction in EEG-based stress detection. The proposed method aims to minimize redundancy and dimensionality while preserving critical information, thereby improving both computational efficiency and classification accuracy. Table 1 summarizes the methodological differences between previous research and our approach, highlighting how the integration of PCA and ICA advances the current state of EEG-based stress classification research.

Table 1. Comparison of previous research and the proposed research

Aspect	Previous researches	This research (proposed)
Feature extraction technique	PCA only [12], ICA only [35], wavelet transform [16]	Combined PCA+ICA
Use of PCA	Applied for dimensionality reduction to preserve variance [12]	Used to reduce dimensionality and integrated with ICA for enriched feature representation
Use of ICA	Mainly used for artifact removal or signal separation [35]	Used as a parallel feature extractor, merged with PCA before classification
Stress classification level	Primarily binary classification tasks (stress vs. no stress) [13], [23]	Multi-level stress classification across three Stroop trials
Dimensionality reduction strategy	Fixed feature sets or PCA-only reduction [12]; no integration or correlation-based selection	Correlation-based selection from merged PCA+ICA features (reduced from 64 to 32 columns)
Evaluation dataset	DEAP, arithmetic EEG tasks [25]; other non-Stroop datasets	SAM-40 dataset using Stroop color-word task [36]
Classification models	Support vector machine (SVM) [13], k-nearest neighbor (KNN) [30], deep learning (multilayer perceptron, convolutional neural network, long short-term memory) [16]	Ensemble, naive Bayes, and SVM applied on PCA, ICA, and combined PCA+ICA feature sets
Novelty	Focus on single-method feature extraction; no systematic integration of PCA and ICA	First to systematically evaluate the PCA+ICA combination for EEG-based multi-level stress classification

However, there is a lack of systematic researches that integrate PCA and ICA for multi-level stress classification using EEG signals, particularly focusing on both dimensionality reduction and independent feature extraction to enhance classification accuracy and computational efficiency. To address these limitations, this research contributes by:

- Combining PCA and ICA for feature extraction aims to obtain a more optimal and informative representation of EEG features. Although both are widely used, PCA may fail to capture non-linear or class-specific features [1], while ICA assumes statistical independence among sources and does not provide an inherent ranking of components [3], [4]. Therefore, this research integrates PCA and ICA to leverage their complementary strengths and enhance EEG-based stress classification.
- Reducing the number of features generated by PCA and ICA from 64 to 32 columns, enhancing computational efficiency while preserving critical information from the EEG signal. The combination of PCA and ICA leverages the strengths of both techniques, dimensionality reduction through principal components and separation of statistically independent sources, resulting in more informative and discriminative EEG features for stress classification tasks.

2. METHOD

This subsection describes the research method and dataset employed in this research. The methodology is designed to ensure that the proposed hybrid extraction feature PCA+ICA framework for EEG feature reduction can be systematically implemented and evaluated. It outlines the stages of data acquisition, feature extraction, and classification, providing a clear overview of how the research objectives are addressed.

2.1. Dataset

This research utilizes EEG data from the publicly available SAM-40 dataset [36], which includes recordings from 40 participants. Each participant completed a series of cognitive tasks, including the Stroop color-word test, which is specifically used in this research to induce varying levels of mental stress. The Stroop task is divided into three trials, each lasting 25 seconds**,** where participants are required to identify the color of words that may or may not match their semantic meaning, a task known to elicit cognitive interference and stress. EEG signals were recorded using a 32-channel Emotiv Epoc Flex gel kit with a sampling rate of 128 Hz. The electrode placement follows the international 10-20 system, and the recorded channels include: Cz, Fz, Fp1, F7, F3, FC1, C3, FC5, FT9, T7, CP5, CP1, P3, P7, PO9, O1, Pz, Oz, O2, PO10, P8, P4, CP2, CP6, T8, FT10, FC6, C4, FC2, F4, F8, and Fp2 as shown in Figure 1.

