Development of a CNN-LSTM Deep Learning Model for Motor Imagery EEG Classification for BCI Applications

Aaqib Raza
Centre for Intelligent Signal & Imaging Research (CISIR), Electrical & Electronic Engineering Department, Universiti Teknologi PETRONAS, Malaysia
aaqib_24000250@utp.edu.my

Mohd Zuki Yusoff
Centre for Intelligent Signal & Imaging Research (CISIR), Electrical & Electronic Engineering Department, Universiti Teknologi PETRONAS, Malaysia
mzuki_yusoff@utp.edu.my (corresponding author)

Received: 22 December 2024 | Revised: 7 February 2025 and 18 February 2025 | Accepted: 21 February 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | *DOI: https://doi.org/10.48084/etasr.9945*

ABSTRACT

Brain-Computer Interface (BCI) systems offer a groundbreaking method for the human brain to directly communicate with external devices, serving applications, such as assistive technology, smart environments, and healthcare. Motor Imagery (MI) brain signals derived from Electroencephalography (EEG) are commonly utilized in various BCI fields. However, accurately classifying MI-based EEG signals remains a significant challenge, with traditional classification techniques struggling to effectively capture both spatial and temporal features, resulting in suboptimal performance. Therefore, this study introduces a novel hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) framework designed for EEG-MI task classification. The model combines adaptive learning with optimal training to significantly improve classification performance using the Berlin BCI Dataset 1 from BCI Competition IV. The proposed CNN-LSTM model achieves a classification accuracy of 98.38% on subject independence evaluation. This research compares subject-dependent and subject-independent evaluation with traditional Machine Learning (ML) methods, such as Support Vector Machines (SVM), Random Forest (RF), and Linear Discriminant Analysis (LDA), as well as Deep Learning (DL) models, such as EEGNet, K-nearest Neighbor (KNN), and Convolutional Neural Network (CNN). Extensive evaluations and cross-validation prove the model's superior performance, thus this work sets a benchmark for real-time MI-EEG classification, offering a scalable solution for practical BCI applications.

Keywords-signal processing; Brain-Computer Interface (BCI); Motor Imagery (MI); Electroencephalography (EEG); Convolutional Neural Networks (CNNs); Long Short-Term Memory (LSTM); hybrid deep learning

I. INTRODUCTION

BCI provide a direct communication pathway between the human brain and external devices, with significant applications in fields, such as neurorehabilitation, assistive technologies, and human-computer interaction. Among the various BCI paradigms, MI based systems have garnered particular interest due to their ability to control devices through the mental rehearsal of movements. MI BCIs have a great impact on improving prosthetic control and facilitating motor recovery in patients with neurological conditions, such as stroke or spinal cord injuries. Despite these advancements, the non-stationary nature of EEG signals, coupled with high inter-subject variability and sensitivity to noise, complicates the task of achieving reliable real-time classification, which is crucial for practical BCI applications [1].

To address the challenges of decoding MI-EEG signals, researchers have proposed several traditional ML and DL approaches. The traditional ML process typically has three steps: preprocessing, feature extraction, and classification. The preprocessing technique can efficiently eliminate interference and artifacts from EEG data while maintaining the original information. Signal processing algorithms that can extract frequency, spatial frequency, or time-frequency features from MI-EEG signals include Power Spectral Density (PSD), Short-time Fourier Transform (STFT), Common Spatial Pattern (CSP), and Filter Bank CSP (FBCSP) [2]. Commonly used feature classification methods include RF [2], SVM [3], LDA [4], and KNN [5], which are used to classify extracted features. However, categorization performance is limited by its labor-intensive nature and the high level of knowledge required for classical ML.

In recent years, research utilizing DL to identify MI tasks has developed rapidly, drawing from the successful applications of DL in other domains [6]. Statistical analysis indicates that among these DL models, CNN has emerged as predominant [7-12]. EEGNet is also one of the prominent techniques in recent times [13-15]. Researchers have worked on various DL models, such as deep neural networks, which have high accuracy but need fine hyperparameter tuning [16-18]. Other models revolve around transfer learning, which has high computational complexity but requires a vast quantity of data [19-21], Autoencoders (AE) and Deep Metric Learning (DML) are unsupervised feature learning techniques, but both possess the potential to generate disintegrating or overstating gradients [22].

