ORIGINAL ARTICLE

Closed Loop Subject-Independent SSVEP Frequency Detection System Using CCA Features and Ensemble Learning Methods

M. Moein Esfahani [1], Hosein Najafi [2], Hossein Sadati [1*]

[1] *Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran*

[2] *Department of Computer Engineering, Islamic Azad University, Tehran, Iran*

*Corresponding Author: Hossein Sadati
Email: Sadati@kntu.ac.ir

Received: 11 November 2022 / Accepted: 22 December 2022

Abstract

Purpose: In recent years, the use of Steady-State Visual Evoked Potentials (SSVEPs) in Brain-Computer Interface (BCI) systems has dramatically increased across several fields, such as rehabilitation, cognitive impairment, and brain disease or disorder detection, as well as artificial limbs, wheelchairs, and biomechanical systems. In this study, a novel method is proposed to help scientists develop more efficient BCI systems for Machine Learning Operations (MLOps). This study proposed a state-of-the-art method for detecting SSVEP-based stimulation frequencies with statistical models to design an optimal BCI system.

Materials and Methods: In this study, the Canonical Correlation Analysis (CCA) method has been implemented to extract features from the accessible-to-the-public Tsinghua University Benchmark dataset. A limited number of subjects are being studied. After completing feature selection methods and selecting the best subset of features using a specified feature selection method, the classification of the best features using machine learning-based classification methods has been completed. Furthermore, it is assumed that scientists will design and implement a system specifically for subjects. Models work for subjects independently. However, because model training is subject-specific, we must execute the proposed methods on each subject separately.

Results: The findings indicate that the novel suggested BCI system achieves an average accuracy of 83±9% in stimulation detection, which is higher than that of the traditional CCA approach with an accuracy of 80±11% ($p<0.05$).

Conclusion: Based on the findings, we demonstrated an increase in accuracy with the novel method. It was also discovered that by using the proposed techniques, it is possible to keep MLOps systems as an advantage.

Keywords: Steady-State Visual Evoked Potentials; Canonical Correlation Analysis; Ensemble Learning; Machine Learning Operations; Brain-Computer Interface; Electroencephalogram.

DOI: https://doi.org/10.18502/fbt.v11i3.15878

1. Introduction

The Brain-Computer Interface (BCI) system is a device that provides a direct route between an individual's brain and a physical device [1]. In the most recent decade, Electroencephalogram (EEG) that is acquired from scalp skin has become a popular solution in Brain-Computer Interface (BCI) systems, in addition to other recording systems like fNIRS and fMRI data [2]. This is because EEG is non-invasive and easy to use. In studies on BCI, the most prevalent EEG signals are event-related synchronization and potentials as well as steady-state visual evoked potentials (SSVEP) and Motor Imagery (MI) signals [3, 4, 5]. Each of these signals operates differently depending on the tasks that are being performed. The notion of Steady-State Visual Evoked Potentials (SSVEP) signals is a periodic response to visual stimuli that are modulated with the frequency of the visual stimulus and its harmonics within the frequency band of 4-40Hz. In addition, information regarding this activity can be gleaned from the visual cortex region of the brain [6]. SSVEP signals have several benefits, some of which include a greater Signal-to-Noise Ratio (SNR), an inherent reaction of the brain, and a shorter amount of time to train people to be able to perform their task/activity better [7], [8]. SSVEP signals are expected to ensure that a person will perform the task. However, previous BCI systems need improvements made to their speed, variability, and ease of use. The conventional SSVEP detection methods are unable to recognize flickering at harmonic frequencies since they only function with the frequency's first harmonics; as a result, this study has attempted to overcome this issue with the use of machine learning algorithms [9]. In recent decades, the huge increase in corporation use of data and the breakthroughs made in Artificial Intelligence (AI) have enabled businesses to use Machine Learning (ML) to tackle real-world challenges. ML Operations (MLOps) is a successful technique for transforming ML models from academic resources to corporate problem-solving instruments [10, 11]. According to frequency ranges, SSVEP can be broken down into low frequency (6-12), medium frequency (13-30), and high frequency (>30) [12]. Studies show that the recognition rate is lowest for high-frequency stimuli. However, it is possible to perform classification of the visual fatigue brought on by stimulation, lower the risk of developing seizures, and make individuals feel more at ease while they are recording tasks. According to the findings of a number of studies, when a participant stares at a visual stimulus that flickers at a consistent rate [13], it generates a continuous reaction in the occipital lobe of the brain in response to the frequency of the stimulus as well as the harmonic frequency. The application of the Canonical Correlation Analysis (CCA) approach was introduced for the first time by Lin *et al.* [14]. It is based on the intention of multi-channel SSVEP stimulus identification with high accuracy in detection rates. Prior to that, the technique of Power Spectrum Density Analysis (PSDA) was utilized in a number of studies to identify stimulation frequencies [9]. The first time that Ruimin Wang *et al.* [15] proposed a method to detect SSVEP signals, it was seen as an approach that combined the CCA and PSDA methods. According to the results of studies, this new strategy delivers more accuracy than any of the two procedures used separately. In this study, a new method that employs the CCA method has been developed, which was first presented by Lin *et al.* [14] to extract features. After selecting the optimal feature set, the classification methods to detect SSVEP were applied. According to the final results, this method led to superior outcomes, while using better ways to detect the SSVEP signals. Additionally, utilizing this method is more feasible in contemporary brain-computer interface devices.

