Interdisciplinary Journal of Computing & AI

IJCAI

RESEARCH

Hybrid Deep Learning and Frequency-Domain Approaches for Robust SSVEP-Based BCI Speller Systems

Srinivas Rao Gorre[1*], Radha Krishna J[2], Vardhan Manjunath B[3], Nishanth Reddy S[4],

[1,2,3,4]*Electronics and Communication Engineering, Vasavi College of Engineering, Hyderabad, India*

[1]*gsrinivasrao@staff.vce.ac.in*; [2]*radhakrishnaj454@gmail.com*; [3]*vardhan.biji@yahoo.co.uk*; [4]*nishanthreddysunkireddy@gmail.com*

ARTICLE INFO.

Keywords:
Steady-State Visual Evoked Potential (SSVEP), Brain–Computer Interface (BCI), Electroencephalography (EEG), Filter Bank Canonical Correlation Analysis (FBCCA), Convolutional Neural Networks (CNN), EEGNet, Hybrid Models, Two-Stage Classification, Domain Adaptation, Temporal Feature Extraction, BCI Speller Systems, Signal Processing, Deep Learning, Human–Computer Interaction
∗Corresponding author email address: gsrinivasrao@staff.vce.ac.in
Manuscript received: 5 May 2026/
Accepted: 30 May 2026
Published online: 10 June 2026
https://doi.org/10.5281/zenodo.20620646
License: Creative Commons Attribution 4.0 International

Published by Aspiration Publishing Trust (DARPAN ID: WB/2025/0887371) | IJCAI.
©Author(s) 2026.

ABSTRACT

Brain-Computer Interfaces (BCIs) have revolutionized the way we directly interface human brain with machines, offering new hope for people with motor disabilities. Steady-State Visual Evoked Potential (SSVEP)-based BCIs have been increasingly attractive due to their high signal-to-noise ratio, negligible training requirements and high information transfer rates. But classification of SSVEP signals is difficult due to noise, individual differences, and the overlap of harmonics. In this paper, we propose a multi-model hybrid approach for classifying SSVEPs in BCI spellers by combining classic frequency-based signal processing models with deep learning models. The proposed framework, in contrast to the traditional approaches that employ single classification models, explores and integrates a range of methodologies, such as Filter Bank Canonical Correlation Analysis (FBCCA), Depthwise EEGNet, subject-specific Convolutional Neural Networks (CNN), fine-tuned hybrid models, and windowed hybrid classifiers. Moreover, a two-stage hierarchical classification approach is proposed to enhance classification performance and robustness. The

Cite as: S. R, Gorre et al., 'Hybrid Deep Learning and Frequency-Domain Approaches for Robust SSVEP-Based BCI Speller Systems', Interdisciplinary Journal of Computing & AI (IJCAI), vol. 1, no. 2. Aspiration Publishing Trust (DARPAN ID: WB/2025/0887371), p. 177-211, Jun. 2026. doi:10.5281/zenodo.20620646.

proposed framework is tested on multi-subject data sets, showing the hybrid methods' abilities to learn both the spectral and spatial-temporal characteristics of the EEG. The findings show the benefits of using deterministic signal processing techniques along with learning-based models to enhance the robustness and transferability of SSVEP classification. In summary, the research lays down the foundation for investigating and improving different classification approaches for SSVEP, which can be used for designing robust and scalable BCI spellers.

1 Introduction

1.1 Background of Brain–Computer Interface Systems

Brain–Computer Interfaces are devices that connect neural activity directly with external technologies or devices, and do not depend on physical movement. BCIs are devices that can convert brain signals into actionable commands, providing a new way for people with severe motor impairments, such as amyotrophic lateral sclerosis (ALS), spinal cord injury and stroke, to communicate [1]. Due to its non-invasive nature, portability, less expensive than most other neurophysiological recording techniques, and its comparatively high temporal resolution, electroencephalography (EEG) is widely used. The basic components of a typical BCI system include signal acquisition, signal preprocessing, feature extraction, classification and decision generation. Although there are benefits to analysis via EEG, there are many challenges due to inter-subject variability, environmental noise and physiological artifacts. Therefore, it is important for classification methods to be able to detect informative representations while being robust to differences in recording conditions.

1.2 Overview of SSVEP-Based Speller Systems

The steady-State Visual Evoked Potential based BCIs are one of the most reliable paradigms for real-time brain–computer interaction, due to a high signal-to-noise ratio, relatively short calibration time and favourable information transfer characteristics [2]. When the user is visually fixated on a repetitive stimulus, presented at a specific frequency of flicker, SSVEP responses occur at the flicker frequency and at harmonic frequencies. Responses in these parts of the brain are mainly seen in the occipital region and are recorded from sites like O1, Oz, and O2. Within SSVEP speller systems, the visual target is actually encoded in terms of different flickering frequencies, which can be used to determine which target is intended by using the spectral properties of the EEG recordings. These systems are typically superior in speed of communication over a number of other paradigms, though the classification problem becomes more difficult as more target frequencies are added. The close frequencies, harmonic interference, recording noise and inter-subject variability can cause confusion in frequency recognition and provide motivation for using more robust methods of signal analysis [3].

1.3 Limitations of Conventional Methods

Classical approaches to SSVEP classification, like Canonical Correlation Analysis (CCA), involve calculating the correlation between EEG signals and reference sinusoidal signals at different frequencies corresponding to the stimulus frequencies [4]. While CCA is fast and does not require training, it has some drawbacks, such as being vulnerable to noise and unable to model non-linear interactions in EEG signals. To overcome these issues, Filter Bank Canonical Correlation Analysis (FBCCA) was proposed to divide the EEG signals into several sub-bands and perform CCA in each sub-band to enhance frequency estimation [5].

1.4 Motivation for Multi-Model and Hybrid Approaches

The conventional frequency-domain-based approach and learning-based approach have complementary benefits for SSVEP recognition. While signal processing methods like FBCCA offer reliable and efficient frequency discrimination and low computation, deep learning models can directly learn the spatial and temporal representation of signals from EEG recordings. Therefore, a combination of these approaches could yield a balanced framework, which combines deterministic frequency analysis and adaptive feature learning. Some important features of the present invention are also temporal segmentation and hierarchical decision strategies, which can enhance the reliability of the classification in practical applications by reflecting the changes in EEG responses over time and removing ambiguity between similar stimulus classes. These observations motivate the investigation of integrated classification strategies for enhancing robustness and cross-subject applicability in the SSVEP-based BCI system in an empirical way.

1.5 Contributions of the Present Work

This study provides a single experimental setup to investigate various classification approaches to SSVEP- based BCI speller systems in the same experimental environment. The work explores hybrid architectures that effectively combine FBCCA and CNN as feature extraction methods with frequency-domain signal processing methods, as well as the integration of frequency-domain signal processing and deep learning techniques. Furthermore, temporal windowing, subject-specific fine-tuning, and domain adaptation approaches are tested, and their effect on classification consistency among subjects is assessed.

Validation of performance is done using cross-trial validation and multiple evaluation measures, such as accuracy of classification, confusion matrix, and receiver operating characteristic (ROC) analysis. To this end, the study offers an empirical and comparative analysis of integrative methods for reliable SSVEP recognition, instead of proposing a new categorization scheme. Figure 1 shows the block diagram of the proposed SSVEP-BCI system.

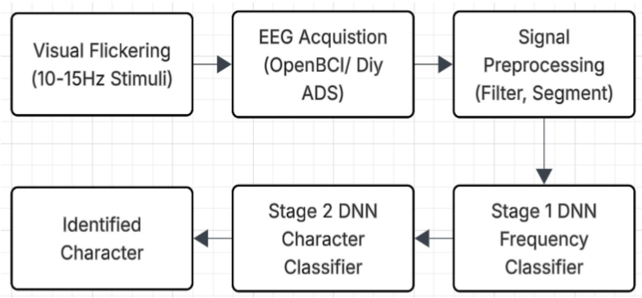

Fig. 1 Block Diagram of the Proposed SSVEP-BCI System

2 Literature Survey

This section reviews representative approaches for SSVEP classification, including conventional signal-processing techniques, deep learning models, hybrid architectures, and domain adaptation strategies relevant to EEG-based BCIs.

2.1 Classical Methods for SSVEP Classification

But CCA is prone to noise and does not take into account harmonics. In order to address these issues, FBCCA method has been introduced, which includes partitioning the EEG signals into bands and performing CCA on these bands. This approach enhances classification performance by considering both fundamental and harmonic frequencies [2]. Previous studies have demonstrated that FBCCA has a better accuracy than CCA, particularly when multiple frequencies are present [4]. Further improvements have been proposed, including Task-Related Component Analysis (TRCA) to maximise the trial-to-trial consistency of the EEG response [3].

2.2 Deep Learning Approaches for EEG and SSVEP

Traditional approaches for SSVEP classification typically use statistical signal processing techniques, based on the frequency information of the EEG signals.

CCA has been widely used because it is easy and does not require training. CCA finds the correlation between the multi-channel signals and a set of reference sinusoidal signals of different frequencies and enables rapid identification of the stimulus frequency [1].

Due to their ability to learn informative signal representations directly from raw or minimally processed EEG signals, deep learning methods have been increasingly investigated in the field of EEG analysis. Among all those techniques, Convolutional Neural Networks (CNNs) are chosen for the classification of SSVEPs because they have the ability to learn spatial and temporal patterns from multichannel EEG signals [4]. EEGNet is one of the most popular lightweight network structures used in classification applications with EEG. EEGNet offers an efficient learning framework for representing signals in the brain with depthwise and separable convolutions, which has relatively low computational complexity [5]. The model's small architecture makes it suitable for use in real-time inference and limited computational resources. Recently, attention-based models, like transformer models, have been explored to enhance the modelling of long-range dependencies in EEG signals. There have been some approaches that have shown promising results across subjects, like SSVEP former [6]; but the models typically need more data and more computing power to be trained well. Recurrent architectures, such as Long Short-Term Memory (LSTM) networks, have also been explored for modelling temporal relationships in EEG time series [7]. While they are flexible to represent, they might not be applicable in real time due to training complexity and computational cost. Thus, despite having great representational power, deep learning models have their limitations when it comes to subject variability, overfitting, and data requirements [8].

