Kent Academic Repository

Full text document (pdf)

Citation for published version

Mzurikwao, Deogratias and Williams Samuel, Oluwarotimi and Grace Asogbon, Mojisola and Li, Xiangxin and Li, Guanglin and Yeo, Woon-Hong and Efstratiou, Christos and Siang Ang, Chee (2019) A Channel Selection Approach Based on Convolutional Neural Network for Multi-channel EEG Motor Imagery Decoding.   In: 2019 IEEE Second International Conference on Artificial

DOI

https://doi.org/10.1109/AIKE.2019.00042

Link to record in KAR

https://kar.kent.ac.uk/76319/

Document Version

Author's Accepted Manuscript

A Channel Selection Approach Based on Convolutional Neural Network for Multi-Channel EEG Motor Imagery Decoding

Deogratias Mzurikwao
School of Engineering and Digital Arts
University of Kent
Canterbury, U.K.
dm521@kent.ac.uk

Oluwarotimi Williams Samuel
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences,
Shenzhen, CHINA
samuel@siat.ac.cn

Mojisola Grace Asogbon
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences,
Shenzhen, CHINA
grace@siat.ac.cn

Xiangxin Li
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences,
Shenzhen, CHINA
lixx@siat.ac.cn

Guanglin Li
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences,
Shenzhen, CHINA P.R
gl.li@siat.ac.cn

Woon-Hong Yeo
Institute for Electronics & Nano, Flexible, Hybrid Electronics
Georgia Institute of Advanced Technology
Atlanta, U.S.A
whyeo@gatech.edu

Christos Efstratiou
School of Engineering and Digital Arts
University of Kent
Canterbury, U.K.
c.efstratiou@kent.ac.uk

Chee Siang Ang
School of Engineering and Digital arts
University of Kent
Canterbury, U.K.
c.s.ang@kent.ac.uk

Abstract— **For many disabled people, brain computer interface (BCI) may be the only way to communicate with others and to control things around them. Using motor imagery paradigm, one can decode an individual's intention by using their brainwaves to help them interact with their environment without having to make any physical movement. For decades, machine learning models, trained on features extracted from acquired electroencephalogram (EEG) signals have been used to decode motor imagery activities. This method has several limitations and constraints especially during feature extraction. Large number of channels on the current EEG devices make them hard to use in real-life as they are bulky, uncomfortable to wear, and takes lot of time in preparation. In this paper, we introduce a technique to perform channel selection using convolutional neural network (CNN) and to decode multiple classes of motor imagery intentions from four participants who are amputees. A CNN model trained on EEG data of 64 channels achieved a mean classification accuracy of 99.7% with five classes. Channel selection based on weights extracted from the trained model has been performed with subsequent models trained on eight selected channels achieved a reasonable accuracy of 91.5%. Training the model in time domain and frequency domain was also compared, different window sizes were experimented to test the possibilities of realtime application. Our method of channel selection was then evaluated on a publicly available motor imagery EEG dataset.**

Index Terms— *BCI, CNN, EEG, Feature maps, Motor imagery, Topographic maps.*
1

I. INTRODUCTION

Brain Computer Interface (BCI) is a technology that allows direct communication between the human brain and an external device [1]. In general, BCI systems has a hardware component incorporate sensors for acquisition of signals generated from brain activity and software component that is composed of computational model (e.g. machine learning model) aimed at decoding the acquired brain signals. The controlled part of the BCI can be hardware such as a prosthetic arm or software such as controlling a cursor on a computer display [2] or playing computer games [3-4]. A standard BCI system implementation involves a sequenc of steps including signal acquisition, data preprocessing or signal enhancement, feature extraction, classification, and the control interface [5]. In the recent years, BCI related research has enticed a lot of attention in industrial applications as well as academia . In fact, for individuals who lack the ability to communicate via conventional means as a result of severe motor disabilities (like spinal cord injuries or amyotrophic lateral sclerosis (ALS) disease), BCI may be a

viable means of re-establishing their communication pathways [6]. Though high level amputees could only produce limited neuromuscular information from their phantom limbs, they have the ability to generate rich set of motor information from their brain when they imagine to move their phantom limbs. Thus, this makes them rely on BCI as a viable means of communication.

Many BCI decoding technologies are centered on intra-cortical implanted microelectrode arrays that measures movement-related signal from the brain[3]. However, this technique is invasive and involves a surgery which may lead to post-surgery complications and infections. Other techniques to acquire brain signal include Functional Near-Infrared Spectroscopy (fNIRS), Positron Electron Tomography (PET), Electroencephalography (EEG), Magnetoencephalography (MEG), and functional Magnetic Resonance Imaging (fMRI) [6]. Amongst these techniques, EEG approach is one of the most common method for non-invasive BCI as data can be seemlessly obtained from the human scalp. Importantly, it should be noted that most of the available EEG devices are bulky, with lot of channels and consumes time to wear. These factors make them hard to use in daily life applications.