This study was conducted based on subject-independent BCI systems. classifiers were trained distinctly from EEG data collected from a group of subjects who were instructed to perform the designed Task [16]. One of the main reasons that the deep learning methods is not used in this study is due to a limited number of samples per subject. it gets poor results in deep learning methods because these methods are highly depending on and are sensitive to the number of data samples to get proportional results other than that it may cause overfitting and poor results in practice (Figure 1).

Other related studies focused on reducing redundant data in EEG recorded signals and then proposed novel methods for feature extraction. In [9], Ma *et al.* proposed a new method that extracts features in a hybrid system and combines CCA and SWT to achieve favorable outcomes in targetless stimuli, but it does not focus on classification optimization in comparison to this method, which has targets during signal recordings. Moreover, in [17], Qianqian *et al.* proposed a novel method that made use of a deep multiset of CCA. It is another approach to the feature

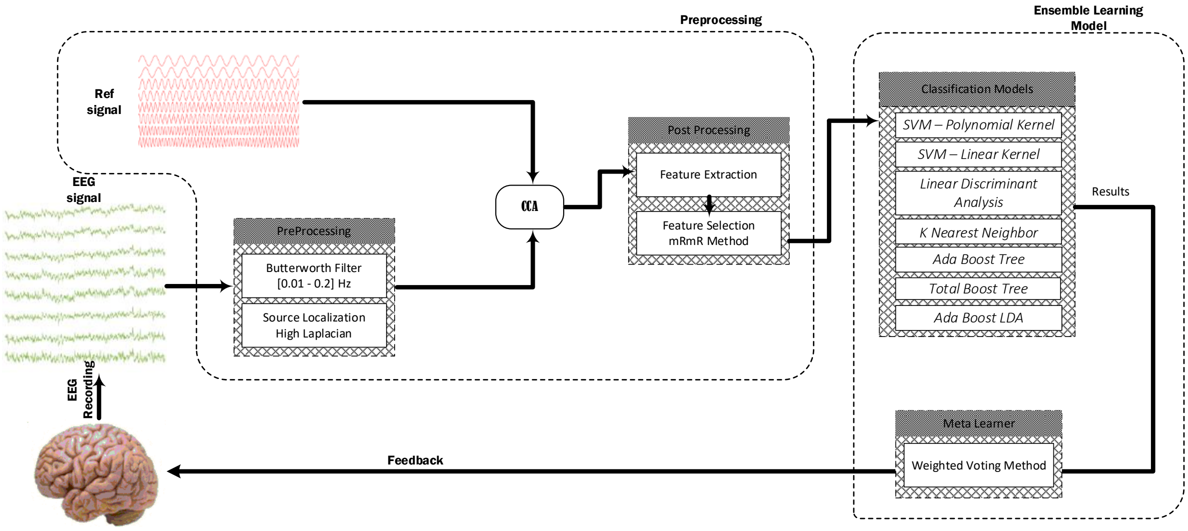

Figure 1. An illustration of the blocks related to the proposed method

extraction section; however, unlike this study, it focuses on feature extraction optimization rather than feature selection and classification, and it only uses a standard classification model. On the contrary, the focus of this study was on classification optimization in order to achieve the best results.

The organization of this paper is structured as follows: in section 2, the benchmark dataset that was employed in this research paper is introduced. Following an introduction to the methodology and flow of the study in Section 3, as well as formulas and a proposed approach to detect SSVEP stimulus, section 4 will be devoted to a discussion of the results that were collected. In the final step, we will wrap up the work and discuss further research in section 5.

1.1. Dataset

This study has made use of a dataset defined as the benchmark dataset, which was originally and publicly released by Tsinghua University [13]. It includes EEG signals on 64 different channels. Eight healthy and experienced subjects are mixed in with the remaining 27 naive individuals to make up a total of 35 subjects. The frequencies of the stimulus range from 8 Hz to 15.8 Hz with an interval of 0.2 Hz. Additionally, all of the data were captured with double accuracy. In this study, based on Figure 2, each trial lasts for six seconds, including a 0.5-sec initial time in the beginning and 0.5 sec as rest time in the last, and the sampling frequency is 250 Hertz; then, a digital filter

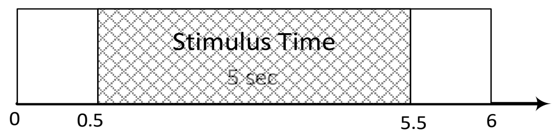

Figure 2. Time schedule of an experiment run

was applied to the dataset by source provider. Finally, we have 1500 different time points structure based on Figure 2. Each array of data is comprised of a four-dimensional matrix with the dimensions 64 x 1500 x 40 x 6, each of which represents the number of electrodes, time points, targets, and blocks, respectively. There are a total of 240 trials for each individual subject because the matrix consists of 6 blocks and 40 targets (Figures 3, 4).

