100 year anniversary
Measurement Science and Technology

PAPER • OPEN ACCESS

Robust CNN architecture for classification of reach and grasp actions from neural correlates: an edge device perspective

View the article online for updates and enhancements.

Robust CNN architecture for classification of reach and grasp actions from neural correlates: an edge device perspective

Hajrah Sultan[1]**, Haris Ijaz**[1]**, Asim Waris**[1]**, Shafaq Mushtaq**[2]**, Khurram Mushtaq**[3]**, Niaz B Khan**[1,4]**, Said Ghani Khan**[4]**, Mehdi Tlija**[5] **and Jamshed Iqbal**[6,*] ●

[1] National University of Sciences and Technology, Islamabad 44000, Pakistan

[2] Pakistan Institute of Medical Sciences, PIMS, Islamabad 44000, Pakistan

[3] School of Automation, Central South University, Changsha, People's Republic of China

[4] Mechanical Engineering Department, College of Engineering, University of Bahrain, Isa Town 32038, Bahrain

[5] Department of Industrial Engineering, College of Engineering, King Saud University, PO Box 800, Riyadh 11421, Saudi Arabia

[6] School of Computer Science, Faculty of Sciences and Engineering, University of Hull, Kingston upon Hull, United Kingdom

E-mail: j.iqbal@hull.ac.uk

CrossMark

Abstract

Brain–computer interfaces (BCIs) systems traditionally use machine learning (ML) algorithms that require extensive signal processing and feature extraction. Deep learning (DL)-based convolutional neural networks (CNNs) recently achieved state-of-the-art electroencephalogram (EEG) signal classification accuracy. CNN models are complex and computationally intensive, making them difficult to port to edge devices for mobile and efficient BCI systems. For addressing the problem, a lightweight CNN architecture for efficient EEG signal classification is proposed. In the proposed model, a combination of a convolution layer for spatial feature extraction from the signal and a separable convolution layer to extract spatial features from each channel. For evaluation, the performance of the proposed model along with the other three models from the literature referred to as EEGNet, DeepConvNet, and EffNet on two different embedded devices, the Nvidia Jetson Xavier NX and Jetson Nano. The results of the Multivariant 2-way ANOVA (MANOVA) show a significant difference between the accuracies of ML and the proposed model. In a comparison of DL models, the proposed models, EEGNet, DeepConvNet, and EffNet, achieved 92.44 ± 4.30, 90.76 ± 4.06, 92.89 ± 4.23, and 81.69 ± 4.22 average accuracy with standard deviation, respectively. In terms of inference time, the proposed model performs better as compared to other models on both the Nvidia Jetson Xavier NX and Jetson Nano, achieving 1.9 sec and 16.1 sec, respectively. In the case of power

* Author to whom any correspondence should be addressed.

consumption, the proposed model shows significant values on MANOVA (p < 0.05) on Jetson Nano and Xavier. Results show that the proposed model provides improved classification results with less power consumption and inference time on embedded platforms.

Keywords: brain computer interface, deep learning, convolutional neural networks, embedded edge devices

1. Introduction

There are millions of people all over the world who have varied abilities and are frequently unable to carry out the activities of everyday life properly. Physical limitations may be due to a specific disability, such as epilepsy, stroke, or spinal cord injuries, as well as the result of multiple physical problems. Such limitations can include reduced ability to stand and walk, poor balance, coordination, and difficulty in bending. According to the disability report published by WHO in 2022, there are over 1.3 billion physically disabled people in the world that are 16% of the global population [1], out of which around one billion are considered to have significant difficulties in functioning. In developing nations, it is estimated that up to 4% of their populations experience significant difficulties in sustaining independent living. These difficulties include making choices and taking actions, communicating with others, and accessing services [2]. Moreover, millions of individuals worldwide have been infected with COVID-19 since 2019, and this illness does not impact everyone similarly. During the 2019 COVID pandemic, situation of the world has shifted. Throughout this outbreak, individuals feel afraid of touching something. A physically impaired individual needs help to do everyday tasks. People were scared of helping a crippled person who offered their arms because they might infect them with the COVID-19 virus. As a result of these fears, many disabled people are scared to accept help from caregivers because they fear that their illness may increase in severity due to contact with caregivers. The ability to guide a disabled person in helping them complete an action is undoubtedly a wonderful thing, but it can be challenging during a pandemic. The advancement in the field of brain–computer interfaces (BCIs) has enabled many people, who are suffering from disabilities, to significantly improve their life quality while performing everyday tasks.

BCI systems are gaining popularity in the healthcare and rehabilitation fields. The objective of the BCI is to encode brain activity with the help of electroencephalography (EEG), magnetic resonance imaging, electrocorticography, and decode it to use it as a control input for a computer [3]. The signals of EEG are more suited & accepted clinically for monitoring the activities inside the brain activity because of its low cost, high signal intensity and substantial fitness to the human subject [4, 5]. Figure 1 shows the basic working of a common BCI system, in which a link is established between the human brain and peripheral devices that can be a combination of robotic arm, computer screen, wheelchair, etc. A BCI requires two main components: an interface that couples electrodes to computers and algorithms capable of decoding acquired EEG data into meaningful information.

However, when employed in a real-world experiment, a BCI system encounters several challenges that include low BCI signal intensity, low signal-to-noise ratio, contamination with multiple artifacts (line noise, muscle artifacts, eye blink components), poor signal classification, limited computational resources, and higher power consumption are some of the most typical BCI systems issues [6]. There are several types of research in which researchers have investigated various techniques for analyzing EEG data for a variety of objectives ranging from the detection of diseases to devising control. Handcrafted feature extraction approaches are traditional procedures that have been applied in numerous EEG data classification systems [3–8]. Although handmade feature extraction approaches are time-consuming, give highly specialized features, are application-specific, and need prior expertise [9]. Deep learning (DL)-based techniques, on the other hand, extract a more general feature set that may be used for a variety of applications [10–12]. And many more researchers used DL in EEG data to build BCI applications by using different embedded platforms. DL huge success with EEG data was due to its capacity to extract significant elements from time-varying input. However, the development of efficient BCI systems is quite complicated, and one of the key challenges is to improve the classification accuracy. As a result, additional research in the BCI system using DL is necessary to achieve a decent level of accuracy.

