SCIENCE AND INFORMATION TECHNOLOGY
SINTECH
JOURNAL

SINTECH JOURNAL | ISSN 2598-7305 | E-ISSN 2598-9642
Vol. 8 No 1 – April 2025 | https://s.id/sintechjournal
DOI: https://doi.org/10.31598
Publishing: Prahasta Publisher

EEG-Based Mental State Classification Using Statistical Features and Convolutional Neural Networks

Gst. Ayu Vida Mastrika Giri[1], Cokorda Pramartha[2], I Ketut Gede Suhartana[3]

[1, 2, 3]Informatics Study Program, Faculty of Mathematics and Natural Sciences, Udayana University
Jl. Kampus Unud Jimbaran, Badung, Indonesia

e-mail: email vida@unud.ac.id[1], cokorda@unud.ac.id[2], ikg.suhartana@unud.ac.id

Received : February, 2025	Accepted : April, 2025	Published : April, 2025

Abstract

The current research examines the effectiveness of Convolutional Neural Networks (CNNs) in classifying mental states from EEG data through statistical features. Conventional EEG classification frequently depends on raw signals or picture transformations, which may insufficiently utilize feature-level differentiations. This research presents a direct statistical feature-based CNN methodology to enhance classification accuracy while minimizing computational complexity. Electroencephalogram data were obtained from 10 subjects utilizing the Muse Headband, resulting in 200 samples (100 in each category). To reduce training instability resulting from the constrained dataset, augmentation approaches expanded the sample size to 1000. The optimized CNN model (32 units, learning rate 0.01, 90 epochs) attained 98% accuracy using enhanced data, illustrating the effectiveness of augmentation in stabilizing training. Comparative analysis with prior studies validates that CNNs can effectively categorize statistical EEG data; however, subsequent research should investigate hybrid architectures (CNN-GRU, Transformers) to enhance temporal modeling. These findings enhance EEG-based mental state classification by demonstrating the application of statistical features in deep learning models and highlighting the significance of dataset size and augmentation.

Keywords: *convolutional neural network, classification, electroencephalogram, mental state, statistical data*

Abstrak

Penelitian saat ini meneliti efektivitas Convolutional Neural Networks (CNN) dalam mengklasifikasikan kondisi mental dari data EEG pada fitur-fitur statistik. Klasifikasi EEG konvensional sering kali bergantung pada sinyal mentah atau konversi gambar, yang mungkin tidak cukup memanfaatkan perbedaan pada tingkat fitur. Penelitian ini menyajikan metodologi CNN berbasis fitur statistik secara langsung untuk meningkatkan akurasi klasifikasi sekaligus meminimalkan kompleksitas komputasi. Data electroencephalogram diperoleh dari 10 subjek dengan menggunakan Muse Headband, menghasilkan 200 sampel (100 sampel di setiap kategori). Agar dapat mengurangi ketidakstabilan pelatihan yang diakibatkan oleh dataset yang terbatas, pendekatan augmentasi memperluas ukuran sampel menjadi 1000. Model CNN yang dioptimalkan (32 unit, laju pembelajaran 0,01, 90 epoch) mencapai akurasi 98% dengan menggunakan data yang diperbanyak, yang menggambarkan keefektifan augmentasi dalam menstabilkan pelatihan. Analisis komparatif dengan penelitian sebelumnya memvalidasi bahwa CNN dapat secara efektif mengkategorikan data statistik EEG; namun, penelitian selanjutnya harus menyelidiki arsitektur hibrida (CNN-GRU, Transformers) untuk meningkatkan pemodelan temporal. Hasil penelitian *ini meningkatkan klasifikasi kondisi mental berbasis EEG dengan mendemonstrasikan penerapan fitur statistik dalam model pembelajaran yang mendalam dan menegaskan pentingnya ukuran dan penambahan dataset.*

Kata kunci: convolutional neural network, klasifikasi, electroencephalogram, kondisi mental, data statistik

1. INTRODUCTION

Electroencephalogram (EEG) technology to classify mental states has gained popularity in brain-computer interaction, cognitive monitoring, and mental health assessments. EEG is a non-invasive technique that detects brain activity with high temporal resolution, making it ideal for identifying specific mental states like concentration and relaxation. Portable and reasonably priced EEG equipment like the Muse Headband has increased research possibilities even more. The Muse Headband employs four electrodes (TP9, AF7, AF8, and TP10) to monitor brain activity [1], offering a practical and accessible tool for mental state.

The nonlinear and non-stationary nature of EEG signals, which frequently results in variability in performance across various individuals and recording conditions, presents significant challenges to EEG-based mental state classification despite its potential. Many research studies have used convolutional neural networks (CNNs) for EEG classification; however, the majority depend on raw signal processing or EEG-to-image conversions [2], which may inadequately capture feature-level distinctions.

Traditional machine learning approaches, such as Support Vector Machines (SVM) and Random Forest, have been widely used but often require extensive feature engineering to extract meaningful representations from EEG signals [3]. The efficacy of various machine learning and deep learning models was assessed in a study conducted by Giri and Radhitya (2024), a classification of emotions from EEG data. The results indicated that the Support Vector Machine (SVM) and Convolutional Neural Network (CNN) models outperformed other models, achieving the maximum accuracy of 0.98 [4].

