Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD28893

EEG-Based PTSD Classification Using Random Forest, CNN, and Ensemble Learning: A Comparative Study

Shenziyang Zhang[1*], Yitong Wang[1], Yizhou Xue[1]

[1]*Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing, China*
**Corresponding Author. Email: p22000221@njupt.edu.cn*

Abstract. Post-traumatic stress disorder (PTSD) poses significant diagnostic challenges due to reliance on subjective assessment. This study presents an objective EEG-based classification framework leveraging handcrafted spectral features and deep learning representations. We extracted absolute power spectral density and coherence metrics from 19-channel EEG recordings across six canonical frequency bands, yielding 114 initial features. Three modeling strategies were evaluated: a Random Forest (RF) classifier with grid-searched hyperparameters and subsequent feature pruning; a modified AlexNet-based convolutional neural network (CNN) trained on 224×224 EEG images; and a soft-voting ensemble combining RF and CNN outputs. The RF model achieved 83.7% accuracy under 5-fold cross-validation, which improved to 88.4% after pruning to the top 45 features. The CNN attained 89.1% accuracy with stable convergence over 30 epochs. The soft-voting ensemble outperformed individual models, reaching 90.9% average cross-validated accuracy and an AUC of 0.92, with balanced precision and recall. Analysis of ROC curves highlighted sensitivity-specificity trade-offs in low false-positive regions, informing potential clinical threshold selection. These results demonstrate that integrating handcrafted and deep features yields a robust, objective approach for PTSD recognition from EEG. Future work will explore end-to-end raw signal modeling, adaptive ensemble weighting, and prospective clinical validation to enhance diagnostic utility. From a clinical perspective, the proposed framework may serve as a non-invasive and quantitative decision-support tool to aid clinicians in early detection and to reduce the likelihood of PTSD misdiagnosis.

Keywords: PTSD, CNN, EEG, Machine Learning, Feature Extraction

1. Introduction

Post-Traumatic Stress Disorder (PTSD) is a complex psychiatric disorder that may develop in the aftermath of exposure to highly stressful and traumatic events, including extreme violence and life-threatening circumstances [1]. Patients commonly experience debilitating distress, resulting in significant impairment in function and comorbidities. Notably, 20%-30% of the individuals who experience trauma eventually develop PTSD [2], and approximately 6% of the American population endures a lifetime prevalence of PTSD [3]. Traditionally, recognition diagnosis relies on clinical interviews and self-reported symptom scales according to DSM-5 criteria. This subjective assessment poses challenges: symptoms can be underreported or overlap with other disorders, and

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD28893

clinicians must differentiate genuine trauma-related pathology from other stress reactions [4]. Thus, there is a critical need for objective biomarkers and assessment tools to assist in the accurate recognition of PTSD.

Brain-Computer Interface (BCI) technologies offer a promising avenue toward more objective PTSD recognition as they involve the direct measurement and interpretation of brain activity, typically translating neural signals into meaningful outputs with the assistance of artificial intelligence [5]. In clinical research, BCIs can detect neurophysiological markers associated with mental states or disorders. In recent years, advances in non-invasive neuroimaging and machine learning have spurred interest in utilizing BCI modalities such as electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) to identify PTSD-related brain patterns [6]. In this work, we develop an EEG-based PTSD recognition framework that integrates band power features with a Random Forest classifier and a CNN. By evaluating individual classifiers and a soft-voting ensemble on multi-channel EEG data, this study aims to improve diagnostic accuracy and provide a non-invasive, objective tool to assist clinicians in early detection and to reduce the risk of PTSD misdiagnosis.