In recent years, hybrid DL models have emerged as promising tools for EEG-based MI classification, combining the spatial feature extraction capabilities of CNN with the temporal modeling strength of Long Short-Term Memory (LSTM) networks [23]. These models have demonstrated robust performance in extracting spatiotemporal features and achieving high classification accuracy by leveraging techniques, such as transfer learning and data augmentation [24-26]. However, these approaches often face challenges related to limited dataset sizes, restricting generalizability across populations. Current studies have sought to address these limitations by incorporating attention mechanisms and lightweight architectures, such as BiLSTM, with inception models, which improve feature extraction and reduce computational demands [27, 28]. Nevertheless, the reliance on fixed hyperparameters in such models limits adaptability to diverse EEG signals, particularly in real-time settings. Similarly, multi-scale feature fusion methods using transfer learning have shown good results but require extensive parameter tuning, reducing their practicality for real-world BCI applications [19], while regarding pre-trained CNNs and multi-layered fusion models, despite having improved classification rates, their scalability and noise robustness remain inadequate for practical deployment [28].

Existing studies, while incorporating advanced techniques, face challenges due to the reliance on fixed hyperparameters, extensive parameter tuning, and inadequate scalability or noise robustness, hindering adaptability to diverse EEG signals and reducing practicality for real-time and real-world BCI applications. To overcome these limitations this paper introduces a novel CNN-LSTM-based framework designed specifically to enhance real-time MI classification. The key contributions of this work are:

- Developing and fine-tuning a CNN-LSTM model to achieve state-of-the-art accuracy in a three-class MI classification task (left hand, right hand, and foot).
- Leverage the CNN for spatial feature extraction and the LSTM for capturing temporal dependencies in EEG signals, ensuring an efficient combination for robust MI decoding.
- Enhancing model generalizability across subjects through domain adaptation, subject-independent training, and regularization strategies.

II. METHODOLOGY

This study develops and evaluates a hybrid CNN-LSTM model for classifying EEG-based MI signals into three distinct categories: foot, left hand, and right hand. The research workflow is outlined in Figure 1.

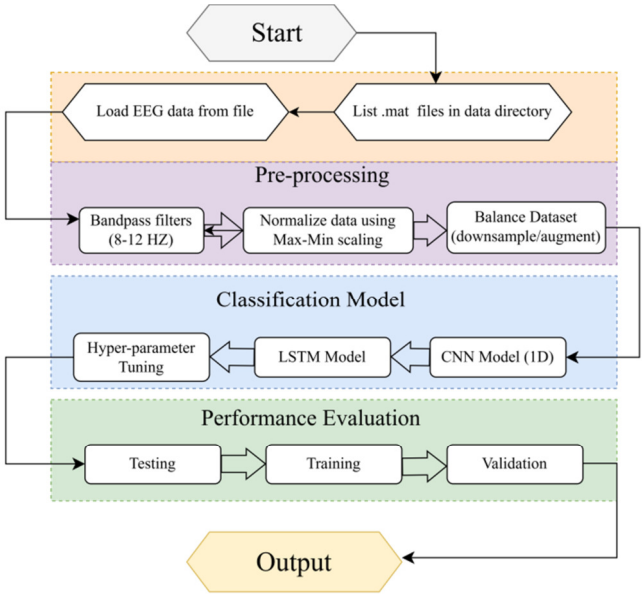

Fig. 1. Flowchart diagram of the proposed work.

Initially, data are prepared for model training by performing signal extraction, data normalization, data balancing, and filtering. Secondly, the model development stage starts with building a CNN model with an optimal number of layers, filters, kernel sizes, and activation functions. The CNN model is then integrated with an LSTM model. Thirdly, performance evaluation is performed separately on subject-dependent and independent bases through the training, validation, and testing phases.

A. Dataset

The Berlin BCI Dataset 1 [1] from the BCI Competition IV was utilized, which contains continuous EEG signals recorded from 59 sensorimotor sites filtered (0.05–200 Hz) and digitized at 1,000 Hz (16-bit, 0.1 µV) over several MI tasks from seven subjects. For each subject, two MI classes were selected from three available categories: left hand, right hand, and foot. During data collection, subjects were instructed to imagine moving their left hand, right hand, or foot. During data collection, subjects were instructed to mentally simulate movements of the left hand, right hand, or foot. Foot MI trials allowed subjects to choose either foot freely; however, the dataset labels foot MI generically as "foot," without specifying laterality. While potential differences exist between left-foot and right-foot MI, the critical requirement is that foot MI can be reliably distinguished from hand MIs. The dataset consists of seven files named a (left hand-foot), b (left-right hand), c (left-right hand), d (left-right hand), e (left-right hand), f (left hand-foot), and g (left-right hand). Dataset files c, d, and e were artificially generated for competition purposes. These

artificial datasets were excluded from this research analysis to maintain a focus on real human data. Therefore, the proposed model and reproduced models were analyzed only on data from the files a, b, f, and g.