In this study, the EEG system alignment was done using the worldwide 10-20 system. Since it is common knowledge that SSVEP takes place in the occipital region of the subject's brain, in Figure 5 nine channels were chosen for this study. Then, more details about these choices are provided in the channel selection section. The ground was positioned exactly in the center of the gap between Fz and FPz and the reference was selected at the vertex. The impedances of the electrodes were kept lower than 10K. A notch filter set at 50 Hz was applied to the recorded data so that the typical noise from power lines could be eliminated. This section is not considered part of the preprocessing section of this study because it is already added to the original dataset (Figure 6)

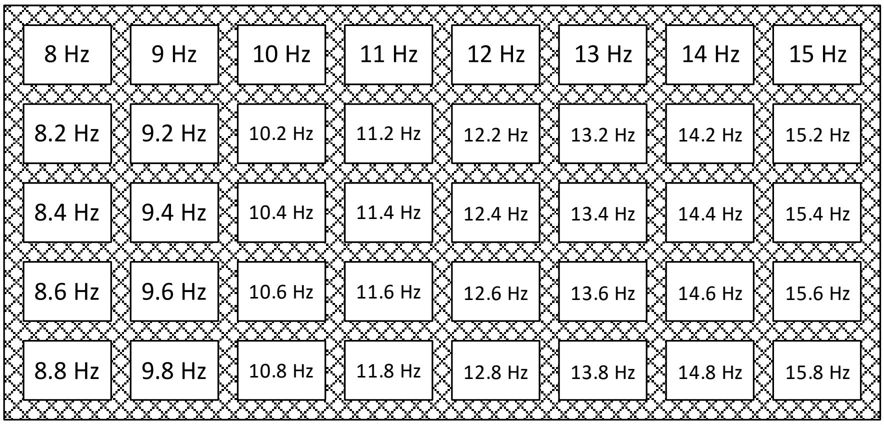

Figure 3. Frequency of stimulation for each trial

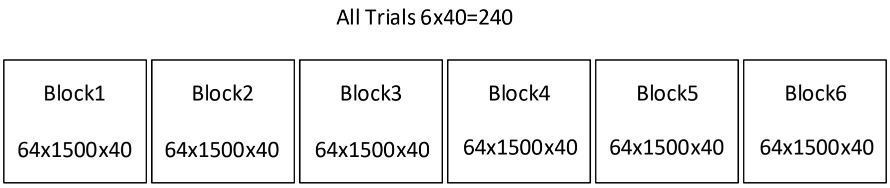

Figure 4. Blocks of time for each subject

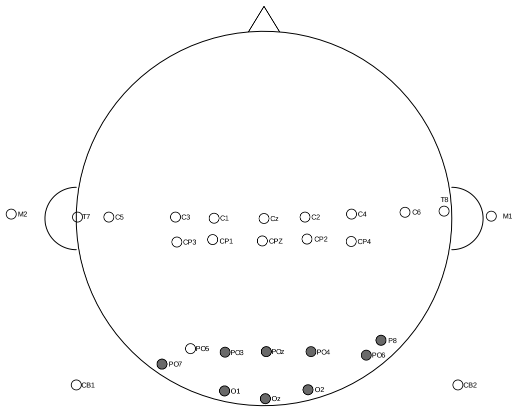

Figure 5. Selected channels in occipital reign of the brain

As shown in the [Figure 3], which is the demonstration of the monitor for subjects, each square block has a width of 200 pixel. The stimulation frequencies in this dataset are quite near, which is one of its major flaws. The fact that stimulation frequencies are not fully unique was a major factor in our decision to use Machine Learning approaches to do classification rather than CCA alone. According to Figure 3, each square represents a specific stimulation frequency, and subjects observe an individual square that flashes at a predetermined frequency; classes of trials are based on these squares.

2. Materials and Methods

In this study, a 40-target dataset with BCI speller was employed. This dataset is freely available to the general public and serves as a benchmark [13]. The frequency ranges between 8 Hz and 16 Hz with a frequency interval of 0.2 Hz in order to prevent the use of harmonic frequencies as stimulus frequencies. In previous techniques for SSVEP recognition systems, the estimation of Power Spectral Density (PSDA) from EEG data is mostly accomplished by the use of the Fast Fourier Transform (FFT). When the Signal-to-Noise Ratio (SNR) of SSVEP signals gets reduced, the recognition accuracy of PSDA quickly drops, whereas the performance of CCA techniques stays consistent throughout this process.

2.1. Proposed State-of-the-Art Technique

Based on the findings of this research, a novel approach was presented to detect and enhance the frequency detection of SSVEPs. In light of this information, we decided to conduct this study using the CCA method, but the approach taken in other studies is different [7, 15, 18]. The CCA arrived at its conclusion by determining the degree of correlation between the EEG signal and the template EEG signal.

