HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 4 (2023)

Comparative Analysis of CNNs and RNNs for EEG-Based Motor Imagery Classification in BCIs

Yat Hei Vanessa Lam[1] and Leo Lui[#]

[1]Heep Yunn School, Hong Kong
[#]Advisor

ABSTRACT

EEG-based motor imagery (MI) classification plays a vital role in brain-computer interface systems (BCIs) to enable the control of external devices with the human brain. However, there is currently limited research focusing on the comparison between different machine learning models for this task. This research paper aims to present a comprehensive comparative analysis of two popular deep learning architectures, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for MI recognition with EEG data. The experiments utilised the EEG Motor Movement/Imagery Dataset v1.0.0 from PhysioNet, which contains EEG signals recorded during a variety of motor imagery tasks. The respective performances of the CNN and RNN architectures were subsequently evaluated and compared based on classification accuracy and computational efficiency. Various metrics and statistics, namely accuracy, precision, training speed, memory usage, etc., were used for assessment. The results revealed that CNN outperforms RNN in terms of accuracy, while RNN demonstrates superior computational efficiency. These findings potentially serve as a valuable guideline for researchers and practitioners in the field of BCIs, aiding them in selecting the most suitable neural network architecture for performing MI related tasks.

Introduction

Over the past decade, brain-computer interfaces (BCIs) have emerged as an exciting technology that enables the human brain to directly communicate with and control external devices without the need for traditional motor pathways. BCIs have the potential to revolutionise various domains, including healthcare, rehabilitation, and assistive technology. Hence, one of the key tasks in BCIs is motor imagery (MI) recognition, which involves classifying users' intentions of limb movements based on their brain signals.

Electroencephalography (EEG) is widely used in BCIs as it directly measures and records the electrical activity in the brain. Thus, EEG allows us to capture the neural signals associated with motor planning and execution, and is considered as a popular choice for performing motor imagery recognition. Through the analysis of EEG signals, it is possible to infer some of the intentions of users, for example, their intention to move a specific limb or perform a particular action. Moreover, EEG's non-invasive, relatively affordable, and portable nature makes it a practical choice of imaging technique for consumer BCIs.

The growing implementation of machine learning (ML) in EEG-based MI recognition is driven by its ability to effectively analyse complex patterns in brain signals, exploit the high temporal resolution of EEG data, handle the noisy nature of the signals, and enable personalised and adaptive BCIs. ML algorithms are known for their ability to learn from large amounts of data, extract meaningful features, and hence improve the accuracy of action recognition. This advancement enhances the usability and effectiveness of EEG-based BCIs, making ML an essential tool for improving the performance of EEG-based action recognition applications.

Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are two popular deep learning architectures that have been successfully applied to a wide range of tasks. Given their success in other domains, there has been growing interest in applying CNNs and RNNs to EEG-based action recognition in BCIs. However, there is a lack of comprehensive comparative studies that directly compare the performance of these architectures specifically for this task. Therefore, this research paper aims to address this gap by conducting a comparative analysis of the CNNs and RNNs for EEG-based motor imagery recognition in BCIs.

This study evaluates the effectiveness of CNNs and RNNs in classifying EEG signals in the case of MI classification. To achieve this, a publicly available dataset of EEG signals recorded during various motor imagery tasks will be utilised. The performance of the two architectures will be assessed based on accuracy, precision, recall, F1-score metrics, and other statistics related to computational efficiency. The findings of this research will be able to provide further insights into the choice of deep learning architectures when it comes to neuroimaging analysis, especially for EEG-based motor imagery classification-related tasks.

Literature Review

Convolutional Neural Networks

A convolutional neural network (CNN) is a deep learning model designed for processing grid-patterned data, such as images. CNN models learn spatial hierarchies of features from low to high levels. Their convolution and pooling layers extract features and translate these features to the final output. Moreover, CNNs can process images efficiently by applying kernels to each image position, and their parameters can be optimised through training with algorithms like backpropagation and gradient descent so as to minimise the difference between outputs and ground truth labels [1].