In recent years, we have seen the development of more portable, low cost, and flexible EEG skin-like sensor such as the Graphene-based epidermal sensor system (GESS) [7]. For these portable devices to work practically in everyday environment, we need to restrict the number of channels significantly, to reduce costs, reduce setup time, and to improve usability and user acceptance. Previous EEG channel reduction methods have been investigated to be inefficient [8]. Channel reduction is a challenging task, because when fewer channels are used, less spatial information is captured, making it difficult to detect artifacts as spatial information obtained in electrode placement are significant for accurate detection of artifacts in the collected EEG data [9].

II. RELATED WORK

Machine learning techniques have been widely used to perform classification of acquired EEG signal for BCI applications [6]. Wang at al. used Fisher Linear Discriminant (FLD) and Support Vector Machine (SVM), to decode motor imagery EEG signals to control a humanoid robot [9]. To perform classification using conventional machine learning algorithms, relevant features must be extracted from EEG recordings to train a built classification model. Since these features are manually extracted from the EEG signals, determining the most representative feature set is still a major challenge. Over the past decades, several handcraft techniques have been used for EEG feature extraction such as standard deviation (SD), variance, Fast Fourier transforms (FFT) [11], Wavelet transform (WT) [12], and power spectral density (PSD) [13]. Factors such as motion artifacts [14], electrooculography (EoG) [15], and EMG interference inherent in EEG recordings makes it even difficult to extract accurate and robust set of features. In motor imagery classification tasks specifically, the imagined movement is often lost in this mixture of signals [16]. In addition, the choice of good discriminative features to train a conventional machine-learning model is time consuming, and requires experts in the field. If feature extractions are not performed well, it may result in consistently low performance and non-robust model for EEG motor imagery decoding [17]. Indeed, the performance of a BCI system using motor imagery is greatly depending on how features are extracted [18]. Additionally, the non-stationary nature of EEG and the amount of data generated with fewer electrodes makes it impractical to perform real time classification by manually extracting features [19]. These limitations of the conventional machine learning based BCI systems necessitate the need for an alternative approach to developing EEG driven BCI systems. To address these limitations, some authors have looked into neural network approach to learn features automatically from the data. However, weights initialization associated with the traditional neural network is a complex task as small weights initialization leads to weight diffusion and large weights leads to poor local minima [20].

Recent development in deep learning has offered a viable approach to extract features automatically through a deep layer of hidden units, which can address the drawbacks of using conventional machine learning. Deep learning models are capable of detecting spatial structure features on a given dataset without the need of handcrafting of features. Deep learning has been reported to have capabilities of detecting necessary features despite of external/unwanted signals interference [21], hence it is becoming a favorable technique for classification of biomedical data including EEG signals which typically contain lot of interference. As significant features are assigned higher weights during the training of a deep learning model, this gives a chance to identify significant channels in EEG for channel reduction.Interestingly, deep learning has been increasingly applied in classification and analysis of EEG signals. Hurbert Cecotti *et al* [6] managed to detect p300 in EEG signals in time domain using deep learning. In healthcare domain, Deep learning has been used to classify EEG patterns of Alzheimer's disease from a prodromal version of dementia using raw EEG signals by Francesco C.M *at al* [17]. Thodorof *at al* [22] used recurrent convolutional neural networks, for automatic seizures detection by capturing spatial, spectral and temporal information from EEG signals. Deep learning-based EEG analysis has also found its application in biometrics, where Mao *at al* [23] successfully performed biometric identification from EEG signals.

In this study, a Convolutional Neural Network (CNN), a type of deep learning algorithm, was implemented to decode multiple classes of imagined upper limb movement from raw EEG signals acquired from transhumeral amputees without the convetional handcrafting of features. This is among the few datasets obtained from real amputees. The possibility of applying deep learning methods for real time application was investigated by experimenting with smaller window sizes. Subsequently, channels reduction was extensively explored using the developed CNN model to reduce the number of EEG channels to the lowest possible number while maintaining good classification accuracy. The best locations for fewer channels' placement on the scalp of the participants corresponding to the multi-class motor imagery tasks were also identified and reported. The channel selection method was then tested on a widely used publicly available motor imagery dataset,

commonly know as the physionet [27] to further prove the effectiveness of our proposed method. The contribution of this paper is on the introduction of a new and more efficient way for EEG channels selection, a method for determining the brain location for motor imaginery activity using topographic maps and the possibility of real time decoding of motor imaginery activity.

III. DATA COLLECTION

A. Study description

Four male transhumeral amputees (with average age of 41.50±7.05 years and mean residual limb of 25.50±4.20 cm as measured from the shoulder blade downwards) were recruited. The participants willingly gave permission for the publication of their photographs/data for scientific and educational purposes in a written approval. The Institutional Review Board of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China, approved all the protocols.

