Transforming EEG Signals into Images for Motor Imagery Classification Using a YOLO11-Based Model

Aya A. Abdullah
Department of Computer Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq
aya.a.abdullh@aliraqia.edu.iq (corresponding author)

Khamis A. Zidan
Vice Rector for Scientific Affairs, Al-Iraqia University, Baghdad, Iraq
khamis_zidan@aliraqia.edu.iq

A. S. Albahri
Technical College, Imam Jaafar Al-Sadiq University, Baghdad, Iraq
ahmed.bahri1978@iips.icci.edu.iq

Received: 15 February 2026 | Revised: 7 April 2026 | Accepted: 17 April 2026

 DOI: https://doi.org/10.48084/etasr.18199

ABSTRACT

A Brain–Computer Interface (BCI) enables direct communication between an individual's brain and external hardware. Potential application areas of BCI systems include assistive technology, smart home environments, healthcare, and many other domains. Electroencephalography (EEG)-based Motor Imagery (MI) signals have been widely used in such applications; however, accurate classification of MI-EEG data remains challenging because traditional EEG classification methods often cannot effectively capture the complex spatiotemporal characteristics of raw EEG signals. To address this limitation, this study proposes a structured EEG-to-RGB representation combined with a YOLO11-L-cls-based classification framework for motor imagery EEG classification. The proposed approach was evaluated using two datasets, namely BCI Competition IV Dataset 2b (BCICIV2b) and Dataset III, under subject-independent evaluation settings. The proposed method achieved a classification accuracy of approximately 99% on BCICIV2b and 97.5% on Dataset III. In addition, the performance of the proposed method was compared with traditional machine learning techniques, including Random Forest (RF) and Linear Discriminant Analysis (LDA), as well as deep learning models such as Convolutional Neural Networks (CNNs). Extensive experimental analyses demonstrated the effectiveness and robustness of the proposed approach for developing MI-EEG-based BCI systems.

Keywords-Motor Imagery (MI); Brain–Computer Interface (BCI); time–frequency image representation; YOLO11-L-cls-based classification

I. INTRODUCTION

Brain–Computer Interface (BCI) systems provide a way for the human brain to communicate directly with external devices and allow a multitude of uses such as neurorehabilitation, assistive technologies, and human–computer interaction. Specifically, Motor Imagery (MI)-based BCI systems provide insight into controlling external devices using imagination or visualization. While these types of BCI systems have the potential to assist in neurorehabilitation, their non-stationary Electroencephalography (EEG) data, inter-subject differences, and vulnerability to noise pose challenges for achieving reliable and accurate real-time MI classification, which is critical for BCI systems to become truly functional and widely accessible [1]. Machine learning techniques have been widely applied for EEG signal classification, typically following a structured pipeline that includes preprocessing, feature extraction, and classification. For example, previous studies such as [2] employed multiple classifiers, including Convolutional Neural Networks (CNNs), Support Vector Machines (SVM), and Random Forest (RF), combined with feature extraction methods to improve classification performance.

The preprocessing stage's primary function is to remove artifacts and noise while preserving the essential features for MI classification. Signal processing techniques such as the extraction of discriminative features from MI-EEG signals include energy-based features, Common Spatial Patterns (CSP), and Autoregressive (AR) features [3]. Popular feature classification techniques include RF [4], Gradient Boosting

(GB) [5], and K-Nearest Neighbors (KNN) [6], which are utilized to classify the extracted features. However, the performance of machine learning techniques is limited principally because they are labor-intensive and rely heavily on expert knowledge. Recent deep learning studies have highlighted the use of lightweight deep learning architectures such as EEGNet and its low-precision variants to address the challenges of power-efficient real-time classification in EEG-based BCIs. While lightweight deep learning models improve hardware efficiency compared to heavier models, the lack of generalization and adaptability across subjects continues to motivate the exploration of alternative deep learning frameworks [7].

Transfer learning methods have improved cross-subject generalization and provide a means for enhancing performance across subjects but they also introduce new challenges including increased computational and training burdens. These limitations motivate the need to explore alternative more efficient deep learning frameworks [8, 9]. In recent years, deep learning approaches have been widely adopted for MI-EEG classification due to their success in various BCI applications [3, 10, 11]. CNN and hybrid deep learning-based models have achieved high accuracy; however, the literature often highlights model complexity, redundant features, and limited generalization to unseen datasets and subjects as key limiting factors [11, 12]. Several recent studies have explored transfer learning, fuzzy logic, and optimization approaches to improve performance; however, these approaches generally require high-quality data, extensive calibration, and significant computational resources [13, 14]. In addition, challenges such as noise sensitivity, overfitting, and limited operational robustness remain persistent obstacles in the deployment of MI-BCI systems [10, 12, 15]

Given these limitations, this study proposes a deep learning-based classification framework using EEG-derived images and a YOLO11-L-cls architecture [16, 17]. YOLO11-L-cls models have demonstrated strong performance in image classification tasks, as well as high processing speed, making them a promising candidate for scalable, real-time MI-BCI applications. Despite their success in other domains, YOLO11-L-cls architectures have not been sufficiently investigated for MI-EEG image classification, representing a clear research gap. To bridge this gap, this work proposes a unified EEG-to-image learning framework that leverages the strengths of YOLO11-L-cls architectures for MI-EEG decoding. Unlike conventional EEG-to-image approaches that rely on standard CNN backbones, the proposed method investigates a YOLO-based classification architecture, enabling efficient multi-scale feature extraction and improved computational efficiency. To the best of our knowledge, the application of YOLO11-L-cls for MI-EEG image classification has not been systematically explored.