Deep learning models, particularly CNNs, have shown improved classification accuracy by automatically extracting features from EEG signals. Previous research has shown that deep learning models can accurately classify EEG signals across a wide range of mental states and cognitive tasks, with some utilizing spectrogram-based transformations to represent EEG signals as images. Kumari et al. offered a CNN model that achieved the maximum accuracy of 98.13%, followed by a Random Forest model at 98.12% and equivalent results from LSTM and GRU models at 97.42% and 97.19%, respectively [3]. Craik et al. (2019). discovered that CNN, RNN, and DBN architectures outperformed simpler MLP and SAE models in several EEG classification tasks, including emotion detection, motor imagery, and mental effort [2].

However, many of these research employ hybrid models such as CNN-LSTM [5], [6], [7], [8], [9] or CNN-GRU [10], which incorporate recurrent layers to capture temporal dependencies to improve classification accuracy, as traditional machine learning approaches have struggled to address these complexities. While hybrid CNN-RNN models improve classification, they require significant computational resources.

Recent developments in deep learning have introduced alternative architectures, like Transformers [11], [12] and Graph Neural Networks (GNNs) [13], [14], for processing EEG signals. Transformer-based models sort EEG data into classes by using attention processes to get and combine different types of information from the signals, whereas GNNs are used to classify EEG data by modeling the relationships between EEG channels as a graph, where nodes represent EEG channels and edges represent the connections or interactions between them. These models generally need large datasets for efficient training and are computationally intensive, making them less appropriate for small-scale EEG research. Considering these limitations, CNNs continue to be an appropriate choice for EEG classification, particularly where computational efficiency is essential.

Another study focused on CNN designs. The study combines Common Spatial Pattern (CSP) for feature extraction, Fast Fourier Transform

Energy Map (FFTEM) for feature selection and a simple CNN with a single convolutional layer and Fourier feature maps to achieve a 0.61 mean kappa value [15], and another by Lawhern (2018) used two blocks of convolutional network called EEGNet to achieve the range of 0.79 to 0.91 AUC values [16]. These two pieces of research highlight the challenge of improving classification performance with basic CNN structures.

This research investigates whether statistical EEG features can enhance CNN-based classification while maintaining computational efficiency by directly utilizing statistical EEG features instead of raw signals or spectrogram-based representations. Unlike prior approaches, which often incorporate temporal modeling through RNNs or require complex preprocessing, this research explores the feasibility of a simpler CNN-based model that classifies EEG data based solely on statistical features.

Three main aspects define this research's uniqueness: first, it evaluates CNN performance using statistical EEG features instead of transformed images or raw signals; second, it methodically investigates the impact of data augmentation on model stability and generalization; and third, it offers ideas on the possible advantages of hybrid CNN-based architectures for future developments.

Electroencephalogram (EEG) data was collected from ten subjects using the Muse Headband to validate the suggested methodology, which captures activity from four electrodes (TP9, AF7, AF8, TP10). Participants were chosen based on self-reported mental state, aged 20 to 35 years, and without a history of neurological problems. The dataset originally had 200 samples, which was increased to 1000 samples through data augmentation methods. This improvement was essential to prevent overfitting and improve training stability. The research evaluates the effectiveness of CNN models trained on both actual and augmented data, focusing the impact of dataset size on classification performance.

This research enhances the understanding of CNN-based EEG classification by addressing these gaps, particularly in the use of statistical EEG features. The results provide useful insights into optimizing CNN architectures for EEG classification while highlighting the significance of dataset augmentation in improving model stability and generalization. Future research may investigate hybrid models to improve classification accuracy by capturing temporal dependencies in EEG signals.

2. RESEARCH METHOD

2.1 EEG Data Acquisition and Preprocessing

Data acquisition was conducted using the Muse Headband, which includes four electrodes (TP9, AF7, AF8, TP10) placed on the frontal and temporal lobes, as shown in Figure 1. The Muse EEG Headband captured EEG data from 10 participants, ranging in age from 20 to 35 (5 males and 5 females). A total of 200 samples were obtained from each participant's 60-second sessions of two mental state tasks: relaxed and concentrated. To preserve important brain activity, the unprocessed EEG signals were preprocessed to eliminate artifacts using a bandpass filter (1-50 Hz) and recorded at a sampling rate of 256 Hz.

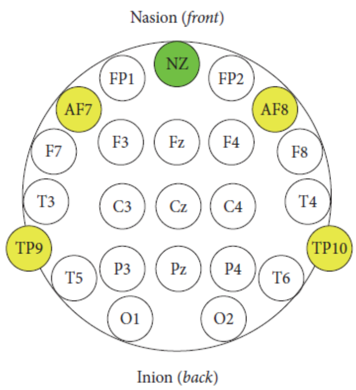

Figure 1. Muse Headband's TP9, AF7, AF8, and TP10 sensors on the international standard EEG layout system [1]

The EEG signals were recorded using the open-source BlueMuse application [17] and stored as raw time-series data. The data was collected in a controlled environment with minimal external disturbances, and participants were asked to perform specific exercises to induce either a concentration state (e.g., reading news articles or attending online classes) or a relaxation state (e.g., deep breathing exercises with eyes closed).

Statistical features were directly extracted from the time and frequency domains rather than converting EEG signals into images. For each of the four electrode channels (TP9, AF7, AF8, TP10), these characteristics are as follows: mean, standard deviation, skewness, kurtosis, power spectral density (PSD), and zero-crossing rate (ZCR). Furthermore, statistical features based on the Fast Fourier Transform (FFT) were calculated to capture frequency-domain characteristics [18].

