ISSN: 1672 - 6553
JOURNAL OF DYNAMICS AND CONTROL
VOLUME 10 ISSUE 04: P1-13

EEG EYE STATE CLASSIFICATION: INVARIANT AND INTERPRETABLE WITH CROSS DATASET ADAPTATION AND NEUROPHYSIOLOGICAL CONSTRAINTS

Taouli Sidi Ahmed

Biomedical Engineering Research Laboratory, Department of Biomedical Genius, Faculty of Technology, University Aboubekr Belkaid University Aboubekr Belkaid, P.O. Box 230, Chetouane, Tlemcen - 13000, Algeria

JOURNAL OF DYNAMICS AND CONTROL
https://doi.org/10.71058/jodac.v10i04001

EEG EYE STATE CLASSIFICATION: INVARIANT AND INTERPRETABLE WITH CROSS DATASET ADAPTATION AND NEUROPHYSIOLOGICAL CONSTRAINTS

Taouli Sidi Ahmed

Biomedical Engineering Research Laboratory,
Department of Biomedical Genius, Faculty of Technology,
University Aboubekr Belkaid
University Aboubekr Belkaid, P.O. Box 230, Chetouane, Tlemcen 13000, Algeria

Correspondence: sidiahmed.taouli@univ-tlemcen.dz
Received on 02-12-2025; Accepted on: 20-03-2026; Published on 03-04-2026

Abstract: *Electroencephalography (EEG)-based eye state classification is a key task in neuroscience and biomedical engineering, widely employed in cognitive monitoring and brain computer interface systems. While many studies report high classification accuracy within a single dataset, performance often deteriorates across different subjects or databases due to inter-subject variability and heterogeneous acquisition conditions. In this study, we introduce a subject-invariant and interpretable approach for EEG eye state classification that explicitly addresses cross-database generalisation. Our approach combines multi-domain EEG features with correlation alignment (CORAL) for domain adaptation and integrates neurophysiological constraints grounded in established EEG principles, including alpha-band reactivity, functional connectivity, and signal complexity. We evaluated the proposed method on three real-world universal EEG databases: PhysioNet, SPIS, and MNNIT. In cross-database evaluations, conventional models achieved accuracies of 63.4–66.1%, revealing limited generalisation. Incorporating CORAL improved performance to 72.1–74.8%, and the addition of neurophysiological constraints further increased accuracy to 81.5–83.2%. Notably, interpretability was maintained through SHAP-based feature importance analysis, identifying alpha-band power, occipital-parietal coherence, and entropy as the most influential features. These findings demonstrate that combining domain adaptation with physiologically informed constraints substantially enhances cross-database EEG eye-state classification, providing a robust and clinically interpretable solution for real-world EEG applications.*

Keywords: *Electroencephalography (EEG); Eye-state classification; CORAL; Neurophysiological constraints; Model interpretability.*

1. Introduction

Electroencephalography (EEG) is a widely used non-invasive technique for recording brain electrical activity through electrodes positioned on the scalp. These signals primarily capture the synchronised firing of large populations of neurons and offer excellent temporal resolution, making EEG an essential tool in neuroscience research, clinical diagnostics, and brain–computer interface (BCI) systems [1,2]. Its portability, cost-effectiveness, and robustness contribute to its widespread adoption for monitoring brain dynamics.

EEG signals are typically analysed within canonical frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (>30 Hz), each linked to specific cognitive and physiological processes [1]. Among these, alpha rhythms are particularly significant in resting state activity and visual information processing, and they are closely modulated by eye state.

A fundamental EEG classification problem is distinguishing between Eyes Open (EO) and Eyes Closed (EC) conditions. This task relies on the well-documented phenomenon of alpha blocking, whereby alpha power increases during eye closure and decreases upon eye opening, especially in occipital and parietal regions [3]. Because of this strong physiological basis, EO/EC classification serves as a benchmark for evaluating EEG signal processing and machine learning approaches [4-10].

Despite numerous studies reporting high intra-dataset accuracies, often exceeding 95% [4,11-13], achieving robust generalisation across different subjects and databases remains challenging. For instance, Laport et al. [11] proposed an eye-state identification method based on the Discrete Wavelet Transform (DWT), achieving promising accuracy on individual datasets by effectively capturing time–frequency features of EEG signals. Reddy

https://doi.org/10.71058/jodac.v10i04001

et al. [12] employed deep learning architectures for online eye-state recognition, demonstrating high accuracy and fast classification performance, though their models were sensitive to inter-subject variability. Narejo et al. [4] implemented Deep Belief Networks (DBN) and Stacked Autoencoders (SAE) for EEG-based eye-state classification, achieving strong intra-dataset performance but exhibiting noticeable degradation when applied to unseen subjects or datasets. More recently, Elshewey et al. [13] proposed an EEG-based eye-state classification approach optimised with a modified BER metaheuristic algorithm, achieving state-of-the-art accuracy on benchmark datasets; however, similar to previous methods, their approach was primarily validated under intra-dataset conditions, leaving cross-subject and cross-database generalisation largely unexplored.