The main contributions of this work are summarized as follows:

- A structured EEG-to-RGB representation integrating time–frequency features and statistical descriptors.
- The adoption of a YOLO11-L-cls-based classification architecture for MI-EEG image classification.
- A complete end-to-end pipeline including preprocessing, feature extraction, image construction, and classification
- Demonstration of high classification performance on benchmark MI-EEG datasets.

II. METHODOLOGY

This study presents an EEG-to-image–based deep learning framework for binary classification of MI-EEG signals corresponding to right- and left-hand movements. The overall workflow of the proposed methodology is shown in Figure 1. Initially, EEG data undergo multiple necessary preliminary signal conditioning and preprocessing steps prior to model training, including bandpass filtering, normalization, and data balancing. To facilitate visual feature extraction from the EEG data, an image representation of the preprocessed EEG signals is generated for visual feature learning. Subsequently, the model development phase utilizes a deep learning architecture based on YOLO11-L-cls for binary classification of two types of MI using the recorded EEG signals (right-hand and left-hand movements). Finally, the classification performance of the developed MI model is evaluated using a train/validation split to ensure reliable estimates of classification and generalization performance.

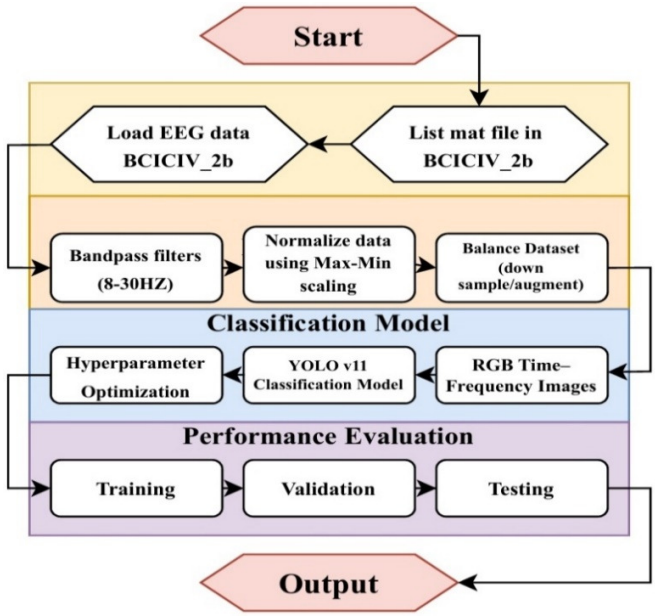


Fig. 1. Flowchart of the proposed YOLO-based EEG RGB image classification framework.

A. Database Description

This study utilizes an EEG-MI dataset from the BCI Competition IV Dataset 2b (BCICIV2b), available at [18] and further discussed in [19]. The dataset includes continuous EEG recordings taken from three sensorimotor channels (C3, Cz, and C4), which are directly related to motor activity. The sampled data were recorded at a rate of 250 Hz from nine

individuals performing motor imagery tasks. There are two MI classes for each individual: left-hand and right-hand movement imagery. During acquisition, participants were instructed to imagine performing either a left- or right-hand movement while seated comfortably and viewing a blank screen. A total of 160 trials were collected per participant.

Two separate recording sessions were conducted. The first was a training session without feedback, in which participants received a brief auditory cue followed by a visual arrow on a blank screen indicating the required MI task for a duration of 4 s. The second was an evaluation session with visual feedback, where subjects were instructed to control a gray smiley displayed at the center of the screen by imagining left- or right-hand movements after a short warning beep. The smiley feedback was presented for 4 s and changed color to green when the imagined movement direction was correct and to red when it was incorrect. This structured protocol ensures the collection of reliable and representative EEG data suitable for binary MI classification tasks.

Dataset II consists of RGB images generated from the EEG feature representations derived from Dataset I, whereas Dataset III consists of EEG data from four individual participants obtained using the EMOTIV EPOC headset [20]. In this study, three datasets are used: (1) Dataset I: BCICIV2b dataset containing raw EEG signals; (2) Dataset II: RGB image dataset generated from Dataset I; (3) Dataset III: external EEG dataset used for testing.

B. Raw Data Preprocessing

The raw EEG-MI signals were first preprocessed to generate discriminative representations for binary MI classification (left-hand vs. right-hand), as shown in Figure 2.

First, the signals were filtered using a fourth-order Butterworth bandpass filter in the range of 8–30 Hz to preserve the sensorimotor rhythms associated with MI (alpha: 8–13 Hz and beta: 14–30 Hz) while attenuating artifacts and irrelevant noise [21]. The filtering operation can be expressed as:

$$y[n] = \sum_{k=0}^{M} b_k x[n-k] - \sum_{k=1}^{N} a_k y[n-k] \quad (1)$$

In this equation, b_k and a_k are the filter coefficients. In the second step, the filtered signals were segmented into multiple time windows to capture temporal variability in MI responses across subjects. Specifically, five time-segment groups were considered (1 s, 2 s, 3 s, 4 s, and 5 s) within the MI interval, producing frames such as (3–4) s, (4–5) s, and (3–8) s.

A generic segmentation form is given by:

$$X_w = X[t_s : t_e] \quad (2)$$

where t_s, t_e define the start and end points of each time window. Next, time–frequency features were extracted from each segment using the Discrete Wavelet Transform (DWT) to capture non-stationary EEG characteristics efficiently. To eliminate redundancy and to reduce dimensionality, the extracted coefficients were listed using statistical descriptors, including maximum, minimum, mean, median, and standard deviation. After combining the features from all time frames, two labels were assigned to each feature set: (I) duration index (segment group) and (II) movement label (left or right) [22].