2. Methodology

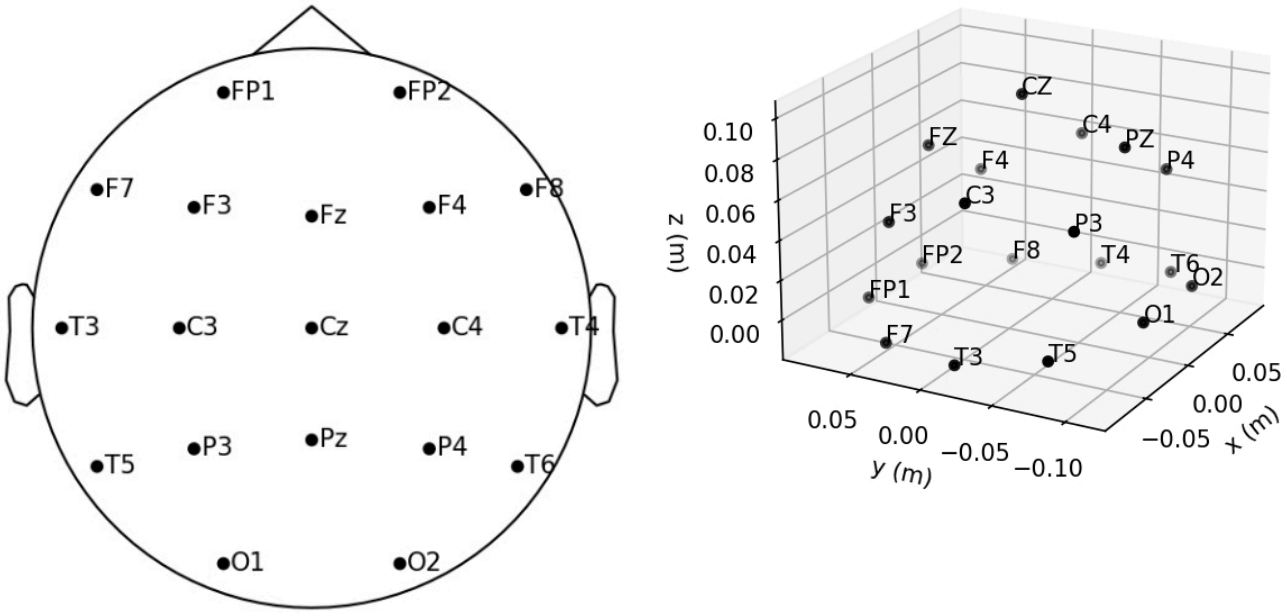
2.1. Data source

We used the EEG dataset available on OSF, which was previously analyzed in Park et al. (2021) [7]. The dataset includes 850 patients from 6 general disorders, including: schizophrenia, mood disorders, anxiety disorders, obsessive-compulsive disorder, addictive disorders, and trauma and stress-related disorders, which are further divided into 9 categories, where 52 PTSD patients lie. The dataset also contains 95 healthy controls without any disorder.

In this dataset, EEG signals were recorded using 19 electrodes placed according to the international 10–20 system, a widely adopted standard for EEG electrode placement in clinical and research settings. The spatial configuration of these electrodes is illustrated in Fig. 1.

The dataset does not contain raw EEG time-series data. Instead, it provides the power spectral density (PSD) features extracted by applying Fast Fourier Transform (FFT) to the original EEG recordings, which are described as absolute power in the dataset. Pairwise coherence (synchronization) metrics were computed across all 19 channels and appended as additional features. For each electrode (channel), the dataset includes spectral power values across six standard frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–25 Hz), high beta (25–30 Hz), and gamma (30–40 Hz). Subjects' gender, ages, and IQ are also provided.

DOI: 10.54254/2755-2721/2025.LD28893



(a) (b)

Figure 1. Layout of the 19 electronodes. (a) 2D topographic layout. (b) 3D spatial coordinates

2.2. Feature extraction

In this study, we used the absolute power values across six frequency bands (delta, theta, alpha, beta, high-beta, and gamma) from the dataset as the main features. For each subject, power values from all 19 EEG channels were concatenated across the six bands, resulting in 114 frequency-domain features.