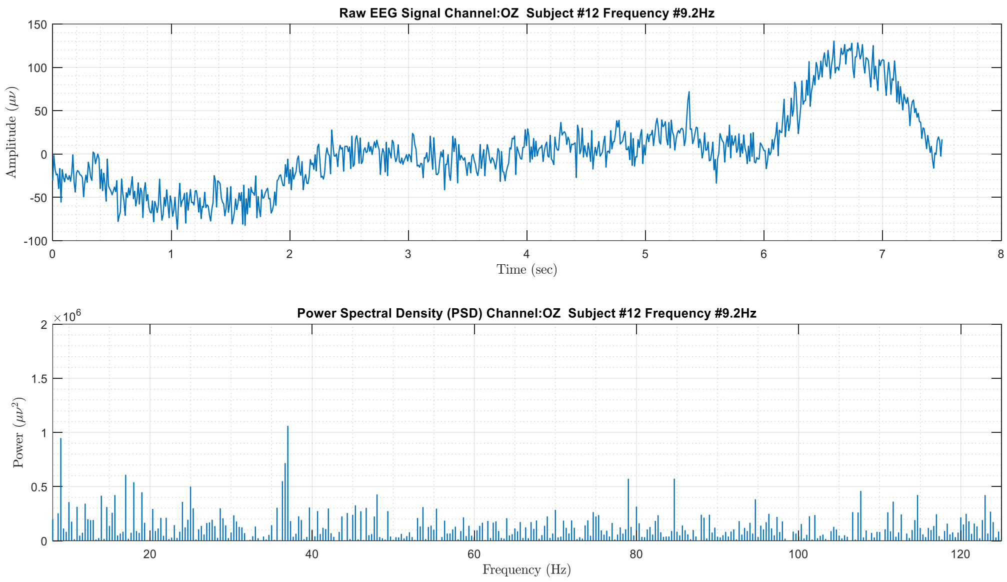

Figure 6. Raw and PSD of EEG signal in specified frequency

On the other hand, with this method that was proposed, we first extracted features using the CCA method, and then, after selecting the best and most optimum features, we utilized the classification algorithm to make the final decision. Moreover, in this study, multiple classification methods were employed and compared, as well as ensemble learning approaches, which describe and function as classification methods with the stacking of multiple low-level classification methods. The dataset used in this study has a limited number of trials, and the main benefit and rationale for employing ensemble learning methods is that these methods perform better with a limited number of samples than deep learning models, which are primarily responsible for worse results and overfitting with a limited number of samples.

2.2. Pre-Processing

The primary objective of the preprocessing stage is to reduce the amount of existing noise and redundant data in our EEG signals. The stimulation frequency domain for SSVEP signals varies from 8 Hz to 18 Hz. One of the best ways to make better and more accurate decisions in SSVEP detection is to clean up our data using digital filtering and source localization techniques.

2.2.1. Digital Butter Worth Filter

In this study, a Butterworth high pass filter with a cutoff frequency of 5 Hz and a Butterworth low pass filter with a cutoff frequency of 20 Hz were employed. Important to note is that stop pass Butterworth could be used for EEG signals. However, with the two distinct filters that we employed, the primary benefit is that principal data that occur in the middle frequency domain are not affected by digital filtering, as opposed to stop-band filtering, which influences the middle data.

2.2.2. Source Localization Methods

Based on the surface EEG recording, there are a variety of ways for source localization with different procedures. In addition, one of the primary benefits of source localization methods is the elimination of redundant data and improve signal-to-noise ratio in EEG signals [19-21], which improves classification accuracy.

After applying a digital filter to EEG signals to eliminate unneeded frequencies, in the second step, a High Laplacian Spatial Filter is implemented. Other alternatives for this are low laplacain and common average reference. The selected method has the highest performance towards source localization techniques. Equations 1 and 2 shows a high Laplacian spatial filter:

$$x_i^{LAP}(t) = x_i(t) - \sum_{j \in S_i} \omega_{ij} x_j(t) \tag{1}$$

$$\omega_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j \in S_i} \frac{1}{d_{ij}}} \tag{2}$$

According to the findings of earlier studies, $x_i^{LAP}(t)$ is the filtered signal of electrode i-th. Additionally, it is essential to point out that we utilized Euclidean distance in this research.

2.3. Channel Selection

After completing the preprocessing section, which should be performed on all of the channels, we will then need to select a subset of channels to use for our dataset, which consists of 64 electrodes or channels. Some of these channels have a greater amount of information than others, while some of them may not have sufficient information at all. According to previous research, information on SSVEPs can be found in channels that are located in the occipital brain region. For this study, nine channels that are mostly located in the occipital region were selected based on Figure 6. However, for the purposes of this study, we employed the same channels for each of the individuals even if it is probable that various channels for each subject contain information that is valuable in comparison to other channels. Eliminating unneeded channels using channel selection techniques that were previously researched in fNIRS signals and EEG signals could be the subject of a study that we will conduct in the future [1, 4] (Table 1).

Table 1. Models and Parameters of Classification

Classification Model	Description
SVM-Linear	Muti class SVM with One VS One
SVM-Polynomial	Muti class SVM with One VS One
KNN	K Nearest Neighbor with K=32 as Optimum Number
LDA	Linear Discreminant Analysis
Adaboost Tree	Number of Trees set as 153
Total Boost Tree	Number of Trees set as 65
Adaboost LDA	Number of Ensembles Set as 41

2.4. Canonical Correlation Analysis (CCA)

CCA has shown a great deal of use in recent years for monitoring the frequency of SSVEPs. CCA is a statistical method that is used to measure the correlation between two variables that have multiple dimensions [7, 9, 14]. The weight vectors found will maximize the correlation between two multivariate arrays to be used in this study. This weight vector is based on an equation involving two multi-dimensional variables and the linear combinations of those variables. The idea of employing the CCA method in SSVEPs for the first time was described by Lin *et al.* [14]. While that method has its own benefits, it has some drawbacks as well, which led to additional studies being carried out to tackle this issue. In order to address the primary limitations of the CCA approaches, several studies such as the following have been carried out: Chen *et al.* [22] presented the filter bank CCA, and their results are based on the findings of experiments described in [22]. Additionally, it demonstrates that Filter bank performs better on results, and additionally, they utilized harmonic frequency components in order to boost performance in SSVEP-based BCIs. In [23], the BCI system is based on SSVEP to be used in the control of robotic arms. In the research that has been done, sinusoidal signals have been employed as reference signals in order to perform frequency detection of the SSVEP in an unsupervised manner. The frequency of stimulation and the number of harmonics make up this component. The CCA algorithm calculates the canonical correlation of multichannel SSVEPs corresponding to each stimulation frequency. Additionally, the frequency of referenced signals with the maximal correlation is considered to be the frequency of SSVEP. Recognizing the frequency of SSVEP requires calculating the canonical correlation of multi-channel SSVEPs.

