A Multi-view CNN with Novel Variance Layer for Motor Imagery Brain Computer Interface

Ravikiran Mane, Neethu Robinson, A. P. Vinod, Seong-Whan Lee, Cuntai Guan

Abstract— **Accurate and robust classification of Motor Imagery (MI) from Electroencephalography (EEG) signals is among the most challenging tasks in Brain-Computer Interface (BCI) field. To address this challenge, this paper proposes a novel, neuro-physiologically inspired convolutional neural network (CNN) named Filter-Bank Convolutional Network (FBCNet) for MI classification. Capturing neurophysiological signatures of MI, FBCNet first creates a multi-view representation of the data by bandpass-filtering the EEG into multiple frequency bands. Next, spatially discriminative patterns for each view are learned using a CNN layer. Finally, the temporal information is aggregated using a new variance layer and a fully connected layer classifies the resultant features into MI classes. We evaluate the performance of FBCNet on a publicly available dataset from Korea University for classification of left vs right hand MI in a subject-specific 10-fold cross-validation setting. Results show that FBCNet achieves more than 6.7% higher accuracy compared to other state-of-the-art deep learning architectures while requiring less than 1% of the learning parameters. We explain the higher classification accuracy achieved by FBCNet using feature visualization where we show the superiority of FBCNet in learning interpretable and highly generalizable discriminative features. We provide the source code of FBCNet for reproducibility of results.**

I. Introduction

Brain-Computer Interface (BCI) systems capture real-time neuronal activity using a signal acquisition device and try to decode user's intentions from the observed brain states [1]. In BCI systems, electroencephalography (EEG) is the most widely used signal acquisition modality and Motor-Imagery (MI) based EEG-BCI, wherein participant performs mental rehearsal of a particular motor movement is one of the frequently investigated protocols. Hence, owing to its widespread use and application in the post-stroke motor rehabilitation, EEG-BCI literature contains many reports on decoding techniques to classify various MI classes with an aim to achieve higher accuracy. Classical machine learning techniques like linear and non-linear classifiers, nearest neighbour classifiers as well as more data-driven techniques like neural networks and deep learning have been explored for this task of MI classification [2]–[6].

EEG, with a high noise content and low signal-to-noise ratio (SNR), offers unique challenges to the classical machine learning techniques. Therefore, many of the classical machine learning strategies of MI classification have paid prominent attention to the extraction of neurophysiologically sound features from the EEG data. These extracted features offer higher SNR and hence result in a more generalizable classifier model. Neuroscientific studies have documented that MI elicit characteristic EEG activation patterns known as sensory-motor rhythms (SMR). SMRs are generally observed at the contralateral and ipsilateral sensory-motor regions in the form of a time-locked, event-related desynchronization/ synchronization (ERD/ERS) [1]. It is also known that different classes of MI differ in the spectro-spatial distribution of SMRs [1]. Taking advantage of exactly this discriminative information, Filter Bank Common Spatial Patterns (FBCSP) has been proven to be one of the most successful algorithms for MI classification [2]. FBCSP first decomposes the EEG into multiple narrowband signals, achieving EEG spectral localisation. The narrowband EEG is then spatially filtered using Common Spatial Patterns (CSP) algorithm which extracts discriminative spatial patterns such that the spatially filtered output will have a maximised class discriminative variance. This results in the spectro-spatially localised features which are then classified using support vector machine algorithm. An efficient and effective embodiment of neurophysiological priors in the classifier design can be considered as a reason behind the success of FBCSP algorithm.

Different from classical machine learning approaches, Deep Learning (DL), which is an extensively data-driven approach to classification has shown very prominent results in the field of natural language processing and computer vision. Hence, there is a growing interest among BCI researchers to use DL methods for achieving higher classification accuracies in MI decoding. In particular, architectures based on Convolutional Neural Network (CNN) have gained popularity in the BCI domain due to their ability of effectively learning the local connectivity patterns from the given data [3]–[7]. Due to this, in recent years, many CNN based architectures have been proposed and validated on various MI datasets. Although these architectures have outperformed state-of-the-art machine learning techniques in the subject-specific classification task, the achieved improvements are still marginal [3]–[5]. These only marginally better results with DL can be traced back to the nature of BCI datasets. One important difference between computer vision and BCI

This work was partially supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (No. 2017-0-00451).

The computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore

Ravikiran Mane, Neethu Robinson, Cuntai Guan are with the Nanyang Technological University, 50 Nanyang Avenue, Singapore (email: ravikian001@e.ntu.edu.sg, nrobinson@ntu.edu.sg, ctguan@ntu.edu.sg)

A. P. Vinod is with Indian Institute of Technology, Palakkad, India (email: vinod@iitpkd.ac.in)

S. -W. Lee is with the Korea University, Seoul 136-713, South Korea (e-mail: sw.lee@korea.ac.kr)

Corresponding author: Cuntai Guan

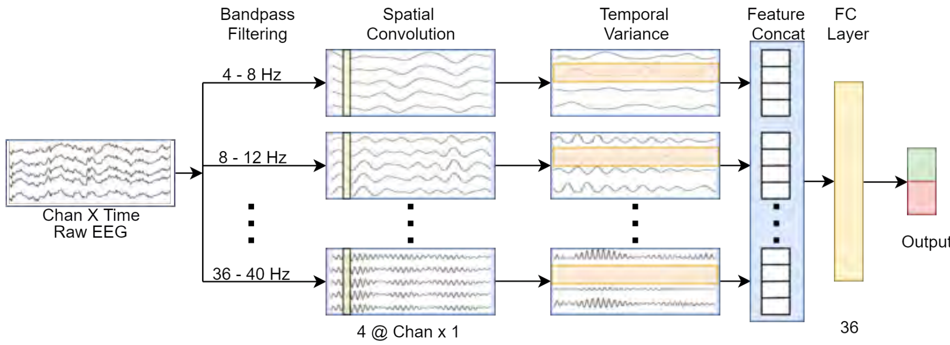

Fig. 1. Proposed network architecture: FBCNet.

field is that BCI datasets are very small and contain very few training samples while having high dimensionality. This creates a huge problem for adaptation of DL methods in the BCI field and results in heavy overfitting of the models. One solution to this overfitting problem is to encode the neurophysiological priors in the classifier architecture.

In this paper, we present Filter-Bank Convolutional Network (FBCNet), which is a novel end-to-end CNN architecture designed with an aim to best encode neurophysiological priors for MI classification. Motivated by the FBCSP algorithm, FBCNet encodes spectro-spatial discriminative information with the help of spectral filtering of the EEG and CNN based spatial filtering. The temporal information is aggregated using a novel variance layer. FBCNet consists of only two trainable layers and hence offers a direct interpretation of learned features. While being simple and interpretable, FBCNet also offers significantly better classification results. We present the classification superiority of FBCNet over other deep learning architectures and FBCSP with the help of large publicly available dataset from Korea University involving the task of classifying right vs left hand MI [8]. We also explain the better classification result with the help of feature map visualization. Lastly, we provide the PyTorch implementation of the FBCNet at `https://github.com/ravikiran-mane/FBCNet`.

II. Related Works

In recent years, numerous DL based architectures have been proposed for EEG-BCI classification [3]–[7]. In [3], Schirrmeister *et* al. proposed two CNN based architectures named Deep Convnet and Shallow Convnet. Also being inspired by FBCSP, both of these architectures consisted of an initial CNN layer acting along the time dimension which was followed by CNN-based spatial filtering. Although proven to be effective, unlike FBCSP, these networks lacked the explicit spectral filtering of the input EEG data. Building upon Deep Convnet, Robinson *et* al. showed that the multi-view representation of EEG using a filter bank is indeed highly effective by achieving higher classification accuracy over broad-band Deep Convnet [6]. EEGNet is another CNN architecture that was also inspired by FBCSP and similar to Deep Convnet, implemented a temporal CNN which was followed by a spatial convolution [4]. By implementing the depth-wise spatial convolution, EEGNet was constructed to learn separate spatial filters for each of the temporally filtered signals. Resenting yet more similarity with FBCSP, EEGNet also achieved better results than both Deep and Shallow Convnets. However, both Deep Convnet and EEGNet lack the explicit multi-view EEG representation and variance-based feature extraction which are core characteristics of FBCSP. In this work, FBCNet incorporates these two missing components.