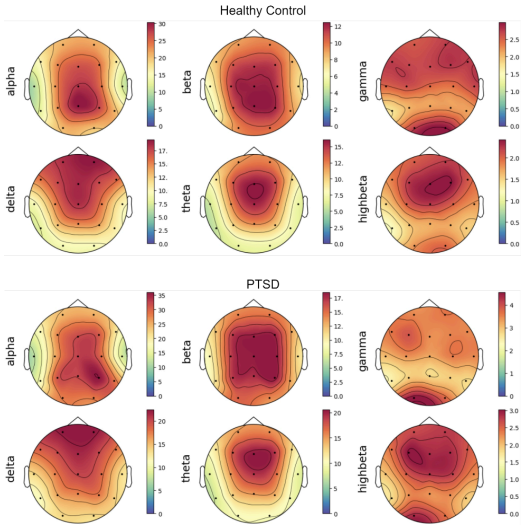


Figure 2. Heatmap of absolute power for healthy controls and PTSD subjects

To visualize group-level differences in spectral power, we computed the average absolute power values across all healthy controls and all PTSD subjects, respectively. Fig. 2 presents a heatmap constructed using these mean values, where the top six graphs correspond to the averaged EEG

DOI: 10.54254/2755-2721/2025.LD28893

features of the healthy control group, and the bottom six graphs represent the averaged features of the PTSD group. All features were then standardized using z-score normalization to ensure comparability across dimensions.

2.3. Classification algorithm

2.3.1. Randon Forest

To classify subjects with and without PTSD based on their EEG-derived features, we employed the Random Forest (RF) algorithm. Random Forest is an ensemble learning method that constructs many decision trees during training and outputs the mode of the classes predicted by individual trees. It is particularly effective for high-dimensional data and can handle non-linear feature interactions. In our experiment, RF was chosen due to its robustness to noise, low risk of overfitting, and ability to provide feature importance measures, which are valuable for interpreting EEG-related biomarkers. The number of trees and other hyperparameters was tuned using grid search with cross-validation on the training set.

We applied 200 estimators and performed a grid search to optimize key hyperparameters, including the maximum tree depth, the minimum number of samples required to split an internal node, and the number of features considered for the best split. The grid search was conducted using 5-fold cross-validation on the training set, and model performance was evaluated using classification accuracy.

2.3.2. Feature pruning

After training the complete RF model on all 114 features, we computed feature importance scores and selected features with importance greater than 1%, resulting in 45 retained features. The RF model was retrained on this reduced feature set under the same cross-validation procedure.

2.3.3. Convolutional neural network

To explore deep learning approaches for EEG-based PTSD recognition, we designed a convolutional neural network (CNN) based on a modified version of the AlexNet architecture [8]. AlexNet was selected for its proven effectiveness in image-based classification tasks and relatively low architectural complexity. It is well-suited for learning spatial patterns from EEG heatmaps with limited dataset sizes. The model accepts a 224×224×1 EEG heatmap as input and consists of five convolutional layers with ReLU activations and max pooling, followed by three fully connected layers ending in a Softmax classifier. The detailed architecture is illustrated in Fig. 3. The network was trained using the Adam optimizer with a learning rate 0.001 for 30 epochs.

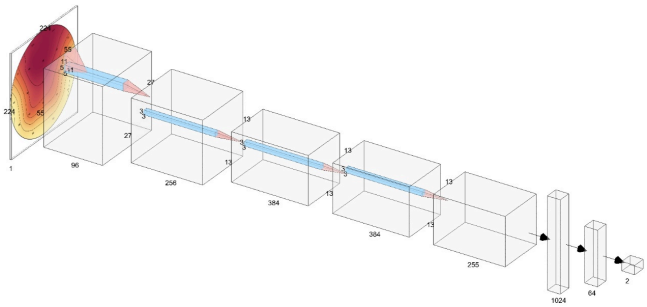


Figure 3. Architecture of the modified AlexNet model used for EEG-based PTSD classification

DOI: 10.54254/2755-2721/2025.LD28893

2.3.4. Soft voting ensemble

We implemented a soft voting ensemble strategy to leverage the complementary strengths of the Random Forest and the modified AlexNet CNN. Let $p_{RF}(x)$ and $p_{CNN}(x)$ denote the class 1 (PTSD) probability outputs from each model, respectively. The combined probability is computed as:

$$p_{\text{ensemble}}(x) = \frac{1}{2}\left[p_{\text{RF}}(x) + p_{\text{CNN}}(x)\right]$$

The final prediction is made by thresholding $p_{\text{ensemble}}(x)$, at 0.5. Using 5-fold cross-validation.

