Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6, 8873-8891
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i6.3884

Exploring motor imagery classification in brain-computer interfaces: Case study of the Vietnamese BCI speller

Chau Ma Thi[1*], Long Vu Thanh[1], Kien Nguyen Minh[1]
[1]University of Engineering and Technology, Vietnam National University; chaumt@vnu.edu.vn (C.M.T.)
21020647@vnu.edu.vn (L.V.T.) 21020638@vnu.edu.vn (K.N.M.).

Abstract: Recent advancements in Brain-Computer Interface technology have led to the development of BCI Spellers, providing innovative communication solutions, particularly for individuals with disabilities. This paper presents a Vietnamese BCI Speller system specifically designed for Vietnamese speakers, addressing the unique linguistic and contextual needs of this population. Our research focuses on two primary contributions: first, we evaluate and implement a motor imagery classification method using a scalable machine learning framework tailored for processing motor imagery signals; second, we propose a user-friendly BCI Speller that facilitates document drafting and communication at an affordable cost. We detail the integration of motor imagery task classification within our Vietnamese BCI Speller, employing both deep learning and traditional machine learning models. Experimental results demonstrate over 90% accuracy in classifying motor imagery tasks with Linear Discriminant Analysis and Common Spatial Parttern, highlighting the system's effectiveness and potential for enhancing user interaction. This study emphasizes the need for continued research to optimize BCI technology and improve accessibility for all users.

Keywords: *Brain-computer Interface (BCI), BCI Speller, Electroencephalogram (EEG), Motor imagery (MI), Vietnamese language.*

1. Introduction

In recent years, the BCI Speller has gained considerable attention in the scientific community [1-6]. This technology provides a modern and flexible means of communication and interaction in the digital age, particularly empowering individuals with disabilities to engage more actively in their lives. By leveraging brain electrical signals, the BCI Speller allows users to control activities through a virtual keyboard, enabling them to draft documents, communicate, and interact.

Furthermore, integrating virtual keyboards with physiological measurement devices offers sensory feedback that enhances user capabilities. However, challenges persist in designing user-friendly and effective interfaces, addressing the high costs of equipment, and overcoming the lack of EEG datasets. Consequently, further research is essential to gain a deeper understanding of the underlying mechanisms of the BCI Speller and to optimize its performance.

This research presents a Vietnamese BCI Speller system specifically developed for the Vietnamese language and its speakers. Tailored to the conditions and resources of the Vietnamese population, this system has shown promising results in initial testing. Our key contributions include:

- Evaluating and selecting a motor imagery classification method that employs an appropriate machine learning mechanism to create a scalable and upgradeable module for processing motor imagery signals in the Vietnamese BCI Speller.
- A flexible and comprehensive BCI Speller system designed for users seeking additional methods for drafting documents, communication, and interaction, offered at an affordable price that aligns with the living conditions of most Vietnamese individuals.

Our research emphasizes the integration of the BCI Speller with motor imagery (MI) task classification. Backgrounds and Related works are introduced in Section 2. Section 3 details the deep

* Correspondence: chaumt@vnu.edu.vn

learning and machine learning models used for classifying MI tasks, while Section 4 discusses the experiments and evaluations related to this classification. The approach and development of the Vietnamese BCI Speller are outlined in Section 5, and conclusions are summarized in Section 6.

2. Backgrounds and Related Works

2.1. Backgrounds

2.1.1. Brain - Computer Interface

The Brain-Computer Interface (BCI) system utilizes Electroencephalography (EEG) signal amplifiers and a computer to monitor and analyze brain activity. It processes these signals to generate inputs for interactive applications. For general users, a BCI system can easily be misinterpreted as a device capable of "reading minds." However, the human brain remains a profound mystery despite significant scientific advancements. Its complexity limits BCIs in their ability to predict specific activities. Thus, at this stage, the BCI system enables users to perform desired actions and commands directly within an interactive application through brain activity, without the need to engage muscles.
BCI systems can be classified into three types:

- ***Active BCIs:*** Asynchronous systems that are directly controlled by commands derived from the user's EEG signals.
- ***Reactive BCIs:*** Synchronous systems in which the user generates appropriate EEG signals in response to stimuli, such as visual (images), auditory (sounds), and tactile (physical stimulation from auxiliary devices like robotic arms or gloves).
- ***Passive BCIs:*** Systems designed to recognize and evaluate the user's cognitive responses based on brain activity, including emotions and levels of concentration.

2.1.2. Electroencephalogram

EEG signals are one type of bio electrical signals obtained from the human body. Essentially, they are electrical potential signals generated during cortical activity. These signals are the instantaneous bio-electrical responses of brain cells.

2.1.2.1. Methods for Acquiring EEG Signals

There are various experimental methods to capture and record EEG signals for developing BCI systems [7]. These methods are generally classified into two types: invasive and non-invasive. In invasive methods, electrodes are implanted directly into the brain tissue. Figure 1 shows common experimental methods for acquiring EEG signals include two invasive methods: Electrocorticogram (ECoG) and Intracortical recordings. Invasive methods provide higher quality signals but pose significant medical safety concerns and high costs. Non-invasive method, commonly used for clinical diagnosis of neurological disorders, sleep studies, and research, safe and easy to set up, and suitable for long-term monitoring. Electrodes are placed on the scalp using a cap or adhesive pads.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

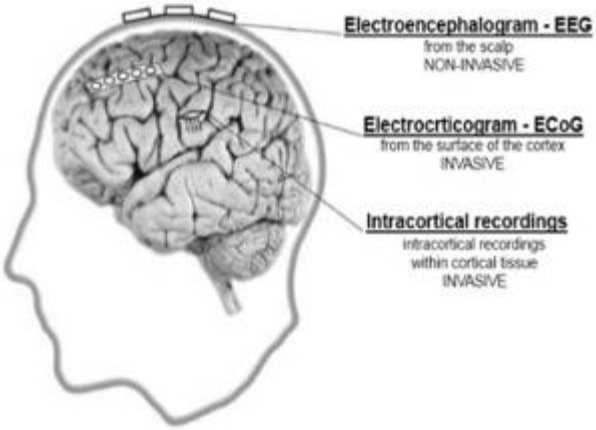

Figure 1:
Methods for acquiring EEG signals [7].

2.1.2.2. Characteristics of EEG Signals

EEG measures electrical activities in the brain by recording potentials at electrodes attached to the user's head. These potentials, ranging from 100μV to +100μV, form continuous signals known as brain waves. EEG signals offer high temporal resolution due to modern devices' sampling rates (up to 1000Hz), but spatial resolution is limited by the number of electrodes, typically between 19 and 256. Figure 2 below illustrates the EEG signals obtained at several electrodes of the measurement device over time.

EEG signals appear as oscillations at specific cerebral cortex locations within particular frequency bands, providing information about the brain's state. When classified by frequency, they fall into six categories: Delta, Theta, Alpha, Mu, Beta, and Gamma, ranging from 0.1Hz to over 100Hz. This classification method helps in filtering out irrelevant signals during data preprocessing. Table 1 below outlines the characteristics and states of the main EEG signal types.