EEG signals are highly individual and sensitive to recording conditions such as electrode placement, acquisition hardware, and experimental protocol.

Consequently, models trained on a single dataset often exhibit degraded performance when applied to unseen subjects or databases, significantly limiting their practical applicability. These limitations emphasise the necessity of developing robust and generalisable models capable of maintaining high performance across varying subjects and acquisition conditions.

To address these issues, recent studies have increasingly employed deep learning architectures, particularly convolutional neural networks (CNNs), which can automatically extract spatial and temporal features from raw or minimally processed EEG data [2,13]. These models have shown improved performance compared to traditional feature-based classifiers within intra-dataset evaluations. However, deep learning systems generally function as black boxes, providing limited interpretability regarding the neurophysiological mechanisms underlying their decisions. Consequently, their alignment with established EEG principles, such as alpha-band modulation and connectivity patterns, is often weak.

This lack of transparency raises concerns about model reliability, interpretability, and clinical applicability, particularly in cross-database contexts.

Models developed without integrating physiological knowledge may yield high accuracy in controlled settings but fail to produce meaningful or generalisable decision patterns in real-world conditions. Ignoring neurophysiological constraints can result in spurious feature importance and inconsistent classification behaviour when facing distributional shifts across databases. Therefore, approaches that integrate neurophysiological priors with accuracy optimisation are needed to ensure biologically plausible representations. Another challenge is the variability between datasets.

Differences in recording hardware, electrode configurations, sampling frequencies, and experimental procedures introduce significant domain shifts, which can impair model performance when transferring across databases. This situation emphasises the importance of combining physiologically informed modelling with domain adaptation techniques to reduce distribution mismatches while maintaining interpretability.

The proposed method is evaluated under realistic and challenging conditions, emphasising cross-database testing using multiple real-world universal EEG databases. Performance is assessed using complementary metrics, including accuracy, sensitivity, specificity, and area under the ROC curve, to provide a comprehensive evaluation. Additionally, the contribution of each component CORAL based domain adaptation [14] and neurophysiological constraint integration is analysed to highlight their individual and synergistic impact on model performance. This design demonstrates both the robustness and generalisation capabilities of the approach while reinforcing its relevance for interpretable EEG applications in line with recent explainable machine learning advances.

The remainder of this paper is structured as follows: Section 2 presents the methodology, including EEG databases, preprocessing steps, multi-domain feature extraction, domain

https://doi.org/10.71058/jodac.v10i04001

adaptation, and neurophysiological constraint integration. Section 3 details the experimental setup and evaluation metrics.

Section 4 discusses the results with a focus on cross-database performance and interpretability analysis. Finally, Section 5 concludes and outlines directions for future work.

2. Materials and methods

The proposed methodology is designed to achieve robust, invariant, and interpretable EEG based eye state classification across heterogeneous datasets. As depicted in Figure 1, the overall process follows a unified processing sequence composed of five sequential stages.

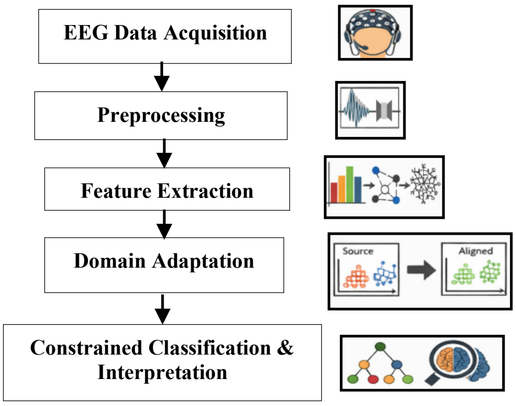

Figure 1. Block diagram of the proposed EEG eye state classification method

First, EEG signals are collected from multiple publicly available datasets recorded under different experimental conditions, ensuring variability in subjects, electrode configurations, and acquisition protocols. Second, a standardised preprocessing procedure is applied to improve signal quality and reduce inter-dataset variability, including artefact suppression, band-pass filtering, and signal normalisation. Third, multi-domain feature extraction is performed to capture complementary neurophysiological information from temporal, spectral, and time–frequency representations, with particular emphasis on features known to be sensitive to eye-opening and eye-closure states. Fourth, a cross-dataset domain adaptation strategy is incorporated to address distributional shifts between training and testing datasets, there by promoting feature invariance and improving generalisation performance in unseen domains. Finally, classification is conducted under explicit neurophysiological constraints, enabling accurate eye state discrimination while preserving interpretability by relating model outputs to established EEG phenomena, such as alpha-band power modulation associated with visual input suppression.

