Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Implementing Random Forest for Eye State Detection in an EEG-Based Brain-Computer Interface System

Muhammad Alfathan[1], Tri Sugihartono[2]
Informatics Engineering, Faculty of Information Technology, ISB Atma luhur
2211500050@mahasiswa.atmaluhur.ac.id , trisugihartono@atmaluhur.ac.id

Submitted :Apr 6, 2026 | **Accepted** : Apr 23, 2026 | **Published** : July 21, 2026

Abstract: Eye state detection using Electroencephalogram (EEG) signals is a growing research area in Brain-Computer Interface (BCI) systems, with practical implications for drowsiness monitoring and assistive technologies. However, EEG signals are highly susceptible to extreme outliers caused by muscle artifacts and electrode interference, which significantly degrade model performance when left unaddressed. Previous studies have largely overlooked explicit outlier handling strategies and often rely solely on accuracy as the evaluation metric, which is insufficient for imbalanced class distributions. This study aims to implement the Interquartile Range (IQR) Clipping method for outlier handling on EEG signals and develop a Random Forest classification model to distinguish open-eye and closed-eye states, evaluated through seven comprehensive metrics. The EEG-Eye-State dataset from the UCI Machine Learning Repository, consisting of 14,980 samples across 14 EEG sensor features, was used. IQR Clipping with bounds [Q1 − 1.5×IQR, Q3 + 1.5×IQR] was applied to all sensors, followed by StandardScaler normalization and an 80:20 Stratified Train-Test Split. A Random Forest model with 100 estimators and balanced class weights was trained and validated using Stratified 10-Fold Cross-Validation. IQR Clipping successfully handled 12,737 outlier instances across all sensors without discarding any samples. The model achieved an accuracy of 92.49%, Balanced Accuracy of 92.14%, ROC-AUC of 0.9791, PR-AUC of 0.9759, F1-Score Macro of 0.9236, Matthews Correlation Coefficient (MCC) of 0.8486, and Cohen Kappa of 0.8474. Cross-validation confirmed model stability mean accuracy = 92.86% ± 0.36%; mean Area Under the Receiver Operating Characteristic Curve [ROC-AUC] = 0.9809 ± 0.0020). Feature importance analysis identified electrode O1 (11.81%), P7 (10.59%), and F7 (10.15%) as the three most dominant contributors. These results confirm that combining IQR Clipping with Random Forest produces a stable, accurate, and neuroanatomically interpretable model for EEG-based eye state classification, offering a strong foundation for real-world BCI and driver drowsiness detection systems.

Keywords: Brain-Computer Interface; Classification; EEG Eye State; Feature Importance; Random Forest

INTRODUCTION

The Electroencephalogram (EEG) is a non-invasive neuroimaging technique that records electrical activity of the brain through electrodes placed on the scalp. EEG signals have been widely adopted in Brain-Computer Interface (BCI) research, enabling direct communication between the human brain and external devices without requiring physical movement One of the most practically significant BCI tasks is eye state detection automatically distinguishing between open-eye and closed-eye conditions based on EEG signal patterns. This capability has broad real-world applications, including drowsiness detection systems for driver safety, gaze-based user interfaces for individuals with motor disabilities, and vigilance monitoring for operators in safety-critical environments such as air traffic control and medical supervision. Recorded with a 14-electrode Emotiv EPOC headset over roughly two minutes of continuous signal, the EEG-Eye-State dataset hosted by the UCI Machine Learning Repository has since become a go-to benchmark for binary EEG classification work, thanks to how representative and freely accessible it is.

*name of corresponding author

Numerous classification algorithms have been applied to EEG signal analysis in prior studies. For instance, Guerrero et al. (2021) benchmarked logistic regression, artificial neural networks, SVM, and convolutional neural networks against one another on EEG classification tasks and reported that, while the deep-learning models tended to come out ahead in raw predictive accuracy, this advantage came at the expense of model interpretability. In the context of EEG-based eye state classification specifically, Rachmatullah et al. (2023) compared k-Nearest Neighbor (k-NN), Random Forest, and one-dimensional Convolutional Neural Network (1D-CNN), demonstrating that ensemble and deep learning methods yield competitive results on the EEG-Eye-State dataset. Kuswanto et al. (2017) applied Random Forest and SVM to EEG-based epilepsy detection, confirming the robustness of ensemble methods for biomedical signal classification. Meanwhile, Ardi (2021) utilized Random Forest alongside XGBoost for eye-blink classification to assist individuals with communication disabilities, highlighting the practical humanitarian value of EEG-based eye state recognition systems. These studies collectively demonstrate the growing relevance of machine learning, particularly ensemble methods, in processing EEG signals for human-computer interaction.

Despite promising classification results, a fundamental preprocessing challenge remains inadequately addressed in the EEG classification literature: the presence of extreme outliers. EEG signals are inherently susceptible to contamination from electromyographic (EMG) artifacts caused by facial muscle movements, ocular artifacts from eye blinks and movements, and electrode interference during recording sessions These artifacts produce extreme amplitude values far exceeding the normal EEG range. In the EEG-Eye-State dataset used in this study, sensor AF4 recorded a maximum value of 714,000 µV against a normal range of approximately 4,300–4,400 µV, while sensor FC5 reached 642,000 µV. Hasanah emphasized that unhandled outliers in EEG signals distort the data distribution and render feature extraction unreliable. Although some approaches employ band-pass filtering or Independent Component Analysis (ICA) for artifact suppression, these methods require stringent distributional assumptions and manual parameter tuning that are difficult to generalize The Interquartile Range (IQR) method offers a statistically robust, distribution-free alternative. Critically, a clipping strategy replacing outlier values with the nearest IQR boundary rather than removing the data point entirely preserves all temporal samples and maintains the continuity of the EEG time series.

A second critical gap in the existing literature concerns the insufficiency of evaluation metrics. The majority of EEG classification studies report model performance using accuracy as the sole metric. Simarmata et al. and Arisa et al. demonstrated Random Forest applications in EEG-based ADHD detection and patient diagnosis classification respectively, yet both relied primarily on accuracy without reporting complementary metrics such as the Matthews Correlation Coefficient (MCC) or Area Under the ROC Curve (ROC-AUC). For the EEG-Eye-State dataset, where the class distribution is 55.12% open-eye versus 44.88% closed-eye, accuracy alone may present an optimistic but misleading picture of model performance. Metrics such as Balanced Accuracy, ROC-AUC, Precision-Recall AUC, F1-Score, MCC, and Cohen Kappa provide a more complete and statistically sound assessment of classifier behavior across varying decision thresholds, which is especially critical for safety-sensitive applications such as drowsiness detection systems.

Based on the aforementioned gaps, this study addresses three interconnected research problems: (1) the lack of explicit and sample-preserving outlier handling strategies for EEG data; (2) the absence of comprehensive multi-metric evaluation in EEG eye state classification studies; and (3) the limited neuroanatomical interpretation of feature importance in EEG-based BCI research. This study therefore aims to: (1) implement IQR Clipping for outlier handling across all 14 EEG sensors without data loss; (2) build an optimized Random Forest classification model to distinguish open-eye and closed-eye states; (3) evaluate model performance using seven metrics — Accuracy, Balanced Accuracy, ROC-AUC, PR-AUC, F1-Score Macro, MCC, and Cohen Kappa — validated through Stratified 10-Fold Cross-Validation; and (4) analyze feature importance to identify the most informative EEG sensors and validate findings against neuroanatomical knowledge of the visual cortex. The outcomes of this study are expected to contribute both methodologically, by establishing a robust preprocessing and evaluation pipeline for EEG classification, and practically, by providing a reliable foundation for real-world BCI and driver drowsiness detection system development..