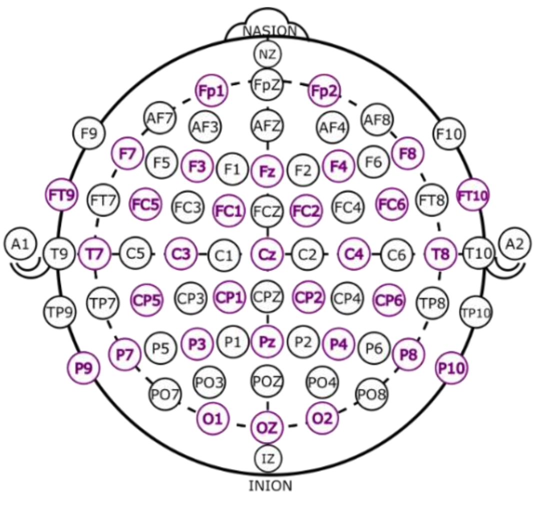

Figure 1. Electrode positioning with 32 electrodes

During data acquisition, participants were seated in a quiet room and instructed to minimize movement to reduce artifacts. Prior to the experiment, all subjects provided informed consent, and the research protocols were approved by the respective institutional ethical review board. EEG signals were recorded continuously while the participants performed each Stroop trial, and the corresponding stress levels were labeled based on task type and difficulty. The resulting dataset consists of a total of 120 EEG recordings (40 subjects×3 Stroop trials), each with 32-channel time-series data and a corresponding stress level classification (low, moderate, and high), which are used as ground truth labels for supervised learning.

2.2. Proposed method

Explanation of Figure 2 is the stages of this research proposal:

- EEG data acquisition: EEG data is obtained from a publicly available dataset. This research specifically utilizes Stroop color-word EEG data from trials 1, 2, and 3.
- Feature extraction: methods such as ICA and PCA. The EEG data, recorded from 32 channels, results in 32 columns per feature extraction method (PCA and ICA). When combined, the PCA+ICA feature set consists of 64 columns.
- Feature reduction to enhance computational efficiency: feature reduction is conducted using the correlation method in (1). The 64 extracted feature columns are reduced to 32 columns by selecting those with the highest correlation values.
- Stress classification: stress levels are classified into high and low categories as defined in Table 1. The classification process employs Ensemble [37], naive Bayes [38], [39], and SVM algorithms.

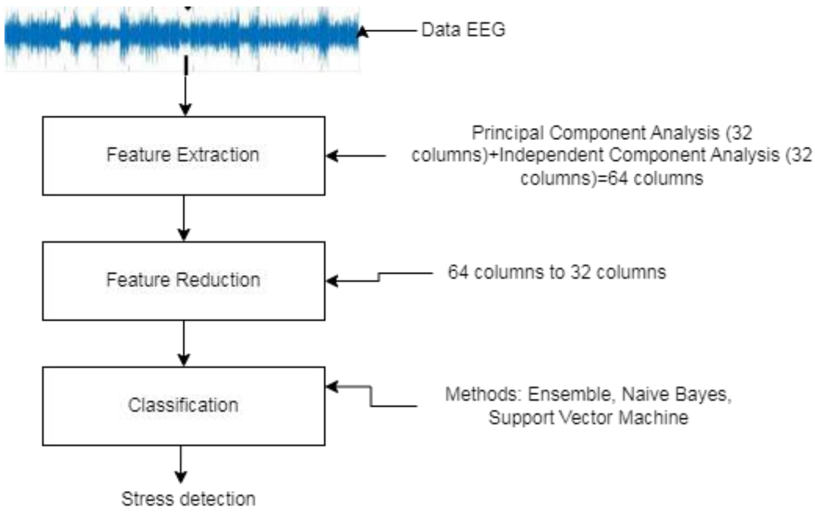

Figure 2. Stages of research

$$r_{x,y} = \frac{n\sum xy-(\sum x\sum y)}{\sqrt{((n\sum x^2)-(\sum x)^2)((n\sum y^2)-(\sum y)^2)}} \quad (1)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