To ensure uniform scaling and stable learning, Dataset I (standardized preprocessing output) was generated using min–max normalization:

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3)$$

where x is the original value, $\min(x)$ is the minimum value in the dataset, $\max(x)$ is the maximum value, and z is the normalized value.

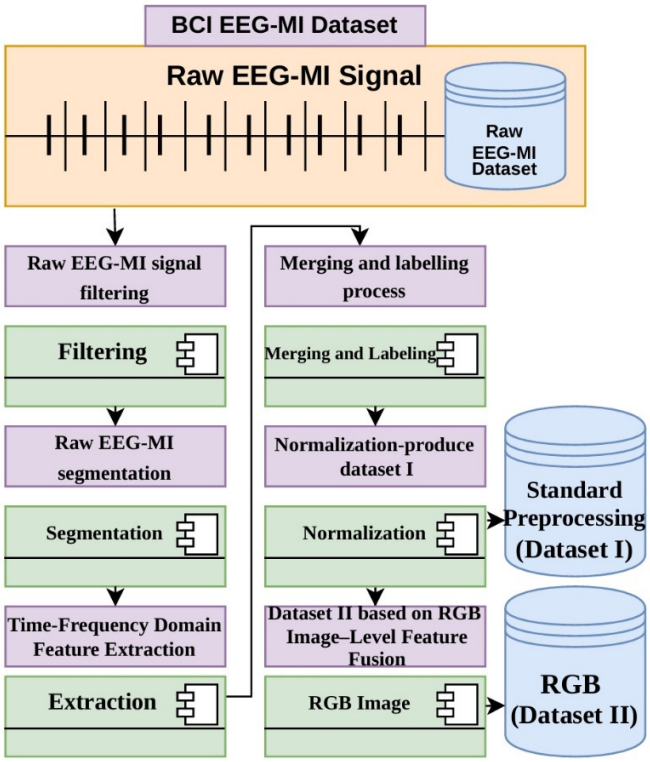


Fig. 2. Dataset description and preprocessing of EEG data.

In conclusion, the preprocessing pipeline was adapted to fit the structure of an image-based deep neural network. The subsequently created Dataset II comprised normalized EEG-MI feature vectors (described earlier in the document), which were converted into RGB images. Therefore, for each normalized feature vector corresponding to a single EEG trial (i.e., Trial n), there is a corresponding single RGB image [23]. Each EEG trial was thus represented as one RGB image, where temporal segments were integrated within the same image representation rather than treated as independent samples. Additionally, this transformation enables image-level feature fusion using convolutional architectures, allowing more effective use of spatial patterns that are not explicitly represented in the original one-dimensional feature space.

Mathematically, this transformation can be expressed as:

$$F\big(\text{reshape}(x_n)\big) = RGB^I \quad (4)$$

where the feature vector $x_n \in \mathbb{R}^d$ represents a single EEG-MI trial. The reshape operation converts the one-dimensional feature vector into a two-dimensional matrix, and $F(\cdot)$ maps this matrix into red, green, and blue channels to produce the final RGB image [24]. This representation preserves the underlying feature relationships while enabling effective learning within the YOLO11-L-cls-based classification framework [25].

Due to the balanced nature of the dataset (i.e., equal numbers of left- and right-hand MI samples), no oversampling techniques (such as SMOTE) were required. The overall EEG-to-RGB transformation and training pipeline can be summarized as follows:

1. Apply bandpass filtering (8–30 Hz) to raw EEG signals.
2. Segment EEG signals into temporal windows.
3. Extract DWT-based features.
4. Compute statistical descriptors.
5. Combine features into a single vector per trial.
6. Reshape feature vector into a 2D matrix.
7. Map matrix into RGB channels to form an image.
8. Train the YOLO11-L-cls model on generated images.
9. Evaluate model performance on validation and external datasets.

C. Proposed Classification Model

The YOLO11-L-cls classification model proposed in this study utilizes a YOLO11-L-cls architecture for image classification of RGB images produced from MI-EEG data. As a preprocessing step, all EEG trials (representing a single MI instance) were normalized and transformed into independent RGB images of fixed size 224 × 224 pixels. These generated images correspond to two MI classes, i.e., left-hand movement or right-hand movement.

The model was implemented using the Ultralytics YOLO framework and initialized with pretrained weights (yolo11l-cls.pt). Training was performed for 50 epochs with a batch size of 16 and an input image size of 224 × 224 pixels. The AdamW optimizer was used with an initial learning rate of 8×10^{-6}, momentum of 0.937, and weight decay of 0.0016. Pretrained weights (yolo11l-cls.pt) were fine-tuned on the EEG-derived RGB image dataset. A dropout rate of 0.28 was applied to reduce overfitting, and a fixed random seed (42) was used to ensure reproducibility. The use of pretrained weights enabled efficient transfer learning and accelerated convergence during training. The model is trained using categorical cross-entropy loss, which improves optimization stability and generalization performance.

The architecture enables the learning of discriminative spatial representations directly from EEG-derived RGB images, avoiding the need for manual feature engineering. The overall proposed YOLO11-L-cls classification pipeline is illustrated in Figure 3.

