Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

The Mechanism of a Multi-Branch Structure for EEG-Based Motor Imagery Classification

Haifeng Zhou ,Xinqiao Zhao , Hongmiao Zhang, *Member, IEEE* ,Shaolong Kuang , *Member, IEEE*

***Abstract*—In the field of motor imagery (MI) classification, it is a great challenge to preserve temporal features and spatial ones of the electroencephalogram (EEG) data simultaneously for effective processing. In our former study, a novel framework for MI classification has been proposed, which represents the EEG data into a 3D form and handles it with a multi-branch 3D convolutional neural network (3D CNN). The 3D CNN is constructed with three different branches and already proved to be better than one branch network. In this paper, to study the mechanism of the multi-branch structure, we modify the framework of the multi-branch network with three same networks as the branch and compare different multi-branch networks with the corresponding one branch network to find whether multi-branch structure is better than one branch structure in MI classification tasks. Experiment results revealed that the modified multi-branch 3D CNN can also reach state-of-the-art classification kappa value level and performs better than other methods in terms of standard deviation including the original network with three different branches. The good performance of the modified network may reflect multi-branch structure is better than one branch structure and shows its application value in EEG classification.**

***Index Terms*—**motor **imagery (MI), electroencephalogram (EEG), 3D convolutional neural network (3D CNN), multi-branch structure, one branch structure**

I. INTRODUCTION

The electroencephalogram (EEG) is one of the most significantly and scientifically exploited signals records for humans. In the field of robotics, EEG is widely used to identify human intentions, so as to control the manipulator to help patients or disabled persons enhance their abilities[1]-[2]. Based on the ERS/ERD phenomenon[3] and EEG, the framework for motor imagery (MI) classification tasks is built and a great amount of methods are proposed. Conventional methods to extract features [4]-[6] are common spatial pattern (CSP) method [7] and a series CSP-based methods which have been applied in classifying two kinds imagery successfully[8]-[9]. To improve the performance of CSP method, the subband CSP (SBCSP) [10] method is proposed to for searching the optimal frequency range. And based on the SBCSP method, popular bilinear methods [11] were proposed including discriminative filter bank CSP (DFBCSP) which can optimizes the filters in spatial and temporal dimensions simultaneously. Several other methods different to CSP-based methods also perform well on MI-classification tasks such as the Principal Component Analysis (PCA) transform [12]. The PCA transform reduces the dimension of EEG data to release resources and retains the difference of data features.

Although the methods mentioned above can complete MI-classification well, they all lack much information resulting from their frameworks which complete feature extraction and classification in two steps and even produce redundant features [13] because of a series of linear transformations in CSP method. In contrast, deep learning methods [14] show great advantages with a concise and integrative framework which can optimize the parameters of feature extraction model and classifier jointly. It is also predictable for deep learning methods to benefit from advances in computer-relevant areas.

To complete MI classification tasks with deep learning methods, it is necessary to represent EEG data to a processable form [15] and preserve multi-dimension information of EEG data. Consequently, methods to represent data with more dimensions are proposed such as the method in [16] which preserves the spectral, temporal and spatial structure of the raw EEG data. The method in [17] revealed that CNN indeed learned to use spectral power modulations and performed effectively in spatially mapping the learned features which indicates that spatial features are significant for EEG classification.

On the basis of the representation methods above, it is feasible to apply deep learning methods to extract features when completing MI classification tasks. In [18], a deep recurrent neural network (RNN) is applied to classify MI and outperforms other methods because of its spatial-frequency-sequential relationships and cropped training strategy[17]. We had proposed a novel architecture that combines multi-branch 3D CNN and 3D representation to achieve a better performance[19]. Branches of this kind of structure are CNNs with different depths and the EEG data is transformed into a 3D array which contains the information of three dimensions. The advantage of this method is spatial and temporal features can be extracted simultaneous so as to fully make use of the relationship between them.

Indeed, multi-branch structure has been used widely in neural networks and proved to be useful in dealing with JPEG format image [20] which has a high signal/noise ratio in contrast with EEG data with a low ratio. In consequence,

This work was supported in part by the National Nature Science Foundation of China No.U1713218，U1613224， and 61375090. (Corresponding author: Shaolong Kuang)

H. Zhou, X. Zhao, H. Zhang, S. Kuang are with the School of Mechanical and Electrical Engineering, Soochow University, Suzhou, 215000, China (e-mail: slkuang@suda.edu.cn).