Each column of the PCA+ICA feature x is calculated for correlation with stress levels (y) (high and low). This research uses a combination of PCA and ICA features to optimize data representation by reducing dimensions while extracting more informative features. Most previous researches on the feature extraction process use PCA [12], multi-domain in the time domain, frequency domain, time-frequency domain [40], and wavelet [15], [16]. The difference between this research and the previous ones is combining PCA features with ICA, with a total of 64 columns, reducing the columns to 32 to increase accuracy in identifying stress. PCA performs feature extraction based on eigenvalues and eigenvectors of the covariance matrix. ICA is used to decompose mixed signals into more independent components by finding non-Gaussian transformations that maximize statistical independence between features, and more independent components can be used as the main features. Several previous studies of stress identification used the KNN method [31], SVM [30], and naive Bayes [13]. Based on previous research, the method that has high accuracy to determine stress is SVM, so this research uses SVM, naive Bayes, and Ensemble methods. Statistical learning theory, founded on the principle of structural risk minimization, is applied in SVM. Based on given labels, an SVM selects a hyperplane that divides the feature space into a control group and a stress group. SVMs offer strong generalization performance and reduce the risk of data overfitting. Based on Bayes' theorem, naive Bayes is a probabilistic classifier. It utilizes the maximum posterior hypothesis from statistics and performs well with high-dimensional input data. As a nonlinear classifier, it is effective in practical applications. Furthermore, the naive Bayes classifier requires only a small amount of training data to approximate the statistical parameters. This research uses accuracy in (2) to evaluate the stress identification method [23].

3. RESULTS AND DISCUSSION

This research classifies stress using feature reduction from PCA and ICA. The data of this research are task Stroop trials 2, 3, and 4 from SAM-40 [36]. The data is divided into training data to create a model and testing data to test the model. The process of dividing training and testing data is random, so that each training and testing process is carried out three times. Table 2 shows the average accuracy of the classification model from the PCA feature. Table 3 shows the average accuracy of the classification model from the ICA feature. Based on Tables 2 and 3, the feature that has accuracy in identifying stress is ICA.

This research combines PCA+ICA features, and the total feature columns is 64. The average accuracy results of the combined PCA+ICA features are in Table 4. Table 5 shows the average accuracy of the PCA+ICA feature reduction to 16 columns. Table 6 shows the average accuracy of the PCA+ICA feature reduction to 32 columns. Based on the average model accuracy in Table 5, the results of the PCA+ICA feature combination did not experience an increase in accuracy compared to Table 3. The SVM method showed the best accuracy in all Stroop tasks, with the highest accuracy in trial 1 (0.772), trial 2 (0.898), and trial 3 (0.753). The naive Bayes method has lower accuracy than SVM, but better than ensemble. The SAM-40 data on the Stroop trial 3 task always had high accuracy, as in Tables 2-6.

Table 2. Model accuracy of PCA features

Methods	Trial 1	Trial 2	Trial 3
Ensemble	0.684	0.883	0.641
Naive Bayes	0.751	0.893	0.730
SVM	0.753	0.896	0.730

Table 3. Model accuracy of ICA features

Methods	Trial 1	Trial 2	Trial 3
Ensemble	0.704	0.891	0.667
Naive Bayes	0.764	0.892	0.732
SVM	0.772	0.898	0.753

Table 4. Model accuracy of PCA+ICA

Methods	Trial 1	Trial 2	Trial 3
Ensemble	0.706	0.892	0.678
Naive Bayes	0.728	0.891	0.702
SVM	0.772	0.898	0.753

Table 5. Model accuracy from PCA+ICA feature reduction (16 columns)

Methods	Trial 1	Trial 2	Trial 3
Ensemble	0.665	0.883	0.655
Naive Bayes	0.732	0.865	0.742
SVM	0.724	0.893	0.743

Table 6. Model accuracy of PCA+ICA feature reduction (32 columns)

Methods	Trial 1	Trial 2	Trial 3
Ensemble	0.684	0.900	0.639
Naive Bayes	0.741	0.908	0.723
SVM	0.753	0.908	0.729

Based on Tables 4 and 5, the PCA+ICA feature reduction process 64 to 16 columns, has an average accuracy decrease. This means that reducing to 16 columns cannot increase the accuracy in stress identification, only in trial 1 (0.4%) and trial 3 (4%); the naive Bayes method increased its accuracy. The decrease in accuracy in Table 4 is due to changes in the distribution of test data that make it difficult for the model to classify. There are adjustments to the parameters/attribute columns that make it difficult for the model to adjust.