Existing research has extensively explored the utilisation of Convolutional Neural Networks (CNNs) for EEG-based action recognition in Brain-Computer Interfaces (BCIs). For instance, it is demonstrated that CNNs are effective in terms of both decoding and visualising EEG signals, and are able to automatically learn spatial features from raw EEG data through convolutional and pooling layers, enabling accurate classification of different action categories [2]. It is also proven that employing a multi-scale CNN architecture to capture both frequency and temporal information from EEG signals can lead to improved classification accuracy [3].

Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are a class of neural networks that excel in modelling sequential data by considering temporal dependencies. RNNs operate on a recurrent basis, allowing them to maintain the memory of past information while processing new inputs. This makes RNNs suitable for tasks involving time series data, such as speech recognition or natural language processing. In the context of action recognition from EEG signals, RNNs have shown promise in capturing the temporal dynamics and sequential patterns present in the data.

Researchers have demonstrated that the incorporation of spatial and temporal information through the use of Spatial and Temporal RNNs can significantly improve the accuracy of EEG-based MI classification, by allowing for a more comprehensive analysis of EEG data, capturing both the individual characteristics of each EEG channel and the interactions between channels [4]. LSTM-based approaches are also proven to outperform other RNN methods in terms of accuracy, suggesting the advantage of capturing long-term dependencies and modelling temporal dynamics in EEG data [5].

Relevant Studies and Research Gap

Several studies have compared the performance of CNNs and RNNs for action recognition tasks or EEG-based analysis. For example, in domains like action recognition with video [6], or EEG-based emotion recognition [7]. However, in the context of EEG-based MI recognition, direct comparative analysis between CNNs and RNNs using the same dataset is relatively limited, as most studies have focused on either CNNs or RNNs individually, or exploring the fusion between the two [8]. Therefore, further investigation is required to determine the strengths and weaknesses of these approaches in this specific context.

Methodology

Dataset Selection and Preprocessing

In this study, the EEG Motor Movement/Imagery Dataset v1.0.0 [9] from PhysioNet [10] will be used. The dataset consists of EEG recordings collected from 109 subjects while they performed different motor tasks or imagined specific movements. The motor tasks in the dataset involve various limb movements, such as left-hand movement and right-hand movement. Each task is associated with a specific class label. The dataset also includes a rest state, where the subjects were instructed to relax and not perform any motor imagery tasks. The EEG signals were recorded using a 64-channel as per the international 10-10 system, and the sampling rate of the recordings was 160 Hz. Raw EEG data from the dataset is used.

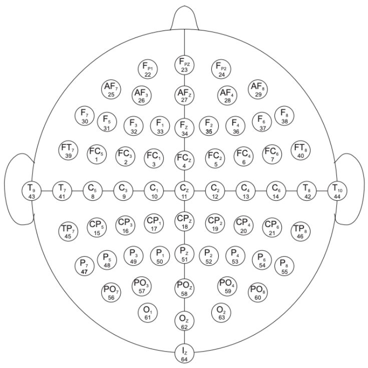

Figure 1. EEG cap according to the international 10-10 system for 64 channels

The experiment will use the data from only part of the subjects in order to maintain cost-effectiveness. Classification training will be based on three motor tasks, namely rest state, left-hand movement, and right-hand movement only. These motor tasks are assigned to the labels T0, T1, and T2 respectively. The raw data from each file in the dataset was originally presented in the form of a continuous two-minute recording of a subject performing one out of the three motor tasks every 4 seconds, hence extraction of the input data files is necessary. 500 labeled input data files are subsequently compiled.

Model Architecture and Training

Both CNN and RNN architectures are implemented using the Keras [11] library in Python.

The CNN architecture employs a sequence of convolutional and max pooling layers to extract spatial features from the input data. The resulting feature maps are flattened and connected to fully connected layers. In the RNN architecture, LSTM (Long Short-Term Memory) layers are employed to handle sequential data. The architectures both conclude with a softmax activation layer for multi-class classification.

