IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

Classification of Electroencephalogram (EEG) based on Deep Learning and Neural Networks-1

Shashwitha Puttaswamy[1], Vishesh S[2]

Research Scientist, Konigtronics Pvt Ltd, Bangalore, India[1]

Research Head, Konigtronics Pvt Ltd, Bangalore, India[2]

Abstract: Electroencephalogram (EEG) is a multi-dimensional time-series brain signal that is highly information packed. While an EEG has high potential to serve in medicine (e.g. disease diagnosis, prognosis, pre-disease risk identification), psycho-physiology (e.g. mood classification, stress monitoring, alertness monitoring, sleep stage monitoring), brain-computer interface application (e.g. thought typing, prosthesis control), and many other areas, the classical design of EEG feature extraction algorithms and EEG classifiers is time-consuming and challenging to fully tap into the vast data embedded in the EEG. Deep learning (or deep neural network) which enables higher hierarchical representation of complex data has been strongly suggested by a wide range of recent research that these deep architectures of artificial neural network generally outperform the classical EEG feature extraction algorithms or classical EEG classifiers. In this paper/ research project, deep neural network architectures have been constructed to perform binary classification on an EEG dataset that was shown by traditional EEG feature extraction methods to have no significant difference between its two data pools (resting EEG recorded before and recorded after listening to music). The Convolutional Neural Network (CNN) model constructed in this project has achieved a validation accuracy of 75±1% using the same EEG dataset. Using the top performing CNN architectures, short duration of relaxing music listening is found to affect the EEG signals generated by the frontal lobe more than the other lobes of the brain; and also to affect the EEG generated by the left cerebral hemisphere more than the right hemisphere.

Keywords: Electroencephalogram (EEG), Deep learning (or deep neural network), Convolutional Neural Network (CNN) model, short duration of relaxing music listening, Activation techniques, epoch, validation accuracy.

I. INTRODUCTION

Electroencephalograms (EEG) are recordings of the electrical potentials of the brain typically measured from the scalp, as signal waveforms of varying frequencies and amplitude (in mV). The EEG is packed with information regarding the electrical activities of the brain, be it pathological or physiological. Hence, EEGs are very useful in the medical field (such as diagnostic purposes, real-time monitoring of clinical progress of patients, prognostic purposes, and the pre-disease identification of prodromal neuro-pathological signals in preventive healthcare of increasing importance), for Brain-Computer-Interface (BCI) applications (such as thought-typing, prosthetic limbs control, and many others which can potentially improve the quality of life of the people with motor disabilities, as well as the normal), and myriad forms of other potential applications such as drowsiness warning system for drivers or lie detection for criminal investigation.

II. PROBLEM STATEMENT

EEG signals are packed with information and hence it requires lots of effort and time to perform manual analysis or decoding of these signals. Siuly and Li (2014) [1] stated that manual design the feature extraction model for multiple-class electroencephalogram (EEG) signal classification is an extremely challenging task because the true representative features/patterns have to be identified and extracted precisely from the multidimensional time series of EEG measured from the brain.

With the advances in the techniques for modelling the deep learning architecture, deep learning has revolutionized the computer's capability for processing information-packed data. For example, convolutional neural network for image processing has provided solutions to challenges previously encountered by the computer vision community, while recurrent neural network has resulted in much improvement in the processing of time-series signals such as speech processing.

It is thus very likely that deep learning will improve the analysis of EEG signals as well. A number of different studies (Ren and Wu, 2014; [8] Behncke et al, 2017; Schirrmeister et al, 2017 [2]) trained and tested various architectures of deep learning for EEG data analysis and reported improved accuracies compared with the state-of-art EEG feature extraction methods. Yet, the research in the application of deep learning on EEG analysis is a new area of study and further analytical accuracy improvement is in need for much more reliable practical application. In this paper, various architectures of DNN for EEG analysis, feature extraction, and classification will be carried out.[3-8]

