SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I9P106

Volume 11 Issue 9, 56-63, September 2024

Original Article

Channel Selection Using Stochastic Diffusion Search Algorithm for Classification in Brain-Computer Interface

Sanoj Chakkithara Subramanian[1*], D. Daniel[1]

[1]*Department of Computer Science and Engineering, School of Engineering and Technology, CHRIST(Deemed to be University), Bengaluru, Karnataka, India.*

[*]*Corresponding Author : sanoj.cs@res.christuniversity.in*

Received: 27 June 2024 Revised: 07 August 2024 Accepted: 30 August 2024 Published: 30 September 2024

Abstract - *Utilization of the Brain-Computer Interfaces (BCI) is done via Electroencephalogram (EEG) signals that provide several environmental interactions among individuals having restricted movements owing to neurodegenerative diseases or strokes. However, the BCI system was based on Motor Imagery (MI). It was not used for any form of real-life application owing to a decrease in the performance of various Common Spatial Pattern (CSP) algorithms, especially while the actual number of channels was high. A multi-channel structure of such EEG signals can increase cost and bring down speed. Due to this, a reduction in the system cost by the detection of active electrodes during the process can increase accessibility. This way, optimization techniques in choosing electrodes can be used to determine other effective channels by employing a method of random selection. For this work, a Stochastic Diffusion Search (SDS) algorithm based on herd optimization techniques was used with four different classifiers, which were the AdaBoost, the Classification and Regression Tree (CART), the Naive Bayes (NB) as well as the K-Nearest Neighbor (KNN). The channels that were chosen frequently were determined to improve the system performance with regard to accuracy and speed. The results proved that the approach proposed was successful in bringing down the channel number and run time without affecting the accuracy of classification.*

Keywords - *Channel selection, Common Spatial Pattern (CSP), Stochastic Diffusion Search (SDS), Naive Bayes (NB), Classification and Regression Tree (CART), K-Nearest Neighbor (KNN), Adaboost.*

1. Introduction

BCI establishes a communication pathway, enabling direct interaction between the brain and various external devices. This is quite distinct from the conventional communication channels, which are dependent on neuromuscular pathways in the peripheral muscles and nerves. Electroencephalogram (EEG)-based BCIs use neural activities like visual evoked potentials, auditory evoked potentials, the ERD/ERS, visual attention, and the event-related potentials [1].

Motor Imagery (MI)-based BCI relies on imagining movement in lieu of physically executing it to send control commands to external devices connected to the BCI. Unlike in the case of the BCI, the neurofeedback will use visual and, in certain cases, auditory feedback that enables participants to ensure self-regulation, especially for the neural substrate systems. Further, there was an improvement in motorways or pathways among post-stroke subjects. MI-based neurofeedback training was useful in amending motor function among healthy [2].

Among these brain signals, the EEG signal was quite commonly used in the BCI systems due to its portableness, low cost, non-invasiveness, and higher temporal resolution. Even though several EEG channels provide better information on neural activity, redundancy is increased owing to noise, thus resulting in high-dimensional data. Aside from these above-mentioned reasons, the BCI implementation and feasibility, the BCI cost reduction, and the BCI performance improvement are the major reasons for the minimization of the actual number of EEG signal channels. Channel selection is the term used for this process.

The extraction of information from these recorded EEG signals was carried out to identify certain patterns of activation that have a major role to play in BCI-based research. A system of the BCI that can be successful will have two requirements that are an efficient set of EEG features that can differentiate any form of the task-induced brain. As both frequency and time information of the EEG is very important, a multi-resolution-based wavelet transform can be well-suited in analyzing the EEG compared to a frequency-based Fourier domain. Several studies have been incorporated using the extraction of wavelet transform-based features for BCI applications. One more very popular method of feature extraction employed among MI applications includes the Common Spatial Patterns (CSP). The CSP constructs some more additive sub-windows that have maximum difference invariance. Such extracted features include a time-domain signal's statistical traits to reduce computational complexity [3]. Classification techniques and their application to statistical features are useful in the determination of a BCI-based EEG activation pattern. Therefore, an evaluation of the various ML tools has been done to identify the best and optimum solution for MI task recognition. There has been a remarkable improvement that was observed among conventional algorithms of classification like the Artificial Neural Network (ANN), the Linear Discriminant Analysis (LDA), the Support Vector Machine (SVM), the Bayesian classifier, the K-Nearest Neighbor (KNN) as well as the Quadratic Discriminant Analysis (QDA).

Studies have identified a set of relevant EEG channels that an earlier neural correlate had determined. However, a more discriminative set of EEG channels was later found, showing inconsistencies with the preselected channels. The selection of EEG channels is a major challenge in BCI research. Such conflicts were looked at as a form of multiobjective optimization problems wherein there was more than a single solution to meet the criteria. At times, certain solutions satisfy a single criterion without rendering the other one worse. There may not be one single dominant solution for this, and in place of this, there can be multiple solutions that are Pareto dominant. These multiobjective problems were addressed by making use of evolutionary algorithms that were explored for a paradigm of EEG channel selection. Another Genetic Algorithm (GA) based evolutionary approach to choose discriminative channels. However, the GA will need to carefully consider parameters that can affect the accuracy of classification. The PSO is yet another popular nature-inspired algorithm used in BCI channel selection [4]. In this work, the Stochastic Diffusion Search (SDS) algorithm is chosen for its efficiency, robustness, scalability, and capability to handle complex multi-objective optimization problems. These characteristics make it suitable for EEG channel selection in BCIs, where optimal solutions need to be found within a noisy and high-dimensional environment.

The focus of this work is channel selection using the SDS for the EEG, which is based on the BCI system. There are trade-off coefficients that must undergo adjustments to pick the optimal channel set, which is not feasible for real-time online BCI. The remainder of this investigation has been structured as follows: Section 2 will detail the related literary works. Section 3 will elaborate on the methods employed. Section 4 will describe the experimental outcomes, while Section 5 will offer the conclusion.

2. Related Works

Antoniou et al. [5] presented another novel system with BCI utilization for the capture of every EEG signal during the human subjects' eye movement that can be classified into six different categories by the application of a Random Forest (RF). The RF refers to an ensemble method of learning constructing a decision tree series wherein every tree gives a class prediction. The definition of the proposed RF-BCI classes was given in accordance with the subject's position of the eyes: open, closed, down, up, left, and right. This approach's key goal was to utilize an EEG-based control system for operating an electromechanical wheelchair. A dataset with 219 records from a total of 10 patients was employed for testing this approach. This system was compared to the NB, the K-NN, the Bayes Network, the SVM, the Decision Tree, the Multi-Layer Perceptron (MLP), and the J48-C4.5. The results of the experiment proved that the RF algorithm could outperform this in comparison with other approaches that have higher accuracy (85.39%) for its 6-class classification.

Li et al. [6] presented the Fisher Linear Discriminant Analysis classification algorithm with the incorporation of the Naïve Bayes (B-FLDA) for ERP-BCI in order to perform a concurrent recognition of the targets for the working state as well as the idle state. This approach employed a visual-evoked potential's spectral traits together with the ERP's time-domain traits for detection of the brain states as well as the target stimulus to acquire the results of the discrimination with probability fusion utilization. The accuracy of the information transfer rates will rise to 98.61% and 62.80 bits/min under 10 repetitions, as well as for one repetition. The receiver operator characteristic curve would involve three different parameters having better performance.

Sharma et al. [7] provided a more comprehensive comparison between the traditional methods of classification and also suggested the significance of deep learning-based techniques, particularly for MLP. To summarize this, the EEG signals in a task of motor imagery, the SVM, show a short training time with prediction speed having good accuracy. There was the classification of a subject-independent generalized MLP model in the form of signals with an accuracy of 90% within half the time required for the other conventional ML-based models' classification. The outcome suggested that the chances of better accuracy with a robust as well as generalized BCI (which, in turn, was independent of the subject) were available in the case of the model's integration with a sophisticated optimization.

Shahlaei et al. [8] presented the Hilbert Transform (HT) that was used to detect the ERPs along with ML classifiers to implement the classification of the left MI task as well as the right MI task. The following two distinct steps constituted the proposed method: the first was a sensorimotor frequency band (8-30 Hz) that was associated with the MI activities and had undergone signal extraction. From the HT, extremely critical features such as ERPs and Band Power (BP) were extracted. Additionally, other key features were acquired prior to being fed into ML classifiers, such as the Linear Discriminant Analysis (LDA), the NB, and the SVM. The BCI-competition 2008 Graz dataset (II-b) was used for the testing of the proposed method. The outcomes proved that HT's extracted features showed better performance (with values of % CA=82.22%, K=0.63, as well as Auc =0.81) compared to other methods. This demonstrates that the approach can enhance BCI.

Rajashekhar et al. [9] presented a two-class motor imagery electroencephalogram signal with several automated ML algorithms. In this, data had been decomposed into frequency bands that were identified using the wavelet transform spanning between the range of 0 and 30 Hz. The statistical measures were applied to such frequency bands to train classifiers. Also, the assessment of parameters such as entropy, mean, SD, and SNR was computed to analyze the proposed work and its performance. It was evident from experimental results that the proposed technique had better classifier accuracy when compared to other more technologically advanced techniques.

Zhang and Wei [10] proposed the binary quantum-behaved Particle Swarm Optimization (PSO), which was yet another novel evolutionary search algorithm. This was employed for channel selection, using a wrapping manner coupled with the CSP (for feature extraction) as well as the SVM (for classification purposes). This algorithm's fitness function was defined in the form of a weighted sum, which constituted a classification error rate and a relative channel set. Evaluation of a binary quantum-behaved PSO-based CSP's classification performance was done on the basis of an electroencephalograph dataset together with an Electrocorticography dataset. Later, it was compared with three other distinct types of CSP methods by making use of channels chosen by the binary PSO, the channels in a raw dataset with manually chosen channels.

The results proved that the proposed binary quantum-behaved PSO-based CSP method was able to outperform three other CSP methods and bring about a decrease in the rate of error as well as the channels in comparison to the CSP method with channels for raw datasets. The method improved the practicability of motor imagery-based BCI systems.

Zhang et al. [11] presented the quantum PSO to ensure a frequency band's sole selection through the selection of a frequency band as well as a time segment, which was realized via the utilization of a wrapping technique that had duly incorporated the CSP to extract features within the classification model. The rate of error classification was used in the form of a fitness function for the quantum PSO. Evaluation of the classification performance of the quantum PSO-based CSP algorithm, which was used for a joint selection of the time segment as well as the frequency band, was done by making a comparison with three CSP algorithms.

The PSO and the quantum-behaved PSO are used for frequency selection. Two MI datasets that had several trials and channels were used for evaluation. The outcomes proved the algorithm's superiority over others in terms of their classification error rates. For the proposed algorithm, the averaged error rate was 7.45%, 2.97%, and 2.05% lower than the CSP with a fixed time segment and frequency band with the chosen frequency bands by the PSO and the chosen bands by the quantum PSO. The algorithm could facilitate all real-world BCI applications.

Abenna et al. [12] put forward a novel method for quality improvement of the classification of the EEG motor imagery by means of the utilization of the BCI competition IV 2a, 2b, as well as the PhysioNet EEG-MI datasets. The study used a bandpass filter for the removal of unused signals so as to improve prediction accuracy from 50% to over 96% for both binary and multi-class scenarios. This was done with the knowledge that the application of the PSO optimizer to the LightGBM classifier parameters would allow for stability in the status of the classification of the EEG signals. The DT algorithm (DT) had permitted the degree of importance of the acquisition electrodes that are used in the stage of classification. Also, the study utilized a correlation matrix to determine the artifacts between various electrodes, such that there was an increase in the prediction accuracy with a higher prediction speed that would continuously stay above 63703 and 2395 samples for each second. A comparison was made, and the maximum accuracy value was found to be 85.5%.

In summary, various algorithms, including Random Forest, GAs, and PSO, have shown effectiveness in EEG signal classification and BCI applications, each demonstrating strengths in specific areas like accuracy, speed, or multi-class classification. The challenges remain in developing generalized models that balance accuracy, computational efficiency, and real-time processing. However, the SDS is used in this work for its efficiency in handling complex search spaces and ability to effectively manage multi-objective optimization problems, making it a suitable choice for EEG channel selection in BCIs.

3. Methodology

The BCI IVa is the dataset employed in this section. Common Spatial Patterns (CSP) [13] technique is used for feature extraction from the EEG data. CSP will discriminate the EEG signals by decomposing them into various spatial patterns to enhance the distinctions between classes. The primary purpose of applying the CSP was to maximize the variance for one class of EEG signals while minimizing it for the others. The section includes a discussion of such classification with the NB, CART, AdaBoost, and KNN.

3.1. Channel Selection Using SDS

However, the method employed for performance improvement of the EEG-based BCIs will utilize suitable channels. This is owing to the fact that both redundant and noisy channels will be removed, and their computation complexity will be reduced. Using large numbers of channels may not always be practical since it may require a long time to be set up. There was a method for identifying any subject-specific and optimal channels that are crucial to the application of the BCI and its performance, as there may be suitable channels that are different from each other [14]. The main challenge in the selection of an EEG channel can be feature selection. The methods involved in channel selection have been characterized primarily as either filter or wrapper. In contrast, the methods of feature selection based on less intensive features for computation that do not select an optimal subset feature are also included. Another EEG-specific approach that makes use of CSP coefficients has been used. The protocol serves efficiently while differentiating two areas that measure the EEG in the BCI. The selection of a channel using the CSP coefficients will not select any optimal subsets with suitable channels. This is because the CSP is very sensitive to the outliers, while the EEG signals are very noisy due to the artifacts.

This work aimed to assess several other techniques, such as the CSP and specify SDS usage. This will indicate a population-based algorithm with the implementation of direct communication patterns that were attributed originally to the Bishop in the year 1989. This was in the form of a population-based matching algorithm using direct patterns of communication in search space evaluation [15].

In the SDS, the population of the agent will have a 'hypothesis' about the possible solutions; this hypothesis will be evaluated to give proper feedback that will ensure convergence on promising solutions. With the SDS's utilization, the agents' communication with their partial evaluation of the hypothesis will play a key role in the performance as well as the emergence of intelligence. Below is the description of the SDS's three different phases in Algorithm 1:

$Initiali\,sin\,g\;\; agents$ ()
$While$ ($stopping\ condition\ is\ not\ met$)
$Testing\ hypotheses$ ()
$Diffusion\ hypotheses$ ()
End

The initialization phase takes place when every agent will pick from the search space a random selection of hypotheses (i.e., the index of the element for a dataset or the instance number). Later, these points were utilized to take the lead in the SDS population's search procedure. Upon completion of the initialization, each agent will be provided with a random hypothesis within the search space as well as in the test space, wherein there is an objective function-based evaluation of the hypothesis of each agent. In case there is a successful hypothesis evaluation, the agent will be set as active. Otherwise, it will get set as inactive. Hence, each agent will adopt either one of the two Boolean possible outcomes at the end of the test phase. In the diffusion phase, all hypothesis information will be exchanged. Here, a passive strategy for recruitment will be used wherein every inactive agent will randomly pick one more agent. In case of this agent being active, there will be diffusion of this agent's hypothesis, and then, a hypothesis will be chosen by the inactive agent from the search space.

3.2. Naïve Bayes (NB) Classifier

The NB classifier refers to a very simple and effective technique used in a classifier algorithm [16]. The Bayes theorem further permits one to compute a posteriori probability (which is the probability for a hypothesis taking into consideration the value of the variable). It is on the basis of a priori probability (a hypothesis's frequency) of the detected data as well as the total data in accordance with (1):

$$P(v_j/A) = \frac{P(A/v_j)\times P(v_j)}{P(A)} \tag{1}$$

Wherein: v_j will indicate the hypothesis, j = 1, 2, 3, for a hypothesis set V, while A will indicate a feature set $< x_1, x_2, \ldots., x_n >$ that will describe the data.

When A is composed of more than a single attribute, it becomes important to evaluate $P = (x_1, x_2, \ldots, x_n|v_j)$ for calculation $P = (v_j|x_1, x_2, \ldots, x_n)$. This estimate P (A | v_j) will be expensive computationally if the samples are large. It may be possible to get a good performance classification if the attributes are not independent in (2).

$$P = (x_1, x_2, \ldots., x_n|v_j) = \prod_i P(x_i/v_i) \tag{2}$$

For this, the classifier output has been given by (3):

$$v_{MAP} = \underset{v_j \in V}{argmax}\{P(v_j) \times \prod_i P(x_i/v_i)\} \tag{3}$$

Wherein, v_{MAP} will indicate the maximum *a posteriori* probability that is evaluated inside the space of hypothesis, V.

3.3. K-Nearest Neighbors (KNN) Classifier

The K-NN algorithm can be well-suited for classifying the EEG data as it has been proven to be a technique that is quite robust for dealing with huge and noisy data [17]. All the data samples would have undergone classification by means of the majority vote in the neighbor's class. To determine such a class, the algorithm will need training data with a predefined k value until it searches the sample space to obtain the k-most similar samples on the basis of a similarity measure as well as a distance function. The k value and the distance metric impact the outcomes of the classification. Euclidean distance refers to a measure of finding distance existing between two distinct points. For Cartesian coordinates, if x and y refer to the two distinct points, the Pythagoras theorem will be used for the definition of the distance (d) either from x to y or from y to x as per (4):

$$d(x, y) = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{4}$$

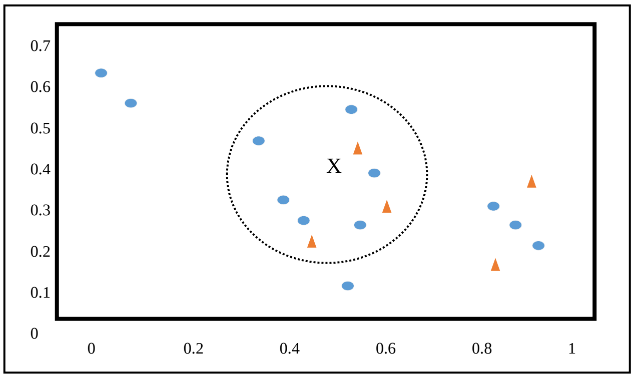

Fig. 1 The K-NN algorithm's classification through utilization of the Euclidean distance and k=9

Figure 1 further illustrates the K-NN algorithms while applying the distance metric to determine the new data's suitable class along with k = 9. There is a classification of the data at point (0.6, 0.45), and this is depicted as an "X" mark. A big circle, as well as a dotted line, will represent a distance metric that makes use of the computation of Euclidean distance. This includes two different classes representing the circle class that has six instances, with a triangle class having three instances. This algorithm further classifies the "X" mark to a new circle class due to the circle class having most of the data within the radius.

3.4. Classification and Regression Tree (CART)

Here, the CART makes use of the Gini Index to engender binary classification trees that make decisions [18]. Firstly, dataset adulteration is assessed. D will refer to the probability, P_n, which is evaluated by |cl, D| / |D| as per Equation (5). In this, the sum over C classes is computed for every instance of $x_n \in D$ class cl.

$$Gini(D) = 1 - \sum_{n=1}^{N} P_n^2 \tag{5}$$

The dataset is split, and D will take into account a binary split with a weighted sum of about 305 adulterations, which will result in the sub-data. As an example, a split of the dataset, D, as D_1as well as D_2 with computation of the D's Gini Index in accordance with the below Equation (6):

$$Gini_A(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \tag{6}$$

For every attribute, A_f will take into consideration that sub-data and this is nominated as a splitting attribute to ameliorate reduction impurities in the following Equation (7):

$$\Delta Gini(A) = Gini(D) - Gini_A(D) \tag{7}$$

As per the Gini Index impurity attribute, there is a split of the dataset as well as the creation of the leaf nodes till all the splitting data fall under an equivalent class. Algorithm 1 depicts the CART algorithm according to this.

3.5. Adaboost Classifier

AdaBoost refers to a technology used to combine weak classifiers to generate a single strong classifier. This weak classifier refers to a simple classifier that will outperform a poor-performing random extraction. Yet another simple example will involve using a person's height to differentiate between men and women. While a man will be considered to be a person above a height of 159 cm, a woman will be considered to be a person with a lower height. In spite of the misclassification of several people, its accuracy is at least about 50%. For a stronger classifier's creation, it can train other weak classifiers and, thus, accordingly combine their results [19].

Firstly, let us take a look at Equation (8), which indicates the last classifier:

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \tag{8}$$

H (x), the last classifier, will include "T" weak classifiers. Here,$h_t$will be the t-th weak classifier's output $\alpha_t$will be the t-th weak classifier's weight by the AdaBoost. So, the final output H (x) will be a linear sum of every one of these weak classifiers. Upon completion of each classifier's training, the probability of training data will be updated to the succeeding classifier. There is training of the first classifier (t = 1), and the entire data will utilize the same probability. Now, each classifier's output weight,α_t, will be computed as below (9):

$$\alpha_t = \frac{1}{2} ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) \tag{9}$$

Output weight,α_t, is based on an error rate ($\in_t$). $\in_t$will be the error or misclassification probability. Once α is computed, the training data weight is updated by using Equation (10):

$$D_{t+1}(i) = \frac{D_t(i)\, exp(-\alpha_{t+1} y_i h_t(x_i))}{Z_t} \tag{10}$$

$Z_t$refers to the sum of $D_t : Z_t = \sum D_t(i)$. D_t which is the vector for every weight in the training data. The variable "i" refers to an index of training data. All equations prove the manner in which the weights for ith training data have been updated.

4. Results and Discussion

In this section, the SDS-NB, SDS-KNN, SDS-CART, and SDS-Adaboost methods are used. Table 1 shows the summary of the results. The classification accuracy, precision, recall, and f measure (features like right and left) are shown in Figures 2 to 5.

Table 1. Summary of results

	SDS-NB	SDS-KNN	SDS-CART	SDS-Adaboost
Classification accuracy	0.8422	0.8356	0.8444	0.8667
Precision for right	0.8683	0.8636	0.8719	0.8926
Precision for foot	0.8116	0.8029	0.8125	0.8365
Recall for right	0.844	0.836	0.844	0.864
Recall for foot	0.84	0.835	0.845	0.87
F measure for right	0.856	0.8496	0.8577	0.8781
F measure for foot	0.8256	0.8186	0.8284	0.8529

Figure 2 shows that SDS-Adaboost demonstrates a higher classification accuracy compared to other methods by 2.86% for SDS-NB, 3.65% for SDS-KNN, and 2.61% for SDS-CART, respectively. The ensemble nature of Adaboost likely contributes to its superior performance, as it combines the strengths of multiple weak classifiers to improve the final prediction.

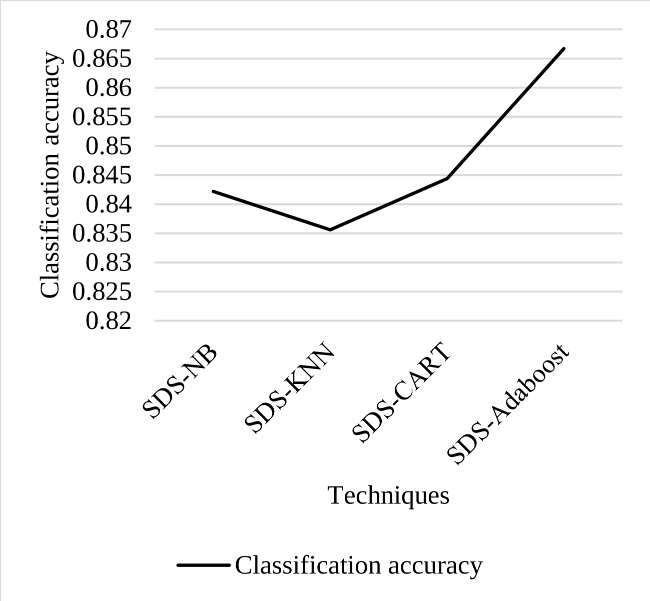

Fig. 2 Classification Accuracy for SDS-Adaboost

Figure 3 shows that SDS-Adaboost demonstrates a higher precision compared to other methods for right by 2.76% for SDS-NB, 3.3% for SDS-KNN, and 2.34% for SDS-CART, respectively. The SDS-Adaboost has higher precision for left by 3.02% for SDS-NB, 4.09% for SDS-KNN, and 2.91% for SDS-CART, respectively. The proposed SDS-Adaboost is more reliable in minimizing false positives.

Figure 4 shows that SDS-Adaboost demonstrates a higher recall compared to other methods for right by 2.34% for SDS-NB, 3.29% for SDS-KNN, and 2.34% for SDS-CART, respectively. The SDS-Adaboost has a higher recall for left by 3.51% for SDS-NB, 4.1% for SDS-KNN, and 2.91% for SDS-CART, respectively. The results show that SDS-Adaboost is effective in identifying most of the true positives.

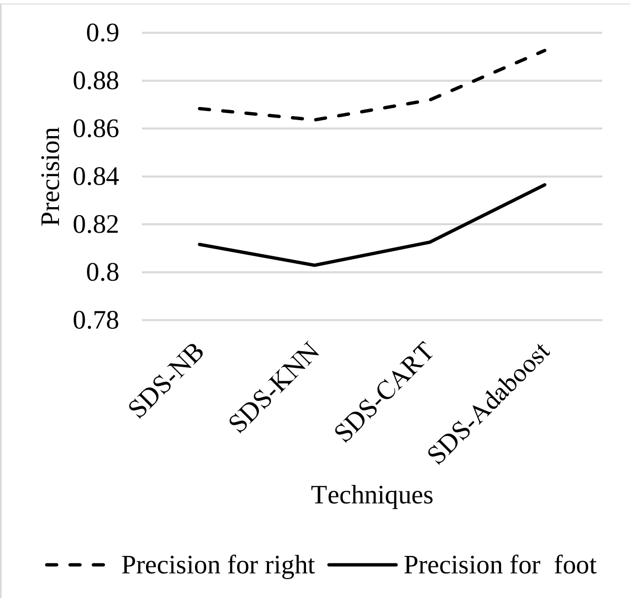

Fig. 3 Precision for SDS-Adaboost

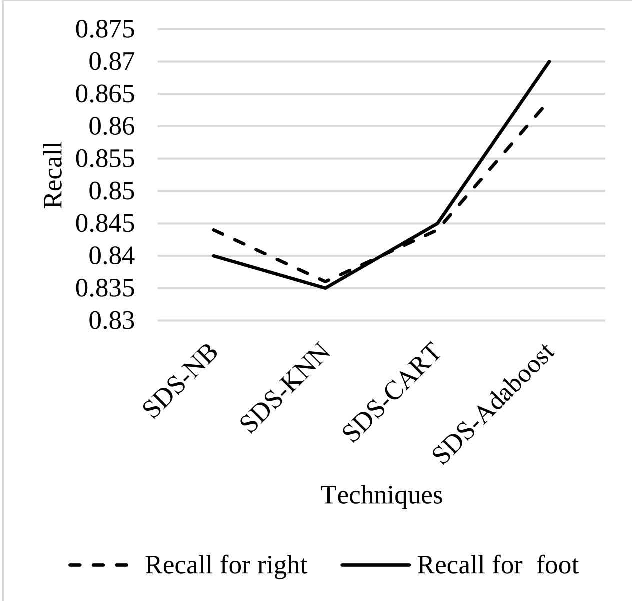

Fig. 4 Recall for SDS-Adaboost

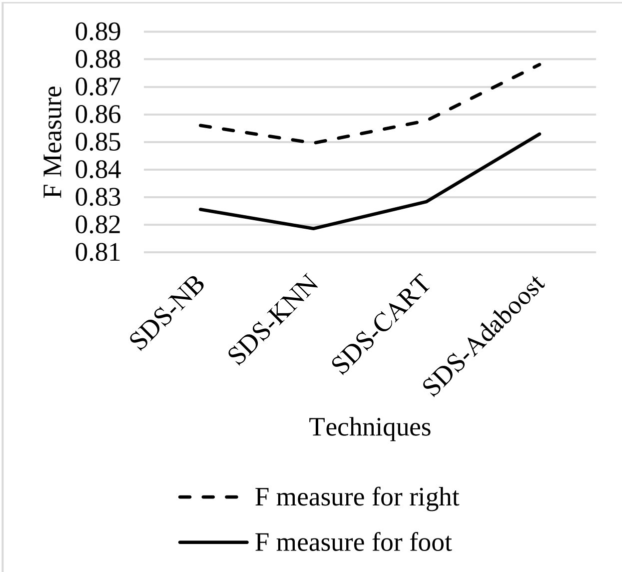

Fig. 5 F measure for SDS-Adaboost

Figure 2 shows that SDS-Adaboost demonstrates a higher f measure compared to other methods for right by 2.54% for SDS-NB, 3.29% for SDS-KNN, and 2.35% for SDS-CART, respectively. The SDS-Adaboost has a higher f measure for left by 3.25% for SDS-NB, 4.1% for SDS-KNN, and 2.91% for SDS-CART, respectively. The SDS-Adaboost achieves the highest f measure, demonstrating its overall performance in both precision and recall.

The SDS algorithm played a critical role in improving EEG channel selection by optimizing the identification of the most relevant channels for classification. SDS enables efficient search across the complex, high-dimensional space of EEG channels, identifying the optimal set of channels leading to better classification. Unlike traditional methods, SDS's population-based approach and its ability to handle multi-objective optimization ensured that both the variance and discriminative power of selected channels were maximized.

5. Conclusion

Several BCI studies have overlooked the optimization of the channel owing to their inherent complexity. However, a careful selection of the channel can increase user comfort and performance, thus reducing cost. Evolutionary meta-heuristics are useful in solving complex problems. Here, the channel selection algorithm with the AdaBoost, NB, CART, and KNN has been proposed. The SDS is an efficient generic search method and serves as a population-based approach for best-fit pattern matching. It employs a one-to-one recruitment system similar to the tandem-running behavior observed in certain ant species. The NB classifier is a probabilistic model based on Bayes' theorem. KNN classifiers, on the other hand, are simple and common but can perform as well as most complex classifiers. Another very important feature of the CART is its ability to generate regression trees. For this, the CART splits and minimizes its prediction squared error. The AdaBoost combines several classifiers and their weighted votes for classification. The results show that an SDS-Adaboost has better accuracy of classification by about 2.86% for the SDS-NB, by 3.65% for the SDS-KNN, and finally by 2.61% for the SDS-CART, respectively.

References

[1] Cili Zuo et al., "Novel Hybrid Brain–Computer Interface System Based on Motor Imagery and P300," *Cognitive Neurodynamics*, vol. 14, pp. 253-265, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Hesam Varsehi, and S. Mohammad P. Firoozabadi, "An EEG Channel Selection Method for Motor Imagery Based Brain–Computer Interface and Neurofeedback Using Granger Causality," *Neural Networks*, vol. 133, pp. 193-206, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Piyush Kant et al., "CWT Based Transfer Learning for Motor Imagery Classification for Brain Computer Interfaces," *Journal of Neuroscience Methods*, vol. 345, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Vikram Shenoy Handiru, and Vinod A. Prasad, "Optimized Bi-Objective EEG Channel Selection and Cross-Subject Generalization with Brain–Computer Interfaces," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 6, pp. 777-786, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[5] Evangelos Antoniou et al., "EEG-Based Eye Movement Recognition Using Brain–Computer Interface and Random Forests," *Sensors*, vol. 21, no. 7, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Mengfan Li et al., "A Fisher Linear Discriminant Analysis Classifier Fused with Naïve Bayes for Simultaneous Detection in an Asynchronous Brain-Computer Interface," *Journal of Neuroscience Methods*, vol. 371, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Rahul Sharma, Minju Kim, and Akansha Gupta, "Motor Imagery Classification in Brain-Machine Interface with Machine Learning Algorithms: Classical Approach to Multi-Layer Perceptron Model," *Biomedical Signal Processing and Control*, vol. 71, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Fatemeh Shahlaei et al., "Detection of Event Related Patterns Using Hilbert Transform in Brain Computer Interface," *2019 IEEE Canadian Conference of Electrical and Computer Engineering*, Edmonton, AB, Canada, pp. 1-4, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] U. Rajashekhar, D. Neelappa, and L. Rajesh, "RETRACTED: Electroencephalogram (EEG) Signal Classification for Brain–Computer Interface Using Discrete Wavelet Transform (DWT)," *International Journal of Intelligent Unmanned Systems*, vol. 10, no. 1, pp. 86-97, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Lei Zhang, and Qingguo Wei, "Channel Selection in Motor Imaginary-Based Brain-Computer Interfaces: A Particle Swarm Optimization Algorithm," *Journal of Integrative Neuroscience*, vol. 18, no. 2, pp. 141-152, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Lei Zhang, Qingguo Wei, and Zongwu Lu, "Combined Optimization of Frequency Band and Time Segment Using Quantum Particle Swarm Algorithm for Brain-Computer Interfaces," *Journal of Circuits Systems and Computers*, vol. 30, no. 13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Said Abenna, Mohammed Nahid, and Abderrahim Bajit, "Motor Imagery Based Brain-Computer Interface: Improving the EEG Classification Using Delta Rhythm and LightGBM Algorithm," *Biomedical Signal Processing and Control*, vol. 71, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Gul Hameed Khan et al., "High Performance Multi-class Motor Imagery EEG Classification," *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020)*, vol. 4, pp. 149-155, 2020. [Google Scholar] [Publisher Link]

[14] A. Franklin Alex Joseph, and C. Govindaraju, "Minimizing Electrodes for Effective Brain Computer Interface," *Biomedical Signal Processing and Control*, vol. 63, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Haya Abdullah Alhakbani, and Mohammad Majid al-Rifaie, "Exploring Feature-Level Duplications on Imbalanced Data Using Stochastic Diffusion Search," *Multi-Agent Systems and Agreement Technologies, Lecture Notes in Computer Science*, vol. 10207, pp. 305-313, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[16] B. Szuflitowska, and P. Orłowski, "Comparison of the EEG Signal Classifiers LDA, NBC and GNBC Based on Time-Frequency Features," *Pomiary Automatyka Robotyka*, vol. 21, no. 2, pp. 39-45, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Nurul E'zzati Md Isa et al., "The Performance Analysis of K-Nearest Neighbors (K-NN) Algorithm for Motor Imagery Classification Based on EEG Signal," *MATEC Web of Conferences*, vol. 140, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] Md. Ochiuddin Miah et al., "Prediction of Motor Imagery Tasks from Multi-Channel EEG Data for Brain-Computer Interface Applications," *BioRxiv*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Kim June-Hyounga, and Yang Yeon-Mob, "An Enhanced Classification Scheme with AdaBoost Concept in BCI," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 1, pp. 63-68, 2018. [CrossRef] [Google Scholar] [Publisher Link]