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

we wonder if the multi-branch structure is authentically effective to EEG signals and whether the structure proposed in [19] is the best for EEG-based MI classification. To study the mechanism behind the multi-branch structure, the framework of multi-branch 3D CNN is modified by constructing a 3D CNN with three same branches considering that more branches influence real-time performance according to current computing power. Small-size convolution filters [21] are adopted instead of big ones to decrease the number of parameters in the case of achieving same receptive field. The structure is evaluated with the BCI competition IV-2a data set [22]. Then we compare it with the CNNs in [19] to find whether multi-branch network is better than one branch network and which multi-branch network is better in MI classification tasks.

In this paper, the structure and the mechanism behind it will be highlighted. In Sec. II, the method proposed for MI classification is described. The experiment and results are reported in Sec. III. We conclude the study in Sec. IV.

II. METHODS

In [19], the raw EEG data was transformed into a 2D array and then the 2D array was expended to 3D array by adding the temporal information to acquire a 3D representation. The spatial distribution of electrodes was taken into consideration when we arranged the 2D array and at the same time the point with no electrode was padded to construct a rectangle for ease of convolution.

It is obvious that the temporal features of EEG data are preserved in this type of representation which contains the sequential data of EEG signals. Furthermore, under the condition in which EEG data can be processed, the spatial features of EEG data is preserved to some extent in this representation which is formed according to the distribution of electrodes.

To achieve a better performance, a multi-branch 3D CNN is constructed accordingly, made up of three CNNs with different sizes of receptive field. These networks are respectively named as small receptive field network (SRF), medium receptive field network (MRF) and large receptive field network (LRF). To increase training examples, a cropped training which is described in [17] is taken. When it comes to network optimization, the way to initialize the weights is a normalized initialization method [23] and the optimization method is the ADAM with default parameters values. To improve the test accuracy, a cropped strategy is also adopted. In each trial, the classification results predicted of all cropped EEG data [24] were calculated and then summed together to be regarded as the final result. Additionally, to enhance the real-time capability of the multi-branch network in test process, the cropping stride was set at 5 to reduce the computation time.

To study the mechanism behind the multi-branch structure, we evaluated the performances of three different multi-branch 3D CNNs which are CNNs with three small receptive field networks (SRFs), MRFs and LRFs. Other structure such as sharing their first convolutional layer and summing their respective soft-max results to another soft-max layer were preserved. The detailed information about the framework of multi-branch 3D CNN is shown in Figure 1. The modified structures are shown in Figure 2, Figure 3 and Figure 4.

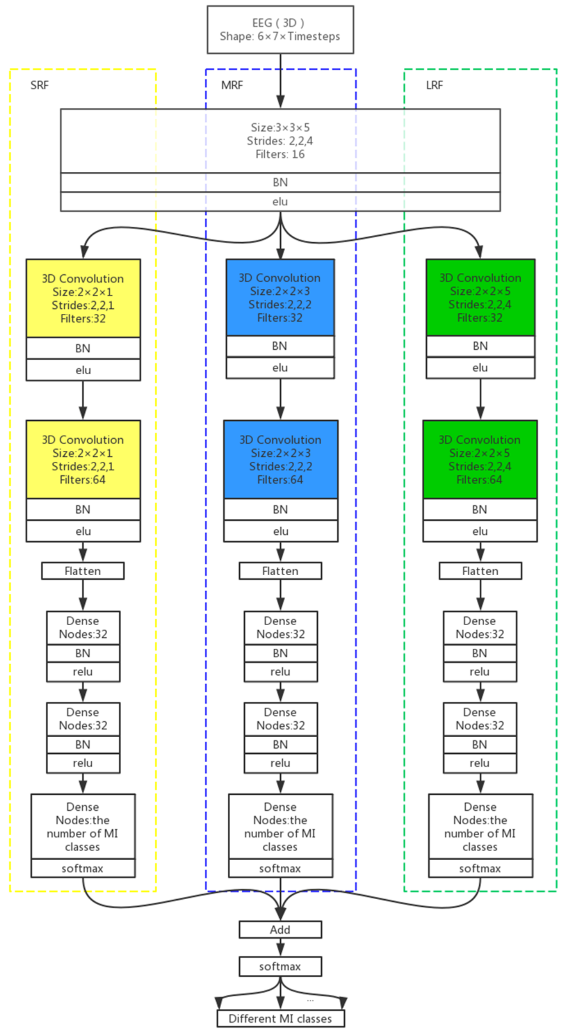

Figure 1. The architecture of multi-branch CNNs. Zero-padding strategy is adopted throughout the convolutional processes. 'Shape' means the width, height and depth of the input array. 'Size' represents the width, height and depth of the 3D convolution window. 'Strides' means the convolutional strides along each dimension. 'Filters' means the number of output filters in the convolution. 'BN' means batch normalization layer. 'Nodes' means the dimensionality of the dense layer's output space. Three branch networks are distinguished with three different colors.