B. Preprocessing Raw Data

The raw EEG signals were first preprocessed to extract the most relevant features for MI classification. The data were passed through a bandpass filter with a frequency range of 8-30 Hz to isolate the Mu (8-12 Hz) and Beta (13-30 Hz) rhythms, known to be associated with MI [29]. The filtering process is mathematically described as:

$$X_{filtered}(t,c) = \sum_{k=0}^{N} b_k X(t-k,c) - \sum_{k=1}^{N} a_k Y(t-k,c) \quad (1)$$

In this equation, b_k and a_k are the filter's coefficients. In the second step, segmenting the data into smaller windows centered on the cue positions for MI tasks is performed using a typical segment length of 200 timesteps:

$$X_{segment} = X_{norm}[\rho - 100 : \rho + 100, :] \quad (2)$$

where ρ represents the cue position. The dataset was also inspected for class imbalance, and a resampling technique called Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the classes by upsampling the minority classes. SMOTE generates synthetic samples for the minority classes, resulting in a balanced class distribution and improving model training effectiveness across all categories.

C. Proposed Classification Model

The designed hybrid CNN + LSTM model extracts both spatial and temporal features from the EEG-MI data. Table I shows the CNN-LSTM model summary, which includes information, such as output shape, layer type, and the number of parameters. Moreover, the proposed model utilized categorical cross-entropy for the loss function and the Adam optimizer.

1) Convolutional Neural Network

The CNN layers extract spatial features from EEG signals, enabling local pattern recognition. The architecture consists of three convolutional layers with a kernel size of 3, progressively increasing the number of filters from 64 to 256. Each convolutional layer is followed by batch normalization to stabilize training and max-pooling layers to reduce dimensionality. The ReLU activation function ensures non-linearity and efficient feature extraction. The CNN module effectively handles spatial dependencies, preparing the data for sequential analysis.

2) Long Short-Term Memory

LSTM layers capture temporal dependencies in EEG signals, crucial for MI classification. The first LSTM layer outputs sequences, passing them to a second LSTM layer for final representation. Each LSTM unit uses Tanh and Sigmoid activations to process sequential patterns while addressing vanishing gradient issues. This structure enables robust modeling of time-series data, critical for distinguishing MI tasks.

TABLE I. MODEL SUMMARY OF CNN-LSTM FRAMEWORK

Layer (type)	Output shape	Parameters
Conv1D-1	(N,198,64)	11,392
Batch normalization	(N,198,64)	256
Max-pooling1D	(N,99,64)	0
Conv1-2	(N,97,128)	24,704
Batch normalization	(N,97,128)	512
Max-pooling1D	(N,48,128)	0
Conv1D-3	(N,46,256)	98,560
Batch normalization	(N,46,256)	1,024
Max-pooling1D	(N,23,256)	0
LSTM-1	(N,23,128)	197,120
LSTM-2	(N,64)	49,408
Dense-1	(N,128)	8320
Dropout	(N,128)	0
Dense-2	(N,3)	387
Total number of parameters	1,173,259	
Trainable parameters:	390,787	
Non-trainable parameters	896	
Optimizer parameters	781,576	

3) Convolutional Neural Network-Long Short-Term Memory Classifier

The EEG data matrix is $X_i \in \mathbb{R}^{(T \times C)}$, where T represents the number of timesteps, and C represents the number of channels (in our case, $C = 59$), while $y_i \in \{-1, 0, 1\}$ are the labels for each time segment, where -1 represents left-hand MI, 1 represents right-hand MI, and 0 represents the foot MI [28].

First, the spatial features are extracted from our EEG data using 1D convolutional layers. The convolution operation for each filter was:

$$h^{(l)}(t) = \sigma(\sum_{k=0}^{K-1} W_k^{(l)} X^{(l-1)}(t+k) + b^{(l)}) \quad (3)$$

where $h^{(l)}(t)$ is the output at time t in the 1st layer, $W_k^{(l)}$ is the filter's weight matrix at the layer l, while k is the kernel size ($k = 3$), σ is the activation function (ReLU), and $b^{(l)}$ is the bias term.

After each convolutional layer, the model applied max-pooling to reduce the dimensions of input but still preserve the most prominent features [23]:

$$h_{pooled}^{(l)}(t) = \max(h^{(l)}(t : t + pool_{size})) \quad (4)$$

Secondly, the model processed the spatial features extracted by CNN-1D over time using LTSM layers. This is used to capture temporal dependencies. The model used 2 layers of LTSM for this purpose. The designed model's LSTM layer 1 has 128 units, and it returns sequences for the next LSTM layer. LSTM layer 2 has 64 units and outputs a single hidden state for the fully connected layer. The following equations define f_t forget gates, i_t input gates, and o_t output gates of the used LTSM layer:

$$f_t = \sigma(W_f . [h_{t-1}, X_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i . [h_{t-1}, X_t] + b_i) \quad (6)$$

$$o_t = \sigma(W_o \,.\, [h_{t-1}, X_t] + b_o) \quad (7)$$

$$\mathbb{C}_t = tanh(W_C \,.\, [h_{t-1}, X_t] + b_C) \quad (8)$$

where:

- X_t: Current input at time step *t*.
- h_{t-1}: Hidden state from the previous time step.
- W_f, W_i, W_i, W_C : Weight matrices for forget, input, output, and candidate cell state.
- b_f, b_i, b_o, b_C: Bias terms for respective gates.
- $tanh$: Outputs values.