III. LITERATURE SURVEY

The classification of EEG signals using various versions of artificial neural networks have been published with higher sensitivity, specificity, and accuracy than other traditional feature extraction and statistical methods.
Patnaik and Manyam (2008) [6] applied the neural network for the identification of epileptic EEG segment from the non-epileptic EEG. They used discrete wavelet transform (DWT) for feature extraction, followed by a feed-forward back propagating artificial neural network (ANN) for classification, with the training set for the ANN model being selected by a genetic algorithm instead of randomization. They improved their classification result by incorporating a post-classification stage using harmonic weights. The training and validation were done using the invasive pre- surgical EEG recording of 21 patients with medically intractable focal epilepsy. The average specificity of 99.19%, sensitivity of 91.29%, and selectivity of 91.14% were obtained. Each patient's EEG recording contained at least 50 min of pre-ictal and 50 min of post-ictal recording and the average duration of EEG with epileptic data was 7.73 min for one patient.[10-12]

Subasi and Ercelebi (2016) compared logistic regression and neural network models for EEG signals (epileptic vs. normal data) classification. [12] They obtained 89% accuracy using logistic-regression based classifier, which was lower than the two neural network models. The Multi- Layer Perceptron Neural Network (MLPNN) trained with common error back propagation algorithm achieved an accuracy of 92%; while the MLPNN trained with Levenberg-Marquardt (L-M) optimization method achieved an even higher accuracy of 93%. The MLPNN models were trained with a total of 300 EEG examples (102 epileptic and 198 normal EEG), and validated with another set of 200 EEG examples (88 epileptic and 112 normal EEG).

Satapathy, Dehuri and Jagadev (2017) performed classification of EEG for epileptic seizure identification using a version of neural network known as Radial Basis Function Neural Network (RBFNN).[9] Their RBFNN was trained for mean square error optimization with a modified Particle Swarm Optimization (PSO). The improved PSO (termed IPSO in the paper) was designed for improving the searching speed of traditional PSO for global optimum. The RBFNN with IPSO had achieved a maximum accuracy of 99%.

Supratak et. al. (2017) constructed a Deep Learning model which utilizes: convolutional neural network (CNN) to extract time-invariant features, and bidirectional long-short-term-memory (bidirectional-LSTM) to learn transition rule among sleep stages from EEG epochs. [13] Their model was trained with a two-step training algorithm which pre-trains the model using over-sampled data to lessen class-imbalance problems, and later fine tunes the weights of the pre-trained model with sequences of EEG epochs to encode the model with necessary patterns for sleep stages classification. The training dataset was from the F4-EOG channel of 62 subjects, giving rise to a total of 58600 EEG epochs, with the total recording duration close to 490 hours. The model achieved an accuracy of 86.2% and the macro F1-score of 81.7 as shown in Table 1.

	Predicted					Per-class Metrics		
	W	N1	N2	N3	REM	PR	RE	F1
W	**5433**	572	107	13	102	87.3	87.2	87.3
N1	452	**2802**	827	4	639	60.4	59.3	59.8
N2	185	906	**26786**	1158	499	89.9	90.7	90.3
N3	18	4	1552	**6077**	0	83.8	79.4	81.5
REM	132	356	533	1	**9442**	88.4	90.2	89.3

Table 1: Confusion Matrix for the Performance of the DeepSleepNet (Supratak et. al., 2017)

Hajinoroozi, Mao and Huang (2015) applied Deep Learning to perform prediction of driver's drowsy or alert states using the EEG data. They introduced the Channel-wise Convolutional Neural Network (CCNN) and a variation of CCNN (termed CCNN-R in the paper) which adopted the Restricted Boltzmann Machine (RBM) in place of the convolutional filter/layers of conventional CNN models.[4]

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

The EEG data set was collected from three studies of the driver's cognitive states using a virtual reality dynamic driving simulator. The simulated driving scenes were night time driving with 100 km/h with perturbation being injected into driving path every 8 to 12 seconds. The reaction times of the drivers were used to determine their alert/drowsy mental states. The dataset was collected from 70 sessions for 37 subjects. The EEG was recorded for 3 seconds before each perturbation was taken into consideration for CNN models training, with a total of 35074 non-overlapping 1s epochs (23074 alert and 15924 drowsy epochs).

In contrast to the conventional CNN which uses 2-D or multi-dimensional convolutional kernels for feature extraction, CCNN applies a 1-D kernel to convolve along each channel (hence channel-wise). After the feature extraction, the categorization with Fully Connected (FC) layers also uses back propagation for weight optimization. The common kernels for CCNN include the Gaussian or Xavier filters.