To optimise the models, the Adam optimiser is used. Both utilise the categorical cross-entropy loss function, as well as incorporate dropout regularisation to prevent overfitting and enhance generalisation. They also employ rectilinear units (ReLU) as the activation function in the fully connected layers, introducing non-linearity to capture complex patterns in the data.

Evaluation

The two models will be evaluated and compared based on two aspects, classification accuracy and computational efficiency. Classification accuracy will be evaluated using the classification report generated with Scikit-learn [12], which includes evaluation metrics such as accuracy, precision, recall, and F1-score. A confusion matrix is also generated with extra unseen samples for each model to provide additional insights into the errors made by the models. Computational efficiency, on the other hand, will be evaluated with the statistics of the respective training time, memory usage, inference speed, and model size.

The models will be evaluated using a separate test dataset that was not used during the training phase. This ensures an unbiased evaluation of their performance. The predictions made by the models will be compared with the true labels from the test dataset to calculate the aforementioned evaluation metrics.

Experimental Design

The experiment will consist of 5 trials, with the input data files being randomly divided into two distinct sets: training and testing sets, in the ratio of 8:2. In each set, the numbers of input data corresponding to each motor imagery will be maintained at a ratio of 1:1:1. After completion of the trials, the average performance metrics across all runs will be calculated. This average performance serves as a more reliable estimate of the models' capabilities, as it accounts for the variability observed across different training instances.

Experimental Results

Table 1. Average figures of Classification Reports generated

	Convolutional Neural Networks			Recurrent Neural Networks		
	Precision	Recall	F1-score	Precision	Recall	F1-score
T0	0.63	0.83	0.72	0.61	0.58	0.60
T1	0.61	0.52	0.56	0.36	0.38	0.37
T2	0.45	0.48	0.46	0.37	0.36	0.36
Accuracy	0.61			0.46		
Macro Avg.	0.56	0.52	0.52	0.44	0.43	0.43
Weighted Avg	0.58	0.59	0.57	0.48	0.45	0.46

