TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 6, December 2026, pp. 1962~1975
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i6.28137

Anxiety detection from EEG signals using discrete wavelet transform and XGBoost

Anisa Putri Setyaningrum, Yusup Miftahuddin, Yuzzar Alpriatna Malik
Faculty of Technology Industry, Informatics, Institut Teknologi Nasional Bandung, Bandung, Indonesia

Article Info

Article history:

Received Jun 30, 2026
Revised Sep 10, 2026
Accepted Oct 11, 2026

Keywords:

Anxiety
Discrete wavelet transform
Electroencephalography
Machine learning
Extreme gradient boosting

ABSTRACT

Anxiety detection from electroencephalography (EEG) signals is challenging because of their non-stationary nature and the risk of overfitting. Although discrete wavelet transform (DWT) and extreme gradient boosting (XGBoost) have each been applied to EEG classification, their combination has often been explored without channel selection or regularization. This study integrates DWT feature extraction with XGBoost classification, adding channel selection and regularization-focused tuning to improve accuracy and generalizability while preserving interpretability. Data came from the database for anxious states based on psychological stimulation (DASPS), comprising 23 subjects and 276 trials across 14 EEG channels, of which 10 (AF3, F7, F3, F4, F8, AF4, FC6, FC5, P7, and P8) were selected. Preprocessing involved a 50 Hz notch filter, a 4–45 Hz band-pass filter, independent component analysis (ICA), and 1-second segmentation. Features were extracted using the Daubechies-4 (db4) wavelet, followed by statistical feature computation and class balancing with the synthetic minority over-sampling technique (SMOTE). XGBoost was tuned on $n_estimators$, max_depth, $learning_rate$, reg_alpha, reg_lambda, and $gamma$. The baseline model achieved 73.8% accuracy, while the optimized model reached 88.9% accuracy, 93.8% precision, 89.9% recall, 91.8% F1-score, and 94.2% area under the curve (AUC). Addressing overfitting and channel relevance yields a more robust, interpretable framework.

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

Corresponding Author:

Anisa Putri Setyaningrum
Faculty of Technology Industry, Informatics, Institut Teknologi Nasional Bandung
P.H.H. Mustofa St., No.23, Cibeunying Kaler, Bandung, West Java, Indonesia
Email: anisaputrisetyaningrum@itenas.ac.id

1. INTRODUCTION

Anxiety disorders represent one of the most pressing mental health problems globally, affecting millions of individuals across various age groups and socio-economic backgrounds. Anxiety disorders represent one of the most pressing mental health problems globally, affecting millions of individuals across various age groups and socio-economic backgrounds. The World Health Organization (WHO) reports that anxiety disorders affected 359 million people worldwide in 2021, making them the most common mental disorder globally [1]. Anxiety not only affects emotional well-being but also impairs cognitive performance, academic achievement, and workplace productivity critical review of 38 studies found consistent evidence linking poor mental health, including anxiety, to lost productivity through absenteeism and presenteeism [2]. In severe cases, anxiety disorders can lead to depression, substance abuse, or even suicidal ideation, underscoring the urgency of early detection and effective management strategies.

Traditional approaches for detecting anxiety rely heavily on clinical interviews, psychological assessments, and self-report questionnaires such as the Hamilton anxiety rating scale (HAM-A) [3]. While widely used, these methods often suffer from subjectivity, inter-rater variability, and recall bias, which can result in misdiagnosis or inconsistent evaluation. Moreover, such assessments cannot always capture real-time fluctuations in anxiety states, making them less effective for continuous monitoring. This limitation has driven interest toward objective, physiological biomarkers that can complement or even replace traditional methods in certain contexts [4].

Electroencephalography (EEG) is one of the most promising physiological techniques for studying brain activity associated with mental states, offering high temporal resolution and real-time recording of neural oscillations linked to emotional and cognitive processes. In anxiety research, frontal and parietal regions involved in attention, emotional regulation, and executive control have been shown to exhibit distinct oscillatory alterations for instance, increased beta activity linked with heightened arousal and reduced alpha activity reflecting impaired relaxation [5], [6].

Despite its advantages, EEG analysis faces significant challenges: signals are inherently non-stationary, and vulnerable to noise from eye blinks, heartbeats, and motion artifacts. Traditional approaches such as fast fourier transform (FFT) and power spectral density (PSD) provide useful frequency information but often fail to capture transient events and the dynamic temporal evolution of brain activity, making them insufficient for the nuanced task of anxiety detection where both spectral content and temporal dynamics matter [7].

To address this, researchers have increasingly turned to time–frequency methods, particularly the discrete wavelet transform (DWT). DWT decomposes EEG signals into multiple levels, separating approximation and detail coefficients corresponding to different frequency bands, offering simultaneous localization in time and frequency a property especially valuable for wavelet packet-based decomposition approaches applied to abnormal and emotion-related EEG classification [8]. Features derived from DWT energy, entropy, variance, skewness, kurtosis have been shown to effectively differentiate anxiety and non-anxiety states [7], [9].

Studies applying EEG-based classification to anxiety and emotion recognition can broadly be grouped into three methodological families, each with distinct trade-offs that the current literature rarely compares directly. Conventional and ensemble learning approaches remain the most widely reported due to their low data requirements and computational efficiency. Muhammad and Al-Ahmadi [10] achieved 94.90% accuracy for binary anxiety classification using random forest (RF) on the database for anxiety states based on physiological signals (DASPS) dataset with only 9 selected features, illustrating that careful channel and feature selection can substitute for large sample sizes. Similarly, Aldayel and Al-Nafjan [7] found RF (87.5%) consistently outperformed adaptive boosting (AdaBoost) bagging (79%) when paired with DWT rather than PSD features, suggesting that ensemble performance is highly sensitive to the feature extraction method rather than the classifier alone. Shen *et al.* [5] reported even higher accuracy (97.83%) using support vector machine (SVM) and backpropagation–bootstrap aggregating (BP-Bagging), but their approach depended on 10-minute resting-state recordings from clinically diagnosed generalized anxiety disorder (GAD) patients a data collection burden that limits real-time or low-cost deployment compared to the short 1-second trial windows used by [11] on the same DASPS benchmark. This reveals a key limitation rarely discussed critically in prior review, reported accuracy figures across studies are not directly comparable, since they depend heavily on trial duration, dataset size, and label source rather than classifier choice alone.

Deep learning approaches have pushed reported accuracy higher still. Li *et al.* [12] achieved 87.8% on binary and 86.9% on quad-level anxiety classification using a convolutional neural network (CNN)–transformer architecture (MSDSTT) on DASPS, while Khan *et al.* [13] reported 99.7% accuracy fusing CNN-extracted spectrogram features with extreme gradient boosting (XGBoost) for emotion classification on database for emotion analysis using physiological signals (DREAMER). However, these gains come with well-documented costs, deep architectures require substantially larger training sets than the 23-subject DASPS benchmark typically provides, are computationally expensive to train and tune, and critically for clinical adoption function largely as black boxes. Fiorini *et al.* [14] explicitly raised this concern, applying explainability techniques to EEG emotion classifiers and showing that model behavior did not always align with theoretically expected neural correlates, meaning high accuracy alone does not guarantee that a deep learning model is learning clinically meaningful patterns.