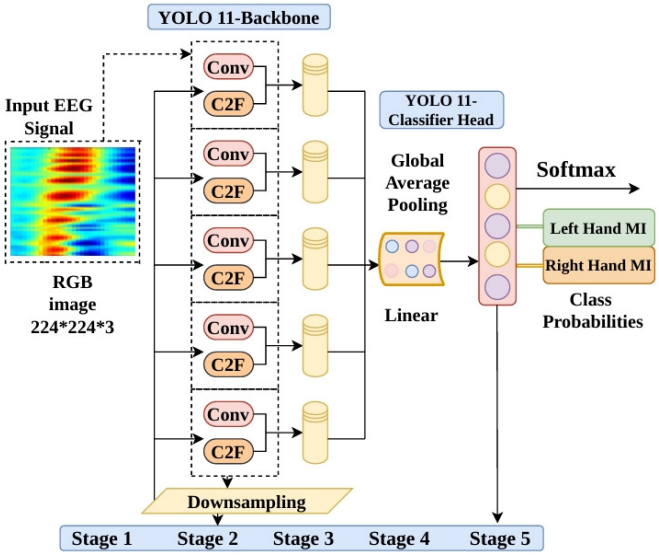


Fig. 3. Proposed YOLO11 classification model.

1) YOLO11-L-cls-Based Convolutional Feature Extraction

The YOLO11-L-cls deep learning framework integrates multiple convolutional stages and C2f blocks to automatically extract multi-scale features from MI-EEG RGB images. These images encode spatial characteristics of EEG signals, which are processed through successive convolutional layers as expressed in (5):

$$f(l) = \varphi\left(\sum_{c=1}^{C} W_{l,c} * X_c + b_l\right) \quad (5)$$

where $f(l)$ represents the spatial convolutional output feature the output feature map of layer l, X_c is the input feature map of channel c, $W_{l,c}$ denotes the convolutional weights, b_l is the bias term, and $\varphi(\cdot)$ is the SiLU activation function, which promotes the preservation of feature representations and gradient propagation within the YOLO11-L-cls architecture.

YOLO11-L-cls employs strided convolutions for downsampling instead of max-pooling layers, preserving spatial discriminative information in MI-EEG images. In addition, C2f blocks enhance feature reuse and improve information flow across network layers.

2) Image-Level Classification

In contrast to sequence-based architectures such as CNN-LSTM networks [26], the proposed method performs image-level classification. In this approach, a single RGB image represents one MI trial. The advantage of this design is that it does not require recurrent or temporal modeling components, yet it still achieves classification performance comparable to that of sequence-based models. Experimental results demonstrate that the proposed YOLO11-L-cls-based classifier effectively learns the spatial representations required for accurate and efficient binary classification of MI from EEG-derived RGB images.

3) YOLO-Based Image Classification Model

MI-EEG are represented as trial-based feature matrices, $X \in \mathbb{R}^{N \times D}$, where N denotes the number of MI-EEG trials and

D denotes the number of distinct features extracted from each trial. Prior to analysis, the MI-EEG feature data undergo preprocessing and normalization. Subsequently, each MI trial is converted into a single RGB image representation, resulting in one image per MI-EEG trial [19]. Each RGB image is represented as follows:

$$I_i \in \mathbb{R}^{3\times W\times H} \quad (6)$$

where W and H represent the image width and height, respectively, both set to 224 pixels in the experimental design. The three channels correspond to the red, green, and blue components of the MI-EEG feature representation. The label for each image is defined as $y_i \in \{0,1\}$, where 0 indicates left-hand MI and 1 indicates right-hand MI.

Spatial feature extraction is performed using the YOLO11-L-cls classification architecture, which employs a deep CNN backbone structure. The backbone contains five convolutional stages composed of strided convolutional layers and C2f blocks that enable hierarchical extraction of spatial features from RGB MI-EEG images. These layers automatically learn discriminative global and local spatial patterns from the transformed EEG representations. The convolutional operations within the YOLO11-L-cls backbone follow the formulation presented in (5).

After spatial feature extraction, the feature maps are pooled into a single feature vector using global average pooling:

$$z_c = \frac{1}{H_f W_f}\sum_{i=1}^{H_f}\sum_{i=1}^{W_f} F_c(i,j) \quad (7)$$

where F_c is the feature map of channel c, and H_f and W_f denote the spatial dimensions of the final pooled feature map. After pooling, this single feature vector is fed to the classification head. A fully connected linear layer followed by a softmax function is used to estimate class probabilities. The probability of class i is computed as follows:

$$\rho = \frac{e^{z_i}}{\sum_{j=1}^{c} e^{z_j}} \quad (8)$$

where z_i is the logit corresponding to class i, and C denotes the total number of classes ($C = 2$).

The model uses categorical cross-entropy loss combined with an AdamW optimizer to achieve stable performance and efficient training. Since the dataset is inherently balanced between left-hand and right-hand MI classes, no additional data balancing techniques were required. The proposed YOLO11-L-cls-based image classification model effectively learns discriminative spatial representations from EEG-derived RGB images, enabling accurate and efficient binary classification of MI [14].

III. RESULTS AND DISCUSSION

The current study investigated the effectiveness of the proposed method for binary MI classification involving two movement classes: left hand and right hand. To accomplish this, the EEG data were preprocessed and normalized, and each EEG trial was converted into a single RGB image before being provided as input to the YOLO11-L-cls classification model. Subsequently, YOLO11-L-cls was used to classify the EEG data into left-hand or right-hand MI classes.

The evaluation metrics used to assess the performance of YOLO11-L-cls included accuracy, precision, recall, and F1-score, as these metrics provide a comprehensive analysis of classification performance. The experiments were conducted using an 80/20 train–validation split. Each EEG trial was treated as an independent sample, and no overlap between the training and validation sets was permitted. The training and validation samples were drawn from the same subject pool while maintaining strict separation between the two sets. Since the dataset was inherently balanced between left-hand and right-hand classes, no additional data balancing techniques were required.