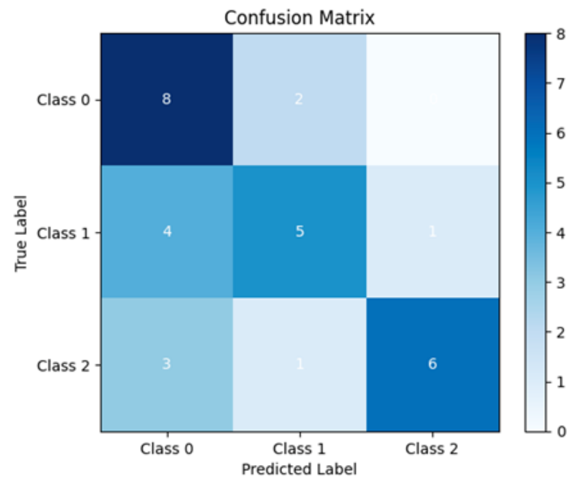

Figure 2. CNN confusion matrix

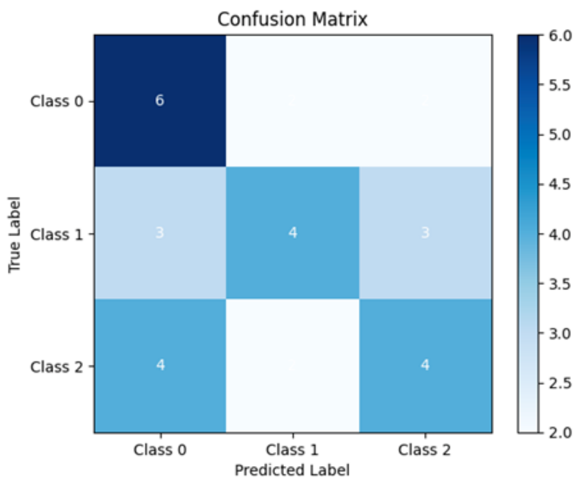

Figure 3. RNN confusion matrix

From Table 1, we can see that the CNN model generally exhibits higher precision, recall, and F1-score for all three MI classes, as well as achieves higher accuracy of 60% compared to 45% of the RNN model. This difference is also reflected in Figures 2 and 3, even though specific figures might vary from those in the classification report, possibly due to differences in the test samples used. From the confusion matrix results, it can be observed that the RNN model has a more balanced distribution of errors.

Table 2. Average computational efficiency statistics of the generated models

	Convolutional Neural Networks	Recurrent Neural Networks
Training time/ seconds	607.5	36.84
Memory usage/ MB	3190	2104
Inference time/ seconds	5.383	1.830
Model size/ MB	707.8	6.344

From Table 2, it is observed that RNN requires significantly less training time than CNN, as it only requires an average of 36.84 seconds to train while CNN requires an average of 607.5 seconds. The RNN model's interference time of 1.83 seconds is also shorter than that of the CNN model (5.383 seconds). The RNN model has a smaller memory usage of 2104 MB and a model size of 6.344 MB compared to 3190 MB and 707.8 MB of CNN respectively.

Discussion

In terms of classification accuracy, the difference between CNN and RNN can be attributed to several factors. One of the major factors would be CNN's model architecture being advantageous in terms of capturing spatial information. The EEG data is recorded over different scalp locations, hence the CNN model's ability to extract spatial features effectively allows it to identify patterns and activations across these different regions. On the other hand, the RNN model might struggle to capture these dynamics present in the EEG data due to its sequential processing nature, resulting in lower performance. However, the data also suggested that overfitting is less likely to occur in RNN architectures due to them having a smaller number of parameters.

It is worth noting that the nature of the chosen task is also an important factor affecting the accuracy of the models. For MI classification, the CNN model's ability to capture spatial information becomes particularly relevant. EEG data during MI tasks often exhibit strong spatial patterns, as different regions of the brain are activated depending on the imagined movement. Hence due to its strength in extracting spatial features and patterns across the channels in the raw data, the CNN model is expected to outperform the RNN model when the input data contain significant spatial information that is crucial for accurate classification. While EEG data used in our experiments do have inherent temporal dependencies (e.g. the progression of motor imagery over time), this further suggests that the spatial patterns and activations across different scalp locations might play a more significant role in distinguishing between MI classes than the temporal dynamics captured by the RNN model, leading to the fact that the RNN model is less effective for this specific task.

In terms of computational efficiency, it was observed that the CNN model require significantly more time to train compared to the RNN model. This can be attributed to the more complex architecture and the larger number of parameters involved in CNN models to process spatial information, which necessitate more computational resources for training. This might also be a major contributing factor to a longer inference time for the CNN model relative to the RNN model. In addition, the RNN model requires less memory compared to CNN, as it primarily focuses on temporal dependencies and does not require as much memory for processing, while CNN models would need to store and process large amounts of spatial information present in the input data. Lastly, it was found that the CNN model has a larger size compared to the RNN model. This is primarily due to the larger number of parameters involved in CNNs compared to RNNs, which generally have a more compact architecture.

Limitations

One of the limitations of this research is the relatively small sample size used for training and evaluation. While the results obtained with the available data are already able to demonstrate the differences between the performances of the two models, a larger sample size would provide more robust evidence of the generalisability of the findings. Another limitation would be the potential influence of preprocessing method choices on model performance. Addressing these limitations in future research would help validate and enhance the reliability of the above-mentioned results.

Conclusion

This research article presented a comparative analysis of convolutional neural networks and recurrent neural networks for EEG-based motor imagery recognition in brain-computer interfaces. Based on our findings, it can be concluded that CNNs demonstrated better classification accuracy compared to RNNs, thanks to their effectiveness in capturing spatial features from EEG signals. On the other hand, RNNs showed superior computational efficiency, possibly making them a feasible choice for real-time applications.