Canonical Correlation Analysis (CCA) is described as below (Equations 3 , 4):

$$Y_{fi} = \begin{bmatrix} \sin(2\pi f_1 t) \\ \cos(2\pi f_1 t) \\ \dots \\ \sin(2\pi N_h t) \\ \cos(2\pi N_h t) \end{bmatrix}, \qquad t = \left[\frac{1}{f_s}, \frac{2}{f_s}, \dots, \frac{N_s}{f_s}\right] \tag{3}$$

$$\max \rho(x,y) = \frac{E[x^T y]}{\sqrt{E[x^T x]E[y^T y]}} = \frac{E[W_x{}^T X Y^T W_y]}{\sqrt{E[W_x{}^T X X^T W_x]E[W_y{}^T Y Y^T W_y]}} \quad (4)$$

Where Y_{fi} represents as reference signals based on Figure 1 and t represents time domain for Y_{fi}array. Moreover, the fundamental premise of CCA is to choose two linear transforms $x = W_x{}^T X$ and $y = W_y{}^T Y$ that maximize the correlation coefficient between one indicator and the other.

The maximum value of $\rho(x,y)$ is equivalent to the canonical correlation between variables x and y being at its highest possible level. Although the CCA is effective for SSVEPs, sine and cosine waves might not be the best choice for use as reference signals when performing correlation analysis. This is due to the fact that sine and cosine waves do not contain any information regarding inter-subject fluctuation, and noise or other information may exist in the signal regardless of the preprocessing section that was applied in the section before it. One technique to do this is to locate more productive reference signals, which is covered in Equation 3. An additional strategy that was utilized in this study was an approach based on machine learning. This indicates that we simply utilize CCA to extract features and do not use maximum values to make a final determination regarding which stimulus this particular factor belongs to. After doing feature selection, which is followed by performing the classification procedure, and our final decision on the classifiers' output, we make use of CCA vectors as the features.

2.5. Feature selection

In order to obtain the optimal subset of features, a feature selection algorithm should be implemented to aid in better data interpretation and to avoid the curse of dimensionality. This is because the CCA method yields extracted features set that is a high-dimensional matrix that affects the classification results. There are two primary ways for selecting features: the Wrapper and Filter methods [3, 4, 17]. The Wrapper technique is a subset of feature combinations, and its analysis is computationally expensive, whereas the Filter approach, which analyzes each feature independently, has the advantage of being computationally efficient and quick. This study employs minimum redundancy maximum relevance (mRmR) to identify whether or not a feature belongs to a feature set [24, 25]. In other words, some features should be removed, while the rest should be accepted for the following section. In the mRmR method, a rank between 0 and 1 is assigned to each feature, with 0 representing irrelevant or rejected features and 1 representing relevant or accepted ones. We implemented the mRmR algorithm described as:

$$I(x,y) = \iint p(x,y) log \frac{p(x,y)}{p(x)p(y)} dxdy \quad (5)$$

I(x, y) is the mutual information that was obtained through probabilistic density, and it can be found in Equation 1. You can learn more about this method by studying [1, 4].

The primary reason why we decided to employ the mRmR method rather than another filter approach such as the ANOVA or chi-squared method was because the results of our comparisons led us to select this particular way. In contrast, the mRmR channels selection approach achieves significantly faster and more accurate results than methods that focus just on statistical feature selection (Tables 2, 3).

Table 2. Comparison of the Methods' Average Classification Accuracy

	CCA	**PSDA**	**CCA**	**CCA-ML**
Subject 01	79.58 %	27.24 %	75.35 %	82.34 %
Subject 02	97.50 %	39.52 %	83.68 %	93.13 %
Subject 08	83.75 %	34.28 %	84.25 %	82.12 %
Subject 09	83.58 %	32.27 %	75.42 %	90.28 %
Subject 11	77.08 %	28.51 %	75.46 %	76.32 %
Subject 16	56.25 %	23.19 %	42.28 %	64.35 %
Subject 18	76.66 %	28.34 %	77.38 %	78.22 %
Subject 19	84.56 %	28.32 %	83.75 %	86.29 %
Subject 21	74.41 %	31.76 %	73.28 %	82.52 %
Subject 33	92.91 %	34.43 %	94.42 %	95.79 %
Mean	80.62	30.79	76.53	83.14
Std	11.21	4.62	13.63	9.11

Table 3. Subjects Classification accuracy based on Models

	SVM-Linear	SVM-Polynomial	KNN	LDA	Adaboost Tree	Total Boost Tree	Adaboost LDA
Subject 01	56.72 %	65.42 %	66.53 %	77.42 %	65.42 %	72.68 %	**82.34** %
Subject 02	52.37 %	56.68 %	62.41 %	89.68 %	54.34 %	62.38 %	**93.13** %
Subject 08	42.38 %	43.57 %	36.54 %	73.57 %	42.23 %	51.34 %	**82.12** %
Subject 09	58.95 %	67.43 %	61.53 %	82.43 %	67.43 %	59.37 %	**90.28** %
Subject 11	54.25 %	57.43 %	52.21 %	68.43 %	57.43 %	63.37 %	**76.32** %
Subject 16	60.42 %	63.36 %	32.73 %	**65.42** %	63.36 %	59.52 %	64.35 %
Subject 18	56.58 %	54.76 %	42.62 %	68.76 %	49.83 %	51.38 %	**78.22** %
Subject 19	75.36 %	73.42 %	65.45 %	**86.42** %	74.32 %	81.32 %	86.29 %
Subject 21	62.46 %	68.35 %	56.85 %	71.35 %	56.67 %	66.34 %	**82.52** %
Subject 33	74.36 %	75.32 %	82.47 %	91.32 %	64.28 %	59.14 %	**95.79** %
Mean	59.39	62.57	55.93	77.48	59.53	62.68	83.14
Std	9.83	9.60	15.23	9.43	9.35	9.14	9.11