2.1 EEG Datasets

To ensure diversity in subject populations, recording hardware, and acquisition protocols, three publicly available EEG datasets were selected for this study. The main characteristics of these datasets are summarised in Table 1. All datasets were acquired under resting-state conditions and include clearly annotated Eyes Open (EO) and Eyes Closed (EC) segments, providing a common experimental paradigm across heterogeneous recording settings. The PhysioNet EEG Motor Movement dataset comprises recordings from 20 subjects acquired using 64 EEG channels at a sampling rate of 160 Hz [15]. The SPIS Resting-State EEG dataset [16-17]; contains EEG recordings from 18 subjects, acquired using 64 channels and sampled at a higher frequency of 512 Hz. This dataset captures resting-state brain activity during eyes-open and eyes-closed periods, making it

https://doi.org/10.71058/jodac.v10i04001

suitable for analyses of spatio-spectral EEG features and vigilance variability. In contrast, the MNNIT Eye State EEG dataset contains recordings from a smaller number of subjects, acquired using 14 EEG channels at a sampling rate of 256 Hz [18]. These datasets were chosen due to their widespread use in the literature and the availability of reliable eye-state annotations, which facilitate reproducible benchmarking and fair comparison with existing approaches. Moreover, the heterogeneity in recording configurations and signal characteristics allows for a rigorous evaluation of the robustness and generalisation capability of the proposed method across datasets, as detailed in Table 1.

Table 1. EEG datasets used in this study

Dataset	Subjects	Channels	Sampling Rate	Experimental Task
Physionet	20	64	160 Hz	EO / EC
SPIS	18	64	512 Hz	EO / EC
MNNIT	2	14	256 Hz	EO / EC

2.2 EEG Preprocessing

EEG preprocessing was standardised across all datasets to minimise methodological bias. Acquired signals were first band-pass filtered between 1 and 45 Hz using a zero-phase Butterworth filter to remove baseline drift and high-frequency noise. A 50 Hz notch filter was subsequently applied to suppress power-line interference. Signals were then average-referenced to reduce spatial bias and segmented into 4-second epochs with 50% overlap, a duration considered sufficient to capture stable spectral features while maintaining adequate temporal resolution [19]. Epochs contaminated by excessive artefacts were automatically rejected based on amplitude thresholding.

2.3 Multi Domain Feature Extraction

To preserve neurophysiological interpretability, only features with established physiological relevance were extracted. For each 4 second epoch, a total of 53 features were computed and organised into five domains, as summarised in Table 2.

Table 2. Multi domain EEG features

Domain	Extracted Features
Time domain	Variance- standard deviation skewness-kurtosis
Frequency domain	Relative power in δ, θ, α, β, γ bands
Non-linear	Sample entropy, multiscale entropy
Time–frequency	Discrete wavelet (db4) sub-band energies
Connectivity	Magnitude-squared coherence- PLV- wPLI

- Time domain: variance, standard deviation, skewness, and kurtosis, capturing the instantaneous distribution, dispersion, and asymmetry of EEG signals. These statistical measures provide information on signal amplitude fluctuations associated with eye-state changes.

- Frequency domain: relative power in the δ (1–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30–45 Hz) bands, computed using the Welch periodogram method. These features reflect classical brain oscillatory activity and are sensitive markers of eye opening and closure, particularly in the alpha band.

- Non-linear domain: sample entropy and multiscale entropy, quantifying the temporal complexity and irregularity of EEG signals. Sample entropy was computed with

https://doi.org/10.71058/jodac.v10i04001

a pattern length of m = 2 and tolerance r = 0.2 × standard deviation, while multiscale entropy was evaluated over five scales to capture signal dynamics at multiple temporal resolutions [20-21].

• Time–frequency domain: sub-band energies derived from discrete wavelet decomposition using the Daubechies 4 (db4) wavelet up to level 5. The decomposition allowed separation of EEG signals into physiologically relevant frequency bands, and the energy of each sub-band was computed as the sum of squared wavelet coefficients, providing localised spectral information over time [22].

• Connectivity domain: functional connectivity features were extracted to characterise inter-regional EEG interactions, including magnitude-squared coherence, phase-locking value (PLV), and weighted phase-lag index (wPLI). Magnitude-squared coherence was computed between all pairs of channels to assess linear coupling within each frequency band [23]. PLV quantifies phase synchronisation between EEG signals [24], while wPLI mitigates the effects of volume conduction and common sources by emphasising non-zero phase-lag interactions, providing more robust connectivity estimates [25].

The selection of these features was motivated by neurophysiological evidence demonstrating that alpha-band activity, entropy measures, and functional connectivity patterns are strongly modulated by eye-state conditions [26]. This comprehensive multi-domain feature representation ensures physiological relevance and interpretability while supporting robust cross-dataset generalisation.

2.4 Cross Dataset Domain Adaptation

Comments EEG signals are inherently non-stationary and exhibit strong dataset dependency, which can hinder the generalisation of machine learning models across different recording conditions. To mitigate this issue, Correlation Alignment (CORAL) was employed to align the second-order statistics between the source and target feature spaces.