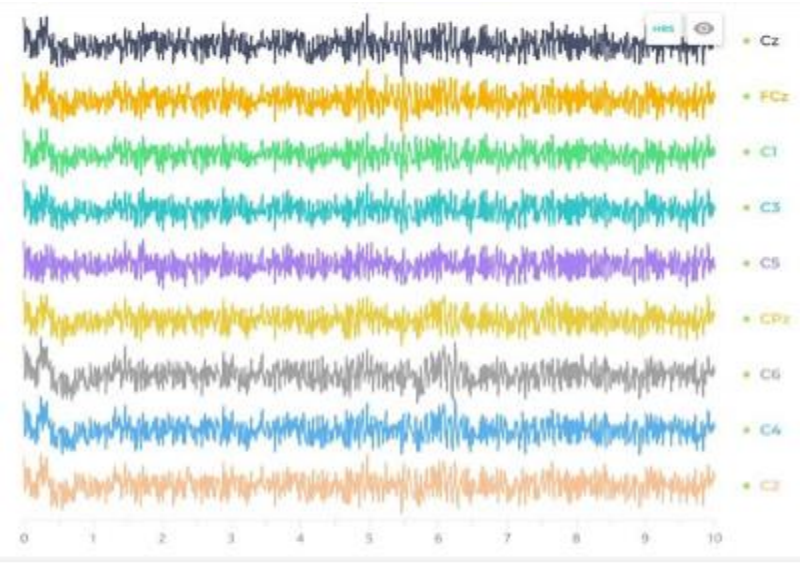

Figure 2:
Samples of EEG signals

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Table 1:
Classification of EEG signals by frequency band

Waveforms	Frequency	Characteristics
Delta	0 - 4Hz	Slow waves, characterized by their low frequency and high amplitude, are commonly observed during deep sleep, dreamless sleep (non-REM), or unconscious states.
Theta	4 - 8Hz	Appears when a person is drowsy, relaxed, or engaged in tasks involving imagination, creativity, and memory recall.
Alpha	8 - 12Hz	Appears during relaxation and with closed eyes. Alpha waves diminish when a person is alert or begins engaging in cognitive activities.
Mu	8 - 12Hz	Mu waves, with a frequency band similar to Alpha waves, primarily appear in the motor cortex region. These waves decrease when a person performs, observes, or imagines movements.
Beta	12 - 30Hz	Commonly appear when a person is awake. Associated with cognitive activities, focus, and influenced by body movements.
Gamma	> 30Hz	Highest frequency, appearing during sensory processing, high-level cognitive functions, or specific movements.

There are several limitations in collecting, preprocessing, and classifying EEG signals that can affect the final system outcomes:

Variability in Brain States: Brain states are continuously changing, resulting in non-stationary EEG signals that can vary across sessions and individuals.

Limited Depth of Signal Capture: Non-invasive EEG systems primarily capture signals from neurons near the cortex, making it challenging to analyze deep brain signals.

Impact of Noise: EEG signals often suffer from a low signal-to-noise ratio, significantly affecting data quality. Recorded potentials represent averages from groups of neurons and are influenced by various types of noise, including body movements, electromagnetic interference, and biological noise (such as eye movements).

2.1.3. Motor Imagery

Motor Imagery (MI) is an active cognitive process in which individuals imagine performing movements of body parts, such as their eyes, tongue, hands, and feet, without any actual movement. Previous research has indicated that the neural mechanisms involved in imagining movements share similarities with those engaged during actual movement, particularly observable in the motor cortex [8]. Thus, classifying MI is a critical challenge that our BCI speller system must address.

Studies have primarily examined the spatial and temporal characteristics of Mu (8-12 Hz) and Beta (12-30 Hz) rhythms during MI tasks. Research conducted by Gert Pfurtscheller and colleagues [9-12] involved participants imagining movements of the left hand, right hand, feet, and tongue in response to visual stimuli. They found that Event-Related Desynchronization (ERD) of the Mu waveform and Event-Related Synchronization (ERS) of the Beta waveform occur briefly in the primary motor cortex, particularly around electrodes C3, C4, and Cz, and are observed in most subjects [12] (Figure 3). These patterns resemble the processes involved in preparing for movement.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

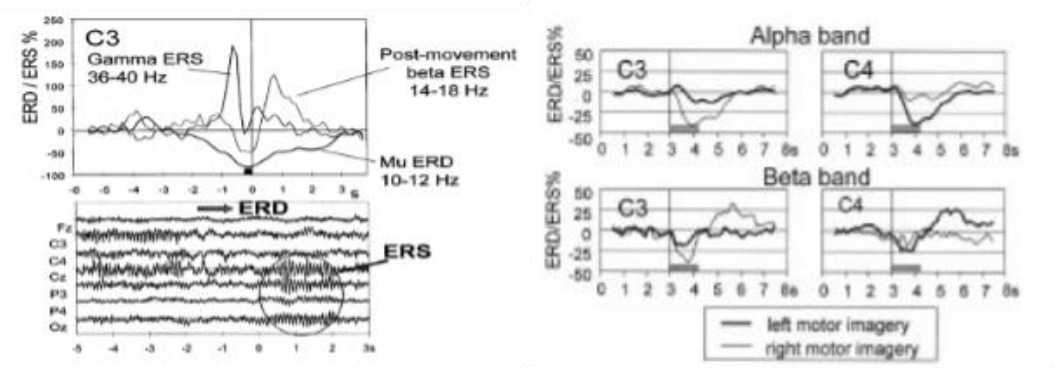

Figure 3:
ERD & ERS of Mu & Beta waveform at C3 & C4 electrode [10].

2.1.4. EEG-Atcnet

As of the time we conducted this research, EEG-ATCNet [13] was one of the models that achieved the highest accuracy on the BCI Competition IV-2a dataset [14] compared to other state-of-the-art deep learning techniques for processing EEG signals. The architecture of ATCNet consists of three main blocks (see Figure 4): Convolutional (CV), Multi-Head Self Attention (AT), and Temporal Convolutional (TC). The CV block encodes low-level spatio-temporal signals from MI-EEG data through three convolutional layers: temporal, channel depth-wise, and spatial. The output from the CV block is a high-level time sequence, which is then segmented into multiple windows using a sliding window approach and fed into separate AT and TC blocks to enhance data efficiency and improve accuracy. The AT blocks emphasize the most relevant information in the time sequence by utilizing Multi-Head Self-Attention (MSA). Finally, the TC blocks extract high-level temporal features from the time sequence using the TCNet architecture described in [15], and these features are passed to a fully connected (FC) layer with a SoftMax classification function.

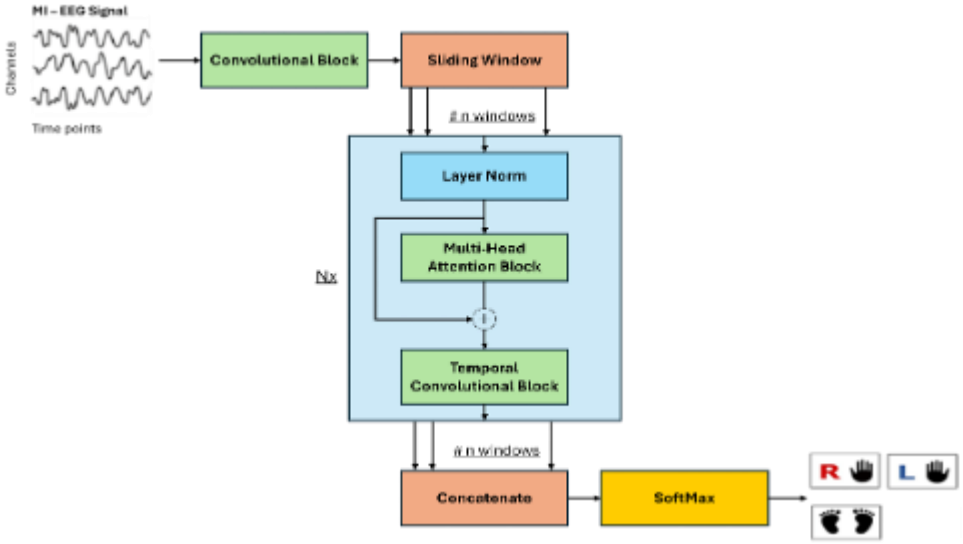

Figure 4:
Overview of EEG-ATCNet architecture

2.1.5. Common Spatial Pattern

Common Spatial Pattern (CSP) [16] is one of the most widely used techniques in biomedical signal processing, particularly in healthcare applications. It functions as a spatial filtering method designed to extract features from multi-channel biomedical signals, specifically EEG signals.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

CSP is grounded in linear algebra and multivariate statistical methods. It transforms EEG data from the time domain to the spatial domain using a spatial filter. This filter enhances the signal variance of one class of signals while simultaneously reducing it for another, generating new features or components that are linear combinations of the original channels. In a typical CSP analysis, EEG data are divided into two classes corresponding to different cognitive states (e.g., imagining the movement of a hand or leg). CSP identifies spatial weights that best separate these two classes based on their covariance matrices.