Until now, the idea of being able to be controlling the devices with our minds was only a fantasy. However, utilizing electrical signals gathered from brain activity, the technology of BCI enables people who are physically disabled and are not able to use their bodies to control assistive equipment.

Just a few researchers have investigated DL-based BCI systems, but they have yet to create an accurate and efficient model that can perform exceptionally well in resource constraints scenarios. All the DL models consist of convolution, pooling, and fully connected layers. The complexity of a model varies depending on the different number of convolution layers, feature maps, and layers. DL models need high resources to train and infer which makes them difficult to implement on edge-embedded devices. Thus, this study aims to construct an efficient lightweight DL architecture to efficiently identify brain signals and implement them on embedded edge devices.

This paper introduces an efficient DL model specifically developed for real-time BCI applications on embedded devices with limited resources. The study showcases a
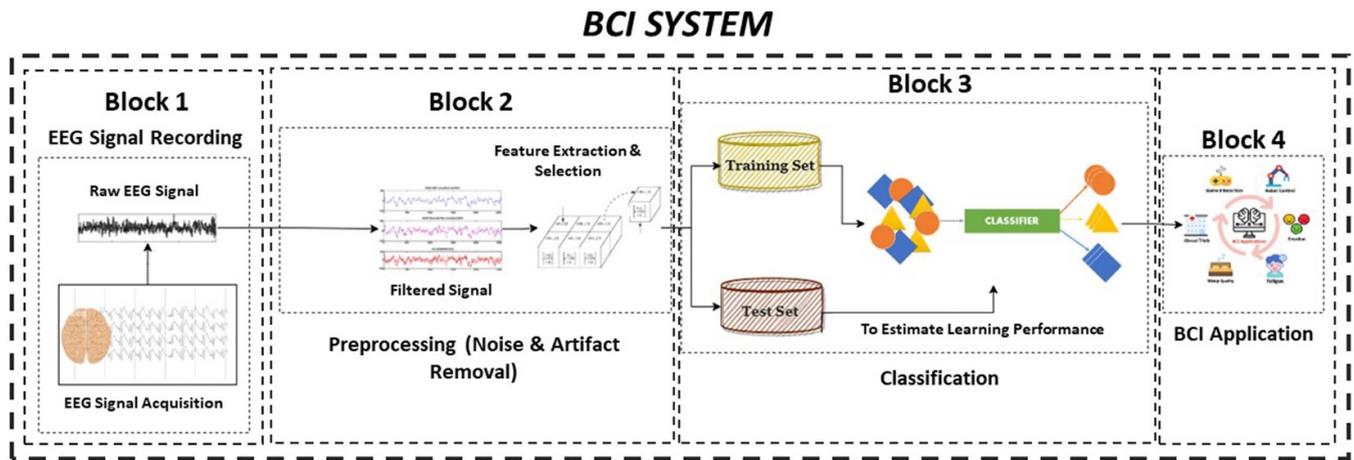
Figure 1. Architecture of brain computer interfacing system using EEG signal.

noteworthy progression in the domain of EEG signal classification, featuring the subsequent pivotal advancements:

- We present a lightweight DL model for EEG signal classification. This innovation is driven by the need to improve BCI accuracy and computational efficiency. Our model redefines the model complexity-performance trade-off, a notable contribution.
- In addition to accuracy, our research emphasizes model efficiency. We analyze edge device power consumption and inference time to ensure our model is suitable for real-time BCI applications. BCI technology resource efficiency is redefined by this approach.
- We closely compare our novel lightweight model to EEGNet, Deep ConvNet, and EffNet on the same EEG dataset. This comparison reveals our model's performance and suitability compared to benchmarks.
- We hardware-compare our model on popular embedded platforms like the Nvidia Jetson Nano and Xavier NX to verify its real-world applicability. This ensures that our innovation can be deployed on common hardware platforms, enabling real-time BCI solutions.

Our work is notable for its innovative design of a lightweight, efficient DL model for EEG signal classification, meticulous comparative analysis, and hardware-based evaluation. These innovations advance BCI technology and promise real-time BCI applications on embedded devices. The remainder of the paper is structured as follows. Section 2 describes the review of the literature on several existing BCI systems and their uses. Section 3 describes the experimental setup, the dataset that was employed, the hardware used for model evaluation, and the proposed technique. Section 4 describes the classification results of DL models and the evaluation of DL models on hardware. Section 5 discusses the results and section 6 brings the paper to a close.

2. Literature review

Table 1 shows the comparison of different DL techniques used for the classification of hand motions from neural correlates and the hardware implementation of the proposed DL model for evaluation of the model in terms of power and inference time. The studies exhibit variability in terms of their datasets, methods of pre-processing, and selection of classification models. In a particular case, a dataset that was recorded by the individual themselves was employed. Prior to analysis, the dataset underwent pre-processing, which involved applying a filtration technique using a software development kit (SDK) known as ABM. Subsequently, a DeepNet classifier was applied to the dataset, yielding an accuracy rate of 63.99%. The specific details regarding the hardware, power consumption, and inference time for this particular case were not provided [13]. In a separate investigation, the BCI competition III and IV datasets underwent independent component analysis (ICA) as a preliminary procedure. A classifier that integrated spatted convolutional neural network (CNN), discrete wavelet transform (DWT), and long short-term memory (LSTM) was employed, resulting in diverse accuracies across different scenarios (D1 = 56.2%, D2 = 89.3%, D3 = 90.2%, and D4 = 86.7%). Nevertheless, the CNN model that was developed was not subjected to hardware testing [14]. Furthermore, the Physionet dataset was subjected to pre-processing using EEGlab and advanced artifact removal (AAR) techniques, employing a neural network toolbox in MATLAB® as the classifier. This process yielded an accuracy rate of 65%. Like the preceding instances, this model was also not subjected to hardware testing, so the assessment of power consumption and inference times was not conducted [15].

Furthermore, within the context of a self-recorded dataset, the pre-processing steps encompassed the use of a bandpass filter and a DWT. subsequently, a feedforward neural network was employed, resulting in an accuracy rate of 90.09%. The hardware utilized for implementation was the Raspberry Pi

Table 1. Comparison of different techniques for the classification a of motor signals from neural correlates.