Given the source feature matrix X_S and the target feature matrix X_t, CORAL performs alignment as follows [27]:

$$X_S^{aligned} = X_S C_S^{-\frac{1}{2}} C_t^{\frac{1}{2}} \tag{1}$$

Where C_S and C_t denote the covariance matrices of the source and target datasets, respectively. This transformation adjusts the source feature distribution to match the covariance structure of the target dataset, effectively reducing inter-dataset statistical discrepancies.

The effectiveness of CORAL is illustrated in Figure 2, where the feature distributions are projected onto two dimensions using PCA for visualisation.

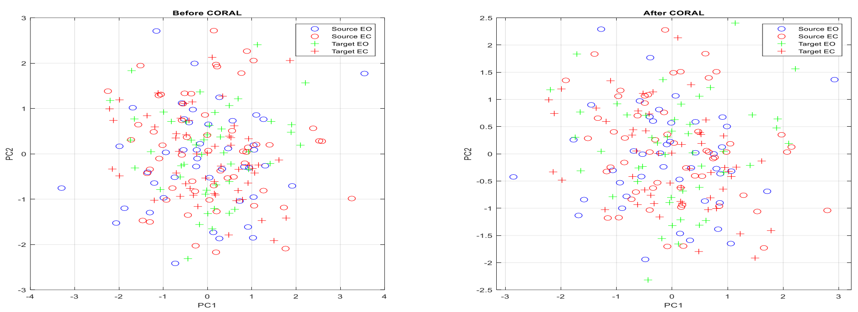

Figure 2. EEG feature distributions before and after CORAL

In the left subplot (Before CORAL), the source and target features are visibly misaligned, with significant overlap between different distributions and classes (EO and EC). After

https://doi.org/10.71058/jodac.v10i04001

applying CORAL (right subplot, After CORAL), the source features are transformed to better match the target distribution, resulting in a substantial reduction in inter dataset mismatch.

This visual evidence demonstrates that CORAL successfully reduces the distributional shift between datasets, allowing the classifier to learn more invariant and physiologically meaningful representations. Consequently, the method enhances cross-dataset generalisation without requiring labels from the target dataset, which is particularly valuable in EEG analysis, where annotation is costly and time-consuming.

2.5 Neurophysiologically informed classification strategy

To ensure that EEG classification is both accurate and physiologically meaningful, domain knowledge was explicitly integrated into the learning process.

Three well-established neurophysiological principles were incorporated as regularisation constraints during model training: (i) alpha-band power increases during eye closure [3], (ii) functional connectivity decreases during eye opening [26], and (iii) signal complexity, measured through entropy, increases during eye opening [28]. These constraints guide feature weighting and penalise classifier solutions that violate known EEG patterns, thereby enforcing physiological plausibility in the predictions.

Building on this neurophysiological guidance, a stacking ensemble classifier was adopted to balance robustness and interpretability. The ensemble combined Random Forest and XGBoost as base learners, with logistic regression serving as the meta-classifier. Model evaluation was performed using intra-dataset 5-fold cross-validation, Leave-One-Subject-Out Cross-Validation (LOSO-CV), and cross-dataset testing to assess generalisation across hete rogeneous recording conditions [29].

2.6 Proposed Methodology

To provide a unified view of the proposed methodology, this section summarises the complete EEG eye-state classification approach and its practical implementation, as illustrated in Figure 3.

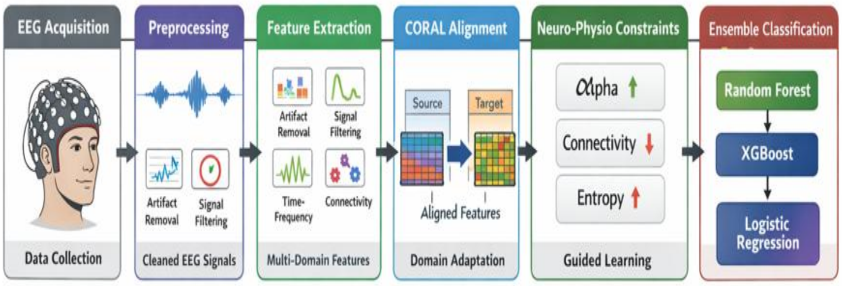

Figure 3. Proposed EEG eye state classification methodology

The method is designed as a modular, sequential process, allowing clear interpretation of each processing stage while maintaining flexibility for cross-dataset evaluation. Implementation begins with EEG data acquisition from heterogeneous public datasets, followed by standardised preprocessing to minimise inter-dataset variability. Rather than detailing individual preprocessing steps already described earlier, this stage ensures signal consistency in terms of frequency content, referencing, and temporal segmentation across datasets. Next, the extracted EEG segments are transformed into a comprehensive multi-domain feature representation, encompassing time–frequency, non-linear, and connectivity-based descriptors. These features collectively capture complementary aspects of EEG dynamics, including localised spectral energy, signal complexity, and inter-regional synchronisation patterns, thereby forming a high-dimensional yet physiologically meaningful feature space. To address domain shift arising from dataset heterogeneity,

https://doi.org/10.71058/jodac.v10i04001

Correlation Alignment (CORAL) is applied to the extracted feature space prior to classification. This alignment step reduces second-order statistical discrepancies between source and target datasets, facilitating more reliable transfer of learned decision boundaries across datasets. Building on the aligned feature representation, neurophysiological constraints are incorporated into the learning process to guide classifier behaviour toward biologically plausible solutions. These constraints act as soft regularisers, reinforcing known EEG patterns associated with eye-state modulation and enhancing both robustness and interpretability.