Other signal-processing techniques have been investigated as well for improving the reliability of the classification, such as Task-Related Component Analysis (TRCA), which aims to increase the consistency of each trial and the representativeness of the signals [3]. These approaches show promise for practical applications, but the ability to preserve the more intricate non-linear relationships in EEG data is still restricted.

2.3 Hybrid Models Combining Signal Processing and Deep Learning

The hybrid methods are proposed as a practical direction in the classification of SSVEP, combining the methods of deterministic signal processing and data-driven learning models. The methods don't only use handcrafted frequency representations or learned representations, but they try to integrate complementary information derived from both domains. Various research has shown that the combination of FBCCA and CNN-based architectures showed better classification performance, with frequency-domain analysis giving robust frequency information and CNN learning the spatial–temporal information of EEG recordings. Such integration can help to increase robustness in noisy and inter-subject datasets. Combinations of signal processing and neural architectures have also been investigated to enhance the reliability of classifications, such as in other approaches using ensemble frameworks and sub-band learning strategies [9]. Temporal segmentation methods, which segment the EEG recording into smaller time windows, have also been explored to better capture transient changes in the signal and to have more stable predictions under practical recording conditions. Although promising outcomes, hybrid systems tend to be more complex in terms of the models and the need to balance the trade-off between computational efficiency and the performance of the classification model. The observations serve as a guide to further research on strong SSVEP recognition integration.

2.4 Domain Adaptation Techniques in BCI Systems

In practical BCI, however, the inter-subject variability of EEG recordings is a major obstacle, making it difficult to implement models for generalizability with subject independence. This has motivated the study of domain adaptation and transfer learning methods to minimize the amount of calibration needed while maintaining classification accuracy. Recently, source-free domain adaptation techniques have received much attention due to the ability of adapting a model to a target subject using unlabeled target-subject data without accessing the original training set [10]. The methods generally aim to match feature distributions between the source and target domains, to enhance generalization to unseen subjects. Additionally, transfer learning techniques have been explored to adapt to specific subjects with small amounts of specific data, which applies to the fine-tuning of pre-trained models and can greatly reduce the amount of training required and enhance the personalization [11]. In practice, however, the non-stationarity of EEG signals and signal variability among different subjects remain significant challenges. Therefore, the problem of cross-subject generalization is still an open challenge in SSVEP-based BCIs.

2.5 Identified Research Gaps

The literature reviewed suggests that traditional signal processing approaches are still computationally efficient, but are less flexible in modelling complex EEG variability. On the other hand, deep learning techniques exhibit excellent representational power but can be highly data and domain-dependent. Hybrid methods have yielded promising results when combining different classes of signals, but inconsistencies in evaluation methods, the complexity of the models and the results across subjects make it difficult to compare results across studies. In addition, temporal segmentation, score fusion and adaptation strategies have rarely been explored in a shared experimental environment. The observations inspire a systematic empirical investigation of integrative classification strategies on the basis of a consistent validation scenario to gain insights into their performance in SSVEP-based BCI applications.

3 Implementation

3.1 Overall System Architecture

This proposed framework is intended as a modular testing bed for both traditional signal processing-based and learning-based BCI classification via SSVEP. Implementation involves incorporating frequency-domain methods, CNN architectures, hybrid fusion methods and domain adaptation methods, and integrating all in one experimental environment. This organization allows comparison of various classification methods under the same preprocessing and validation procedure.

It is implemented in MATLAB with the Signal Processing and Deep Learning toolboxes. The Processing Pipeline consists of acquiring EEG signals, preprocessing the data, extracting features, training the model, making predictions, and assessing the model performance. Special attention is given to the uniformity of the experiments by having all models run on the same pre-processed dataset and use the same validation protocols. For reliable evaluation, predictions of every fold and subject are combined and global performance measures like classification accuracy, confusion matrices, ROC curves, and AUC values are computed. This unified framework allows the systematic analysis of the relative strengths and weaknesses of the various SSVEP classification strategies [12], [13].

3.2 Dataset Description and Experimental Configuration

The experimental evaluation is carried out on the publicly available BETA SSVEP dataset featuring EEG recordings from 70 subjects engaged in visual stimulation tasks for 40 target frequencies. The target stimuli are repeated four times each (160 trials per subject). The recordings are made with 64 channels at a sampling rate of 250 Hz. Stimulus frequencies are in the range of about 8 Hz to 15.8 Hz, and the frequency span between each pair of stimuli is fixed, with the aim of reducing the overlap in the spectrum. In each trial, subjects look at a flickering target and generate SSVEP responses to this target at the frequency of the stimulus flicker and its harmonics. Visual response latency is addressed with a temporal offset before the feature extraction process so that the analyzed EEG segment will mostly consist of steady-state activity instead of onset-related activity. The EEG signals are reorganized to form trial-wise representations appropriate for both traditional signal processing methods and deep learning representations, to achieve computational efficiency.

3.3 Cross-Trial Validation Strategy

Crucially, the implementation uses a cross-trial validation strategy, enabling us to evaluate model performance. We use a deterministic fold selection, instead of a random partition, where each trial has a deterministic allocation to the training or test sets [14].

We divide the 4 trials for each of the 40 stimulus frequencies for each participant into 4 folds. We select a single trial for each frequency to be used for testing in each fold and use the remaining three trials for training. We then obtain 4 folds, where each trial is tested only once.

Mathematically, for a given class c, the trial indices are defined as:

$$\text{Test Index} = (c-1) \times 4 + f \ldots\ldots\ldots\ldots\ldots\ldots\ldots.\ldots.(1)$$
$$\text{Train Index} = (c-1) \times 4 + \{1,2,3,4\} \setminus f \ldots\ldots\ldots.(2)$$

where $f \in \{1,2,3,4\}$represents the fold number.

This structured splitting ensures that:

- All classes are equally represented in both training and testing sets
- No data leakage occurs between training and testing
- Performance metrics are consistent across folds

We then compute the accuracy and the AUC for each subject as the average of all folds. The metrics for all subjects are then averaged to obtain the global metrics. A structured cross-trial validation strategy was adopted to ensure fair comparison among all evaluated models. Similar validation methodologies have been employed in previous EEG classification studies [15], [16].

3.4 Channel Selection and Signal Representation

Channel selection is important in SSVEP-based systems since the responses to the visual stimuli are mainly generated by the visual cortex. In our design, different models are designed to select different channels. For example, in classical and single-channel models, we select the Oz channel for its responsiveness to visual stimuli. This makes the model simpler and faster to train, but yields valuable information. For multiple channels, such as the Depthwise EEGNet hybrid, three occipital channels (O1, Oz and O2) are chosen. These channels incorporate spatial information into the EEG data and enable the model to learn spatial features using depthwise convolutions. When performing domain adaptation, more channels are selected, including multiple occipital and parietal channels. This larger set of channels allows for better feature representation and generalisation. The EEG signals are reshaped into different formats depending on the model:

- Single-channel models: $1 \times T \times N \ldots\ldots\ldots\ldots...(3)$
- Multi-channel models: $C \times T \times N \ldots\ldots\ldots.\ldots...(4)$
- Deep learning models: reshaped into 4D tensors for convolutional processing

This enables the model to be compatible with conventional and deep learning models with scalability.

3.5 Signal Preprocessing

EEG signals are subject to various types of noise, such as physiological and environmental artifacts. To combat this problem, a complex preprocessing pipeline is implemented.

3.5.1 Band-Pass Filtering

The band-pass filter is used to extract the frequencies of interest in SSVEP signals. In particular, a fourth-order infinite impulse response (IIR) filter is used with cutoff frequencies at 6 Hz and 90 Hz. This allows the pass-through of the fundamental frequencies of the stimuli and their harmonics, while low-frequency drift and high-frequency noise are eliminated.

Zero-phase forward-backwards filtering is used to apply the filter, preventing phase distortion. In the case of multiple channels, each channel is filtered separately [17].

3.5.2 Filter Bank Processing (Domain Adaptation)

In the domain adaptation approach, a filter bank is used to divide the EEG signals into several sub-bands. The sub-bands are filtered with Chebyshev Type-I band-pass filters, each with varying lower frequencies. The sub-band signals are then used to extract frequency-specific features, which are integrated to improve classification results. Filter banks help in extracting both fundamental and harmonic frequencies.

3.5.3 Temporal Segmentation (Windowed Hybrid Model)

The windowed hybrid model segments the EEG signals into temporal windows to track changes in SSVEP responses. The length of each window is 1.5 seconds with a shift of 0.5 seconds. This means that each trial is broken down into multiple segments, which are then independently analysed. This overlapping approach provides smooth feature extraction and helps to filter out any noise.

The segmentation process can be expressed as:

$$\text{Start Points} = 1,\ 1 + \Delta,\ 1 + 2\Delta, \ldots \qquad \ldots\ldots\ldots.(5)$$

where $\Delta$is the step size.

These are then processed by the CNN to extract features, which are combined to make the final classification.

3.5.4 Normalization

Features or scores are normalised before classification to provide comparable scales. Outputs of FBCCA and CNN are normalised using min-max normalisation:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad \dots\dots\dots.(6)$$

This avoids bias in hybrid models, which use fused scores.