III. Methodology

A. Proposed Architecture : Filter-Bank Convolutional Network

FBCNet is designed to effectively extract the spectro-spatial discriminative information which is a signature of MI while avoiding the problem of overfitting in the presence of small datasets. There are four main components of FBCNet, viz., 1. The multi-view representation of the EEG data that is obtained by spectrally filtering the raw EEG with multiple narrowband filters, 2. The spatial CNN layer that learns discriminative spatial patterns, 3. The variance layer, which computes temporal variance of input signal representing the power in the signal, 4. The final fully connected (FC) layer that classifies features from the Variance layer into given classes. The multi-view EEG representation followed by the spatial filtering allows extraction of spectro-spatial features and Variance layer provides a compact representation of the temporal information. With this brief design philosophy, this section provides details of the FBCNet architecture and the architecture is presented in Fig. 1.

1) Data Representation: FBCNet uses the multi-view representation of the EEG data in which each view represents a narrowband localised EEG. Consider a set of single-trial epoched raw EEG data with C channels and T time points from N_c classes that is represented as $x \in \mathbb{R}^{C\times T}$ and its corresponding label to be $y \in \{0,1,..,N_c\}$. FBCNet utilises the multi-view representation of the EEG data, x_{FB}, which is generated by spectrally filtering the raw EEG x with a filter bank $F = \{f_i\}_{i=1}^{N_b}$ consisting of N_b number of narrow band temporal filters. Following the filtering operation, x_{FB}

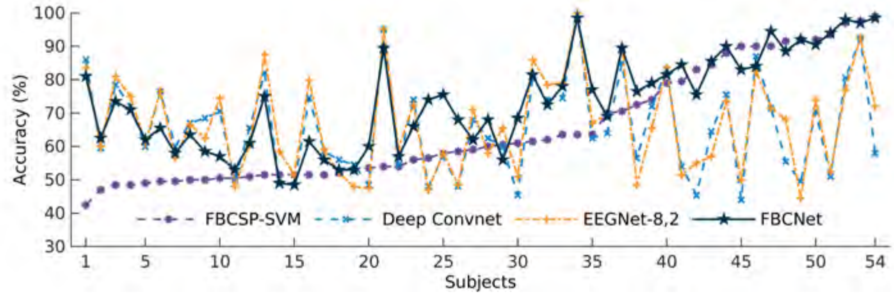

Fig. 2. Average accuracy in the binary classification of right hand vs left hand MI using 10-fold cross-validation. (sorted by FBCSP-SVM acc.)

belongs in a $\mathbb{R}^{C \times T \times N_b}$ space where the time-series along the second dimension is spectrally localised.

In this particular work, as proposed in the FBCSP algorithm, the filter bank F was constructed using $N_b = 9$ non-overlapping frequency bands, each of 4Hz bandwidth, spanning from 4 to 40 Hz (4-8, 8-12, ..., 36-40 Hz). The filtering is done using Chebyshev Type II filter with transition bandwidth of 2Hz and stopband ripple of 30dB.

2) Architecture: The convolutional layer of FBCNet consists of N_b parallel Spatial Convolution Blocks (SCB). Each convolutional block is composed of a spatial CNN module (Conv2D) having m filters of size $(C, 1)$, a Batch-normalisation layer, and finally, an Exponential Linear Unit (ELU) activation function all arranged in a sequential manner. Each SCB receives one view of the multi-view EEG representation i.e. the k_{th} SCB receives $x_{FB_{:,:,k}} \in \mathbb{R}^{C \times T \times 1}$ as an input. Owing to the kernel that spans across all the channels, each SCB acts as a spatial filter and outputs m time-series $x_{SCB_{:,:,k}} \in \mathbb{R}^{m \times T \times 1}$. The output of a SCB is then passed to a Variance layer (VL) which computes the temporal variance of the individual time-series as given by equation 1 in the forward pass.

$$x_{VL_{i,-,k}} = \frac{1}{T} \sum_{j=1}^{T} (x_{SCB_{i,j,k}} - \mu_{i,-,k})^2 \tag{1}$$

where, $\mu_{i,-,k}$ is the temporal mean of $x_{SCB_{i,:,k}}$

Following the VL, features from all parallel branches are concatenated and given to a FC layer with linear activation. The output of the FC is then passed through the softmax layer to get the output probabilities of each class.