Finally, classification is performed using a stacking ensemble architecture, in which complementary base learners generate intermediate predictions that are integrated by a meta-classifier. This ensemble strategy improves generalisation by balancing model diversity, while evaluation is carried out using intra-dataset, inter-subject, and cross-dataset validation protocols to comprehensively assess performance under realistic deployment conditions.

3. Performance measures

The classification performance was evaluated using multiple complementary metrics to provide a comprehensive assessment:

- Accuracy (ACC): overall correctness of the predictions.
- Sensitivity (SE): proportion of correctly identified Eyes Closed (EC) epochs.
- Specificity (SP): proportion of correctly identified Eyes Open (EO) epochs.
- Area Under the ROC Curve (AUC): overall discriminative capability of the classifier.

These metrics are mathematically defined as follows:

$$\boldsymbol{Acc} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \tag{2}$$

$$\boldsymbol{SE} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \tag{3}$$

$$\boldsymbol{SP} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}} \tag{4}$$

where TP and TN denote true positives and true negatives, and FP and FN denote false positives and false negatives, respectively.

To assess statistical significance between different methods, the Wilcoxon signed-rank test was employed, with a significance level of $p<0.05$.

4. Results and discussion

4.1. Cross-Dataset Classification Accuracy

The primary goal of this study is to evaluate the robustness of EEG eye-state classification under cross-dataset conditions, which represent a realistic and challenging deployment context. The classification performance obtained when training on one dataset and testing on another is summarised in Table 3.

As reported in Table 3, baseline models exhibit limited generalisation capability, achieving accuracies of 66.1% for SPIS → PhysioNet and 63.4% for PhysioNet → SPIS. This performance degradation highlights the strong influence of inter-dataset variability, despite the apparent simplicity of the eye-state classification task.

When CORAL-based domain adaptation is applied, a clear improvement is observed in both cross-dataset configurations, with accuracy increasing to 74.8% and 72.1%, respectively. These results confirm that aligning second-order feature statistics effectively

https://doi.org/10.71058/jodac.v10i04001

reduces domain mismatch between heterogeneous EEG datasets.

Table 3. Cross dataset performance of baseline and adapted EEG models

Training→ Testing	Method	*Acc*	*Se*	*SP*	*AUC*
SPIS→ PhysioNet	Baseline	66.1	64.8	67.2	0.68
	+ CORAL	74.8	73.5	76.0	0.78
	+ CORAL + Neuro	**83.2**	**81.6**	**84.5**	**0.88**
PhysioNet → SPIS	Baseline	63.4	61.2	65.1	0.66
	+ CORAL	72.1	70.4	73.9	0.76
	+ CORAL + Neuro	**81.5**	**80.1**	**82.7**	**0.86**

The highest performance is achieved when neurophysiological constraints are combined with CORAL. As shown in Table 3, the proposed approach attains accuracies of 83.2% (SPIS → PhysioNet) and 81.5% (PhysioNet → SPIS), demonstrating a substantial and consistent improvement over both baseline and CORAL-only approaches across all evaluated performance metrics, including sensitivity, specificity, and AUC.

4.2. Per-Class Performance Assessment

Beyond overall accuracy, it is essential to assess the class-wise detection capability of the proposed method, particularly the balance between correctly identifying Eyes Closed (EC) and Eyes Open (EO) states. The sensitivity and specificity values obtained for the different adaptation strategies are illustrated in Figure 4 using grouped bar plots.

As shown in Figure 4, baseline classifiers exhibit only moderate performance, with sensitivity and specificity values of 64.8% and 67.2%, respectively. These results indicate an imbalance in class discrimination and unstable decision boundaries under cross-dataset conditions. The introduction of CORAL-based domain adaptation leads to noticeable improvements in both metrics, reflecting enhanced robustness and more consistent class separation.

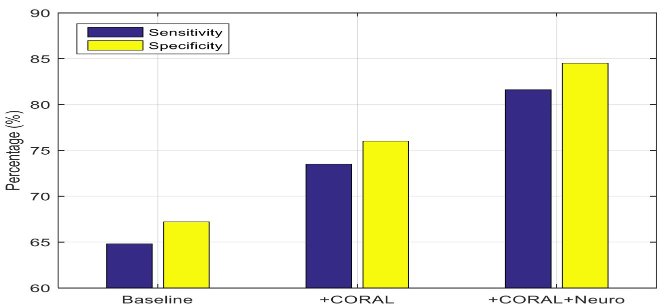

Figure 4. Performance of EEG eye-state classification

The proposed CORAL + Neuro model achieves the most balanced performance, with both sensitivity and specificity exceeding 80%, as illustrated in Figure 4. This balanced improvement is particularly important for practical EEG applications, where reliable detection of both eye states is crucial. In real-world cognitive monitoring and clinical

https://doi.org/10.71058/jodac.v10i04001

settings, misclassification of either EO or EC conditions may significantly affect subsequent analysis or decision making processes.