2.2 EEG Data Augmentation and Normalization

As part of the data augmentation process, three primary techniques were employed to enhance generalization and resolve the limited dataset size: Additive Gaussian Noise, Feature Scaling, and Sample Mixing.

a. Additive Gaussian Noise: To replicate real-world variability while maintaining the underlying signal structure, statistical EEG features were supplemented with random noise with a small variance (0.01).
b. Feature Scaling: To incorporate controlled variations in amplitude while maintaining relative feature distributions, each feature was scaled using a random factor within a predefined range (0.8 to 1.2).
c. Sample Mixing: Synthetic samples were produced by combining two randomly selected EEG feature vectors with a weighted ratio 0.5. This method generated new, realistic samples while maintaining essential statistical properties.

These augmentation strategies increased the dataset from 200 to 1000 samples. Subsequently, the augmented data were normalized using Robust Scaler, which scales features based on the median and interquartile range [19], rendering it more resilient to outliers than min-max scaling.

2.3 CNN Model for Classification

The CNN model was implemented using TensorFlow and Keras. Figure 2 presents the model architecture diagram, providing a clearer understanding of its structure. The CNN model consists of an input layer, two convolutional layers with ReLU activation, and max-pooling layers for feature extraction. The first Conv1D layer applies numbers filters with a kernel size of 3, followed by a max-pooling layer with a pool size of 2. The second Conv1D layer doubles the filters from the first layer while maintaining the same kernel size, followed by another max-pooling layer. After convolutional blocks, the output is flattened and passed through a fully connected Dense layer with 64 units and ReLU activation. A dropout layer with a rate of 0.5 is added to prevent overfitting before the final Dense output layer. This layer uses a softmax activation function to categorize mental states into two classes. The model was trained using the Adam optimizer and sparse categorical cross-entropy loss.

2.4 Hyperparameter Tuning and Model Training

Manual Hyperparameter Tuning was conducted by testing different configurations of convolutional units (32, 64, 128), learning rates (0.001, 0.005, 0.01), and epochs (60, 75, 90). The dataset was split into 80% training and 20% validation. Each combination was tested on actual data (200 samples) and augmented and normalized data (1000 samples) to determine the best-performing configuration.

Following the hyperparameter tuning phase, the model is trained using the optimal set of hyperparameters. Training stability was analyzed using learning curves to assess fluctuations in accuracy and loss over epochs. The dataset was split into 80% training and 20% validation, and training was conducted using a mini-batch size of 32.

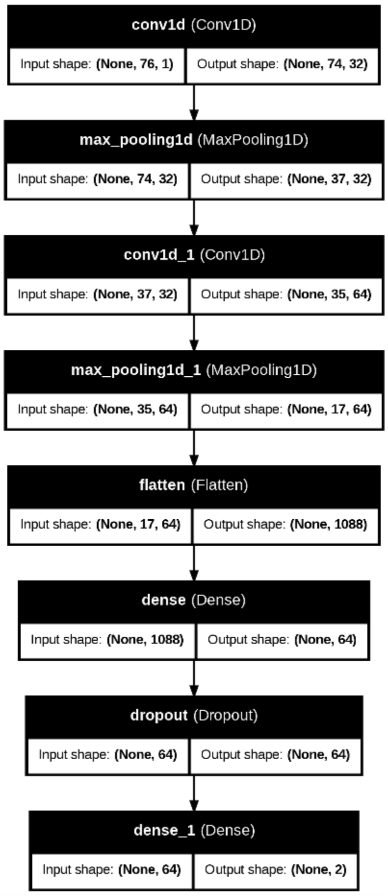

Figure 2. Initial Model Architecture Diagram

2.5 Model Evaluation

The trained CNN model was evaluated using precision, recall, F1-score, and accuracy. Precision assesses the accuracy of positive predictions, whereas recall evaluates the model's ability to identify all instances of each class. The F1-score is the harmonic mean of precision and recall, providing a balanced performance assessment[18].

A confusion matrix visually represents true positives, true negatives, false positives, and false negatives. It was also used to analyze classification performance across the two classes. To further investigate training stability, learning curves were plotted to visualize fluctuations in accuracy and loss over epochs.

Performance was also compared between models trained on the actual dataset (200 samples) and the augmented dataset (1000 samples). Additional comparisons were made with the use of GRU in statistical EEG Data [20].

3. RESULTS AND DISCUSSIONS

3.1 EEG Data Acquisition and Preprocessing Results

The Muse EEG Headband was employed to record the EEG signals of all 10 participants properly. Figure 3 provides sample EEG waveforms for both mental states (concentration and relaxation), visually representing the unprocessed dataset. The raw signals were then processed with bandpass filtering (1-50 Hz) and artifact removal to maintain important brain activity and minimize noise.

Each EEG sample yielded a total of 76 statistical features, including time-domain features (mean, standard deviation, skewness, kurtosis, median, minimum, maximum, peak-to-peak amplitude, mean absolute deviation, power spectral density, zero-crossing rate, and root mean square) and frequency-domain features (FFT mean, FFT standard deviation, FFT skewness, FFT kurtosis, FFT minimum, FFT maximum, and FFT peak-to-peak amplitude) for each of the four electrodes (TP9, AF7, AF8, TP10). These features capture various signal characteristics, ensuring an exhaustive representation of EEG patterns. Subsequently, the extracted features were employed as input for CNN-based classification.