The spatial filter in CSP is typically derived by solving a generalized eigenvalue problem, which involves identifying the eigenvectors of a matrix that most effectively distinguishes between the two EEG data classes. This process includes calculating the eigenvectors of the composite covariance matrices of both classes. These eigenvectors (spatial filters) project the EEG signals into a new space where the classes are optimally separated in terms of variance. Essentially, the resulting spatial filters serve as weights that can be applied to the EEG data, producing a new set of spatially filtered data that emphasizes the differences between the two classes.

2.1.6. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) [17] works by identifying a linear combination of features (a projection) that maximizes the ratio of between-class scatter to within-class scatter. In other words, LDA seeks a direction where the classes are as dispersed as possible while keeping the data points within each class as close together as possible. After projecting the data, LDA can establish a decision boundary in this lower-dimensional space, allowing new data points to be classified based on which side of the boundary they fall on.

The combination of CSP and LDA is particularly effective for classifying MI tasks. The components extracted by CSP are designed to maximize class separation, while LDA finds a linear boundary that further distinguishes the classes. Additionally, LDA is computationally efficient compared to other methods, such as deep learning models, making it well-suited for real-time applications.

Furthermore, EEG data is often high-dimensional but typically has relatively few training samples. LDA performs well in such scenarios because it reduces dimensionality and helps avoid overfitting, making it more robust for small datasets compared to deep learning models.

2.2. Related Works

2.2.1. BCI Speller

Over the years, several types of BCI spellers have been developed, categorized by their GUI designs and BCI paradigms. Based on BCI paradigms, there are three main classes: P300 [18, 21], steady-state visual evoked potential (SSVEP) [19] and MI [20, 22].

The first BCI speller was a P300-based system called the Matrix Speller, introduced by Farwell and Donchin [18]. Their GUI was arranged as a 6 x 6 matrix of flashing symbols on a monitor. By randomly intensifying the light in each row and column, the GUI aimed to evoke an "oddball" paradigm. Subjects were instructed to focus on the desired letter and count the number of flashes, which generated a P300 wave in the EEG signals. The order of occurrence of rows and columns was then correlated with the P300 signal, allowing the identification of the selected letter at the intersection of the chosen row and column. This Matrix Speller achieved a maximum Information Transfer Rate (ITR) of 12 bits and served as the foundation for most P300-based BCI spellers, which vary in stimuli, symbol representation, and letter arrangement. Geospell [21] is one such variation. It uses an N x N matrix, further divided into 2 x N sets of squares containing N characters. The rows and columns were rearranged into separate boxes to enhance letter recognition, with flashes occurring at fixed points for each box to minimize eye movement and focus on covert attention. However, this system's results were not as promising as its counterparts.

Another approach to BCI spellers utilizes SSVEP. The advantage of this method is its independence from calibration or subject training, generally resulting in faster responses than P300 systems. An early example of SSVEP spellers is the Bremen Speller [19], featuring a diamond-shaped keyboard with 32 symbols representing English characters, punctuation marks, and commands. Five boxes, four arrows,

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

and one "Select" button control a cursor that moves along the keyboard to indicate the character to be selected. Each box flickers at a specific frequency to generate an SSVEP response. With the later addition of a prediction model and improvements in signal processing, this speller achieved an impressive ITR of over 32 bits per minute.

Despite the advantages of the aforementioned paradigms for BCI spellers, they also have notable drawbacks. Firstly, both P300 and SSVEP systems are gaze-dependent, requiring subjects to maintain their gaze on the monitors or displays for optimal results. Secondly, the flashing visual stimuli from these systems can cause irritation for users after a short session, interfering with communication processes that often last longer than typical conversations. Therefore, we focus our research on the third paradigm, MI, which is gaze-independent and requires no visual stimulation.

Hex-o-spell [20] is a prominent example of this paradigm. First developed in 2006, its interface consists of six hexagons arranged around a circle, with an arrow pointing toward the hexagons at the center. By imagining movements, such as right-hand and foot movements, the subject can rotate the arrow or select the hexagon currently being pointed to. After selection, the hexagons fill with characters from the selected segment, each corresponding to one hexagon. The user then repeats these rotating and selecting operations to type the desired text.

Despite extensive research into BCI speller development [23], almost all existing systems cater exclusively to English-speaking users. Noticing a distinct lack of localization for Vietnamese users, this paper focuses on designing a BCI speller specifically for Vietnamese patients, aiming to effectively draft Vietnamese text.

2.2.2. Motor Imagery Classification

In [1], the authors introduce a novel CSP-AM-BA-SVM approach that leverages bio-inspired algorithms for feature selection and classifier optimization to enhance motor imagery (MI) BCI classification accuracy. This method optimally selects time intervals for each subject and extracts features from EEG signals using the CSP technique, extending binary CSP to multi-class problems through a one-vs-one strategy. A hybrid attractor metagene (AM) algorithm, combined with the Bat optimization algorithm (BA), identifies the most discriminative CSP features and optimizes support vector machine (SVM) parameters. The approach was evaluated on three datasets, achieving 78.55% accuracy and a mean kappa of 0.71 for the BCI Competition IV dataset 2a, 86.6% accuracy and a mean kappa of 0.82 for the BCI Competition III dataset IIIa, and 85% accuracy for the binary class in the BCI Competition III dataset IVa. In multi-class datasets, it outperformed the winners of BCI Competition IV 2a, and BCI Competition III, IIIa, with kappa values of 0.14 and 0.17, respectively. For the binary class BCIC III, IVa, it exceeded existing studies by approximately 0.5%. Overall, the CSP-AM-BA-SVM approach surpasses traditional CSP-SVM methods and other studies in the field.

The research in [2] investigates motor imagery classification in brain-computer interfaces (BCIs) using machine learning classifiers. The BCI process involves two main steps: feature extraction and classification. Fast Fourier Transform (FFT) features are derived from EEG signals to convert them into the frequency domain. To manage the high dimensionality of the data, LDA is employed to enhance class separability. The study evaluates five classifiers: SVM, K-Nearest Neighbors (KNN), Naïve Bayes, Decision Tree, and Logistic Regression, using Dataset 1 from BCI Competition IV, which contains EEG data from imagined hand and foot movements. The SVM, Logistic Regression, and Naïve Bayes classifiers achieved the highest accuracy, with an area under the curve (AUC) of 89.09%.

The authors in [3] present a focused survey on various machine learning (ML) and deep learning (DL) techniques applied in EEG-based BCIs, specifically examining three EEG paradigms: motor imagery, P300, and steady-state evoked potential. They also address challenges faced by current EEG-based BCI systems related to signal processing, functionality, performance assessment, and commercialization. EEG is widely used due to its simplicity and non-invasiveness. The assessment of an EEG-based BCI system primarily involves classifying EEG signals for specific applications. The rise of artificial intelligence has led researchers to adopt ML and DL techniques for classifying EEG data, enabling BCIs to learn from each session and adapt classification rules to enhance system efficiency. The

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

authors aim to provide insights that will assist in selecting suitable machine learning techniques and serve as a foundation for future BCI research.