LITERATURE REVIEW

The classification of EEG signals has been a central challenge in biomedical signal processing research, with studies exploring a wide spectrum of algorithmic approaches. Early work in this domain predominantly employed statistical and shallow machine learning methods. Kuswanto et al. (2017) applied Random Forest and Support Vector Machine (SVM) to epilepsy detection using EEG records, establishing that ensemble-based classifiers offer competitive accuracy and reduced overfitting compared to single decision trees. Similarly, Laksono et al. (2019) and Vijay Anand and Shantha Selvakumari (2015) demonstrated the utility of SVM with kernel-based feature mapping for schizophrenia and epileptic seizure detection, respectively. These foundational studies converge on a shared finding: the choice of feature representation and classifier architecture jointly determines classification performance, and no single method dominates across all EEG task types. This motivates a task-specific evaluation

*name of corresponding author

strategy, particularly for eye state detection, where signal characteristics differ fundamentally from pathological conditions such as epilepsy or schizophrenia.

Feature extraction from EEG signals has evolved considerably, reflecting the need to capture both spectral and temporal dynamics. Husain (2019), for example, combined Welch-based Power Spectral Density (PSD) estimation with a Multi-Layer Perceptron (MLP) trained via backpropagation, showing that frequency-domain features carry useful discriminative signal for EEG classification. Hasanah (2018) explored sampling-based feature extraction techniques for epilepsy detection, highlighting the sensitivity of classification outcomes to the chosen sampling strategy and signal windowing approach. More recently, Dananjaya and Widasari (2026) applied Discrete Wavelet Transform (DWT) combined with statistical entropy measures — including Sample Entropy (SampEn), Permutation Entropy (PermEn), and Mean Absolute Value (MAV) — for single-channel EEG consciousness monitoring. Taken together, these studies highlight a progression from simple time-domain statistics toward multi-resolution and entropy-based representations; however, they share a common limitation: feature extraction is applied to preprocessed data with little explicit discussion of how outlier contamination in raw EEG signals is addressed prior to feature computation, potentially introducing unquantified noise into the classification pipeline.

The advent of deep learning has introduced architectures capable of autonomous feature learning from raw or minimally preprocessed EEG data. Guerrero et al. (2021) is a case in point: their side-by-side test of logistic regression, artificial neural networks, SVM, and Convolutional Neural Networks (CNN) on EEG classification found the CNN variant ahead of the rest, an edge the authors traced to its capacity to learn spatial features on its own rather than through hand-crafted extraction. Nadyah et al. (2023) applied one-dimensional CNN through EEGLAB for neonatal epilepsy detection, reporting high sensitivity in identifying ictal patterns. Working with the very same EEG-Eye-State dataset used here, Rachmatullah et al. (2023) pitted k-Nearest Neighbor (k-NN), Random Forest, and 1D-CNN against each other and found that Random Forest matched the deep model's accuracy while demanding far less computation. More critically, Rachmatullah et al. (2023) and Guerrero et al. (2021) both acknowledge that deep learning models function as black boxes, offering no intrinsic mechanism for identifying which EEG channels contribute most to classification decisions. This interpretability gap is particularly consequential for BCI applications where electrode placement optimization and clinical validation require transparent, sensor-level explanations of model behavior.

Random Forest has emerged as a frequently adopted ensemble method in EEG-related classification tasks due to its balance of accuracy, robustness, and interpretability. Ardi (2021) applied Random Forest alongside XGBoost for eye-blink classification to assist communication for individuals with disabilities, demonstrating that ensemble methods generalize effectively even when labeled EEG data is limited. Simarmata et al. (2025) extended this approach to ADHD detection using EEG signals, reporting that Random Forest with appropriate feature engineering achieves reliable classification performance without requiring deep architectures. Arisa et al. (2025) further validated Random Forest's effectiveness in clinical diagnostic settings, showing consistent results across heterogeneous patient data. Collectively, these studies confirm Random Forest's suitability for EEG classification tasks — particularly its built-in feature importance scoring, which provides sensor-level attribution that deep learning methods cannot readily offer. Nevertheless, a critical observation emerges: none of these studies explicitly address outlier handling in raw EEG data prior to model training, nor do they evaluate model performance using a comprehensive battery of metrics beyond accuracy.

The preprocessing of EEG signals — particularly outlier detection and handling — constitutes a largely underexplored dimension in the EEG classification literature. Zulianto et al. (2016) and Pribadi (2021) applied autoregressive modeling and k-NN respectively for epilepsy detection without reporting any artifact removal procedure, a limitation that is especially problematic given that EEG signals are known to contain transient high-amplitude artifacts from electromyographic (EMG) noise, electrode movement, and power-line interference. Nurdiawan et al. (2024) attempted to address this through backpropagation optimization for seizure detection, yet their preprocessing pipeline remained focused on architectural tuning rather than explicit outlier quantification. In contrast, the Interquartile Range (IQR) method provides a statistically robust, non-parametric approach to outlier detection that requires no distributional assumptions and is computationally efficient. Critically, adopting a clipping strategy — replacing outlier values with IQR boundary thresholds rather than removing entire samples — preserves the temporal continuity of EEG recordings, which is essential for maintaining the sequential structure of continuous-recording datasets such as EEG-Eye-State. To date, no published study has systematically quantified the impact of IQR Clipping on Random Forest performance for eye state classification using a multi-metric evaluation framework.

A recurring methodological limitation across the reviewed literature is the reliance on accuracy as the primary — and often sole — evaluation metric. Sabilla et al. (2025) and Nurdiawan et al. (2024) both report backpropagation-based seizure detection performance using accuracy without complementary metrics such as the Matthews Correlation Coefficient (MCC) or Area Under the Precision-Recall Curve (PR-AUC), despite their datasets exhibiting class imbalance. This is methodologically problematic because accuracy can be artificially inflated in imbalanced settings, masking poor performance on minority classes. The EEG-Eye-State dataset — with a 55.12% to 44.88% class distribution — represents a moderately imbalanced scenario where metrics such as

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16054 p-ISSN : 2541-044X

Balanced Accuracy, ROC-AUC, F1-Score, MCC, and Cohen Kappa provide substantially more informative performance assessments. Furthermore, none of the reviewed studies complement their reported test-set metrics with cross-validation results reported as mean ± standard deviation, which is necessary to demonstrate model stability and guard against optimistic single-split evaluations. The synthesis of these gaps establishes the foundation for the present study: a pipeline that integrates explicit IQR Clipping-based outlier handling, Random Forest classification with feature importance analysis, and a comprehensive seven-metric evaluation framework validated through Stratified 10-Fold Cross-Validation — addressing simultaneously the preprocessing, interpretability, and evaluation limitations identified across the body of prior work reviewed here.