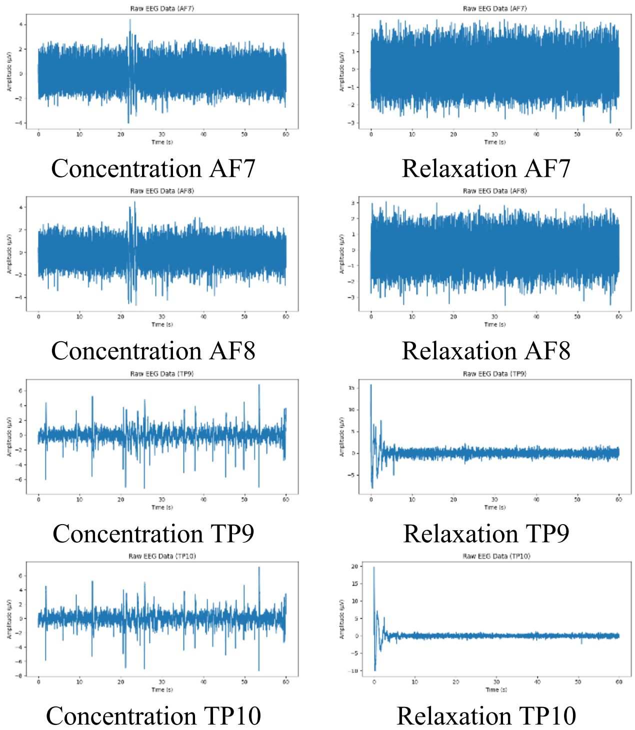

Figure 3. EEG Raw Signal Samples for Concentration and Relaxation States

Boxplots were created using the real dataset to assess extracted features. These boxplots, as shown in Figure 4, show EEG feature distribution across samples. Mean absolute deviation, kurtosis, and peak-to-peak amplitude have a broad interquartile range (IQR), indicating considerable EEG signal variability, while others have a shorter IQR, indicating better stability. Outliers in mean absolute deviation, kurtosis, and peak-to-peak amplitude may indicate brain activity variations or data noise. Skewness indicates long-tailed distributions, while symmetric distributions have a center median.

Features with a lot of variation may add noise, and features whose distributions meet may make classification less accurate. Adding more data to the dataset can help get a more accurate picture of how the EEG patterns are distributed, which can help fix these problems. Normalizing the extracted features also makes sure that all of them add equally to the learning process. This stops dominant features from making the model less accurate.

3.2 EEG Data Augmentation and Normalization Results

Applying augmentation techniques significantly increased the dataset size from 200 to 1000 samples to improve model generalization. Additive Gaussian noise introduced small variations to mimic real-world EEG fluctuations. Feature scaling adjusted amplitudes within a controlled range to account for signal intensity variations across participants. Sample mixing created new synthetic samples by blending real EEG feature vectors, preserving underlying statistical relationships.

Normalization using Robust Scaler has also been implemented. This ensures consistent feature distributions leading to more stable training.

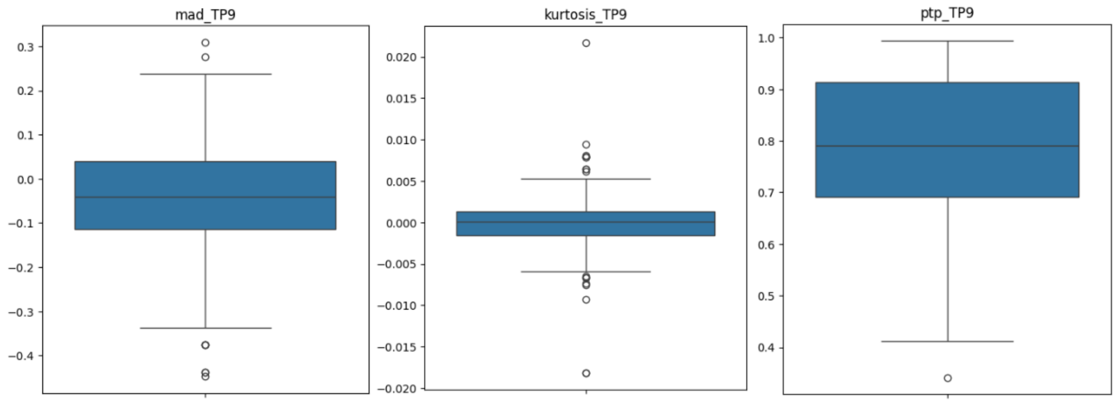

Figure 4. Sample Boxplots of Real Data

In Figure 5, the boxplots show how the normalized features are spread out. There are big changes in feature consistency between these plots and those from the original dataset. In the original dataset, features like mean absolute deviation, kurtosis, and peak-to-peak amplitude had various interquartile range and outliers.

The augmentation and normalization method centers the feature distributions, which makes it easier to compare EEG recordings that were recorded at different times. This transformation makes sure that all features add equally to classification.

It's interesting that the augmentation normalization process has found more outliers in certain features while also making them more consistent. This might be because of the rescaling and augmenting effect, which makes variations that were not so noticeable before stand out more. The fact that there are more outliers indicates that normalization makes extreme values easier to see.