Recognizing that effective rehabilitation training depends on accurate feedback, the authors in [4] reviewed literature on current and past EEG-based BCI frameworks that report online classification of upper limb movements in both healthy individuals and stroke patients. They found that recent deep learning methods do not outperform traditional machine learning algorithms, with similar accuracy observed for both patient and healthy subject classifications. Notably, functional electrical stimulation (FES) systems demonstrated superior performance compared to non-FES systems.

The authors in [5] propose a novel approach that combines multivariate iterative filtering (MIF) with CSP (MIFCSP) to automatically select optimal frequency bands for discriminant feature extraction. MIF decomposes EEG signals into multivariate intrinsic mode functions, from which features are extracted using CSP. This method aims to identify the minimum number of significant features that achieve the highest classification accuracy, subsequently using a LDA classifier for task classification. Experimental results from the BCI Competition IV dataset 2a and BCI Competition III-IIIa show that MIFCSP achieves average accuracies of 83.18% and 84.44% for left-hand versus right-hand MI classification, respectively, indicating its potential as a promising approach for MI BCI applications.

Mahbod N. et.al. 6] introduces a novel subject-independent (SI) BCI framework called CCSPNet (Convolutional Common Spatial Pattern Network), trained on a large-scale EEG database consisting of 400 trials from 54 subjects performing two-class hand movement MI tasks. The framework employs a wavelet kernel convolutional neural network (WKCNN) and a temporal convolutional neural network (TCNN) to extract spectral features from EEG signals. A CSP algorithm is used for spatial feature extraction, followed by feature reduction via a dense neural network. The final class label is determined using an LDA classifier. Evaluation results indicate that CCSPNet achieves competitive BCI performance, rivaling both subject-dependent (SD) and SI state-of-the-art models, thus proving its effectiveness and efficiency.

3. Motor Imagery Classification

3.1. Datasets

3.1.1. BCI Competition IV-2a dataset

BCI Competition IV-2a dataset [14] contains brain signals from 9 subjects, who participated in a total of 4 MI tasks: left hand, right hand, feet, and tongue. Each subject attended two sessions on different days, labeled as Training (T) and Evaluation (E). Each session consists of 6 runs, separated by resting periods. During each run, subjects perform 48 trials, with 12 trials dedicated to each MI task. At the beginning of each trial, a cross appears at the center of the screen, accompanied by an audio notification for 2 seconds. Following this, a cue image—a directional arrow pointing left, right, up, or down—represents the MI task and is displayed for 1.25 seconds. The subject then performs the corresponding MI task until the cue image disappears. The electrodes used in this dataset are arranged as shown in Figure 5 and record data at a sampling rate of 250 Hz.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

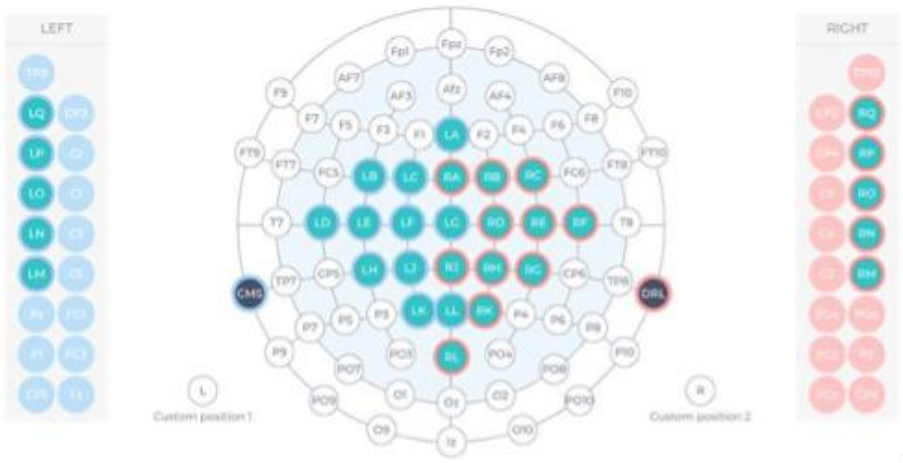

Figure 5.
Electrodes layout.

3.1.2. VKIST dataset

We conduct our study using the VKIST dataset, which includes data from different subjects over two sessions on separate days, each lasting for one hour. Each session consists of 20 runs, separated by a 20-second rest period. In each run, the subject performs 12 trials, with 3 trials for each MI task: raising the right hand, raising the left hand, raising the right foot, and raising the left foot. At the start of each trial, a notification sound is played. Two seconds later, two cue images appear for 2 seconds: one in the form of an arrow pointing in one of two directions (left or right), and the other indicating which limb (hand or foot) the subject should use for the MI task. The subject then performs the corresponding MI for 4 seconds. This dataset was recorded using the Emotiv EPOC Flex system, with electrodes arranged as shown in Figure 5 and a sampling rate of 128 Hz. These recording sessions serve both as a training phase to help subjects understand how to perform MI tasks and as a data collection phase for training the MI task classifier and calibrating bias for each specific user.

3.2. Deep Learning Models in Motor Imagery Tasks Classification

Our approach to evaluating a deep learning model begins with testing on a public dataset, specifically the BCI Competition IV-2a dataset, to identify the best-performing models. Subsequently, we conduct a second test on our private VKIST dataset to assess model performance in various real-world scenarios. Each deep learning model was evaluated using two methodologies: subject-dependent and subject-independent. In the subject-dependent approach, we utilized the original training and testing dataset from the BCI Competition IV-2a, which includes 288 trials per session for each subject. The first session was used for training, while the second session served for testing. For the subject-independent approach, we employed the Leave-One-Subject-Out (LOSO) method. In this method, data from one subject is excluded from the training set for testing purposes, while the remaining data from the other eight subjects is combined for training. Performance metrics were averaged across all nine subjects in the dataset. In addition to evaluating the ATCNet model, we also compared it with other deep learning methods, including EEGNet [17], EEG-TCNet [15], MBEEG_SENet [24], TCNet_Fusion [25], DeepConv [26], and ShallowConv [27].

3.3. LDA And CSP in Motor Imagery Classification

To improve compatibility with hardware components such as EEG signal measurement devices and processing computers, and to enhance system performance for real-time continuous processing, we decided to implement the operating mechanism of MI task analysis. Instead of using a deep learning model for feature extraction and EEG signal classification, we will employ digital signal processing (DSP) methods to extract features from the EEG signals. These features will then be input into a machine learning model for classifying the MI tasks. This new approach aims to provide better

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

computational performance for real-time processing, thereby reducing system delays. As noted in the Related Work section, we will use CSP for feature extraction and LDA as the classifier.

Additionally, rather than applying a single set of deep learning model weights across all test subjects, we will utilize the EEG signal recording module to create personalized models through BCI speller training sessions. The data collected from these sessions will allow the system to generate individual machine learning model weights for each user, potentially enhancing the classification performance of MI tasks for each individual.

4. Experiments and Evaluation

4.1. Performance Metrics

4.1.1. Accuracy

Accuracy is the ratio of correct predictions made by a model and is calculated using the following formula:

$$Accuracy = \frac{Correct\ predictions}{Total\ cases} (1)$$

Accuracy tells us the percentage of data correctly classified, without specifying how each class is classified, which class is most accurately classified, or which class's data is most frequently misclassified into other classes.