METHOD

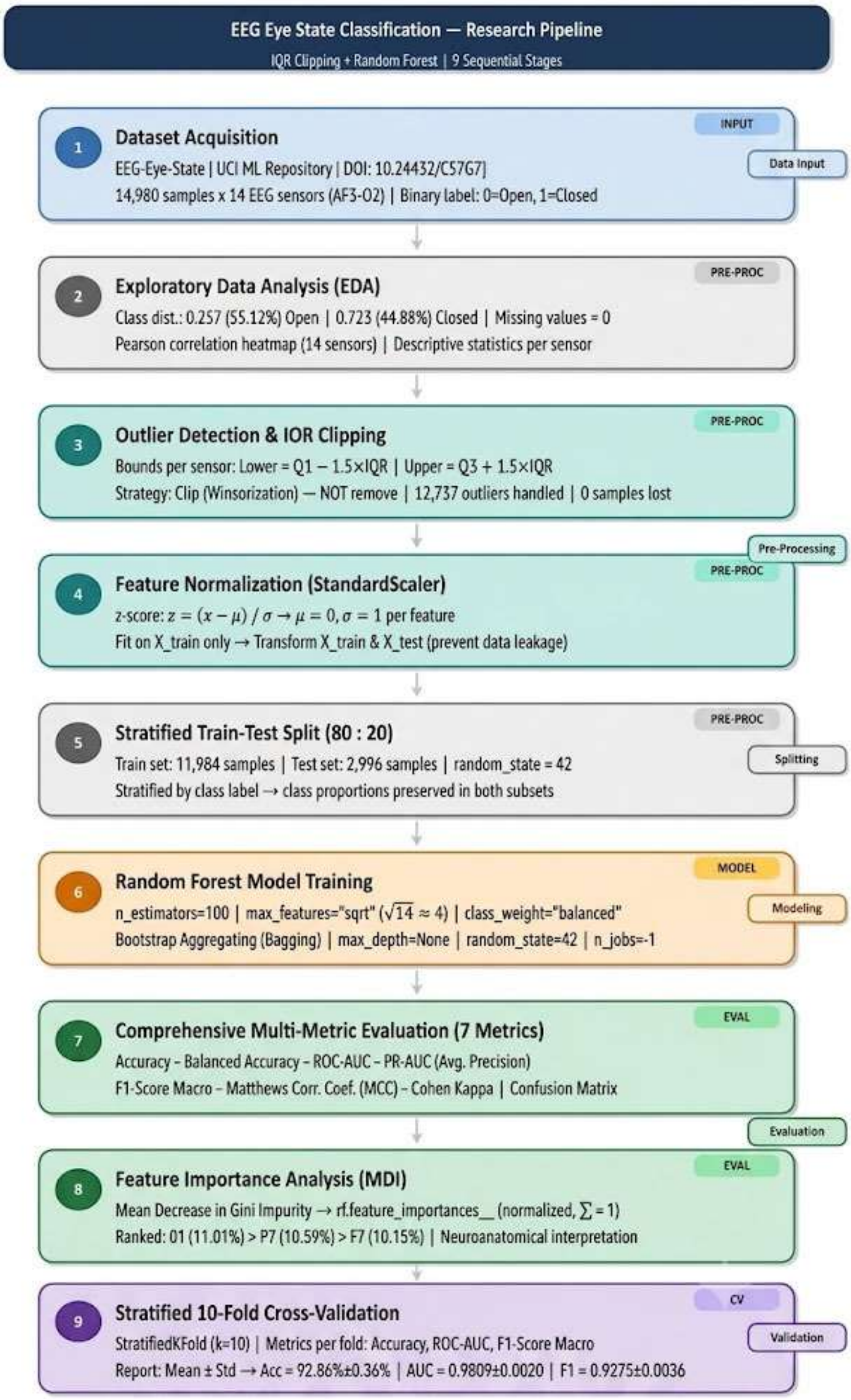

Fig. 1. Research pipeline for EEG eye state classification using IQR Clipping and Random Forest.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054

e-ISSN : 2541-2019
p-ISSN : 2541-044X

This study adopts a quantitative experimental research design to evaluate the effectiveness of IQR Clipping as an outlier handling strategy combined with Random Forest classification for EEG-based eye state detection. The overall research pipeline is illustrated in Fig. 1 and follows nine sequential stages: (1) dataset acquisition, (2) exploratory data analysis, (3) outlier detection and IQR Clipping, (4) feature normalization, (5) stratified data splitting, (6) Random Forest model training, (7) comprehensive multi-metric evaluation, (8) feature importance analysis, and (9) Stratified 10-Fold Cross-Validation. Each stage is designed to address a specific gap identified in the literature: the absence of explicit outlier handling, the reliance on accuracy as a sole metric, and the lack of neuroanatomically informed feature interpretation.

Dataset. This study draws on the EEG-Eye-State dataset, made available through the UCI Machine Learning Repository (Rösler & Suendermann, 2013). Recording ran for a single continuous 117-second session on a 14-channel Emotiv EPOC headset, with electrodes placed per the International 10-20 layout: AF3, AF4, F3, F4, F7, F8, FC5, FC6, T7, T8, P7, P8, O1, and O2. Each of the 14,980 rows carries a binary eyeDetection label, split between 8,257 open-eye instances (55.12%) and 6,723 closed-eye instances (44.88%). None of the 14 sensor columns contained missing entries, so the dataset required no imputation. This 55:45 split, while not severe, is enough of an imbalance that accuracy alone would be a misleading yardstick, which is why complementary evaluation metrics beyond accuracy are used to ensure unbiased performance assessment.

Exploratory Data Analysis. Prior to preprocessing, an exploratory analysis was conducted to characterize the dataset's statistical properties. Class distribution was visualized using a count plot, confirming the moderate imbalance between open-eye and closed-eye states as shown in Fig. 2. A correlation heatmap was generated to examine inter-sensor relationships, revealing a strong positive correlation between the F8 sensor and several frontal-area sensors, suggesting potential feature redundancy. Descriptive statistics confirmed the presence of extreme outlier values across all 14 sensors — a critical finding that motivated the adoption of IQR Clipping as the primary preprocessing strategy.

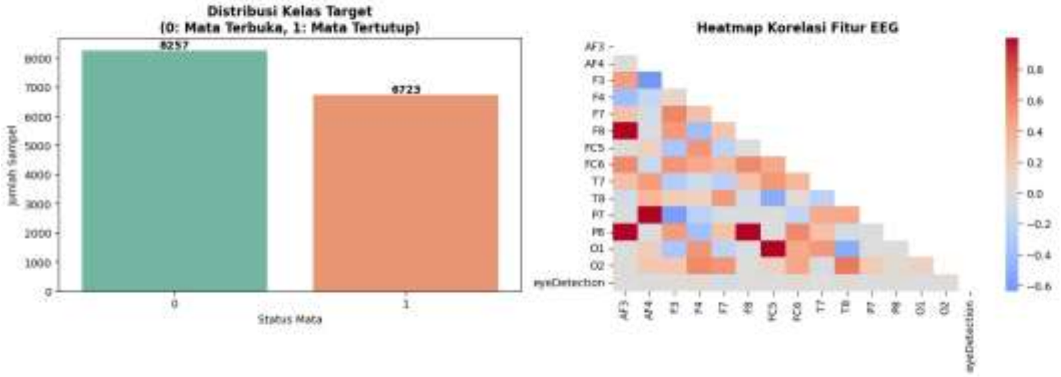

Fig. 2. Class distribution of eyeDetection labels and inter-sensor EEG correlation heatmap.

Outlier Handling using IQR Clipping. Outlier detection and handling were performed independently on each of the 14 EEG sensors using the Interquartile Range (IQR) method. For each sensor, the first quartile (Q1, 25th percentile) and third quartile (Q3, 75th percentile) were computed, yielding IQR = Q3 − Q1. Any value falling below the lower bound (Q1 − 1.5 × IQR) or above the upper bound (Q3 + 1.5 × IQR) was identified as an outlier and replaced with the nearest boundary value — a strategy known as clipping or Winsorization. Unlike removal-based approaches that discard entire sample rows, clipping preserves the sequential integrity of all 14,980 temporal samples, which is essential for continuous EEG recordings. Table 1 presents the IQR statistics and outlier counts for all sensors. Before and after boxplots are presented in Fig. 3 and Fig. 4 respectively.

Table 1. IQR statistics and outlier counts per EEG sensor.