III. EPERIMENTS AND RESULTS

The performance of the method was evaluated on the dataset 2a of BCI competition IV. The EEG data was composed of 4 classes of motor imagery respectively named as left hand, right hand, foot and tongues motor imagery. In two sessions, nine subjects were recorded with twenty-two electrodes placed according to the international 10-20 system. Each session contained two hundred and eighty-eight trials, one session for training and the other for evaluation. The signals were sampled at 250 Hz and then addressed with a band-pass filter between 0.5 Hz and 100 Hz. To suppress line noise, an additional 50Hz notch filter

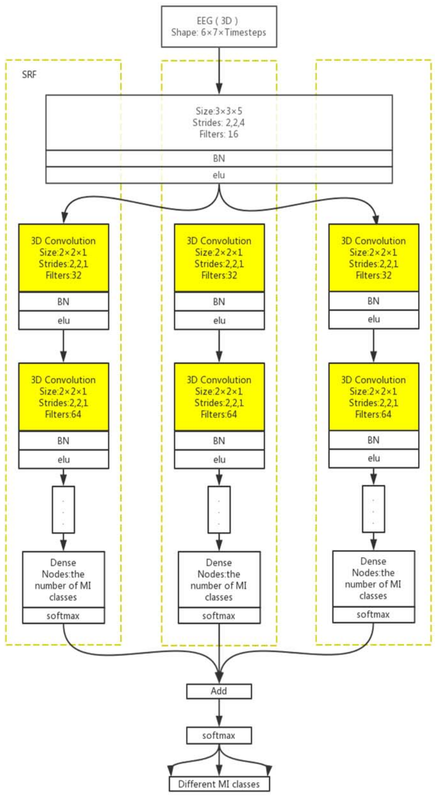

Figure 2. The multi-branch CNNs with three SRFs

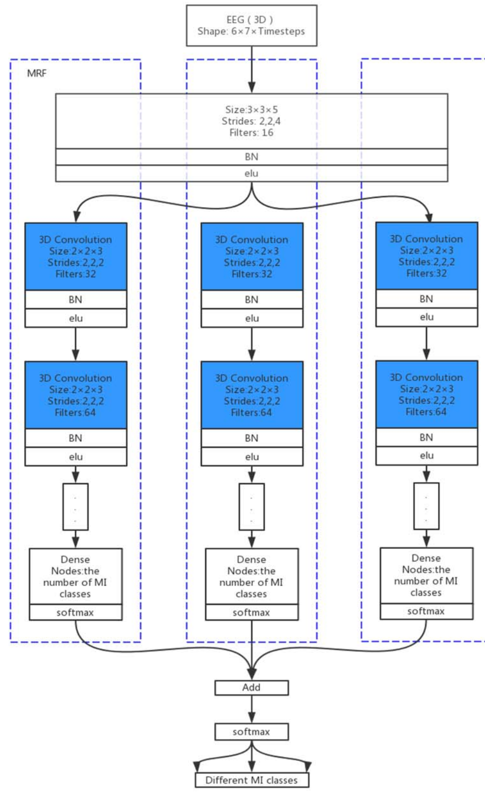

Figure 3. The multi-branch CNNs with three MRFs

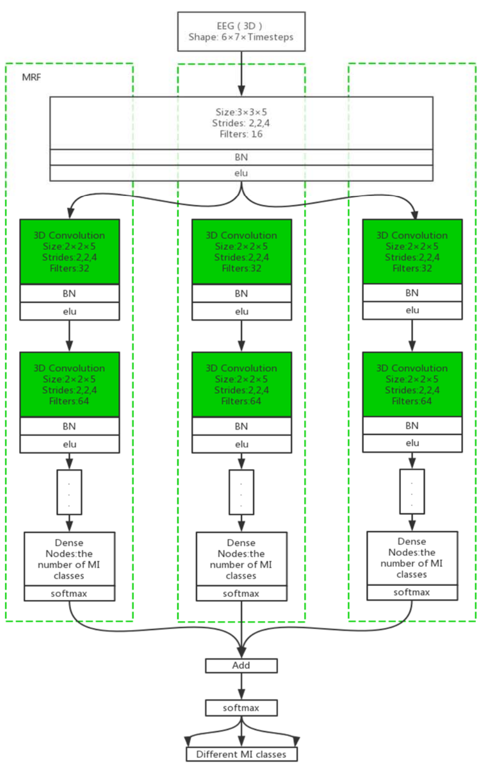

Figure 4. The multi-branch CNNs with three LRFs

was adopted. In each trial, 2s for fixation, 1.25s cue time, and then 4s for MI classification was cost.