3.6 FBCCA Baseline Method

Our baseline classification method is the Filter Bank Canonical Correlation Analysis method because it performs well in frequency recognition, and does not require any training. FBCCA does not involve learning of parameters, and is based on statistical correlation between EEG signals and sinusoidal templates. For the current implementation, the pre-processed EEG signal $X \in \mathbb{R}^{1\times T}$ of the Oz channel is used. It is first bandpass filtered and then converted into trial-wise format. For each trial, FBCCA calculates the correlation with a series of reference signals for all stimulus frequencies. The reference signals are generated by taking a number of harmonics of the stimulus frequencies. For a particular frequency f_i, the reference matrix is:

$$Y_i = \begin{bmatrix} \sin(2\pi f_i t) \\ \cos(2\pi f_i t) \\ \sin(4\pi f_i t) \\ \cos(4\pi f_i t) \\ \vdots \\ \sin(2N\pi f_i t) \\ \cos(2N\pi f_i t) \end{bmatrix} \quad \dots\dots\dots.(7)$$

with N being the number of harmonics (set to 5 in this case). This allows the model to learn both the fundamental and higher harmonics of the SSVEP signals [18], [19].

The EEG signal is also divided into sub-bands via a filter bank. For each sub-band j, canonical correlation analysis (CCA) is applied between the EEG signal and the reference signals. The resulting correlation coefficients $\rho_{i,j}$ are combined using weighted summation:

$$S_i = \sum_{j=1}^{M} w_j \cdot \rho_{i,j}^2 \quad \dots\dots\dots.(8)$$

where M is the number of sub-bands and w_j are sub-band weights computed using a power-law function:

$$w_j = j^{-a} + b \quad \ldots\ldots\ldots..(9)$$

with empirically chosen parameters $a = 1.25$and $b = 0.25$.

The final prediction is obtained by selecting the class with the maximum score:

$$\hat{y} = \arg\max_i S_i \quad \ldots\ldots\ldots.(10)$$

This approach harnesses the information in the frequency domain as well as the harmonics, making FBCCA a robust baseline for SSVEP classification. But it suffers from being unable to capture non-linear relationships in EEG signals and its sensitivity to noise and inter-subject variations [20], [21].

3.7 Depthwise EEGNet–FBCCA Hybrid Model

The depth wise EEGNet-FBCCA hybrid model combines deep learning for feature extraction with FBCCA. It takes multi-channel EEG signals from occipital electrodes (O1, Oz, O2), enabling it to account for spatial variability. The EEG signals are first filtered and reshaped into a 4D tensor of the form:

$$X \in \mathbb{R}^{C\times T\times 1\times N}$$

where $C = 3$ represents the number of channels, $T$is the number of time samples, and $N$is the number of trials.

The CNN model is similar to EEGNet. The network is comprised of spatial and temporal convolutional layers, which are applied separately to each channel. Depthwise convolution allows efficient extraction of spatial features by convolving each channel using different filters, followed by pointwise convolution to merge the features.

The network is trained using Adam optimiser (Learning rate - 10^{-3} and several epochs to ensure convergence. The network's output is a distribution over all classes:

$$P_{CNN}(\, y \mid X\,)$$

Simultaneously, FBCCA is applied to the same EEG signal to obtain frequency-domain scores:

$$S_{FBCCA}$$

Both outputs are normalised using min–max scaling and combined using weighted fusion:

$$S_{final} = \alpha \cdot S_{FBCCA}^{norm} + \beta \cdot P_{CNN}^{norm} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots.(11)$$

where $\alpha = 0.4$and $\beta = 0.6$ are empirically chosen weights that balance the contributions of frequency-domain and learned features.

The prediction is made by choosing the class with the highest fused score. This combination of methods enhances classification accuracy by merging the deterministic frequency detection of the DLT with the adaptability of feature learning in the CNN, thus increasing tolerance to noise and variability [22], [23].

3.8 Subject-Specific CNN Model

The subject-specific CNN model aims to capture subject-specific changes in the EEG by training a neural network model for each subject. This model uses only deep learning, without frequency domain features.
The input EEG signal is extracted from the Oz channel and reshaped into a 4D tensor:

$$X \in \mathbb{R}^{1\times T\times 1\times N}$$

The CNN model comprises a convolutional layer with a kernel size equal to the length of the signal, followed by batch normalization and a rectified linear unit (ReLU) activation. It's then followed by a dense layer to output class probabilities.
The model is trained using categorical cross-entropy loss and optimised using the Adam optimiser. During inference, the network produces class probabilities:

$$P_{CNN}(y \mid X)$$

and the predicted class is obtained using the argmax function.

While this approach captures subject-specific features, it fails to generalise across subjects and is data-intensive. The lack of frequency information also makes it difficult to classify different stimulus frequencies [24].

3.9 Fine-Tuned Hybrid Model (CNN + FBCCA)

The fine-tuned hybrid model utilises the ability of CNN to learn features and FBCCA to analyse the frequency domain. In this case, the CNN learns from raw EEG signals to extract spatial–temporal features, and the FBCCA complements this with frequency information.
The CNN is similar to the subject-specific model, but fine-tuned with subject-specific data. The signal is reshaped to a 4D tensor, and then it is convolved using the convolutional layers, batch normalisation, and activation functions.
The CNN outputs class probabilities:

$$P_{CNN}(y \mid X)$$

which are combined with FBCCA scores using weighted fusion:

$$S_{final} = \alpha \cdot S_{FBCCA}^{norm} + \beta \cdot P_{CNN}^{norm}$$

where $\alpha = 0.7$and $\beta = 0.3$.

The larger weight for FBCCA is due to the relevance of the features extracted from the frequency domain, and the CNN provides complementary information. This enables the model to attain high performance with good generalisation across subjects. This fine-tuning allows the model to learn subject-specific details, enhancing its performance without additional time-consuming training. This is a viable balance between generalisation and individualisation [25].

3.10 Windowed Hybrid SSVEP Classifier (FBCCA + Temporal CNN)

The windowed hybrid model incorporates temporal windows as an additional feature in the EEG signal. Rather than analysing the whole EEG signal at once, it is broken down into a number of time windows. Each window is processed independently by a CNN, which produces a probability distribution over all classes:

$$P_{CNN}^{(k)}(y \mid X_k)$$

where $k$denotes the window index.

The final CNN score for a trial is obtained by averaging the outputs across all windows:

$$P_{CNN}(y \mid X) = \frac{1}{K}\sum_{k=1}^{K} P_{CNN}^{(k)}(y \mid X_k) \dots\dots\dots (12)$$

where $K$is the number of windows.
Simultaneously, FBCCA is applied to the full signal to obtain frequency-domain scores. The final prediction is obtained using weighted fusion:

$$S_{final} = \alpha \cdot S_{FBCCA}^{norm} + \beta \cdot P_{CNN}^{norm}$$

with $\alpha = 0.75$and $\beta = 0.25$.

Temporal segmentation enables the model-to-model temporal dynamics in EEG signals and helps to minimise the effect of noise in different segments. This leads to better stability and robustness especially in online applications [26].

3.11 Source-Free Domain Adaptation Framework

Variations in EEG signals among individuals are one of the most important issues in SSVEP-based BCIs. BCI models developed using one group of subjects usually do not work well with a new group of subjects due to variations in brain responses, head shape, physiological states, etc. To overcome this problem, a source-free domain adaptation approach is adopted, where the model can adapt to new subjects without any labels from the target domain. A two-step training procedure is adopted, which involves pre-training and subject adaptation. During the pre-training phase, a deep neural network is trained with data from multiple subjects, leaving one subject out. For each subject, the network is trained on the data of all other subjects, leaving out the target subject's data. In this way, a collection of pre-trained models is obtained, which are cross-subject generalised.

The input to the network is a multi-channel EEG tensor obtained after filter bank preprocessing. The preprocessing stage decomposes the EEG signal into multiple sub-bands, allowing the model to capture frequency-specific features across different spectral ranges. The resulting data representation has the form:

$$X \in \mathbb{R}^{C \times T \times S \times N}$$

where Crepresents the number of channels, $T$is the number of time samples, $S$is the number of sub-bands, and $N$is the number of trials.

The neural network model has several convolutional layers to extract both spatial and spectral information. The initial layer transforms the channels, followed by convolutional layers that combine information from different channels and sub-bands. Dropout layers are used to avoid overfitting, and a fully connected layer is used to transform the extracted features to the class probabilities. Following pre-training, the model is adapted in a source-free manner for each subject. During this step, only the unlabelled data of the target subject is accessible. To adapt the model in a label-free manner, pseudo-labels are used. We make preliminary predictions with both FBCCA and the pre-trained neural network:

$$\hat{y}_{FBCCA}, \hat{y}_{DNN}$$

A consensus strategy is then used, comparing the predictions. If the predictions are consistent, it is deemed reliable and is used as a pseudo-label. If they don't, the FBCCA prediction is used as a backup since it is always stable. To enhance stability, a confidence filtering technique is proposed. The model provides confidence scores for each prediction, and only those with confidence scores exceeding a certain threshold are used to train the model. This avoids the use of incorrect p.

This adaptation is done in a cyclic manner where the model is trained with pseudo-labelled data. This results in improved predictions, and uncertain predictions are substituted with FBCCA predictions. Furthermore, a collapse detection method is used to avoid model collapse. When the number of predicted classes is less than a threshold, FBCCA predictions are used instead.

This iterative learning approach allows the model to progressively learn the distribution of features of the subject of interest, leading to better classification performance without the need for labels. The reliance on FBCCA for guidance provides stability, and the neural network learns the subject's characteristics. This approach leads to a stable and scalable approach for cross-subject classification of SSVEP [27], [28].

3.12 Enhanced Source-Free Domain Adaptation with Multi-Sub Band Fusion

The source-free domain adaptation approach is extended with the addition of a multi-sub band filter bank and a non-linear fusion rule with confidence weighting. The goal of this enhancement is to enhance the ability of SSVEP-based BCIs to achieve good generalisation and classification. Here, the EEG is divided into five sub-bands to capture both fundamental and harmonics of SSVEP responses. The selected frequency ranges are 6–14 Hz, 14–22 Hz, 22–30 Hz, 30–45 Hz, and 45–70 Hz. This more detailed spectral division offers better frequency resolution than typical filter bank methods. The EEG signals are first pre-processed with a notch filter to remove 50 Hz power line interference and then normalised to have the same magnitude across channels. Normalisation is carried out as:

$$X_{norm} = \frac{X - \mu}{\sigma + \epsilon} \quad \ldots\ldots\ldots(13)$$

where $\mu$and σrepresent the mean and standard deviation of the signal, respectively.