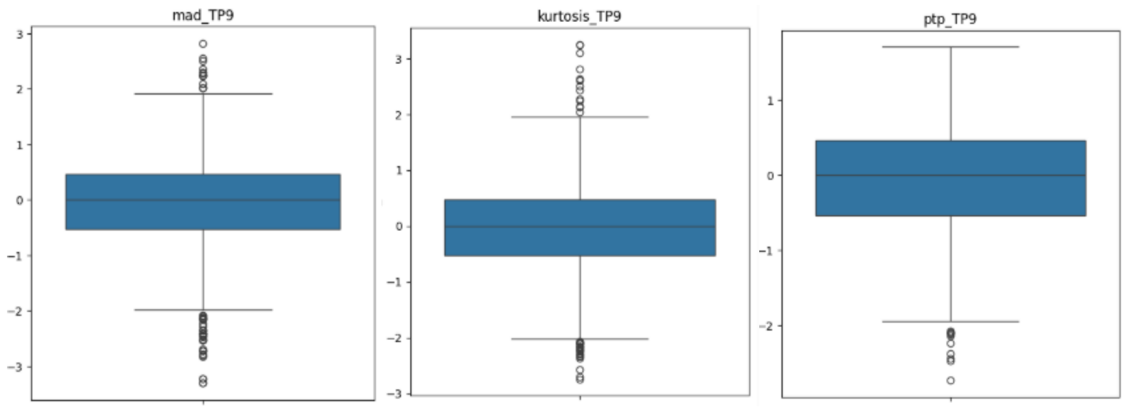

Figure 5. Sample Boxlplots of Augmented and Normalized Data

3.3 CNN Model Implementation

The model developed in the research methodology has been successfully implemented to classify statistical EEG data into two mental states: concentration and relaxation. The CNN model was implemented using TensorFlow and Keras, following a structured architecture. The model consists of two Conv1D layers with ReLU activation, followed by max-pooling for dimensionality reduction. A Flatten layer transforms extracted features into a dense representation, which is then processed by a fully connected Dense layer with 64 units and a Dropout layer (0.5) to prevent overfitting. Finally, a softmax output layer classifies mental states into two categories. The model was compiled using the Adam optimizer and trained using sparse categorical cross-entropy loss.

3.4 Hyperparameter Tuning and Model Training Results

The CNN model went through to an exhaustive hyperparameter tuning process in order to identify the optimal configuration in this research. The learning rate, the number of epochs for training, and the number of units in convolutional layers were all adjusted. The most effective model configuration was determined through systematic experimentation, utilizing an augmented and normalized dataset of 1000 samples. The optimal hyperparameters were 32 CNN units in first convolutional layer, 64 in the second convolutional layer, a learning rate of 0.01, and 90 training epochs.

The results of the model training show the substantial differences between the use of actual data and augmented data. The model demonstrated instability when trained on a smaller actual dataset (200 samples), as evidenced by the fluctuating validation accuracy, which suggests poor generalization. The model was unable to achieve the optimal validation accuracy, despite the convergence of training and validation loss.

The training and validation loss and accuracy plots for the actual (real) dataset of 200 samples, which are shown in Figure 6, showing apparent signs of instability. The validation accuracy fluctuates considerably, whereas the training accuracy increases over epochs, suggesting that the model experiences difficulty generalizing to unseen data. This implies that the dataset is insufficiently diverse to enable effective learning, resulting in a significant amount of performance variance.

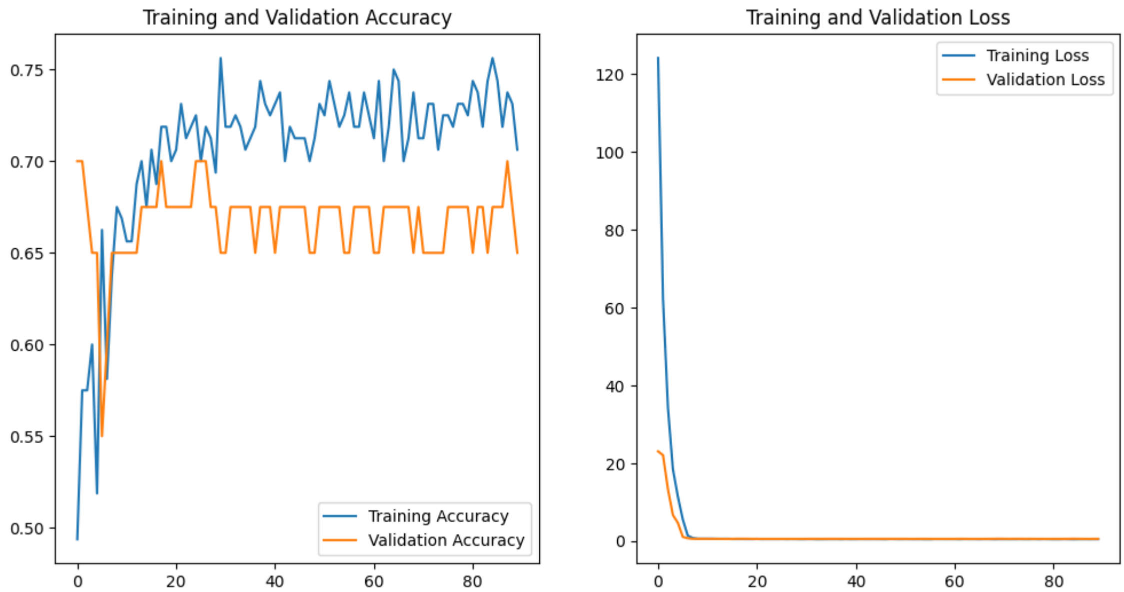

Figure 6. Training and Validation Accuracy and Loss for Real Data (200 Samples)

In contrast, the training and validation accuracy curves were more consistent as a result of training on the augmented and normalized dataset (1000 samples), which is shown in Figure 7. The model consistently exhibited high validation accuracy, having smoother accuracy trends. The augmented dataset produced a rapid convergence of the loss values when examining the training and validation loss. The model was learning effectively, as evidenced by the consistently low validation loss and the rapid decrease in the training loss. These findings suggest that the model's capacity to extract meaningful features was improved by data augmentation and normalization, which also ensured greater generalization across unseen data and improved robustness against noise.