References	Dataset	Pre-processing	Classifier	Accuracy	Hardware	Power consumption (Watt)	Inference time (mSec)
[13]	Self-recorded	Filtration using SDK (ABM)	DeepNet	63.99%	—	—	—
[14]	BCI competition III & IV	ICA	Spatted CNN+ DWT+ LSTM	D1 = 56.2% D2 = 89.3% D3 = 90.2% D4 = 86.7%	—	—	—
[15]	Physionet	EEGlab +AAR (MATLAB® toolboxes)	Neural network toolbox MATLAB®	65%	—	—	—
[16]	Self-recorded	Bandpass filter + DWT	Feedforward NN	90.09%	Raspberry Pi 3B	5.77	—
[17]	Physionet	—	EEGNet	82.43%	ARM cortex M4F MCU	4.28 mJ	101
					ARM cortex M7	18.1 mJ	44
[18]	—	—	LSTM	87.89%	Mindreading photonic ULQ	0.2155	1500
[20]	BNCI Horizon 2020 reach and grasp dataset	Bandpass filter + CAR + ICA	Ensembled KNN	85.13%	—	—	—
[19]	Self-recorded	—	CCA	93.9%	ALINX xiliax zyng 7000	8–10	—
[21]	BCI competition IV a & III a	DWT + CSP	TSGL-EEGNet	81.34% 88.89%	—	—	—
[22]	Self-recorded + scientific data dataset	Low pass + band rejection + band stop filters	Hybrid model (LSTM + CNN)	84.96% 79.7%	Emotive EPOC + brainwear ®	—	750
[23]	HGD+ BCI IV 2a	—	TCNet-fusion (CNN)	94.1% 83.73%	—	—	—
[24]	BCI IV 2a + HGD[a]	—	MPEEGCBAM [b]	82.85% 95.45%	—	—	—
[25]	BCI competition IV	CAR + bandpass filter	LFANN [c]	91.58%	—	—	—
[26]	BCI competition IV 2a	3rd order butter worth bandpass filter	DSCNN [d]	85%	—	—	—
[27]	BCI competition III & IV a	Linear adaptive Filter	CSDA [e]	91.6% 91.1%	—	—	—

[a] High gamma dataset.
[b] Multibranch EEGNet convolutional block attention module.
[c] Light weight feature fusion network.
[d] Double branch CNN.
[e] Competitive swarm dragonfly algorithm.

3B, with a power consumption of 5.77 watts. However, the specific inference time was not provided [16]. In a separate investigation utilizing the physionet dataset, the EEGNet classifier demonstrated a level of accuracy amounting to 82.43%. The hardware employed in this study was an ARM cortex M4F MCU, which consumed 4.28 mJ of energy and had an inference time of 101 milliseconds [17]. The utilization of an ARM cortex M7 was also observed, exhibiting a power consumption of 18.1 mJ and an inference time of 44 milliseconds [17]. In an alternative methodology, a research investigation employed a LSTM classifier without explicitly mentioning the dataset used, resulting in an accuracy rate of 87.89%. The hardware utilized in this study was the mindreading photonic ULQ, which had a power consumption of 0.2155 Watts and an inference time of 1500 milliseconds, as reported in [18]. Moreover, a dataset that was recorded by the researchers themselves was used in a separate study. However, specific information regarding the pre-processing methods employed was not disclosed. Nevertheless, the use of canonical correlation analysis (CCA) as a classifier yielded a noteworthy accuracy rate of 93.9%. The hardware utilized in the study was the ALINX Xiliax Zyng 7000, which had a power consumption range of 8–10 Watts. Unfortunately, the specific data regarding inference time was not provided [19]. The research conducted utilized the BNCI Horizon 2020 reach and grasp dataset. The pre-processing stage involved applying a bandpass filter, common average reference (CAR), and ICA. The classification task employed an ensemble K-nearest neighbors (KNN) algorithm, resulting in an accuracy rate of 85.13% [20]. In an alternative situation, the BCI competition IV and III datasets were subjected to pre-processing using the DWT and common spatial patterns (CSP). A classifier named TSGL-EEGNet achieved accuracies of 81.34% and 88.89% [21]. However, the two aforementioned studies did not include an evaluation of the performance of the developed CNN models on hardware in terms of inference time and power consumption.

Moreover, in both a self-recorded dataset and a scientific data dataset, the pre-processing stage involved the application of low-pass filtering, band rejection filtering, and band-stop filtering. Additionally, a hybrid model that combined LSTM and CNNs was employed, leading to accuracies of 84.96% and 79.7% for the respective datasets. The hardware components utilized in this case study consisted of the emotive EPOC and brainwear® devices, which were found to have a power consumption of 750 Watts. However, it is important to note that no specific information regarding the inference time was provided [22]. In a separate investigation utilizing the HGD+ BCI IV 2a dataset, specific information regarding pre-processing procedures was not provided. However, a TCNet-fusion (CNN) classifier demonstrated an impressive accuracy rate of 94.1% [23]. In an alternative scenario, the BCI IV 2a and HGD1 datasets were employed without the inclusion of pre-processing techniques. However, the classifier MPEEGCBAM 2 demonstrated accuracies of 82.85% and 95.45% for the respective datasets [24]. In the BCI competition IV dataset, the pre-processing steps consisted of applying the CAR technique and a bandpass filter. The classification task was performed using a linear feedforward artificial

neural network (LFANN) classifier, which achieved an accuracy of 91.58% [25]. The BCI competition IV 2a dataset utilized a 3rd-order butterworth bandpass filter as a pre-processing technique. The classification model, DSCNN, achieved an accuracy rate of 85% [26]. The BCI competition III and IV studies employed a linear adaptive filter for pre-processing purposes. Additionally, a common spatial discriminant analysis (CSDA) classifier was utilized, resulting in accuracies of 91.6% and 91.1% [27]. However, it is important to note that the existing studies have not conducted hardware testing of the developed CNN models. As a result, there is a lack of information regarding power consumption and inference time, leaving this aspect as an unanswered inquiry. In this study, we addressed the research question by designing a lightweight CNN architecture. We evaluated the performance of our architecture on embedded devices, specifically the Jetson Nano and Jetson Xavier. Additionally, we compared the accuracy, power consumption, and inference time of our model with state-of-the-art models commonly employed for motion classification using neural correlates.

3. Methodology

In this section, the methods that are used in this research work are discussed in detail. For a fair comparison, the data is classified using both machine learning (ML) and DL techniques with a 70:30 train test split ratio. For the ML method, the data is firstly pre-processed by using the 4th order butterworth bandpass filter and 50 Hz. Notch filter. For further filtration of EEG signals the technique of ICA is used which removes the unwanted signals artifacts such as eye blink, muscle, line noise etc.