Equations (9) and (10) define c_t cell and h_t hidden state of the used LTSM layer:

$$c_t = f_t \,.\, c_{t-1} + i_t \,.\, \mathbb{C}_t \quad (9)$$

$$h_t = o_t \,.\, \tanh(c_t) \quad (10)$$

where c_{t-1} is the previous cell state and $\tanh(c_t)$ is the non-linear transformation of the cell state and o_t output gates.

After the LSTM layers, the model included a fully connected layer with 128 units and a ReLU activation function to further capture complex features from the LSTM outputs. The dropout rate was set to 0.5. The proposed model has $h_{dense} = \sigma(W_{dense} \,.\, h_{LTSM} + b_{dense})$ layer to add nonlinearity to the model. Lastly, the dropout layer is applied to prevent overfitting. The final layer is a softmax output layer that classifies the data into 3 classes:

$$\mathcal{Y}_i = \frac{e^{z_i}}{\sum_{j=1}^{3} e^{z_j}} \quad (11)$$

where z_i is the log output for class i, $\mathcal{Y}_i$ is the probability of class i, and the sum of all class probabilities equals 1. The model also used (12) sparse categorical cross-entropy as the loss function [29]:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} \mathcal{Y}_{i,j} \log(\hat{y}_{i,j}) \quad (12)$$

where:

- N: The number of samples.
- C: Total number of classes.
- $\mathcal{Y}_{i,j}$: The true class label for the i^{th} sample and j^{th} class.
- $\hat{y}_{i,j}$: Predicted probability for the i^{th} sample belonging to the j^{th} class.
- $\log \hat{y}_{i,j}$: Natural logarithm of the predicted probability for the true class.

The model used an Adam optimizer with a learning rate $\alpha = 0.0001$. The optimization process with model parameter θ updates the weights using gradients of the loss function:

$$\theta_{t+1} = \theta_t - \alpha \nabla_\theta L(\theta_t) \quad (13)$$

The model implemented early stopping to prevent overfitting and saved the model with the lowest validation loss in Keras. The proposed CNN-LSTM classifier is shown in Figure 2.

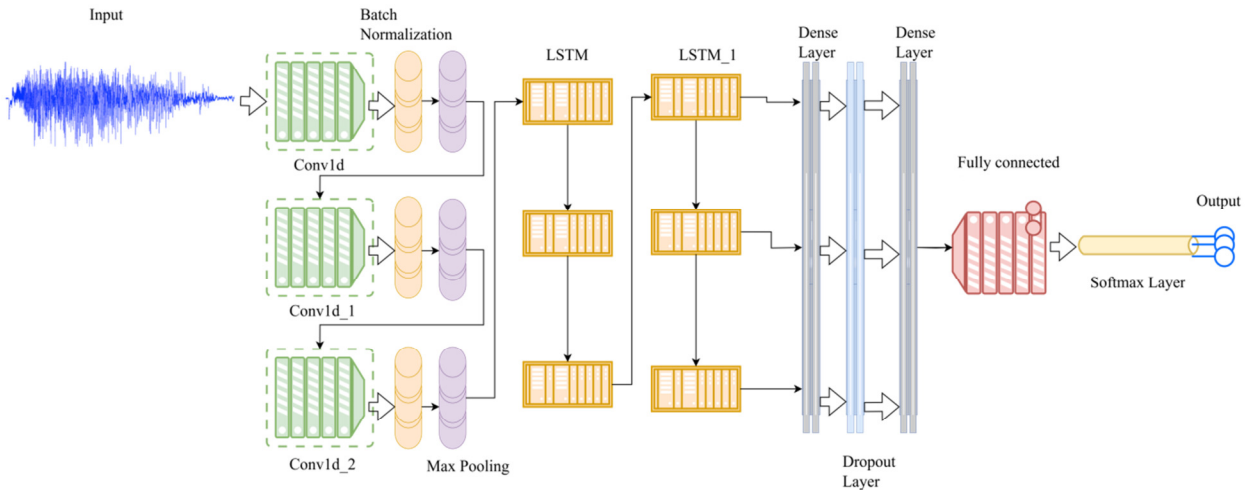

Fig. 2. The proposed CNN-LSTM classifier.