For classification, the YOLO11-L-cls-based model produces a probability distribution over the two MI classes using the softmax activation function. The model was trained by minimizing the categorical cross-entropy loss, expressed as:

$$L = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{C} y_{n,i}\log\left(p_{n,i}\right) \quad (9)$$

where N is the number of samples, $y_{n,i}$ is the ground-truth label of sample n for class i, and $p_{n,i}$ is the predicted probability.

Optimization was performed using the AdamW optimizer**,** which updates the model parameters according to:

$$\theta_{t+1} = \theta_t - \eta\nabla_\theta \mathcal{L}(\theta_t) \quad (10)$$

where θ denotes the model parameters, η is the learning rate, and $\nabla_\theta \mathcal{L}(\theta_t)$ is the gradient of the loss at iteration t. To prevent overfitting, early stopping was applied based on the validation loss, and the model with the best validation performance was retained.

Experimental results demonstrate that the proposed YOLO11-cls-based classification framework achieves high and stable performance in binary MI classification, confirming the effectiveness of transforming EEG features into RGB images and leveraging modern image-based deep learning architectures for MI-EEG analysis.

A standard analysis setup was adopted to evaluate the performance of the proposed method in comparison with several well-established machine learning classifiers, including RF, KNN, and GB. These classifiers are representative of widely used approaches for MI classification using EEG signals and provide a comprehensive basis for evaluating the performance of the YOLO11-L-cls-based framework. All models were trained and evaluated under identical experimental conditions and data partitions to ensure fair comparison. The selected classifiers represent different learning paradigms, including ensemble learning, distance-based learning, and margin-based learning, enabling a broad comparison between the proposed YOLO11-L-cls framework and traditional machine learning approaches.

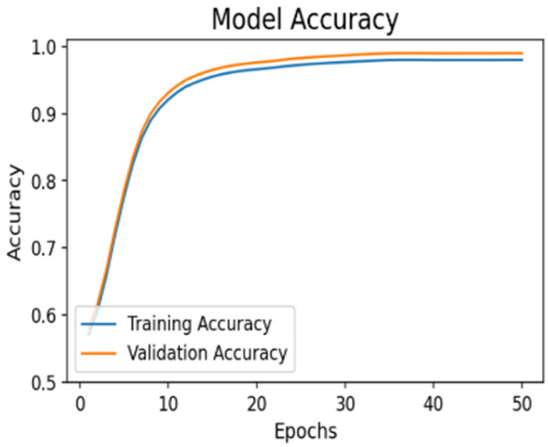
A. Training and Validation Performance

An early stopping strategy was utilized during training of the proposed model, resulting in 50 training epochs without

signs of overfitting. At the beginning of the training process, when the model was still learning, the results for Dataset II indicated an accuracy of approximately 59%, as shown in Figure 4(a). During training, the cross-entropy loss steadily decreased, indicating increasing confidence in the model predictions, as shown in Figure 4(b).

By the end of training, the validation accuracy for Dataset II increased to 99%, as shown in Figure 4(a), and the final recorded validation loss was 0.09185 (Figure 4(b)). These results indicate that the model converged effectively. Furthermore, the results for Dataset III demonstrated a classification accuracy of 97.5%.

(a)



(b)

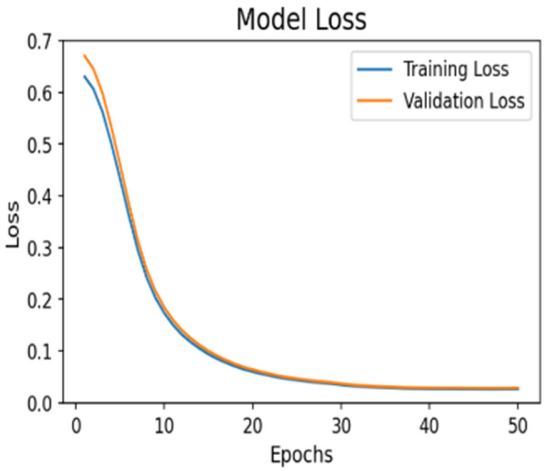


Fig. 4. Training and validation curves: (a) accuracy, (b) cross-entropy loss.

B. Performance Evaluation Metrics

The performance of the proposed YOLO11-L-cls classification model was evaluated against conventional machine learning models, including RF, KNN, and GB, through comparative experiments conducted on three datasets: Dataset I, Dataset II, and Dataset III. Dataset I represents the standardized EEG dataset described in Figure 2. Each model was trained and tested under the same experimental protocol for all datasets, enabling direct and fair comparison among the evaluated methods.

The experimental performance results for each dataset are presented in Tables I and II. Performance was assessed using several quantitative evaluation metrics, including accuracy, precision, recall, F1-score, specificity, and LogLoss. The evaluation metrics were computed using the following equations:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (11)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (12)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (13)$$

$$\text{F1} - \text{score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \quad (14)$$

$$\text{LogLoss} = -\frac{1}{N}\sum_n \sum_c y_{n,c} \log\left(p_{n,c}\right) \quad (15)$$

C. Performance Comparison

To provide an unbiased assessment of each model against its competitors, the same experimental protocols were applied to all models during each phase of the analysis. In addition to classical machine learning models, recent deep learning approaches have been widely investigated for MI-EEG classification. For instance, the MTFB-CNN model utilizes multi-scale time–frequency representations combined with channel attention mechanisms, achieving an accuracy of 84.48%. Similarly, the AORNDL-MIC framework integrates MSPCA, CWT, RetinaNet, and optimization algorithms, achieving an accuracy of 86.53%. These approaches demonstrate the effectiveness of combining signal processing techniques with deep learning methods.