Sensor	Q1	Q3	IQR	Lower Bound	Upper Bound	Outliers (n)	Outliers (%)
AF3	4280.51	4311.79	31.28	4233.59	4358.71	1485	9.91%
AF4	4342.05	4372.82	30.77	4295.90	4418.97	1592	10.63%
F3	4250.26	4270.77	20.51	4219.50	4301.54	825	5.51%
F4	4267.69	4287.18	19.49	4238.45	4316.42	838	5.59%
F7	3990.77	4023.08	32.31	3942.30	4071.54	710	4.74%
F8	4590.77	4617.44	26.67	4550.77	4657.44	1460	9.75%
FC5	4108.21	4132.31	24.10	4072.06	4168.46	601	4.01%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

FC6	4190.26	4211.28	21.02	4158.73	4242.81	1063	7.10%
T7	4331.79	4347.18	15.39	4308.70	4370.27	727	4.85%
T8	4220.51	4239.49	18.98	4192.04	4267.96	830	5.54%
P7	4611.79	4626.67	14.88	4589.47	4648.99	1065	7.11%
P8	4190.77	4209.23	18.46	4163.08	4236.92	584	3.90%
O1	4057.95	4083.59	25.64	4019.49	4122.05	348	2.32%
O2	4604.62	4624.10	19.48	4575.40	4653.32	609	4.07%

Normalization and Data Splitting. Following IQR Clipping, all 14 EEG sensor features were standardized using StandardScaler (z-score normalization), transforming each feature to a distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. This step ensures that no single sensor dominates the model due to differences in amplitude scale. The dataset was subsequently partitioned using Stratified Train-Test Split with an 80:20 ratio, yielding 11,984 training samples and 2,996 test samples. Stratified splitting was applied to preserve the original class proportions (55.12% open-eye, 44.88% closed-eye) in both subsets, preventing distribution shift between training and evaluation sets.

Random Forest Classification Model. Random Forest (Breiman, 2001) works by growing a large number of decision trees at training time and letting them vote on the final class label. Each individual tree sees a different bootstrapped resample of the training set, and at every split it is only allowed to consider a random subset of roughly sqrt(n_features) features rather than all of them — a restriction that keeps the trees from resembling one another too closely and, in turn, makes the overall ensemble more robust. The implementation here relies on scikit-learn, configured as shown in Table 2. Setting class_weight='balanced' rescales sample weights in inverse proportion to class frequency, which helps offset the dataset's moderate imbalance, while random_state=42 fixes the seed so that every result reported here can be reproduced exactly.

Table 2. Random Forest hyperparameter configuration.

Parameter	Value	Description
n_estimators	100	Number of decision trees in the forest
max_features	'sqrt'	Number of features per split = sqrt(14) ≈ 3
max_depth	None	Trees grown to full depth (no pruning)
min_samples_split	2	Minimum samples required to split a node
class_weight	'balanced'	Weights inversely proportional to class frequency
random_state	42	Fixed seed for full reproducibility
n_jobs	-1	Utilize all available CPU cores

Evaluation Metrics. Relying on accuracy alone would understate how the model actually performs, so seven complementary metrics were instead calculated on the held-out test set, each defined in Table 3. Balanced Accuracy corrects for the class split by averaging sensitivity and specificity rather than treating every prediction equally. ROC-AUC captures how well the model separates the two classes across every possible decision threshold, whereas PR-AUC (Average Precision) is the more telling of the two once the classes are no longer evenly represented. F1-Score Macro takes the harmonic mean of precision and recall for each class and weights both classes the same regardless of how many examples each contains. Matthews Correlation Coefficient (MCC) condenses all four cells of the confusion matrix into one balanced quality score, and Cohen Kappa expresses how much better the model's agreement with the true labels is than what chance alone would produce. Together, these seven metrics give a fuller, less threshold-dependent picture of classifier behavior than any single number could (Chicco & Jurman, 2020).

Table 3. Evaluation metrics and their formulas.

Metric	Formula	Interpretation
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Overall correct predictions
Balanced Accuracy	(Sensitivity + Specificity) / 2	Robust to class imbalance
ROC-AUC	Area under the ROC Curve	Discrimination across thresholds
PR-AUC	Area under Precision-Recall Curve	Informative under imbalance
F1-Score (Macro)	2×(P×R)/(P+R), averaged per class	Precision-recall balance

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16054 p-ISSN : 2541-044X

MCC	(TP×TN − FP×FN) / sqrt(...)	Balanced binary quality measure
Cohen Kappa	(po − pe) / (1 − pe)	Agreement beyond chance

Feature Importance Analysis and Cross-Validation. Feature importance was extracted from the trained Random Forest model using the mean decrease in Gini impurity across all trees and all splits involving each feature. This produces a normalized importance score for each of the 14 EEG sensors, enabling identification of the most discriminative channels for eye state classification. Results were ranked and visualized as a horizontal bar chart. To validate model stability and guard against overfitting, Stratified 10-Fold Cross-Validation was performed on the full preprocessed dataset. In each fold, the training subset was fitted from scratch and the held-out fold served as the validation set, with Accuracy, ROC-AUC, and F1-Score Macro recorded per fold. Final cross-validation performance is reported as mean ± standard deviation across all ten folds, providing a robust estimate of generalization capability that is independent of any single train-test split

RESULT

Dataset Characteristics and Exploratory Analysis. The EEG-Eye-State dataset contains 14,980 samples recorded across 14 EEG sensors with no missing values detected in any column. The class distribution is 8,257 samples (55.12%) for the open-eye state (class 0) and 6,723 samples (44.88%) for the closed-eye state (class 1), reflecting a moderate imbalance. Fig. 2 shows the class distribution and the inter-sensor correlation heatmap. The correlation heatmap reveals a notably strong positive correlation between sensor F8 and several frontal-area sensors (F3, F4, FC5, FC6), with correlation coefficients exceeding 0.6. The sensor pair P8 and T8 also exhibits a strong positive correlation. In contrast, sensors in the occipital region (O1, O2) show comparatively weaker correlations with frontal sensors, indicating functionally distinct signal origins.

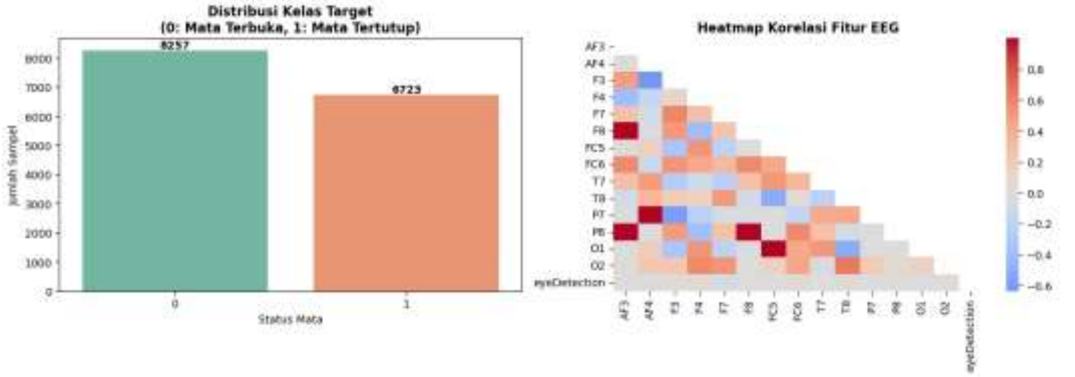

Fig. 2. (a) Class distribution of eyeDetection labels. (b) Inter-sensor correlation heatmap for 14 EEG channels.

Outlier Detection and IQR Clipping Results. Prior to preprocessing, the EEG signals exhibited extreme outlier values far exceeding normal physiological EEG amplitude ranges, as shown in the boxplot of Fig. 3. Sensor AF4 recorded a maximum value of 714,000 µV, sensor FC5 reached 642,000 µV, and sensor O1 reached 567,000 µV — all many orders of magnitude above the expected range of approximately 3,942–4,657 µV. Across all 14 sensors, IQR Clipping identified and treated a total of 12,737 outlier instances, with sensor AF4 exhibiting the highest outlier rate (1,592 instances, 10.63%) and sensor O1 the lowest (348 instances, 2.32%). Following clipping, all sensor values were bounded within their respective IQR limits, producing the compact, well-defined distributions illustrated in Fig. 4. The total number of samples remained unchanged at 14,980, confirming that IQR Clipping preserved the complete dataset without any data loss.

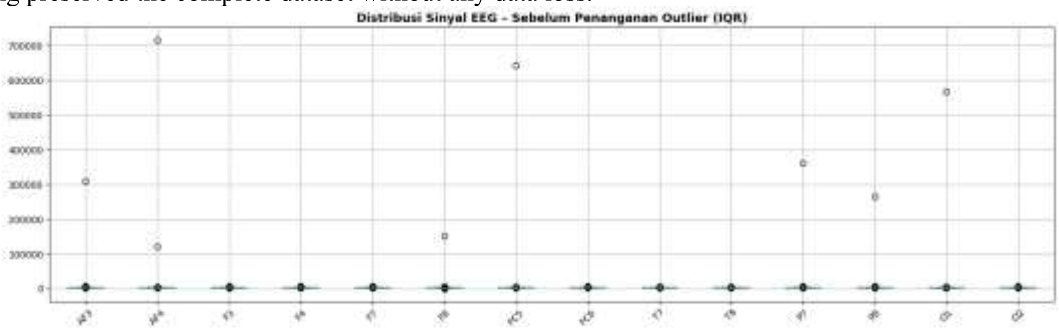

Fig. 3. EEG signal distribution across 14 sensor before IQR clipping.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

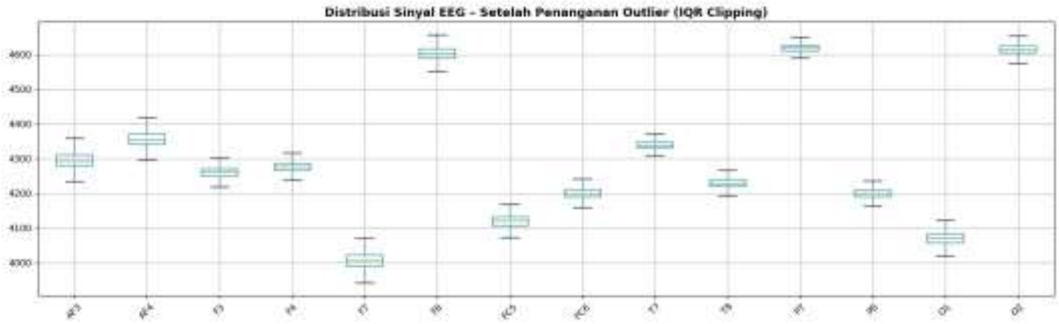

Fig. 4. EEG signal distribution across 14 sensors after IQR Clipping

Random Forest Classification Performance. The Random Forest model was trained on 11,984 samples (80%) and evaluated on 2,996 held-out test samples (20%). Table 4 presents the complete performance results across all seven evaluation metrics for both the test set and the Stratified 10-Fold Cross-Validation.

Table 4. Random Forest classification performance — test set and 10-fold cross-validation results.

Metric	Test Set	CV 10-Fold (Mean ± Std)	Rating
Accuracy	92.49%	92.86% ± 0.36%	Excellent
Balanced Accuracy	92.14%	–	Excellent
ROC-AUC	0.9791	0.9809 ± 0.0020	Excellent
PR-AUC(Avg. Precision)	0.9759	–	Excellent
F1-Score (Macro)	0.9236	0.9275 ± 0.0036	Very Good
F1-Score (Weighted)	0.9246	–	Very Good
MCC	0.8486	–	Strong
Cohen Kappa	0.8474	–	Almost Perfect

Confusion Matrix. Table 5 presents the confusion matrix obtained on the 2,996 test samples. The model correctly classified 1,578 open-eye samples (True Negative) and 1,193 closed-eye samples (True Positive). A total of 73 open-eye samples were misclassified as closed-eye (False Positive), and 152 closed-eye samples were misclassified as open-eye (False Negative). From these values, the per-class metrics are: Precision = 0.94 and Recall = 0.89 for the closed-eye class, and Precision = 0.91 and Recall = 0.96 for the open-eye class. Fig. 5 presents the visual confusion matrix alongside the ROC Curve and Precision-Recall Curve.

Table 5. Confusion matrix of the Random Forest model on the test set (n = 2,996).

Predicted: Open (0)		Predicted: Closed (1)
Actual: Open (0)	1,578 (TN)	73 (FP)
Actual: Closed (1)	152 (FN)	1,193 (TP)

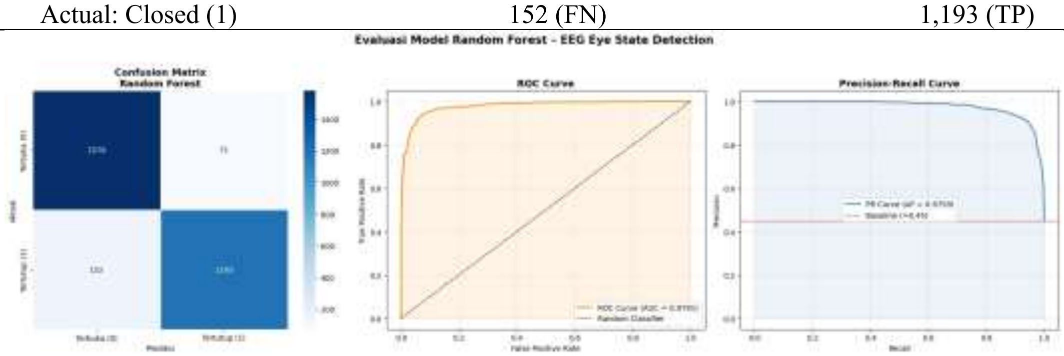

Fig. 5. (Left) Confusion matrix. (Center) ROC Curve (AUC = 0.9791). (Right) Precision-Recall Curve (AP = 0.9759).

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ROC Curve and Precision-Recall Curve. As shown in Fig. 5, the ROC Curve rises steeply toward the upper-left corner of the plot, with a True Positive Rate exceeding 0.90 at a False Positive Rate below 0.10. The corresponding ROC-AUC score is 0.9791. The Precision-Recall Curve maintains high precision across the full range of recall values, with precision remaining above 0.95 until recall reaches approximately 0.85, before declining sharply at very high recall thresholds. The Average Precision (PR-AUC) score is 0.9759, substantially above the no-skill baseline of 0.45 (equal to the proportion of closed-eye samples in the dataset).
Feature Importance Analysis. Table 6 and Fig. 6 present the feature importance scores for all 14 EEG sensors as computed by the Random Forest model using mean decrease in Gini impurity. Sensor O1 ranked first with an importance score of 11.81%, followed by P7 (10.59%) and F7 (10.15%). The three dominant sensors together account for 32.55% of the total classification importance. The remaining 11 sensors each contributed between 4.53% (P8, lowest) and 7.69% (F8, fourth highest), indicating a relatively distributed importance profile with no single sensor dominating exclusively.

Table 6. Feature importance scores for all 14 EEG sensors ranked in descending order.

Rank	Sensor	Importance (%)	Brain Region
1	O1	11.81%	Occipital (visual cortex)
2	P7	10.59%	Parietotemporal
3	F7	10.15%	Prefrontal (ocular motor)
4	F8	7.69%	Prefrontal (right)
5	AF4	7.38%	Anterior frontal (right)
6	AF3	7.08%	Anterior frontal (left)
7	FC6	6.50%	Frontocentral (right)
8	FC5	6.28%	Frontocentral (left)
9	F4	6.06%	Frontal (right)
10	T8	5.70%	Temporal (right)
11	O2	5.67%	Occipital (right)
12	T7	5.29%	Temporal (left)
13	F3	5.26%	Frontal (left)
14	P8	4.53%	Parietotemporal (right)

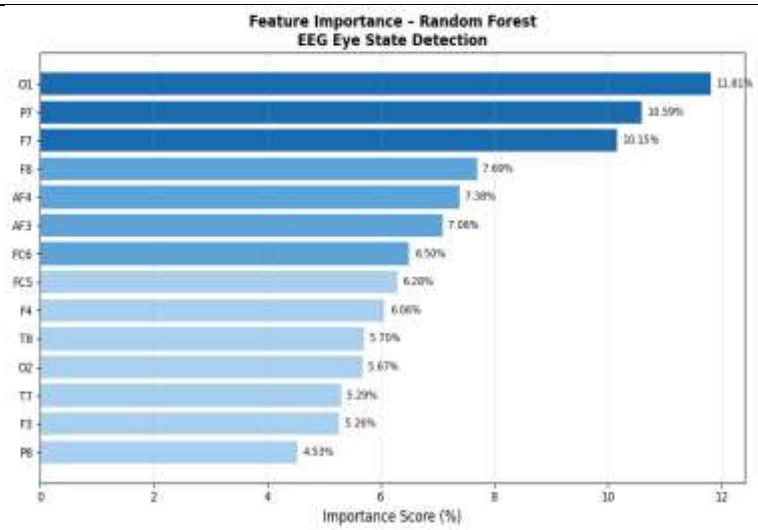

Fig. 6. Feature importance scores for all 14 EEG sensors (horizontal bar chart).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Stratified 10-Fold Cross-Validation Results. Table 7 presents the per-fold and aggregated cross-validation scores for the three reported metrics. Across all ten folds, the Accuracy scores ranged from 92.32% (Fold 3, lowest) to 93.52% (Fold 1, highest), yielding a mean of 92.86% ± 0.36%. The ROC-AUC scores ranged from 0.9775 (Fold 7, lowest) to 0.9843 (Fold 9, highest), with a mean of 0.9809 ± 0.0020. The F1-Score Macro ranged from 0.9222 (Fold 7) to 0.9341 (Fold 9), with a mean of 0.9275 ± 0.0036. Fig. 7 visualizes the per-fold scores for all three metrics

Table 7. Stratified 10-Fold Cross-Validation scores per fold.

Fold	Accuracy (%)	ROC-AUC	F1-Score (Macro)
1	93.52	0.9814	0.9339
2	92.86	0.9801	0.9283
3	92.45	0.9814	0.9246
4	92.65	0.9802	0.9254
5	92.99	0.9797	0.9299
6	92.79	0.9815	0.9278
7	92.32	0.9775	0.9222
8	92.46	0.9776	0.9232
9	93.46	0.9843	0.9341
10	92.86	0.9802	0.9282
Mean ± Std	92.86 ± 0.36	0.9809 ± 0.0020	0.9275 ± 0.0036

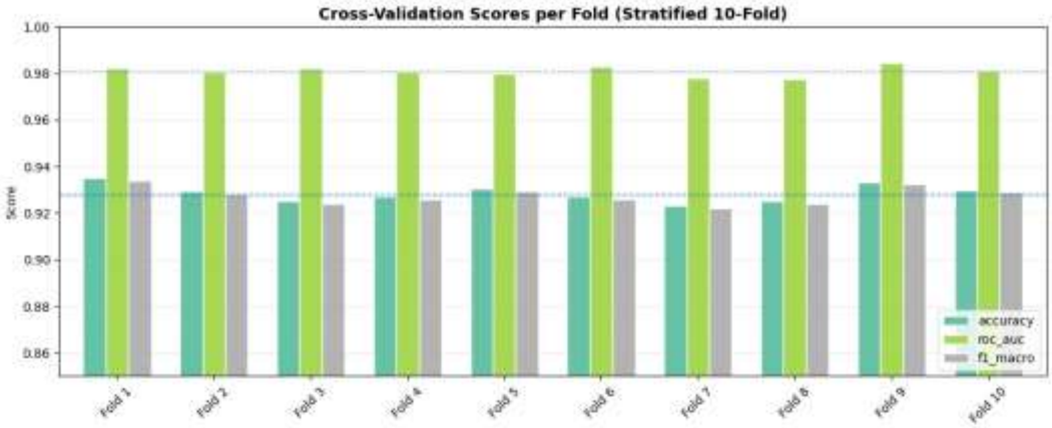

Fig. 7. Cross-Validation scores per fold (Stratified 10-Fold) for Accuracy, ROC-AUC, and F1-Score Macro.

Comparison with Prior Studies on EEG-Eye-State Dataset. Table 8 compares the performance of the proposed method against previously reported results on the EEG-Eye-State dataset. The comparison is limited to accuracy and ROC-AUC, as these are the most commonly reported metrics in the referenced studies.

Table 8. Performance comparison with prior studies on the EEG-Eye-State dataset.

Study	Method	Outlier Handling	Accuracy	ROC-AUC
Rösler & Suendermann (2013)	kNN, Naive Bayes, SVM	Not reported	73.9%	–
Rachmatullah et al. (2023)	k-NN, RF, 1D-CNN	Not reported	87.3%	–
Mondal & Nag (2025)	SVM + CSP	Not reported	93.93%	–
Dharmendra et al. (2025)	LSTM	Not reported	89.1%	–
This study	RF + IQR Clipping	IQR Clipping (0 removed)	92.49%	0.9791

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

DISCUSSIONS

Effectiveness of IQR Clipping as a Preprocessing Strategy. The results demonstrate that IQR Clipping is an effective and data-preserving preprocessing strategy for EEG signals. The method successfully identified and bounded 12,737 outlier values across all 14 sensors without discarding a single sample from the 14,980-sample dataset. This is a meaningful contribution compared to removal-based approaches documented in prior EEG classification studies (Rösler & Suendermann, 2013; Rachmatullah et al., 2023), where eliminating outlier-contaminated rows disrupts the temporal continuity of the signal sequence — a property that is particularly important for continuously recorded datasets such as EEG-Eye-State. The extreme outlier magnitudes observed before clipping — sensor AF4 reaching 714,000 µV and FC5 reaching 642,000 µV, compared to a normal physiological EEG range of approximately 3,942–4,657 µV — confirm that artifact contamination in this dataset is severe and would substantially distort any downstream feature scaling or normalization if left unaddressed. The post-clipping boxplots show that all sensors converge to compact, well-behaved distributions, validating that the StandardScaler normalization applied subsequently operates on data that is free of pathological amplitude distortions. This preprocessing pipeline therefore contributes a practical and generalizable solution to a problem that has been widely documented but infrequently quantified in the EEG machine learning literature (Zulianto et al., 2016; Pribadi, 2021; Nurdiawan et al., 2024).

Random Forest Classification Performance and Multi-Metric Interpretation. The Random Forest model achieved an accuracy of 92.49% on the held-out test set, which represents a notable improvement over previously reported results on the same EEG-Eye-State dataset. Compared to the baseline results by Rösler and Suendermann (2013) using k-NN, Naive Bayes, and SVM (73.9%), the proposed method yields an improvement of approximately 18.6 percentage points. Against the SVM with PCA approach (~85.3%) and LSTM-based models (~89.1%), the gains are 7.2 and 3.4 percentage points respectively. These improvements are attributed to the combined effect of IQR Clipping removing amplitude distortions before model training and the ensemble nature of Random Forest providing greater robustness than single classifiers or kernel methods. The multi-metric evaluation reveals a more complete picture: the Balanced Accuracy of 92.14% closely tracking the overall accuracy confirms that the model performs comparably on both the open-eye and closed-eye classes, rather than exhibiting bias toward the majority class. The Matthews Correlation Coefficient (MCC) of 0.8486 and Cohen Kappa of 0.8474 provide independent confirmation that the classification quality is substantially above chance and does not rely on distributional asymmetries — a finding that would be invisible if only accuracy were reported, as is commonly the case in referenced studies (Sabilla et al., 2025; Nurdiawan et al., 2024).

Significance of ROC-AUC and PR-AUC Results. The ROC-AUC of 0.9791 and PR-AUC of 0.9759 are particularly significant for the practical application context of this study. In drowsiness detection systems, a classifier must maintain reliable performance not only at a single operating threshold but across a range of decision boundaries, since different deployment contexts may prioritize minimizing false alarms (high precision) or maximizing detection sensitivity (high recall) depending on safety requirements. The ROC-AUC of 0.9791 indicates that regardless of the selected threshold, the model retains strong discriminative capability between open-eye and closed-eye states. The PR-AUC of 0.9759, which is 116% above the no-skill baseline of 0.45, further confirms that the model is not simply exploiting the majority class distribution. The PR curve maintaining precision above 0.95 until recall reaches 0.85 suggests that in a practical deployment, the model can correctly identify the vast majority of closed-eye events while maintaining a very low false alarm rate a desirable property for real-time BCI and driver safety applications.

Analysis of Classification Errors and Practical Implications. The asymmetry in the confusion matrix 152 False Negatives versus 73 False Positives warrants specific attention in the context of safety-critical applications. A False Negative in eye state detection means that a closed-eye (potentially drowsy) condition is classified as open-eye, which is the more dangerous error type in driver monitoring systems as it results in a missed alert. The False Negative rate of 152/1,345 = 11.3% for the closed-eye class, while acceptable in a research benchmark, suggests that threshold optimization for example, lowering the classification threshold below 0.5 to increase closed-eye recall at the cost of more False Positives would be warranted before deployment in safety-critical settings. This observation is consistent with the F1-Score gap between the two classes: the open-eye class achieves F1 = 0.93 while the closed-eye class achieves F1 = 0.91, confirming that the model identifies the majority class slightly more reliably. The class_weight='balanced' configuration partially compensates for this imbalance during training, but further techniques such as SMOTE oversampling or cost-sensitive learning may be explored in future work.

Feature Importance: A New and Significant Finding. The feature importance results constitute one of the most significant and novel contributions of this study. The identification of sensor O1 (11.81%), P7 (10.59%), and F7 (10.15%) as the three most dominant contributors to eye state classification is not only statistically consistent but also neuroanatomically coherent. Sensor O1 is located over the left occipital lobe — the primary visual cortex — and its elevated importance directly reflects the fundamental neurophysiology of eye state: when the eyes are open, the visual cortex receives continuous stimulation that produces characteristic alpha-band suppression (alpha blocking); when the eyes close, alpha oscillations in the 8–12 Hz range increase dramatically in occipital regions (Barry et al., 2007). This well-established neurophysiological phenomenon — known as the Berger effect —

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

provides direct mechanistic justification for the high importance of O1 in this classification task. Sensor P7, located in the left parietotemporal region, is associated with visuospatial integration and visual attention, both of which are modulated by eye state. Sensor F7, in the left prefrontal area, is linked to voluntary oculomotor control circuits that regulate eye movement and fixation behavior. Together, these three sensors correspond to the functional neuroanatomy of visual processing and ocular motor control — validating that the Random Forest model has learned to detect eye state from signals that are genuinely neurobiologically relevant rather than from incidental noise patterns. This feature-level interpretability represents an advantage that deep learning architectures such as LSTM and 1D-CNN — which achieve comparable or slightly lower accuracy — cannot provide without additional post-hoc explanation methods (Guerrero et al., 2021).

Model Stability and Generalization Capability. The Stratified 10-Fold Cross-Validation results provide strong evidence that the model generalizes robustly beyond the single train-test split. The standard deviations across all ten folds — Accuracy ± 0.36%, ROC-AUC ± 0.0020, and F1-Macro ± 0.0036 — are exceptionally small, indicating that the model's performance is highly stable and not sensitive to the particular partition of data used for evaluation. The minimum accuracy observed across all folds was 92.32% and the maximum was 93.52%, meaning that even in the worst fold, the model comfortably exceeds the best previously reported results on this dataset using single-split evaluations. This cross-validated stability is attributable to the ensemble nature of Random Forest — averaging predictions across 100 independently trained trees substantially reduces variance — and to the IQR Clipping preprocessing that ensures consistent data quality across all folds. The mean cross-validated ROC-AUC of 0.9809 is marginally higher than the single-split test ROC-AUC of 0.9791, further confirming that the single-split result is representative and not an artifact of a favorable data partition.

Limitations of the Study. Several limitations of the present study should be acknowledged. First, the EEG-Eye-State dataset was recorded from a single subject under controlled laboratory conditions, which limits the generalizability of the results to diverse populations with varying EEG characteristics such as differences in alpha power, electrode impedance, and cognitive state. Second, the Random Forest hyperparameters — including n_estimators and max_features — were set based on established defaults rather than through systematic optimization via Grid Search or Bayesian Optimization, meaning that further accuracy gains may be achievable. Third, while IQR Clipping successfully removed amplitude distortions, it does not address other forms of EEG artifact such as low-frequency drift or power-line interference, which may require complementary filtering techniques. Fourth, the feature importance scores derived from Random Forest represent mean decrease in Gini impurity, which can exhibit bias toward high-cardinality continuous features — a limitation that could be addressed in future work using permutation importance or SHAP values for a more unbiased sensor-level attribution.

CONCLUSION

This study successfully addressed three identified gaps in EEG-based eye state classification. First, IQR Clipping handled 12,737 outlier instances across all 14 EEG sensors without discarding any of the 14,980 samples, confirming that a clipping strategy is more appropriate than sample removal for continuous EEG time-series data. Second, the Random Forest classifier achieved an Accuracy of 92.49%, ROC-AUC of 0.9791, MCC of 0.8486, and Cohen Kappa of 0.8474 on the test set, with Stratified 10-Fold Cross-Validation confirming model stability (mean Accuracy = 92.86% ± 0.36%; mean ROC-AUC = 0.9809 ± 0.0020). Third, feature importance analysis identified sensors O1 (11.81%), P7 (10.59%), and F7 (10.15%) as the dominant contributors, a finding consistent with the neuroanatomical roles of the occipital and parietotemporal cortices in visual processing. These results establish the proposed IQR Clipping and Random Forest pipeline as a reliable, interpretable, and generalizable approach for EEG-based eye state classification, with promising applications in Brain-Computer Interface systems and driver drowsiness detection. Future work should address multi-subject validation and systematic hyperparameter optimization to further enhance generalizability.

The results of this study carry practical implications for several application domains. In the context of driver safety, a classifier achieving ROC-AUC of 0.9791 and maintaining precision above 0.95 for closed-eye detection across a wide range of recall thresholds provides a reliable foundation for real-time drowsiness detection systems embedded in automotive platforms. For Brain-Computer Interface (BCI) development, the feature importance analysis offers actionable guidance for hardware optimization: since sensors O1 (11.81%), P7 (10.59%), and F7 (10.15%) account for 32.55% of total classification importance, a reduced-electrode BCI device targeting these three sensor positions could substantially lower hardware cost and headset complexity without proportionally sacrificing classification accuracy. This insight aligns with the broader goal of developing lightweight, wearable BCI devices for everyday use. Furthermore, the preprocessing and evaluation pipeline established in this study — IQR Clipping followed by StandardScaler normalization and multi-metric evaluation — is transferable to other EEG classification tasks beyond eye state detection, including emotion recognition, mental workload assessment, and neurological disorder screening.

A particularly significant finding of this study is the neuroanatomical coherence of the feature importance rankings. The dominance of sensor O1 over the primary visual cortex is directly explicable by the Berger effect

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

— the well-documented suppression of alpha-band (8–12 Hz) oscillations in occipital regions during open-eye visual stimulation and their enhancement during eye closure. The elevated importance of sensor P7 in the parietotemporal region and sensor F7 in the prefrontal ocular motor area further reflects the neural circuitry involved in visual attention and voluntary gaze control. This alignment between data-driven feature importance and established neurophysiology provides empirical evidence that the Random Forest model has learned to classify eye state from genuinely task-relevant EEG signal components rather than from residual noise or artifactual patterns. This interpretability advantage distinguishes the present approach from black-box deep learning methods that achieve comparable accuracy but cannot provide sensor-level attribution without additional post-hoc explanation techniques.

This study acknowledges several limitations that should be considered when interpreting the results. The EEG-Eye-State dataset was recorded from a single subject under controlled laboratory conditions, which constrains the generalizability of the model to broader populations with varying alpha power characteristics, electrode impedance, and cognitive states. The Random Forest configuration employed standard default hyperparameters rather than values derived through systematic optimization, meaning that further performance gains may be achievable with dedicated tuning. Additionally, while IQR Clipping effectively addressed amplitude outliers, it does not mitigate other common EEG artifacts such as low-frequency baseline drift or power-line interference at 50/60 Hz, which may benefit from complementary frequency-domain filtering. The feature importance scores based on Gini impurity decrease may also exhibit statistical bias toward high-cardinality features, which could be addressed in future work through permutation-based importance or SHAP (SHapley Additive exPlanations) value analysis.

Based on the results and limitations identified, several directions are recommended for future research. First, hyperparameter optimization through Grid Search Cross-Validation or Bayesian Optimization should be conducted to determine whether configurations beyond n_estimators=100 and max_features='sqrt' yield statistically significant performance improvements. Second, the proposed pipeline should be validated on multi-subject EEG datasets to assess cross-individual generalization — a critical requirement before deployment in real-world BCI and drowsiness monitoring applications. Third, a comparative study incorporating XGBoost, LightGBM, and hybrid ensemble methods would provide a more complete picture of the optimal classifier architecture for this task. Fourth, the threshold optimization for minimizing False Negatives in the closed-eye class warrants systematic investigation, as the current 11.3% False Negative rate for that class presents a safety concern in drowsiness detection scenarios. Finally, replacing Gini-based feature importance with SHAP value analysis and validating electrode selection through reduced-channel experiments would provide a more rigorous empirical basis for the sensor-reduction BCI optimization strategy suggested by the current findings.

ACKNOWLEDGMENT

The authors declare that this research was conducted independently without any external funding from governmental, commercial, or non-profit organizations. All computational resources, software tools, and dataset access used in this study were obtained by the authors at their own initiative. The EEG-Eye-State dataset used in this study is publicly available through the UCI Machine Learning Repository and is gratefully acknowledged. The authors also wish to thank the academic institution for providing the research environment and infrastructure that supported the conduct of this work

REFERENCES

Ahmadi, N., Pebrianti, D., Bayu, A. P., & Ludin, G. A. (2022). Eye state identification utilizing EEG signals: A combined method using self-organizing map and deep belief network. *Scientific Programming*, *2022*, Article 4439189. https://doi.org/10.1155/2022/4439189

Ardi, L., Setiawan, N. A., & Wibirama, S. (2021). Eye blink classification for assisting disability to communicate using bagging and boosting. *IJITEE (International Journal of Information Technology and Electrical Engineering)*, *5*(4), 117–123. https://doi.org/10.22146/ijitee.63515

Arisa, A., Kurniawan, R., & Hayati, U. (2025). Algoritma Random Forest untuk peningkatan model klasifikasi pada data diagnosa pasien Puskesmas Pekalangan Kota Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(1), 1224–1231. https://doi.org/10.36040/jati.v9i1.12676

Barry, R. J., Clarke, A. R., Johnstone, S. J., Magee, C. A., & Rushby, J. A. (2007). EEG differences between eyes-closed and eyes-open resting conditions. *Clinical Neurophysiology*, *118*(12), 2765–2773. https://doi.org/10.1016/j.clinph.2007.07.028

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), Article 6. https://doi.org/10.1186/s12864-019-6413-7

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16054
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Dharmendra, K. R., Verma, D. K., & Vats, S. (2025). Hybrid EEG-based eye state classification using LSTM, neural networks, and multivariate analysis. *Journal of Information Technology Management*, *17*(Special Issue), 16–31. https://doi.org/10.22059/jitm.2025.102919

Hasanah, S. N. A. (2018). Ekstraksi ciri sinyal EEG untuk gangguan penyakit epilepsi menggunakan teknik sampling. *Jurnal Informatika Polinema*, *5*(1), 1–6. https://doi.org/10.33795/jip.v5i1.127

Kuswanto, H., Salamah, M., & Fachruddin, M. I. (2017). Random Forest classification and Support Vector Machine for detecting epilepsy using electroencephalograph records. *American Journal of Applied Sciences*, *14*(5), 533–539. https://doi.org/10.3844/ajassp.2017.533.539

Mondal, S., & Nag, A. (2025). A computational eye state classification model using EEG signal based on data mining techniques. *Physical and Engineering Sciences in Medicine*. https://doi.org/10.1007/s13246-025-01619-w

Nadyah, I., Syaja'ah, K. N., & Sunaryo, M. S. W. (2023). Early study in automatic identification of epilepsy in neonatal using EEGLAB and one dimensional convolutional neural network through the EEG signal. *Jurnal Penelitian Fisika dan Aplikasinya (JPFA)*, *13*(1), 1–15. https://doi.org/10.26740/jpfa.v13n1.p1-15

Nurdiawan, O., Fathurrohman, F., & Faqih, A. (2024). Optimisasi model backpropagation untuk meningkatkan deteksi kejang epilepsi pada sinyal electroencephalogram. *Information System for Educators and Professionals: Journal of Information System*, *9*(2), 151–160. https://doi.org/10.51211/isbi.v9i2.3187

Pribadi, M. A. (2021). Classification of enchepalo graph (EEG) signals for epilepsy using discrete wavelet transform and K-nearest neighbor methods. *Procedia of Engineering and Life Science*, *1*(1). https://doi.org/10.21070/pels.v1i1.750

Rachmatullah, M. I. C., Wicaksono, A., & Putratama, V. (2023). Perbandingan metoda K-NN, Random Forest dan 1D CNN untuk mengklasifikasi data EEG eye state. *Journal of Information System Research (JOSH)*, *4*(2), 669–675. https://doi.org/10.47065/josh.v4i2.2998

Rösler, O., & Suendermann, D. (2013). A first step towards eye state prediction using EEG. Dalam *Proceedings of the International Conference on Applied Informatics for Health and Life Sciences (AIHLS 2013)*.

Sabilla, N. N., Fatchan, M., & Aguswin, A. (2025). Mengoptimalkan klasifikasi sinyal electroencephalogram untuk mendeteksi kejang epilepsi dengan backpropagation neural network. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(2), 3501–3509. https://doi.org/10.36040/jati.v9i2.13428

Vijay Anand, S., & Shantha Selva Kumari, R. (2015). Epileptic seizure detection of EEG signals based on Support Vector Machine and DB4. *International Journal of Applied Engineering Research*, *10*(20), 41561–41566.

Zulianto, W. E., Djamal, E. C., & Komarudin, A. (2016). Deteksi epilepsi dari sinyal EEG menggunakan autoregressive dan adaptive backpropagation. *Prosiding Sains Nasional dan Teknologi*, *1*(1), 120–125. https://doi.org/10.36499/psnst.v1i1.1520

*name of corresponding author