Explainable, tree-based gradient boosting (XGBoost) occupies a middle ground that has received comparatively less critical attention in EEG-anxiety literature despite its demonstrated advantages elsewhere. Albaqami *et al.* [8] showed that wavelet packet decomposition (WPD) combined with gradient boosting decision trees (categorical boosting (CatBoost), XGBoost, light gradient boosting machine (LightGBM)) achieved 87.68% accuracy on abnormal EEG classification while outperforming prior state-of-the-art by more than 1% without requiring the computational infrastructure of deep learning. Beyond raw accuracy, XGBoost's built-in feature importance scores allow researchers to identify which specific EEG features and

frequency bands drive classification decisions, a property that neither black-box deep learning models nor most conventional classifiers provide as directly. This interpretability has proven valuable outside anxiety detection specifically Wei and Mooney [15] used XGBoost for seizure detection in clinical EEG precisely because false-positive interpretability mattered for physician trust, while Zeghlache *et al.* [16] combined XGBoost with autoencoder-based feature reduction for driver vigilance estimation, demonstrating the algorithm's flexibility across EEG applications beyond emotion or anxiety recognition alone.

Taken together, this comparison suggests that the choice between ensemble learning, deep learning, and XGBoost-based approaches for EEG-based anxiety detection should not be framed purely in terms of which achieves the highest reported accuracy, since accuracy figures across studies are confounded by dataset size, trial duration, and labeling protocol. Instead, the relevant trade-off lies between the data and computational demands of deep learning and the competitive accuracy with built-in explainability that gradient boosting frameworks such as XGBoost offer. However, prior applications of XGBoost to EEG classification, including those combined with DWT, have generally applied it either across the full EEG channel montage or with limited attention to regularization tuning [8], [15], [16] leaving open how much of XGBoost's potential for small, high-dimensional EEG datasets depends on addressing these two factors directly, which motivates the present study.

This study is guided by the following research question: can integrating DWT feature extraction with XGBoost classification, enhanced by channel selection and regularization-focused tuning, improve the accuracy and interpretability of EEG-based anxiety detection on small clinical datasets such as DASPS? based on these considerations, this study proposes a comprehensive framework for EEG-based anxiety detection that integrates DWT feature extraction with XGBoost classification, incorporating channel selection and regularization-focused hyperparameter tuning to jointly address overfitting risk and channel relevance. The framework includes preprocessing steps such as notch filtering, band-pass filtering, and independent component analysis (ICA) for artifact removal applied using the FastICA algorithm (scikit-learn), followed by segmentation into one-second windows. The extracted DWT features are then classified using XGBoost, with systematic hyperparameter tuning to balance bias and variance. By addressing the limitations of traditional methods and critically weighing the trade-offs of ensemble, deep learning, and explainable boosting approaches, this study aims to provide a robust, efficient, and interpretable framework for EEG-based anxiety detection one that improves generalizability on small clinical datasets without sacrificing the accuracy competitiveness demonstrated by more computationally intensive alternatives.

2. RESEARCH METHOD

This chapter presents the overall research methodology applied in this study, describing in detail how the research was designed, conducted, and evaluated. It explains the experimental framework used to develop the proposed model, including the data acquisition process, dataset characteristics, preprocessing procedures, feature extraction techniques, and model training strategy. The research procedure is systematically outlined through algorithmic steps or pseudocode to ensure clarity and reproducibility. Figure 1 shows the block diagram of the proposed research process. It illustrates the overall workflow, starting from EEG data acquisition, followed by preprocessing, feature extraction, model training, and performance evaluation. This diagram provides a clear overview of the methodological framework used in this study. The complete workflow, from channel selection to model evaluation, is summarized in Algorithm 1.

2.1. Dataset

This study utilized the DASPS, which contains physiological recordings from 23 participants across 276 trials. Each trial corresponds to a session designed to elicit varying levels of anxiety, ranging from low to high states. The EEG subset of DASPS was used, originally recorded with 14 channels placed according to the international 10–20 system. For this research, 10 channels were selected due to their strong association with frontal and parietal brain activity relevant to anxiety detection, namely AF3, F7, F3, F4, F8, AF4, FC6, FC5, P7, and P8. Each trial was labeled using the HAM-A, a widely adopted clinical tool for quantifying anxiety severity. Labels were subsequently binarized into anxious and non-anxious categories, forming the ground truth for supervised classification tasks. The dataset provided a balanced representation of different levels of anxiety states, but preprocessing was required to address class imbalance and noise contamination [11].

2.2. Data merging and selection

Each participant's EEG data was stored in separate trial files. These files were first merged into a unified dataset to ensure consistent formatting and facilitate subsequent analysis. Among the 14 available channels, only 10 channels were retained: AF3, F7, F3, F4, F8, AF4, FC6, FC5, P7, and P8. The selection was based on findings in neuroscience that highlight the role of frontal and parietal lobes in emotional

processing and anxiety-related activity. Frontal channels (AF3, F7, F3, F4, F8, and AF4) were retained due to their established role in frontal EEG asymmetry, a marker consistently linked to approach–avoidance motivational tendencies and trait anxiety [17], [18]. Parietal channels (P7 and P8) were retained given evidence that frontal and parietal asymmetries interact to predict attentional bias toward threat-related stimuli, implicating a front-parietal network in anxiety-related attentional processing [19]. Fronto-central channels (FC5 and FC6) were included as intermediate electrodes bridging the frontal and central regions relevant to this network. The remaining four channels (occipital and central-midline regions) were excluded, as they are less directly implicated in this frontal–parietal anxiety-processing network; omitting them also reduced feature dimensionality, mitigating overfitting risk given the dataset's limited sample size (23 subjects).