III. RESULTS AND DISCUSSION

The proposed method is used to classify three classes with combinations of two MI tasks (left hand - foot, and left hand - right hand). The accuracy, precision, recall, F1-score, and the Kappa value are calculated for each file and the overall combined files. The performance comparison presents subject dependence and subject independence performed on the same standard analysis setup with various reproduced models, such as SVM, RF, LDA, EEGNet, KNN, Standard CNN, and the proposed model. For cross-validation, four previously researched techniques that were close to the proposed classifier were selected and studied.

A. Training and Validation Performance

The model was trained for a maximum of 50 epochs, with early stopping implemented to halt training when the validation loss ceased improving, as depicted in Figure 3. The initial training epoch exhibited relatively low performance, with an accuracy of 36.77% and a loss of 4.1301. As the training progressed, the cross-entropy loss decreased, indicating that the predicted probabilities for the correct classes improved. The validation accuracy peaked at 98.34%, while the corresponding validation loss reached 0.2534. These results suggest that the model generalized well to unseen validation data and avoided overfitting due to regularization techniques, such as dropout and L2 regularization.

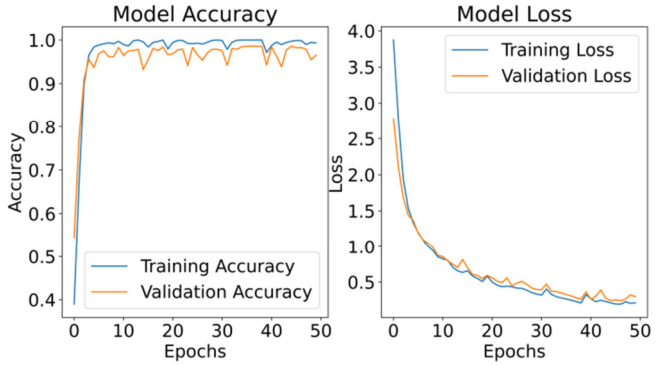

Fig. 3. Evolution of training and validation accuracy and losses.

B. Accuracy, Precision, Recall, F1-score, and Kappa Value

In this section, the performance of the proposed CNN-LSTM model was compared with other reproduced models, including, RF [2], SVM [3], LDA [4], KNN [5], CNN [7-9], and EEGNet [13-15]. These models were preprocessed, trained, and evaluated according to the methods specified in this study, while their results depend on the parameters specified within the original papers. Table II demonstrates that the proposed model consistently outperforms other approaches, achieving superior accuracy, precision, recall, F1-score, and Kappa values. Notably, the proposed model attains a maximum accuracy of 98.38% and a Kappa value of 0.95, signifying its robustness and reliability in MI-EEG classification. Table III provides the average performance of the proposed model across each file independent of subjects. The calculations are based on the following equations:

$$Accuracy = \frac{\sum_{i=1}^{n} TPi/Ii}{} \quad (14)$$

$$Precision = \frac{n}{i}\sum_{i=1}^{n}\frac{TP_i}{TP_i+FP_i} \quad (15)$$

$$Recall = \frac{n}{i}\sum_{i=1}^{n}\frac{TP_i}{TP_i+FN_i} \quad (16)$$

$$F1 = \frac{n}{i}\sum_{i=1}^{n}\frac{2\times TP_i}{2\times TP_i+FP_i+FN_i} \quad (17)$$

$$Kappa = \frac{1}{n}\sum_{a=1}^{n}\frac{Pa-Pe}{1-Pe} \quad (18)$$

Support represents the total number of samples used for evaluation in each file. It indicates the count of true instances in the dataset for which the accuracy, precision, recall, and F1-score are computed.

TABLE II. SUBJECT DEPENDENT CLASSIFICATION EVALUATION OF THE PROPOSED CNN-LSTM MODEL AND REPRODUCED MODEL

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Support	Kappa (%)
File a						
SVM	70	72	67	69.3	780	0.60
RF	75	77	76	77.1	780	0.65
LDA	81	90	75	81.8	780	0.70
EEGNet	86	89	81	84.7	780	0.75
KNN	87	88	90	89.1	780	0.78
CNN	90	92	89	90.4	780	0.85
Proposed	**97**	**97**	**96**	**97.0**	**780**	**0.94**
File b						
SVM	71	71	68	69.9	800	0.65
RF	76	78	77	77.9	800	0.69
LDA	80	89	74	80.8	800	0.74
EEGNet	84	87	89	86.7	800	0.72
KNN	88	89	91	90.2	800	0.81
CNN	92	94	91	92.4	800	0.89
Proposed	**98**	**98**	**97**	**97.5**	**800**	**0.95**
File f						
SVM	70.8	72.5	67	69.3	810	0.62
RF	75.7	77.9	76	77.3	810	0.64
LDA	82	91	76	81.9	810	0.78
EEGNet	86.3	89.5	81	84.7	810	0.79
KNN	86	87	89	89.4	810	0.76
CNN	91	92.1	88	90.9	810	0.89
Proposed	**99**	**99**	**99**	**99.0**	**810**	**0.98**
File g						
SVM	70.1	72.8	66	69.8	780	0.61
RF	75.9	77.6	77	76.1	780	0.60
LDA	81	90	75	81.8	780	0.70
EEGNet	85	88	80	83.1	780	0.70
KNN	87.9	88.1	90	89.7	780	0.77
CNN	89	91	88	90.9	780	0.82
Proposed	**98**	**98**	**97**	**98.0**	**780**	**0.95**