After filtration, EOG channels are removed, and trails of reach and grasp actions are epoch from continuous signals. From the epoch trails, the time domain features are extracted and used as input for ML classifiers.

For DL, data is minimally filtered and given to the DL model as an input in the form of time points into channels. The DL models are trained and tested on the raw signals and finally, these pre-trained models are used for evaluating the performances of our proposed model with other DL models on Jetson Xavier NX and Nano embedded platforms. Figure 2 shows the proposed methodology of this research work.

3.1. Dataset description

In this research, the dataset used is BNCI Horizon 2020 reach and grasp action decoding from Gel-based EEG electrodes [28]. Institute of Neural Engineering, Graz University of Technology, recorded the public dataset and G.tec USBamp/ladybird system gel-based EEG was used for recording data [28]. 15 people—5 women and 10 men—collected data. Each participant was 15–30 and right-handed.

EOG electrodes recorded eye movement, and EEG electrodes were carefully placed on a 5% grid. For data preprocessing 0.01–100 Hz 8th-order chebyshev filter was used.
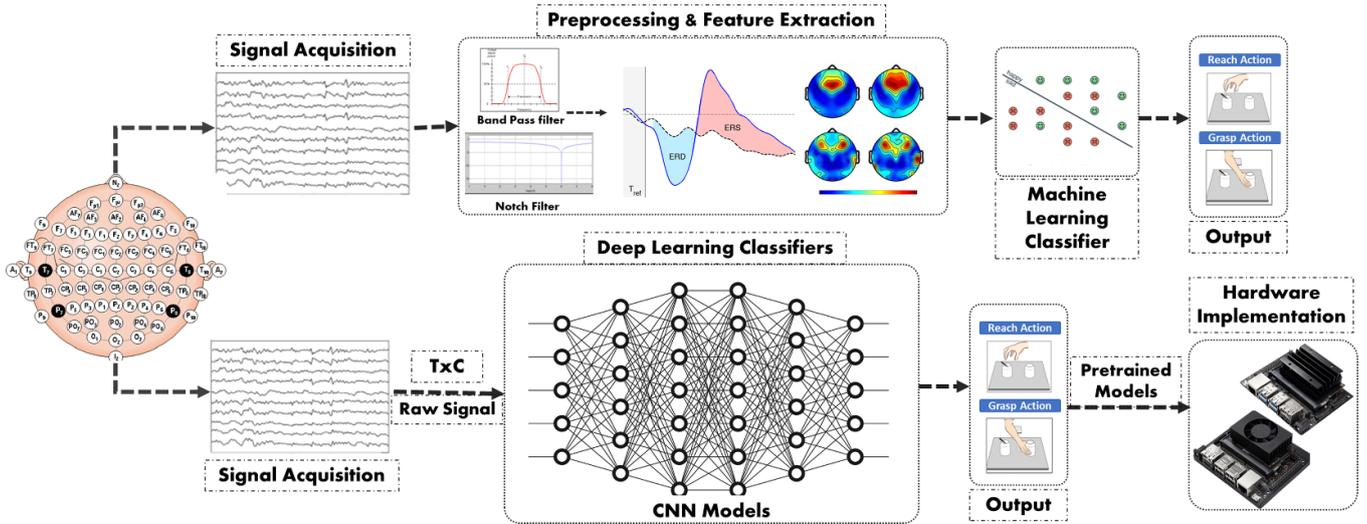
Figure 2. A flowchart of brain–computer interfacing classification of reach and grasp actions using machine learning and deep learning methods and implementation of dl models on embedded platforms.
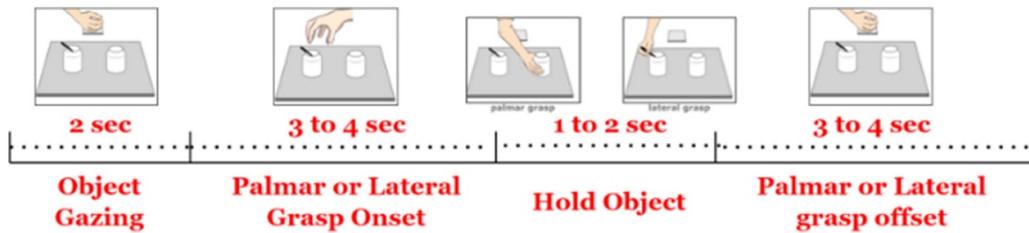

Figure 3. Timeline of each trial. The subject gazed at an object for 2 s, then performed the motion of reach & grasp. Grasp the object for 1–2 s. Finally returned the hand to starting position and take a rest of 4 s & start the next trail.

Sampling frequency: 256 Hz, 50 Hz notch filter applied [29]. Movement onset and grasping time recorded using FSR sensors [30]. Same experimental format as [31]: subjects seated, right hand on sensorized table. Objects used: jar with spoon (lateral grasp), empty jar (palmar grasp) [30]. Subjects gaze at object, perform self-initiated reach and grasp action [30] as shown in figure 3. Data recorded for 80 trials per condition in 4 runs [30]. Rest data recorded for 3 min at start of 1st run, after 2nd run, and at end of 4th run [29]. 6 EOG channels used for eye movement recordings [31].

3.2. Hardware

Edge devices are utilized to evaluate the performance of the suggested approach. This research employs the Nvidia Jetson Nano and Nvidia Jetson Xavier NX devices. The Nvidia Jetson Nano works on two power modes that are 5 W and 10 W with 4GB RAM. While on the other side, Jetson Xavier NX is a smaller size module as compared to Nano and it is also an energy-efficient AI platform. It operates on 3 power modes that are 10 W, 15 W, and 20 W WI5H 8 GB RAM, making it a more efficient embedded platform for developing remote BCI systems.

In this research, both hardware is used for evaluating the performance of our DL models in the form of inference time and power consumption. Table 2 shows the specification of both Jetson Nano and Jetson Xavier NX modules.