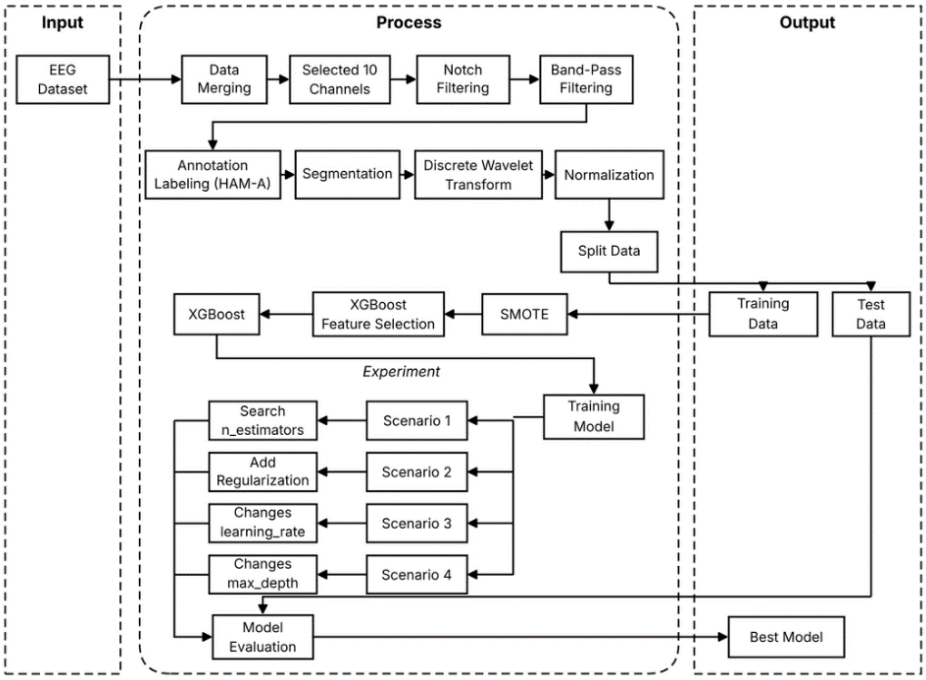


Figure 1. Flowchart of the proposed EEG-based anxiety detection

```
Algorithm 1. EEG-based anxiety detection pipeline
Input:  Raw EEG recordings R (DASPS dataset, 14 channels, fs=128 Hz), HAM-A scores H
Output: Trained XGBoost classifier M, performance metrics
1:  procedure ANXIETY_DETECTION(R, H)
2:      C ← SelectChannels(R, {AF3,F7,F3,F4,F8,AF4,FC6,FC5,P7,P8})
                                              ▷ based on neurological relevance
3:      C_notch ← NotchFilter(C, f0=50 Hz)
4:      C_bp ← BandpassFilter(C_notch, 4–45 Hz)
5:      C_clean ← FastICA(C_bp)                    ▷ remove ocular/muscular artifacts
6:      for each trial t in C_clean do
7:          label[t] ← Anxious if H[t] > 20 else Non-anxious
8:      end for
9:      S ← Segment(C_clean, window=1 s, 128 samples/segment)
10:     for each segment s in S do
11:         {A4, D4, D3, D2, D1} ← DWT(s, wavelet=db4, levels=4)
12:         F[s] ← ComputeStatistics(A4, D4, D3, D2, D1)  ▷ mean, std, quartiles
13:     end for
14:     (F_train, F_test, y_train, y_test) ← StratifiedSplit(F, labels, ratio=80:20)
15:     (F_train_bal, y_train_bal) ← SMOTE(F_train, y_train)  ▷ training set only
16:     M ← XGBoost_GridSearchCV(F_train_bal, y_train_bal,
            params={n_estimators, max_depth, learning_rate, reg_alpha, reg_lambda, gamma})
17:     P ← Evaluate(M, F_test, y_test)             ▷ accuracy, precision, recall,
                                                      F1, AUC, 95% CI (Wilson)
18:     I ← FeatureImportance(M)                    ▷ rank channel–frequency features
19:     return M, P, I
20: end procedure
```

2.3. Notch filtering

The EEG recordings are often contaminated with electrical interference from power lines, typically at 50 Hz in Indonesia. A notch filter was applied to suppress this narrowband noise. The transfer function of the notch filter is given by (1).

$$H(z) = \frac{1-2cos(\omega_0)z^{-1}+z^{-2}}{1-2rcos(\omega_0)z^{-1}+r^2z^{-2}} \quad (1)$$

where $\omega_0 = \frac{2\pi f_0}{f_s}$ with $f_0 = 50$ and $f_s = 128$ Hz. The parameter r determines the width of the notch, with values closer to 1 producing a sharper filter. This step ensured that periodic noise from electrical sources did not overlap with meaningful neural oscillations, particularly in the beta and gamma ranges [16].

2.4. Bandpass filtering

After filtering, annotation and labelling were performed to establish the ground truth for supervised learning. The DASPS dataset provides trial-level metadata, including psychological assessments conducted using HAM-A. The HAM-A evaluates anxiety severity across 14 items encompassing somatic and psychological dimensions. Scores are categorized into different levels ranging from mild to severe anxiety. In this study, the original continuous scores from HAM-A were binarized into two classes: "anxiety" and "non-anxiety," following thresholds defined in recent DASPS-based research [10]. The binary labels were assigned at the trial level and synchronized with the segmented EEG data. Each one-second segment inherited the label of its corresponding trial, ensuring consistency across preprocessing and feature-extraction stages. This annotation process aligned physiological data with psychological ground truth, providing a reliable target variable. The use of HAM-A as a clinical reference strengthens validity, as it remains an internationally recognized instrument for anxiety assessment [11], [13]. The 4–45 Hz band-pass range was selected to encompass the theta, alpha, beta, and gamma frequency bands most consistently associated with anxiety-related neural activity (see Table 1), while excluding the delta band (<4 Hz) associated with sleep and deep relaxation states not relevant to conscious anxiety response, and attenuating frequencies above 45 Hz to avoid overlap with the 50 Hz powerline interference targeted by the notch filter. This range is consistent with prior findings linking anxiety to alterations in beta activity (heightened arousal) and alpha activity (relaxation/vigilance) [5], as well as broader evidence that anxiety-related processing spans theta through gamma oscillations reflecting attention, memory, and cognitive control [7].

Table 1. Level decomposition of EEG signals using Daubechies-4 (db4) wavelet

Level decomposition	Coefficient	Band frequency	Description
A4	Approximation	0.5–4 Hz	Delta band - brain activity in a state of sleep or deep relaxation.
D4	Detail	4–8 Hz	Theta band - related to light relaxation and meditation.
D3	Detail	8–16 Hz	Alpha band - relaxed but alert state
D2	Detail	16-32 Hz	Beta band - brain activity in a state of focus or stress.
D1	Detail	32-64 Hz	Gamma band - related to high-level cognitive processing.

2.5. Annotation labelling

After the EEG signals were filtered through notch and band-pass filters, the next step involved annotation and labelling to establish the ground truth for supervised learning. The DASPS dataset provides trial-level metadata, including the results of psychological assessments conducted with the HAM-A. The HAM-A is a widely used clinical instrument that evaluates anxiety severity across 14 items, covering both somatic and psychological symptoms. Scores are then categorized into different levels ranging from mild to severe anxiety. In this study, the original continuous scores from the HAM-A were binarized into two classes: "anxiety" and "non-anxiety". This binarization followed thresholds defined in prior studies on DASPS [7], where scores below the clinical cut-off were labeled as non-anxious, while scores equal to or above the threshold were labeled as anxious. The binary labels were assigned at the trial level and synchronized with the segmented EEG data. Each one-second segment inherited the label of its corresponding trial, ensuring consistency across preprocessing and feature extraction stages [7]. Anxiety labeling was based on the HAM-A, with scores assigned by a psychotherapist following clinical evaluation of each subject's response during the experiment. A threshold of 20 was used to binarize continuous HAM-A scores: subjects scoring above 20 were labeled as anxious, reflecting clinically significant anxiety severity, while subjects scoring 20 or below were labeled as non-anxious. This threshold falls near the upper boundary of the moderate anxiety range under standard HAM-A severity conventions, meaning the non-anxious

category in this study encompasses both individuals with no clinically significant anxiety and those with mild-to-moderate symptoms below this cutoff, while the anxious category captures cases at the higher end of the clinical severity spectrum. This threshold choice prioritizes identifying clinically pronounced anxiety states over finer-grained distinctions between mild and moderate cases, which is appropriate given the binary classification framing adopted in this study.

The annotation process was essential to align physiological data with psychological ground truth. Without this step, the classification model would lack a reliable target variable. Moreover, the use of HAM-A as a clinical reference strengthens the validity of the dataset, as it is an internationally recognized scale for diagnosing and assessing anxiety disorders. By incorporating annotation and labelling after filtering, the dataset was transformed into a structured form where each EEG segment had a clearly defined label, ready to be used for feature extraction, balancing, and classification [3], [10].

2.6. Segmentation

Following annotation, EEG signals were segmented into shorter epochs to increase the number of training samples and improve temporal resolution. Segmentation is crucial in EEG analysis because anxiety-related neural dynamics can fluctuate within short time frames. In this study, each trial was divided into non-overlapping one-second windows (128 samples per segment, given fs = 128, Hzf_s = 128, $Hzfs$ = 128 Hz). The one-second window was chosen as a balance between temporal and spectral resolution [20], [21].

The one-second segment inherited its parent trial label, resulting in a larger dataset with thousands of labelled epochs that improved statistical robustness for machine learning [11]. Segmenting also ensured compatibility with the subsequent DWT process. Because DWT performs best on short, quasi-stationary segments, dividing EEG into one-second windows enhanced decomposition accuracy and allowed extraction of features across theta, alpha, beta, and gamma bands.

2.7. DWT

After preprocessing and segmentation, feature extraction was carried out to transform raw EEG signals into numerical representations that capture meaningful information for classification. In this study, the DWT was employed because of its ability to analyze non-stationary signals such as EEG in both the time and frequency domains. Unlike the FFT, which only provides frequency information without temporal localization, DWT decomposes signals into components that preserve both temporal resolution and spectral details. This makes it particularly suitable for capturing transient neural events associated with anxiety. The DWT operates by passing the EEG signal through a series of high-pass and low-pass filters, producing approximation coefficients (low-frequency components) and detail coefficients (high-frequency components) at successive levels of decomposition. In this study, the db4 wavelet was selected as the mother wavelet. The db4 wavelet is widely used in EEG research due to its compact support, orthogonality, and similarity to EEG waveform morphology, and has specifically been associated with capturing the fast transient patterns linked to anxiety-related cortical reactivity [7], [22]. Table 1 presents the level decomposition of EEG signals using the db4 wavelet. It describes how the original EEG signal is decomposed into multiple frequency sub-bands at different levels to capture relevant time–frequency characteristics for feature extraction and subsequent analysis. The statistical descriptors used (mean, standard deviation, and quartiles) were a deliberate choice prioritizing interpretability and efficiency, consistent with this study's focus on channel selection and regularization rather than feature engineering novelty. More advanced descriptors, such as entropy or fractal dimension measures, represent a promising direction for future work.

2.8. Synthetic minority over-sampling technique (SMOTE)

The SMOTE implemented using the imbalanced-learn library was employed to address the imbalance among anxiety classes in the EEG dataset. SMOTE generates new synthetic samples for minority classes by interpolating feature vectors between existing samples and their nearest neighbors. This method increases the representation of minority classes without merely duplicating data, thus improving the classifier's ability to generalize across all anxiety levels. After applying SMOTE, the class distribution became approximately uniform across low, moderate, and high anxiety categories, ensuring that the subsequent XGBoost training process was not biased toward the majority class [23]. SMOTE was selected over alternative balancing methods such as adaptive synthetic sampling (ADASYN) because prior work on this same dataset demonstrated its effectiveness for addressing anxiety-state class imbalance [23], whose safe-level SMOTE variant achieved improved classification performance on DASPS. While ADASYN adaptively generates more synthetic samples for harder-to-classify minority instances, it can also amplify noise near class boundaries given that EEG features are already susceptible to artifact-related noise despite preprocessing, standard SMOTE's uniform interpolation was considered a more conservative and stable choice for this dataset.

2.9. XGBOOST

The extracted DWT coefficients were then used as input features for classification using the XGBoost algorithm. XGBoost is an ensemble learning method based on gradient boosting decision trees that provides high accuracy and robustness while preventing overfitting through regularization [15]. Each tree in the ensemble is built sequentially to minimize the classification error from the previous tree, and the final prediction is obtained by aggregating the outputs of all weak learners. In this study, XGBoost was implemented in Python using the xgboost library. The hyperparameters optimized during model tuning included the learning rate (η), maximum tree depth, number of estimators, subsample ratio, and column sampling rate. The optimization process was performed using grid search cross-validation to determine the configuration that maximized performance and minimized overfitting. Recent research has demonstrated the effectiveness of XGBoost in biomedical signal classification and emotion recognition, outperforming conventional classifiers such as SVM and RF in terms of both accuracy and interpretability. Furthermore, explainability methods integrated with gradient boosting have been increasingly used to identify which EEG features most strongly contribute to the detection of anxiety or stress, supporting transparency in clinical decision-making [4], [24].

The hyperparameter search ranges were chosen with consideration for the dataset's limited size (4,608 training samples after SMOTE balancing) to reduce the risk of overfitting inherent to more complex configurations. For max_depth, values were explored in the range of 3 to 6, rather than deeper trees typically used for large-scale datasets; shallower depths (3–6) limit the number of splits each tree can make, constraining model complexity relative to the number of available training samples and the moderately high-dimensional, redundant nature of the DWT-derived feature space (see section 3.1). Similarly, the search ranges for n_estimators, $learning_rate$, and the regularization parameters (reg_alpha, reg_lambda, and $gamma$) were selected to prioritize configurations capable of balancing model capacity against the generalization constraints of a small clinical EEG dataset, following the progressive tuning strategy described in section 3 (Table 2), where regularization and depth were incrementally introduced and evaluated for their effect on the train–test performance gap.

2.10. Model evaluation

The classification performance was evaluated using common metrics applied in EEG-based affective computing research: accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (under the curve / AUC). These evaluation metrics were computed based on the confusion matrix derived from the test set, ensuring a comprehensive assessment of model generalization capability [13]. To ensure robustness, the experiment followed a stratified 80:20 data split for training and testing. Data imbalance was addressed using the SMOTE procedure described previously [23]. A limitation of this study is its reliance on a stratified 80:20 train–test split rather than subject-independent validation. Although this approach is consistent with several prior DASPS-based studies, it may not fully capture the model's ability to generalize to entirely new subjects. Future work should incorporate leave-one-subject-out cross-validation (LOSO-CV) or other cross-subject evaluation schemes to more rigorously assess generalizability across the full DASPS subject pool. The overall workflow, including preprocessing, feature extraction, class balancing, and model training, is summarized in Figure 1. Such evaluation practices are consistent with modern standards in EEG-based machine learning, where multiple metrics are preferred over accuracy alone to ensure balanced performance across both major and minority classes [11].

To assess the statistical reliability of the reported accuracy values, 95% confidence intervals (CI) were computed using the Wilson score interval method [25]. Brown *et al*. [26] demonstrated that the conventional Wald interval exhibits erratic coverage behavior, particularly for smaller sample sizes, and recommended the Wilson interval as a more reliable alternative a consideration directly relevant to this study's test set (n = 828) [26]. The Wilson interval was computed for the baseline and optimized model configurations, allowing a comparison of whether their respective accuracy ranges overlap non-overlapping intervals provide evidence that the observed performance difference is unlikely to result from sampling variability in the single train–test split used.

3. RESULTS AND DISCUSSION

Five cumulative experimental scenarios were performed to progressively optimize model performance, with each scenario building on the configuration of the previous one rather than testing parameters in isolation. This design allows the interaction between capacity-increasing parameters ($n_estimators$ and max_depth) and regularization parameters (reg_alpha, reg_lambda, and $gamma$) to be observed directly, rather than evaluating each hyperparameter's effect independently from the baseline. As shown in Table 2, test accuracy increased from 73.8% (scenario 1) to 88.9% (scenario 5), but the train–test

gap across scenarios reveals a more nuanced pattern than accuracy alone suggests. The baseline model (scenario 1) showed a comparatively small train–test gap (78.7% vs. 73.8%, a 4.9-point difference) but with high loss on both sets (0.59/0.62), indicating the model had not yet learned sufficient discriminative structure from the DWT features consistent with underfitting rather than a generalization problem. Adding estimators in scenario 2 sharply increased model capacity (train accuracy increases to 97.3%), but the train–test gap widened to 9.9 points, signaling the onset of overfitting as the model began memorizing training-set-specific patterns. Introducing regularization (reg_alpha, reg_lambda, and $gamma$) in scenario 3 reduced train accuracy to 90.2% while test accuracy fell only slightly to 85.0%, narrowing the gap to 5.2 points direct evidence that the regularization terms constrained the model's capacity to overfit, at the cost of a small reduction in peak performance. Scenario 4 (learning rate adjustment) maintained a similar gap (6.1 points) while improving test accuracy to 86.8%, suggesting that a lower learning rate improved convergence stability without reintroducing overfitting.

Table 2. Overall model performance

Experiment	Accuracy (train/test) (%)	Precision (train/test) (%)	Recall (train/test) (%)	F1-score (train/test) (%)	AUC (train/test) (%)	Description
1 – baseline	78.7/73.8	80.6/79.7	75.4/71.7	78.3/79.2	86.4/83.0	Basic model, significant gap between train and test, indication of overfitting.
2 – $n_estimators$	96.1/85.3	97.8/91.4	94.4/87.0	96.1/89.1	99.3/92.8	Adding more trees improves accuracy but starts overfit.
3 – Regularization	90.2/85.0	91.9/92.5	88.3/85.4	90.0/88.8	96.7/91.3	Regularization suppresses overfitting, resulting in smaller gaps and stable models.
4 – Learning rate	92.9/86.8	94.4/93.0	91.3/87.7	92.8/90.3	98.3/93.0	A higher learning rate accelerates convergence, balanced results.
5 – Max depth	97.0/88.9	98.8/93.8	95.3/89.9	97.0/91.8	99.7/94.2	A greater tree depth produces higher accuracy but tends to cause overfitting.

Scenario 5 (increased max_depth) then raised test accuracy further to 88.9%, but the train–test gap widened again to 8.1 points (97.0% vs. 88.9%) nearly returning to the gap magnitude observed in scenario 2. This indicates that increasing tree depth reintroduces overfitting risk that the fixed regularization strength ($reg_lambda = 0.5$) inherited from scenario 3 was not fully able to offset. In other words, deeper trees increase model complexity multiplicatively, while the regularization penalty applied here scales additively with tree structure; the result is a configuration that achieves the highest test accuracy in absolute terms but does so by trading back some of the generalization gain achieved in scenario 3–4. This is a critical distinction the accuracy figure alone does not convey: the final model is not the scenario with the best bias–variance balance it is the scenario with the highest test accuracy despite a re-widened generalization gap. We report this explicitly rather than presenting Scenario 5 as free of overfitting and note that scenarios 3 or 4 may represent a more conservative choice if minimizing the train–test gap, rather than maximizing raw test accuracy, is prioritized for deployment on unseen subjects.

Examining AUC alongside accuracy reinforces the bias–variance trade-off discussed above: scenario 3 (regularization) achieved the narrowest train–test AUC gap (96.7% vs. 91.3%, a 5.4-point difference), while scenario 5 (max_depth) achieved the highest test AUC in absolute terms (94.2%) but with a comparable train–test gap (99.7% vs. 94.2%), reflecting the same overfitting pattern observed in the accuracy analysis.

Table 3 contextualizes the proposed framework's performance against prior DASPS-based anxiety classification studies. Direct comparison should be interpreted cautiously, as these studies differ in channel selection, segmentation, and label binarization thresholds. Nonetheless, the proposed framework outperforms studies using the full 14-channel montage without channel selection [11], [27] and performs comparably to Daud *et al.* [23] who applied SMOTE with a k-nearest neighbors (KNN) classifier. It does not exceed the 94.9% accuracy reported by Muhammad and Al-Ahmadi [10], who used RF with a more aggressive 6-channel selection. This suggests that further channel or feature reduction beyond the present 10-channel configuration, combined with the regularization strategy demonstrated in this study, represents a promising direction for future improvement.

To assess whether the observed improvement reflects a statistically meaningful difference rather than variance from a single train–test split, 95% CI were computed for the baseline and optimized model accuracies using the Wilson score interval method described in section 2.10. The baseline model achieved 73.8% accuracy (95% CI: 70.7%–76.7%), while the optimized model achieved 88.9% accuracy (95% CI: 86.6%–90.9%). The non-overlapping intervals indicate that the observed improvement is unlikely to be

attributable to chance variation in the test set composition, supporting the practical significance of the reported performance gain alongside the bias–variance trade-off discussed above.

Table 3. Comparison with prior DASPS-based anxiety classification studies

State of the art	Method	Channels	Accuracy
Baghdadi *et al.* (2019) [11]	Stacked sparse autoencoder	14 (full)	83.5%
Chatterjee *et al.* (2023) [27]	Spatio-temporal feature transition	14 (full)	83.8%
Djadoudi *et al.* (2024) [6]	KNN (frontal channels only)	6 (frontal)	84.2%
Daud *et al.* (2023) [23]	KNN + safe-level SMOTE	14 (full)	89.5%
Muhammad and Al-Ahmadi (2022) [10]	RF, channel + feature selection	6 (selected)	94.9%
This study	DWT + XGBoost + SMOTE, regularization tuning	10 (selected)	88.9%

3.1. Optimized model configuration

The optimised configuration of the XGBoost model was obtained with $learning_rate = 0.1$, $max_depth = 6$, $n_estimators = 200$, $subsample = 0.8$, and $reg_lambda = 0.5$. This depth can be interpreted in relation to the dimensionality and structure of the DWT-derived feature space with 10 selected channels, each decomposed into 5 sub-bands and multiple statistical descriptors extracted per sub-band, the resulting feature space is moderately high-dimensional but contains substantial redundancy, as adjacent statistical descriptors from the same channel and sub-band (e.g., mean and q75 of AF3_detail2) tend to be correlated. A tree depth of 6 allows sufficient partitioning to capture interactions among the small subset of channel–sub-band combinations that later feature-importance analysis confirmed to carry most of the discriminative signal, without the tree over-partitioning on redundant or low-information features that shallower or deeper configurations would under- or over-utilize. Figure 2 presents the accuracy and loss curves of this configuration, illustrating the model's training behaviour across boosting iterations. The validation loss steadily decreased and stabilised around 0.30 after approximately 80 iterations, indicating stable convergence rather than continued learning or divergence.

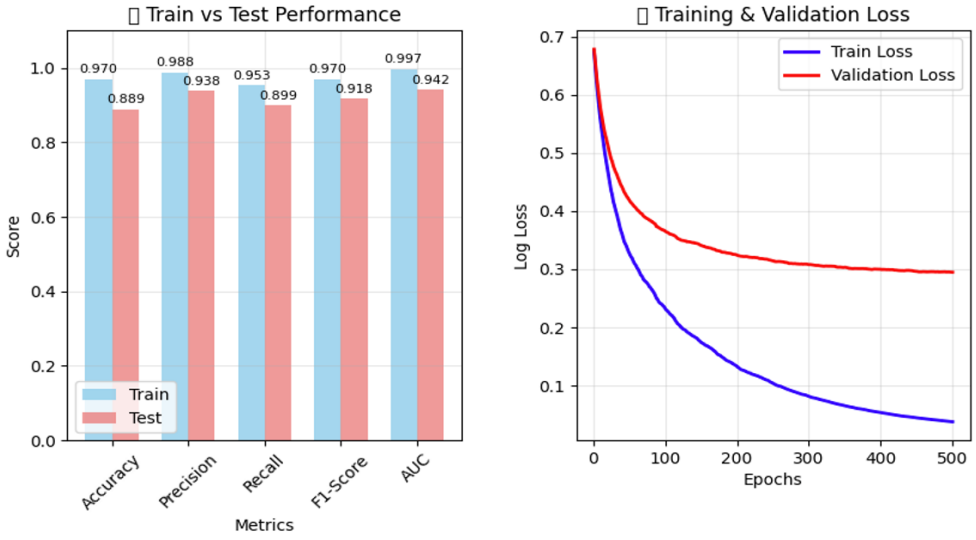


Figure 2. Accuracy and loss model best scenario

However, consistent with the train–test gap analysis in the preceding section, the training and validation accuracy curves do not fully converge, retaining a gap of approximately 8 percentage points (97.0% vs. 88.9%) indicating that while this configuration achieved the highest test accuracy among the tested scenarios, some degree of overfitting remains present rather than fully eliminated. This suggests that max_depth=6 approaches, but does not exceed, the point at which model capacity begins to outpace the amount of available training data ($n = 4{,}608$ segments) relative to the effective dimensionality of the discriminative feature subset. Although runtime was not formally benchmarked in this study, training on the available dataset (4,608 samples after SMOTE balancing) completed efficiently given XGBoost's CPU-based

tree ensemble structure, without requiring GPU acceleration in contrast to the substantially longer training times typically required by deep learning alternatives. Formal latency and scalability benchmarking remain necessary future work to establish deployment feasibility on wearable or edge hardware.

3.2. Class wise performance analysis

Figure 3 shows the confusion matrices of the optimized XGBoost model for both training and testing datasets. The training matrix demonstrates a high accuracy of 97.0 %, while the testing matrix achieves 88.9 %, indicating good generalization capability with only a moderate reduction in performance between training and unseen data. For the testing set, the model correctly identified 518 anxious and 218 non-anxious samples, misclassifying only a small portion of each class. The slightly higher number of misclassifications in the non-anxious, anxious direction suggests that the model is more sensitive to anxiety-related EEG patterns, which is desirable for avoiding false negatives in psychological state detection. The relatively balanced performance across both classes confirms that the application of SMOTE effectively addressed class imbalance during training, enabling the model to learn representative decision boundaries. Overall, the results verify that the tuned model not only achieved high global accuracy but also maintained stable recognition across anxiety and non-anxiety conditions, demonstrating its robustness for EEG-based anxiety detection.

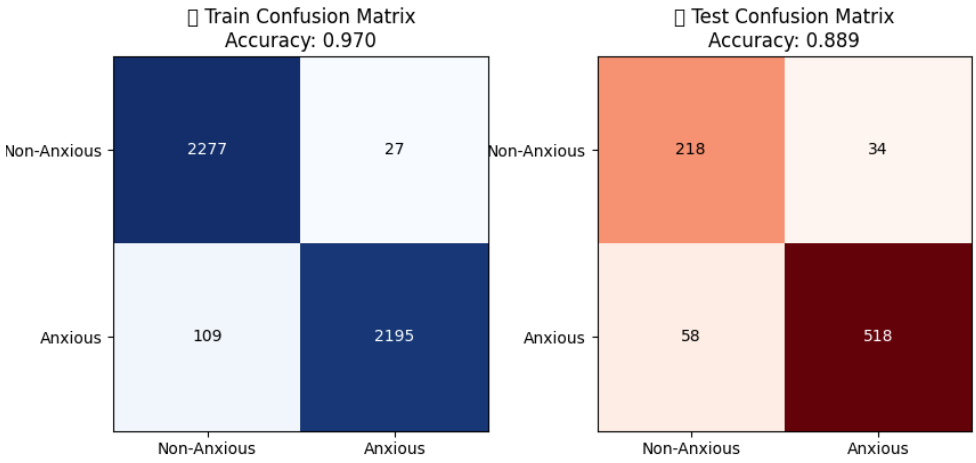


Figure 3. Confusion matrix model best scenario

3.3. Feature importance analysis

Figure 4 shows the feature-importance analysis of the optimised XGBoost model, illustrating how each EEG feature contributed to the classification of anxiety and non-anxiety states. The results reveal that the most discriminative features were predominantly derived from the frontal (AF3 and F8) and parietal (P7 and P8) electrodes, followed by smaller contributions from the central regions such as FC5 and FC6. Among the top 20 features, statistical measures such as the mean, standard deviation, and upper quartile (q75) of the detail-2 and detail-3 coefficients obtained through the DWT decomposition were found to be the most influential. These coefficients correspond to the beta (13–30 Hz) and gamma (30–64 Hz) frequency bands, which are known to reflect neural activities associated with emotional arousal, attention, and cognitive control. The most dominant features $AF3_detail2_mean$, $AF3_detail2_q75$, and $P8_detail3_std$ indicate that high-frequency oscillations in these brain regions carry critical information for detecting anxiety-related changes in EEG signals. The channel-wise analysis further confirms this pattern, where the AF3 electrode exhibited the highest overall importance (0.16), followed by P8 (0.14) and F8 (0.124), suggesting that both frontal and parietal lobes play key roles in modulating anxiety-related neural responses. This frontal–parietal activation pattern aligns with prior neurophysiological evidence indicating that frontal asymmetry is linked to emotional regulation, while parietal engagement reflects altered attentional processing under anxious states. Overall, the feature-importance results validate that the proposed DWT–XGBoost framework not only achieved high predictive accuracy but also successfully extracted meaningful and interpretable neurophysiological markers of anxiety, providing both diagnostic and explanatory value for EEG-based affective computing and mental-health monitoring.

Returning to the research question posed in the Introduction, the findings of this study indicate that integrating DWT feature extraction with XGBoost, enhanced by channel selection and regularization-focused

tuning, does improve both accuracy and interpretability of EEG-based anxiety detection on the small DASPS dataset. Accuracy improved from 73.8% to 88.9%, a gain supported by non-overlapping CI, while regularization measurably narrowed the train–test gap relative to the unregularized configuration. Interpretability was preserved and enhanced through feature-importance analysis, which identified specific channel–frequency combinations (e.g., AF3 detail-2 coefficients in the beta band) driving classification decisions a level of transparency not inherently available in deep learning alternatives. However, the results also indicate this improvement is partial rather than complete: the optimized configuration retained a residual overfitting gap, indicating that channel selection and regularization mitigated, but did not fully resolve, the generalization challenge posed by the dataset's small size.

Precision (93.8%) exceeded recall (89.9%), reflecting more missed anxious cases (58 false negatives) than misclassified non-anxious cases (34 false positives) consistent with the test set retaining its original class imbalance (69.6% anxious) despite SMOTE being applied only to training data, causing the model to apply a comparatively conservative decision boundary at test time. This pattern aligns with the feature-importance findings Figure 4: the dominant detail-2 (beta, 16–32 Hz) and detail-3 (alpha, 8–16 Hz) coefficients reflect established anxiety signatures elevated beta activity linked to cortical arousal and hypervigilance, and altered alpha activity reflecting reduced relaxation under threat while the smaller detail-1 (gamma) contribution aligns with heightened cognitive control during anxious states. This convergence between the model's discriminative features and known anxiety-related neural patterns suggests the classifier is learning physiologically meaningful representations rather than dataset-specific artifacts.

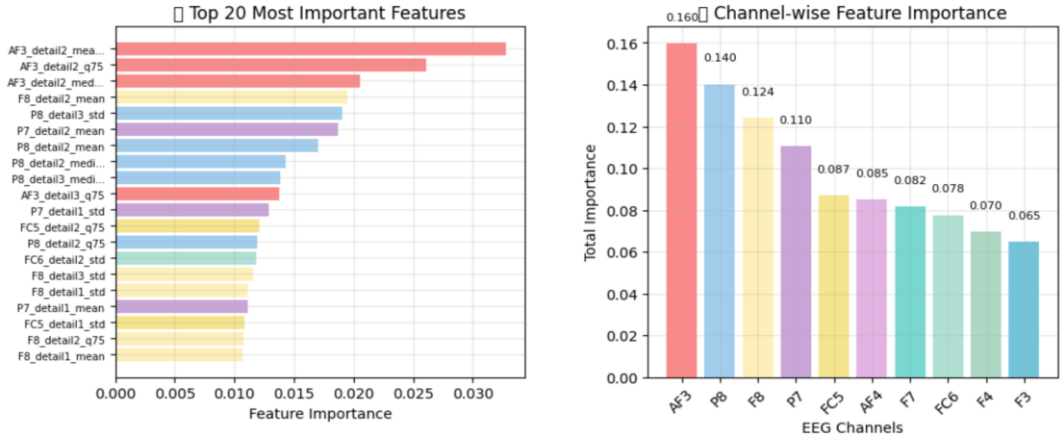


Figure 4. Feature-importance ranking and channel-wise importance

4. CONCLUSION

This study developed an EEG-based anxiety detection framework combining DWT feature extraction, using the db4 mother wavelet for simultaneous time–frequency decomposition, with XGBoost classification. Through systematic hyperparameter tuning, test accuracy improved from 73.8% to 88.9%, a gain supported by non-overlapping 95% CI. The optimal configuration (deeper trees with regularization) achieved the highest accuracy but retained a residual train–test gap of approximately 8 percentage points, indicating overfitting was reduced but not fully eliminated a trade-off to consider for deployment. Feature-importance analysis showed the most discriminative information came from detail-2 (beta) and detail-3 (alpha) coefficients, with a smaller contribution from detail-1 (gamma), concentrated in frontal (AF3 and F8) and parietal (P7 and P8) regions linked to emotional regulation and attention, confirming the neurophysiological relevance of the extracted features.

Beyond classification performance, the framework offers practical engineering value, reducing electrode requirements from 14 to 10 channels lowers hardware complexity for wearable EEG devices, while XGBoost's computational efficiency makes it comparatively lightweight for resource-constrained deployment relative to deep learning alternatives. Real-time integration, however, was not evaluated in this study and would require further engineering work, such as streaming-compatible artifact handling. From a clinical perspective, the reduced 10-channel montage and lightweight classifier position this framework as a candidate for low-cost anxiety screening e.g., preliminary triage in primary care or community settings using low-cost EEG headsets, flagging individuals for further clinical assessment. However, this would require

prospective clinical validation against clinician-administered diagnoses in real-world settings, which we identify as future translational work rather than a capability demonstrated here.

This study's generalizability is bounded by the DASPS dataset's scope 23 subjects, binary anxiety labeling, and a single train–test split without subject-independent or cross-dataset validation constraints common to DASPS-based studies in this area. Addressing these through larger cohorts, multi-level classification, and subject-independent evaluation represents a natural next step building on the framework established here.

ACKNOWLEDGMENTS

The authors would like to thank the creators of the DASPS dataset for making their data publicly available. This research did not receive any specific research grant or contract; it was self-funded by the authors.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

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
Anisa Putri Setyaningrum	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Yusup Miftahuddin	✓	✓		✓		✓		✓		✓		✓		
Yuzzar Alpriatna Malik			✓	✓	✓		✓		✓			✓		

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

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

INFORMED CONSENT

This study used the publicly available DASPS dataset. The authors did not collect any new data from human participants; therefore, informed consent was not required for this secondary analysis of existing anonymized data.

ETHICAL APPROVAL

This study used the publicly available DASPS dataset: https://ieee-dataport.org/open-access/dasps-database, which was collected and released by its original authors. The authors did not recruit participants or collect new data from human subjects; therefore, institutional review board approval was not required for this secondary analysis of existing anonymized data.

DATA AVAILABILITY

The raw EEG data that support the findings of this study are openly available in the IEEE DataPort repository as the DASPS database at https://doi.org/10.21227/barx-we60, described by Baghdadi *et al*. [11].

Derived data supporting the findings of this study (extracted wavelet features and model outputs) are available from the corresponding author, A.P.S., upon reasonable request.

REFERENCES

[1] World Health Organization, "Anxiety disorders," 2026. https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders (accessed Aug. 07, 2026).

[2] C. de Oliveira, M. Saka, L. Bone, and R. Jacobs, "The Role of Mental Health on Workplace Productivity: A Critical Review of the Literature," *Applied Health Economics and Health Policy*, vol. 21, no. 2, pp. 167–193, Nov. 2022, doi: 10.1007/s40258-022-00761-w.

[3] M. A. X. Hamilton, "The assessment of anxiety states by rating," *British Journal of Medical Psychology*, vol. 32, no. 1, pp. 50–55, Mar. 1959, doi: 10.1111/j.2044-8341.1959.tb00467.x.

[4] L. Ancillon, M. Elgendi, and C. Menon, "Machine Learning for Anxiety Detection Using Biosignals: A Review," *Diagnostics*, vol. 12, no. 8, p. 1794, 2022, doi: 10.3390/diagnostics12081794.

[5] Z. Shen *et al.*, "Aberrated Multidimensional EEG Characteristics in Patients with Generalized Anxiety Disorder: A Machine-Learning Based Analysis Framework," *Sensors*, vol. 22, no. 14, p. 5420, 2022, doi: 10.3390/s22145420.

[6] M. S. Djadoudi, S. Soulimane, and K. F. Arbi, "The Role of Frontal Electrodes in Human State Anxiety Detection Using Wavelet Features and Machine Learning," in *2024 2nd International Conference on Electrical Engineering and Automatic Control (ICEEAC)*, May 2024, pp. 1–7, doi: 10.1109/iceeac61226.2024.10576469.

[7] M. Aldayel and A. Al-Nafjan, "A comprehensive exploration of machine learning techniques for EEG-based anxiety detection," *PeerJ Computer Science*, vol. 10, p. e1829, Jan. 2024, doi: 10.7717/peerj-cs.1829.

[8] H. Albaqami, G. M. Hassan, A. Subasi, and A. Datta, "Automatic detection of abnormal EEG signals using wavelet feature extraction and gradient boosting decision tree," *Biomedical Signal Processing and Control*, vol. 70, p. 102957, 2021, doi: 10.1016/j.bspc.2021.102957.

[9] A. Subasi, T. Tuncer, S. Dogan, D. Tanko, and U. Sakoglu, "EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier," *Biomedical Signal Processing and Control*, vol. 68, p. 102648, 2021, doi: 10.1016/j.bspc.2021.102648.

[10] F. Muhammad and S. Al-Ahmadi, "Human state anxiety classification framework using EEG signals in response to exposure therapy," *PLOS ONE*, vol. 17, no. 3, p. e0265679, Mar. 2022, doi: 10.1371/journal.pone.0265679.

[11] A. Baghdadi, Y. Aribi, R. Fourati, N. Halouani, P. Siarry, and A. M. Alimi, "DASPS: A Database for Anxious States based on a Psychological Stimulation," 2019, *arXiv*. doi: 10.48550/ARXIV.1901.02942.

[12] Q. Li, Y. Sun, Y. Xie, and Y. Zhou, "EEG-based anxiety emotion classification using an optimized convolutional neural network and transformer," *Signal, Image and Video Processing*, vol. 19, no. 6, Apr. 2025, doi: 10.1007/s11760-025-04067-x.

[13] M. S. Khan, N. Salsabil, M. G. R. Alam, M. A. A. Dewan, and M. Z. Uddin, "CNN-XGBoost fusion-based affective state recognition using EEG spectrogram image analysis," *Scientific Reports*, vol. 12, no. 1, Aug. 2022, doi: 10.1038/s41598-022-18257-x.

[14] L. Fiorini, F. Bossi, and F. Di Gruttola, "EEG-based emotional valence and emotion regulation classification: a data-centric and explainable approach," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, doi: 10.1038/s41598-024-75263-x.

[15] L. Wei and C. Mooney, "Epileptic Seizure Detection in Clinical EEGs Using an XGboost-based Method," in *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec. 2020, pp. 1–6, doi: 10.1109/spmb50085.2020.9353625.

[16] R. Zeghlache, M. A. Labiod, and A. Mellouk, "Driver vigilance estimation with Bayesian LSTM Auto-encoder and XGBoost using EEG/EOG data," *IFAC-PapersOnLine*, vol. 55, no. 8, pp. 89–94, 2022, doi: 10.1016/j.ifacol.2022.08.015.

[17] J. A. Coan and J. J. B. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological Psychology*, vol. 67, no. 1–2, pp. 7–50, Oct. 2004, doi: 10.1016/j.biopsycho.2004.03.002.

[18] R. Thibodeau, R. S. Jorgensen, and S. Kim, "Depression, anxiety, and resting frontal EEG asymmetry: A meta-analytic review.," *Journal of Abnormal Psychology*, vol. 115, no. 4, pp. 715–729, Nov. 2006, doi: 10.1037/0021-843x.115.4.715.

[19] G. M. Grimshaw, J. J. Foster, and P. M. Corballis, "Frontal and parietal EEG asymmetries interact to predict attentional bias to threat," *Brain and Cognition*, vol. 90, pp. 76–86, Oct. 2014, doi: 10.1016/j.bandc.2014.06.008.

[20] O. Balli, "Use of XGBoost Algorithm in Classification of EEG Signals," *1st International Conference on Engineering, Natural and Social Sciences*, 2022.

[21] Y. Chen, R. Chang, and J. Guo, "Emotion Recognition of EEG Signals Based on the Ensemble Learning Method: AdaBoost," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, Jan. 2021, doi: 10.1155/2021/8896062.

[22] F.-C. Adochiei, A. Ioniță, I.-R. Adochiei, O.-I. Stirbu, G. Petroiu, and F. C. Argatu, "Computational Analysis of EEG Responses to Anxiogenic Stimuli Using Machine Learning Algorithms," *Applied Sciences*, vol. 16, no. 3, p. 1504, Feb. 2026, doi: 10.3390/app16031504.

[23] S. N. S. S. Daud, R. Sudirman, and T. W. Shing, "Safe-level SMOTE method for handling the class imbalanced problem in electroencephalography dataset of adult anxious state," *Biomedical Signal Processing and Control*, vol. 83, p. 104649, May 2023, doi: 10.1016/j.bspc.2023.104649.

[24] M. Xing, J. M. Fitzgerald, and H. Klumpp, "Classification of Social Anxiety Disorder With Support Vector Machine Analysis Using Neural Correlates of Social Signals of Threat," *Frontiers in Psychiatry*, vol. 11, Mar. 2020, doi: 10.3389/fpsyt.2020.00144.

[25] E. B. Wilson, "Probable Inference, the Law of Succession, and Statistical Inference," *Journal of the American Statistical Association*, vol. 22, no. 158, pp. 209–212, 1927, doi: 10.1080/01621459.1927.10502953.

[26] L. D. Brown, T. T. Cai, and A. DasGupta, "Interval Estimation for a Binomial Proportion," *Statistical Science*, vol. 16, no. 2, pp. 101–133, May 2001, doi: 10.1214/ss/1009213286.

[27] D. Chatterjee, R. Gavas, and S. K. Saha, "Detection of mental stress using novel spatio-temporal distribution of brain activations," *Biomedical Signal Processing and Control*, vol. 82, p. 104526, Apr. 2023, doi: 10.1016/j.bspc.2022.104526.

BIOGRAPHY OF AUTHORS



Anisa Putri Setyaningrum received her Master of Engineering degree from Institut Teknologi Bandung (ITB). She is currently a lecturer at Institut Teknologi Nasional (ITENAS), where she teaches courses in signal and speech processing, digital image processing, and computer vision. Her research interests focus on artificial intelligence and machine learning, particularly in signal-based analysis, biomedical signal processing, pattern recognition, and intelligent vision systems. Her recent work involves the development of machine learning and deep learning models for signal classification, feature extraction, and explainable AI applications. She can be contacted at email: anisaputrisetyaningrum@itenas.ac.id.



Yusup Miftahuddin received his Master of Engineering degree from Institut Teknologi Bandung (ITB). He is currently a lecturer at Institut Teknologi Nasional (ITENAS). His research interests include software engineering, signal processing, and artificial intelligence. His work focuses on the development of intelligent systems, system design and optimization, and the integration of machine learning techniques into software-based solutions. He is actively involved in research related to applied AI and signal-based system development. He can be contacted at email: yusufm@itenas.ac.id.



Yuzzar Alpriatna Malik received his Bachelor's degree in Informatics from Institut Teknologi Nasional (ITENAS). His research interests focus on machine learning and signal processing, particularly in the development of intelligent models for signal analysis and pattern recognition. His undergraduate research involved the application of machine learning techniques for signal-based classification and data-driven system development. He can be contacted at email: yuzzar.al@mhs.itenas.ac.id.