B. Equipment setup and data acquisition

For data collection, we utilized a commercial EEG signal acquisition system (EasyCap, Herrsching, Germany) incorporated with the Neuroscan software (version 4.3), sampled at 1000Hz. The equipment has 64-channels of Al-AgCl electrodes on a cap. The electrodes were distributed over the scalp of a participant based on the 10-20 international system standard. Prior to the placement of the electrode cap, each participant's hair was properly washed to ensure that high signal quality is acquired. In addition, the impedance between each electrode and the scalp was maintained below 10 kΩ. After setting up the signal acquisition system, participants were asked to sit straight facing a computer screen where five different motor imagery (MI) tasks were displayed. The five MI tasks includes hand close (HC), hand open (HO), wrist pronation (WP), wrist supination (WS), and no movement (NM). Each MI task was displayed for 5 seconds. Participants were asked to perform the action contionously through out the 5 seconds period. To avoid mental fatigue, which often affect the quality of the EEG signals, a rest session of 5 seconds was introduced between any two consecutive MI tasks. To acquire a significant amount of dataset for training, validation, and testing of the proposed CNN model, each participant accomplished five experimental trials. In every trial, each MI task was repeated ten times thus producing 50 seconds data recordings per MI task per trial. Fig. 1 shows the experiment setup.
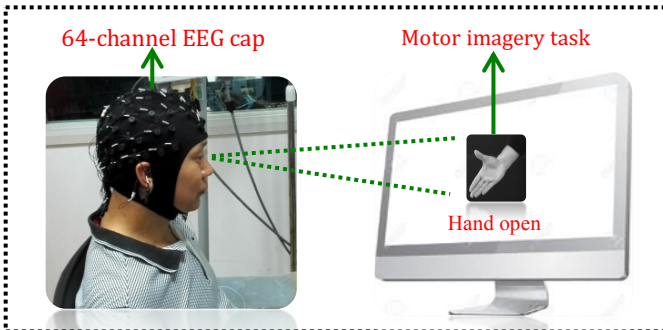

Figure 1. Experiment setup adopted for data collection

IV. METHODOLOGY

A. Data preprocessing

The data set is in time domain and considering each participant's data, a moving window (non-overlapping) was used to slice the input resulting into samples of dimension (W_l, 64). Different window length W_l were tested to imitate the possibilities of real time classification. 2-D image like, were the width (64) represents the spatial structure whilst the height (W_l) represents the temporal structure[20]. For training purpose, 80 % of the dataset were used, while the remaining 20% were used for testing. Batches of 128 were fed into CNN model.

B. CNN architecture

Generally, CNN has two main parts that includes the feature extractor and the trainable component. The feature extractor part contains multiple layers of convolution and pooling. The convolution layer is able to learn and extract features from the raw data automatically and the pooling layer is used for down sampling. The trainable part contains fully connected multilayer perceptron, which perform classification based on the features learned in the feature extractor part [17].
The CNN was trained, validated, and tested with each participant's data separately. As real time application is one of the main aims of this study, small and simple architecture was preferred out of other tested models as smaller models are computionally affordable to run on low cost smart devices during real time appications.

The network had two convolutional layers, and each layer had filters of size (5, 5) , with a stride of 1, to form feature maps after dot product between the filter and the input. These filters are aimed to capture different local spatial, spectral, and temporal patterns/features related to imagined motor activity [22]. 32 and 64 filters were used in convolutional layer 1 and convolutional layer 2 respectively, hence 32 and 64 feature maps were formed in the first and second convolutional layers respectively. After every convolution layer, ReLu activation function was applied. That is, Let W represent the filter weights and W_k denotes the weight of filter k, were $k = 1, ...32$ for first convolution layer and $k = 1, ...64$ for second convolution layer. Let $V \in P^{M*1}$ denote a vector input with $M = 500 * 64$. W is a hyper-parameter to be learned [23]. The feature map output in the second convolution layer is as shown in (1):

$$Convolution(V)_k = RELU(W_k V_k) \qquad (1)$$
$$for\ k = 1, ...64$$

Max pooling was used after every convolutional layer as it is considered to better improve the network performance [17].
One fully connected layer was used to perform classification. The output layer had five nodes with soft-max nonlinearities and the output class was assigned to the one with maximum among the five. The fully connected layer combined all the feature maps at the output of the last convolution layer. To avoid overfitting, a dropout regularization technique was applied on the fully connected layer. Fig.2 shows the CNN architecture.
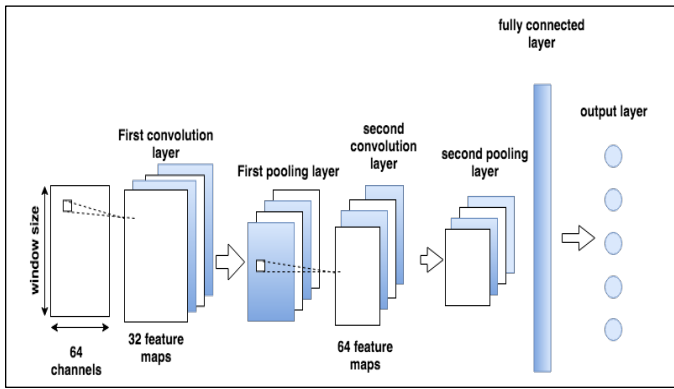
Fig.2 The proposed CNN architecture