3.3. Proposed model

This study proposes a DL CNN model that classifies the reach and grasp motion from neural correlates. Figure 4 illustrates the general layout of the new model that has been proposed. The model consists of 2D convolutional layers along with average pooling layers, and separable convolutional layers, followed by the fully connected layer. Firstly, the input layer is used in the model in which the input shape used for the model is time series data in the form of channels x time points. After the first layer, first 2D convolutional block which is the 2D convolutional layer along with the batch normalization (BN) layer is used. It consists of 8 filters of size (1, 64) which moves along the time axis for extracting the feature values the feature map of the first block serves as input for the second block. The second block consists of a separable convolutional block, utilizing a 2D separable convolutional layer with 32 filters of size (1, 64). This layer performs depth and point-wise convolutions, extracting time features in the process. Following the separable 2D convolutional layer is a BN layer, accompanied by the activation function 'ELU'. After the BN layer, an average pooling layer is applied with a filter size of (1, 8). This layer reduces the complexity of the feature map then a

Table 2. Specifications of Nvidia Jetson Nano and Xavier specifications.

	NVIDIA Jetson Nano	NVIDIA Jetson Xavier
GPU	128-CORE NVIDIA MAXWELL™ ARCHITECTURE-BASED GPU	NVIDIA VOLTA ARCHITECTURE WITH 384 NVIDIA CUDA® CORES AND 48 TENSOR CORES
CPU	QUAD-CORE ARM® A57	6-CORE NVIDIA CARMEL ARM®V8.2 64-BIT CPU 6 MB L2 + 4 MB L3
Memory	4 GB 64-BIT LPDDR4; 25.6 GIGABYTES/SECOND	8 GB 128-BIT LPDDR4X @ 51.2GB/S
OS supported	LINUX FOR TEGRA®	LINUX FOR TEGRA®
Module size	70 MM × 45 MM	103 MM × 90.5 MM × 34.66 MM
Power	5 W & 10 W	10 W, 15 W & 20 W
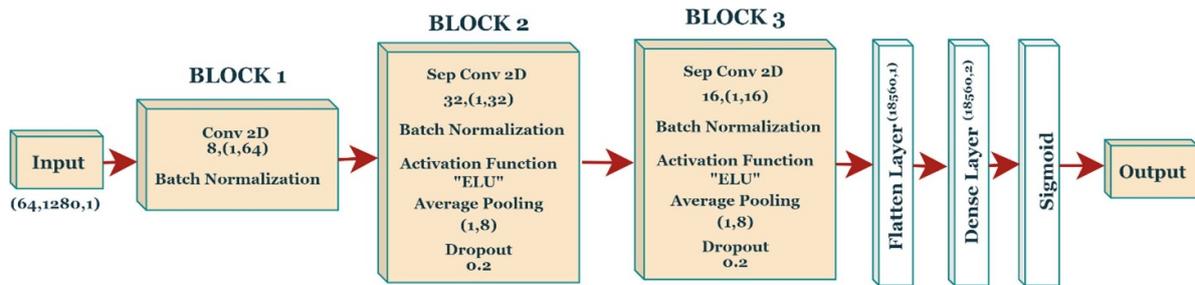
Figure 4. The architecture of the proposed CNN model.

dropout layer is applied with a 0.2 dropout value. 3rd block is again a separable convolutional block in which a 2D separable convolutional layer with 16 filters of size (1, 16) is used and then again apply BN layer with activation function 'ELU'. Then again average pooling layer is applied for reducing the complexity of the feature map. The next layer is a flattened layer (FL) that combines the output of block 3 into a vector & the layer after FL is a fully connected layer. Lastly, the 'Sigmoid' layer is added in the model that predicts the probability distribution of output classes i.e., reach & grasp actions.

4. Results

This section describes the results of the proposed DL model, a comparison of classification accuracies of our proposed model with previously used lightweight DL models along with ML techniques that are commonly used for the EEG signals classification, and finally the results of the evaluation of DL models on hardware based on power consumption and inference time.

For comparison purposes, the typically used ML models for EEG classification are also used for the classification of reach and grasp action. For the ML classification, the data is firstly filtered using Butterworth 4th order bandpass filter with cutoff frequencies 0.1–35 Hz, then CAR filter is applied, and after that for removing the eye blink, line noise, muscle, heart artifacts ICA. The trails of reach and grasp actions are epoch using [−2 3] sec window of interest and after that time domain features are extracted from epoch trails. The hand-crafted feature matrix is then used as an input for LDA, SVM, and KNN classifiers.

4.1. Classification

After extracting the features, ML classifiers are applied to them one by one for each subject. LDA, SVM, and K-NN models are used in this research. With 70:30 split ratios, all features are extracted from the preprocessed data subject of each subject. The test and training accuracies of all 15 participants on LDA, SVM, and K-NN classifiers are shown in table 3. The findings reveal that the K-NN classifier has the highest overall classification accuracy using hand-crafted features when compared to the LDA and SVM classifiers. Table 3 shows a comparison of accuracies of ML classifiers.

Table 3. Classification accuracies of 15 subjects using, LDA, SVM, and K-NN machine learning models.

Classifiers/subjects	Linear discriminant analysis classification (%)		Support vector machine classification (%)		K-nearest neighbour classification (%)	
	Training	Test	Training	Test	Training	Test
G01	65.2	63.7	67.8	66	79.8	77.6
G02	63.4	60.8	68	67.8	77.4	76.0
G03	64.2	60.8	66.4	65.2	84.4	81.4
G04	71.2	69.1	69.6	68.8	80.2	79.1
G05	63.4	60.3	67.4	68.3	81.3	79.6
G06	65.3	64.9	70.2	69.1	81.4	80.7
G07	60.6	56.2	67.4	67.2	76.9	75.8
G08	55.2	53.8	62.9	61.0	78.5	83.8
G09	64.5	62.8	70.6	68.4	74.4	74.6
G10	60.2	60.1	65.4	62.9	72.6	76.3
G11	60.2	61.6	68.5	66.5	77.4	82.8
G12	58	55.7	66.6	68.0	76.3	80.9
G13	66.2	60.7	73.5	72.1	81.4	82.2
G14	62.7	58.6	67.7	67.1	81.7	84.2
G15	66.2	62.4	65.3	62.6	79.8	82.2
Average	**63.1 $\pm$ 3.86**	**60.77 $\pm$ 3.80**	**67.82 $\pm$ 2.53**	**66.73 $\pm$ 2.86**	**78.9 $\pm$ 3.11**	**79.81 $\pm$ 3.11**

Table 4. Comparison of CNN model sizes and no of parameters.