Based on Tables 4 and 6, the process of reducing PCA+ICA features from 64 to 32 columns experienced an increase in accuracy. Stroop trial 2 data in all models experienced an increase in accuracy: ensemble (0.8%), naive Bayes (1.7%), and SVM (1%). And the naive Bayes method always experienced an increase in accuracy, Stroop trial 1 data (1.3%), trial 2 (1.7%), and trial 3 (2.1%). It was concluded that by reducing the PCA+ICA features from 64 to 32 columns, there was an increase only in the Stroop trial 2 data, and the method that experienced an increase was naive Bayes.

Among all tested configurations, the highest accuracy was achieved in trial 2 using the PCA+ICA feature extraction method with 32 features, reaching 90.8%. This indicates that 32 features provide an optimal balance between information richness and dimensionality reduction for EEG-based stress classification. Comparison of model accuracy on SAM-40 data in previous research is the accuracy of the SVM model of PCA features on the arithmetic task trial 3 data is 58% [12]. Table 7 shows the accuracy of the SVM and deep learning models on the SAM-40 Stroop task data [23]. Table 7 shows the results of the model accuracy on the SAM-40 dataset, but it should be noted that the dataset, preprocessing, and data division in this research are not the same as previous research. This research uses the SAM-40 dataset with the Stroop color word task, while previous research used the arithmetic task. In addition, the feature extraction method applied is a combination of PCA and ICA, different from other methods that may have been used previously, such as PCA or wavelets. The data division in this research was done randomly, while previous research may have used the k-fold cross-validation method. Therefore, the accuracy results shown in Table 7 are only a general overview and cannot be used as a direct comparison with the results of this research. To further assess the competitiveness of this method, Table 7 compares the accuracy obtained in this research with several related works.

Table 7. Model accuracy in research [23]

Feature	Classification methods	Accuracy
Frequency domain (BP)	SVM	87.667
Time domain (HP)	SVM	86.500
Non-linear features (HFD)	SVM	89.830
EEGNet	EEGNet	92.500
Shallow ConvNet	Shallow ConvNet	90.440
Deep ConvNet	Deep ConvNet	75.750

The results of column reduction from PCA+ICA features from 64 to 32 columns from each of the Stroop trial 1, 2 and 3 task data are columns from Stroop trial 1: 35, 50, 34, 38, 51, 52, 39, 63, 1, 54, 58, 59, 3, 5, 55, 47, 10, 46, 62, 45, 60, 48, 6, 7, 19, 13, 18, 24, 14, 36, 15, 30; Stroop trial 2: 52, 36, 57, 62, 59, 46, 45, 60, 50, 61, 58, 38, 51, 49, 41, 33, 1, 8, 34, 42, 18, 13, 3, 11, 21, 35, 29, 23, 24, 12, 7, 30; Stroop trial 3: 40, 60, 58, 55, 36, 41, 59, 38, 45, 62, 3, 1, 6, 63, 42, 7, 48, 52, 35, 53, 17, 18, 46, 16, 14, 51, 26, 29, 30, 10, 2, 32. Figure 3 presents a comparison of the accuracy of the stress classification process using the combined PCA+ICA features and the reduced PCA+ICA features. Stroop 1 data, the accuracy of the naive Bayes and SVM methods remains relatively stable before and after feature reduction, while the ensemble method experiences a slight increase in accuracy after feature reduction. Stroop 2 data, the naive Bayes method shows an increase in accuracy after feature reduction, indicating that feature reduction is beneficial for improving the accuracy of stress classification using naive Bayes. Additionally, all classification methods show improved accuracy after feature reduction in Stroop 2 data.