4.1.2. Cohen's Kappa Coefficient

Given the imbalanced nature of the input data, relying solely on accuracy can lead to misleading results. Cohen's kappa coefficient is a widely used statistical measure that assesses the level of agreement between two raters, making it particularly useful for evaluating performance on imbalanced datasets. Unlike accuracy, Cohen's kappa accounts for the distribution imbalance between classes. The coefficient is calculated using a confusion matrix and the following formula:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} (2)$$

4.2. Motor Imagery Classification Using Deep Learning Models

4.2.1. Signals Preprocessing

Each signal segment undergoes conversion from the time domain to the frequency domain using the Fast Fourier Transform (FFT) method. Then, it is upsampled from 128Hz to 250Hz to match the deep learning model's input layer. A 4th order Butterworth bandpass filter is applied to capture frequencies between 8Hz and 30Hz. Finally, the EEG signals are converted back to the time domain using the Inverse FFT method.

4.2.2. Training and Results

Each deep learning models were trained on a single GPU system, Nvidia A100 40GB, using Adam Optimizer of the Tensorflow library and the following training parameters: batch size = 32, learning rate = 0.001, number epoch = 1000, patience = 300. These parameters were determined through many sets of parameters based on the results of r My initial testing results are shown in Table 4.7 below. In the experimental results of subject-independent approach, the ATCNet model we trained gave better results than the original study. This may be because we used batch size = 32 compared to 64 in the original paper. Reducing the batch size helps minimize overfitting, faster convergence which may contribute to better model performance. Table 2 shows the comparison of some deep learning model performance on the competition dataset.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Table 2:
Comparison of deep learning model performance on the Competition IV-2a dataset.

Model	Subject-dependent		Subject-independent	
	Accuracy	K-Score	Accuracy	K-Score
ATCNet	**81.75**	**0.757**	**71.97**	**0.62**
ATCNet - Original paper	_	_	70.97	0.613
EEGNet	70.33	0.604	59.34	0.458
EEGTCNet	66.09	0.548	59.18	0.456
MBEEEG_SENet	67.4	0.565	57	0.427
TCNet_Fusion	71.53	0.62	61.44	0.486
DeepConv	39.81	0.198	58.18	0.442
ShallowConv	72.49	0.633	60.11	0.468

The VKIST dataset is separated into 2 parts: 80% as training data and 20% as testing data. Table 3 shows the comparison of some deep learning model performance on the VKIST dataset.

Table 3:
Comparison of deep learning model performance on the VKIST dataset.

Model	Without PreProcessing		With PreProcessing	
	Accuracy	K-Score	Accuracy	K-Score
ATCNet	20.26	-0.069	23.91	-0.033
EEGNet	**23.54**	**0.011**	**50**	**0.333**
EEGTCNet	26.15	-0.019	_	_
MBEEEG_SENet	23.54	-0.019	_	_
TCNet_Fusion	23.08	-0.042	_	_
DeepConv	23.85	-0.028	_	_
ShallowConv	23.96	-0.014	_	_

Table 3 represents testing result from this phase. Compared to the results from the BCI Competition IV-2a in Table 2, the classification performance of all mentioned deep learning models is significantly lower. This might be because our EEG recording device lacks the necessary sampling rate, signal bit depth and sensor quality compared to more commonly used devices to capture stable EEG data. Among the deep learning models, only EEGNet achieved a positive Cohen's kappa score when tested with raw EEG data. Therefore, we modified the input layers of both EEGNet and ATCNet to test them with preprocessed VKIST data. The modified EEGNet model significantly outperforms the ATCNet.

4.3. Motor Imagery Classification Using LDA And CSP

We extract EEG feature using DSP methods to derive more meaningful features from these raw EEG signals. Subsequently, the logarithm of the variance for each electrode will be calculated to serve as feature vectors for an MI task classifier. Besides that, according to a study by Huang G et al [27], the time-frequency response of EEG signal within-subject is more consistent than cross-subject and the CSP feature has a significant difference between subjects.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

The CSP method optimizes the variance of one class of signals while minimizing the variance of the other class. The CSP filter produces new signal components (new signal channels) that are linear combinations of the original signal components (original signal channels), resulting in decreasing variance for one class and increasing variance for the other. This characteristic of the CSP filter is illustrated in Figure 6.

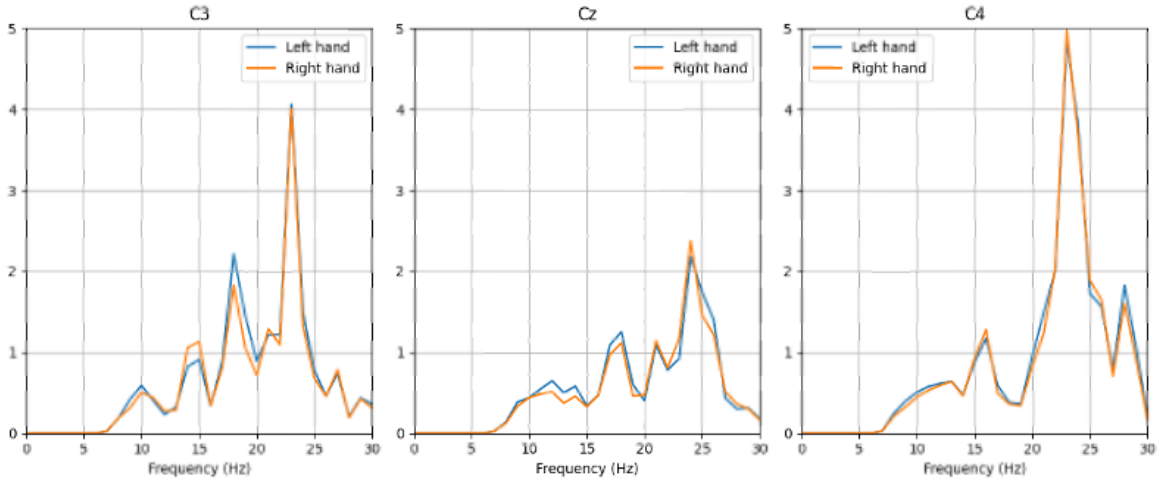

Figure 6:
CSP analysis of the left and right hand classes

We consider the characteristics of the CSP features and choose to implement and train an LDA classifier individually for each test subject. Figure 7 illustrates the classification of data using the LDA algorithm. It demonstrates that the LDA algorithm effectively distinguishes between the signals, separating them into two distinct regions corresponding to the two signal classes.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