TABLE III. AVERAGE SUBJECT-INDEPENDENT CLASSIFICATION EVALUATION OF THE PROPOSED CNN-LSTM MODEL

File	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Kappa (%)
a	0.97	0.97	0.96	0.97	0.94
b	0.97	0.98	0.97	0.96	0.95
f	0.99	0.99	0.99	0.99	0.98
g	0.98	0.98	0.97	0.97	0.95

C. Performance Comparison

Each model was evaluated based on identical preprocessing, training, and testing protocols to ensure a fair comparison. The results in Table IV demonstrate that the proposed model significantly outperforms the reproduced models across all evaluation metrics. It also provides a comprehensive performance analysis that highlights the robustness and efficiency of the proposed architecture in achieving state-of-the-art results.

TABLE IV. SUBJECT INDEPENDENT EVALUATION OF REPRODUCED MODELS AND PROPOSED CNN-LSTM MODEL

Models	Accuracy	Precision	Recall	F1-Score (%)	Support
SVM	0.70	0.72	0.67	69.3	792.5
RF	0.75	0.77	0.76	77.1	792.5
LDA	0.81	0.90	0.75	81.8	792.5
EEGNet	0.86	0.89	0.81	84.7	792.5
KNN	0.87	0.88	0.90	89.1	792.5
CNN	0.90	0.92	0.89	90.4	792.5
Proposed	**0.98**	**0.99**	**0.98**	**97.87**	**792.5**

D. Confusion Matrices

The performance metrics are evaluated using confusion matrices to provide detailed insights into how well each model predicted the respective classes. The confusion matrices offer a class-wise breakdown of the predictions, highlighting the strengths and weaknesses of all the reproduced models and the proposed CNN-LSTM model. In Figure 4, the confusion matrices are presented individually to demonstrate the performance of each model under the same experimental setup, enabling a clear comparison of their classification accuracy and error distribution.

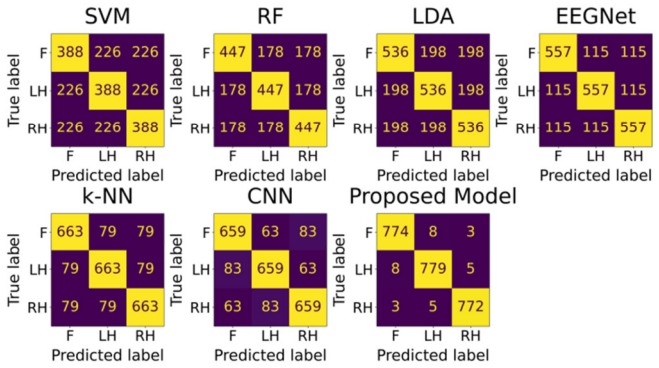

Fig. 4. Confusion matrix of reproduced models and proposed model. Foot (F), Left Hand (LH), Right Hand (RH).

E. Cross-Validation Accuracy Comparison

This section presents the cross-validation of the proposed CNN-LSTM model against four prior research techniques, shown in Table V, which closely align with the characteristics of this study's classifier, as shown in Figure 5. The proposed model with an average accuracy of 98% surpasses the performance of the models in [24, 26, 28, 29] which achieved an accuracy of 95.62%, 86%, 87.68%, and 93.3%, respectively. This comparison emphasizes the proposed model's adaptability and performance in decoding MI tasks, thus reinforcing its robustness.

IV. CONCLUSION AND FUTURE WORK

The classification of Motor Imagery (MI)-based Electroencephalography (EEG) signals remains a significant challenge due to variability in EEG signals, subject-dependence, and the complexity of real-time classification. This study proposed a hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) Deep Learning (DL) architecture for multi-class classification of MI-EEG signals in Brain-Computer Interface (BCI) applications. The robustness of the approach is evident in its high average accuracy (0.98) precision (0.99), recall (0.99), and F1-score (97.87%) across all classes, with minimal misclassifications observed in the confusion matrix.