V. CLASSIFICATION RESULTS

With the proposed CNN model, an overall average accuracy of 99.7% was recorded with a 0.5 second moving window (no overlap), as the performance metrics shown in Table 1 across all participants. Accuracy was measured as a percentage of total number of samples correctly classified over the total number of samples classified. Different window sizes were used to test the possibilities of real time classification. As shown in Table 1, one second window and two seconds window were tested, and the implement model maintained a highly consistent classification accuracy for both windows with lower optimization speed as larger windows contain more information to learn.

Table 1
Window Performance metrics for CNN

Window size (seconds)	Average Accuracy (%)	Average precision (%)	Average Recall (%)	Average F1 score (%)
0.5	99.0	99.0	100.0	99.0
1.0	99.0	98.0	99.0	99.0
2.0	99.0	100.0	100.0	99.0

The learning curves for all four participants, for 0.5 Second, 1 Second and 2-seconds window sizes are shown in Fig 2, 3, and 4 respectively. It took a little bit longer for the model to converge when a smaller window was used as it can be seen in Fig.3, 4, and 5. One possible reason for this observed phenomenon is because the model has relatively few information to learn from a small window data points.
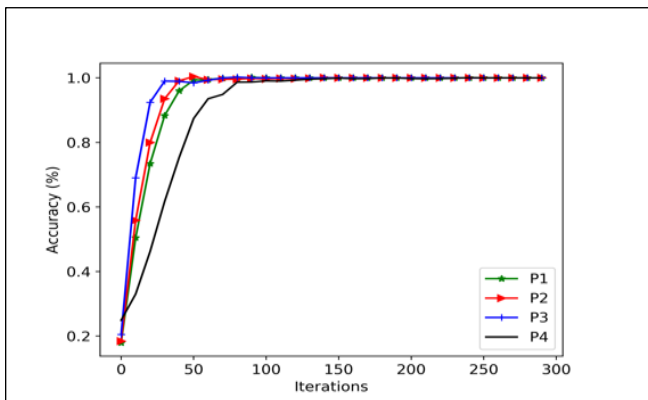

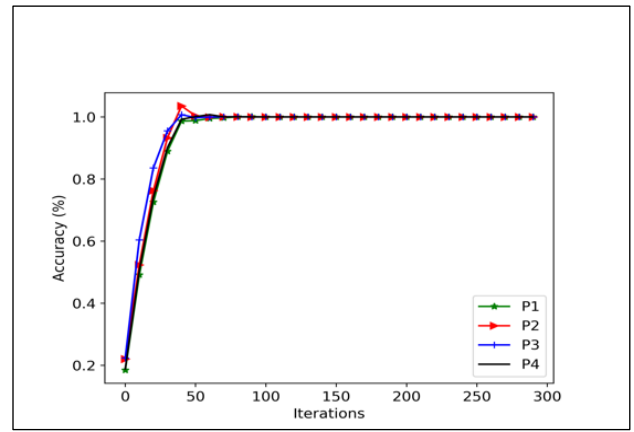Fig.3 Learning curve for 0.5 second window length

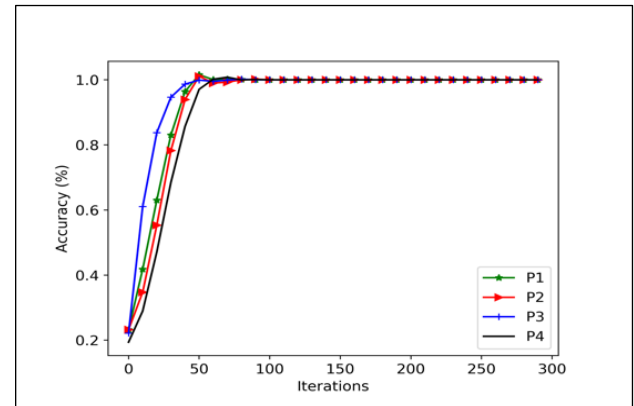Fig.4. Learning curve for 1 second window length

Fig.5. Learning curve for 2 seconds window length

Notably, a deep learning model can be trained with time domain or frequency domain data. Transforming the obtained data from time domain to frequency domain requires preprocessing the obtained data with fast Fourier transform (FFT). FFT adds additional computational costs, hence classifying the obtained data in time domain will avoid this computational expense especially when real time classification is being considered. As shown in Table 2, no much additional performance was gained when the model was trained with frequency domain data across participant, and the learning curve was shown in figure 6. As it can be seen in the learning curves in time domain (figure 2-4), our model generalized well between 50-100 iterations across all participants while it took more than 100 iterations to generalize in frequency domain.