2.6. Classification

The final stage of this study is the classification of data based on Figure 1. After feature extraction and selecting the optimal subset of features, classifiers are trained. In this study, we employed and compared a variety of classification techniques. We could also use deep learning techniques, such as the 1-D CNN method, but the main drawback of such models is that, due to the limited number of trials, we obtain poor results and our model overfits most of the time. We can use ensemble learning and other classification methods to avoid overfitting with a limited number of samples. In this study, seven classification models were implemented. Using Table 1 as a guide, we described each classification model we employed.

2.6.1. Support Vector Machine (SVM)

Support Vector Machines (SVM) are supervised models that use classification algorithms for two-class classification problems (Figure 7). We implemented the One Versus One (OVO) method to classify multiclass data, as this study has more than two classes and thus this method cannot be used (Figure 8). In addition, we used polynomial and linear kernels for SVM classification models, and the results are based on these kernels.

2.6.2. K Nearest Neighbor (KNN)

The K-Nearest Neighbors (KNN) algorithm is a supervised machine learning model that can be used to solve both classification and regression problems. It

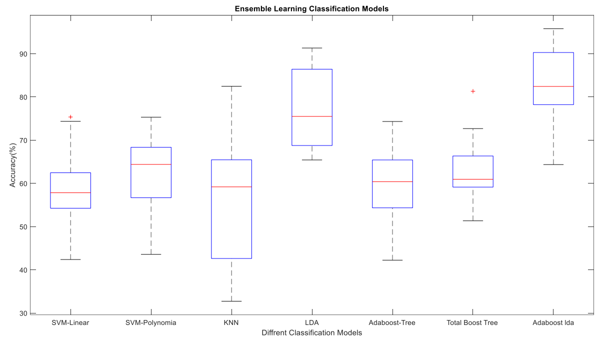

Figure 7. Performance evaluation using classification models

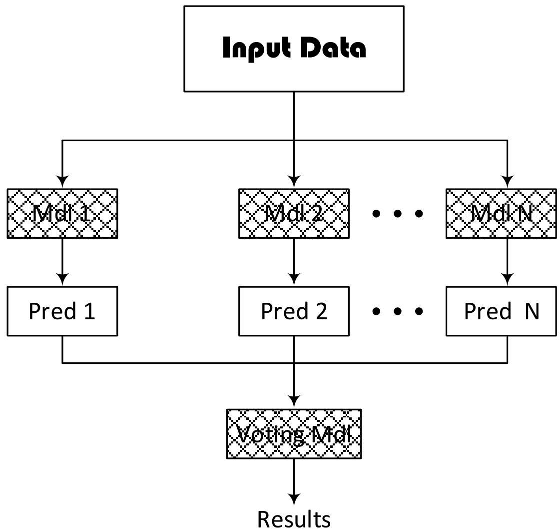

Figure 8. The Proposed Ensemble Method Diagram

becomes dramatically slow when we have large datasets with many dimensions. Consequently, it is not an appropriate machine learning model to stack and operate as an ensemble learning model.

2.6.3. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is an algorithm for multi-class classification based on predictive modeling. It can also be used as a dimensionality reduction technique by providing a projection of a training dataset that most effectively separates examples by their assigned class.

2.6.4. Ensemble Learning Models

Ensemble learning is the process of combining and stacking multiple classification models to solve a specific machine learning problem [26]. The proposed Ensemble Learning Diagram is depicted in Fig.8. The first layer is made up of seven weak learner models, and the second layer uses the first layer's predictions to make the final decision using the voting method. The voting method makes the final decision based on which label has the most weight, and the number of weak learner methods is odd. Due to the limited number of trials in this study, ensemble learning models can be used to improve performance. We employed Adaboost with a decision tree and linear discriminant analysis kernel as well as total boost with a decision tree kernel. In the second layer, also called meta-classification, a voting model has been employed.

3. Results and Discussion

In this section, we will analyze the results obtained in the present research. In the proposed dataset, there are 30 participants, but for this study, only data for 10 participants were used. Because some of the subjects had been trained while others had not, we chose a subset of subjects that mainly have poor accuracy in final results in comparison to trained subjects, who gain high accuracy all the time, and this caused bias in the results. Based on the results of the research done in [7] the summary and conclusion of all the results are represented in Table 2. The ten rows in Table 2 correspond to the ten selected subjects. The channel selection method is based on the results of the conducted research in [7]. Since some subjects always give high accuracy, they are not appropriate for comparative purposes. Therefore, we had to eliminate these subjects from consideration. Column 1 contains the findings of the experiments that were carried out and reported in published research [7]. The classification is then shown in column 2, which employs Power Spectral Density Analysis (PSDA) features. According to the findings, we can see that the PSDA results degrade. The primary reason for this is that the stimulation frequencies on this dataset are only separated by 0.2 Hz, which means that this method cannot function well enough to provide satisfying results. In addition, Canonical Correlation Analysis (CCA), which we performed in column No. 3, and the results that we obtained were compared to those in Column 1. Moreover, column 4 contains the results of the CCA features that were classified using the Machine Learning algorithms that are described in Table 3.