TABLE V. SUBJECT-INDEPENDENT PERFORMANCE COMPARISON BETWEEN PROPOSED CNN-LSTM AND RECENT STUDIES

Reference	Dataset	Classification techniques	Accuracy
[26]	BCI competition IV dataset 2a	CNN-LSTM	86%, accuracy 81% mean Kappa value
[24]	BCI competition IV dataset 2a	CNN-LSTM	95.62% accuracy 94.62% Kappa value
[28]	BCI competition IV dataset 2a	CNN-LSTM	87.68%
[29]	BCI competition IV 2b	2dCNN-LSTM	93.3%

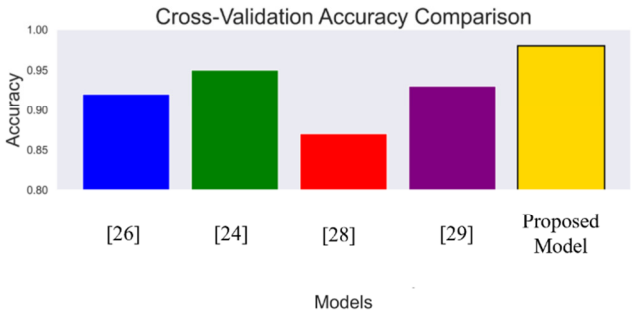

Fig. 5. Cross-validation between recent studies based on CNN-LSTM classifier and the proposed CNN-LSTM model.

The contribution of the proposed framework is an adaptive learning mechanism that dynamically adjusts feature extraction and temporal dependencies based on subject-specific MI-EEG signal variations. This adaptability enables the model to generalize effectively across diverse datasets by capturing individual variability in MI-EEG patterns. The training strategy incorporates an optimized hybrid loss function combining categorical cross-entropy with a regularization term to mitigate overfitting, ensuring robust generalization across subjects. Additionally, the model's lightweight design, features reduced parameter complexity, and optimized inference latency enhance real-time applicability, while existing frameworks struggle with fixed hyperparameters, extensive tuning, and limited scalability.

Future work on dataset diversity and inter-subject adaptation will ensure the continued refinement and practical deployment in real-world BCI systems.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi PETRONAS (UTP) for providing a Graduate Assistant Scheme, and the Centre for Intelligent Signal & Imaging Research (CISIR) for providing all the necessary facilities to complete this research.

REFERENCES

[1] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin Brain–Computer Interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, Aug. 2007, https://doi.org/10.1016/j.neuroimage.2007.01.051.

[2] 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.

[3] M. Norizadeh Cherloo, H. Kashefi Amiri, and M. R. Daliri, "Ensemble Regularized Common Spatio-Spectral Pattern (ensemble RCSSP) model for motor imagery-based EEG signal classification," *Computers in Biology and Medicine*, vol. 135, Aug. 2021, Art. no. 104546, https://doi.org/10.1016/j.compbiomed.2021.104546.

[4] P. Rithwik, V. K. Benzy, and A. P. Vinod, "High accuracy decoding of motor imagery directions from EEG-based brain computer interface using filter bank spatially regularised common spatial pattern method," *Biomedical Signal Processing and Control*, vol. 72, Feb. 2022, Art. no. 103241, https://doi.org/10.1016/j.bspc.2021.103241.

[5] A. Apicella, P. Arpaia, M. Frosolone, and N. Moccaldi, "High-wearable EEG-based distraction detection in motor rehabilitation," *Scientific Reports*, vol. 11, no. 1, Mar. 2021, Art. no. 5297, https://doi.org/10.1038/s41598-021-84447-8.

[6] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "EEG Classification of Motor Imagery Using a Novel Deep Learning Framework," *Sensors*, vol. 19, no. 3, Jan. 2019, Art. no. 551, https://doi.org/10.3390/s19030551.

[7] Z. Cai, T. Luo, and X. Cao, "Multi-branch spatial-temporal-spectral convolutional neural networks for multi-task motor imagery EEG classification," *Biomedical Signal Processing and Control*, vol. 93, Jul. 2024, Art. no. 106156, https://doi.org/10.1016/j.bspc.2024.106156.

[8] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network," *Neural Networks*, vol. 136, pp. 1–10, Apr. 2021, https://doi.org/10.1016/j.neunet.2020.12.013.

[9] Y. K. Musallam *et al.*, "Electroencephalography-based motor imagery classification using temporal convolutional network fusion," *Biomedical Signal Processing and Control*, vol. 69, Aug. 2021, Art. no. 102826, https://doi.org/10.1016/j.bspc.2021.102826.

[10] Y. Han, B. Wang, J. Luo, L. Li, and X. Li, "A classification method for EEG motor imagery signals based on parallel convolutional neural network," *Biomedical Signal Processing and Control*, vol. 71, Jan. 2022, Art. no. 103190, https://doi.org/10.1016/j.bspc.2021.103190.