Table 8 presents the classification accuracy results of the best SVM model using different feature extraction and reduction methods across three trials. The methods compared include PCA, ICA, a hybrid PCA+ICA approach, and dimensionality reduction with 16 and 32 columns based on PCA+ICA. In trial 1, the accuracy values obtained with PCA (0.753), ICA (0.772), and PCA+ICA (0.772) are relatively close, while the 16-column reduction yields a slightly lower accuracy (0.724). However, reducing to 32 columns provides a comparable accuracy (0.753) to PCA alone. In trial 2, the results are consistently higher across all methods, with PCA+ICA (0.898) slightly outperforming PCA (0.896) and ICA (0.898). The 16-column reduction achieves 0.893, while the 32-column reduction yields the highest performance (0.908). This indicates that moderate dimensionality reduction (32 columns) can preserve or even enhance classification performance. In trial 3, ICA and PCA+ICA again provide the same accuracy (0.753), slightly higher than PCA alone (0.730). The 16-column reduction gives 0.743, while the 32-column reduction decreases slightly to 0.729. Overall, the results suggest that the PCA+ICA hybrid approach consistently achieves equal or better performance compared to PCA or ICA individually. Furthermore, dimensionality reduction to 32 columns shows potential for improving classification accuracy in trial 2, although results vary across trials. This highlights that the choice of feature reduction dimension plays an important role in balancing accuracy and efficiency in EEG-based stress classification.

Compared to previous research presented in Table 7, the proposed method combining PCA and ICA with 32 features achieves a competitive classification accuracy of 90.8% using the naive Bayes and SVM classifiers. While EEGNet reported the highest accuracy (92.5%) in [23], it relies on deep convolutional neural networks, which are more computationally intensive and less interpretable. In contrast, our approach maintains a relatively simple architecture with lower complexity while still achieving comparable performance. Furthermore, traditional feature-based methods in [23] using band power (BP), Hjorth parameters (HP), or Higuchi fractal dimension (HFD) yielded lower accuracies ranging from 86.5 to 89.8%. This demonstrates that the hybrid PCA+ICA feature representation offers a more compact yet discriminative structure for stress classification, especially when combined with lightweight classifiers such as naive Bayes. To evaluate the relative performance of the proposed method, Table 8 presents a comparison of classification accuracy against selected previous research that utilized various feature extraction and classification techniques in EEG-based stress detection.

Table 8. Comparison of the accuracy of research proposals

Method	Feature	PCA	ICA	PCA+ICA	16 column reduction (PCA+ICA)	32 column reduction (PCA+ICA)
SVM	Trial 1	0.753	0.772	0.772	0.724	0.753
	Trial 2	0.896	0.898	0.898	0.893	0.908
	Trial 3	0.730	0.753	0.753	0.743	0.729

4. CONCLUSION

This research identifies stress EEG data using feature extraction from combined PCA+ICA reduction with Ensemble, naive Bayes, and SVM methods. The experimental results can be concluded that feature reduction using PCA+ICA from 64 to 32 columns only results in increased accuracy in Stroop trial 2. In addition, the method that experienced increased performance due to this feature reduction was naive Bayes, while other methods did not show significant improvements. This shows that dimensionality reduction does not always improve overall performance, but in certain conditions, it can provide benefits for certain methods, such as naive Bayes. Future research could expand this research in several directions. One approach is the integration of multimodal data sources, such as combining EEG with physiological signals like heart rate variability (HRV), to achieve a more comprehensive understanding of stress responses. Another important direction is the application of artificial intelligence techniques to mobile platforms to improve the transparency and interpretability of stress prediction models, which are invaluable for healthcare and clinical decision-making.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Setyorini	✓	✓	✓	✓	✓				✓	✓	✓			✓
Ilham Ari Elbaith Zaeni	✓			✓	✓		✓	✓		✓		✓		
Hakkun Elmunsyah	✓			✓	✓	✓				✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

All author declares no conflicts of interest related to this research and dissertation.

DATA AVAILABILITY

Data are available from the research [36], at https://figshare.com/articles/dataset/SAM_40_Dataset_of_40_Subject_EEG_Recordings_to_Monitor_the_Induced-Stress_while_performing_Stroop_Color-Word_Test_Arithmetic_Task_and_Mirror_Image_Recognition_Task/14562090.

REFERENCES

[1] K. Boby and S. Veerasingam, "Depression diagnosis: EEG-based cognitive biomarkers and machine learning," *Behavioural Brain Research*, vol. 478, 2025, doi: 10.1016/j.bbr.2024.115325.

[2] A. T. Hermawan, I. A. E. Zaeni, A. P. Wibawa, Gunawan, N. Hartono, and Y. Kristian, "EEG-based lie detection using autoencoder deep learning with Muse II brain sensing," *International Journal of Robotics and Control Systems*, vol. 4, no. 3, pp. 1403–1428, 2024, doi: 10.31763/ijrcs.v4i3.1497.