Compared to these models, the proposed YOLO11-L-cls framework provides a simpler end-to-end image-based pipeline while achieving superior classification performance. It should be noted that direct comparison across studies is challenging due to differences in datasets, preprocessing pipelines, and evaluation protocols.

The results presented in Tables I and II indicate that the proposed YOLO11-L-cls architecture outperforms the benchmark classifiers (RF, GB, and KNN) across all evaluation metrics, including accuracy, F1-score, precision, recall, LogLoss, and specificity. This indicates that the proposed architecture has learned robust discriminative features from EEG-derived images, making YOLO11-L-cls well-suited for reliable MI classification.

Receiver Operating Characteristic (ROC) analysis was performed to further evaluate the discriminative performance of the proposed model. The ROC curve illustrates the trade-off between the true positive rate (TPR) and false positive rate (FPR). The proposed YOLO11-L-cls model achieved an Area Under the Curve (AUC) value of 0.999, indicating excellent classification capability and strong separability between the two motor imagery classes. Figure 6 shows the ROC curve of the proposed model.

The performance of the proposed YOLO11-L-cls model was evaluated using confusion matrices, which provide a detailed view of classification performance for the two MI

classes. The confusion matrices illustrate the classification behavior of the proposed model under identical experimental conditions and are presented in Figure 5, including both absolute values and normalized values. These results indicate a very low misclassification rate for both classes, confirming the high reliability of the proposed model.

Receiver Operating Characteristic (ROC) analysis was also performed to further evaluate the discriminative performance of the proposed model. The ROC curve illustrates the trade-off between the true positive rate (TPR) and the false positive rate (FPR). The proposed YOLO11-L-cls model achieved an Area Under the Curve (AUC) value of 0.999, indicating excellent classification capability and strong separability between the two MI classes. Figure 6 shows the ROC curve of the proposed model.

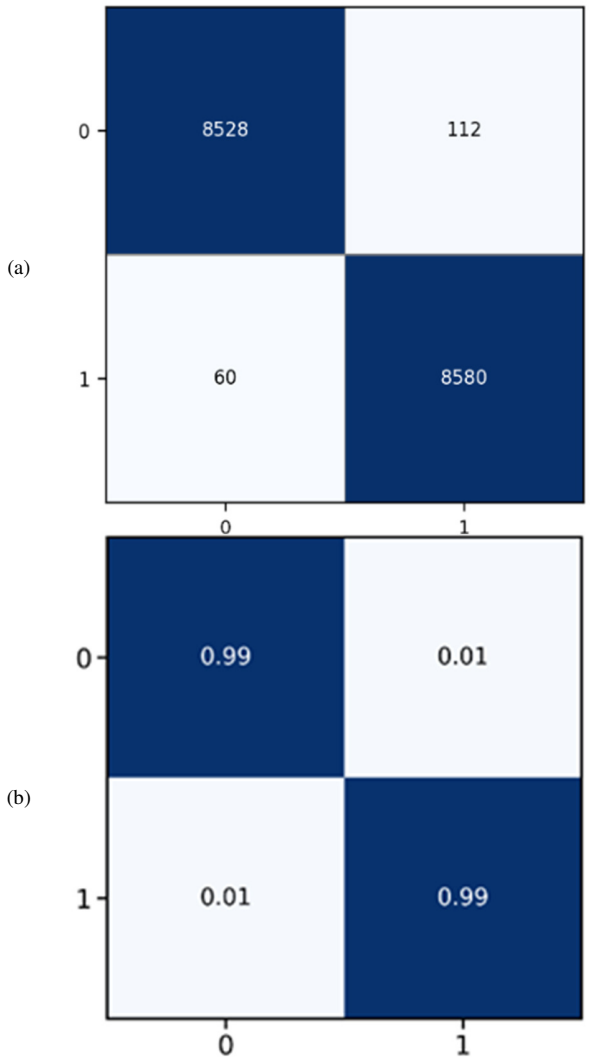


Fig. 5. Confusion matrix of the YOLO11-L-cls classifier: (a) absolute values, (b) normalized values.

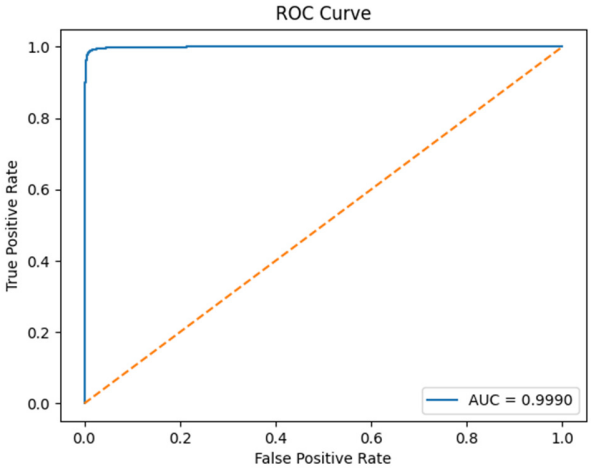


Fig. 6. ROC curve of the proposed model.