Model name	Size (KB)	No. of parameters
Proposed model	319	20 833
EEGNet	107	2753
DeepConvNet	2210	184 726
EffNet	4760	399 073

For the DL model, raw EEG data is used which is minimally pre-processed by applying a 4th-order butterworth bandpass filter with cutoff frequencies of 0.1–35 Hz. After that, the DL CNN model uses segmented data in the form of timepoint x channels with a 70:30 training/test split ratio.

This dataset is also tested on EEGNet [32], EffNet [33], and DeepConvNet [34] CNN models that are lightweight, and EffNet is specially designed for edge device perspective. Table 4 compares the size of the proposed models along with the other 3 DL models in KB's and a total number of parameters. EffNet model which is specially designed for edge device perspective shows a larger size as compared to EEGNet that is normally used for the classification of EEG signals and our proposed model also the total number of parameters of EEGNet is very small as compared to EffNet and DeepConvNet models and the number of parameters of our proposed model is very close to EEGNet model.

Table 5 represents the classification accuracies along with the F1 score, precision, and recall of our proposed DL model on all 15 subjects (subject-wise). Results show that our model shows maximum accuracy of 97.83% on subject 15 and minimum accuracy of 84.37 on subject 6. Table 6 shows the classification accuracies of the same subjects using EEGNet, EffNet, and DeepConvNet models. The average accuracies achieved on the proposed model, EEGNet, EffNet, and DeepConvNet models are 92.44, 90.76, 81.69, and 92.89

respectively. Results show that our model achieves the best accuracy with a smaller number of parameters than all 4 DL models.

For comparison previously proposed DL models for the classification of motor signals from neural correlates are also evaluated on the same dataset the models used are EEGNet [32], DeepConvNet [34] and EffNet [33].

One-way ANOVA is applied to the accuracies of all DL models and results show that there is a significant difference between accuracies ($p < 0.001$). For finding the multiple comparisons between different groups Tukey honestly significant difference (HSD) test is applied which shows that (95% CI [-2.785:6.139], [-4.1985:4.0038], $p = 0.752, .993$)

there is no significant difference between the accuracies of the proposed model and EEGNet and between the proposed model and DeepConvNet model results. In the results of EffNet and the proposed model (95% CI [6.288:15.21], $p < 0.01$) there is a significant difference between the classification accuracies. The overall proposed model performs better than the EEGNet and EffNet but there is a very small significant difference between classification results with the DeepConvNet model.

4.2. Hardware implementation

The DL models are then evaluated on the Nvidia Jetson Nano and Nvidia Jetson Xavier NX embedded platforms. Both the development boards are operated on maximum power mode, and the power consumption and inference time are calculated for all the subjects on the proposed and remaining 3 DL models.

4.2.1. Power consumption. Table 7 shows the power consumption results of 15 subjects when the models are tested on Jetson Nano and Jetson Xavier NX on maximum power mode.

Table 5. Classification accuracies of all the 15 subjects on the proposed DL CNN model.

Subjects	Accuracy	Precision	Recall	F1-score
G01	94.73	95	95	95
G02	96.67	96	96	96
G03	93.75	93	92	92
G04	91.18	90	90	90
G05	88.12	79	78	78
G06	84.37	84	84	84
G07	93.55	93	92	92
G08	97.56	96	96	96
G09	90.42	91	90	90
G10	95.81	96	96	96
G11	89.06	91	91	91
G12	95.83	96	96	96
G13	85.11	85	85	85
G14	92.56	92	91	91
G15	97.83	98	98	98

Table 6. Comparison of classification accuracies of the proposed model, EEGNet, DeepConvNet and EffNet.

Subjects	Proposed model	EEGNet	DeepConvNet	EffNet
G01	94.73	92	95.31	89.06
G02	96.67	90	95	80.88
G03	93.75	90.62	90.62	89.06
G04	91.18	88.36	98.36	77.05
G05	88.12	93.75	90.62	75
G06	84.37	92.19	84.32	79.69
G07	93.55	90.32	98.89	85.48
G08	97.56	98.44	87.5	88.74
G09	90.42	80	86.83	84.132
G10	95.81	93.55	93.55	84.84
G11	89.06	92.19	92.19	84.38
G12	95.83	91.88	95.13	82.19
G13	85.11	88.44	93.68	79.37
G14	92.56	92.86	94.64	69.64
G15	97.83	86.77	96.77	75.8
Average	**92.44 ± 4.30**	**90.76 ± 4.06**	**92.89 ± 4.23**	**81.69 ± 4.22**

The results of one-way ANOVA show that there is a significant difference ($p < 0.001$) in power consumption on Jetson Nano. The multiple comparisons Tukey HSD test shows that there is no significant difference between the proposed model and EEGNet (95% CI [$-4.91.61:340.14$], $p = 0.963$) power consumption results on the Jetson Nano, while there is a significant difference of power consumption shows between the proposed model and EEGNet (95% CI [342.09:750.85], $p < 0.001$) on Jetson Xavier platform. For DeepConvNet and EffNet with the proposed model (95% CI [$-952.74:-120.99$], [$-2220:-1389$], $p = 0.06, < 0.001$) there is a significant difference between the results of EffNet and the proposed model but the value of p for DeepConvNet is also less for Jetson Nano. For Jetson Xavier, the power consumption results of EffNet and DeepConvNet with the proposed model (95% CI [$-1195:-786$], [$-75.58:333.18$], $p < 0.001$, $p = 0.350$) shows that EffNet shows a significant difference but the DeepConvNet shows no significant difference. The average power consumption on Jetson Nano and Jetson Xavier

of the proposed model, EEGNet, DeepConvNet, and EffNet are 5600.2 mW, 5770.4 mW, 5675.93 mW, 5223.93 mW, 6137.07 mW, 5641.6 mW, and 7405.2 mW, 6761.4 mW respectively.

4.2.2. Inference time. Table 8 shows the inference time of the proposed model along with DeepConvNet, EffNet, and EEGNet. The DL models use less time of inference on Jetson Xavier as compared to Nano because of the high RAM and the multiple numbers of processors. The result of one-way ANOVA shows that there is a significant difference ($p < 0.001$) in inference time for Jetson Nano. The multiple comparisons Tukey HSD test shows that (95% CI [$-7.34: 3.07$], [$-11.07: -66$], $p = 0.7, .021$) there is no significant difference between the inference time of our proposed model with EEGNet and DeepConvNet on Jetson Nano platform while there is a significant difference between (95% CI [$-16.14:-5.73$], $p < 0.001$) the proposed model and EffNet

Table 7. Power consumption of Nvidia Jetson Nano and Nvidia Jetson Xavier NX on max power mode of proposed CNN model along with EEGNet, DeepConvNet, and EffNet.

| Subjects | Proposed model | | EEGNet | | DeepConvNet | | EffNet | |
	NVIDIA JETSON NANO (mW)	NVIDIA JETSON XAVIER NX (mW)	NVIDIA JETSON NANO (mW)	NVIDIA JETSON XAVIER NX (mW)	NVIDIA JETSON NANO (mW)	NVIDIA JETSON XAVIER NX (mW)	NVIDIA JETSON NANO (mW)	NVIDIA JETSON XAVIER NX (mW)
G01	5980	5843	5906	5802	6724	6021	7251	6640
G02	5388	5657	5888	5381	6299	5739	7440	6675
G03	5680	5779	4729	5259	5218	5861	7228	6675
G04	5475	5890	4654	5614	5633	5779	7022	6715
G05	5955	5064	5662	5292	6219	5739	7280	6716
G06	5802	5239	6047	5639	5590	5503	7440	6795
G07	5599	4779	6047	5055	5846	5732	7371	6832
G08	5477	5073	6067	5762	6279	5290	7652	6675
G09	5721	4570	5437	5139	6178	5503	7262	6759
G10	5955	5739	5952	5055	6914	5342	7747	6836
G11	4816	5687	5824	5055	5833	5983	7560	6878
G12	5876	5698	5558	5055	6097	5422	7291	6756
G13	5874	5799	5864	5883	6026	5299	7587	6793
G14	6222	5698	5864	5136	6138	5580	7400	6838
G15	4183	5666	5640	4932	7062	5831	7547	6838
Average	**5600.2**	**5478.0**	**5675.9**	**5337.3**	**6137.1**	**5641.6**	**7405.2**	**6761.4**

Table 8. Inference time of Nvidia Jetson Nano and Nvidia Jetson Xavier NX on max power mode of proposed CNN model along with EEGNet, DeepConvNet, and EffNet.

| Subjects | Proposed model | | EEGNet | | DeepConvNet | | EffNet | |
	NVIDIA JETSON NANO (sec)	NVIDIA JETSON XAVIER NX (sec)	NVIDIA JETSON NANO (sec)	NVIDIA JETSON XAVIER NX (sec)	NVIDIA JETSON NANO (sec)	NVIDIA JETSON XAVIER NX (sec)	NVIDIA JETSON NANO (sec)	NVIDIA JETSON XAVIER NX (sec)
G01	16	2	18	2	24	4	30	6
G02	19	2	20	2	21	4	21	5
G03	19	2	19	2	21	4	36	6
G04	16	2	18	2	40	4	32	5
G05	13	2	15	4	20	4	31	5
G06	14	2	13	2	22	4	19	5
G07	16	2	19	2	40	4	23	5
G08	15	2	18	2	18	4	20	5
G09	16	2	12	2	16	4	20	5
G10	16	2	15	3	16	4	28	5
G11	16	2	16	2	18	4	26	5
G12	18	2	23	2	16	4	26	6
G13	16	2	31	2	18	4	36	5
G14	14	1	20	2	16	3	28	5
G15	18	2	17	2	24	4	30	5
Average	**16.1**	**1.9**	**18.3**	**2.2**	**22**	**3.9**	**27.1**	**5.2**

inference time. The results of One way ANOVA on Jetson Xavier also show the same trend for EffNet and EEGNet models (95% CI $[-3.53: -3.00]$, $[-.33: .20]$, $p < 0.001$, $p = 9.11$), but for DeepConvNet and proposed model there is a significant difference between the inference time on Jetson Xavier platform (95% CI $[-2.21: -1.73]$, $p < 0.001$). From the results, it is clear that the proposed CNN model achieves lesser inference time on Nvidia Jetson xavier nx (subject-wise), the average inference time on jetson Nano and Xavier NX of the proposed

model, EEGNet, EffNet, and DeepConvNet are 16.13 sec, 1.9 sec, 27.07 sec, 2 sec, 18.27 sec, 3.9 sec and 22 sec, 5.2 sec respectively.

5. Discussion

The main aim of this research is the development of an efficient and lightweight DL model that can be used for the classification of hand motion using neural correlates. For this
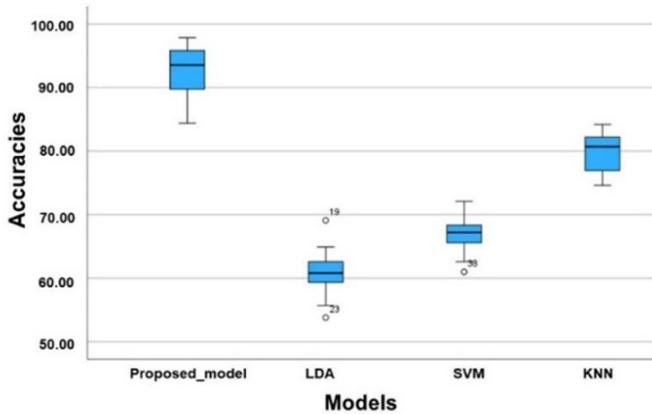
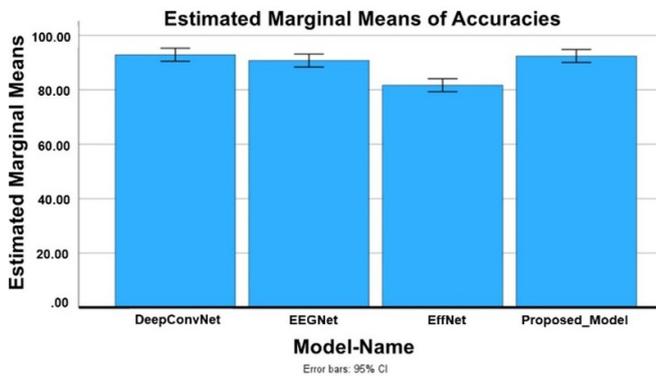**Figure 5.** Boxplot of classification accuracies of proposed DL model with machine learning models (LDA, SVM, KNN).

Figure 6. Average classification results across 15 subjects on the proposed CNN model along with EEGNet, DeepConvNet, and EffNet.