[3] Y. Badr, U. Tariq, F. Al-Shargie, F. Babiloni, F. Al Mughairbi, and H. Al-Nashash, "A review on evaluating mental stress by deep learning using EEG signals," *Neural Computing and Applications*, vol. 36, no. 21, pp. 12629–12654, 2024, doi: 10.1007/s00521-024-09809-5.

[4] A. S. Chaudhari and H. Shrivastava, "Comprehensive review on stress detection using EEG signals and machine learning techniques," *Journal of Electrical Systems*, vol. 20, no. 11, pp. 3310–3327, 2024, doi: 10.52783/jes.8087.

[5] Pawan and R. Dhiman, "Machine learning techniques for electroencephalogram based brain-computer interface: a systematic literature review," *Measurement: Sensors*, vol. 28, 2023, doi: 10.1016/j.measen.2023.100823.

[6] Pooja, S. Pahuja, and K. Veer, "Recent approaches on classification and feature extraction of EEG signal: a review," *Robotica*, vol. 40, no. 1, pp. 77–101, 2022, doi: 10.1017/S0263574721000382.

[7] X. Zhang, L. Yao, X. Wang, J. Monaghan, D. Mcalpine, and Y. Zhang, "A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers," *Journal of Neural Engineering*, vol. 18, no. 3, 2021, doi: 10.1088/1741-2552/abc902.

[8] 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, doi: 10.1109/ACCESS.2021.3085502.

[9] P. Samal and R. Singla, "EEG based stress level detection during gameplay," *2021 2nd Global Conference for Advancement in Technology, GCAT 2021*, 2021, doi: 10.1109/GCAT52182.2021.9587468.

[10] R. K. Rai and D. K. Singh, "Stress detection through wearable EEG technology: a signal-based approach," *Computers and Electrical Engineering*, vol. 126, 2025, doi: 10.1016/j.compeleceng.2025.110478.

[11] S. Pichandi, G. Balasubramanian, and V. Chakrapani, "Hybrid deep models for parallel feature extraction and enhanced emotion state classification," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-75850-y.

[12] K. K. Ng and D. Y. Y. Sim, "Effects of PCA-enabled machine learning classification of stress and resting state EEGs," in *International Conference on Brain Inspired Cognitive Systems*, 2023, pp. 282–290, doi: 10.1007/978-981-97-1417-9_26.

[13] S. M. U. Saeed, S. M. Anwar, H. Khalid, M. Majid, and U. Bagci, "EEG based classification of long-term stress using psychological labeling," *Sensors (Switzerland)*, vol. 20, no. 7, 2020, doi: 10.3390/s20071886.

[14] P. Kaushik *et al.*, "Modelling radiological features fusion and explainable AI in pneumonia detection: a graph- based deep learning and transformer approach," *Results in Engineering*, vol. 26, 2025, doi: 10.1016/j.rineng.2025.105225.

[15] L. Malviya and S. Mal, "A novel technique for stress detection from EEG signal using hybrid deep learning model," *Neural Computing and Applications*, vol. 34, no. 22, pp. 19819–19830, 2022, doi: 10.1007/s00521-022-07540-7.

[16] B. Roy *et al.*, "Hybrid deep learning approach for stress detection using decomposed EEG signals," *Diagnostics*, vol. 13, no. 11, 2023, doi: 10.3390/diagnostics13111936.

[17] D. Risqiwati *et al.*, "Constructing mamdani-intuitionistic fuzzy rules set to detect the relaxed state by transforming spatio-temporal EEG data," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 6, pp. 583–596, 2024, doi: 10.22266/ijies2024.1231.45.

[18] L. Yang, R. Zhou, G. Li, Y. Yang, and Q. Zhao, "Recognizing and explaining driving stress using a Shapley additive explanation model by fusing EEG and behavior signals," *Accident Analysis and Prevention*, vol. 209, 2025, doi: 10.1016/j.aap.2024.107835.