3. Results

3.1. Random Forest performance

As shown in Table 1, the model yielded a precision of 0.86, a recall of 0.93, and an F1-score of 0.89 for the PTSD class, while achieving a precision of 0.88, a recall of 0.78, and an F1-score of 0.82 for the healthy control class. The overall accuracy reached 87% under 5-fold cross-validation, with a macro-averaged F1-score of 0.86 and a weighted F1-score of 0.87, indicating balanced classification capability across both classes.

The ROC curve in Fig.4 demonstrates the model's discriminative ability, with an area under the ROC curve (AUC) of 0.919. However, the ROC curve exhibits reduced sensitivity in the low false positive rate (FPR) region, which suggests that the model may struggle to detect PTSD cases when prioritizing a low false alarm rate.

Table 1. Classification report for PTSD detection using Random Forest

	Precision	Recall	F1-score	Support
PTSD	0.86	0.93	0.89	27
Health Control	0.88	0.78	0.82	18
Accuracy	-	-	0.87	45
Macro Avg	0.87	0.85	0.86	45
Weighted Avg	0.87	0.87	0.87	45

DOI: 10.54254/2755-2721/2025.LD28893

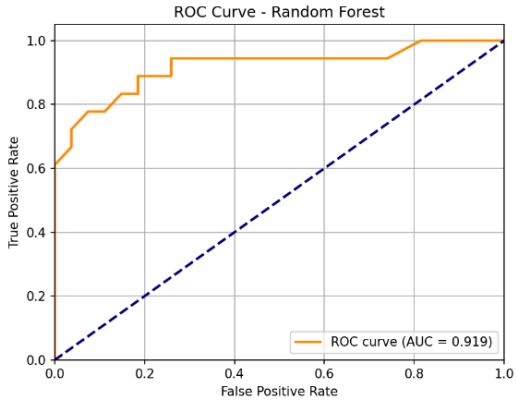


Figure 4. ROC curve of the Random Forest classifier

3.2. Pruned Random Forest

We analyzed the feature importance scores of all input channels, as shown in Fig. 5. The ranked feature importance chart shows that the most influential feature, delta.T3, accounts for 2.8% of the total importance, followed by gamma.FP2 at 2.3% and the third-ranked delta.O2 at 2.1%. Although individual importance scores are relatively low, this distribution indicates that classification relies on the combined contribution of many modestly informative features rather than a single dominant predictor.