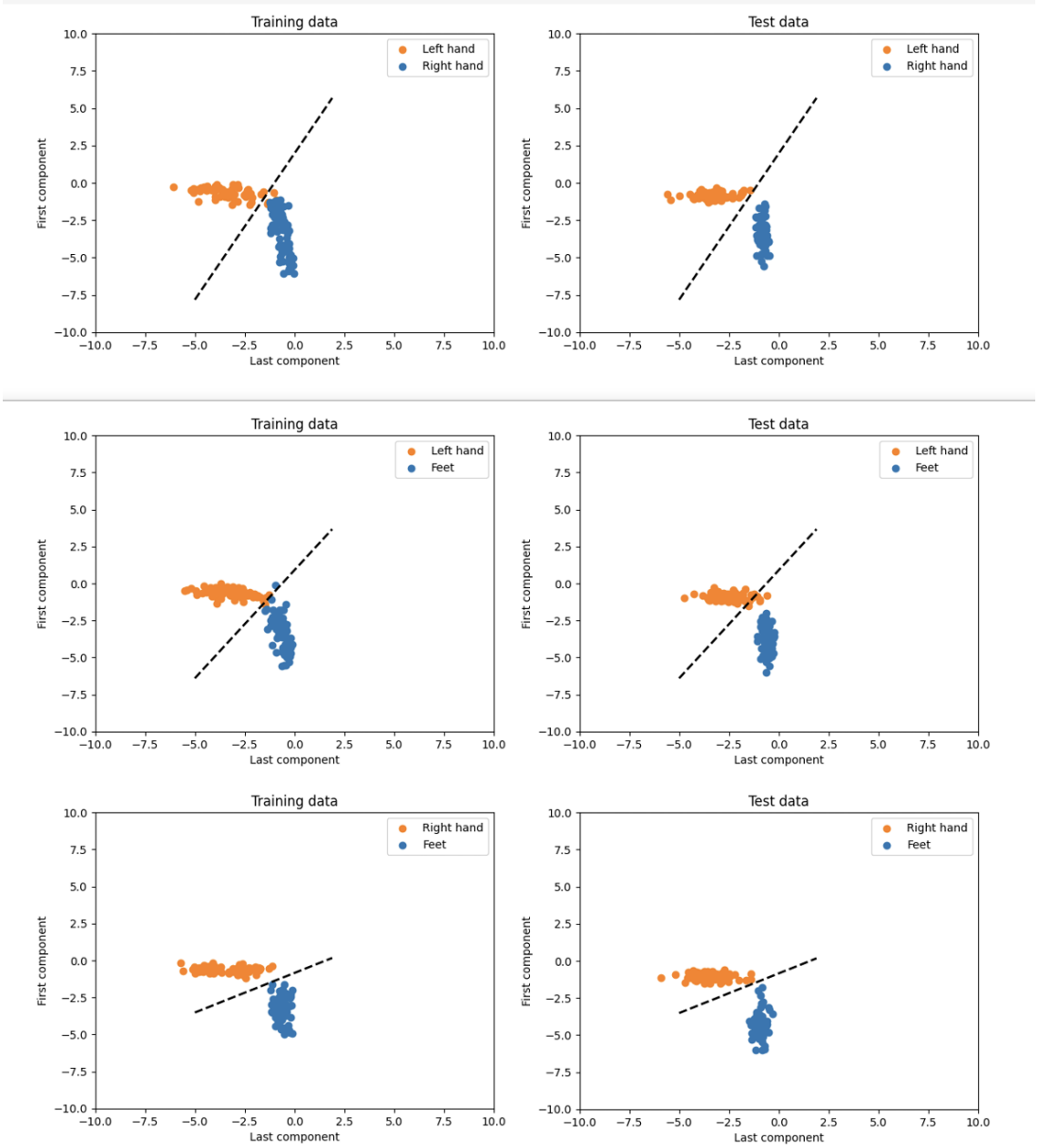

Figure 7:
Motor imagery classification using LDA.

So, we used the BCI Competition IV-2a to test the model performance using LDA with different parameters combinations denoted as Combination A, B, C and D as clarified in Table 4 below. The training results are as shown in Table 5 in term of classification accuracy and K-Score.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Table 4:
Parameter configuration.

Combination	MI tasks	Time window(s)	Band-pass filter(Hz)
A	Left hand - Right hand	2.5 - 3.5	8 - 15
B	Left hand - Right hand	2.5 - 4.5	8 - 15
C	Right hand - Feet	2.5 - 3.5	8 - 15
D	Right hand - Feet	2.5 - 3.5	8 - 30

Table 5:
Comparison of LDA classifier performance (accuracy (%) and K-score) on the BCI Competition IV-2a dataset using CSP features

Sub	Combination A		Combination B		Combination C		Combination D	
	%	K-Score	%	K-Score	%	K-Score	%	K-Score
1	96.45	0.9291	80.85	0.6167	71.94	0.4403	93.53	0.8706
2	38.03	-0.2394	39.44	-0.2113	97.14	0.9429	98.57	0.9714
3	91.97	0.8393	99.27	0.9854	94.93	0.8984	88.41	0.7688
4	59.48	0.194	99.14	0.9828	81.9	0.6355	83.62	0.6742
5	71.85	0.4424	91.85	0.8358	86.13	0.7242	99.27	0.9854
6	78.7	0.5726	100	1	99.08	0.9817	100	1
7	66.43	0.3242	95.71	0.9142	93.57	0.8714	99.29	0.9857
8	80.6	0.6137	99.25	0.9851	92.7	0.8539	82.48	0.6487
9	63.08	0.4374	100	1	76.12	0.5287	94.03	0.8805
Avg	71.83	0.4374	89.5	0.7898	89.04	0.7814	93.24	0.865

As can be seen in Table 4, combination B uses a longer time window than combination A, providing more data, which can lead to higher average classification accuracy in Table 5. However, for certain subjects, the difference in EEG signals between performing left-hand and right-hand MI tasks is not distinct enough, resulting in poor classification performance. Additionally, a 2-second time window resulted in a more delayed BCI system, making a training session longer for the patient. Hence, for subsequent testing with combination C and D, we opted for a 1-second time window but changed the MI tasks and adjusted the band-pass filter to find a more optimal parameters configuration.

In combination C, which involves right-hand and feet MI tasks and applies the same band-pass filter as combination A and B, the average classification result is slightly lower compare to combination B. However, it delivers more consistent performance across all 9 subjects, with no subject having an accuracy below 50%. To provide more input data to the classifier, in combination D, we use a wider band-pass filter ranging from 8 to 30Hz. This configuration yields the highest average accuracy and K-score among all 4 combinations. Consequently, we chose combination D for further testing with the VKIST dataset and for implementation in our system. The classification result in VKIST data set is as shown in the Table 6.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Table 6:
Classification result using VKIST data set.

Labels	Accuracy	K-Score
Right hand - Feet	89.04	0.7814
Left hand - Feet	82.19	0.6455
Right hand - Left hand	100	1
Avg	9.04	8.10

5. Case study of the Vietnamese BCI Speller

5.1.EEG Signals Recording Module

To develop a Vietnamese BCI Speller, it is essential to have a suitable dataset, specifically the VKIST dataset, for training the Motor Imagery classifier. An initial training phase is also necessary to guide users in generating appropriate input signals and to fine-tune the system for real-time processing.

To achieve these goals, our system requires a module capable of creating custom or predefined training scenarios for users while simultaneously recording their EEG signals. This module will also automatically label the recorded EEG signals and store them in a database for further testing and classifier training.

5.2. Vietnamese BCI Speller

5.2.1. Vietnamese Characters

The standard Vietnamese alphabet consists of 29 letters, divided into 17 consonants (b, c, d, đ, g, h, k, l, m, n, p, q, r, s, t, v, x) and 12 vowels (a, ă, â, e, ê, i, o, ô, ơ, u, ư, y). Additionally, there are 11 recognized consonant clusters (ch, gh, gi, kh, ng, ngh, nh, th, tr, qu, ph) and 3 diphthongs formed by vowel clusters (ia-yê-ie, ua-uô, ưa-ươ).

Each vowel can also carry up to 5 diacritics: the acute accent (á), grave accent (à), hook (ả), tilde (ã), and dot below (ạ). Numbers are represented by digits from 0 to 9. Furthermore, there are 12 punctuation marks: period (.), comma (,), question mark (?), exclamation mark (!), colon (:), semicolon (;), parentheses (()), quotation marks (""), dash (-), ellipsis (...), apostrophe (’), and slash (/).

Given the large number of symbols, we decided to categorize them into "parent" and "child" groups based on similar functions in a hierarchical structure. This organization ensures that (i) symbols of similar functions belong to the same group, and (ii) more frequently used symbols are placed at lower levels of the hierarchy to facilitate quicker selection. Additionally, at least four functional keys are necessary: Delete, Enter (for adding a line), Back (to return to previous levels in case of mistakes), and Spacebar.

5.2.2. Vitenamese BCI Speller User Interface.

To better understand our proposal, we define a "key" as a segment on the interface that contains a symbol or a group of symbols. When a key with a single symbol is selected, that symbol becomes the command for that trial. The term "layer" refers to the depth of the hierarchy, while "group" describes symbols or groups of symbols with similar functions associated with a particular key.