The model is initially trained using a leave-one-subject-out strategy, ensuring that the target subject’s data is not included during pre-training. For adaptation, a small calibration subset of labelled samples is used to fine-tune the network parameters, enabling rapid personalisation. During inference, predictions are obtained from both the deep neural network and the FBCCA method. The confidence of the neural network prediction is defined as the maximum class probability:

$$conf = \max_{c} P_{CNN}(c) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(14)$$

To ensure stability, the confidence value is constrained within a predefined range. A non-linear fusion weight is then computed as:

$$\alpha = conf^{1.5} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(15)$$

The final classification score is obtained by combining the normalised outputs of the CNN and FBCCA as follows:

$$S_{final} = \alpha \cdot S_{CNN} + (1 - \alpha) \cdot S_{FBCCA} \quad \ldots\ldots\ldots(16)$$

This adaptive fusion mechanism allows the model to rely more on the neural network when confidence is high, while defaulting to the more stable FBCCA predictions when uncertainty is present. The final predicted class is determined using the maximum score criterion:

$$\hat{y} = \arg\max_{c} S_{final}(c) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (17)$$

The use of multi-sub band filtering, adaptive recalibration, and non-linear fusion transforms the system to be more robust and allows for better generalisation across subjects without access to source domain data.

3.13 Score Fusion and Decision Strategy

To improve the system's performance, model fusion is a crucial part of the proposed approach. Rather than using a single model to classify, multiple models are combined for better resilience and performance. For hybrid models, fusion is at the score level. The prediction scores of FBCCA and CNN are first scaled to the same range using the min-max scaling:

$$S_{norm} = \frac{S - S_{min}}{S_{max} - S_{min}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (18)$$

The normalised scores are then combined using weighted summation:

$$S_{final} = \alpha \cdot S_{FBCCA}^{norm} + \beta \cdot S_{CNN}^{norm}$$

The values of α and β are selected empirically based on the models. For example, greater weight is given to the FBCCA in cases where frequency information predominates, whereas greater weight is given to CNN in models that model complex spatial–temporal information.

The final prediction is obtained using the maximum score criterion:

$$\hat{y} = \arg\max_{i} S_{final}(i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (19)$$

This approach leads to the combination of complementary information from models, leading to better classification.

3.14 Performance Evaluation and Metrics Computation

We assess the performance of the proposed system using various metrics such as classification accuracy, confusion matrix and ROC curves.

3.14.1 Accuracy Calculation

The accuracy for each subject and fold is calculated as:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Predictions}} \ldots\ldots\ldots.(20)$$

The overall accuracy for each subject is the average accuracy over all folds and the accuracy of the system is the average accuracy over all subjects.

3.14.2 Confusion Matrix

The global confusion matrix is created by pooling together the predictions and labels of all subjects and folds. This matrix gives an insight into the class-level performance, including correct and incorrect predictions. Row and column normalised summaries aid in examining the distribution of predictions and detect biases.

3.14.3 ROC Curve and AUC

ROC curves are calculated to assess the classification accuracy of the models, with a one-vs-all approach. For each class, binary labels are created and the true positive rate (TPR) and false positive rate (FPR) are calculated for various thresholds.
The ROC curve is interpolated to a fixed grid to ensure the ROC curve is comparable across different classes, and is averaged across all classes to give the mean ROC curve. The Area Under the Curve (AUC) is used as a performance measure:

$$\text{AUC} = \int_0^1 TPR(FPR)\, dFPR \ \ldots\ldots\ldots\ldots..\ldots\ldots\ldots.(21)$$

This is a reliable classification performance metric, especially in multi-class classification tasks.

3.15 Complete System Workflow

Our system adopts a workflow where the raw EEG signals are gradually converted to the final classification result. The proposed system is implemented on the BETA data, which includes 70 subjects' EEG data, 40 classes of stimulus, and 4 trials of each class. It is reshaped to a trial-wise format, and labels are defined. A cross-trial validation scheme is adopted, where one trial of each class is set aside for testing, while the rest are used for training, for each subject. This is done in four iterations to ensure that all trials are tested and class proportions are balanced. The signals are pre-processed with band-pass filtering to include the desired frequency bands. Filter bank decomposition is also used for models that incorporate multiple bands. For the windowed hybrid model, the signals are also divided into non-overlapping temporal windows and processed separately.

Different approaches are taken for feature extraction and classification. FBCCA provides correlation scores with template signals, while CNN-based approaches provide probabilities for a given class from resampled EEG signals. In the case of hybrid approaches, both are computed. The scores are min-max normalised before fusion. FBCCA and CNN scores are weighted and summed to produce the final score and the maximum score criterion is used to determine the prediction.

The overall performance is assessed by combining predictions of all subjects and folds. The accuracy is evaluated as the percentage of correct predictions, and confusion matrices and ROC curves are used to assess the performance. In domain adaptation, an extra refinement step is involved, where pseudo-labelling and confidence filtering is applied to adapt to new subjects. In general, the system has a uniform and modular framework, which ensures the reproducibility of the results and allows us to compare different classification techniques

4 Results and Discussion

4.1 Experimental Setup and Evaluation Metrics

To assess the performance of the proposed SSVEP-based BCI system, we use a multi-subject EEG dataset that includes data from 70 subjects. This includes multi-channel EEG recordings for multiple stimulus frequencies, allowing for a thorough assessment of both individual and across-subject models. The models are trained and tested using the same EEG preprocessing procedures, such as band-pass filtering (8-30 Hz), segmentation, and normalisation. The features are extracted using Fast Fourier Transform (FFT) and filter banks, depending on the type of classification model. Model performance is assessed through various measures such as classification accuracy, ROC curves, Area Under the Curve (AUC) and confusion matrices. Accuracy gives a general indication of model performance, and ROC curves assess the discriminative power of the model. High AUC values (close to 1) signify good discriminative performance. Confusion matrices are employed to display the performance for each class, and to detect error patterns. Finally, we study the subject-wise accuracy to test the generalization ability of the model. This is crucial in SSVEP applications, as the performance varies between subjects.

4.2 FBCCA Baseline Performance Analysis

The FBCCA method is used as the baseline model for all other methods. As a non-trained method, FBCCA is based on a correlation in the frequency domain between the EEG and sinusoidal templates. The confusion matrix for the

FBCCA model is relatively diagonally dominant, which shows that the model is capable of correctly classifying a good share of the desired frequencies (Figure 2). But there are also significant elements off the diagonal, indicating confusion between nearby frequencies. This is possibly due to frequency aliasing and noise, which affect the correlation-based classification.

The FBCCA model's ROC curve shows a fairly high level of classification accuracy with an average AUC of 0.88 (Figure 3). The ROC curve indicates that the model has a good trade-off between true positive rate and false positive rate at various thresholds. But the ROC curve is not as steep as hybrid models, suggesting that the classes are not as well separated. Based on the experimental results, the average classification accuracy of FBCCA is around 48%. Although this accuracy is on par with other training-free methods, it is much lower than hybrid models and deep learning methods. The main drawback of FBCCA is its failure to capture subject-specific characteristics and non-linearity in the EEG signals. Despite its shortcomings, FBCCA offers a robust baseline performance thanks to its simplicity, efficiency and frequency-based features extraction. It also serves as a key component in hybrid models, aiding in feature extraction and frequency estimation.

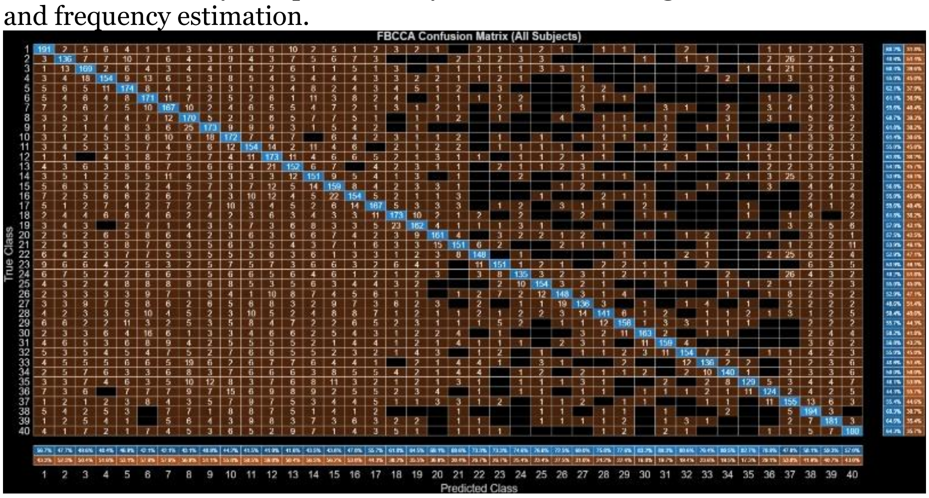

Fig. 2 Confusion Matrix of the FBCCA Baseline Model

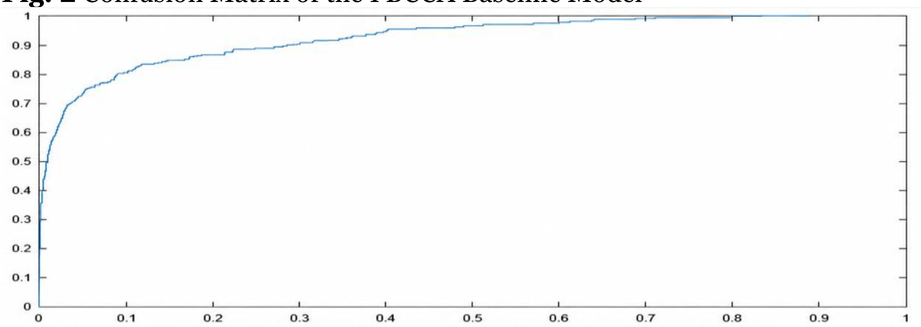

Fig. 3 ROC Curve of the FBCCA Model

4.3 Depthwise EEGNet and FBCCA Hybrid Analysis

The hybrid model Depthwise EEGNet-FBCCA uses both frequency analysis and deep learning feature extraction to improve its accuracy and robustness. This confusion matrix for the hybrid model has a stronger diagonal than that of the FBCCA model, reflecting higher accuracy (Figure 4). The lower off-diagonal values indicate that the model better separates frequencies that are close to each other. This is due to EEGNet's capacity to learn features from the EEG signals, in addition to the frequency information obtained with FBCCA.