In this experiment, the value of m was set to 4 which also corresponds to the number of CSP filters extracted from each filter band in FBCSP.

B. Training

The FBCNet architecture was trained using Adam optimiser at default settings [9]. Log-cross-entropy loss was used for gradient updates. L_2 norm of network weights was constrained to 2 to avoid the problem of exploding weights. Training was performed with the early stopping criteria whereby the validation set loss was monitored and training was stopped if there is no decrease in the validation loss for last 200 epochs. After reaching the stopping criteria, network parameters with best validation loss were restored [3].

TABLE I

AVERAGE CLASSIFICATION RESULTS (MEAN $\pm$ STD)

Algorithm	Accuracy (%)	Sensitivity (%)	Specificity (%)
FBCSP-SVM	65.32 $\pm$ 16.85	64.98 $\pm$ 17.02	65.67 $\pm$ 17.21
Deep Convnet	65.72 $\pm$ 13.96	65.89 $\pm$ 17.56	65.56 $\pm$ 17.88
EEGNet-8,2	66.75 $\pm$ 14.25	64.11 $\pm$ 16.95	69.39 $\pm$ 12.67
FBCNet	**73.44 $\pm$ 14.37**	**76.37 $\pm$ 12.63**	**70.50 $\pm$ 18.47**

C. Dataset

We evaluate the performance of FBCNet on a 54 subject MI dataset from Korea university [8]. The dataset consists of binary classification of left hand vs right hand MI. In this work, we have used the data from the first session that consists of 200 EEG trials, each of 4s in length. As done in the original work [8], here we have selected 20 channels in the motor region for the classification task (FC-5/3/1/2/4/6, C-5/3/1/z/2/4/6, and CP-5/3/1/z/2/4/6). Also, the data is down-sampled by factor of 4 to have a sampling frequency of 250 Hz.

D. Experiment

We calculate the average subject-specific classification accuracy in 10-fold cross-validation (CV) settings to evaluate the performance of FBCNet and other baseline methods. In a 10-fold CV, every time 8 folds were used for training, 1 fold for testing and 1 for validation. The trials were randomly allocated to a particular fold and this allocation was maintained constant while evaluating the results across all the methods. We compare the results of FBCNet with three baseline methods, viz. FBCSP-SVM [2], Deep Convnet [3] and EEGNet-8,2 [4]. All the methods were used in the most optimal settings as recommended by respective authors.

IV. RESULTS & DISCUSSION

Table I presents the average classification results for the proposed and baseline methods. It can be observed that all baseline methods achieved very similar classification accuracy when averaged over all subjects. FBCNet outperformed all methods by a large margin by achieving +8.1%, +7.7%, and +6.7% classification accuracy in comparison to FBCSP-SVM, Deep Convnet, and EEGNet-8,2 respectively. Also, the difference in classification accuracy between FBCNet and all baseline methods was statistically significant (3 pairwise

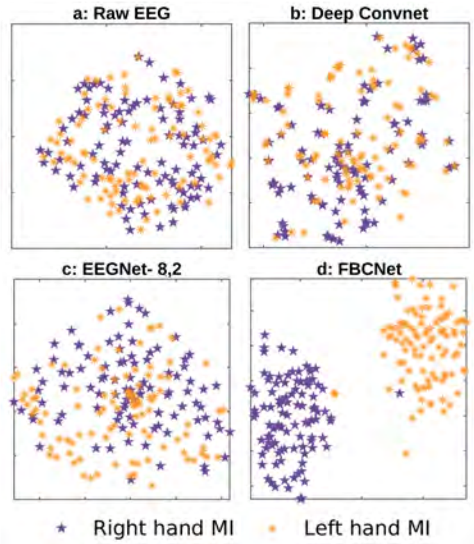

Fig. 3. Visualisation of EEG features for one subject in 2 dimensions using t-SNE. Part (a) is the visualisation of the raw EEG data. Part (b), (c), and (d) present the visualisation of EEG features at the input of the final fully connected layer in trained Deep Convnet, EEGNet-8,2, and FBCNet respectively. Due to the clear superiority of FBCNet in the extraction of highly generalizable features, FBCNet achieved 98.5% classification accuracy for this subject, whereas Deep Convnet, and EEGNet-8,2 resulted in 58.0% and 72% respectively.