TABLE I. PERFORMANCE COMPARISON BETWEEN THE PROPOSED YOLO11-L-CLS AND BENCHMARK MODELS ON DATASET I & DATASET II

Model	YOLO11-L-cls	RF	GB	KNN
Accuracy	0.99	0.714	0.719	0.670
F1-score	0.99	0.714	0.719	0.670
Precision	0.993	0.714	0.719	0.670
Recall	0.987	0.714	0.719	0.670
LogLoss	0.035	0.312	0.342	0.685
Specificity	0.993	0.714	0.719	0.670

TABLE II. PERFORMANCE COMPARISON BETWEEN THE PROPOSED YOLO11-L-CLS AND BENCHMARK MODELS ON DATASET III

Model	YOLO11-L-cls	RF	GB	KNN
Accuracy	0.975	0.801	0.776	0.911
F1-score	0.975	0.801	0.786	0.941
Precision	0.970	0.803	0.779	0.942
Recall	0.980	0.801	0.776	0.941
LogLoss	0.069	0.609	0.610	0.185
Specificity	0.97	0.801	0.777	0.941

IV. CONCLUSION AND FUTURE WORK

This study presented a YOLO11-L-cls-based image classification framework for Motor Imagery Electroencephalography (MI-EEG) signal classification. The proposed approach converts normalized EEG feature vectors into RGB images. The YOLO11-L-cls model facilitates direct utilization of the spatial structures incorporated within the EEG data. Therefore, spatial information can be exploited without the need for handcrafted feature engineering or explicit temporal sequence modeling (e.g., timestamp-based dependencies), which aligns with the learning capabilities of the YOLO11-L-cls network. In this study, a YOLO11-L-cls-based neural network is proposed that accurately and reliably discriminates between left-hand and right-hand MI classes.

As demonstrated by the experimental findings, the proposed YOLO11-L-cls-based classification method generalizes well to new datasets and different noise and

filtering configurations. Due to the use of categorical cross-entropy loss combined with the AdamW optimization algorithm, the model converges efficiently and stably, reducing the risk of overfitting. The compact architecture and computational efficiency of the YOLO11-L-cls model reduce computational cost, improve inference speed, and support the development of practical real-time Brain–Computer Interface (BCI) systems.

Future work will focus on extending the proposed framework to larger and more diverse MI-EEG datasets. In addition, advanced strategies for cross-subject generalization will be investigated. Subject-adaptive learning methods and multi-class MI classification will also be explored to improve the robustness and applicability of the YOLO11-L-cls-based image classification approach in real-world BCI systems.

DECLARATION OF COMPETING INTERESTS

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

ACKNOWLEDGMENT

This research received no external funding.

DATA AVAILABILITY

Dataset I of the BCI Competition IV Dataset 2b is publicly available from the official BCI Competition IV repository at [18]. Dataset II consists of RGB images generated from EEG feature representations derived from Dataset I. Dataset III represents an additional external testing dataset and is not publicly available due to privacy and ethical restrictions. Requests for access to the processed RGB image dataset, derived data, and implementation details may be directed to the corresponding author.

CODE AVAILABILITY

The implementation of the proposed method is available from the corresponding author upon reasonable request.

AI USE AND DECLARATION OF GENERATIVE AI USE

During the preparation of this work, generative AI tools were used for the following purposes: (1) to assist with language editing and grammatical improvement; and (2) to assist with academic writing. Following completion of all literature review and editorial modifications, the authors assume full responsibility for the content of the published article. The authors further state that generative AI tools were not used to generate or fabricate research data, experimental results, illustrations, images, or scientific conclusions.

REFERENCES

[1] A. S. Albahri *et al.*, "Trust and explainability in robotic hand control via adversarial multiple machine learning models with EEG sensor data fusion: A fuzzy decision-making solution," *Computers in Biology and Medicine*, vol. 196, Sept. 2025, Art. no. 110922, https://doi.org/10.1016/j.compbiomed.2025.110922.

[2] M. A. Alsuwaiket, "Feature Extraction of EEG Signals for Seizure Detection Using Machine Learning Algorthims," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9247–9251, Oct. 2022, https://doi.org/10.48084/etasr.5208.

[3] Y. Lu, W. Wang, B. Lian, and C. He, "Feature Extraction and Classification of Motor Imagery EEG Signals in Motor Imagery for Sustainable Brain–Computer Interfaces," *Sustainability*, vol. 16, no. 15, Aug. 2024, Art. no. 6627, https://doi.org/10.3390/su16156627.

[4] R. Zhang *et al.*, "A New Motor Imagery EEG Classification Method FB-TRCSP+RF Based on CSP and Random Forest," *IEEE Access*, vol. 6, pp. 44944–44950, 2018, https://doi.org/10.1109/ACCESS.2018.2860633.

[5] X. Wang *et al.*, "Enhancing effluent biochemical oxygen demand prediction in wastewater treatment plants: An ensemble model of gradient boosting and LSTM," *Journal of Environmental Chemical Engineering*, vol. 13, no. 5, Oct. 2025, Art. no. 117897, https://doi.org/10.1016/j.jece.2025.117897.

[6] X. Tang, T. Wang, Y. Du, and Y. Dai, "Motor imagery EEG recognition with KNN-based smooth auto-encoder," *Artificial Intelligence in Medicine*, vol. 101, Nov. 2019, Art. no. 101747, https://doi.org/10.1016/j.artmed.2019.101747.

[7] J. Cao *et al.*, "An optimized EEGNet processor for low-power and real-time EEG classification in wearable brain–computer interfaces," *Microelectronics Journal*, vol. 145, Mar. 2024, Art. no. 106134, https://doi.org/10.1016/j.mejo.2024.106134.

[8] H. Wang *et al.*, "EEG-Based Motor Imagery Recognition Framework via Multisubject Dynamic Transfer and Iterative Self-Training," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 8, pp. 10698–10712, Aug. 2024, https://doi.org/10.1109/TNNLS.2023.3243339.