The ROC analysis indicates a high mean AUC of around 0.9189, suggesting a good separation of classes (Figure 5). The curve's slope is greater than that of FBCCA, indicating better class separation. This implies that the hybrid model is less affected by variations and noise in EEG features. But, although the AUC is high, the accuracy of the model is found to be around 53.13%. This suggests that although the model is good at classifying (ranking) the probabilities of different classes, it may not be good at classifying (predicting) the actual classes. This suggests that while the model may be effective in distinguishing between classes, it may not always be able to accurately predict the class.

In conclusion, the Depthwise EEGNet and FBCCA hybrid approach shows promising performance and is more robust and has better feature representation than the baseline model, but needs to be optimised for better accuracy.

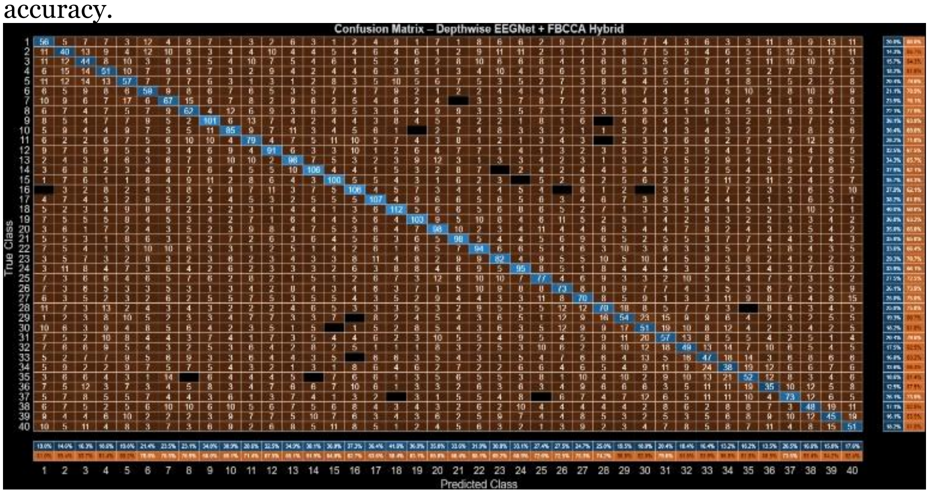

Fig. 4 Confusion Matrix of the EEGNet–FBCCA Hybrid Model

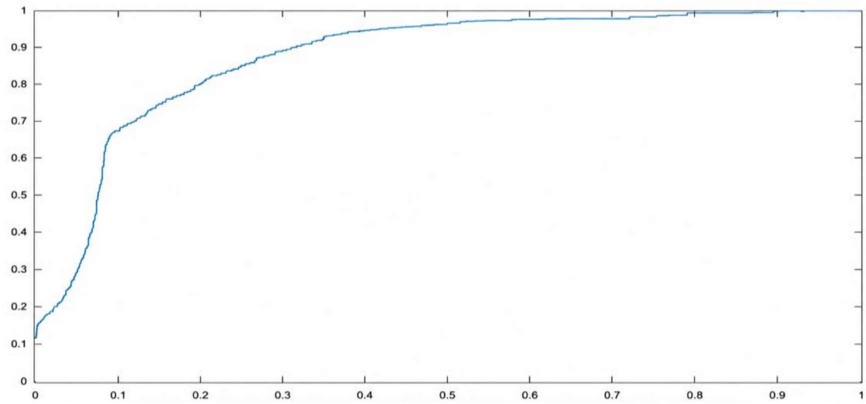

Fig. 5 ROC Curve of the EEGNet–FBCCA Hybrid Model

4.4 Subject-Specific CNN Analysis

The confusion matrix of the subject-specific CNN model exhibits a lower degree of dominance along the diagonal than hybrid models, implying a lower performance (Figure 6). There are strong off-diagonal elements, indicative of considerable cross-class misclassification. This suggests the model is easily influenced by noise and struggles to learn from other subjects. The ROC curve suggests a reasonable discriminative performance, with an average value of the area under the curve (AUC) being around 0.85-0.89 (Figure 7). The model has good discriminative power, but is not as robust as the hybrid models.

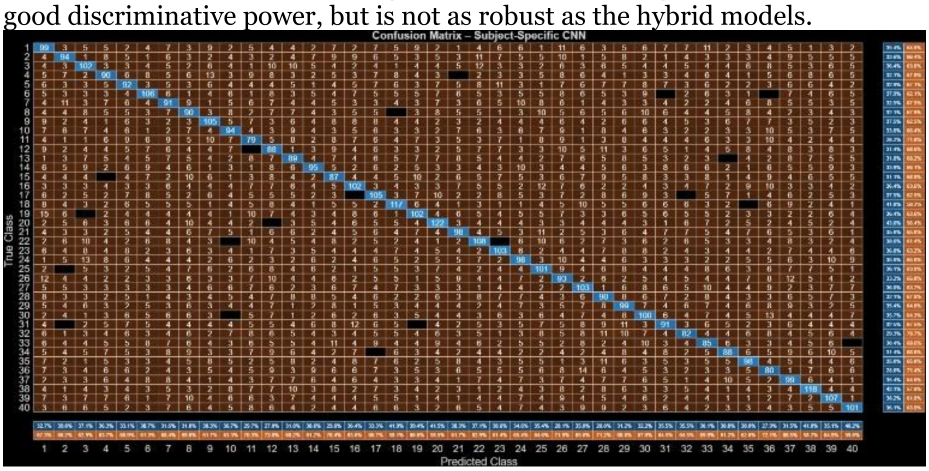

Fig. 6 Confusion Matrix of the Subject-Specific CNN Model

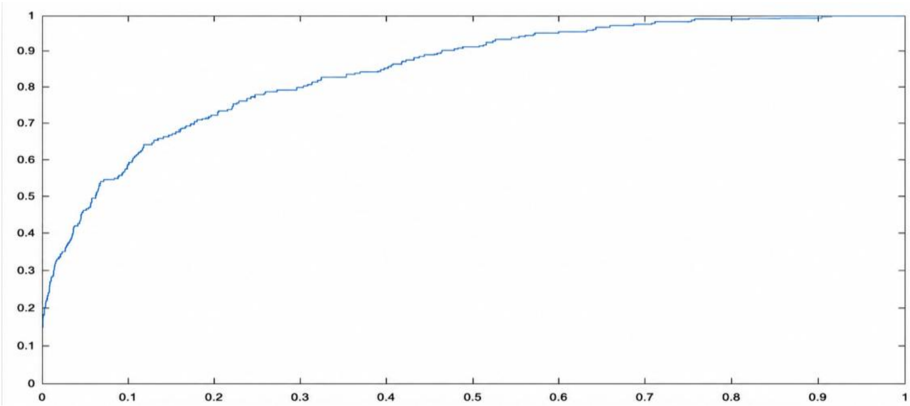

Fig. 7 ROC Curve of the Subject-Specific CNN Model

4.5 Fine-Tuned Hybrid Model (CNN + FBCCA) Analysis

The confusion matrix of the fine-tuned hybrid model is greatly dominated by the diagonal, showing an improvement of the classification performance (Figure 8). The non-diagonal entries are considerably smaller, reflecting the model's capacity to better discriminate between different frequencies and/or target classes. The ROC curve demonstrates good class separability with a mean area under the curve (AUC) of about 0.9255 (Figure 9). It is steep and approaches the top-left corner, suggesting good discrimination. The CNN model (subject-specific) has an overall classification accuracy of around 46%. While the CNN model is specifically designed for individual subjects, it does not necessarily capture consistent features across trials, decreasing the model's robustness. The outcome highlights the need to incorporate other approaches in deep learning-based models to make them more stable.

The hybrid model achieves a classification accuracy of 59.03% when fine-tuned, making it one of the most accurate models. The boost in performance is due to the feature extraction in the frequency domain of FBCCA and the spatial-temporal representation learning of CNN. The fine-tuning enables it to learn from individual subject-to-subject variations, increasing its accuracy. This hybrid model shows that such models not only perform better, but are also more stable and robust than the independent models.