Table 2

Performance metrics in frequency domain

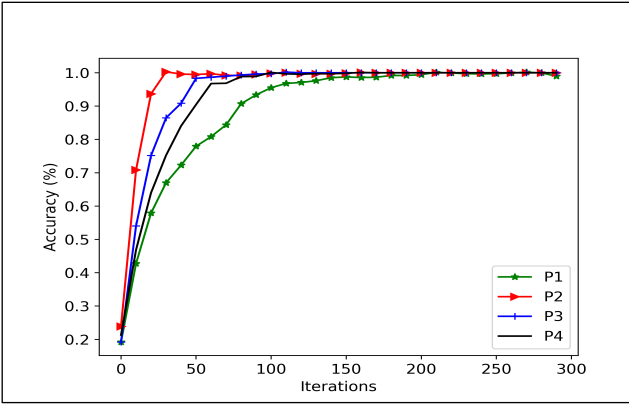| Participants	P1	P2	P3	P4
Accuracy(%)	100.0	99.0	99.0	99.0
Precision(%)	100.0	100.0	99.0	100.0
Recall(%)	100.0	100.0	100.0	100.0
F1_Score(%)	100.0	100.0	99.0	100.0

Fig. 6. Learning curve in frequency domain

Table 3. Shows the performance of conventional machine learning models on the same data set as reported in previous work [10], in which spectral domain features extracted from the EEG recordings in frequency domain were used to train conventional machine learning classifiers, i.e. linear discriminant analysis (LDA), artificial neural network (ANN) and k-nearest neighbors (kNN) to decode multiple classes of MI tasks. The conventional machine learning models reported were trained with a 150ms window with an overlap of 50ms. It shows that our proposed CNN model outperformed all previous models.

Table 3.
Accuracy based on extracted features

	LDA	ANN	KNN
Aaccuracy	97.81	96.44	96.92

VI. FEATURE MAPS AND WEIGHT ANALYSIS

A. Topographic maps

To identify the best discriminant channels for EEG classification, we extracted weights from the trained CNN model for each participant and constructed the topographic maps. For weight extraction, to avoid mixing up features contained in different channels, we retrained our model with the same architecture used above but we changed the filters size from a matrix of 5*5 to a vector of 20*1 and stride of (1,1) in order for the filter to convolve along a single channel at a time. We then extracted the weights from the feature maps formed in the first convolution layer. Sum of weights of all the feature maps was obtained as we have 32 feature maps in the first convolution layer. We used the first convolution layer as the feature maps are still in the same shape since the dimension of the channels are still preserved as the original dimension of the input prior to max-pooling. The weights were calculated using the following formula in (2). W_l Stands for window length, which is 500 for this study.

$$W_p = \sum_{q=0}^{q=W_l} W_{(p,q)} \qquad (2)$$

Where $0 \leq p < 64$ and $0 \leq q < W_l$

The topographic maps created from the weights extracted from the trained model shows that the regions with the highest weight are common among participants and they are situated around the motor cortex region of the human brain, as seen in Fig 7. It should be noted that the third amputee reported phantom limb pain (PLP) when he was performing motor imagery task during data collection. Since the left motor cortex region of the human brain controls the right side of the body and his amputated arm was on the right side, we believe that the participant was not performing enough motor imagination due to the pain reported [24]. This can be seen in Figure 7 for the third participant, which shows high weight values on only one side of his brain. And it has been previously reported that amputees with PLP generally have worse motor control over their phantom hand [25-26].

B. Channel reduction and performance analysis

Training a model with all available channels takes time and hence affects the learning speed of the model as it learns from a large number of channels, some of which are not related to the motor activity of the brain. Identifying the necessary channels will help in desing and manufacturing of smaller, more portable and cheaper EEG devices with fewer channels. Using these portable EEG devices with fewer channels, we will need precise placement of the channels on the scalp. From the topographic maps generated based on the extracted weights from the trained CNN model, it can be seen that the channels with the highest weights are around the motor cortex region of the brain, indicating that these channels are the most discriminating ones.