It is obvious from the comparison of these training results that data preprocessing, particularly augmentation and normalization, is essential for CNN performance. The model was able to learn more effectively, which in turn improved reliability and stability, due to the larger, well-prepared dataset. The accompanying training and validation accuracy plots further substantiate these findings, demonstrating the benefits of a more comprehensive dataset.

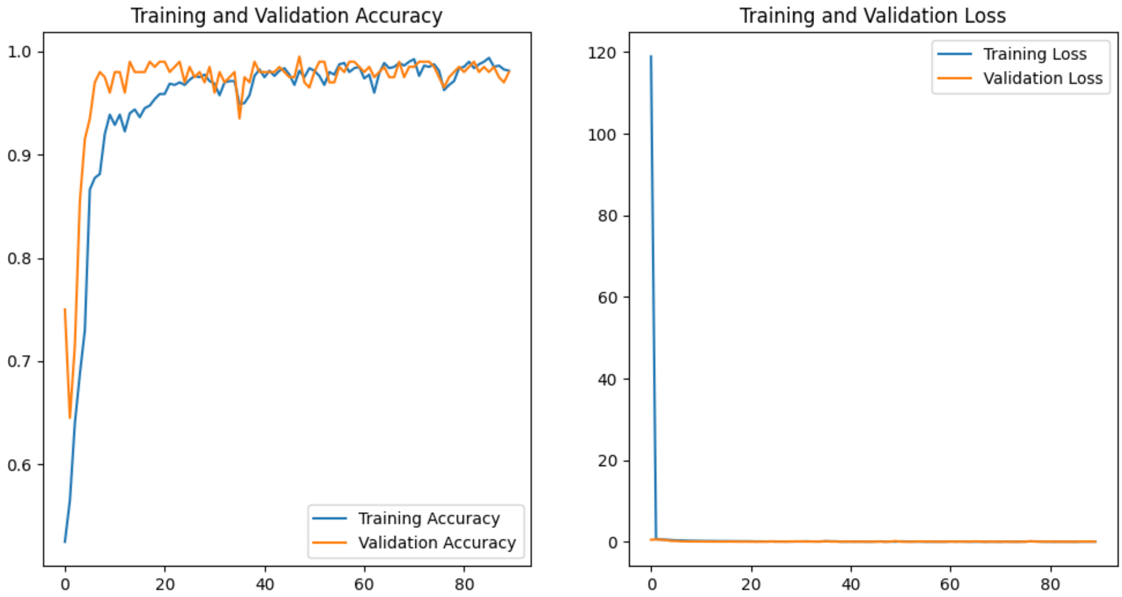

Figure 7. Training and Validation Accuracy and Loss for Augmented and Normalized Data (1000 Samples)

3.5 Model Evaluation Results

The evaluation of the model was conducted using confusion matrices and classification reports, providing a comparative analysis of the model's performance on the actual and augmented datasets. The assessment focuses on accuracy, precision, recall, and F1-score to determine the impact of data augmentation on classification performance. The evaluation was done with the actual data and augmented data.

Using the actual data, the model correctly classified 18 instances of the relaxed state, as seen in the confusion matrix in Figure 8. However, it misclassified 2 instances as concentrated. The model correctly identified 8 Instances in the concentrated state, but it incorrectly classified 12 as relaxed.

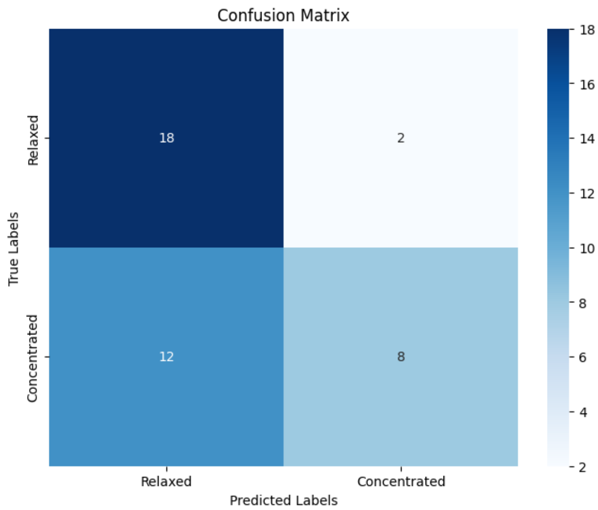

Figure 8. Confusion Matrix of Model Evaluation Using Actual Data

Classification Report:

	precision	recall	f1-score	support
Relaxed	0.60	0.90	0.72	20
Concentrated	0.80	0.40	0.53	20
accuracy			0.65	40
macro avg	0.70	0.65	0.63	40
weighted avg	0.70	0.65	0.63	40

Figure 9. Classification Report of Model Evaluation Using Actual Data

This suggests a substantial imbalance in the model's capacity to differentiate between the two states, with the concentrated state being particularly difficult to distinguish. This observation is further substantiated by the classification report shown in Figure 9, which indicates an accuracy of 65%. The model is unable to accurately identify the majority of concentrated states, as the recall for the concentrated state is only 40%. This implies that either the model is having difficulty learning its distinguishing features or that the dataset may not contain enough variation in the concentrated state.