However, there are still areas that require further investigation. An in-depth comparison of other aspects of the models, e.g. feature extraction, sensitivity to noise, etc., would potentially help us understand more about the architectures 'respective characteristics when it comes to learning from EEG MI datasets. Furthermore, the generalisation capability of the models should also be evaluated across different subjects and datasets to ensure their robustness in real-world scenarios.

The results and insights gained from this study can guide researchers and practitioners in the field of BCI in selecting the most appropriate deep-learning architecture for their specific applications. Our findings can potentially contribute to the development of more accurate and efficient novel MI classification models, ultimately improving the quality of life for BCI users, especially those with motor impairments.

Acknowledgments

I would like to thank my advisor for providing valuable insights into the topic and guidance on experimental design.

References

[1] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9. https://insightsimaging.springeropen.com/articles/10.1007/s13244-018-0639-9

[2] Schirrmeister RT, Springenberg JT, Fiederer LDJ, Glasstetter M, Eggensperger K, Tangermann M, Hutter F, Burgard W, Ball T. Deep learning with convolutional neural networks for EEG decoding and visualization. Hum Brain Mapp. 2017 Nov;38(11):5391-5420. DOI: 10.1002/hbm.23730. Epub 2017 Aug 7. PMID: 28782865; PMCID: PMC5655781. https://pubmed.ncbi.nlm.nih.gov/28782865/

[3] Chang Z, Zhang C, Li C. Motor Imagery EEG Classification Based on Transfer Learning and Multi-Scale Convolution Network. Micromachines (Basel). 2022 Jun 10;13(6):927. DOI: 10.3390/mi13060927. PMID: 35744539; PMCID: PMC9228168. https://pubmed.ncbi.nlm.nih.gov/35744539/

[4] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–Temporal Recurrent Neural Network for Emotion Recognition," in IEEE Transactions on Cybernetics, vol. 49, no. 3, pp. 839-847, March 2019, DOI: 10.1109/ TCYB.2017.2788081.https://ieeexplore.ieee.org/document/8275511

[5] Wang, Ping & Jiang, Aimin & Liu, Xiaofeng & Shang, Jing & Zhang, Li. (2018). LSTM-based EEG Classification in Motor Imagery Tasks. IEEE Transactions on Neural Systems and Rehabilitation Engineering. PP. 1-1. 10.1109/ TNSRE.2018.2876129. https://ieeexplore.ieee.org/document/8496885

[6] C. Zhao, J. G. Han, and X. Xu, "CNN and RNN Based Neural Networks for Action Recognition," Journal of Physics: Conference Series, vol. 1087, no. 6, 2018, art. no. 062013. DOI: 10.1088/1742-6596/1087/6/062013. https://iopscience.iop.org/article/ 10.1088/1742-6596/1087/6/062013#:~:text=In%20order%20to%20classify%20the,the%20extracted%20features%20by%20CNNs.

[7] Wang X, Ren Y, Luo Z, He W, Hong J, Huang Y. Deep learning-based EEG emotion recognition: Current trends and future perspectives. Front Psychol. 2023 Feb 27;14:1126994. DOI: 10.3389/fpsyg.2023.1126994. PMID: 36923142; PMCID: PMC10009917.

https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1126994/full

[8] Hongli Li, Man Ding, Ronghua Zhang, Chunbo Xiu. Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network. Biomedical Signal Processing and Control, Volume 72, Part A, 2022, 103342, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2021.103342.

[9] Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. IEEE Transactions on Biomedical Engineering 51(6):1034-1043, 2004. https://ieeexplore.ieee.org/document/1300799

[10] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. E215–e220. https://physionet.org/content/eegmmidb/1.0.0/

[11] Chollet, F. et al. (2015). Keras. https://keras.io/

[12] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830,2011. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html