To perform channel reduction, the summed weights obtained from each channel were first arranged in descending order. Then the first twenty (20) channels with the highest weights were selected for each participant. For each participant, the network was then retrained using these 20 channels. We then carried out a systematic analysis, reducing the number of channels down to two (2) in ascending order of the weights across all the participants. Table 4 shows the accuracy results of channels reduction for all participants. The model trained with fewer channels maintained reasonable accuracy when fewer electrodes up to eight channels with higher weights are selected. This is because the models learn from less but significant channels' data. When fewer channels are used, the training time was significantly reduced as now the model learn from fewer but significant channels. Our results show that by selecting fewer channels located around the motor cortex region we can still obtain a reasonably high classification accuracy in comparison to using all the 64 channels.
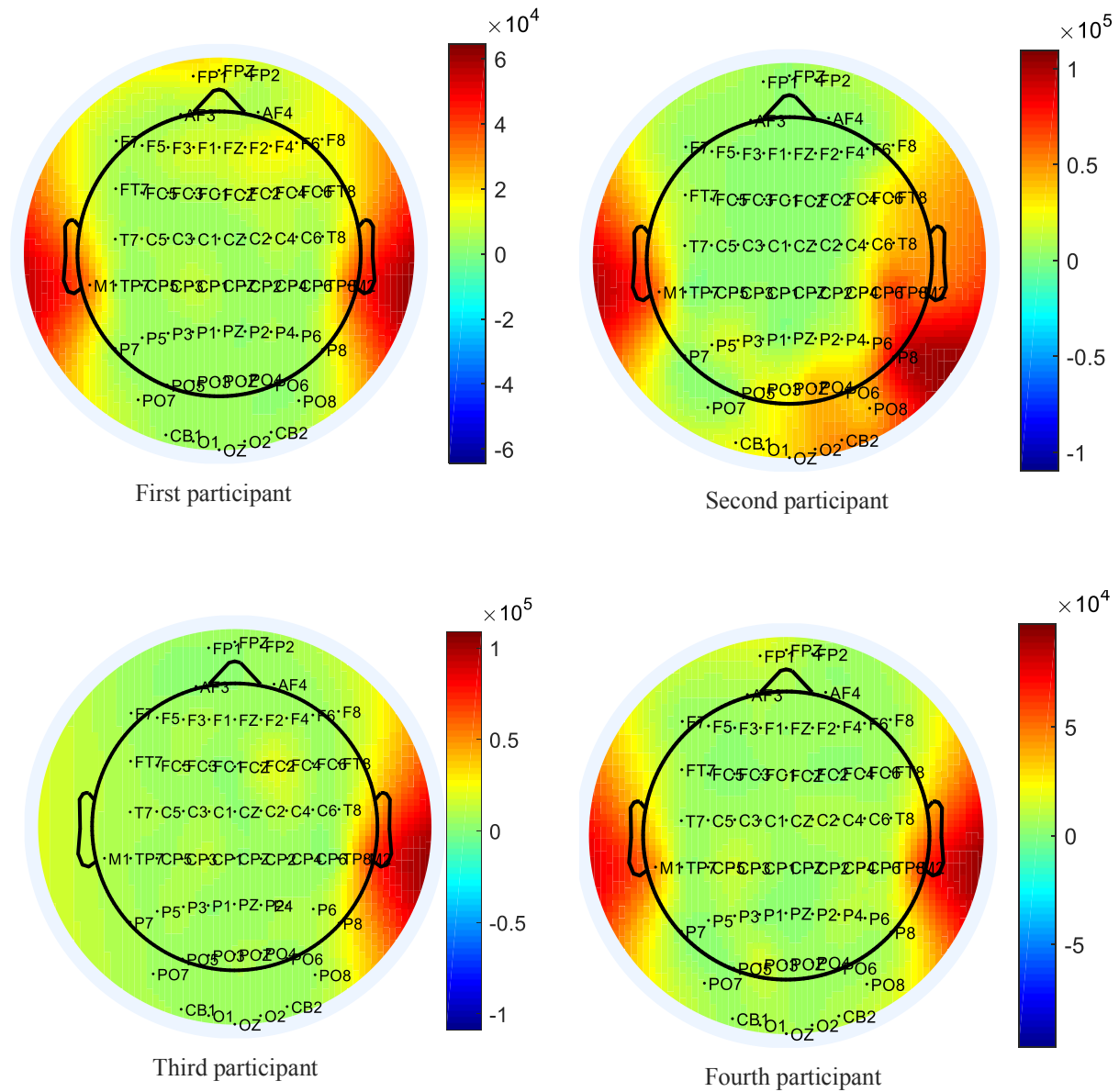
Fig. 7. Topographic maps showing the region of the brain with the most contributive channels of the four participants

Table 4
Channel selection based on our acquired dataset

Participants	Number of Channels and Accuracy(%)						
	64	**20**	**16**	**12**	**8**	**4**	**2**
P1	100.0	99.0	99.0	98.0	92.0	66.0	47.0
P2	99.0	100.0	98.0	97.0	93.0	71.0	86.0
P3	99.0	99.0	99.0	98.0	92.0	91.0	44.0
P4	99.0	98.0	95.0	96.0	89.0	69.0	62.5

VII. EVALUATION OF OUR METHOD ON PUBLICLY AVAILABLE DATA SET

This performance of the proposed CNN model for channel selection was further evaluated on a different EEG dataset which is publicly available from PhysioNet eegmmidb (EEG motor movement/imagery database) database [27]. The data set was collected using a 64 channels BCI200 EEG system at 160Hz sampling rate, with 109 participants. Each participant performed four tasks with a rest between each task. The participants performed three trials while sitting in front of a computer screen. A resting state was included as a class in our case. To perform the tasks, the participants were instructed as explained in the PhysioBank Automated Teller Machine [28]. For our experiment, four participants were randomly selected. 32,000 labelled instances of EEG data was used from each participant.