[20]On the other hand, the augmented dataset showed a significant improvement. The model now accurately classifies 99 instances of relaxed and only misclassifies one. Similarly, the concentrated state has 97 instances that are accurately identified, with only three misclassified as shown in confusion matrix in Figure 10.

The classification report shown in Figure 11 demonstrates this enhanced performance, with an overall accuracy of 98%. The model's capacity to generalize has been considerably improved by data augmentation, as evidenced by the near-perfect precision, recall, and F1-scores of both classes.

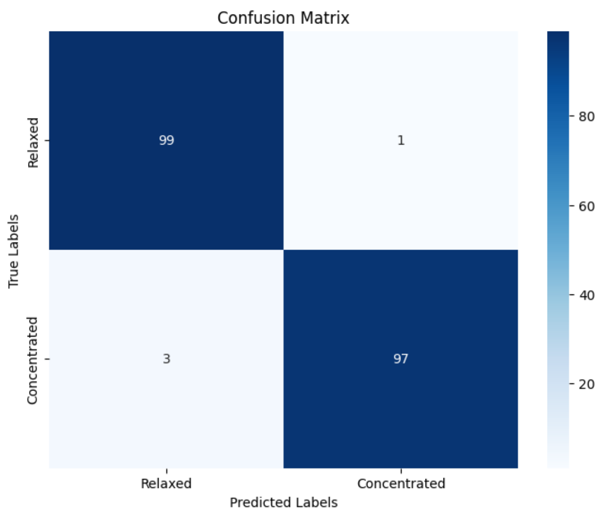

Figure 10. Confusion Matrix of Model Evaluation Using Augmented and Normalized Data

Classification Report:

	precision	recall	f1-score	support
Relaxed	0.97	0.99	0.98	100
Concentrated	0.99	0.97	0.98	100
accuracy			0.98	200
macro avg	0.98	0.98	0.98	200
weighted avg	0.98	0.98	0.98	200

Figure 11. Classification Report of Model Evaluation Using Augmented and Normalized Data

The improvements observed in classification performance can be linked to the approaches taken during the research. The preprocessing approaches, including as feature extraction, and data augmentation, were essential in enhancing the input data for improved learning. The feature extraction phase, which entailed calculating statistical features from both original and FFT-transformed data, guaranteed that the model had access to significant information. The limited dataset resulted in an unstable model, constraining the model's generalization capabilities.

The implementation of data augmentation addressed this problem by expanding the sample size and introducing variability, enabling the model to acquire more stable representations.

In order to put the suggested CNN model's performance in context, it was compared to other EEG classification methods, as shown in Table 1.

Table 1. Result Comparison with Other Research

Researcb	Method	Dataset	Accuracy
This Research	CNN (Statistical Features + Augmentation)	1000 samples (Muse EEG)	98%
Lawhern et al. (2018)	EEGNet (Compact CNN)	Multiple BCI dataset	0.79-0.91 (AUC)
Abbas & Khan (2018)	CNN with FFT Energy Maps	BCI Competition IV	0.61 (Kappa)
Giri & Radhitya (2024)	CNN, LSTM, GRU (Emotion Classification)	Feeling Emotions EEG	98% (CNN), 82% (LSTM), 97% (GRU)
Giri, et al., (2025)	GRU with Statistical Features	400 samples (Muse EEG)	95%

This comparison demonstrates that our CNN-based methodology, employing statistical features and data augmentation, attains accuracy that is either comparable to or exceeds that of other models. Specifically, our model surpasses EEGNet [16] and CNN-FFTEM [15], which utilized alternative feature extraction methods. Moreover, our results align with the emotion classification models proposed by [4][4], indicating the efficacy of statistical EEG features across diverse EEG classification tasks. Additionally, our GRU model in the previous research [20], despite achieving marginally lower accuracy than the CNN, illustrates the potential of recurrent architectures in EEG classification. The findings support the hypothesis that data augmentation significantly enhances model performance and stability.

Although proposed CNN model demonstrated high accuracy, there are some limitations remain in this research; CNNs do not completely understand how EEG data is linked in a sequential way. To improve temporal feature learning, more research should look into CNN-GRU or Transformer-based methods in the future. The dataset size was artificially increased via augmentation. Future studies should include a bigger, more diverse participant group to validate results, and while Transformers and other hybrid methods were mentioned as a potential alternative, they were not directly tested in this research.

4. CONCLUSION

This research examined into how Convolutional Neural Networks (CNNs) can be used to classify mental states based on EEG data using statistical features. Unlike earlier methods that used raw EEG signals or image-based transformations, our method directly used statistical features from EEG signals and used data augmentation to make the model more stable and useful in other situations. The results show that a CNN trained on statistical features and improve with augmentation got 98% accuracy, doing better than several modern methods, such as EEGNet and FFT-based CNN models.

Furthermore, our results show that statistical feature representation is a good alternative to more complicated EEG processing methods, providing a low cost and effective way to classify EEG signals. Compared to other deep learning architectures, like GRUs and hybrid CNN-RNN models, CNNs seem to be very good at classifying objects, especially when they are joined with a lot of extra data.