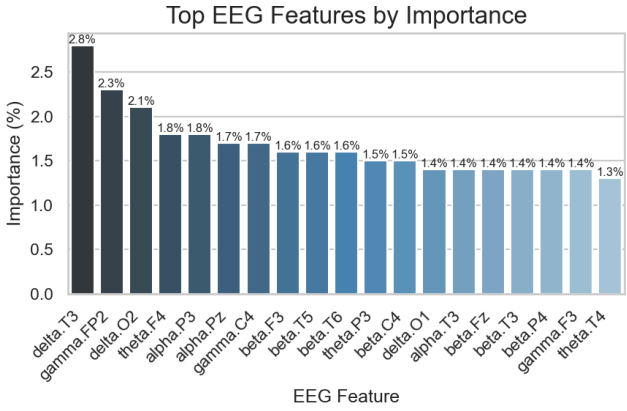


Figure 5. Ranked feature importance scores from the Random Forest classifier

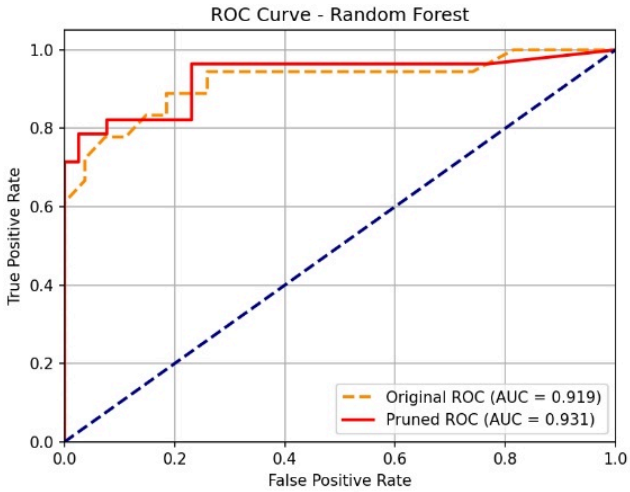


Figure 6. ROC curves before and after feature selection

We selected features with an importance score greater than 1% based on the results, resulting in 45 features chosen out of the original 114. The Random Forest classifier was retrained using only the selected features, and classification performance was re-evaluated accordingly. The pruned model demonstrated improved performance across most thresholds and achieved a higher 5-fold cross-validated accuracy of 88.4% (Fig. 6). Although the pruned RF's actual positive rate (TPR) is slightly lower than the original RF at a FPR of 0.2, it outperforms the original model across most other operating points.

DOI: 10.54254/2755-2721/2025.LD28893

3.3. CNN performance

The modified AlexNet CNN achieved a final test accuracy of 89.1%, outperforming the Random Forest baseline (88.4%).Fig. 7 shows the training process, including the loss and accuracy trends of 30 epochs. Notably, both metrics plateaued after 5 epochs, suggesting that the network had effectively converged early. The early convergence is likely attributable to the limited dataset size, which may have constrained the model's capacity and accuracy.

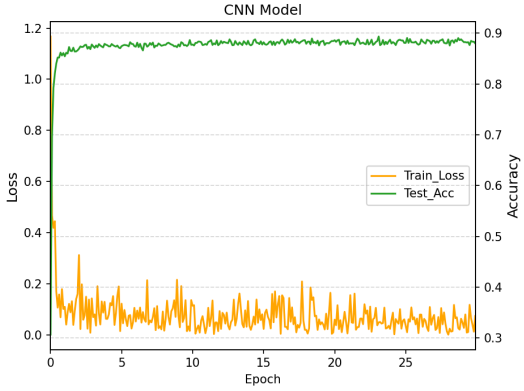


Figure 7. Training loss and accuracy curves of the CNN model across epochs

3.4. Ensemble results

The soft-voting ensemble outperformed individual models, achieving an average 5-fold cross-validated accuracy of 91.0%. The soft voting ensemble achieved an average accuracy of 91.0% (Table 2), outperforming both individual models

4. Discussion

Table 2. Classification performance metrics of the soft voting ensemble model

	Precision	Recall	F1-score	Support
PTSD	0.88	0.93	0.92	27
Health Control	0.89	0.82	0.89	18
Accuracy	-	-	0.91	45
Macro Avg	0.88	0.87	0.89	45
Weighted Avg	0.88	0.91	0.90	45

This study highlights the complementary strengths of different machine learning strategies for EEG-based PTSD recognition. The Random Forest classifier, particularly after feature pruning, demonstrated robust and balanced classification, while the CNN achieved slightly higher accuracy by leveraging its capacity for hierarchical feature learning. However, the soft-voting ensemble surpassed both individual models, reaching 91% accuracy with balanced precision and recall across groups. This result underscores the synergistic effect of combining handcrafted spectral features with deep-learned representations, suggesting that heterogeneous models can compensate for each other's weaknesses and yield a more reliable diagnostic system.

DOI: 10.54254/2755-2721/2025.LD28893

Our results indicate that the ensemble strategy achieved the best performance, outperforming both Random Forest and CNN individually. This suggests that handcrafted spectral features and deep-learned representations provide complementary information: RF captures interpretable frequency-domain patterns, while CNN automatically extracts nonlinear spatiotemporal features. Their combination thus yields a more reliable diagnostic tool.

Feature importance analysis further highlighted Delta.T3, Gamma.FP2 and Delta.O2 are relatively strong predictors, consistent with prior evidence that delta and gamma activity in temporal and frontal regions is altered in PTSD. While no single feature was dominant, the distributed importance across multiple channels reinforces the view that PTSD involves network-level dysfunctions rather than isolated abnormalities.