Accuracy was used as a performance measure for evaluating proposed channel selection method on this data set. The same methodology was used to perform channel selection, as described in section V. We eventually reduced the number of channels from 64 to 12 and even further lower down to 2 based on the computed weights associated with channels. A drop of accuracy in the range of $2 \sim 3\%$ was observed when the channels were reduced from 64 to 12, as seen in Table 5. Meanwhile, at 12-channels, the accuracy seems to be reasonably high and stable and begins to drop when the channel numbers was below 12. Note that the proposed CNN model was trained per instance record of EEG record as it was done in [29], the model learned most of the spatial features and less temporal features. We believe this affected the performance when less than 12 channels were utilized compared to the case of amputees data in which moving window was used.

Table 5
Channel selection based on Physionet data

Participants	Number of Channels and Accuracy(*100)						
	64	20	16	12	8	4	2
P1	96.0	93.0	94.0	92.0	70.0	30.0	24.0
P2	95.0	95.0	94.0	93.0	80.0	29.0	22.0
P3	98.0	96.0	93.0	92.0	79.0	49.0	42.0
P4	99.0	96.0	93.0	90.0	79.0	50.0	22.0

VIII. CONCLUSION

In this study, we demonstrated the use of CNN for decoding of multiple classes of motor imagery activities from raw EEG signals obtained from amputees, avoiding the need of hand crafting features. The proposed CNN possess a number characteristics to addresses the drawbacks found associated with the conventional machine learning methods. The ability of proposed CNN model to learn its own features from raw data provide a significant advantage especially in real world application due to its high performance compared to conventional machine learning which rely on hand-fed features. Handcrafting of features from raw data by human experts can sometimes involve extraction of features which are not related to the classes involved hence adversely affects the model's performance. Furthermore, as no significant difference in model performance was seen between time domain and frequency domain data, this suggest that we can avoid the computational costs and the time of transforming the EEG recordings into frequency domain by utilizing FFT, especially in real time applications.

Using a small moving window of 0.5 second, as shown in this study, allows the possibilities of performing real time classification on low cost smart devices. The size of the networks used with fewer parameters to learn, makes it easy for a pre-trained CNN model to be exported and embedded onto smart mobile devices. This also saves computation cost during training as well as running the trained model. Furthermore, as our model maintained high classification performance with fewer channels, the channel reduction technique with CNN demonstrated that it is possible to use portable EEG devices with fewer channels which can reduce the cost of devices while still achieving reasonably high and consistent accuracy.

Many of the previous methods for EEG channel reduction are performed on the extracted features [8], this will involve manual extraction of features from the obtained data before performing channels selection. This is not the case in our method as all the draw backs and limitations for features extraction are avoided. In addition, our method of channel

selection can be used to locate the important/significant region on a human scalp for a particular activity, providing a systematic guidline for electrode placement for EEG devices, without the need of an expert to locate them. This becomes the case especially when a fewer channel EEG device is to be built and utilized, it will require precise placement of the device.

The main limitation of our method is the pooling layer in CNN. We are currently extracting learned weights from the first convolution layer as the learned features across the channels are intact. The first layers in CNN learn only low level features henced the weights learned for high level features in the deeper layers were not accounted for in performing channels selection. Although this limitation can be avoided by removing the pooling layers, this will results in a high computational coasts. Thus, we hope to conduct further studies in this direction in our future work.

REFERENCES

[1] L. Jingwei, Y. Cheng, and W. Zhang, "Deep learning EEG response representation for brain computer interface," *In Control Conference (CCC), 2015 34th Chinese,* pp. 3518-3523. IEEE, 2015.

[2] A. Kostov and M. Polak, "Parallel man-machine training in development of EEG-based cursor control," IEEE *Transactions on Rehabilitation Engineering* 8, no. 2, pp. 203-205, 2000.

[3] M. Marchesi, "From Mobie to Neu: 3d animated contents controlled by a Brain-Computer Interface." *In Proceedings of the 2012 Virtual Reality International Conference*, pp. 28. ACM, 2012.

[4] N. Kosmyna, F. Tarpin-Bernard, and B. Rivet, "Conceptual priming for in-game BCI training," *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, no. 5 (2015): 26.

[5] O.W. Samuel, X. Li, X. Zhang, H. Wang, and G. Li, "A hybrid non-invasive method for the classification of amputee's hand and wrist movements," *In International Conference on Biomedical and Health Informatics (pp. 161-166). Springer, Singapore.*

[6] H. Cecotti 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), pp.433-445, 2011.