[11] J. F. Hwaidi and T. M. Chen, "Classification of Motor Imagery EEG Signals Based on Deep Autoencoder and Convolutional Neural Network Approach," *IEEE Access*, vol. 10, pp. 48071–48081, 2022, https://doi.org/10.1109/ACCESS.2022.3171906.

[12] X. Liu, S. Xiong, X. Wang, T. Liang, H. Wang, and X. Liu, "A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification," *Biomedical Signal Processing and Control*, vol. 81, Mar. 2023, Art. no. 104456, https://doi.org/10.1016/j.bspc.2022.104456.

[13] D. Park *et al.*, "Spatio-Temporal Explanation of 3D-EEGNet for Motor Imagery EEG Classification Using Permutation and Saliency," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 4504–4513, 2023, https://doi.org/10.1109/TNSRE.2023.3330922.

[14] X. Deng, B. Zhang, N. Yu, K. Liu, and K. Sun, "Advanced TSGL-EEGNet for Motor Imagery EEG-Based Brain-Computer Interfaces," *IEEE Access*, vol. 9, pp. 25118–25130, 2021, https://doi.org/10.1109/ACCESS.2021.3056088.

[15] R. Peng *et al.*, "TIE-EEGNet: Temporal Information Enhanced EEGNet for Seizure Subtype Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2567–2576, 2022, https://doi.org/10.1109/TNSRE.2022.3204540.

[16] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, vol. 14, no. 1, Feb. 2017, Art. no. 016003, https://doi.org/10.1088/1741-2560/14/1/016003.

[17] S. A. C. Yohanandan, I. Kiral-Kornek, J. Tang, B. S. Mshford, U. Asif, and S. Harrer, "A Robust Low-Cost EEG Motor Imagery-Based Brain-Computer Interface," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2018, pp. 5089–5092, https://doi.org/10.1109/EMBC.2018.8513429.

[18] N. S. Suhaimi, M. Z. Yusoff, and M. N. M. Saad, "Artificial Neural Network Analysis On Motor Imagery Electroencephalogram," in *2022 IEEE 5th International Symposium in Robotics and Manufacturing Automation (ROMA)*, Malacca, Malaysia, Aug. 2022, pp. 1–6, https://doi.org/10.1109/ROMA55875.2022.9915671.

[19] A. M. Roy, "Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for EEG MI multiclassification in brain–computer interface," *Engineering Applications of Artificial Intelligence*, vol. 116, Nov. 2022, Art. no. 105347, https://doi.org/10.1016/j.engappai.2022.105347.

[20] 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.

[21] D. Li, J. Wang, J. Xu, X. Fang, and Y. Ji, "Cross-Channel Specific-Mutual Feature Transfer Learning for Motor Imagery EEG Signals Decoding," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 10, pp. 13472–13482, Oct. 2024, https://doi.org/10.1109/TNNLS.2023.3269512.

[22] P. Autthasan *et al.*, "MIN2Net: End-to-End Multi-Task Learning for Subject-Independent Motor Imagery EEG Classification," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 6, pp. 2105–2118, Jun. 2022, https://doi.org/10.1109/TBME.2021.3137184.

[23] Y. Hu *et al.*, "A Cross-Space CNN With Customized Characteristics for Motor Imagery EEG Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1554–1565, 2023, https://doi.org/10.1109/TNSRE.2023.3249831.

[24] C. Uyulan, "Development of LSTM&CNN based hybrid deep learning model to classify motor imagery tasks," *Communications in Mathematical Biology and Neuroscience*, vol. 2021, no. 4, 2021, https://doi.org/10.28919/cmbn/5265.

[25] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A review of critical challenges in MI-BCI: From conventional to deep learning methods," *Journal of Neuroscience Methods*, vol. 383, Jan. 2023, Art. no. 109736, https://doi.org/10.1016/j.jneumeth.2022.109736.

[26] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," *Computers in Biology and Medicine*, vol. 143, Apr. 2022, Art. no. 105288, https://doi.org/10.1016/j.compbiomed.2022.105288.

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

[28] H. Li, M. Ding, R. Zhang, and C. Xiu, "Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network," *Biomedical Signal Processing and Control*, vol. 72, Feb. 2022, Art. no. 103342, https://doi.org/10.1016/j.bspc.2021.103342.

[29] J. Wang, S. Cheng, J. Tian, and Y. Gao, "A 2D CNN-LSTM hybrid algorithm using time series segments of EEG data for motor imagery classification," *Biomedical Signal Processing and Control*, vol. 83, May 2023, Art. no. 104627, https://doi.org/10.1016/j.bspc.2023.104627.