Moreover, Table 3 describes the parameters that are mainly used in the classification results, noting that the SVM classification methods were used in the OVO method for the classification task. As shown in columns 1 and 2 in Table 2, using the SVM method led to poor results in comparison to other results, the main reason being that the primary purpose for the development of the SVM method was not for multi-class classification purposes, but for binary classification. In addition, we can see two classification methods in Table 2 as KNN and LDA, which were primarily developed for multi-class classification, leading to better results compared to other methods than the three ensemble learning classification methods that were used. We optimized the number of stacks in ensembles by using experimentation and reported the results in Table 1. Two of the ensemble learning classification methods have a decision tree kernel, and one of them has an LDA kernel.

In order to check the accuracy of the results, we relied on cross-validation methods. After giving each subject 250 separate trials, we now have 250 samples. For the first four classification models, we only needed one model to classify data; then, the 10-fold cross-validation was used to validate the accuracy of our research results. In this section, our samples were randomly divided into 10 separate sections, each of which has 25 trials, and one of these 10 sections was chosen to evaluate or test the models, while the other 9 sections were used to train the models. Then this procedure was repeated 10 times with different selections ($p<0.05$). In addition, we have implemented the hold-out method for the last three models. We implemented ensemble learning methods, which are methods that combine a number of different methods, and the decision that was obtained was based on the combination of all models' results. Therefore, the hold-out method might be more effective in this study to reduce the amount of time that has managed to pass.

Furthermore, as Figure 7 demonstrates, we can determine which classification approach is more effective. Based on the numerical data in Figure 5 and Table 3, the first option is Adaboost with an LDA kernel model, followed by and LDA model, which performs better than other approaches. However, it is important to mention that the adaboost LDA model increased the results, but it requires more memory than the LDA model.

Lastly, the results of the conducted research are presented in the table. Classification models that use adaboost LDA achieve better results, and the LDA model itself also achieves satisfactory results. The primary benefit of this method is that it is implementable in MLOps systems. Furthermore, this method can achieve better performance if the stimuli are of varying frequencies, in contrast to the selected dataset, which is comprised of the selected dataset whose frequencies are relatively close.

4. Conclusion

The recognition of SSVEP-based Brain-Computer Interfaces (BCIs) is a cutting-edge topic that is primarily supported by research carried out in the most recent decades. The majority of the frequency detection in the other studies was done with PSDA and CCA. However, these two primary approaches make unsupervised decisions and have merits over machine learning models. One of these merits is fast decision-making when there is a limited number of channels available. On the other hand, in order to attain improved performance, machine learning models, and then the deep learning models, require a substantial number of trials. However, as previously noted, CCA performs better and is more stable than PSDA. In this study, SSVEP stimulation frequencies were classified using machine learning methods with CCA features. In addition to the benefits that have already been described, such as its application in MLOps systems, this method has a detection accuracy that is comparable to or even superior to that of CCA methods. It is likely that greater results can be achieved using alternative optimization techniques, such as frequency optimization, in combination with machine learning model optimization strategies. It is recommended that future studies continue the development of channel selection methods in order to choose the most appropriate channels for each individual subject, and then the proposed method be put into practice. This will bring about a number of benefits, including a reduction in the total number of channels and, in some instances, an improvement in performance. In addition, employing other strategies or methods that have been optimized to extract features, can be quite useful. Some examples of such methods include the filterBank CCA, or a combination of PSDA and CCA.

References

1- M. Moein Esfahani and H. Sadati, “Cross-Subject fNIRS Signals Channel-Selection based on Multi-Objective NSGA-II Algorithm,” in *2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME)*, Nov. (2021), pp. 242–247, doi: 10.1109/ICBME54433.2021.9750364.

2- M. M. Esfahani and H. Sadati, “fNIRS Signals Classification with Ensemble Learning and Adaptive Neuro-Fuzzy Inference System,” in *2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, Dec. (2021), pp. 1–5, doi: 10.1109/ICSPIS54653.2021.9729388.

3- A. T. Dehkordi, H. Ghasemi, and M. J. V. Zoej, “Machine Learning-Based Estimation of Suspended Sediment Concentration along Missouri River using Remote Sensing Imageries in Google Earth Engine,” in *2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, Dec. (2021), pp. 1–5, doi: 10.1109/ICSPIS54653.2021.9729382.

4- M. Moein Esfahani and H. Sadati, “Application of NSGA-II in Channel Selection of Motor Imagery EEG Signals with Common Spatio-Spectral Patterns in BCI Systems,” in *2022 8th International Conference on Control, Instrumentation and Automation (ICCIA)*, Mar. (2022), pp. 1–6, doi: 10.1109/ICCIA54998.2022.9737199.

5- A. Yousefiankalareh, M. A. Manoochehri, S. A. Khoshnevis, and S. R. Surakanti, “Improve the Frequency Identification in SSVEP based BCI Systems with Moving Windows Algorithm,” *Proc. - 2019 6th Int. Conf. Control. Instrum. Autom. ICCIA 2019*, pp. 1–5, 2019, doi: 10.1109/ICCIA49288.2019.9030862.