4.3 AUC-Based Comparative Evaluation

To further assess the discriminative power of the proposed method, the Area Under the ROC Curve (AUC) was analysed and summarised in Figure 5. As depicted in Figure 5, baseline models achieve an AUC of 0.68, reflecting limited separability between EO and EC classes under cross-dataset testing. CORAL-based adaptation improves the AUC to 0.78, confirming enhanced class discrimination.

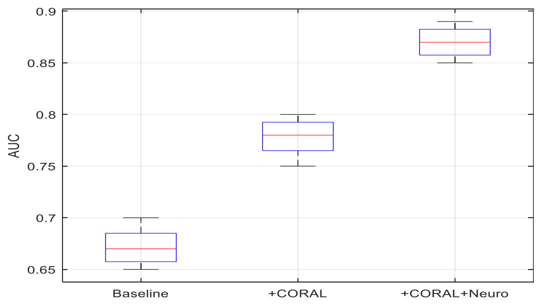

Figure 5. AUC comparison across methods

The proposed method achieves an AUC of 0.88, indicating strong and reliable separation between eye states. This result demonstrates that combining domain adaptation with neurophysiological constraints not only improves accuracy but also strengthens the overall discriminative structure of the feature space.

4.4 Quantitative AUC Gain Analysis

To better understand the contribution of each component of the proposed technique, the average accuracy gain introduced by CORAL and neurophysiological constraints is illustrated in Figure 6. As shown in Figure 6, CORAL alone contributes an average gain of approximately 8–10%, confirming its effectiveness in reducing dataset-induced variability. The addition of neurophysiological constraints provides a further gain of 7–9%, indicating that physiological priors offer complementary information beyond purely statistical alignment.

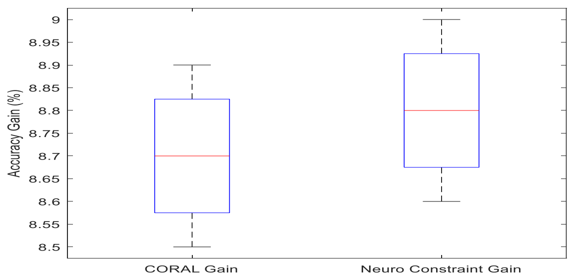

Figure 6. Comparison of Accuracy Gains

This analysis clearly demonstrates that domain adaptation and physiological constraints play distinct yet synergistic roles, resulting in robust cross-dataset performance.

4.5 Comparison with State of the Art Methods

To position the proposed procedure within the existing literature, a comparison with representative state-of-the-art methods is presented in Figure 7.

https://doi.org/10.71058/jodac.v10i04001

As illustrated in Figure 7, traditional machine learning methods (e.g., SVM) and deep learning approaches (e.g., CNN, DANN) exhibit limited performance under cross-dataset evaluation, with accuracies typically below 75%. In contrast, the proposed method achieves an accuracy of 83.2%, outperforming existing approaches while maintaining interpretability.

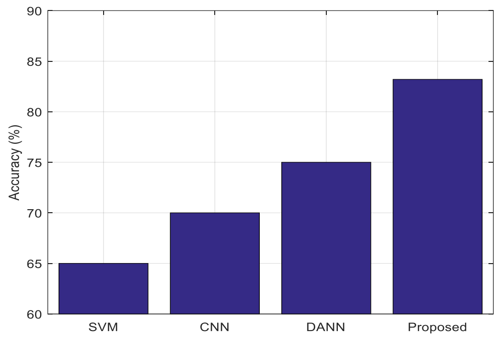

Figure 7. Accuracy comparison of the proposed method with other methods

This result highlights a key advantage of the proposed methodology: it achieves competitive or superior performance without relying on black-box deep learning models, making it more suitable for clinical and real world EEG applications.

4.6 Neurophysiological Interpretation and Discussion

The proposed method's performance stems from its ability to capture true neural biomarkers rather than statistical noise, with feature importance analysis highlighting three key neurophysiological pillars driving its decisions:

- Alpha Band Power and the Alpha Blocking Phenomenon: Relative power in the alpha band (8–13 Hz) emerged as the most influential feature. This aligns with the well-documented alpha blocking phenomenon, where the transition from eyes-open to eyes-closed states triggers a massive synchronisation of cortical neurons, significantly increasing alpha power in the occipital and parietal regions.

- Occipital–Parietal Functional Connectivity: The model relies heavily on connectivity metrics, specifically magnitude-squared coherence and the weighted phase-lag index (wPLI). This confirms that eye-state transitions involve a large-scale reorganisation of communication networks between visual and associative brain areas, rather than merely localized amplitude shifts.