We propose a multi-level keyboard GUI, where each interface level contains keys representing symbols, groups of symbols, or functions. An outward-spinning needle, starting from the center and pointing towards the keys, helps users visualize the current selection and the results of their motor imagery tasks.

The proposed UI includes 12 vowels, 17 consonants, 4 functional keys for text drafting, 7 vowel clusters representing three diphthongs, 10 consonant clusters, 10 digits, and 7 punctuation marks, along with diacritics for vowels and vowel clusters. To enhance typing speed, a Vietnamese language model is integrated for character and word suggestions. A suggestion key reveals a list of relevant results upon selection.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Due to the varied appearance of Vietnamese character groups, the number of keys on each layer is flexibly arranged. The keys are organized in a circular pattern around the needle to ensure even distribution of movement time across the keys. With this circular design, the number of keys can range from 7 to 9 per layer. The characters represented include single vowels, single consonants, double vowels, compound consonants, numeric characters, diacritics for single and double vowels, and punctuation marks according to Vietnamese standards.

When a character group is not fully displayed, an expansion key opens a new layer for that group, revealing the remaining symbols. Control keys such as Delete, Enter, Backspace, and Space are represented by familiar symbols for keyboard users.

Figure 8 illustrates the working window of the virtual keyboard, divided into three areas from left to right: Area 1 shows part of the text being drafted, Area 2 displays the interface layers of the virtual keyboard, and Area 3 presents suggested words from the language model based on the drafted Vietnamese content. These suggestions are located under the "Gợi ý" key and are generated using a small language model trained on Vietnamese communication data.

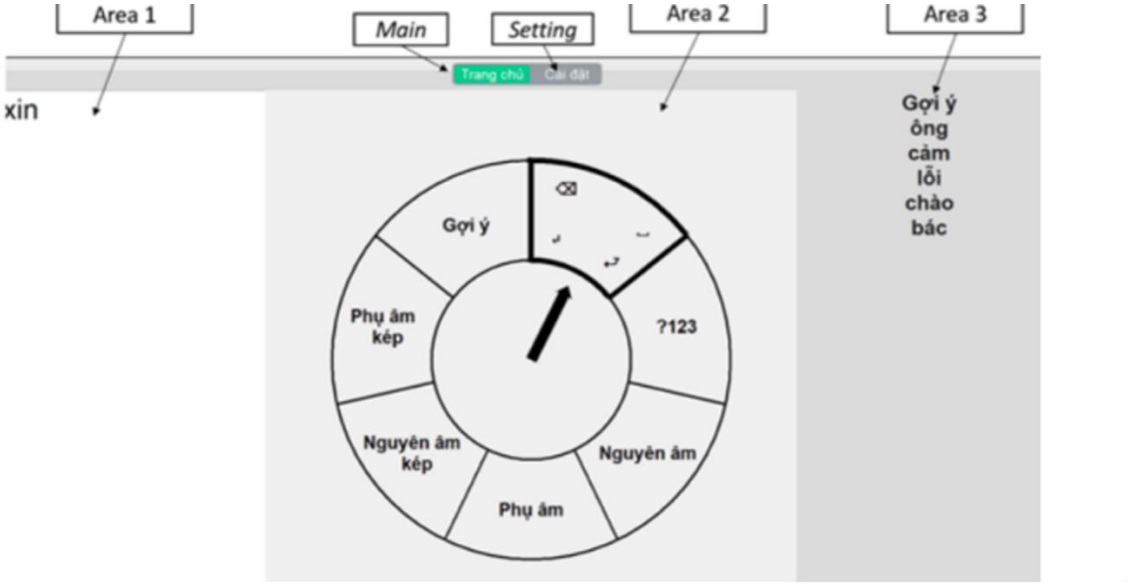

Figure 8:
Vietnamese BCI speller interface.

5.3. Vietnamese BCI Speller Architecture

The BCI-vSpeller system is designed using an asynchronous BCI mechanism that employs EEG signals. Its architecture consists of three main components: the EEG signal acquisition block, the MI task classification block based on EEG signals, and the MI task processing block. This architecture is illustrated in Figure 9 below.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

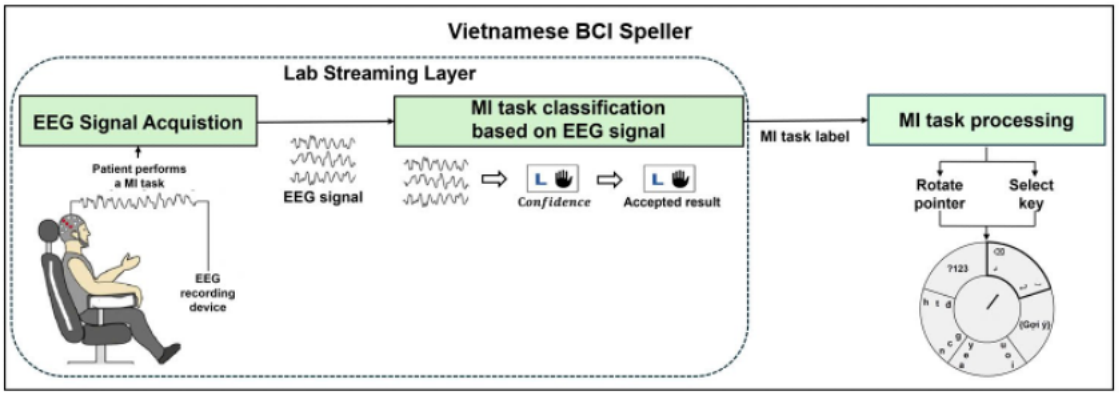

Figure 9.
Vietnamese BCI Speller architecture.

The operating process of the system is detailed as follows:

- ***Acquisition of EEG Signals:*** The patient wears a cap designed to measure EEG signals and opens the Lab Streaming Layer server within the EmotivPro application. The EEG measurement cap automatically connects to the Vietnamese BCI Speller application. The user actively engages in imagined motor imagery tasks corresponding to their desired actions, such as moving the mouse pointer or selecting a character. The device records signals from the electrodes attached to the cap and transmits this data to the Vietnamese BCI Speller system for processing via the Lab Streaming Layer platform.
- ***Classification of MI Tasks Based on EEG Signals:*** This module is developed using the LDA and CSP classification models previously described. The EEG signals sent from the acquisition block are preprocessed and input into the classification model to identify the corresponding imagined motor imagery tasks, which include: 0 (left hand), 1 (right hand), and 2 (foot). Each MI task, associated with an action (Right Hand Movement for rotation and Foot Movement for selection) is predicted by the MI task classification.
- ***Processing of MI Tasks:*** The predicted imagined motor imagery task is compared against a confidence threshold (T). If the confidence level for a task meets or exceeds T, the corresponding keyboard action (moving the mouse pointer or selecting a character) is executed. Conversely, tasks with a confidence level below T are disregarded by the system.

5.4. Results

A performance metric we used is characters per minute (characters/min), which is essentially dividing the length of the text by the time to draft it. Another one is average minimum actions per character, obtained by averaging the minimum numbers of MI tasks need to select and output a character or perform a function. This ends up being 13.52 actions/char, with the least action required being 2, and the most being 25.

From our online experiment on healthy subjects (Figure 10), the subject was instructed to type the phrase "xin chào các bạn", made up of 16 characters including spaces. Optimally, (no wrong classification, no rest, all suggestions after the first word are relevant), the typing rate would be average of 12.31 characters/min. In the first scenario, the target phrase is typed without suggestion. Due to several mistakes made by the classifiers, the subject managed to achieve a rate of only 2.37 characters/min. In the second scenario, suggestions are used whenever possible. This significantly boosted the performance average of the subjects to 6.15 characters/min.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 8873-8891, 2024
DOI: 10.55214/25768484.v8i6.3884

Figure 10:
Using Vietnamese BCI speller.