6- S. Aggarwal and N. Chugh, “Signal processing techniques for motor imagery brain computer interface: A review,” *Array*, vol. 1–2, no. August, p. 100003, (2019), doi: 10.1016/j.array.2019.100003.

7- M. Norizadeh Cherloo, H. Kashefi Amiri, and M. R. Daliri, “Spatio-Spectral CCA (SS-CCA): A novel approach for frequency recognition in SSVEP-based BCI,” *J. Neurosci. Methods*, vol. 371, no. January, p. 109499, (2022), doi: 10.1016/j.jneumeth.2022.109499.

8- W. Yan, Y. Wu, C. Du, and G. Xu, “Cross-subject spatial filter transfer method for SSVEP-EEG feature recognition,” *J. Neural Eng.*, vol. 19, no. 3, (2022), doi: 10.1088/1741-2552/ac6b57.

9- P. Ma *et al.*, "A classification algorithm of an SSVEP brain-Computer interface based on CCA fusion wavelet coefficients," *J. Neurosci. Methods*, vol. 371, p. 109502, (2022), doi: 10.1016/j.jneumeth.2022.109502.

10- M. Testi *et al.*, "MLOps: A Taxonomy and a Methodology," *IEEE Access*, vol. 10, pp. 63606–63618, (2022), doi: 10.1109/access.2022.3181730.

11- M. M. Esfahani and H. Sadati, "Application of Transfer Learning in Optimized Filter-Bank Regularized CSP to Classification of EEG Signals with Small Dataset," in *2022 30th International Conference on Electrical Engineering (ICEE)*, May (2022), pp. 963–967, doi: 10.1109/ICEE55646.2022.9827057.

12- W. Liu, L. Zhang, and C. Li, "A method for recognizing high-frequency steady-state visual evoked potential based on empirical modal decomposition and canonical correlation analysis," *Proc. 2019 IEEE 3rd Inf. Technol. Networking, Electron. Autom. Control Conf. ITNEC 2019*, no. Itnec, pp. 774–778, (2019), doi: 10.1109/ITNEC.2019.8729005.

13- Y. Wang, X. Chen, X. Gao, and S. Gao, "A Benchmark Dataset for SSVEP-Based Brain-Computer Interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, (2017), doi: 10.1109/TNSRE.2016.2627556.

14- Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-Based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, (2007), doi: 10.1109/TBME.2006.889197.

15- R. Wang, W. Wu, K. Iramina, and S. Ge, "The combination of CCA and PSDA detection methods in a SSVEP-BCI system," *Proc. World Congr. Intell. Control Autom.*, vol. 2015-March, no. March, pp. 2424–2427, (2015), doi: 10.1109/WCICA.2014.7053101.

16- A. M. Ray *et al.*, "A subject-independent pattern-based brain-computer interface," *Front. Behav. Neurosci.*, vol. 9, no. OCTOBER, pp. 1–15, (2015), doi: 10.3389/fnbeh.2015.00269.

17- Q. Liu *et al.*, "Efficient representations of EEG signals for SSVEP frequency recognition based on deep multiset CCA," *Neurocomputing*, vol. 378, pp. 36–44, (2020), doi: 10.1016/j.neucom.2019.10.049.

18- P. Ma *et al.*, "A classification algorithm of an SSVEP brain-Computer interface based on CCA fusion wavelet coefficients," *J. Neurosci. Methods*, vol. 371, p. 109502, (2022), doi: 10.1016/j.jneumeth.2022.109502.

19- D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 3, pp. 386–394, (1997), doi: 10.1016/S0013-4694(97)00022-2.

20- D. J. McFarland and J. R. Wolpaw, "Brain-computer interfaces for communication and control," *Commun. ACM*, vol. 54, no. 5, pp. 60–66, (2011), doi: 10.1145/1941487.1941506.

21- J. Lu, D. J. McFarland, and J. R. Wolpaw, "Adaptive Laplacian filtering for sensorimotor rhythm-based brain-computer interfaces," *J. Neural Eng.*, vol. 10, no. 1, (2013), doi: 10.1088/1741-2560/10/1/016002.

22- X. Chen, Y. Wang, S. Gao, T. P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface," *J. Neural Eng.*, vol. 12, no. 4, (2015), doi: 10.1088/1741-2560/12/4/046008.

23- X. Chen, B. Zhao, Y. Wang, S. Xu, and X. Gao, "Control of a 7-DOF Robotic Arm System with an SSVEP-Based BCI," *Int. J. Neural Syst.*, vol. 28, no. 8, (2018), doi: 10.1142/S0129065718500181.

24- A. Taheri Dehkordi, M. J. Valadan Zoej, H. Ghasemi, E. Ghaderpour, and Q. K. Hassan, "A New Clustering Method to Generate Training Samples for Supervised Monitoring of Long-Term Water Surface Dynamics Using Landsat Data through Google Earth Engine," *Sustainability*, vol. 14, no. 13, p. 8046, Jun. (2022), doi: 10.3390/su14138046.

25- A. C. Ramos, R. G. Hernández, and M. Vellasco, "Feature Selection methods applied to Motor Imagery task classification," *2016 IEEE Lat. Am. Conf. Comput. Intell. LA-CCI 2016 - Proc.*, no. Mi, (2017), doi: 10.1109/LA-CCI.2016.7885731.

26- O. Sagi and L. Rokach, "Ensemble learning: A survey," no. December 2017, pp. 1–18, (2018), doi: 10.1002/widm.1249.