[9] R. Zhang, Q. Zong, L. Dou, X. Zhao, Y. Tang, and Z. Li, "Hybrid deep neural network using transfer learning for EEG motor imagery decoding," *Biomedical Signal Processing and Control*, vol. 63, Jan. 2021, Art. no. 102144, https://doi.org/10.1016/j.bspc.2020.102144.

[10] I. Velasco, A. Sipols, C. S. De Blas, L. Pastor, and S. Bayona, "Motor imagery EEG signal classification with a multivariate time series approach," *BioMedical Engineering OnLine*, vol. 22, no. 1, Mar. 2023, Art. no. 29, https://doi.org/10.1186/s12938-023-01079-x.

[11] F. Y. Assiri and M. Ragab, "Boosted Harris Hawks Shuffled Shepherd Optimization Augmented Deep Learning based motor imagery classification for brain computer interface," *Plos One*, vol. 19, no. 11, Nov. 2024, Art. no. e0313261, https://doi.org/10.1371/journal.pone.0313261.

[12] Y. An, H. K. Lam, and S. H. Ling, "Multi-classification for EEG motor imagery signals using data evaluation-based auto-selected regularized FBCSP and convolutional neural network," *Neural Computing and Applications*, vol. 35, no. 16, pp. 12001–12027, June 2023, https://doi.org/10.1007/s00521-023-08336-z.

[13] D. Gwon and M. Ahn, "Motor task-to-task transfer learning for motor imagery brain-computer interfaces," *NeuroImage*, vol. 302, Nov. 2024, Art. no. 120906, https://doi.org/10.1016/j.neuroimage.2024.120906.

[14] R. S. Chowdhury, S. Bose, S. Ghosh, and A. Konar, "Attention Induced Dual Convolutional-Capsule Network (AIDC-CN): A deep learning framework for motor imagery classification," *Computers in Biology and Medicine*, vol. 183, Dec. 2024, Art. no. 109260, https://doi.org/10.1016/j.compbiomed.2024.109260.

[15] A. Falcon-Caro, J. Filipe Ferreira, and S. Sanei, "Cooperative Identification of Prolonged Motor Movement From EEG for BCI Without Feedback," *IEEE Access*, vol. 13, pp. 11765–11777, 2025, https://doi.org/10.1109/ACCESS.2025.3528539.

[16] A. Ahmed, A. I. Siam, A. E. M. Atwa, M. A. Atwa, E. M. Abdelrahim, and E.-S. Atlam, "An Explainable YOLO-Based Deep Learning Framework for Pneumonia Detection from Chest X-Ray Images," *Algorithms*, vol. 18, no. 11, Nov. 2025, Art. no. 703, https://doi.org/10.3390/a18110703.

[17] R. Kailasam and S. Balasubramanian, "Deep Learning for Pneumonia Detection: A Combined CNN and YOLO Approach," *Human-Centric Intelligent Systems*, vol. 5, no. 1, pp. 44–62, Mar. 2025, https://doi.org/10.1007/s44230-025-00091-9.

[18] "BCI Competition IV." Bbci.de. https://bbci.de/competition/iv/.

[19] M. Tangermann *et al.*, "Review of the BCI Competition IV," *Frontiers in Neuroscience*, vol. 6, July 2012, Art. no. 55, https://doi.org/10.3389/fnins.2012.00055.

[20] I. I. Turulin and H. Sh. Mogheer, "Method and Algorithm for Synthesis of Controlled Digital Low-Pass Butterworth Filters on the Example of A 4th Order Filter," in *2022 International Ural Conference on Electrical Power Engineering (UralCon)*, Magnitogorsk, Russian Federation, 2022, pp. 336–340, https://doi.org/10.1109/UralCon54942.2022.9906684.

[21] S. Mallat, E. Hkiri, A. M. Albarrak, and B. Louhichi, "A Synergy of Convolutional Neural Networks for Sensor-Based EEG Brain–Computer Interfaces to Enhance Motor Imagery Classification," *Sensors*, vol. 25, no. 2, Jan. 2025, Art. no. 443, https://doi.org/10.3390/s25020443.

[22] Y. Xu, H. Zhang, L. Cao, X. Shu, and D. Zhang, "A Shared Control Strategy for Reach and Grasp of Multiple Objects Using Robot Vision and Noninvasive Brain–Computer Interface," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 1, pp. 360–372, Jan. 2022, https://doi.org/10.1109/TASE.2020.3034826.

[23] M. Miao, W. Hu, H. Yin, and K. Zhang, "Spatial-Frequency Feature Learning and Classification of Motor Imagery EEG Based on Deep Convolution Neural Network," *Computational and Mathematical Methods in Medicine*, vol. 2020, no. 1, July 2020, Art. no. 1981728, https://doi.org/10.1155/2020/1981728.

[24] A. A. Zayed and A. I. Siam, "A YOLO-based deep learning approach for vibrationbased rotating shaft imbalance detection," *Journal of Engineering Research*, vol. 9, no. 1, pp. 1–12, Mar. 2025, https://doi.org/10.21608/jer.2025.499011.

[25] A. Raza and M. Z. Yusoff, "Development of a CNN-LSTM Deep Learning Model for Motor Imagery EEG Classification for BCI Applications," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 22705–22711, June 2025, https://doi.org/10.48084/etasr.9945.

[26] X. Wang, M. Hersche, M. Magno, and L. Benini, "MI-BMInet: An Efficient Convolutional Neural Network for Motor Imagery Brain–Machine Interfaces With EEG Channel Selection," *IEEE Sensors Journal*, vol. 24, no. 6, pp. 8835–8847, Mar. 2024, https://doi.org/10.1109/JSEN.2024.3353146.