- Signal Complexity and Entropy: High importance scores for sample and multiscale entropy validate the principle that signal complexity increases during the eyes-open state. This reflects the increased temporal irregularity of EEG signals associated with the active processing of external visual stimuli.

The results demonstrate a clear hierarchy in classification performance across heterogeneous datasets. Baseline models, which lack domain adaptation, show limited generalisation, with accuracies ranging between 63.4% and 66.1% when tested on unseen recording environments. These low figures highlight the sensitivity of traditional machine learning to differences in acquisition hardware and electrode configurations.

The integration of CORAL-based domain adaptation significantly mitigates these discrepancies by aligning the second-order statistics of the datasets, raising accuracy to 72.1–74.8%. However, the most robust results are achieved by the full proposed method, which reaches accuracies of 81.5–83.2% and an AUC of 0.88. By enforcing neurophysiological constraints, the model acts as a "physiologically informed" regulariser, penalising solutions that contradict established biological principles, such as a decrease in

https://doi.org/10.71058/jodac.v10i04001

alpha power during eye closure, thereby preventing the classifier from overfitting to spurious, non-biological patterns present in the training data.

Ultimately, this approach transforms the classification task from a "black-box" process into a transparent diagnostic tool. By maintaining high accuracy while remaining faithful to established cortical mechanisms, this methodology provides a reliable and clinically relevant solution for real-world applications in cognitive monitoring and brain–computer interfaces.

5. Conclusion

This study addressed one of the most critical challenges in EEG-based eye state classification: achieving robust generalisation across subjects and datasets while preserving physiological interpretability. Unlike most existing approaches that rely on intra-dataset evaluation and black-box learning models, we proposed a subject-invariant and interpretable approach that explicitly targets cross-dataset variability.

The proposed methodology integrates multi-domain EEG features, CORAL-based domain adaptation, and neurophysiological constraints derived from established EEG knowledge, including alpha-band reactivity, functional connectivity modulation, and signal complexity changes. This combination enables the model to simultaneously reduce statistical domain mismatch and enforce physiologically meaningful decision patterns.

Extensive experiments conducted on multiple public EEG datasets demonstrated that conventional baseline classifiers suffer from significant performance degradation under cross-dataset evaluation, achieving accuracies below 67%. The introduction of CORAL resulted in a substantial improvement of approximately 8–10%, confirming the effectiveness of covariance alignment in mitigating dataset-induced variability. Most importantly, the incorporation of neurophysiological constraints further boosted performance, allowing the proposed method to achieve cross-dataset accuracies exceeding 81–83%, along with balanced sensitivity, specificity, and high discriminative capability (AUC $\approx$ 0.88). Beyond performance gains, the proposed approach preserves model interpretability, as evidenced by feature importance analysis highlighting alpha-band power, occipital–parietal connectivity, and entropy measures as dominant contributors. These findings are fully consistent with known neurophysiological mechanisms underlying eye-state modulation, reinforcing the clinical and neuroscientific relevance of the model.

Overall, this work demonstrates that combining domain adaptation with physiological priors constitutes an effective and principled solution for deployable EEG systems. The proposed approach bridges the gap between data-driven machine learning and neurophysiological interpretability, making it particularly suitable for real-world applications such as cognitive monitoring, vigilance assessment, and brain–computer interfaces.

Future work will focus on extending this approach to larger and more diverse EEG cohorts, incorporating adaptive or learnable physiological constraints, and generalising the methodology to other cognitive and neurological states, including mental workload, fatigue, and pathological conditions.

Acknowledgments

The author thanks colleagues at the Biomedical Engineering Research Laboratory, University of Aboubekr Belkaid, for their guidance and support. Publicly available EEG databases were used in this study. No specific funding was received for this research.

Funding

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

https://doi.org/10.71058/jodac.v10i04001

Declaration of interests

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] *E. Niedermeyer, F. Lopes da Silva, "Electroencephalography: Basic principles, clinical applications, and related fields", 5th ed. Lippincott Williams & Wilkins, 2005.*

[2] *S. Sanei, J. A. Chambers, "EEG Signal Processing", John Wiley & Sons, 2007.*

[3] *W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis," Brain Research Reviews., vol. 29, no. 2–3, pp. 169–195, 1999.*

[4] *S. Narejo, E. Pasero, and F. Kulsoom, "EEG based eye state classification using deep belief network and stacked autoencoder," Int. J. of Electrical and Computer Engineering., vol. 6, no. 6, pp. 3131–3141, 2016.*

[5] *M. Nilashi, R. A. Abumalloh, H. Ahmadi, S. Samad, A. Alghamdi, M. Alrizq, S. Alyami, and F. Nayer, "Electroencephalography (EEG) eye state classification using learning vector quantization and bagged trees," Heliyon., vol. 9, no. 4, e15258, 2023.*