Beyond accuracy, other considerations are noteworthy. RF offers low computational cost and good interpretability, which is valuable in clinical settings. CNNs, although less transparent and more resource-intensive, demonstrated strong representational power. The ensemble balances these trade-offs by combining interpretability with predictive robustness. The CNN's rapid convergence within five epochs also reflects both efficiency and the constraint of limited sample size.

Nevertheless, the study is limited by the small dataset (n=52) and reliance on preprocessed features. Future work should validate the approach on larger and more diverse cohorts, explore adaptive ensemble weighting, and investigate real-time diagnostic feasibility to support clinical applicability.

5. Conclusion

The ensemble achieved the highest accuracy (91.0%) because RF and CNN contributed complementary strengths: RF exploited distributed spectral features with clear interpretability, while CNN extracted nonlinear spatial–temporal patterns. Their probability-level integration reduced variance and improved recall of PTSD cases compared to single models.

Feature ranking revealed Delta.T3 (2.8%), Gamma.FP2 (2.3%) and Delta.O2 (2.1%) are the top contributors. These regions correspond to temporal and frontal activity previously linked with hyperarousal and emotional regulation in PTSD, supporting the neurophysiological relevance of the selected predictors. The fact that most features carried >1% importance indicates that PTSD classification depends on multichannel network patterns rather than single biomarkers.

Beyond accuracy, RF had a lower computational cost and higher transparency, while CNN required more resources but converged rapidly within 5 epochs. The ensemble preserved clinical interpretability while improving robustness. However, the small sample size (n=52) limits generalizability. Larger datasets, adaptive ensemble weighting, and end-to-end modeling on raw EEG will be necessary for real-world validation.

References

[1] Lancaster, C. L., Teeters, J. B., Gros, D. F., & Back, S. E. (2016). Posttraumatic Stress Disorder: Overview of Evidence-Based Assessment and Treatment. Journal of Clinical Medicine, 5(11), 105. https: //doi.org/10.3390/jcm5110105

[2] Lynn M. Almli, Negar Fani, Alicia K. Smith, Kerry J. Ressler, Genetic approaches to understanding post-traumatic stress disorder, International Journal of Neuropsychopharmacology, Volume 17, Issue 2, February 2014, Pages 355–370, https: //doi.org/10.1017/S1461145713001090

[3] Peddi, A., Sendi, M. S. E., Minton, S. T., Langhinrichsen-Rohling, R., Hinojosa, C. A., West, E., Ressler, K. J., Calhoun, V. D., & van Rooij, S. J. H. (2025). Towards predicting posttraumatic stress symptom severity using portable EEG-derived biomarkers. Scientific Reports, 15, Article 5344. https: //doi.org/10.1038/s41598-025-88426-1

DOI: 10.54254/2755-2721/2025.LD28893

[4] Salazar-Castro, J. A., Peluffo-Ordóñez, D. H., & López, D. M. (2025). Advances in electroencephalography for post-traumatic stress disorder identification: A scoping review. IEEE Open Journal of Engineering in Medicine and Biology, 6, 332–344. https: //doi.org/10.1109/ojemb.2025.3538498

[5] Edlinger G, Rizzo C, Guger C: "Brain Computer Interface." Springer Handbook of Medical Technology. Springer, Berlin Heidelberg, 2011. 1003–1017.

[6] Peddi, A., Sendi, M.S.E., Minton, S.T. et al. Towards predicting posttraumatic stress symptom severity using portable EEG-derived biomarkers. Sci Rep 15, 5344 (2025). https: //doi.org/10.1038/s41598-025-88426-1

[7] S. M. Park et al., “Identification of Major Psychiatric Disorders from Resting-State Electroencephalography Using a Machine Learning Approach, ” *Frontiers in Psychiatry*, vol. 12, art. 707581, Aug. 2021. [Online]. Available: https: //osf.io/8bsvr

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks, ” in *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012. [Online]. Available: https: //dl.acm.org/ doi/pdf/10.1145/3065386