[19] V. S. Babu and A. Ramakrishna, "A multimodal approach harnessing EEG and ECG signals for advanced sleep stage classification," *International Journal of Intelligent Engineering and Systems*, vol. 18, no. 1, pp. 68–80, 2025, doi: 10.22266/ijies2025.0229.06.

[20] K. Aqeel and K. Ali, "An efficient EEG based emotion classification using optimized supported vector machine based on dynamic GOA algorithm," *International Journal of Intelligent Engineering and Systems*, vol. 18, no. 2, pp. 406–416, 2025, doi: 10.22266/IJIES2025.0331.30.

[21] A. Hemakom, D. Atiwiwat, and P. Israsena, "ECG and EEG based machine learning models for the classification of mental workload and stress levels for women in different menstrual phases, men, and mixed sexes," *Biomedical Signal Processing and Control*, vol. 95, 2024, doi: 10.1016/j.bspc.2024.106379.

[22] T. S. Gill, S. S. H. Zaidi, and M. A. Shirazi, "Attention-based deep convolutional neural network for classification of generalized and focal epileptic seizures," *Epilepsy and Behavior*, vol. 155, 2024, doi: 10.1016/j.yebeh.2024.109732.

[23] E. A. A.-Ghaffar and M. Salama, "The effect of stress on a personal identification system based on electroencephalographic signals," *Sensors*, vol. 24, no. 13, 2024, doi: 10.3390/s24134167.

[24] M. Y. T. Sulistyono, E. S. Pane, E. M. Yuniarno, and M. H. Purnono, "Hybrid significant stroke feature: a novel stroke feature analysis approach for stroke severity classification of EEG signals based on time domain, frequency domain, and signal decomposition domain," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 6, pp. 1241–1267, 2024, doi: 10.22266/ijies2024.1231.91.

[25] M. A. Hafeez and S. Shakil, "EEG-based stress identification and classification using deep learning," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 42703–42719, 2024, doi: 10.1007/s11042-023-17111-0.

[26] A. S. S. Prasath, S. Lokesh, N. J. Krishnakumar, T. Vandarkuzhali, D. N. Sahu, and P. C. S. Reddy, "Classification of EEG signals using machine learning and deep learning techniques," *International Journal of Health Sciences*, pp. 10794–10807, 2022, doi: 10.53730/ijhs.v6ns1.7595.

[27] A. Sawan, M. Awad, and R. Qasrawi, "Machine learning-based approach for stroke classification using electroencephalogram (EEG) signals," pp. 111–117, 2022, doi: 10.5220/0010774200003123.

[28] M. Gupta and S. Vaikole, "Recognition of human mental stress using machine learning paradigms," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3571754.

[29] N. Phutela, D. Relan, G. Gabrani, P. Kumaraguru, and M. Samuel, "Stress classification using brain signals based on LSTM network," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/7607592.

[30] S. A. Abdulrahman, M. Roushdy, and A. B. M. Salem, "Using k-nearest neighbors and support vector machine classifiers in personal identification based on EEG signals," *International Journal of Computer Science and Information Security*, vol. 18, no. 5, pp. 29–37, 2020.