[7] S.K. Ameri, R. Ho, H. Jang, Y. Wang, D. M. Schnyer, D. Akinwande, and N. Lu, "Thinnest transparent epidermal sensor system based on grapheme," *In Electron Devices Meeting (IEDM), 2016 IEEE International*, pp. 18-4. IEEE, 2016.

[8] H. Sun, Y. Zhang, B.J. Gluckman, X. Zhong, and X. Zhang, "Optimal-channel Selection Algorithms in Mental Tasks based Brain-computer Interface". In Proceedings of the 2018 8th International Conference on Bioscience, Biochemistry and Bioinformatics (pp. 118-123). ACM.

[9] F. Wang, C. Zhou, X. Hao, S. Wang, and G. Yang, "BCI control system for humanoid robot based on motor imaginary," In Control and Decision Conference (CCDC), 2013 25th Chinese (pp. 5140-5143). IEEE, 2013.

[10] O.W. Samuel, Y. Geng, X. Li, and G. Li, "Towards Efficient Decoding of Multiple Classes of Motor Imagery Limb Movements Based on Eeg Spectral and Time Domain Descriptors," *Journal of medical systems* 41, no. 12 (2017), pp.194, 2017.

[11] G. Chen, "Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features," *Expert Systems with Applications* 41, no. 5 (2014): pp. 2391-2394, 2014.

[12] S.J. Schiff, A. Aldroubi, M. Unser, and S. Sato, "Fast wavelet transformation of EEG," *Electroencephalography and clinical neurophysiology*, *91*(6), pp.442-455, 1994.

[13] S.M. Zhou, J.Q. Gan, and F. Sepulveda, Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface. *Information Sciences*, *178*(6), pp.1629-1640, 2008.

[14] S.N. Abdulkader, A. Atia, and M.S.M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Informatics Journal, 16*(2), pp.213-230, 2015.

[15] J. Frey, J. Mladenović, F. Lotte, C. Jeunet, and L. Pillette, "When HCI Meets Neurotechnologies: What You Should Know about Brain-Computer Interfaces," In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1253-1256. ACM, 2017.

[16] Wang, Ker-Jiun, Lan Zhang, Bo Luan, Hsiao-Wei Tung, Quanfeng Liu, Jiacheng Wei, Mingui Sun, and Zhi-Hong Mao. "Brain-computer interface combining eye saccade two-electrode EEG signals and voice cues to improve the maneuverability of wheelchair." In *Rehabilitation Robotics (ICORR), 2017 International Conference on*, pp. 1073-1078. IEEE, 2017.

[17] F.C. Morabito et al., "Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings," In *Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), 2016 IEEE 2nd International Forum on* (pp. 1-6). IEEE, 2016.

[18] H.J. Park, J. Kim, B. Min, and B. Lee, "Motor imagery EEG classification with optimal subset of wavelet based common spatial pattern and kernel extreme learning machine," In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE* (pp. 2863-2866). IEEE, 2017.

[19] M.P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for EEG big data and seizure prediction BCI via internet of things," *IEEE Transactions on Big Data*, *3*(4), pp.392-404, 2017.

[20] G.E. Hinton and R.R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, *313*(5786), pp.504-507, 2006.

[21] N.S. Kwak, K.R. Müller, and S.W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PloS one*, *12*(2), p.e0172578, 2017.

[22] P. Thodoroff, J. Pineau, A. Lim, "Learning robust features using deep learning for automatic seizure detection," In *Machine learning for healthcare conference* (pp. 178-190), 2016.

[23] Z. Mao, W.X. Yao, and Y. Huang, "EEG-based biometric identification with deep learning," In *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on* (pp. 609-612). IEEE, 2017.

[24] O.W. Samuel, X. Li, Y. Geng, P. Feng, S. Chen, and G. Li, "Motor imagery classification of upper limb movements based on spectral domain features of EEG patterns," In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, pp. 2976-2979. IEEE, 2017.

[25] S. Kikkert, M. Melvin, D.H. Slater, h. Johansen-Berg, I. Tracey, and T.R Makin, "Motor correlates of phantom limb pain." Cortex 95 (2017):29-36, available: https://www.sciencedirect.com/science/article/pii/S001094521730240X

[26] X. Li, O.W. Samuel, X. Zhang, H. Wang, P. Fang, G. Li, "A motion-classification strategy based on sEMG-EEG signal combination for upper-limb amputees," *Journal of neuroengineering and rehabilitation,* 14(1), 2, 2017.

[27] A.L. Goldberger et al "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, *101*(23), e215-e220, 2000.

[28] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, J.R. Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Transactions on Biomedical Engineering* **51**(6):1034-1043, 2004.

[29] X. Zhang, L. Yao, Q.Z. Sheng, S.S. Kanhere, T. Gu, and D. Zhang, "Converting your thoughts to texts: Enabling brain typing via deep feature learning of EEG signals," In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (pp. 1-10). IEEE, 2018.