Wilcoxon Signed-Rank tests, all $p < 0.05$). Moreover, FBCNet was able to achieve this performance while having only a fraction of learnable parameters (l_n = 902) compared to Deep Convnet (l_n = 283,577) and EEGNet (l_n = 3,114).

Fig. 2 presents the classification accuracy (10-folds averaged) for every subject. The graph is sorted according to the results achieved by FBCSP-SVM algorithm. It can be observed from the figure that the baseline deep learning architectures achieve significantly less accuracy for subjects wherein FBCSP attains >80% 'classification accuracy (avg. acc.: 65%, 67% vs 91%) and FBCNet matches the accuracy of the FBCSP for these subjects (90% vs 91%) resulting in far better performance than that of other deep learning architectures. On the lower end of the distribution, where FBCSP-SVM is struggling with <60% accuracy (avg. 52%), deep learning architectures perform far better (both 64%) and FBCNet matches the accuracy of deep learning architectures in this region (64%). This indicates that, by incorporation of neurophysiological priors in the network architectures, as done in FBCNet, we can bring the benefits of deep learning methods to the BCI domain.

Next, we compare the capabilities of deep learning architectures in extraction of generalizable features from the noisy EEG data. Fig. 3 presents the visualisation of extracted features from all trials in the testing phase for one subject in 2 dimensions using t-SNE algorithm [10]. Fig. 3(a) expresses the non-separability of MI classes from the raw EEG data and highlights the necessity of effective feature extraction. Fig. 3(b), and 3(c) display a 2D representation of a feature map at the FC layer of Deep Convnet, and EEGNet-8,2 and points towards their inability in extraction of separable features. Lastly, the clearly separable features presented in Fig. 3(d) are extracted by FBCNet. This indicates the superiority of FBCNet in extraction of generalizable features and it explains the higher classification results achieved by FBCNet.

Finally, the shallow nature of FBCNet also enables direct interpretation of learned filters. Similar to the CSP feature visualization, the CNN filters can be directly inspected to understand the knowledge extracted by FBCNet. In the future, we will explore these features to offer more insights about FBCNet learning. Lastly, the novel variance layer is the other important part of FBCNet. While acting as a temporal information aggregator, it produces a compact representation of the time-series and extracts a feature that is proportional to the power in the time-series. We will investigate the exact working and the mathematical analysis of the variance layer in the future.

V. Conclusion

This paper proposed a neurophysiologically motivated CNN architecture for classification of motor imagery EEG data. While being completely interpretable, proposed architecture offered a significant increase of +6.7% in classification accuracy. Moreover, visualization of extracted features demonstrated the superiority of the proposed architecture in the extraction of highly generalizable discriminative features. Overall, the results indicate that inclusion of neurophysiological priors while designing deep learning architectures, as done in this work, may offer better classification results in the field of brain-computer interfacing.

References

[1] G. Pfurtscheller, C. Brunner *et al.*, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, may 2006.

[2] K. K. Ang, Z. Y. Chin *et al.*, "Filter Bank Common Spatial Pattern (FBCSP)," *2008 Int. Jt. Conf. Neural Networks (IJCNN 2008)*, pp. 2391–2398, 2008.

[3] R. T. Schirrmeister, J. T. Springenberg *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[4] V. J. Lawhern, A. J. Solon *et al.*, "EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, pp. 1–30, 2018.

[5] O.-y. Kwon, M.-H. Lee *et al.*, "Subject-Independent Brain-Computer Interfaces Based on Deep Convolutional Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. XX, no. X, pp. 1–14, 2019.

[6] N. Robinson, S.-w. Lee *et al.*, "EEG Representation in Deep Convolutional Neural Networks for Classification of Motor Imagery," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, oct 2019, pp. 1322–1326.

[7] S. Sakhavi, C. Guan *et al.*, "Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5619–5629, nov 2018.

[8] M. H. Lee, O. Y. Kwon *et al.*, "EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, pp. 1–16, 2019.

[9] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[10] L. Van Der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2625, 2008.