[31] H. Elmunsyah, R. Mu'awanah, T. Widiyaningtyas, I. A. E. Zaeni, and F. A. Dwiyanto, "Classification of employee mental health disorder treatment with k-nearest neighbor algorithm," in *2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, 2019, vol. 6, pp. 211–215. doi: 10.1109/ICEEIE47180.2019.8981418.
[32] M. Grobbelaar *et al.*, "A survey on denoising techniques of electroencephalogram signals using wavelet transform," *Signals*, vol. 3, no. 3, pp. 577–586, 2022, doi: 10.3390/signals3030035.
[33] R. Ghosh, S. Phadikar, N. Deb, N. Sinha, P. Das, and E. Ghaderpour, "Automatic eyeblink and muscular artifact detection and removal from EEG signals using k-nearest neighbor classifier and long short-term memory networks," *IEEE Sensors Journal*, vol. 23, no. 5, pp. 5422–5436, 2023, doi: 10.1109/JSEN.2023.3237383.
[34] S. Phadikar, N. Sinha, R. Ghosh, and E. Ghaderpour, "Automatic muscle artifacts identification and removal from single-channel EEG using wavelet transform with meta-heuristically optimized non-local means filter," *Sensors*, vol. 22, no. 8, 2022, doi: 10.3390/s22082948.
[35] Y. Li *et al.*, "Electromyogram (EMG) removal by adding sources of EMG (ERASE)—a novel ICA-based algorithm for removing myoelectric artifacts from EEG," *Frontiers in Neuroscience*, vol. 14, 2021, doi: 10.3389/fnins.2020.597941.
[36] R. Ghosh *et al.*, "SAM 40: dataset of 40 subject EEG recordings to monitor the induced-stress while performing Stroop color-word test, arithmetic task, and mirror image recognition task," *Data in Brief*, vol. 40, 2022, doi: 10.1016/j.dib.2021.107772.
[37] Purnawansyah, A. Adnan, H. Darwis, A. P. Wibawa, T. Widyaningtyas, and Haviluddin, "Ensemble semi-supervised learning in facial expression recognition," *International Journal of Advances in Intelligent Informatics*, vol. 11, no. 1, pp. 1–24, 2025, doi: 10.26555/ijain.v11i1.1880.
[38] T. Widiyaningtyas, I. A. Elbaith Zaeni, and R. Al Farisi, "Sentiment analysis of hotel review using N-gram and naive Bayes methods," in *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 2019, pp. 1–5. doi: 10.1109/ICIC47613.2019.8985946.
[39] M. F. A. Saputra, T. Widiyaningtyas, and A. P. Wibawa, "Illiteracy classification using k means-naive Bayes algorithm," *International Journal on Informatics Visualization*, vol. 2, no. 3, pp. 153–158, 2018, doi: 10.30630/joiv.2.3.129.
[40] A. Hag *et al.*, "Enhancing EEG-based mental stress state recognition using an improved hybrid feature selection algorithm," *Sensors*, vol. 21, no. 24, 2021, doi: 10.3390/s21248370.

BIOGRAPHIES OF AUTHORS

Setyorini is an alumnus of S1 Informatics Engineering, University of Muhammadiyah Malang, class of 2006. After graduating from S1, the author continued his S2 study of Information Systems Management at Merdeka University Malang in 2011. The author is an active lecturer teaching at the Asian Institute of Technology and Business Malang and serves as the Head of the Career Center of the Asia Malang Institute. This Software Design and Engineering Book is the 5th Book. Some of the books previously written are about the Basic Concepts of Management Information Systems, Introduction to Accounting Information Systems, 3 in 1 Vector and Bitmap, and Books from research Design and Build InfoKu GPMB News Portal Malang Regency. In addition, the author is also actively writing articles in several national print media and has created several research software programs recorded in Intellectual Property Rights (IPR). She can be contacted at email: setyorini.2205349@students.um.ac.id.

Ilham Ari Elbaith Zaeni was born in Lamongan, Indonesia, on June 26, 1981. He received the bachelor's degree in Electrical Engineering from Brawijaya University, Indonesia, in 2005 and the master's degree in Electrical Engineering from Brawijaya University, Indonesia, in 2012. In 2017, he received a Ph.D. degree in Electrical Engineering from Southern Taiwan University of Science and Technology, Taiwan. He has been working as a lecturer in the Department of Electrical Engineering, Universitas Negeri Malang since 2006. His research interests include biomedical electronics, signal processing, machine learning, and embedded systems. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) Indonesia Section. He can be contacted at email: ilham.ari.ft@um.ac.id.

Hakkun Elmunsyah is a member of the National Association of Vocational Teachers and Lecturers from 2009 until now. He often works as a reviewer for international conferences, including the International Conference on Electrical, Electronics, and Information Engineering (ICEEIE), International Conference on Vocational Education and Electrical Engineering (ICVEE), and Advanced Technology, Applied Science, and Engineering Conference (ATASEC). He also works as a senior lecturer at the Department of Electrical and Informatics Engineering, Universitas Negeri Malang, East Java, Indonesia. Apart from that, he is the Editor-in-Chief of the journals Journal of Educational Science (JIP) and Letters in Information Technology Education (LITE), which are national journals indexed by Sinta level 2 and 4. He can be contacted at email: hakkun@um.ac.id.