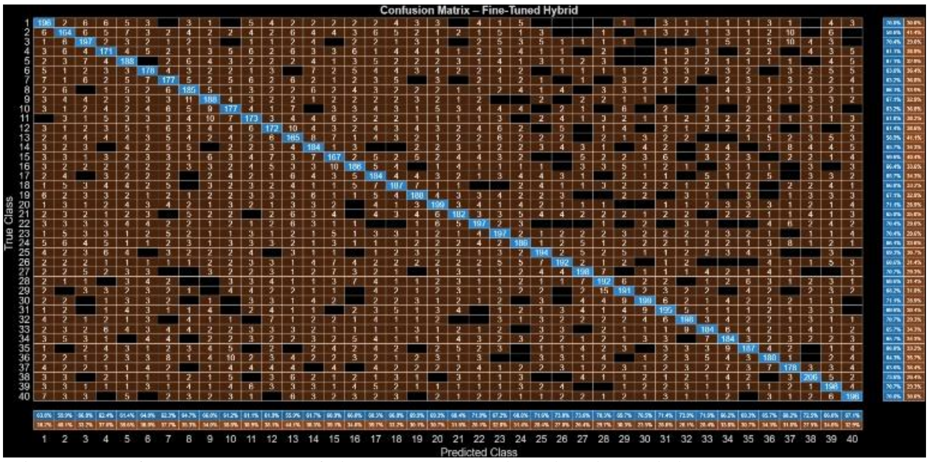

Fig. 8 Confusion Matrix of the Fine-Tuned Hybrid Model

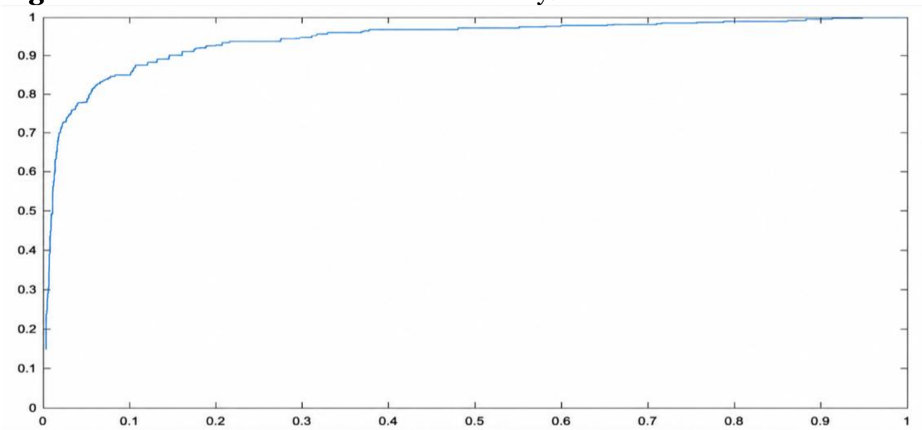

Fig. 9 ROC curve of the Fine-Tuned Hybrid Model

4.6 Windowed Hybrid SSVEP Classifier Analysis (FBCCA + Temporal CNN)

The diagonal entries of the confusion matrix of the windowed hybrid model are dominant compared to the baseline and subject-specific models, suggesting improved classification accuracy (Figure 10). The use of temporal windows in the EEG signals enables the model to adapt to the dynamic changes in the SSVEP responses, thereby decreasing the classification errors. The ROC curve illustrated good discriminative power, as indicated by the mean area under the ROC curve (AUC) of 0.9189 (Figure 11). The curve has a steep slope for all the classes, suggesting good discrimination between different responses. The hybrid model's overall classification accuracy is 53.13%. Although this is marginally lower than the fine-tuned hybrid model, the proposed windowed approach increases the stability of the model. Using multiple short windows, the model allows the reduction of noise and enhances the prediction stability. This approach shows how including temporal information in the analysis of EEG signals can be beneficial and its potential for online BCI applications.

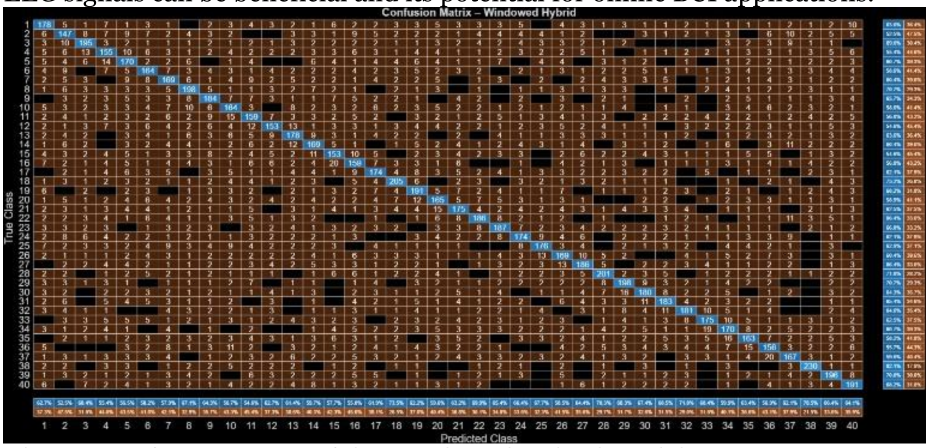

Fig. 10 Confusion Matrix of the Windowed Hybrid Model

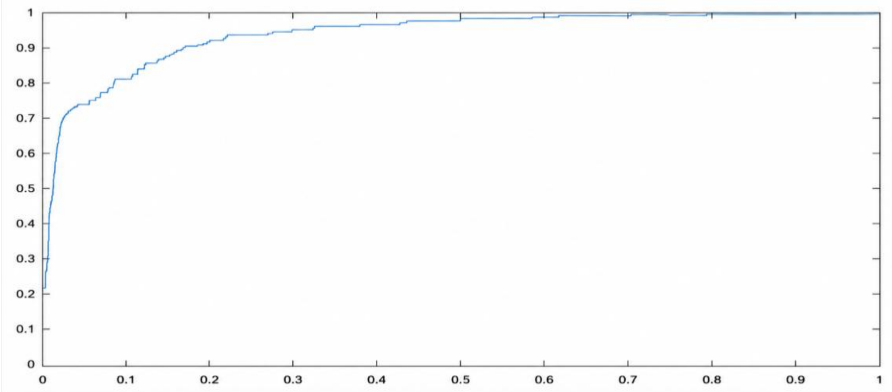

Fig. 11 ROC Curve of the Windowed Hybrid Model

4.7 Source-Free Domain Adaptation Analysis

The source-free domain adaptation model aims to enhance the generalisation to new subjects without access to the target subject's data. The findings suggest that the model has high information transfer rates, suggesting that it is successful in generalising to new subjects. The results demonstrate that the model has an information transfer rate of 201 bits per minute (bpm) in benchmark scenarios and 145 bpm in practical scenarios (Figure 12). These findings suggest that domain adaptation methods can greatly decrease the amount of calibration needed while keeping high performance. But the complexity of applying domain adaptation methods is a challenge, as they need to align the feature distributions of the source and target domains. However, the technique shows great potential for real-world application of BCIs, where subject-independent performance is crucial.

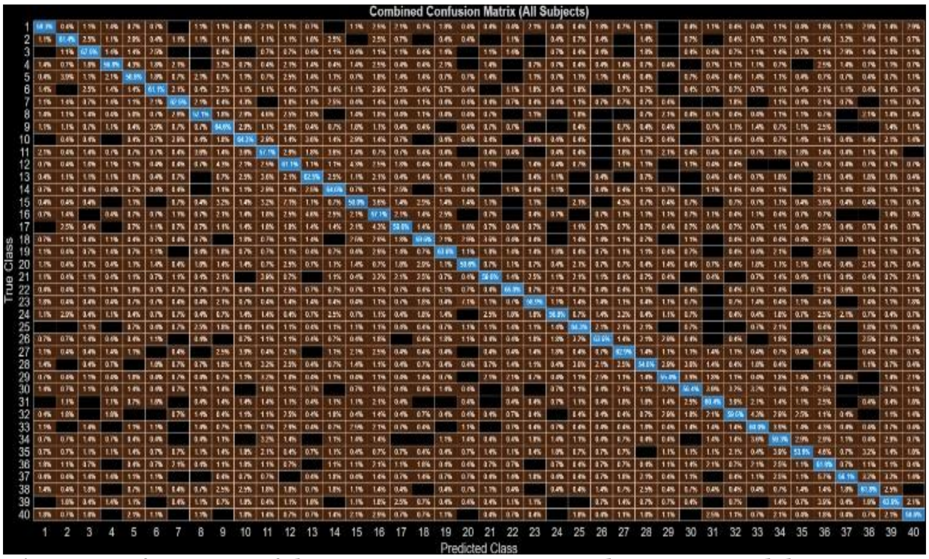

Fig. 12 Performance of the Source-Free Domain Adaptation Model

4.8 Enhanced Source-Free Domain Adaptation with Multi-Subband Fusion

The improved source-free domain adaptation with a multi-subband fusion approach was tested on the BETA data set (70 subjects, 40 classes). It obtained a mean recognition accuracy of 71.38%, showing good generalisation performance. The distribution of subject-wise accuracy reveals some variability, but overall satisfactory results. The confusion matrix has a high diagonal dominance, showing correct classification of most classes (Figure 13). There are some misclassifications between neighbouring frequencies due to harmonics (Figure 14). The model also shows a mean Information Transfer Rate (ITR) of 68.72 bits/min, suggesting the possibility for effective real-time BCIs (Figure 15).