A model variation mentioned above (CCNN-R) uses a more complicated feature extraction layers (RBM). The FC layers' backpropagation method has to be adjusted accordingly.
The algorithms' performance were evaluated using Az-score, with the CCNN having achieved the Az-score of 79.63% and the CCNN-R 82.78%. The prediction performance of other popular methods were investigated too, with the LDA achieving 52.81%, SVM achieving 50.38%, and CNN 71.41%.

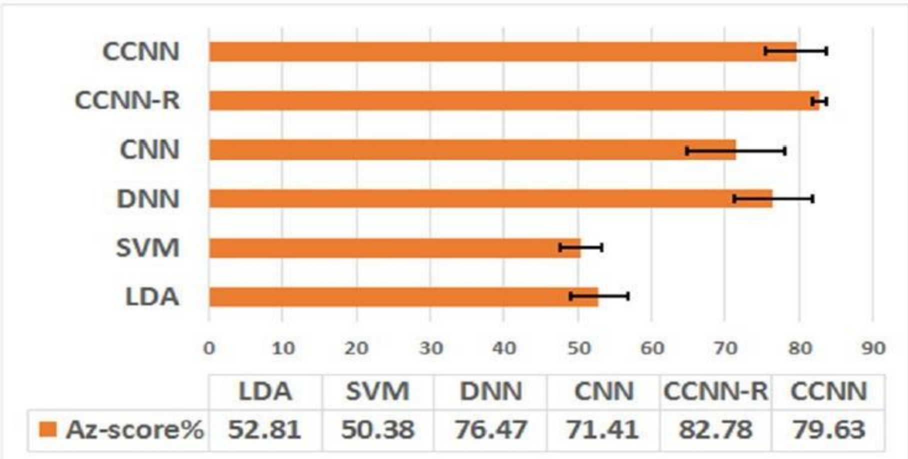

Figure 1: Performance (Az-score) of various machine learning methods at predicting drivers' alertness using raw EEG data (Hajinoroozi, Mao and Huang, 2015)

Behncke et al (2017) attempted to classify the EEG signals of humans observing robot action into two classes (observing a successful robotic operation or observing a robotic failure).[2] The classification task was performed with deep convolutional neural network (deep ConvNets), regularized Linear Discriminant Analysis (rLDA), and filter bank common spatial patterns (FB-CSP) combined with rLDA. Deep ConvNets achieved accuracies of 75% ± 9%, significantly higher than both the other two commonly used EEG classifiers, with the rLDA of 65% ± 10% and the FB-CSP combined with rLDA of 63% ± 6%, as shown in Table 2.

paradigm	interval	mean accuracy ± standard deviation		
		ConvNet	*rLDA*	*FB-CSP*
KPO error	2.5-5s	(78.2 ± 8.4) %	(67.5 ± 8.5) %	(60.1 ± 3.7) %
KPO error	3.3-7.5s	(71.9 ± 7.6) %	(63.0 ± 9.3) %	(66.5 ± 5.7) %
RGO error	4.8-6.3s	(59.6 ± 6.4) %	(58.1 ± 6.6) %	(52.4 ± 2.8) %
RGO error	4-7s	(64.6 ± 6.1) %	(58.5 ± 8.2) %	(53.1 ± 2.5) %

Table 2: Accuracies of ConvNet, rLDA and FB-CSP at identify EEG of human observing robotic failure (Behncke et. Al., 2017)

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

Ren and Wu (2014) also compared the performance of a deep learning architecture using Convolutional Restricted Boltzmann Machines (CRBM), to other state-of-art classical feature extraction methods including power band, multivariate adaptive autoregressive (MVAAR), and common spatial pattern (CSP). For 2-class and 4-class classification, the deep learning model achieved accuracies of 83% - 88% which is in general higher than the classical feature extraction methods (80% - 86%). The accuracy of the deep learning method in particular increased as the number of training samples increased from 80 to 240, as shown in Table 3 and Table 4.[8]

Training Samples	Correct Rate (%)			
	CSP	*MVAAR*	*Band Power*	*CDBN*
80	**85.38±2.24**	80.35±4.07	81.90±2.75	83.63±1.82
120	85.25±1.94	84.88±3.87	83.59±2.90	**85.94±1.77**
160	85.74±2.33	85.46±2.54	85.00±2.55	**86.04±2.09**
200	85.56±3.10	84.81±4.06	84.63±4.24	**86.06±3.38**
240	85.75±6.18	85.88±5.80	86.13±6.36	**88.25±5.70**

Table 3: Mean 2-class motor imagery EEG classification accuracy of various methods (Ren and Wu, 2014)

Training Samples	Correct Rate (%)			
	CSP	*MVAAR*	*Band Power*	*CDBN*
140	80.45±1.62	81.08±1.50	80.54±2.21	**82.02±1.88**
160	81.02±1.50	80.70±2.00	81.53±1.61	**82.41±1.44**
180	85.47±1.44	85.64±2.36	86.09±2.36	**87.33±1.74**

Table 4: Mean 4-class motor imagery EEG classification accuracy of various methods (Ren and Wu, 2014)

IV. METHODOLOGY

Two main varieties of deep learning architecture investigated in this project are pure multilayer perceptron (MLP) models and convolutional neural networks (CNN). The impact of modelling techniques and hyperparameters of deep learning models on the model's performance are also investigated, which include the effect of different optimizers, activation functions and dropout rates. The EEG data used for training and validation of the deep learning models are 14-channel EEG of 26 participants of a music-based neuro-feedback training previously conducted by a UTAR FYP student (Phneah, 2017).[14]

In neurofeedback training, the measurement of brain activity (EEG in this case) is used as the feedback information to the participant for the purpose of attaining desired regulation of the brain function. The EEG data used for this project is a portion of the EEG recorded at the very initial phase of the neurofeedback study, with each participant having undergone only a single short session of listening to favorite and relaxing music.

Three-minute EEG signal was recorded, at sampling frequency of 128Hz, before and during each of the 26 participants listened to their favourite and relaxing music, generating 52 EEG recordings (26 before listening to music and another 26 after listening to music). Each of the EEG recordings, after artifact removal and data cleaning, has different lengths ranging from 80-100 seconds. Hence, only the first 10000 sampling points (about 78 seconds) of each pre-processed EEG are used as the dataset of this project.

Each of the 52 cleaned EEG recordings is then split into 40 sub-segments, generating 2080 EEG recording segments (1040 before music and 1040 after music). Each of the sub-segments has the time span of about 1.95 seconds (250

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

sampling points). This dataset is shuffled and divided randomly into the training set and validation set at the ratio of 9:1, giving 1872 EEG segments (942 before music and 930 after music) as the training data set and 208 EEG segments (98 before music and 110 after music) as the validation set, as shown in Table 5.

EEG Dataset	Training set	Validation set
Before listening to music	942 (45.3%)	98(4.71%)
After listening to music	930(44.7%)	110(5.29%)
Total	1872(90%)	208(10%)

Table 5: The numbers of categorized EEG data contained in the training set and the validation set.

The computer system used for the training and validation of the models is a Dell Inspiron 7567 laptop, with the following specifications:
1. CPU: Intel Core i5-7300HQ 2.50 GHz
2. RAM: 4GB DDR4, plus an extra 8GB upgrade
3. GPU: NVIDIA Geforce GTX 1050 4GB graphic RAM

The capability of GPU is of utmost importance because the fundamental design of GPUs allows huge amount of parallel computation of the same instructions. This suits the requirement of running deep learning models which are generally designed with large matrix of repetitive computational nodes.

In fact, the NVIDIA GTX 1050 GPU used in this project is designed for gaming purpose and is a rather low end GPU for deep learning research.
The programming language used in this project is the Python language, version 3.6.4, under Anaconda distribution. Anaconda enables convenient creation and management of Python environment (conda environment), under which we can selectively run different tools specifically installed to the particular environment. The scientific programming Python libraries used in this project include the numpy library, scikit-learn (sklearn) library, and matplotlib library. The Python 'os' library is used to move around, read from, and write to the system's directories. The 'mne' library is used to handle EEG data. And last but not of any less, the 'tensorflow' machine learning library is used for the constructing and running the deep learning models (Abadi, et al, 2016).[15] Table 3.2 summarizes the Python libraries used.

Deep learning models can learn or be trained through unsupervised or supervised learning process. Unsupervised learning of a deep learning model will enable the model to divide the dataset into classifiable clusters, without any indication as to which group any training or validation data belongs to. On the other hand, supervised learning, which is the training method used in this project, requires each example (x) of the training data to be associated or encoded with a label (y). After repeated observation of the paired examples of data x and label y, the model learns to predict y from data x. Conceiving and constructing a deep learning model involve specifying the type of feature extraction operation to be incorporated (the application of convolutional kernels in convolutional neural network as in this project or the application of feedback loop in the neural network forming a recurrent neural network), the number of network layers (the depth of the model), the number of neuron at each layer (the width of the model), the type of activation function for the neural layers, the model regularization methods such as the drop-out mechanism to prevent overfitting, the choice of error back-propagation optimizer and the learning rate.

The training data set is divided from a total of 1872 segments of EEG into 16 smaller mini-batches, each containing 117 segments of EEG. During the model training stage, the mini-batches are fed batch-by-batch to the model-under-training.

V. RESULTS AND DISCUSSIONS

In order to investigate the impact of different optimizers, activation functions, and dropout rates on the progress of deep learning training process, all the different modelling techniques are independently tested on the same single convolutional neural network architecture.

Two optimization algorithms (the basic gradient descent algorithm and the adaptive moment estimation (Adam)) are tested for their effectiveness in searching the minimal point of the cost function of the deep learning model for EEG classification. The activation function is fixed as ReLU and the dropout rate is fixed at 50% for either of the optimization techniques. This is to ensure that the changes in the performance of the model are all the result of the change in optimizer, instead of being the combined effect of changing various different modelling techniques.

The Adam optimizer incorporates the operation of both the momentum optimizer and the RMSprop optimizer. Both the momentum optimizer and the RMSprop optimizer allow speeding up of the optimization process towards the minimal loss, by ignoring noises in the parameter updating process.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

Adam optimizer significantly outperformed the other optimizers (including stochastic gradient descent, RMSprop, AdaGrad and AdaDelta optimizers) in training both the multilayer neural network and convolutional neural network models, using MINST and CIFAR-10 data. Adam optimizer is able to achieve a much lower training cost (training error) than the other optimizers. Adam has also markedly increased the optimization convergence speed.

The performance of the multilayer neural network and convolutional neural network with different optimizers as shown in figure 2.

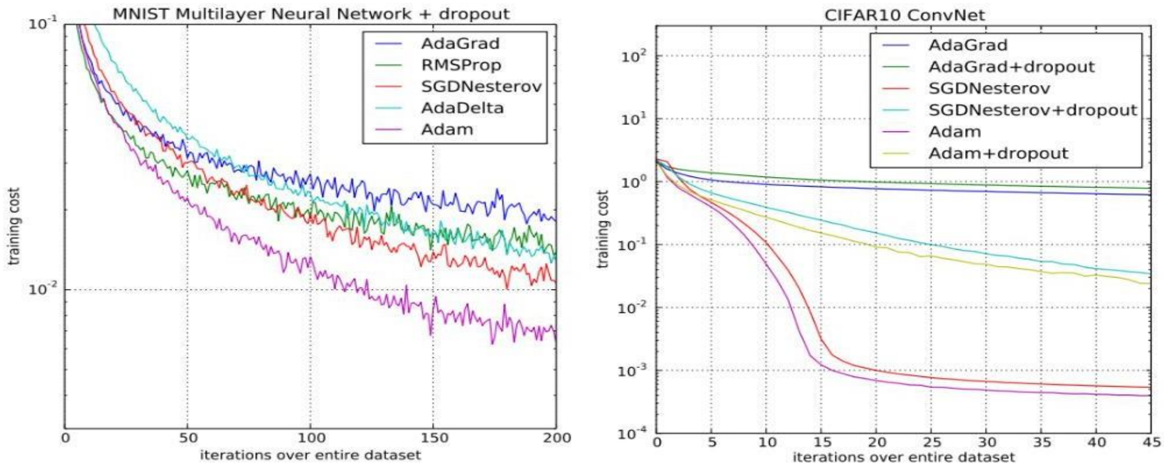

Figure 2: Performance of multilayer neural network (left) and convolutional neural network (right).

When the gradient descent optimizer is used, the model fails to learn from the EEG training dataset. The optimization process may have been trapped at a very early local minimum, or the deep model may have a cost function with extremely low gradient which has caused the gradient descent optimizer to learn too slowly.

On the other hand, when the Adam optimizer is used, the model has successfully learned and extracted the distinctive features between the two groups of EEG (before and after listening to music), enabling its classification accuracy to improve above 70% over the training iterations.

The sigmoid and ReLU activation functions both resulted in an increased classification accuracy, as compared to the model without activation function. ReLU activation function is the most suitable, among the tested functions, for the designed model to perform classification on the EEG data. ELU activation function has negatively impacted the model's optimization, resulting in a performance worse than that without any activation function. However, the reason for ELU's negative impact is not clear.

One of the main challenges in the design of a deep learning model is the requirement for the model to perform with an almost equal accuracy on previously unobserved data (such as the test dataset), as on the training data set. This is a desired ability of the learning model, termed as generalization. The models that can generalize well are usually models with large capacity that are properly regulated.

Dropout mechanism is one of the regularization methods. In dropout method, a percentage of neurons (or computational nodes) of certain layers of the neural network is specified to be randomly blocked out during the training steps. Each training step will make a different combination of computational nodes available, instead of the full network. Hence, the model-under-training will not be able to rely too much on any selective few features propagated by certain computational nodes. Instead, every partial combination of the network will be more sufficiently trained, having their weights been updated more properly through backpropagation of error.

The best dropout rate among the examined is 40-50% dropout. The model with no dropout mechanism overfits the earlier. Extremely high dropout rate (such as 70% dropout) throttled the learning speed too much although overfitting is avoided. The classification accuracy achieved using the temporal, parietal and occipital channels combined without the frontal channels is significantly lower than that achieved using six frontal channels. The model is able to classify frontal lobe EEG signals better than the signals from the other lobes.

This is probably because the short session of relaxing music listening has a greater impact on the frontal lobe than the other regions of the brain, causing the EEG generated by the frontal lobe to differ more significantly (before and after listening to music) than the EEG from the other regions.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14712

The CNN model trained and validated with the left hemisphere EEG signals has achieved significantly higher classification accuracy than the model trained and validated with the right hemisphere EEG.
This indicates that the short session of relaxing music has affected the left cerebral hemisphere more than it does to the right cerebral hemisphere. This finding is in contrary to our expectation that the right cerebral hemisphere, which is in charge of our emotion, should be affected more by the music than the left hemisphere.

VI. CONCLUSIONS AND FUTURE RECOMMENDATIONS

For the task of binary classification of EEG, one of the 14-channel CNN model has achieved the top validation accuracy of 75±1%. This performance is closely followed by another 6-frontal-channel CNN model, which has achieved validation accuracy of 71.5±2%.

This finding is significant as the models are operating on the EEG dataset that was shown by previous classical manual feature extraction methods to have no statistical significant difference.

Basic Gradient Descent Optimizer is not sufficient for training the deep learning models for the task of EEG data classification. Using basic gradient descent optimization algorithm to minimize the cost function, deep learning models have failed to learn from the EEG data, causing the validation accuracy to stay below 50%. Adam Optimizer performs significantly better at training the deep learning model for EEG data classification, with the validation accuracy to reach up to and above 67±2%. ReLU is the most suitable activation function for deep learning model for EEG classification, followed by the sigmoid function. The model with ELU activation function performs worse (with validation accuracy below 50%) than the model without any activation function.

The most suitable dropout rate is around 40% to 50%. Too low the dropout rate (0% to 30%) does not help much in preventing overfitting of the model to the training data. Too high the dropout rate (70%) will slow down the model learning speed excessively.

Convolutional layers significantly improve the performance of the deep learning model for EEG classification, elevating the validation accuracy from below 64% up to above 75%. A short session of listening to relaxing music has greater degree of impact to the frontal region than the other regions of the brain, and also greater impact to the left cerebral hemisphere than the right, inferring from the discrepancy at the classification accuracy as discussed in the results section of the paper/ research project.

REFERENCES

[1]. Siuly and Li, Y. (2014). A novel statistical algorithm for multiclass EEG signal classification. [online]. *Engineering Applications of Artificial Intelligence*, **34**(2014), pp.154–167. Available at: http://www.sciencedirect.com.libezp.utar.edu.my/science/article/pii/S0952197614 001092 [Accessed 6 Jul. 2017]
[2]. Behncke, J., Schirrmeister, R. T., Burgard, W. and Ball, T. (2017). The signature of robot action success in EEG signals of a human observer: Decoding and visualization using deep convolutional neural networks. [online]. *2018 6th International Conference on Brain-Computer Interface (BCI)*, GangWon, South Korea, January 15-17, 2018. IEEEXplore, pp 1-6. Available at:https://arxiv.org/ftp/arxiv/papers/1711/1711.06068.pdf [Accessed 4 Feb. 2018]
[3]. Clevert, D., Unterthiner, T. and Hochreiter, S. (2016). *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. [online]. Available at: https://arxiv.org/abs/1511.07289 [Accessed 20 Nov. 2017]
[4]. Hajinoroozi, M., Mao, Z. and Huang, Y. (2015). Prediction of driver's drowsy and alert states from EEG signals with deep learning. In: *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. [online]. Cancun: IEEE Conference Publications, pp. 493–496. Available at: http://ieeexplore.ieee.org.libezp.utar.edu.my/document/7383844/ [Accessed 20 Jul. 2017]
[5]. Ide, H. and Kurita, T. (2017). Improvement of learning for CNN with ReLU activation by sparse regularization. [online]. *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, May 14-19, 2017. IEEEXplore, pp. 2684-2691. doi: 10.1109/IJCNN.2017.7966185 [Accessed 10 Dec. 2017]
[6]. Patnaik, L. M. and Manyam, O. K. (2008). Epileptic EEG detection using neural networks and post-classification. *Computer Methods and Programs in Biomedicine*. [online]. **91**(2), pp. 100–109. Available at: http://www.sciencedirect.com.libezp.utar.edu.my/science/article/pii/S0169260708 000539 [Accessed 10 Jul. 2017]
[7]. Ramachandran, P, Zoph, B. and Le, Q. V. (2018). Searching for Activation Functions. [online]. *6th International Conference on Learning Representations*, Vancouver Convention Center, Vancouver, BC, Canada, April 30 - May 3, 2018. Available at: https://openreview.net/forum?id=Hkuq2EkPf [Accessed 8 Apr. 2018]

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ≍ Peer-reviewed & Refereed journal ≍ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14712

[8]. Ren, Y. and Wu, Y. (2014). Convolutional Deep Belief Networks for Feature Extraction of EEG Signal. *2014 International Joint Conference on Neural Networks*. [online]. Beijing: IEEE Conference Publications, pp.Available at: http://ieeexplore.ieee.org.libezp.utar.edu.my/document/6889383/ [Accessed 10 Jul.

[9]. Satapathya, S. K., Dehuri, S. and Jagadev, A. K. (2016). EEG signal classification using PSO trained RBF neural network for epilepsy identification. [online]. *Informatics in Medicine Unlocked*, **6**(2017), pp. 1–11. Available at: http://www.sciencedirect.com/science/article/pii/S2352914816300387 [Accessed 5 Aug. 2017]

[10]. Schirrmeister, R. T., Gemein, L., Eggensperger, K., Hutter, F. and Ball, T. (2017). *Deep learning with convolutional neural networks for decoding and visualization of EEG pathology*. [online]. Available at: https://arxiv.org/abs/1708.08012 [Accessed 30 Mar. 2018]

[11]. Schirrmeister, R. T., Springenberg, J. T., Fiederera, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W. and Ball, T. (2017). *Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human EEG.* [online]. Available at: https://arxiv.org/abs/1703.05051 [Accessed 4 Feb. 2018]

[12]. Subasi, A. and Ercelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. [online]. *Computer Methods and Programs in Biomedicine*, **78**(2), pp.87–99. Available at: http://www.sciencedirect.com/science/article/pii/S0169260705000246 [Accessed 10 Jul. 2017]

[13]. Supratak, A., Dong, H., Wu, C. And Guo, Y. (2017). DeepSleepNet: a Model for Automatic Sleep Stage Scoring based on Raw Single-Channel EEG. [online]. *IEEE Transactions on Neural Systems and RehabilitatioEngineering,* **Volume PP**(99), pp.1–1.Available at: http://ieeexplore.ieee.org.libezp.utar.edu.my/document/7961240/ [Accessed 25 Jul. 2017]

[14]. Phneah, S. W. and Nisar, H. (2017). EEG-based alpha neurofeedback training for mood enhancement. [online]. *Australasian Physical & Engineering Sciences in Medicine*. **40**(2), pp.325-336. Available at: https://link.springer.com/article/10.1007/s13246-017-0538-2 [Accessed 5 Apr. 2018]

[15]. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. and Zheng, X. (2016). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. [online]. Available at: https://arxiv.org/abs/1603.04467 [Accessed 15 Nov. 2017]