Even though these results look good, there are still a few challenges. First, the CNN model doesn't have any temporal modeling tools. This problem could be fixed by adding recurrent architectures like GRUs or Transformer-based models. Second, this study only used a small EEG dataset. While augmentation made the results more general, these results should be confirmed in future studies using bigger, more diverse EEG datasets. Lastly, more research into advanced deep learning methods such as attention mechanisms and hybrid models could make EEG classification even more accurate and reliable.

ACKNOWLEDGEMENT

We extend our sincere gratitude to Universitas Udayana, LPPM, and Fakultas MIPA for their support in funding this research through the 2024 PNBP funds.

REFERENCES

[1] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekárt, and C. D. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," Complexity. Accessed: Dec. 13, 2023. [Online]. Available: https://dx.doi.org/10.1155/2019/4316548

[2] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," 2019. doi: 10.1088/1741-2552/ab0ab5.

[3] N. Kumari, S. Anwar, and V. Bhattacharjee, "A Comparative Analysis of Machine and Deep Learning Techniques for EEG Evoked Emotion Classification," Wireless Personal Communications. Accessed: Dec. 13, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s11277-022-10076-7

[4] Gst. A. V. M. Giri and M. L. Radhitya, "Electroencephalogram-Based Emotion Classification Using Machine Learning and Deep Learning Techniques," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems*, vol. 18, no. 3, pp. 1–5, Jun. 2024, doi: 10.22146/ijccs.xxxx.

[5] S. M. Omar, M. Kimwele, A. Olowolayemo, and D. M. Kaburu, "Enhancing EEG signals classification using LSTM-CNN architecture," *Engineering Reports*, vol. 6, no. 9, Sep. 2024, doi: 10.1002/eng2.12827.

[6] N. Jiwani, K. Gupta, M. H. U. Sharif, N. Adhikari, and N. Afreen, "A LSTM-CNN Model for Epileptic Seizures Detection using EEG Signal," in *2022 2nd International Conference on Emerging Smart Technologies and Applications, eSmarTA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/eSmarTA56775.2022.9935403.

[7] C. Lee and J. An, "LSTM-CNN model of drowsiness detection from multiple consciousness states acquired by EEG," *Expert Syst Appl*, vol. 213, Mar. 2023, doi: 10.1016/j.eswa.2022.119032.

[8] B. Chakravarthi, S. C. Ng, M. R. Ezilarasan, and M. F. Leung, "EEG-based emotion recognition using hybrid CNN and LSTM classification," *Front Comput Neurosci*, vol. 16, Oct. 2022, doi: 10.3389/fncom.2022.1019776.

[9] A. Shoeibi *et al.*, "Automatic Diagnosis of Schizophrenia in EEG Signals Using CNN-LSTM Models," *Front Neuroinform*, vol. 15, Nov. 2021, doi: 10.3389/fninf.2021.777977.

[10] S. Gao, J. Yang, T. Shen, and W. Jiang, "A Parallel Feature Fusion Network Combining GRU and CNN for Motor Imagery EEG Decoding," *Brain Sci*, vol. 12, no. 9, 2022, doi: 10.3390/brainsci12091233.

[11] H. J. Ahn, D. H. Lee, J. H. Jeong, and S. W. Lee, "Multiscale Convolutional Transformer for EEG Classification of Mental Imagery in Different Modalities," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, 2023, doi: 10.1109/TNSRE.2022.3229330.

[12] Y. Song, Q. Zheng, B. Liu, and X. Gao, "EEG Conformer: Convolutional Transformer for EEG Decoding and Visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, 2023, doi: 10.1109/TNSRE.2022.3230250.

[13] D. Klepl, M. Wu, and F. He, “Graph Neural Network-Based EEG Classification: A Survey,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, 2024, doi: 10.1109/TNSRE.2024.3355750.
[14] P. Zhong, D. Wang, and C. Miao, “EEG-Based Emotion Recognition Using Regularized Graph Neural Networks,” *IEEE Trans Affect Comput*, vol. 13, no. 3, 2022, doi: 10.1109/TAFFC.2020.2994159.
[15] W. Abbas and N. A. Khan, “DeepMI: Deep Learning for Multiclass Motor Imagery Classification,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2018. doi: 10.1109/EMBC.2018.8512271.
[16] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces,” *J Neural Eng*, vol. 15, no. 5, 2018, doi: 10.1088/1741-2552/aace8c.
[17] J. J. Bird, L. J. Manso, E. P. Ribeiro, A. Ekart, and D. R. Faria, “A Study on Mental State Classification using EEG-based Brain-Machine Interface,” 9th International Conference on Intelligent Systems 2018: Theory, Research and Innovation in Applications, IS 2018 - Proceedings. Accessed: Dec. 13, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/8710576/
[18] G. Ayu, V. Mastrika Giri, N. Agus, S. Er, I. Ketut, and G. Suhartana, “Application of Gated Recurrent Unit in Electroencephalogram (EEG)-Based Mental State Classification,” 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC
[19] A. Khoirunnisa and N. G. Ramadhan, “Improving malaria prediction with ensemble learning and robust scaler: An integrated approach for enhanced accuracy,” *JURNAL INFOTEL*, vol. 15, no. 4, 2023, doi: 10.20895/infotel.v15i4.1056.
[20] G. A. V. Giri, N. A. S. ER, and I. K. G. Suhartana, “Application of Gated Recurrent Unit in Electroencephalogram (EEG)-Based Mental State Classification,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9 No. 1, 2025, [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC