IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Improved Accuracy for Subject-Dependent and Subject-Independent Deep Learning-based SSVEP BCI Classification: A User-Friendly Approach

ADEEL WAHAB[1], UMAR S. KHAN[1,2], TAHIR NAWAZ[1,2], HASSAN AKBAR[2], S.T. HUSSAIN[1], AZFAR KHALID[3], ALI R. ANSARI[4], RAHEEL NAWAZ[5].

[1]Department of Mechatronics Engineering, National University of Sciences and Technology (NUST), Islamabad, 44000, Pakistan
[2]National Centre of Robotics and Automation (NCRA), College of Electrical and Mechanical Engineering (CEME), National University of Sciences and Technology (NUST), Islamabad, 44000, Pakistan
[3]Department of Engineering, School of Science & Technology, Nottingham Trent University, Clifton, Nottingham, NG11 8NS, UK
[4]Department of Mathematics and Natural Sciences, Gulf University for Science, Mishref, Kuwait
[5]Staffordshire University Stoke-on-Trent, UK

Corresponding author: Adeel Wahab. (e-mail: adeel.whb@ceme.nust.edu.pk)

This work is funded by the Higher Education Commission of Pakistan under grants titled "Establishment of National Centre of Robotics and Automation (DF-1009-31) and this research is partially supported by the Gulf University for Science and Technology."

ABSTRACT In brain-computer interfacing, the SSVEP (steady-state visual evoked potential) method serves to foster collaboration between humans and robots. SSVEP-based detection methods require complex multichannel data acquisition, making them difficult to deploy due to discomfort during extended use and the complexity of the algorithms involved. On the other hand, single-channel setup offers simplicity and ease of use. However, in a single channel, achieving encouraging performance in the SD (subject-dependent) scenario is challenging, and accuracy drops further in the SI (subject-independent) scenario. This requires the development of a generalized approach to improve performance in both scenarios. This study proposes (VMD-DNN) to detect SSVEP in single-channel setups for SD and SI scenarios. The novelty of the proposed method lies in utilizing VMD (Variational Mode Decomposition) as a preprocessor, leveraging harmonic information and Kurtosis of the cross-correlation function to select harmonics from VMD decomposed signal. The preprocessed reconstructed signal uses complex spectrum features as input to the DNN for classification. The results show an average accuracy of 93%, 95.3% in SD and 79%, 92.33% in SI scenarios tested on two publicly available datasets, respectively. The ITR (Information transfer rate) was 67.50 bit/min, 92.31 bit/min for SD, and 46.13 bit/min, 85.94 bit/min for SI for both datasets, respectively. In SD, accuracy is improved by 3.34% and 5%, and ITR by 8.87% and 12.91% over baseline methods for both datasets respectively. The proposed VMD-DNN model is effective, with improved performance and lower computational complexity. The robust single-channel approach makes it user-friendly for human-robot collaboration.

INDEX TERMS Steady-state visual evoked potential, Single-channel, Human-robot collaboration, Deep neural network, Variational mode decomposition, EEG measurement and classification technique, Harmonics.

I. INTRODUCTION

The Brain-Computer Interface (BCI) translates bioelectric signals of the brain in order to communicate between humans and computers or machines. It allows the machine to receive and respond to commands directly from the brain [1]. Increasing technological advancements, Industry 4.0, and cyber-physical systems are reshaping the interaction between humans and machines [2]. The role of humans in highly automated systems has evolved from simple roles to integral components [3]. BCI can serve as an efficient interface, enhancing communication between humans and machines or robots, and is suitable to implement in noisy environments [4]. In a human-robot interface, humans and robots collaborate to accomplish shared tasks, utilizing human intelligence along with robot precision and repeatability [5]. Whenever the human makes a decision, the

robot, as a co-worker, recognizes and executes it accordingly.

The preferable approach for Human-robot Collaboration (HRC) with BCI uses naturally produced brain signals and does not require extra training or active thinking by humans. In addition, it also requires better communication between humans and robots. Steady-state visual evoked potential (SSVEP) is a reactive electroencephalogram-(EEG-) based BCI that occurs as a result of periodic visual stimulation in the occipital cortex, manifested by brain responses to frequency and harmonics [6]. SSVEP-based BCI is preferable for its advantages, such as high signal-to-noise ratio, high information transfer rate, and minimal to no subject or operator training requirement [7]. For the seamless design of human-robot collaboration, it is important to develop an accurate and robust algorithm for SSVEP signals classification with improved Information Transfer Rate (ITR).

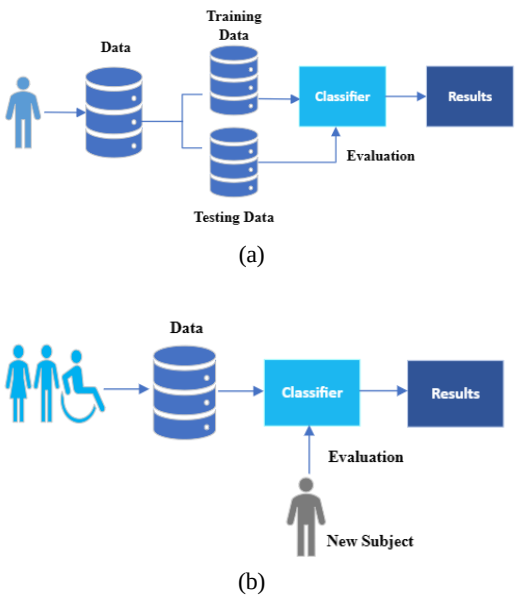


Figure 1. Two training model (a) Subject-dependent training (b) Subject-independent training.

When HRC is used with BCI, two learning agents are involved: the human operator and the algorithm. Operator training for SSVEP-based BCIs is minimal or nonexistent compared with other BCIs such as motor imagery. Generally, the available algorithms for extracting and classifying SSVEP features can be categorized into three groups: training-free methods, subject-dependent training methods, and subject-independent training methods [8]. In the training-free method, there is no need for training data, so new users can begin using BCI immediately [9]. Due to subject-to-subject variability, training-free algorithms are not necessarily robust [10]. Subject-dependent (SD) training methods involve training a particular subject to extract optimal features for the same subject as shown in Fig. 1(a). The subject's training data reduced the impact of spontaneous background EEG signals on SSVEP response [11]. In SSVEP identification, training methods produce higher recognition accuracy than training-free methods because a machine learning algorithm is trained on specific subject data and can predict the outcome later [12]. It is highly desirable to build a general training model that is applicable for new unseen operators, known as the subject-independent (SI) training model. The SI approach requires training data from different subjects to develop a model that can be used for general purposes. In SI models, there is no need to collect training data from new unseen subjects as shown in Fig. 1(b). Once the SI model is trained, it can be applied to predict data from unseen users [10], [12].

Several factors affect SSVEP signal classification accuracy, including the number of channels used for data acquisition, frequency detection algorithm, and signal-to-noise ratio [9], [13]. To achieve high accuracy in SSVEP detection, most studies prefer a multi-channel setup, which, though effective, pose challenges in terms of complexity and discomfort [14], especially for operators wearing EEG caps for prolonged periods. This discomfort may lead to irritation, potentially impacting communication between collaborative robots and operators. Conversely, single-channel setups offer simplicity and comfort [15], a crucial advantage, particularly when operators are required to wear safety helmets simultaneously. This is often not feasible with multichannel setups, also in multichannel data acquisition the processing algorithmic complexity increases due to data from multiple electrodes.

However, achieving satisfactory performance in SD scenario with single-channel setups poses a challenge, exacerbated by further accuracy drops in SI scenario. Addressing this challenge necessitates the development of a generalized approach to enhance SSVEP single-channel performance across both scenarios. To address this need, our study extends efforts by incorporating effective preprocessing using VMD and a lightweight DNN model. Additionally, training data from multiple subjects is integrated, followed by evaluation. To the best of our knowledge, the SI training scenario has not been implemented using single-channel SSVEP setups. Additionally, single-channel configurations have not been utilized in smart industries for Human-Robot Collaboration (HRC).

Our proposed method employs VMD as a preprocessing step, leveraging harmonic information along with Kurtosis of the cross-correlation function (KCCF) to select necessary harmonics from the VMD decomposed signal. The resultant reconstructed signal, obtained through preprocessing, utilizes complex spectrum features as input to the DNN for classification. This preprocessing step effectively denoises the SSVEP signal by extracting first and second harmonics while disregarding irrelevant and noisy information. The integration of a lightweight and improved DNN model compared to more elaborate neural network architectures already presented in the literature [16], aims to generalize the study, making it applicable to both SD and SI scenarios. The introduction of the VMD and DNN result in improved accuracy as reflected by a higher performance of the proposed method (VMD-

DNN) compared to several existing related methods such as Power spectral density analysis (PSDA), Discrete wavelet transform-PSDA (DWT-PSDA), Empirical mode decomposition-PSDA (EMD-PSDA) [17], Canonical correlation analysis (CCA)-DNN, and Filter bank CCA (FBCCA)-DNN on two publicly available datasets: AVI-SSVEP [18] and [19]. The proposed approach VMD-DNN achieved an average accuracy of 93% and ITR of 67.50 bit/min (bpm) for four subjects on the AVI-SSVEP dataset and 95.3% average accuracy with 92.31 bpm ITR on a second dataset with 10 subjects when tested on SD training scenario. Furthermore, it achieved an accuracy of 79% with 46.13 bpm ITR for AVI-SSVEP and 92.33% accuracy with 85.94 bpm ITR for second Dataset when tested on SI training scenario. Overall, this robust single-channel approach makes it user-friendly for various applications in the automation industry and robotics.

The paper is organized as follows. Section II describes the related work. Section III details the materials and methods, including the proposed methodology. Section IV shows and discusses the results, whereas the scope and application of the work are explained in Section V. The paper concludes in Section VI.

II. RELATED WORK

A brain-computer interface has historically been used to support disabled people or patients, such as wheelchairs [20], electric prostheses [21], and feeding robots [22]. During some industrial processes, when operator hands are involved, and robotic collaboration is required to automate the process, BCI can help humans transfer their decisions to robots. In BCI systems based on EEG, a variety of brain responses are used, including SSVEP [5], P-300 [23], and motor imagery responses [24]. Motor imagery requires additional user training, whereas SSVEP and P300 rely on visual stimuli. Due to the SSVEP aligning with the stimulus frequency and possessing notable advantages such as a high signal-to-noise ratio (SNR), fast information transfer rate, and minimal training requirements, it can be widely used in various industrial applications.

Target identification is an essential step in constructing SSVEP-based BCIs, which translate SSVEP signals into commands. As the SSVEP frequency aligns with the stimuli frequency, methods such as PSDA and Canonical Correlation Analysis (CCA) are used when we have prior knowledge of the target frequency. Also, the CCA method is commonly utilized for single-channel data [25]. However, both PSDA and CCA methods can be affected by background noise [8]. Additionally, the CCA method often inadequately incorporates harmonic information [26]. Given that SSVEP signals contain both the target frequency and its harmonics, the filter bank CCA (FBCCA) was proposed based on a filter bank approach that effectively integrates the target frequency and its harmonics information. FBCCA is widely utilized in various training-free SSVEP methods. The FBCCA method requires optimization of three parameters: the number of harmonics in reference signals, the weight vector for sub-bands, and the number of filter banks for sub-bands. As the performance of the FBCCA method depends on these factors, it is a critical consideration in incorporating a BCI system [10]. Many single-channel SSVEP studies use DWT as a preprocessing method because it allows detailed time and frequency localization within the signal. However, choosing the right mother wavelet is a key limitation to optimizing SSVEP performance. EMD overcomes DWT's limitations by eliminating the need for decision-making in wavelet selection [17], [27]. However, the EMD technique has limitations in mixing intermediate modes [28]. VMD effectively addresses the mode mixing problem and has a solid mathematical foundation for signal decomposition. VMD has the ability to distinguish between two harmonics with frequencies that are very close, and this separation effect remains unaffected by variations in the sampling frequency [29]. This characteristic of VMD makes it a suitable choice for integrating SSVEP target frequency information with harmonics.

Many studies in a multichannel setup utilize subject-dependent training scenarios such as individual template-based CCA methods, task-related component analysis (TRCA), and task-discriminant component analysis (TDCA). These methods contribute significantly to SD classification scenarios. It was noted that, in SSVEP processing, TRCA performs was not optimal when dealing with asynchronously processed data [30].

In the industry, SSVEP is used for HRC using multichannel to perform different tasks such as assembly tasks [2], [5], picking and placing defective components [31] etc. On the other hand, a single channel SSVEP based classification is already practiced in a few applications in SD scenarios specifically, the convolutional neural network is utilized to perform an application of BCI spellers [15]. The fuzzy feature threshold algorithm (FFTA) is utilized for single-channel control of the mobile robot [32]. Numerous efforts have been made by utilizing conventional state-of-the-art methods of multichannel systems when applied to SI scenarios, but the results are still not satisfactory [33] because SSVEP signals display non-stationary properties and vary in characteristics among subjects [10], [11]. Hence, the primary objective of this research is to address the challenges inherent in single-channel setups for the SI as well as SD scenarios.

The main challenge in SSVEP-BCI classification is the signal from the same subject for the particular task at different times instants may have different patterns due to several factors such as electromagnetic interference and background noise etc. The stated challenge greatly affects the performance the of conventional machine learning method in the SD scenario which is even worse in SI scenario [11]. Comparing deep learning with other SSVEP classification methods, it offers numerous advantages. Deep learning leverages neural networks with multiple layers to uncover hidden patterns in EEG signals, essential for accurate classification. A network learns more complex

features as data flows through it [34]. The network is fed with preprocessed SSVEP signals, eliminating the need for additional feature extraction. Due to this property of deep learning, it is used for SSVEP signal classification. To meet the demands of research, we utilized a single-channel SSVEP approach using VMD coupled with a customized deep-learning model. This strategy enhances the accuracy of classification in both (SD) and (SI) training scenarios, making it well-suited for implementation in Human-Robot Collaboration (HRC) systems.

Reviewing the state-of-the-art literature, it is evident there are numerous techniques available to improve SSVEP performance with multichannel systems in both training scenarios, but they are less preferable due to their complexity and uncomfortable use. The need therefore remains for a single-channel SSVEP classification approach that is accurate and robust with a better tradeoff in accuracy and ITR, which is addressed in the proposed method.

III. MATERIALS AND METHODS

A. DATASETS

In this study, the performance of the proposed method was evaluated using two publicly available datasets.

1) DATASET A

This dataset is known as the AVI-SSVEP Dataset [18]. There are four subjects in this dataset, consisting of three males and one female, with an age range of 27-32 years. The trials for all four subjects were performed for seven frequencies flickering at 6 Hz, 6.5 Hz, 7 Hz, 7.5 Hz, 8.2 Hz, 9.3 Hz, and 10 Hz. The single electrode was placed at the 'Oz' position, while the reference and ground electrode positions were 'Fz' and 'Fzp' respectively. The electrodes were set as per 10-20 international standards for the placement of electrodes. The duration of each dataset trial consisted of 30 seconds at a 512 Hz sampling rate. The data for each frequency was collected at least three times for all subjects. For four subjects, 92 trials of data were conducted using seven frequencies. For Dataset A, the train-test split was set to 75% training and 25% testing in the SD scenario. Table 1 provides detailed information about Dataset A.

Table 1. Dataset A: Details of the AVI-SSVEP dataset.

Dataset A	No of channels	Subject wise No of trial					Trial duration (sec)	Total classes
		1	2	3	4	total		
AVI-SSVEP	01	24	26	21	21	92	30	7

2) DATASET B [19]:

This dataset was collected for 12 classes ranging from 9.25 Hz to 14.75 Hz with a step size of 0.5 Hz. A total of 10 subjects with normal or correlated normal vision were selected to sit 60 cm away from a 27-inch LCD monitor in a dim room. The eight channels (O1, O2, Oz, PO3, PO4, PO7, PO8, and POz) were used to acquire EEG data at the occipital region. A sampling rate of 2048 Hz was used for data collection, which was then down sampled to 256 Hz for computation. For each class, 15 trials were conducted, and each trial duration was 4sec.

The data was placed in 4-D matrix format (Number of targets, Number of channels, Number of data points, Number of trials). We used single-channel (Oz) data for this study. In Dataset B, 12 out of 15 trials were used for training, and the remaining 3 for testing in the SD scenario. However, in the SI training scenario, we tested the subject with the lowest accuracy in the SD scenario as the unseen subject in each dataset to evaluate the performance of the proposed method. The DNN model was trained using combined data from the remaining subjects in the same dataset. Table 2 provides details on Dataset B.

Table 2. Details of the Dataset B.

Dataset	Maximum channels	Selected channel/s	Subjects	Total trials per class	Trial duration (sec)	Total classes	Total trial
B	08	01 (Oz)	10	15	4	12	1800

Besides the fundamental frequency, the SSVEP signal also contains its harmonic components, which are its integral multiples [6], [7]. Harmonics analysis is also effective for feature extraction and improved classification performance. The SSVEP signal of 10 Hz of subject 4 of Dataset A is shown in Fig. 2(a). The frequency spectrum of the 10 Hz SSVEP signal represents the fundamental frequency (first harmonics), 2nd and 3rd, harmonics with respect to normalized amplitude as shown in Fig. 2(b).

The first harmonic's amplitude is higher than the second and third harmonics. However, in some cases, the second harmonic's amplitude is more than the first harmonic [7]. While the third harmonic holds minimal significance in comparison to the first and second harmonics.

The SSVEP signals contain the stimulus frequency, its harmonics, and noise. To reduce the noise, a moving average filter (MAF) is applied because VMD performance is affected by the noise in the signal. MAF is a time domain finite impulse response filter commonly used for smoothing [35]. The MAF reduces noise without disturbing sharp steps [36]. Initially, the recorded EEG signal is normalized by its absolute maximum amplitude [27]. We scale the values based on the maximum signal value. EEG signal normalization can be expressed as follows.

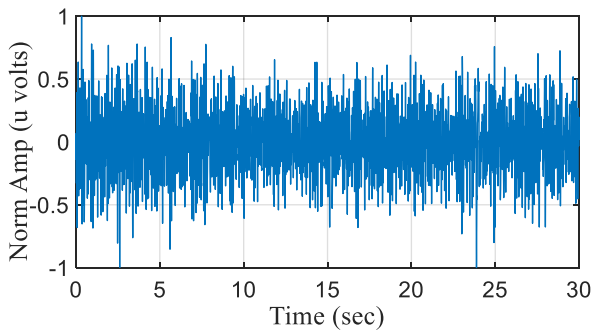
$$O_i = \frac{R_i}{\|\bar{R}\|_\infty} \tag{1}$$

Where, R is the recorded EEG signal and $\|\bar{R}\|_\infty$ is the maximum value of the recorded signal. The normalized signal is represented as O. Then, MAF is applied to suppress

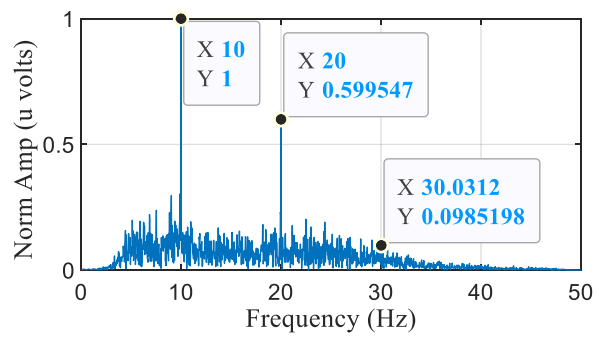
noise added during recording. The MAF is expressed as follows.

$$z_{i,k} = \frac{1}{N_k} \sum_{j=0}^{N-1} O(i-j) \quad (2)$$

Where z is the MAF output signal, and N is the filter length. The $O(i-j)$ and i represent the impulse delayed by j sample and current index value, respectively. It is important to choose an appropriate MAF filter window length for optimal results.



(a)



(b)

Figure 3. (a) Normalized acquired EEG 10Hz signal of Subject 4 (b) FFT-based PSDA of EEG signal.

A shorter filter length may lower the SNR, while a longer filter length can distort the signal and incur higher computational costs. We optimized the MAF length based on the maximum SNR of the EEG signal. The SNR can be represented as below [37].

$$\mathrm{SNR_k} = 10 \log_{10} \frac{S_k}{N_k} \quad (3)$$

Where S_k and N_k are the signal power and noise power respectively. In this study, we consider the power of both the flickering frequency (10 Hz) and its second harmonic (20 Hz) as the power of the required signals, while treating the power of all other frequencies as noise power. Initially, the SNR of the recorded EEG signal is -5.85 dB which increases to -3.27 dB by applying MAF with a filter length of 22 as shown in Fig. 3. The MAF improves the SNR value by 44.19%. Even though the SNR has improved following MAF, the negative sign indicates that the signal power is less than the noise

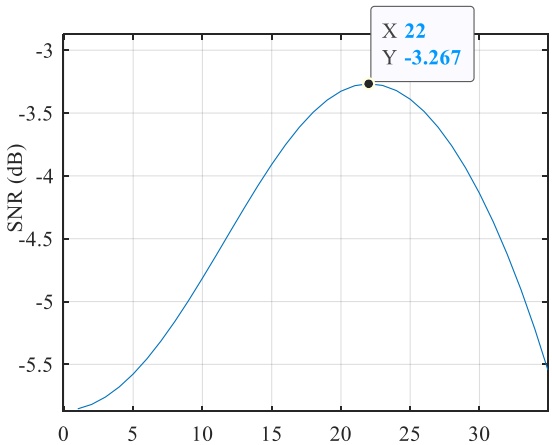


Figure 2. An optimal MAF window length is selected by the maximum SNR of the SSVEP signal.

power. Therefore, VMD is applied to further enhance the SNR of the SSVEP signal.

B. VARIATIONAL MODE DECOMPOSITION

VMD technique is used to decompose the non-stationary and nonlinear signal into various components called intrinsic mode functions (IMFs). Each IMF central frequency and bandwidth are determined by an iterative search for optimization results [38]. The original signal is the sum of all individual IMFs. The harmonics and inter harmonics of the original signal are found by VMD application with the setting of suitable mode numbers [39]. In this study, the EEG signal is converted into its harmonics and inter harmonics using the VMD technique. The $\boldsymbol{z(t)}$ is the output of MAF, which can be decomposed using VMD. The equation of IMF produces as a result of VMD is written as:

$$u_k(t) = A_k(t)cos\big(\emptyset_k(t)\big) \quad (4)$$

Where $A_k(t)$ is the instantaneous amplitude of signal $u_k(t)$ of k_{th} mode and $A_k(t) \geq 0$. $\emptyset_k(t)$ is a non-decreasing phase function and $\emptyset_k(t) \geq 0$. Each IMF bandwidth is found by adopting the following steps:

1. The analytical signal is determined for each mode using the Hilbert transform to get a unilateral spectrum.
2. The mode spectrum is shifted to the baseband by combining with an exponential tuned to central frequency.
3. The signal bandwidth is obtained by applying H^1 Gaussian smoothness. The constraint problem is depicted as:

$$min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * u_k(t)] e^{-j\omega_k(t)} \right\|_2^2 \right\} \quad (5)$$

Subjected to: $\sum_k u_k(t) = f$ where $u_k(t)$ is the K IMFs produced by VMD and $\omega_k(t)$ is the central frequency of each component. δ (t) is the unit pulse function and * is represented as the convolution symbol. Time index is

represented by t, $\{u\} = \{u_1, u_2, \dots, u_k\}$ and $\{\omega\} = \{\omega_1, \omega_2, \dots, \omega_k\}$ are denoted as the modal number and center frequencies after decompositions respectively.

1) VMD PARAMETER OPTIMIZATION

VMD decomposes the signal into different modes number K and the decomposed components are called IMFs. With proper parameter setting, the VMD method shows better robustness, otherwise, it will significantly affect the accuracy of the decomposition results. VMD parameter optimization mainly depends upon two parameters, which are mode number K and plenty factor α [40]. If the value of K is very small, mode aliasing will occur and if K is very large, it will produce false decomposed components. Each IMF has a limited bandwidth size for a particular mode number K and plenty factor α. The bandwidth size mainly depends on the plenty factor α. If α is small, then bandwidth is also small and the larger bandwidth shows the lager value of α [41].

Many methods are available for VMD parameter optimization, but we used the Kurtosis maximum method for optimization [42]. In Kurtosis maximum method, plenty factor α is initially assigned a value and then optimal mode number K is found based on Kurtosis maximum. Afterward, the α value is optimized using the optimized K value. The optimized plenty factor value is determined as the value that produces the maximum Kurtosis within the defined range of α. The steps of Kurtosis maximum method are applied to Dataset A of subject 4 of 10 Hz frequency (shown in Fig. 2) and these steps are as follows.

First, mode number K is optimized. We assigned K = 2 because we are interested in selecting the first two harmonics).

1. Based on the two harmonics approach). Initialize the plenty factor α and bandwidth τ with default values such as α = 2000 & τ=0. Select the range of the K (K ∈ [2,20]). Each IMF kurtosis value (Ku) is calculated under this mode number K, which is written as

$$\mathrm{Ku_m} = (\mathrm{Ku_1}, \mathrm{Ku_2}, \dots, \mathrm{Ku_K}) \quad (6)$$

Where $\mathrm{Ku_m}$ is the Kurtosis value for each IMF and the subscript "m" indicates the Ku value during the mode number optimization process. The local maximum value for each mode number K is written as

$$\mathrm{Ku_{K,l}^{max}} = \max(\mathrm{Ku_1}, \mathrm{Ku_2}, , \dots, \mathrm{Ku_K}) \quad (7)$$

We select the K as K∈ [2,20] so there are 19 local maxima. Therefore, it can be written as

$$\mathrm{Ku_{m,l}^{max}} = \mathrm{Ku_{2,l}^{max}}, \mathrm{Ku_{3,l}^{max}}, \dots, \mathrm{Ku_{20,l}^{max}} \quad (8)$$

The global maximum of the Kurtosis is depicted as

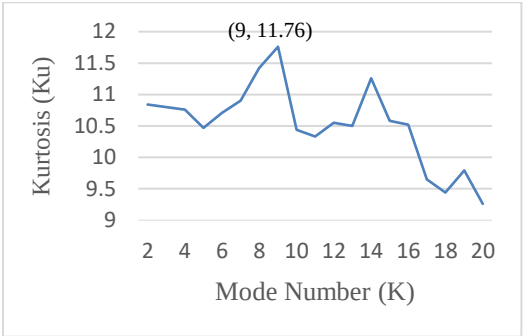


Figure 4. Relationship between Kurtosis and mode number.

$$\mathrm{Ku_{9,g}^{max}} = \mathrm{Ku_{m,g}^{max}} = \max(\mathrm{Ku_{2,l}^{max}}, \mathrm{Ku_{3,l}^{max}}, \dots, \mathrm{Ku_{20,l}^{max}}) \quad (9)$$

The mode value K against this Kurtosis value $\mathrm{Ku_{m,g}^{max}}$ is called the optimized mode number *K'*. The global maximum Kurtosis value is 11.76 under the mode number 9 as shown in Fig. 4. The optimized mode number is *K'* = 9.

2. Afterwards the plenty factor α is optimized. The *K'* value is used to optimize the plenty factor α. Set the value of K as the optimized value *K'* and the bandwidth value as τ=0. The range of the plenty factor α is selected as α ∈ [10, 2000] and the search step size is 10. The kurtosis value of each IMF under the plenty factor α is written as

$$\mathrm{K_p} = (\mathrm{Ku_1^p}, \mathrm{Ku_2^p}, \dots, \mathrm{Ku_{K'}^p}) \quad (10)$$

The local maximum of the $Ku_{p,l}^{max}$ against each plenty factor α is given as

$$\mathrm{Ku_{p,l}^{max}} = \max(\mathrm{Ku_{1,l}^{max}}, \mathrm{Ku_{2,l}^{max}}, \dots, \mathrm{Ku_{K',l}^{max}}) \quad (11)$$

There are 201 values of the plenty factor α, the optimized value of α is found when the kurtosis value is maximum. Therefore, we obtained.

$$\mathrm{Ku_p^{max}} = \mathrm{Ku_{10,l}^{max}}, \mathrm{Ku_{20,l}^{max}}, \dots, \mathrm{Ku_{2000,l}^{max}} \quad (12)$$

The global maximum $Ku_{p,g}^{max}$ can be written as

$$\mathrm{Ku_{p,g}^{max}} = \max(\mathrm{Ku_{10,l}^{max}}, \mathrm{Ku_{20,l}^{max}}, \dots, \mathrm{Ku_{2000,l}^{max}}) \quad (13)$$

The plenty factor α at which Kurtosis value is $Ku_{p,g}^{max}$ is called the optimized value of the plenty factor α and it is denoted as *α'*. The global Kurtosis maximum value is 12.97 when *α'* is 770. The optimized value obtained from step 1 and step 2 is [*K'*, *α'*] = [9, 770].

C. PROPOSED METHOD

The proposed method (VMD-DNN) used to detect the target frequency is represented as a flowchart in Fig. 5. The recorded EEG signal is normalized by its absolute maximum amplitude. Normalization means dividing each component value by the

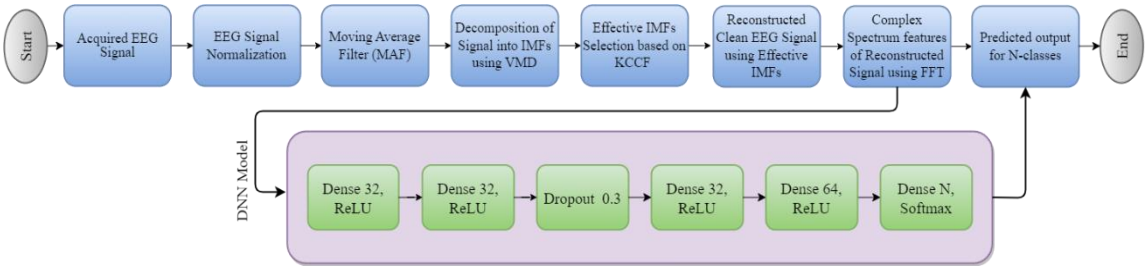


Figure 5. Flow chart for the VMD-DNN (Proposed Method).

maximum component value of the signal, scaling all values relative to the maximum. This approach facilitates optimization processes, such as improving deep learning model convergence, by distinguishing signal components close to the maximum value or significantly smaller. To improve the signal-to-noise ratio (SNR), normalized EEG signals are filtered with MAF. Then, the filtered signal is decomposed into an optimized number of IMFs using VMD. The kurtosis of cross-correlation function (KCCF) is used to segregate the effective VMD-IMFs, which contain the target frequency and its second harmonic. Then the EEG signal is reconstructed by adding those effective VMD-IMFs containing the first and second harmonics of the target frequency. The SNR of the restored signal is enhanced by removing noise from the EEG signal. The FFT of the reconstructed EEG signal is computed to obtain the complex spectrum and subjected to a deep neural network for classification.

There are two challenges associated with the application of the VMD method. Firstly, it requires the optimization of two parameters to achieve accurate signal decomposition: mode number and penalty factor, as explained in Section III (B). Secondly, for SSVEP signal reconstruction, it is necessary to select effective IMFs that contain harmonics.

Various methods have been developed for the detection of harmonics of the VMD decomposed signal such as Strong tracking extended Kalman filter (STEKF) [43], Teager energy operator (TEO) [44], Permutation Entropy (PE) [45], Pearson Correlation Coefficient (PCC) [46]. Using these methods, harmonics were picked based on thresholds. Selection of an effective threshold is challenging and can lead to incorrect harmonic identification. Therefore, we introduce the kurtosis of the cross-correlation function (KCCF) for harmonic selection. The KCCF method efficiently identifies the harmonics of the decomposed SSVEP signal based on its minimum value.

Initially, this method calculates the cross-correlation function (CCF) between each VMD-IMF and its parent signal to find similarities between them. The following equation can be used to calculate the CCF of each IMF with its parent signal.

$$\mathrm{CCF}_{z,u_k(r)} = \sum_{-\infty}^{+\infty} z(t)\, u_k\,(t-r)dt \qquad (14)$$

The subscript z, u_k of cross-correlation, indicates the sequences are correlated, and r is the time shift (lag) parameter. In (13) when r changes from positive to negative, only the signal u_k (t) shifts from right to left. There is a special case when $z(t) = u_k(t)$, known as autocorrelation. Afterwards, the kurtosis value of each CCF is obtained. The Kurtosis value defines the signal's impulsiveness. If kurtosis is greater than three, the signal is impulsive, otherwise, it is sinusoidal [47]. Accordingly, IMFs with less than three KCCF values were ignored. The KCCF (κ) is defined as the following equation.

$$\kappa_m = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_{i,m}-\mu_m)^4}{{\sigma_m}^4} \qquad (15)$$

Where $1 \le m \le \kappa$. The κ represents the KCCF value, y is the output of (14) and N is the sampling length of y. The μ and σ are the mean and standard deviation of y, respectively. The filtered signal is decomposed into 9 IMFs based on VMD using optimized parameters. The VMD-IMFs and their FFT-based PSDA of VMD-IMFs are shown in Fig. 6. The VMD-IMFs produce the impulse response of the first and second harmonics: IMF-7 and IMF-4 respectively.

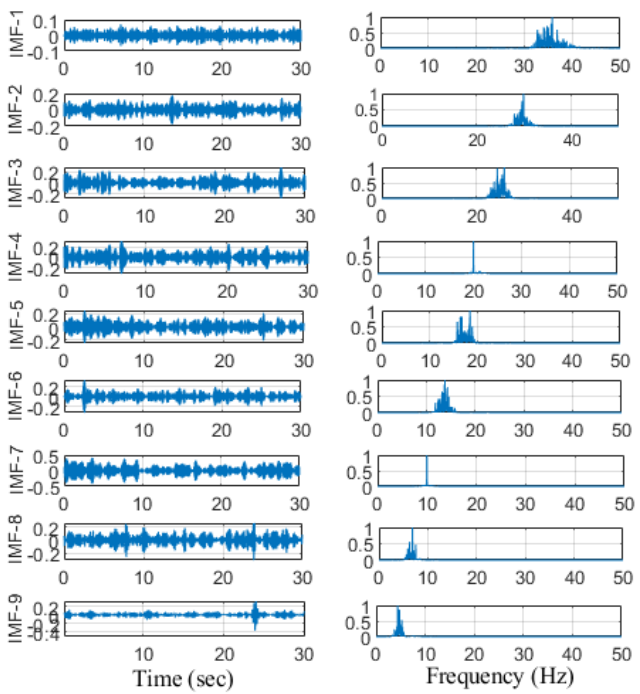


Figure 6. Dataset A: (a) 09 IMFs of VMD decomposed signal of 10Hz (b) and their FFT-based PSDA responses.

Unlike impulsive IMFs, noisy IMFs have more peaks and a higher kurtosis value [48]. Therefore, those VMD-IMFs are declared effective IMFs with greater than 3 KCCF values since less than 3 KCCF values are sinusoidal in nature and not impulsive. The KCCF value of each VMD-IMFs is represented in Fig. 7.

The IMF-7 and IMF-4 represent the first and second harmonic values, respectively, and have the lowest KCCF value. Thus, IMF-7 and IMF-4 were chosen as effective IMFs since this study considers only two harmonics.

All the other IMFs showing higher KCCF values are due to noise. The equations for selecting effective IMFs containing first and second harmonics based on KCCF technique are as follows:

$$\boldsymbol{\kappa} = \{\kappa_i\} \tag{16}$$

Where κ_i is the KCCF value of the u_i (IMF corresponding to κ_i); $1 \leq i \leq K'$, and K' is the total number of optimized IMFs.

$$Rs = u_i\ (\min\{\kappa_i\}) + u_j\ (\min\{\kappa_j\}); \tag{17}$$

Where $\kappa_i\ and\ \kappa_j > 3$, Where $i \neq j$ and $1 \leq i, j \leq K'$ and Rs represents the reconstructed signal with a higher SNR than the recoded signal.

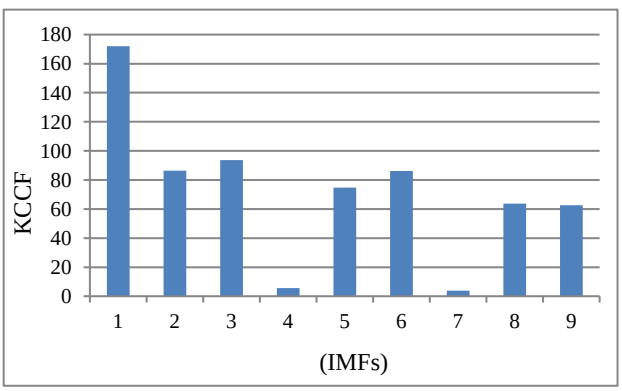


Figure 7. Dataset A: Kurtosis of the Cross-Correlation Function (KCCF) of 9 VMD-IMFs.

1) DEEP NEURAL NETWORK MODEL

The two available SSVEP datasets were used to test the proposed method of VMD-DNN. The pre-processing steps included SSVEP signal normalization, MAF to enhance the SNR, selection of effective VMD-IMFs based on the stimulus frequency and its second harmonics. Subsequently effective IMFs were added to restore the clean SSVEP signal. The complex spectrum features of the reconstructed signal are obtained and given to the DNN model as input. The DNN model was trained for each subject with 1sec, 2sec, and 3sec time windows.

The single feature vector (x) in time domain of the reconstructed signal has dimensions $N_c \cdot N_{dp}$. Where N_c is the number of channels used for recording EEG data, and N_{dp} is the length of the data points. The value of N_c is equal to 1 because we used single channel to record the EEG data. Whereas N_{dp} value is different for different time window (1sec, 2sec and 3sec). The shape of the $N_{dp} = f_s \cdot t_w$. Whereas f_s represents the sampling frequency and t_w is the time window. We provide the model with frequency domain data by performing FFT on the time domain data. The output of the FFT can be expressed as

$$\text{FFT} = \text{Re}[\text{FFT}(x)] + iIm\ [\text{FFT}(x)] \tag{18}$$

Where x is the time domain signal contains N_{dp} values and i represent the imaginary part. The Re[.] and Im[.] represents the real and the imaginary value of FFT. The magnitude spectrum X_{mag} can be written as

$$X_{mag} = \sqrt{\text{Re}[\text{FFT}(x)]^2 + Im[\text{FFT}(x)]^2} \tag{19}$$

Then the complex spectrum can be obtained by concatenating the real and the imaginary parts into a single vector.

$$X_{comp} = \text{Re}[\text{FFT}(x)]\ ||\ \text{Im}[\text{FFT}(x)] \tag{20}$$

The magnitude spectrum contains information related to magnitude, but phase related information is missing. On the other hand, the complex spectrum provides both magnitude and phase-related information. Thus, inspired from the previous studies, we use complex spectrum to preserve both magnitude and phase related information, also complex features perform better than when considering only magnitude spectrum [16], [30]. The complex feature input given to the model is expressed as

$$I_{comp} = [X_{comp}\ (\text{CH}_{O_z})] \tag{21}$$

Where CH_{O_z} describe the data from the EEG O_z channel.

In this study, the proposed deep neural network (DNN) model consists of 5 layers and an output layer. The input layer is a dense layer containing 32 neurons with ReLU activation functions. This layer receives the input data and applies a linear transformation by computing a weighted sum of its inputs, adding a bias term, followed by a non-linear activation function (ReLU). This allows the model to learn more complex patterns. Neurons in dense layers transform the input features into a higher-dimensional space, allowing the model to capture complex patterns and relationships within the data.

In the second layer, 32 neurons are activated by ReLU, followed by a third dropout layer with a 30% dropout rate as in Fig. 5. Dropout is a regularization technique used to prevent overfitting as during training, neurons are dropped randomly, so the network is forced to learn redundant representations, which improves generalization.

The fourth layer consists of 32 neurons with ReLU activation, while the fifth layer has 64 neurons with ReLU

activation. This increase in neurons adds more complexity to the network, helping it capture more complex interactions in the data but with an increase in classification time.
Finally, an output layer with the number of neurons aligned with the total number of classes (Dataset A: 7 classes, and Dataset B: 12 classes). The output layer activation function is Softmax, for multi-class classification. Softmax produces a probability distribution for each class in the outcome and the class with the maximum probability value is the final output of the model.
The number of neurons in a neural network layer affects its learning capabilities and performance. More neurons capture complex patterns, enhancing performance but risking overfitting and requiring more computational resources. In contrast, fewer neurons reduce the risk of overfitting and demand less computation but may lead to underfitting and poor performance. The proposed model balances accuracy and complexity, highlighting this tradeoff.

Table 3. DNN model summary in terms of trainable parameters and FLOPs for 2sec time windows.

Serial No	Layer (type)	Output Shape	Dataset A		Dataset B	
			Parameters (K)	FLOPs (K)	Parameters (K)	FLOPs (K)
1	Dense 1 (Input Layer)	(32)	16.448	32.832	8.256	16.448
2	Dense 2	(32)	1.056	2.048	1.056	2.048
3	Dropout (0.3)	(32)	0	0	0	0
4	Dense 3	(32)	1.056	2.048	1.056	2.048
5	Dense 4	(64)	2.112	4.096	2.112	4.096
6	Dense 5 (Output Layer)	(NC)	0.455	0.896	0.780	1.536
Total			21.127	41.920	13.260	26.176

Table 3 summarizes the DNN model according to layer type, output shape, trainable parameters, and FLOPs (floating point operations) of each layer for both datasets. The output shape specifies the shape of the output produced by each layer and depends on the number of neurons associated with the specific layer of the DNN. The (NC) in the final layer represents the total number of classes for each dataset, which is 7 for Dataset A and 12 for Dataset B respectively. The parameters indicate the number of trainable parameters (weights and biases) associated with each layer of the model when trained on each dataset. The FLOPs represent the number of floating-point operations required for each layer of the model when processing data from each dataset. FLOPs are a measure of the computational complexity of the model during inference or training on the dataset. The total number of parameters for Dataset A are 12.935K, 21.127K, and 37.511K with FLOPs 25.536K, 41.920, 74.688K for the time window of 1sec, 2sec, and 3sec respectively, For Dataset B, the total number of parameters are 9.164K, 13.260K, and 21.452K, with FLOPs 17.984K, 26.176K, and 42.560K for time windows of 1sec, 2sec, and 3sec respectively. Overall, considering the number of trainable parameters and FLOPs, the proposed DNN model appears to be lightweight and improved compared to more elaborate neural network architectures already presented in the literature [16]. It is evident that the more complex the deep learning model is, the more computationally expensive it will be and requires large amount of data for sufficient training [34]. The proposed study underlines a balance between accuracy and computational efficiency, which may also be suitable for certain applications where computational resources are limited or where a simpler model is preferred.
We have presented a custom model that represents a feedforward neural network with multiple hidden layers, employing dense, dropout regularization, ReLU and Softmax activation for multi-class classification tasks. The specific values for the number of neurons, activation functions, and dropout rate were chosen empirically can be adjusted based on the requirements of the task and the characteristics of the data.

2) TRAINING PARAMETERS

We performed several training experiments with several learning rate settings to choose the best hyper-parameters. The number of training epochs and optimizer was fixed to 100 and 'Nadam' (Nesterov-accelerated Adaptive Moment Estimation) respectively. Table 4 shows the few training iteration results regarding the model's validation loss, accuracy, and overall testing accuracy. A backpropagation technique minimizes categorical cross-entropy loss functions to train the network. 'Nadam' was used to train the model as

Table 4. The choice and fine-tuning of hyper-parameters.

Epochs	Learning rate	Validation loss	Validation accuracy (%)	Testing accuracy (%)
100	0.01	1.9380 - 0.5961	0.3650 - 0.7953	76
100	0.001	1.9350 - 0.1547	0.3922 - 0.9521	91
100	0.0001	1.9265 - 0.4579	0.3954 - 0.8115	81
100	0.001	1.9364 - 0.0258	0.2895 - 0.9957	95

an optimizer to minimize errors (loss functions) or maximize accuracy. 'Nadam' is an extension of Adaptive Movement Estimation (Adam) that adds Nesterov's Accelerated Gradient (NAG), an improved version of momentum. The validation split was set to 20 percent of the training data. The batch size (B_s) was set a 2^s, $s \in \{5\}$. The split of both datasets for training and testing is described in Section III (A).

IV. RESULTS AND DISCUSSION

The raw SSVEP signal consists of an effective signal along with noise. The MAF is applied in such a way as to improve the SNR of the SSVEP signal without disturbing the signal characteristics [36]. It is necessary to suppress noise with MAF as a first step since VMD is affected by noise, and background noise can influence decomposition [49]. The optimum MAF window length for EEG signals is determined based on the maximum SNR. As a result, different window

lengths and SNR values were iteratively checked to meet the requirement. A plot of SNR versus window size reveals that MAF optimum window size is 22, as represented in Fig. 3.
The VMD approach is capable of effectively separating harmonics of extremely close frequencies [38]. Due to this capability, VMD is used to identify target frequency and its higher harmonics by SSVEP signal decomposition. In this work, two harmonics are chosen for SSVEP features extraction. We select the two VMD-IMFs based on the lowest KCCF which contains the first and second harmonic peaks. The cross-correlation function between each VMD mode and its parent signal indicates their similarity [50], and the kurtosis defines the impulsiveness of the function [51]. As SSVEP's harmonic order rises, the energy at harmonic frequency points decreases gradually [12], [52]. First and second harmonic IMFs are more impulsive and contain more energy than other IMFs. In addition, kurtosis values greater than three indicate impulsive behavior. Therefore, we select harmonic IMFs with the least KCCF value but greater than three. Based on (3), the SNR after each technique, such as MAF and VMD, is shown in Fig. 8.

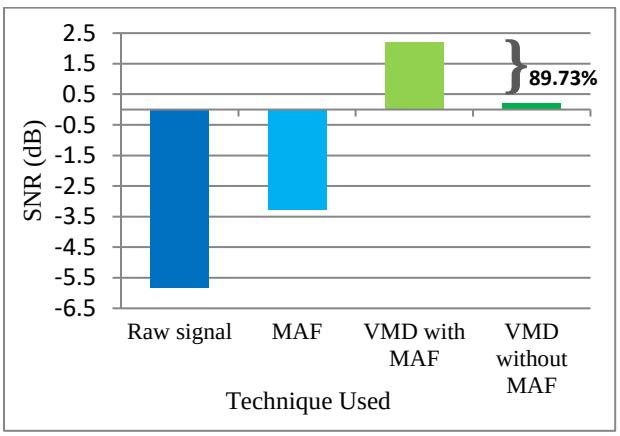


Figure 8. SNR of the SSVEP signal after each technique.

According to Fig. 8, VMD significantly improves the SNR of filtered SSVEP signals. The improvement in SNR from MAF to VMD is 5.49 dB and the overall improvement is 8.08 dB. Thus, in terms of percentage the MAF enhanced the SNR by 44.19% whereas VMD enhanced it by 168.04%. Hence, the overall SNR improved to 212.23%. For SSVEP, the VMD technique demonstrated a significant increase in SNR despite MAF's improvement. Using the VMD technique directly on raw EEG signals without MAF reduces the SNR to 0.23 dB, resulting in an 89.73% drop in SNR. Therefore, VMD performance can be improved with MAF.
The performance of BCI systems was evaluated based on classification accuracy (CA) and information transfer rate (ITR). CA indicates the percentage of correctly classified prediction, and it is defined as the ratio of correctly classified frequencies to the overall number of frequency classes. CA is determined by evaluating the trained model on unseen data splits (SD and SI training scenarios) as described in Section III (A). CA can be expressed as [53]

$$\mathrm{CA} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{22}$$

Where, T_P is the true positive (number of correctly classified target frequency), T_N is the true negative (number of correctly classified non-target frequency), F_P is the false positive (number of incorrectly classified non-target frequency as target frequency) and F_N is false negative (number of incorrectly classified target frequency as non-target frequency). Whereas the ITR measures the communication speed of the BCI system, and it depends mainly on its classification accuracy and processing time. The below equation was proposed by Walpow et al [54] to calculate the ITR of the BCI system.

$$ITR = \frac{60}{T}\left[\log_2 C + (\mathrm{CA})\log_2(\mathrm{CA}) + (1-(\mathrm{CA})\log_2\left(\frac{1-(\mathrm{CA})}{C-1}\right)\right] \tag{23}$$

Where C is the total number of classes, and T is the BCI system processing time. The custom-built model is capable of identifying the unique pattern in each class of the dataset. Initially, the proposed neural network model was trained and tested on subject-dependent scenarios; that is, the model was trained and validated using the same subject with a certain split of the entire dataset into training and testing data. Additionally, the proposed model was trained and validated on all the subjects of Dataset A and Dataset B for different time windows such as 1sec, 2sec, and 3sec. The time window length determines the size of the feature vector utilized for classification purposes. Results of the proposed method (VMD-DNN) on SD-trained model for both datasets are given in Table 5 and Table 6.

Table 5. Dataset A: classification accuracy of SD-trained models for different time windows.

Subject	Classification accuracy for different time window (%)		
	1sec	2sec	3sec
1	77.00	93.00	88.00
2	76.00	93.00	92.00
3	73.00	91.00	87.00
4	79.00	95.00	93.00
Mean value	**76.25**	**93.00**	**90.00**

Based on the above results, it is evident that the classification accuracy for the 2sec time window decomposed signal is better than the 1sec and 3sec for both datasets. For extensive experimental evaluation, we further analyze the 2sec split signal in comparison with several methods such as PSDA, DWT-PSDA, EMD-PSDA, CCA-DNN, and FBCCA-DNN.
In Tables 7 and 8, it is evident that the proposed method (VMD-DNN) achieved an average accuracy of 93% with 67.50 bpm ITR bpm in Dataset A and 95.3% with 92.31 bpm ITR bpm in Dataset B, indicating that it performs better in terms of accuracy and ITR than other baseline methods. As compared to EMD-PSDA, the proposed method improved accuracy by 3.34% and ITR by 8.87% for Dataset A.

Additionally, accuracy improved by 5% and ITR by 12.91% over FBCCA-DNN for Dataset B. The baseline methods such as PSDA and CCA showed deteriorating performance because both are influenced by noise [8]. Additionally, harmonic information is often not sufficiently incorporated into CCA [26]. In the case of DWT, the selection of the right mother wavelet is the key limitation for SSVEP optimized performance [17], [27]. However, EMD is challenged by mixing intermediate modes, [28] which affects its performance in both datasets. Moreover, FBCCA showed comparable performance to our proposed method. The key challenge associated with the FBCCA method is the optimization of three parameters: the number of harmonics in the reference signals, the weight vector for sub-bands, and the number of filter banks for sub-bands. Choosing suitable values for these parameters is critical, as only the appropriate values will ensure optimized performance in FBCCA, making this a complex task [10]. On the other hand, VMD properly decomposes the EEG signal using optimized parameters (mode number and plenty factor). In the proposed method, MAF reduced the noise to 44.19% in the EEG signal and VMD further suppressed the noise to 168.04% (as shown in Fig. 8) by picking relevant harmonics IMFs using the KCCF method. The DNN facilitates robust classification of target frequencies. Therefore, compared to baseline methods, the proposed method performed better in terms of classification accuracy and ITR.

Table 6. Dataset B: classification accuracy of SD-trained models for different time windows.

Subject	Classification accuracy for different time window (%)		
	1sec	2 sec	3 sec
1	85.00	96.00	91.00
2	88.00	96.00	93.00
3	88.00	95.00	91.00
4	84.00	92.00	90.00
5	91.00	96.00	96.00
6	84.00	95.00	92.00
7	92.00	99.00	96.00
8	85.00	93.00	89.00
9	92.00	97.00	93.00
10	87.00	94.00	91.00
Mean value	**87.60**	**95.30**	**92.20**

We conducted a detailed comparative analysis of the KCCF method with other methods on two subjects from each dataset, as shown in Table 9. The subjects were chosen based on the highest accuracy (Dataset A: subject 4, Dataset B: subject 7) and lowest accuracy (Dataset A: subject 3, Dataset B: subject 4) values across each dataset. The KCCF method selects only the minimum value and sets criteria for choosing effective IMFs for SSVEP signals, while other methods use either a set of values or a threshold value, which can vary with different signals. The KCCF method outperforms other related methods in selecting the number of harmonics (NoH), as shown in Table 9. KCCF has a computational cost (CC) comparable to TEO, PCC, and STEKF, while PE exhibits relatively high computational costs. Overall, the KCCF method demonstrates superior performance than other methods analyzed. Table 9 shows the results for selected frequencies from both datasets. Dataset A includes 42 trials for 7 classes of subjects 3 and 4, as shown in Table 1. Similarly, Dataset B includes 360 trials for 12 classes of two subjects, as represented in Table 2. As the other frequencies exhibited similar behavior, we focused on these frequencies to avoid generating excessively large data in the Table.

It is observed that classification results are improved when signals in frequency domain [16] because in the frequency domain, the magnitude spectrum is distinguishable according to each class which improves the classification capability of the DNN.

In Fig. 6(b), the FFT based PSDA of the VMD decomposed signal is represented, while the time domain VMD decomposed signal is shown in Fig. 6(a). In frequency domain, the magnitude spectrum for each class has higher strength at a target frequency and its second harmonics whereas other frequencies are suppressed, and this effect is not present for a signal in time domain.

For extensive evaluation of the proposed VMD-DNN several tests are performed for subject-independent scenario. To evaluate the performance of the proposed method on unseen subjects, we isolated the data of the subject with the lowest accuracy in the SD training scenario from that of the remaining subjects. The DNN model was then trained using the combined data from the remaining subjects. In dataset A, subject 3 showed degraded performance as compared to the other subjects in the SD evaluation scenario, as illustrated in Table 7. As a result, Subject 3 was checked as unseen using a DNN-trained model based on the combined data of other subjects. In dataset B, data from nine subjects were combined to train the model; this model's performance was evaluated for the lowest performer (subject 4) in the SD case, as shown in Table 8. In the same manner, the second lowest (subject 8) and the third lowest (subject 10) were also evaluated as unseen subjects separately. The result of the SI training scenario for different time windows is represented in Table 10.

The proposed method is evaluated across different time windows, such as 1sec, 2sec and 3sec for SI training scenarios, as shown in Table 10. In Dataset A, Subject 3 exhibited encouraging performance with 79% classification accuracy and 46.13 bpm ITR when evaluated against the trained model as unseen for a 2sec time window. Furthermore, in Dataset B, containing 12 classes and 10 subjects, the three subjects tested as unseen by the DNN model showed significantly improved average accuracy of 92.33% with 85.94 bpm ITR for the 2sec time window. However, the CA and ITR of the proposed method for the 1sec and 3sec time windows are lower than for the 2sec across both datasets. From Table 5, Table 6, and Table 10, it is evident that in the SD scenario where training and evaluation utilize the same subject data, the CA and ITR are notably superior to the SI scenario. This is because, in the SI

Table 7. Dataset A (comprising of 7 classes): Comparison of proposed method VMD-DNN with different methods in terms of classification accuracy (CA) in percentage and Information transfer rate (ITR) in bpm

Subject	PSDA		DWT-PSDA		EMD-PSDA		CCA-DNN		FBCCA-DNN		VMD-DNN	
	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)
1	84.70	53.92	88.88	60.43	91.66	65.42	75.00	80.99	87.00	56.29	**93.00**	**66.48**
2	78.90	45.62	86.15	56.14	89.23	61.10	73.00	76.08	87.00	55.84	**93.00**	**67.81**
3	80.64	47.96	78.73	45.33	85.71	55.42	68.00	64.55	84.00	52.78	**91.00**	**64.15**
4	87.30	57.90	87.93	58.99	92.06	66.06	77.00	86.09	88.00	58.89	**95.00**	**71.57**
Mean	82.88	51.35	85.42	55.22	89.66	62.00	73.25.	76.93	86.50	55.95	**93.00**	**67.50**

Table 8. Dataset B (comprising of 12 classes): Comparison of proposed method VMD-DNN with different methods in terms of classification accuracy (CA) in percentage and Information transfer rate (ITR) in bpm.

Subject	PSDA		DWT-PSDA		EMD-PSDA		CCA-DNN		FBCCA-DNN		VMD-DNN	
	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)
1	63.74	41.57	70.42	49.58	83.79	68.14	85.00	71.89	91.00	81.06	**96.00**	**91.55**
2	75.27	57.39	77.66	61.07	81.24	66.85	83.00	69.83	92.00	86.75	**96.00**	**95.65**
3	50.49	26.10	68.72	48.07	69.41	49.03	81.00	66.62	91.00	84.90	**95.00**	**93.53**
4	53.9	29.84	55.26	31.36	61.21	38.39	75.00	57.26	84.00	71.91	**92.00**	**87.18**
5	80.65	64.59	83.41	69.16	88.61	78.42	88.00	77.28	93.00	87.13	**96.00**	**93.78**
6	77.81	58.96	79.41	61.41	83.92	68.68	79.00	60.78	90.00	79.52	**95.00**	**89.73**
7	81.94	67.69	86.19	75.09	89.22	80.76	89.00	80.33	92.00	86.32	**99.00**	**103.06**
8	58.64	34.25	70.68	49.45	77.24	58.95	77.00	58.58	88.00	76.91	**93.00**	**86.71**
9	76.84	57.78	81.59	65.17	82.54	66.72	80.00	62.63	92.00	83.83	**97.00**	**94.81**
10	65.42	41.67	75.41	55.13	80.56	62.91	78.00	58.96	90.00	79.14	**94.00**	**87.14**
Mean	68.47	47.98	74.87	56.55	79.77	63.88	81.50	66.42	90.30	81.75	**95.30**	**92.31**

Table 9. Comparison of the KCCF method with different methods in terms of Number of harmonics (NoH) selection and computational costs (CC).

Description			KCCF		PCC		TEO		STEKF		PE	
Dataset	Subject	Class (Hz)	NoH	CC (msec)	NoH	CC (msec)	NoH	CC (msec)	NoH	CC (msec)	NoH	CC (msec)
A	4	10	2	1.5	1	1.2	1	1.7	1	1.0	1	6.6
A	3	7.5	1	1.9	1	1.0	0	2.1	0	1.1	0	5.4
B	7	9.75	2	1.3	1	0.9	1	1.5	1	0.8	1	4.7
B	4	13.25	2	1.4	0	1.2	1	1.2	0	0.9	1	4.7

Table 10. Classification accuracy (CA) in percentage and ITR in bpm subject-independent (SI) training scenario on different time window

Dataset	unseen Subject	Performance of SI training scenario for different time windows					
		1sec		2sec		3sec	
		CA (%)	ITR (bpm)	CA (%)	ITR (bpm)	CA (%)	ITR (bpm)
B	4	74.00	102.32	**91.00**	**83.04**	88.00	51.44
	8	77.00	105.86	**92.00**	**84.64**	87.00	50.38
	10	72.00	90.30	**94.00**	**90.15**	88.00	49.51
Mean Value		74.33	99.49	**92.33**	**85.94**	87.67	50.44
A	3	69.00	58.06	**79.00**	**46.13**	76.00	27.03

scenario, training and evaluation involve different subjects. The challenge in achieving better performance in SI scenario lies in the fact that SSVEP signals display non-stationary properties and vary in characteristics among subjects for the same task [10], [11]. However, integrating the proposed DNN architecture aims to enhance the study's generalizability. Results from the SI scenario indicate that incorporating training data from a diverse and larger number of subjects can significantly improve accuracy and ITR, as demonstrated in Table 10 for dataset B. This underscores the potential for enhancement even in the SI scenario.

From Tables 5, 6, and 10, it is found that the classification accuracy of the 2sec time window is better than the 1sec and 3sec. Usually, shorter time windows lead to lower accuracy but higher ITR and vice versa [30]. A shorter window length, such as 1sec, may encompass less SSVEP information or fewer flickering cycles compared to a longer window length. As a result, 1sec has a lower accuracy than 2sec and 3sec However, in our case with longer time window lengths such as 3sec there is more noise and more feature information. Because we add two effective IMF (containing the first and second harmonics) during the reconstruction phase. This addition not only improves the signal component but also increases the noise component. An increase in time window size means an increase in the length of a feature vector. Each time window length is linked to a temporal aspect. Shorter windows offer enhanced temporal resolution, allowing for

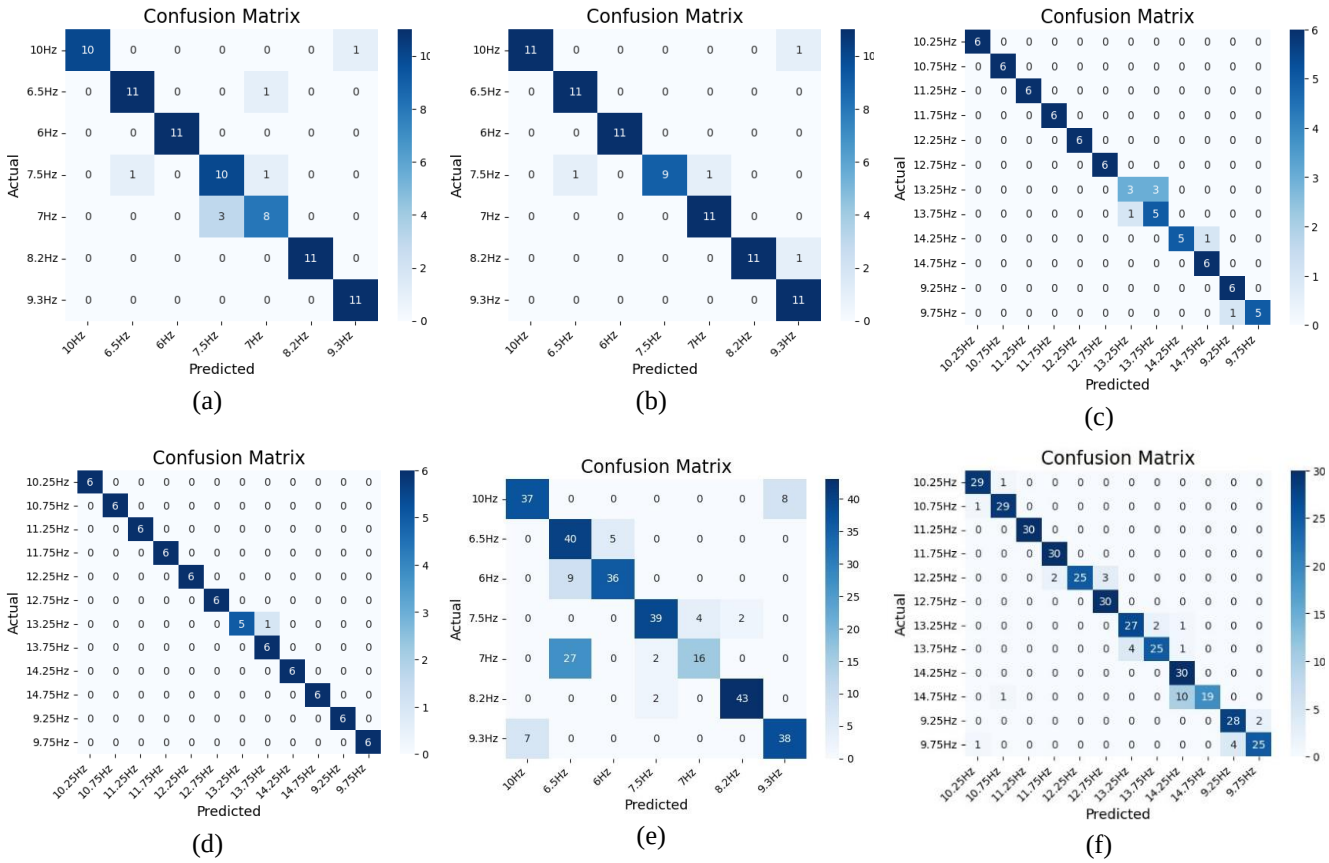


Figure 9. **(a) Dataset A: Confusion matrix of Subject 3 (minimum accuracy) in SD scenario (b) Dataset A: Confusion matrix of Subject 4 (maximum accuracy) in SD scenario (c). Dataset B: Confusion matrix of Subject 4 (minimum accuracy) in SD scenario (d). Dataset A: Confusion matrix of Subject 7 (maximum accuracy) in SD scenario (e). Dataset A: Confusion matrix of unseen Subject 3 in SI (f). Dataset B: Confusion matrix of unseen Subject 4 in SI.**

more precise capture of rapid changes in the SSVEP signal. Conversely, longer windows provide a broader perspective of the signal and a chance of more noise because of the addition of effective IMFs. This can potentially lead to missing quick-changing elements. We found that the classification accuracy of the 3sec window is slightly lower than that of the 2sec window because of this effect. This is also evident in Table 5, 6, and 9. Furthermore, one additional factor is the number of feature vectors provided to the DNN model. Since in the case of the 3sec window, we are dividing the feature vector length by 3, hence a smaller number of feature vectors are produced and used in training of DNN model.

However, this is not the only case, in some individual cases the classification accuracy of the 3sec is higher as compared to 1sec and 2sec but considering the overall performance of algorithm for our datasets, we found that selection of 2sec window is better. Moreover, the ITR rate is a crucial factor for practical implementation which is highest for the case of 1sec but has lower classification accuracy when compared to 2sec and 3sec windows. Thus, the 2sec window is the optimal window size for classification accuracy and ITR for practical implementation.

We have added four representative confusion matrices based on the minimum and the maximum accuracy values from both datasets in SD scenario. Additionally, two confusion matrices are also incorporated one from each dataset (Dataset A: subject 3 and Dataset B: subject 4) to represent the SI scenario classification details of the unseen subject (lowest performer in SD scenario). These confusion matrices are added to show the accurate classified and misclassified samples for each class and for both datasets, as shown in Fig. 9.

However, despite the advantages of the proposed study in accuracy and ITR, some limitations still exist. The limitation of our proposed method is that it generally detects two harmonics successfully. However, further improvements are needed to detect more numbers of harmonics. Furthermore, the study is evaluated on two publicly available datasets. To address this limitation, future research will involve the inclusion of additional SSVEP single-channel custom datasets, as well as an evaluation of the method's performance in an online scenario. Considering the importance of SSVEP-based BCI systems, the technology needs to be widely adopted in industrial applications. Therefore, optimized setups are required to be developed to leverage the benefits of this study.

V. SCOPE AND APPLICATION

Initially, the SSVEP-based BCI system found application in various domains to assist disabled individuals, including controlling robotic wheelchairs, prostheses, and robots for patient care, such as assistive feeding robot etc. With the emergence of Industry 4.0 and cyber-physical systems, there have been remarkable advancements in human-robot collaboration, especially in leveraging SSVEP-based interfaces for interaction between humans and robots. HRC with SSVEP-based BCI is widely employed in various smart industries for tasks such as assembly [5] and segregation of defective parts [31] , utilizing a multichannel EEG cap for high accuracy.

The proposed work focuses on a single-channel setup, which offers more comfort for the operator while maintaining better accuracy. For smart industries and automation, where real-time monitoring and control are crucial, the study's emphasis on single-channel setups and efficient classification methods could be particularly beneficial. The ability to extract meaningful information from SSVEP signals with high accuracy and speed suggests that SSVEP-based BCIs could contribute to improved human-machine interaction and automation processes. Considering the importance of automation there is huge demand for extensive research in this domain. In the future, we can expect even more diverse applications and broader adoption of SSVEP-based BCI.

For practical implementation, the following guidelines and mitigation measures can be used for better performance. A single electrode (Oz) should be installed for data acquisition rather than data extraction from an EEG cap containing multiple electrodes. Additionally, the LED monitor used to display the flickering frequency should have a high refresh rate (60 Hz or above) because it provides more comfort for the user by ensuring accurate frequency presentation. It is important to keep the lab and real environment as close as possible, such as maintaining a quiet, dimly lit room to minimize noise. The operator sits at a suitable distance from the LCD, and for classification, uses the optimum window size to reduce noise and improve the SNR. In the case of an inexperienced user, additional measures should be taken like explaining him/her complete guidelines before use and a suitable rest period should be provided for the user between two data acquisition trials. More comprehensive information for practical implementation of multichannel setup is provided in reference [5], [31].

VI. CONCLUSION

In this work, we propose an effective and efficient method (VMD-DNN) for detecting SSVEP frequencies using a single-channel setup. The SSVEP signal is decomposed into its harmonic IMFs and noisy IMFs by VMD. Based on the lowest value of the KCCF, effective IMFs are selected that contain the first and second harmonics of the target frequency. This method employs VMD as a preprocessor and DNN for effective feature extraction and classification of the SSVEP signal.

The efficacy of the VMD-DNN (proposed method) was subsequently validated by utilizing two publicly available datasets in two training scenarios: subject-dependent and subject-independent. Additionally, the proposed method is evaluated against existing relevant methods and results further validates the effectiveness of the VMD-DNN.
Following are the conclusions drawn from the study.

- By incorporating KCCF, the signal-to-noise ratio is enhanced by selecting the stimulus frequency and its harmonics. This can be achieved by dropping the irrelevant noisy VMD-IMFs from the EEG signal.
- An average accuracy of 93% and ITR of 67.50 bpm for Dataset A is achieved. For Dataset B, it attains 95.30% accuracy and 92.31 bpm ITR for the SD training scenario.
- The proposed approach was also tested for the SI training scenario, achieving 79% accuracy with 46.13 bpm ITR (Dataset A) and 92.33% average accuracy with 85.94 bpm ITR (Dataset B).
- According to the results, the proposed approach outperforms existing related methods such as PSDA, DWT-PSDA, EMD-PSDA, CCA-DNN, and FBCCA-DNN in terms of accuracy and ITR.
- The results highlight the generalization capability of the deep learning model and eliminate the need for additional training sessions for new unseen operators.
- The proposed method, VMD-DNN, is based on single-channel SSVEP frequency detection, which provides more comfort for the operator during longer durations of work.
- The proposed VMD-DNN model appears to be lightweight and more effective in performance with fewer trainable parameters and FLOPs and unified approach for both scenarios that is SD and SI.

This makes our proposed method user-friendly and a suitable choice for Human-Robot Collaboration (HRC) applications within the smart industry. In future work, we will implement the proposed method (VMD-DNN) for seamless HRC to enable a single robot to perform multiple tasks.

ACKNOWLEDGEMENT

The authors wish to acknowledge the support provided by NCRA (National Centre of Robotics and Automation) under RDDL (Robot Design and Development Lab) of College of Electrical and Mechanical Engineering, NUST Islamabad Pakistan. This work is funded by the Higher Education Commission of Pakistan under grants titled “Establishment of National Centre of Robotics and Automation (DF-1009-31)”. Furthermore, the authors are also very thankful for the partial support provided by Gulf University for Science and Technology.

REFERENCES

[1] J. Peksa and D. Mamchur, "State-of-the-Art on Brain-Computer Interface Technology," Sensors, vol. 23, no. 13, pp. 1-28, 2023.
[2] U. Asad, M. Khan, A. Khalid, and W. A. Lughmani, "Human-Centric Digital Twins in Industry: A Comprehensive Review of Enabling Technologies and Implementation Strategies," Sensors, vol. 23, no. 8, pp. 1-27, Apr. 2023.
[3] B. C. Pirvu, C. B. Zamfirescu, and D. Gorecky, "Engineering insights from an anthropocentric cyber-physical system: A case study for an assembly station," Mechatronics, vol. 34, pp. 147–159, Mar. 2016.
[4] Leopoldo Angrisani et. al, "Wearable Augmented Reality and Brain Computer Interface to improve human-robot interactions in smart industry", in Proc. IEEE Int. Conf: Research and Technologies for Society and Industry, 2018, pp. 1-5.
[5] Y. Dmytriyev, F. Insero, M. Carnevale, and H. Giberti, "Brain–Computer Interface and Hand-Guiding Control in a Human–Robot Collaborative Assembly Task," Machines, vol. 10, no. 8, pp. 1-16, Aug. 2022.
[6] Y. Jiao, Y. Zhang, Y. Wang, B. Wang, J. Jin, and X. Wang, "A novel multilayer correlation maximization model for improving CCA-Based frequency recognition in ssvep brain-computer interface," Int J Neural Syst, vol. 28, no. 4, pp. 1-14, May 2018.
[7] V. Çetin, S. Ozekes, and H. S. Varol, "Harmonic analysis of steady-state visual evoked potentials in brain computer interfaces," Biomed Signal Process Control, vol. 60, pp. 1-6, Jul. 2020.
[8] Y. Zhang, S. Q. Xie, H. Wang, and Z. Zhang, "Data Analytics in Steady-State Visual Evoked Potential-Based Brain-Computer Interface: A Review," IEEE Sens J, vol. 21, no. 2, pp. 1124–1138, Jan. 2021.
[9] 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, Jun. 2007.
[10] R. Zerafa, T. Camilleri, O. Falzon, and K. P. Camilleri, "To train or not to train? A survey on training of feature extraction methods for SSVEP-based BCIs," J Neural Eng, vol. 15, no. 5, pp. 1-25, Jul. 02, 2018.
[11] J. Chen, Y. Zhang, Y. Pan, P. Xu, and C. Guan, "A transformer-based deep neural network model for SSVEP classification," Neural Networks, vol. 164, pp. 521–534, Jul. 2023.
[12] 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, pp. 1-13, 2022.
[13] Y. Te Wang, M. Nakanishi, Y. Wang, C. S. Wei, C. K. Cheng, and T. P. Jung, "An Online Brain-Computer Interface Based on SSVEPs Measured from Non-Hair-Bearing Areas," IEEE Trans on Neural Sys and Rehabilitation Engineering, vol. 25, no. 1, pp. 14–21, Jan. 2017.
[14] C. Farmaki, M. Krana, M. Pediaditis, E. Spanakis, and V. Sakkalis, "Single-channel SSVEP-Based BCI for robotic car navigation in real world conditions," In Proc. IEEE Int. Conf. on Bioinformatics and Bioengineering, Oct. 2019, pp. 638–643.
[15] T. H. Nguyen and W. Y. Chung, "A single-channel SSVEP-based BCI speller using deep learning," IEEE Access, vol. 7, pp. 1752–1763, 2019.
[16] D. Zhao, T. Wang, Y. Tian, and X. Jiang, "Filter bank convolutional neural network for SSVEP classification," IEEE Access, vol. 9, pp. 147129–147141, 2021.
[17] M. K. Ojha and M. K. Mukul, "A Novel Approach Based on EMD to improve the Performance of SSVEP Based BCI System," Wirel Pers Commun, vol. 118, no. 4, pp. 2455–2467, Jun. 2021.
[18] "AVI SSVEP Dataset – Adnan Vilic." Accessed: Aug. 11, 2022. [Online]. Available: https://www.setzner.com/avi-ssvep-dataset/
[19] M. Nakanishi, Y. Wang, Y. Te Wang, and T. P. Jung, "A comparison study of canonical correlation analysis-based methods for detecting steady-state visual evoked potentials," PLoS One, vol. 10, no. 10, pp. 1-18, Oct. 2015.
[20] T. Carlson and J. Del R. Millan, "Brain-controlled wheelchairs: A robotic architecture," IEEE Robot Autom Mag, vol. 20, no. 1, pp. 65–73, 2013.
[21] G. R. Müller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," IEEE Trans Biomed Eng, vol. 55, no. 1, pp. 361–364, Jan. 2008.
[22] S. C. Chen, C. M. Wu, I. A. E. Zaeni, and Y. J. Chen, "Applying fuzzy decision for a single channel SSVEP-based BCI on automatic feeding robot," Microsystem Technologies, vol. 24, no. 1, pp. 199–207, Jan. 2018.
[23] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The Mental Prosthesis: Assessing the Speed of a P300-Based Brain-Computer Interface, IEEE tran, on rehabilitation engineering, vol. 8, no. 2, pp. 174-179, June 2000.
[24] G. Pfurtscheller and C. Neuper, "Motor Imagery and Direct Brain-Computer Communication." Proc. IEEE Conf. 2001, pp. 1123-1134.
[25] A. Talha and M. B. Öz, "Subject-Specific Sinusoid Approach for A Brain–Computer Interface Based on Single-Channel Steady-State Visual Evoked Potential", Journal of Science of Engineering, pp. 1-12, 2021.
[26] 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, pp. 1-11, Apr. 2022.
[27] M. K. Ojha and M. K. Mukul, "Detection of Target Frequency from SSVEP Signal Using Empirical Mode Decomposition for SSVEP Based BCI Inference System," Wirel Pers Commun, vol. 116, no. 1, pp. 777–789, Jan. 2021.
[28] H. Song, A. Men, and Z. Jiang, "Breast tumor detection using empirical mode decomposition features," IEEE Access, pp. 1-9, Aug. 2017.
[29] Z. Li, Y. Jiang, Q. Guo, C. Hu, and Z. Peng, "Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving-speed variations," Renew Energy, vol. 116, pp. 55–73, Feb. 2018.
[30] A. Ravi, N. H. Beni, J. Manuel, and N. Jiang, "Comparing user-dependent and user-independent training of CNN for SSVEP BCI," J Neural Eng, vol. 17, no. 2, pp. 1-14, Apr. 2020.
[31] Y. Li and T. Kesavadas, "SSVEP-Based Brain-Computer Interface for Part-Picking Robotic Co-Worker," J Comput Inf Sci Eng, vol. 22, no. 2, pp. 1-8, Apr. 2022.
[32] S. C. Chen, Y. J. Chen, I. A. E. Zaeni, and C. M. Wu, "A Single-Channel SSVEP-Based BCI with a Fuzzy Feature Threshold Algorithm in a Maze Game," International Journal of Fuzzy Systems, vol. 19, no. 2, pp. 553–565, Apr. 2017.
[33] 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, pp. 1-13, 2022.
[34] D. Xu, F. Tang, Y. Li, Q. Zhang, and X. Feng, "An Analysis of Deep Learning Models in SSVEP-Based BCI: A Survey," Brain Sciences, vol. 13, no. 3, pp. 1-23, Mar. 2023.
[35] Jason Forbes, Martin Ordonez, and Matias Anun, "Improving the dynamic response of power factor correctors using simple digital filters: moving average filter comparative evaluation," in Proc. IEEE Int. Conf. Energy Conversion Congress and Exposition, 2013, pp. 4814-4819.
[36] J. Lu, X. Qu, D. Wang, J. Yue, L. Zhu, and G. Li, "Signal filtering method of variational mode decomposition and Euclidean distance based on optimizing parameters of classification particle swarm optimization algorithm," Transactions of the Institute of Measurement and Control, vol. 43, no. 9, pp. 2018–2029, Jun. 2021.
[37] C. Li, H. Deng, S. Yin, C. Wang, and Y. Zhu, "sEMG signal filtering study using synchro squeezing wavelet transform with differential evolution optimized threshold," Results in Engineering, vol. 18, pp. 1-11, Jun. 2023
[38] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," IEEE Transactions on Signal Processing, vol. 62, no. 3, pp. 531–544, Feb. 2014.
[39] K. Cai, W. Cao, Z. Liu, W. Wang, and G. Li, "Application of the Variational Mode Decomposition for Power Quality Analysis," Electric Power Components and Systems, vol. 47, no. 1–2, pp. 43–54, Jan. 2019.
[40] X. Yun and R. Jian, "Features Method for Selecting VMD Parameters based on Spectrum without Modal Overlap," in Journal of Physics, Institute of Physics Publishing, pp. 1-10, Aug. 2020.

[41] T. Liang, H. Lu, and H. Sun, "Application of parameter optimized variational mode decomposition method in fault feature extraction of rolling bearing," Entropy, vol. 23, no. 5, pp. 1-21, May 2021.
[42] H. Li, T. Liu, X. Wu, and Q. Chen, "Application of optimized variational mode decomposition based on kurtosis and resonance frequency in bearing fault feature extraction," Transactions of the Institute of Measurement and Control, vol. 42, no. 3, pp. 518–527, Feb. 2020.
[43] L. Wang, "A new harmonic measurement method based on VMD and STEKF," in Proc. IEEE Int. Conf. Chinese Control and Decision, Jul. 2017, pp. 1917–1922.
[44] L. Guoxin, Z. Tangyu, L. Guoqing, and W. Jian, "Harmonic detection algorithm based on adaptive variational modal decomposition," in Proc. Int. Conf. on Artificial Intelligence and Electromechanical Automation, May 2021, pp. 375–379.
[45] Z. Xing-Li, C. Lian-Yue, C. Yan, J. Rui-Sheng, and L. Xin-Ming, "Microseismic signal denoising by combining variational mode decomposition with permutation entropy," Applied Geophysics, vol. 19, no. 1, pp. 65–80, Mar. 2022.
[46] Y. Huang, H. Bao, and X. Qi, "Seismic random noise attenuation method based on variational mode decomposition and correlation coefficients," Electronics (Switzerland), vol. 7, no. 11, pp. 1-13, Nov. 2018.
[47] H. Mahgoun, R. E. Bekka, and A. Felkaoui, "Gearbox fault diagnosis using ensemble empirical mode decomposition (EEMD) and residual signal," Mechanics and Industry, vol. 13, no. 1, pp. 33–44, 2012.
[48] X. Zhang, Z. Luan, and X. Liu, "Fault diagnosis of rolling bearing based on kurtosis criterion VMD and modulo square threshold," The Journal of Engineering, vol. 2019, no. 23, pp. 8685–8690, Dec. 2019.
[49] Z. Wang, J. Wang, and W. Duv, "Research on fault diagnosis of gearbox with improved variational mode decomposition," Sensors, vol. 18, no. 10, pp. 1-15, Oct. 2018.
[50] M. L. Al-dabag, H. T. S. ALRikabi, and R. R. O. Al-Nima, "Anticipating Atrial Fibrillation Signal Using Efficient Algorithm," International journal of online and biomedical engineering, vol. 17, no. 2, pp. 106–120, 2021.
[51] K. Bouaouiche, Y. Menasria, and D. Khalfa, "Detection of defects in a bearing by analysis of vibration signals," Diagnostyka, vol. 24, no. 2, pp. 1–7, Mar. 2023.
[52] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain-computer interface," Proc Natl Acad Sci U S A, vol. 112, no. 44, pp. 6058–6067, Nov. 2015.
[53] J. C. Yin, Y. Zhu, J. Fei, and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," IEEE Access, vol. 5, pp. 21954–21961, Oct. 2017.
[54] J. R. Wolpaw, N. Birbaumer, D. J. Mcfarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," pp. 767-791, 2002

ADEEL WAHAB received the B.Sc. degree (Hons.) in Mechanical Engineering from the University of Engineering and Technology (UET) Taxila, Pakistan, in 2004, and the M.S. degree in mechatronics engineering from the College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST), Pakistan, in 2013. He is currently pursuing the Ph.D. degree in mechatronics engineering with the National University of Sciences and Technology (NUST), Islamabad, Pakistan. His research interests include artificial intelligence, electroencephalogram (EEG) based brain-computer interface.

UMAR SHAHBAZ KHAN did his bachelors in mechatronics engineering from National University of Sciences and Technology, Pakistan in 2005. He completed his PhD in electrical engineering and electronics from University of Liverpool, UK in 2010. Currently he is working as a Tenured Professor at the department of mechatronics engineering, National University of Sciences and Technology and his research interests include Image processing, embedded systems, prosthesis, Robotics and Automation. He is also the Director of the National Centre of Robotics and Automation a 1.3 billion PSDP Project, besides the centre he has funding of over 40 million from various projects. He has filed 1 trademark, 2 copyrights and 12 design patents out of which 2 have been granted. He is co-founder of two spinoffs and has over 80 publications.

TAHIR NAWAZ has a strong demonstrated track record of research and development in the areas of computer vision (visible/thermal imagery) and artificial intelligence with more than 12 years of experience of working in academic and industrial sectors across multiple European countries in prestigious organizations. Since 2021, he has been working on an Assistant Professor position at the Department of Mechatronics Engineering, College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST), Pakistan, where his core interests focus around multi-modal sensing techniques, particularly investigating cutting-edge technologies pertaining to automated video surveillance and autonomous vehicles. He completed a Ph.D. with a specialization in computer vision, which was a joint Doctorate program under the highly prestigious Erasmus Mundus Fellowship in Queen Mary University of London (UK) and Alpen-Adria University of Klagenfurt (Austria). He also received an M.Sc. in Computer Vision and Robotics under the Erasmus Mundus Scholarship, a joint master program at Heriot-Watt University (UK), University of Girona (Spain) and University of Burgundy (France). In 2005, he also represented Pakistan as a Team Leader in Asia-Pacific Broadcasting Union (ABU) Robocon Contest (an international robot competition), held at Beijing, China. He has published numerous papers in prestigious publication venues and has been involved in several international funded projects.

HASSAN AKBAR received bachelor's degree in Electrical Engineering, specializing in Computer Engineering. Currently, he holds the position of Research Associate at the National Center of Robotics and Automation (NCRA), NUST Pakistan. His research interests primarily lie in the fields of Computer Vision, Deep Learning, and Embedded Systems. He is committed to conducting research and development on advanced technologies, with a focus on creating solutions for real-world problems using AI. Specifically, he emphasizes devising AI solutions tailored for edge devices to address modern challenges effectively.

SYED TAYYAB HUSSAIN SHAH He is a PhD in mathematics and working as associate professor in the Basic Sciences and Humanities Department, National university of sciences and technology, Islamabad, Pakistan. His area of research is mathematical modelling, image processing, data analysis and fluid dynamics**.**

AZFAR KHALID received the B.Sc. degree in Mechanical Engineering from the GIK Institute, Pakistan, in 2000, and the Ph.D. degree in precision engineering and robotics from The University of Manchester, U.K., in 2009. He is currently a Senior Lecturer of industrial robotics and Autonomous Systems with Nottingham Trent University, U.K. Prior to this, he has worked in the Capital University of Science and Technology (CUST) and the National University of Science and Technology (NUST), Islamabad, Pakistan. He has also worked as a Research Associate in the Bremen Institute for Production and Logistics (BIBA), Bremen, Germany, where his research focus was cyber physical system development for human robot collaboration in the context of Industry 4.0. His research interests include precision engineering, robotics, cyber physical systems, and human robot collaboration. He is a Chartered Engineer (CEng), a member of Institute of Mechanical Engineers (MIMechE), U.K., and Fellow of Higher Education Academy (FHEA).

ALI R ANSAI has an academic with over 30 years-experience in teaching and research in the area of applied mathematics. He was the Dean of the College of Arts and Sciences, Gulf University for Science and Technology, Kuwait, and served for ten years. He is currently a Professor of applied mathematics and Head of Accreditation and Quality Assurance. He has more than 100 publications in the area of applied mathematics.

RAHEEL NAWAZ is an international thought leader in higher education and industry-academia co-design, who has advised national governments, policy organizations, and parliamentary committees on AI, Digital Education and Work-based Learning. At Staffordshire University he leads transformative action across the university to deliver enhanced academic experience.

With 4 books and over 200 research articles, he ranks among the top-10 most cited scholars in the world in the fields of Applied AI, Digital Education, Educational Data Science, and Digital Transformations.

His sector-leading work on pedagogical design and delivery led to outstanding student outcomes, with first generation university students with minimal prior attainment outperforming graduates of top UK universities (including Oxbridge) in both average salaries and job security; he was awarded National Teaching Fellowship (NTF) for this work. He has advised on the establishment and launch of work-based learning degree programs across three continents and received Principal Fellowship of the Higher Education Academy (PFHEA) for this work.

The impact of his academic leadership has been endorsed by heads of state, government ministers, and industry leaders. He passionately believes that kindness in the most impactful leadership trait and being an intrinsic giver in life is the most sustainable route to happiness, as outlined in his TEDx talk**.**