In our study, the EEG data of 1.25s period after the visual cue was taken as the experimental data. With the representation method proposed above, these data were then addressed for convolution without any further preprocessing. In the process of training each cropped EEG data, the one-hot-vector method was taken to present corresponding label.

In the step of evaluation, the 10-fold cross-validation was taken by dividing the dataset 2a of BCI IV into ten equal subsets randomly. In each run, nine subsets were selected randomly for training and the rest was for validation. By averaging fifty results with different initializations, the final accuracies were achieved and then reported in the following evaluations. To find the significant difference between each kind of network, the p-values calculated from two-tailed paired t-test are shown in Table I. The evaluations of SRFs, MRFs, LRFs and original Multi-branch architecture proposed above are shown in Table I. Performance of one branch network such as SRF is also presented in this table.

In our former study [19], we have concluded that the performance of different single branch network such as SRF on the same subject is different with each other. In this study, however, the performance of SRFs, MRFs and LRFs is similar on the same subject. Comparing a single branch network and the corresponding multi-branch network, such as SRF and SRFs, the SRFs is 0.373% higher than SRF on

Table I

COMPARISON OF CROSS-VALIDATION RESULTS OF NETWORKS WITH DIFFERENT ARCHITECTURES

Subject	Multi-branch 3D CNN	SRFs 3D CNN	MRFs 3D CNN	LRFs 3D CNN	SRF 3D CNN	MRF 3D CNN	LRF 3D CNN
1	77.397	73.960	75.342	72.975	73.738	76.112	74.743
2	60.140	61.438	60.614	60.494	57.079	58.048	54.936
3	92.927	79.038	79.928	79.173	81.172	82.559	80.405
4	72.288	71.201	71.311	69.778	70.439	70.748	66.484
5	75.836	68.843	68.498	68.136	73.990	75.117	72.173
6	68.988	68.634	69.755	67.515	66.355	67.402	62.711
7	76.036	75.701	75.397	74.659	75.265	73.637	70.754
8	76.855	77.857	77.625	76.815	76.425	75.226	74.369
9	84.665	83.147	84.386	83.640	81.996	82.647	82.188
Mean	75.015	73.313	73.651	72.577	72.940	73.500	70.974
Standard deviation	7.344	6.553	7.013	6.96	7.667	7.597	8.570
p-value (Multi-Branch)	-	9.04E-02	1.54E-01	1.99E-02	3.43E-04	1.14E-04	6.18E-05
p-value (SRFs)	9.04E-02	-	2.53E-01	8.01E-03	6.89E-01	8.65E-01	8.43E-02
p-value (MRFs)	1.54E-01	2.53E-01	-	3.65E-03	4.65E-01	8.86E-01	4.51E-02
p-value (LRFs)	1.99E-02	8.01E-03	3.65E-03	-	6.86E-01	3.89E-01	1.81E-01
p-value (SRF)	3.43E-04	6.89E-01	4.65E-01	6.86E-01	-	2.19E-01	1.38E-02
p-value (MRF)	1.14E-04	8.65E-01	8.86E-01	3.89E-01	2.19E-01	-	8.08E-04
p-value (LRF)	6.18E-05	8.43E-02	4.51E-02	1.81E-01	1.38E-02	8.08E-04	-

p-value (“NETWORK”) means the significant difference (p-value) of network performance between 3D CNN network with “NETWORK” architecture and each other 3D CNN network with different architecture.

the mean value and 1.134% lower than SRF on the standard deviation. Other multi-branch networks also outperform the corresponding single branch networks to some extent and the degree of improvement is different such as LRFs is 1.603% much higher than LRF on the mean value. The analysis may reflect the multi-branch structure is better than single branch structure.

Comparing the performance of multi-branch network with three same branches and the original multi-branch network, in many circumstances, the Multi-branch 3D CNN performs best on all trials which can be seen obviously in subject 3. When it comes to standard deviation, however, the Multi-branch 3D CNN performs no better than other multi-branch networks, such as SRFs, which is 0.791% much lower than the Multi-branch 3D CNN. From an overall perspective, however, all multi-branch networks achieved higher mean and much lower standard deviation than corresponding single branch networks. The results may reflect multi-branch structure is more robust and more effective than one branch network.