6. Conclusions

In summary, we evaluated several state-of-the-art deep learning classification models on the BCI Competition IV-2a dataset. The ATCNet model demonstrated the highest performance, achieving excellent accuracy and K-score. However, the classification of MI signals yielded better results on both the BCI Competition IV-2a dataset and our Vkist dataset.

To enhance our Vietnamese BCI speller, we developed a module specifically for classifying MI signals. Initial testing of the Vietnamese BCI speller indicates that the system offers sufficient accuracy and speed. The interface supports drafting Vietnamese text, accommodating all characters and diacriticsan essential feature given the complexity of the language. Additionally, we implemented a suggestion system based on a language model, which significantly improves typing speed by generating relevant suggestions based on the context of communication.

Copyright:

References

[1] S. Selim, M. M. Tantawi, H. A. Shedeed and A. Badr, "A CSP\AM-BA-SVM Approach for Motor Imagery BCI System," in IEEE Access, vol. 6, pp. 49192-49208, 2018, doi: 10.1109/ACCESS.2018.2868178.

[2] N. E. Md Isa, A. Amir, M. Z. Ilyas, M. S. Razalli. Motor imagery classification in Brain computer interface (BCI) based on EEG signal by using machine learning technique. Bulletin of Electrical Engineering and InformaticsVol.8, No.1, March 2019, pp. 269~275ISSN: 2302-9285, DOI: 10.11591/eei.v8i1.1402269Journal homepage: http://beei.org/index.php/EEI

[3] Aggarwal, S., Chugh, N. Review of Machine Learning Techniques for EEG Based Brain Computer Interface. Arch Computat Methods Eng 29, 3001–3020 (2022). https://doi.org/10.1007/s11831-021-09684-6.

[4] Vavoulis, Athanasios & Figueiredo, Patrícia & Vourvopoulos, Athanasios. (2023). A Review of Online Classification Performance in Motor Imagery-Based Brain-Computer Interfaces for Stroke Neurorehabilitation. Signals. 4. 10.3390/signals4010004.

[5] K. Das and R. B. Pachori, "Electroencephalogram-Based Motor Imagery Brain–Computer Interface Using Multivariate Iterative Filtering and Spatial Filtering," in IEEE Transactions on Cognitive and Developmental Systems, vol. 15, no. 3, pp. 1408-1418, Sept. 2023, doi: 10.1109/TCDS.2022.3214081.

[6] Mahbod Nouri, Faraz Moradi, Hafez Ghaemi, Ali Motie Nasrabadi. Towards real-world BCI: CCSPNet, a compact subject-independent motor imagery framework. Digital Signal Processing, Volume 133, 2023, 103816, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2022.103816. (https://www.sciencedirect.com/science/article/pii/S105120042200433X)

[7] Kawala-Sterniuk, Aleksandra & Browarska, Natalia & Al-Bakri, Amir. (2021). Summary of over Fifty Years with Brain-Computer Interfaces—A Review. Brain Sciences. 11. 10.3390/brainsci11010043.

[8] Mulder, Th. "Motor imagery and action observation: cognitive tools for rehabilitation." Journal of neural transmission 114 (2007): 1265-1278. doi:10.1007/s00702007-0763-z

[9] Pfurtscheller, Gert, and Christa Neuper. "Motor imagery activates primary sensorimotor area in humans." Neuroscience letters 239.2-3 (1997): 65-68. doi:10.1016/S0304-3940(97)00889-6

[10] Pfurtscheller, Gert, and Christa Neuper. "Motor imagery and direct braincomputer communication." Proceedings of the IEEE 89.7 (2001): 1123-1134.doi:10.1109/5.939829

[11] Pfurtscheller, Gert, et al. "Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks." NeuroImage 31.1 (2006): 153-159. doi:10.1016/j.neuroimage.2005.12.003

[12] Pfurtscheller, Gert, et al. "Foot and hand area mu rhythms." International Journal of Psychophysiology 26.1-3 (1997): 121-135. doi:10.1016/S0167-8760(97)00760-5

[13] Altaheri, Hamdi, Ghulam Muhammad, and Mansour Alsulaiman. "Physics-informed attention temporal convolutional network for EEG-based motor imagery classification." IEEE transactions on industrial informatics 19.2 (2022): 2249-2258. doi:10.1109/TII.2022.3197419

[14] Brunner, Clemens, et al. "BCI Competition 2008–Graz data set A." Institute for knowledge discovery (laboratory of brain-computer interfaces), Graz University of Technology 16 (2008): 1-6.

[15] Ingolfsson, Thorir Mar, et al. "EEG-TCNet: An accurate temporal convolutional network for embedded motor-imagery brain–machine interfaces." 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2020.doi:10.1109/SMC42975.2020.9283028

[16] Ramoser, Herbert, Johannes Muller-Gerking, and Gert Pfurtscheller. "Optimal spatial filtering of single trial EEG during imagined hand movement." IEEE transactions on rehabilitation engineering 8.4 (2000): 441-446. doi:10.1109/86.895946.

[17] (Online access: November 10th, 2024) Holtel, Frederik. Linear Discriminant Analysis (LDA) Can Be So Easy. Medium. Published in Towards Data Science. Link: https://towardsdatascience.com/linear-discriminant-analysis-lda-can-be-so-easy-b3f46e32f982. 2023.

[18] Farwell, L.A., and E. Donchin. "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials." Electroencephalography and Clinical Neurophysiology, vol. 70, no. 6, Dec. 1988, pp. 510–523, doi:10.1016/00134694(88)90149-6

[19] Volosyak, Ivan, et al. "Evaluation of the bremen SSVEP based BCI in real world conditions." 2009 IEEE International Conference on Rehabilitation Robotics, June 2009, doi:10.1109/icorr.2009.5209543

[20] Blankertz, Benjamin, et al. "A note on brain actuated spelling with the Berlin Brain-Computer interface." Lecture Notes in Computer Science, 2007, pp. 759–768, doi:10.1007/978-3-540-73281-583

[21] Aloise, Fabio, et al. "A covert attention p300-based brain–computer interface: Geospell." Ergonomics, vol. 55, no. 5, 29 Mar. 2012, pp. 538–551, doi:10.1080/00140139.2012.661084

[22] Ramoser, H., et al. "Optimal spatial filtering of single trial EEG during Imagined Hand Movement." IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 4, Dec. 2000, pp. 441–446, doi:10.1109/86.895946.

[23] Rezeika, Aya, et al. "Brain–Computer interface spellers: A Review." Brain Sciences, vol. 8, no. 4, 30 Mar. 2018, p. 57, doi:10.3390/brainsci8040057.

[24] Altuwaijri, Ghadir Ali, et al. "A multi-branch convolutional neural network with squeeze-and-excitation attention blocks for EEG-based motor imagery signals classification." Diagnostics 12.4 (2022): 995. doi:10.3390/diagnostics12040995

[25] Musallam, Yazeed K., et al. "Electroencephalography-based motor imagery classification using temporal convolutional network fusion." Biomedical Signal Processing and Control 69 (2021): 102826. doi:10.1016/j.bspc.2021.102826

[26] Schirrmeister, Robin Tibor, et al. "Deep learning with convolutional neural networks for EEG decoding and visualization." Human brain mapping 38.11 (2017): 5391-5420. doi:10.1002/hbm.23730

[27] Huang, Gan, et al. "Discrepancy between inter-and intra-subject variability in EEG-based motor imagery brain-computer interface: Evidence from multiple perspectives." Frontiers in neuroscience 17 (2023): 1122661. doi:10.3389/fnins.2023.1122661.