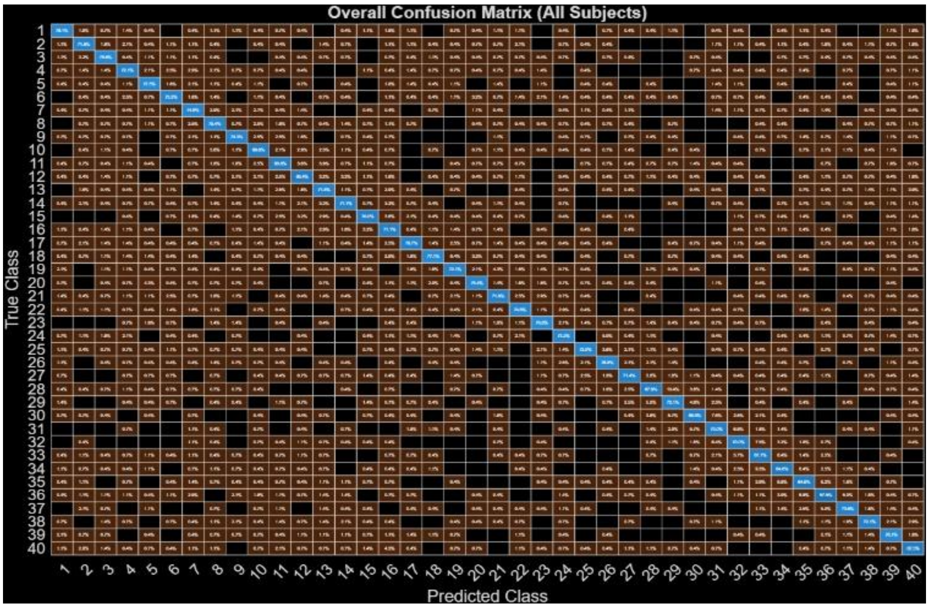

Fig. 13 Overall Confusion Matrix for Enhanced Domain Adaptation Model

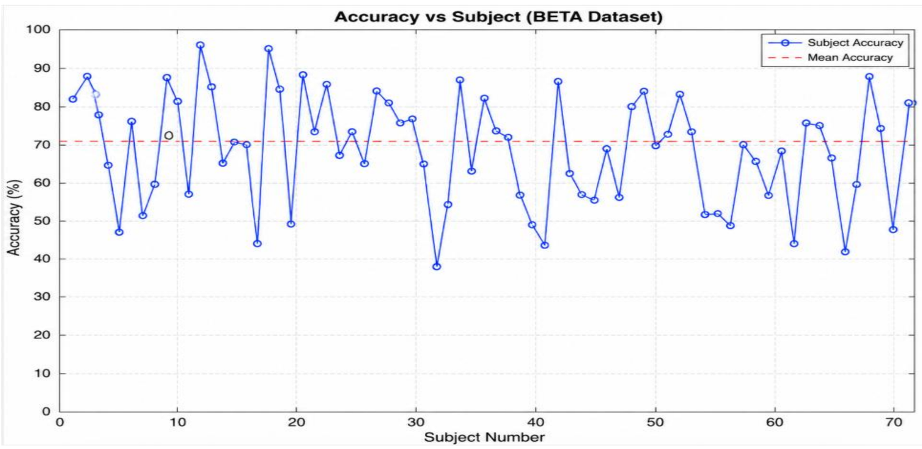

Fig. 14 Accuracy vs Subject for Enhanced Domain Adaptation Model

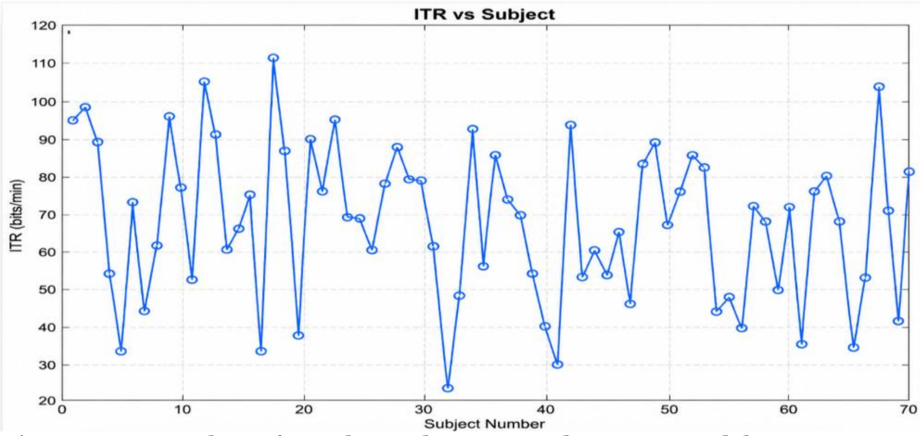

Fig. 15 ITR vs Subject for Enhanced Domain Adaptation Model

4.9 Comparative Performance Evaluation

The performance of all the models can be compared to understand the trends (Table 1). The FBCCA baseline model attains moderate accuracy as it is based on frequency domain correlation, and the subject-specific CNN model performs poorly due to a lack of generalising capability. Hybrid models systematically outperform other models, highlighting the benefits of integrating signal processing with deep learning. The fine-tuned hybrid model shows the best accuracy of nearly 59.03%, followed by the windowed hybrid model and hybrid models based on EEGNet. This is also evident in the ROC-AUC results, where the hybrid models have values greater than 0.91, suggesting good class separability. In contrast, lower values of AUCs are obtained for FBCCA and CNN models, signifying poor discriminative ability.

Table 1 Comparative Performance of Different SSVEP Classification Methods

Method	**Accuracy (%)**	**Mean AUC**	**Generalization**
FBCCA	48.00	0.88	High
Subject-Specific CNN	46.00	0.85 – 0.89	Low
Depthwise EEGNet + FBCCA	53.13	0.9189	Medium
Fine-Tuned Hybrid (CNN + FBCCA)	59.03	0.9255	High
Windowed Hybrid (FBCCA + Temporal CNN)	53.13	0.9189	High
Source-Free Domain Adaptation	66	—	Very High
Enhanced Source-Free Domain Adaptation with Multi-Sub Band Fusion	71.38	—	Very High

4.10 Comparison with Recent State-of-the-Art Approaches

Recent advances in SSVEP-based BCI research have increasingly focused on deep learning, transformer-based architectures, and hybrid signal-processing frameworks for improving classification accuracy and cross-subject generalization. Conventional methods such as CCA, FBCCA, and Task-Related Component Analysis (TRCA) continue to remain important because of their computational efficiency and strong frequency discrimination capability. However, these approaches may exhibit reduced flexibility when modeling complex spatial–temporal EEG variations across subjects.

Deep learning approaches, particularly CNNs, have demonstrated improved representation learning by automatically extracting discriminative spatial and temporal features from EEG recordings. Lightweight architectures such as EEGNet have shown encouraging performance while maintaining relatively low computational complexity. More recently, transformer-based models such as SSVEP former have been proposed to capture long-range temporal dependencies and improve cross-subject learning capability. Although these methods provide strong classification performance, they generally require larger training datasets and increased computational resources.

Hybrid approaches integrating frequency-domain processing with deep learning have emerged as a practical alternative for balancing interpretability, computational efficiency, and classification performance. By combining deterministic spectral analysis with adaptive feature learning, hybrid frameworks attempt to improve robustness against noise, harmonic interference, and inter-subject variability. Table 2 presents a concise comparison between representative approaches reported in recent literature and the models evaluated in the present study.

Table 2 Comparison with Recent State-of-the-Art SSVEP Classification Approaches

Method	Core Approach	Strengths	Limitations
CCA	Correlation-based frequency recognition	Simple and training-free	Sensitive to noise and harmonics
FBCCA	Multi-subband frequency-domain analysis	Improved harmonic utilization and computational efficiency	Limited nonlinear feature modeling
TRCA	Trial consistency optimization	Strong within-subject performance	Reduced cross-subject generalisation

EEGNet	Lightweight CNN architecture	Efficient spatial–temporal feature extraction	Requires sufficient training data
SSVEPformer	Transformer-based EEG modeling	Improved long-range dependency learning	Higher computational complexity
Proposed Hybrid Models	FBCCA + CNN-based fusion	Balanced frequency and learned feature representation	Moderate training complexity
Enhanced Source-Free Domain Adaptation	Multi-subband adaptive hybrid framework	Improved cross-subject robustness and reduced calibration	Increased adaptation complexity

The findings of the present study indicate that hybrid and adaptation-based approaches provide a favourable balance between classification performance and practical deployment feasibility. In particular, the enhanced source-free domain adaptation framework demonstrates competitive cross-subject performance while maintaining stable frequency discrimination through integration of FBCCA-guided adaptation and multi-subband fusion. Rather than proposing a completely new classification paradigm, the present work provides a unified empirical evaluation of multiple integrative strategies under a common validation framework, enabling systematic comparison of robustness, generalisation capability, and practical applicability for SSVEP-based BCI systems.

4.11 Discussion on Model Generalisation and Robustness

The experimental findings demonstrate clear performance differences among conventional, deep learning, and hybrid SSVEP classification approaches. Frequency-domain methods such as FBCCA provide computational efficiency and stable frequency recognition, although their ability to model nonlinear EEG variability remains limited. Hybrid architectures combining FBCCA and CNN-based learning exhibit improved classification consistency due to the integration of deterministic spectral analysis and adaptive feature representation learning. In particular, fine-tuned and adaptation-based models demonstrate improved robustness under cross-subject evaluation conditions. The enhanced source-free domain adaptation framework achieved the highest overall recognition performance, indicating the importance of adaptive learning strategies for reducing inter-subject variability in practical BCI systems. Nevertheless, computational complexity and calibration stability remain important considerations for real-time deployment. Transformer-based methods generally achieve strong feature representation performance but often require larger datasets and higher computational resources for stable training. Conventional methods such as TRCA and FBCCA remain computationally efficient but may exhibit reduced flexibility when modelling complex EEG variability. Hybrid frameworks attempt to balance these complementary strengths through integration of frequency-domain processing and learned representations. In comparison, the present study focuses primarily on a unified empirical evaluation of multiple integrative classification strategies under a consistent validation framework. The enhanced source-free domain adaptation model achieves competitive performance while maintaining practical feasibility for cross-subject SSVEP-based BCI applications.

4.12 Computational Complexity and Real-Time Deployment Feasibility

The computational complexity of the considered models depends on the classification technique used. The main advantage of FBCCA is its comparatively low computational cost, since it is based on a deterministic correlation between the EEG recordings and the reference templates, which are sinusoidal signals. Hence, FBCCA is still applicable for real-time inferences and in low-resource BCI environments. Convolutional operations and parameter tuning during training add more computational complexity to deep learning architectures, such as models based on EEG signals and hybrid CNN models. But the complexity of inference is still moderate thanks to compact architectures that reuse depthwise and separable convolutions. Therefore, lightweight CNN structures are beneficial for practical application in online SSVEP-based BCI systems.