IV. CONCLUSION AND FUTURE WORK

The comparative analysis above indicates multi-branch structure achieves a better result than one branch network and the network with three different branches outperforms the one with three same branches. The multi-branch structure is effective in dealing with EEG data and it has priority in MI classification tasks which contributes to the development of rehabilitation robot. A network with more branches should be construct to evaluate the mechanism behind the multi-branch structure in future work.

.

REFERENCES

[1] Sun, Aiqin , B. Fan , and C. Jia . "Motor imagery EEG-based online control system for upper artificial limb." International Conference on Transportation IEEE, 2012.

[2] Gandhi, Vaibhav, et al. "EEG-Based Mobile Robot Control Through an Adaptive Brain–Robot Interface." IEEE Transactions on Systems Man & Cybernetics Systems 44.9(2014):1278-1285.

[3] Neuper, C , Wörtz M, and G. Pfurtscheller . "ERD/ERS patterns reflecting sensorimotor activation and deactivation. " *Progress in Brain Research* 159.1(2006):211-222.

[4] Nixon, Mark, and A. S. Aguado. "Feature Extraction and Image Processing." *Feature Extraction & Image Processing for Computer Vision* (2002):37–82.

[5] Udhaya Kumar, S. , and H. Hannah Inbarani . "PSO-based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task." *Neural Computing and Applications* (2016).

[6] Hsu, Wei Yen . "EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features." *Journal of Neuroscience Methods* 189.2(2010):295-302.

[7] Guger, C. , H. Ramoser , and G. Pfurtscheller . "Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI). " Rehabilitation Engineering IEEE Transactions on 8.4(2000):447-456.

[8] Ricardo Ron-Angevin…. "Brain-Computer Interface application: auditory serial interface to control a two-class motor-imagery-based wheelchair." *Journal of NeuroEngineering and Rehabilitation* 14.1(2017).

[9] Wang, Hong Tao , Y. Q. Li , and T. Y. Yu . "Coordinated control of an intelligentwheelchair based on a brain-computer interface and speech recognition." *Journal of Zhejiang University SCIENCE C* 15.10(2014):832-838.

[10] Lin, Chin Teng , et al. "A Novel Mechanism to Fuse Various Sub-Aspect Brain-Computer Interface (BCI) Systems with PSO for Motor Imagery Task." *2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* IEEE, 2015.

[11] H. Higashi and T. Tanaka, "Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification," IEEE Trans. Biomed. Eng., vol. 60, no. 4, pp. 1100–1110, Apr. 2013.

[12] Geethanjali, and P. "Comparative study of PCA in classification of multichannel EMG signals." *Australasian Physical & Engineering Sciences in Medicine* 38.2(2015):331-343.

[13] Cecotti, Hubert , and A. Graser . "Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.3(2011):433-445.

[14] Philipp, Kainz , P. Michael , and U. Martin . "Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization." *PeerJ* 5(2017):e3874-.

[15] Weiss, S, and P. Rappelsberger. "Long-range EEG synchronization during word encoding correlates with successful memory performance. " *Cognitive Brain Research* 9.3(2000):299-312.

[16] Bashivan P, Rish I, Yeasin M, et al. Learning representations from EEG with deep recurrent-convolutional neural networks[J]. arXiv preprint arXiv:1511.06448, 2015.

[17] Schirrmeister, Robin Tibor , et al. "Deep learning with convolutional neural networks for decoding and visualization of EEG pathology." *Human Brain Mapping* 38.11(2017):5391-5420.

[18] Ma, Xuelin, et al. "Improving EEG-Based Motor Imagery Classification via Spatial and Temporal Recurrent Neural Networks." 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018.

[19] Xinqiao Zhao,Hongmiao Zhang; Guilin Zhu, Fengxiang You, Shaolong Kuang, A multi-branch 3D convolutional neural network for EEG-based motor imagery classification. IEEE Transactions 2019, 10.1109/TNSRE.2019.2938295 .Available online:https://ieeexplore.ieee.org/document/8820089(25,9,2019)

[20] Li, Bin , et al. "A multi-branch convolutional neural network for detecting double JPEG compression." (2017).

[21] Jiang, Wenbin , et al. "A Novel GPU-Based Efficient Approach for Convolutional Neural Networks with Small Filters." Journal of Signal Processing Systems 86.2-3(2017):313-325.

[22] Ang, K. K., et al. "Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b." Frontiers in Neuroscience 6(2012):39.

[23] Uličný, Matej, J. Lundström, and S. Byttner. "Robustness of Deep Convolutional Neural Networks for Image Recognition." International Symposium on Intelligent Computing Systems 2016.

[24] Luo T, Chao F. Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network[J]. BMC bioinformatics, 2018, 19(1): 344.