[6] *A. Tahmassebi, J. Martin, A. Meyer-Baese, and A. H. Gandomi, "An interpretable deep learning framework for health monitoring systems: A case study of eye state detection using EEG signals," in 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 211–218, 2020.*

[7] *Z.-T. Liu, C.-S. Jiang, S.-H. Li, M. Wu, W.-H. Cao, and M. Hao, "Eye state detection based on weight binarization convolution neural network and transfer learning," Applied Soft Computing., vol. 109, 107565, 2021.*

[8] *L. Zhao, Z. Wang, G. Zhang, Y. Qi, and X. Wang, "Eye state recognition based on deep integrated neural network and transfer learning," Multimedia Tools and Applications., vol. 77, no. 15, pp. 19415–19438, 2018.*

[9] *C.-H. Han, G.-Y. Choi, and H.-J. Hwang, "Deep convolutional neural network based eye states classification using ear-EEG," Expert Systems with Applications., vol. 192, 116443, 2022.*

[10] *A. Roy, A. Banville, I. Albuquerque, A. Gramfort, and T. Falk, "Deep learning-based electroencephalography analysis: A systematic review," J. of Neural Engineering., vol. 16, no. 5, 2019.*

[11] *F. Laport, P. M. Castro, A. Dapena, F. J. Vazquez-Araujo, and O. Fresnedo, "Eye state identification based on discrete wavelet transforms," Applied Sciences., vol. 11, 5051, 2021.*

[12] *T. Reddy and L. Behera, "Online eye state recognition from EEG data using deep architectures," in Proc. 2016 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC), pp. 712–717, Budapest, Hungary, 2016.*

[13] *A. M. Elshewey, A. A. Alhussan, D. S. Khafaga, E.-S. M. Elkenawy, and Z. Tarek, "EEG-based optimization of eye state classification using modified-BER metaheuristic algorithm," Scientific Reports., vol. 14, 24489, 2024.*

[14] *B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," in Proc. AAAI Conf. on Artificial Intelligence, pp. 2058–2065, 2016.*

[15] *N. Abo Alzahab, et al., "Auditory evoked potential EEG-Biometric dataset (version 1.0.0)," PhysioNet, 2021.*

[16] *M. Torkamani-Azar, S. D. Kanik, S. Aydin, and M. Cetin, "Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state EEG in a long sustained attention task," IEEE J. of Biomedical and Health Informatics., vol. 24, no. 9, pp. 2550–2558, 2020.*

[17] *T. Mastaneh, "SPIS-Resting-State-Dataset [Dataset]," GitHub repository, 2025.*

https://doi.org/10.71058/jodac.v10i04001

[18] *MNNIT Eye State EEG Dataset – UCI Machine Learning Repository, 2013.*

[19] *N. Bigdely-Shamlo, T. Mullen, C. Kothe, K. M. Su, and K. A. Robbins, "The PREP pipeline: Standardized preprocessing for large scale EEG analysis," Front. Neuroinformatics., vol. 9, 16, 2015.*

[20] *J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," Am. J. Physiol.-Heart Circ. Physiol., vol. 278, no. 6, pp. H2039–H2049, 2000.*

[21] *M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," Phys. Rev. Lett., vol. 89, no. 6, 068102, 2002.*

[22] *A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," Expert Syst. Appl., vol. 32, no. 4, pp. 1084–1093, 2007.*

[23] *C. J. Stam, "Functional connectivity patterns of human magnetoencephalographic recordings," Neurosci. Lett., vol. 355, no. 1–2, pp. 25–28, 2004.*

[24] *J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," Human Brain Mapping., vol. 8, no. 4, pp. 194–208, 1999.*

[25] *M. Vinck, R. Oostenveld, M. V. Wingerden, F. Battaglia, and C. M. A. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," NeuroImage., vol. 55, no. 4, pp. 1548–1565, 2011.*

[26] *R. J. Barry, A. R. Clarke, S. J. Johnstone, C. A. Magee, and J. A. Rushby, "EEG differences between eyes-closed and eyes-open resting conditions," Clin. Neurophysiol., vol. 118, no. 12, pp. 2765–2773, 2007.*

[27] *B. Sun and K. Saenko, "Deep CORAL: Correlation alignment for deep domain adaptation," in European Conf. on Computer Vision (ECCV), pp. 443–450, 2016.*

[28] *F. Vecchio, F. Miraglia, C. Pappalettera, A. Orticoni, F. Alù, E. Judica, M. Cotelli, and P. M. Rossini, "Entropy as measure of brain networks' complexity in eyes open and closed conditions," Symmetry., vol. 13, no. 11, 2178, 2021.*

[29] *I. R. Munthe, B. H. Rambe, F. Hanum, A. T. Amanda, A. S. R. Hutagaol, and R. Harianto, "Implementation of stacking technique combining machine learning and deep learning algorithms using SMOTE to improve stock market prediction accuracy," J. of Applied Data Sciences., vol. 5, no. 4, pp. 2079–2091, 2024.*