The windowed hybrid model requires more computation since the multiple temporal segments are processed separately prior to score fusion. Likewise, domain adaptation methods follow pseudo-label refinement and confidence-based updating in an iterative manner, which makes the adaptation process more complex during the calibration phases. These operations, however, are offline or are carried out periodically and do not have significant impact on the real-time inference performance. The experimental results presented herein indicate that by combining the advantages of the frequency-domain and deep learning methods, a practical compromise between classification accuracy and computational complexity can be obtained. The proposed framework is readily adaptable for online BCI and real time assistive communication applications with optimized implementation and GPU assisted processing.

Despite the encouraging performance obtained in this study, several limitations remain. Hybrid and adaptation-based frameworks introduce increased computational complexity compared with conventional frequency-domain approaches. In addition, classification performance may still vary across subjects because of EEG non-stationarity and recording variability. Future work may therefore focus on lightweight adaptive architectures, reduced calibration requirements, and embedded real-time implementation for practical assistive BCI systems.

5 Conclusion

In this study, we present a robust multi-model approach to SSVEP-based Brain–Computer Interface speller systems, which combines conventional signal processing and deep learning techniques. Various classification algorithms, such as FBCCA, subject-specific CNN, hybrid approaches based on EEGNet, fine-tuned hybrid models, hybrid models with windowed data, and domain adaptation, were rigorously implemented and compared within a common framework. Our findings show that hybrid methods consistently achieve better results than individual methods. The FBCCA baseline method offers reliable frequency analysis but is not flexible to complex EEG patterns. The subject-specific CNN model, while able to learn subject-specific features, has poor generalization. Hybrid models, on the other hand, successfully integrate spectral and spatial–temporal features, resulting in enhanced classification accuracy.

The fine-tuned hybrid model has the highest accuracy and AUC, demonstrating the benefits of the combination of FBCCA and deep learning, along with subject-specific fine-tuning. The windowed hybrid model also highlights the benefits of time windowing in enhancing resilience to noise and non-stationary features of the EEG recordings. Also, source-free domain adaptation methods have the potential to offer subject-independent performance, essential for real-world applications.

In conclusion, the research demonstrates that a hybrid approach with multiple models offers a practical and scalable approach for SSVEP-based BCI. The combination of various approaches into a single model leads to better accuracy, flexibility and generalization, enabling the system to be applied in assistive communication devices.

Acknowledgment

The authors would like to thank the Electronics and Communication Engineering Department of Vasavi College of Engineering, Hyderabad, for providing research support, laboratory resources, and guidance throughout the project.

Contribution of each Author

Srinivas Rao Gorre is responsible for Supervision, Conceptualization, Project Administration, Technical Guidance, Review and Editing, Final Approval of Manuscript, Radha Krishna J is responsible for Conceptualization, Methodology, Software Development, Data Curation, Experimental

Investigation, Writing – Original Draft, Vardhan Manjunath B dealt with Methodology, Software Implementation, Validation, Performance Evaluation, Writing – Review and Editing, and Nishanth Reddy S has done Data Analysis, Visualization, Experimental Validation, Literature Survey, Writing – Review and Editing.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

1. S. Pan, K. Sun, J. Zhang, and X. Gao, "A Review of Deep Learning Approaches for SSVEP-Based Brain-Computer Interfaces," IEEE Access, vol. 9, pp. 120953-120972, 2021.
2. D. Xu, Y. Li, F. Tang, and X. Chen, "An Analysis of Deep Learning Models for SSVEP-Based Brain-Computer Interfaces," Brain Sciences, vol. 13, no. 2, 2023.
3. J. Chen, Y. Zhang, Y. Pan, et al., "SSVEPformer: A Transformer-Based Model for SSVEP Classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022.
4. O. B. Guney, M. Oblokulov, and H. Ozkan, "A Deep Neural Network for SSVEP-Based Brain-Computer Interfaces," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021.
5. X. Xiong, Y. Li, and Z. Wang, "Deep Transfer Learning for SSVEP-Based Brain-Computer Interfaces," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024.
6. Y. Wei, H. Li, and X. Gao, "Hybrid eTRCA and CNN-Based Framework for High-Accuracy SSVEP Classification," IEEE Transactions on Biomedical Engineering, 2025.
7. H. Zhang, J. Liu, and X. Chen, "Temporal Convolutional Networks for EEG-Based Brain-Computer Interfaces," IEEE Access, 2022.
8. M. Hossan, S. Mahmood, and M. I. Molla, "Advances in Deep Learning-Based SSVEP Classification: A Survey," IEEE Access, 2025.
9. Y. Wang, X. Chen, and S. Gao, "Benchmarking SSVEP-Based BCI Systems: Recent Advances and Future Directions," IEEE Transactions on Cognitive and Developmental Systems, 2021.
10. O. B. Guney, M. Oblokulov, and H. Ozkan, "Source-Free Domain Adaptation for SSVEP-Based Brain-Computer Interfaces," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023.
11. J. Li, Z. Zhang, and L. Wang, "Cross-Subject EEG Classification Using Domain Adaptation Techniques," IEEE Transactions on Biomedical Engineering, 2022.
12. Y. Zhang, X. Gao, and S. Gao, "Enhancing SSVEP Detection Using Multi-Scale Deep Neural Networks," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021.
13. X. Chen, Y. Wang, and S. Gao, "Filter Bank Methods for SSVEP Detection: Recent Improvements and Applications," Journal of Neural Engineering, 2020.

14. R. Mane, S. Chew, and K. K. Ang, "EEGNet-Based Deep Learning Framework for Brain-Computer Interface Systems," IEEE Access, 2021.
15. Z. Lin, Y. Zhang, and X. Gao, "Advanced Frequency Recognition Methods for SSVEP-Based BCIs," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020.
16. Y. Li, X. Chen, and S. Gao, "Deep Learning-Based Feature Extraction for SSVEP Signal Classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021.
17. J. Wang, H. Li, and X. Gao, "Robust SSVEP Detection Using Multi-Domain Feature Fusion," IEEE Access, 2022.
18. H. Liu, Y. Zhang, and S. Gao, "Improved EEG Signal Preprocessing Techniques for BCI Applications," IEEE Transactions on Biomedical Engineering, 2021.
19. P. Roy, M. Banville, I. Albuquerque, et al., "Deep Learning-Based Electroencephalography Analysis: A Systematic Review," Journal of Neural Engineering, 2021.
20. S. Craik, Y. He, and J. L. Contreras-Vidal, "Deep Learning for EEG-Based Brain-Computer Interfaces: A Review," Journal of Neural Engineering, 2020.
21. X. Chen, Y. Wang, and S. Gao, "An Improved FBCCA Method for SSVEP-Based Brain-Computer Interfaces," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020.
22. V. J. Lawhern et al., "EEGNet Revisited: Improved Architectures for EEG-Based BCIs," IEEE Access, 2021.
23. Y. Wei, X. Gao, and S. Gao, "Hybrid Deep Learning and Signal Processing Methods for SSVEP Classification," IEEE Transactions on Cognitive and Developmental Systems, 2023.
24. R. Schirrmeister, J. Springenberg, and T. Ball, "Deep Learning for EEG Decoding: Recent Advances," Journal of Neural Engineering, 2021.
25. X. Xiong, Y. Li, and Z. Wang, "Fine-Tuning Strategies for Cross-Subject EEG Classification," IEEE Transactions on Biomedical Engineering, 2024.
26. H. Zhang, J. Liu, and X. Chen, "Temporal Window-Based CNN Models for EEG Signal Classification," IEEE Access, 2022.
27. O. B. Guney, M. Oblokulov, and H. Ozkan, "Unsupervised Domain Adaptation for EEG-Based BCIs," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023.
28. L. Zhang, Y. Wang, and X. Gao, "Hierarchical Classification Strategies for Multi-Class EEG Signals," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021.

Srinivas Rao Gorre, is presently working as an Assistant Professor at the Department of Electronics and Communication Engineering, Vasavi College of Engineering, Hyderabad, Telangana, India. He is pursuing Ph.D. from S R University, Warangal, Telangana, India. He has 18 years of experience in teaching and research. He published 6 international journals and 4 International Conferences. His areas of interests are Biomedical Signal Processing, BCI, IoT, Machine Learning and Deep Learning

Radha Krishna is a student in Electronics and Communication Engineering and has a keen interest in Brain-Computer Interfaces (BCI), VLSI, Machine Learning, Signal Processing and Embedded Systems. He has been working on SSVEP-based BCI systems, with an emphasis on enhancing the classification accuracy by combining deep learning and signal processing techniques. His experience includes working with EEG datasets, convolutional neural networks and classification methods based on FBCCA. He assesses system performance in terms of accuracy and Information Transfer Rate (ITR). Outside of BCI research, he has also worked on projects concerning IoT systems, FPGA/Verilog design, AXI-based System-on-Chip architectures, and sensor fault detection. Keen to learn new technologies and use engineering principles to develop innovative, smart and practical solutions to real world problems.

Vardhan Manjunath B graduated with a Bachelor of Engineering degree in Electronics and Communication Engineering from Vasavi College of Engineering. His academic interests span Brain-Computer Interfaces (BCI), machine learning, signal processing, and intelligent assistive technologies. He has been actively involved in research on Steady-State Visual Evoked Potential (SSVEP)-based BCI systems, with a focus on improving classification accuracy and developing efficient human-computer interaction frameworks. Beyond research, he has participated in innovation and entrepreneurship initiatives, contributing to projects that bridge theoretical concepts with practical applications.

Nishanth Reddy Sunkireddy graduated with a Bachelor of Engineering degree in Electronics and Communication Engineering from Vasavi College of Engineering. His areas of interest include Brain-Computer Interfaces (BCI), signal processing, semiconductor technology, and emerging intelligent systems. His research focuses on the development of Steady-State Visual Evoked Potential (SSVEP)-based Brain-Computer Interface systems for enhancing human-computer interaction. He is passionate about applying advanced engineering concepts to solve real-world challenges and contributing to innovative research in the field of Electronics and Communication Engineering.