Figure 7. Average power consumption across 15 subjects of the proposed CNN model along with EEGNet, DeepConvNet, and EffNet with 95% CI on NVIDIA Jetson Xavier NX.

estimated marginal means of accuracies of all 4 models with a 95% confidence interval the results clearly show that there is a slight difference between the average actuaries of proposed model and Deep ConvNet model. For testing the best performance between the 4 DL in terms of accuracy, inference time, and power consumption 2-way multivariant ANOVA (MANOVA) between subjects is applied which shows that $p < 0.001$ for accuracy, inference time, and power. The results shows that the proposed model and DeepConvNet performs better in term of classification as compared to EEGNet and EffNet models (95% CI $[-4.9185 : 4.00]$, $[-2.78 : 6.14]$, $[-15.21 : -6.29]$, $p = .993, .752, <0.001$).

The performance of the DL models is then evaluated on the hardware in terms of power consumed by the models and inference time. On the Nvidia Jetson Nano platform, the proposed model and EEGNet model show no significant difference as compared to EffNet and DeepConvNet models (95% CI $[-49\,161 : 340.14]$, $[-2220.88 : -1389.12]$, $[-952.74 : -120.99]$, $p = .963, <0.05, <0.001$) while on Nvidia Jetson Xavier proposed model shows no significant difference with EEGNet and DeepConvNet while it shows significance difference with EffNet models (95% CI $[-141.72 : 424.66]$, $[-446.05 : 120.32]$, $-1565.86 : -999.48]$, p = .553, .431, <0.001$).

Overall, the best model that shows the least performance on both hardware in terms of power consumption is the EffNet model as compared to the proposed model ($p < 0.001$). In terms of Inference time on Nvidia Jetson Nano the proposed model performs better than the DeepConvNet and EffNet models as there is a significant difference between them (95% CI $[-11.07 : -.66]$, $[-16.14 : -5.72]$, $p < 0.05$, p < 0.05$). While there is no significant difference between proposed model and EEGNet (95% CI $[-7.34 : 3.07]$, $p = 0.7$). On the Nvidia Jetson.

Xavier NX the same trend is repeated between the proposed model and DeepConvNet, EffNet models ($p < 0.001$) but the significant difference between EEGNet and the proposed Model is less ($p = 0.25$).

So, in terms of accuracy, the proposed model and DeepConvNet models perform better as shown in figure 6, in terms of power consumption and inference time on Nvidia

purpose, a lightweight model of DL is designed and tested on the raw EEG signals of reach and grasp actions. For the comparison purpose, 3 states of the art ML models are also used that are previously used in the literature for the classification of brain signals of different hand movements i.e., LDA, SVM, and K-NN models with 70:30 split ratios. 2-way ANOVA is applied to the classification accuracies of these 3 ML classifiers with the proposed DL model. The results show that there is a significant difference (95% CI [28.22:35.11], [22.25:29.15], [9.17: 16.07], $p < 0.002$, <0.001, <0.001) present between accuracies of proposed model and LDA, SVM, K-NN. According to the results, our models performed better than the conventionally used ML models for EEG signals classification. So the method of DL is more efficient as compared to the conventional ML approaches and also it uses less preprocessing of data so overall it requires less computational power. After testing the results of the proposed model with ML models, the model is then tested with three DL proposes that were previously proposed in the literature i.e., EEGNet, DeepConvNet, and EffNet. The results of the classification accuracies show that the proposed model performs better than the EEGNet and EffNet models (figure 5). But there is a very slight difference between the average accuracies of DeepConvNet and the proposed model as figure 6 shows the
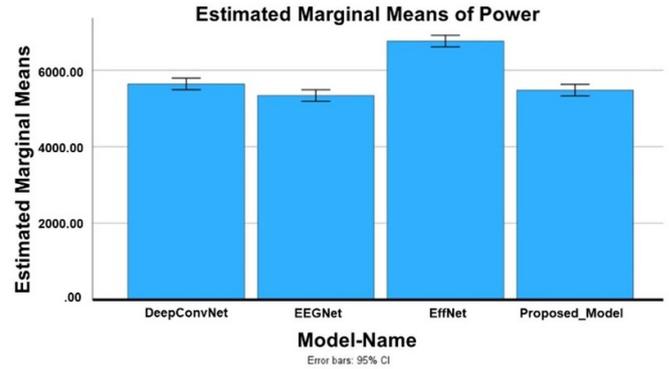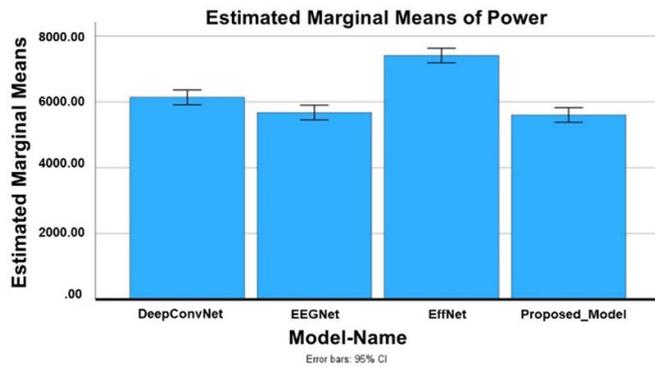
Figure 8. Average power consumption across 15 subjects of the proposed CNN model along with EEGNet, DeepConvNet, and EffNet with 95% CI on NVIDIA Jetson Nano.
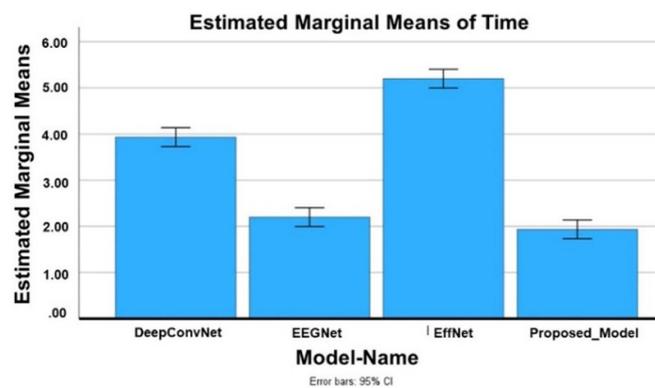

Figure 9. Average Inference across 15 subjects of the Proposed CNN model along with EEGNet, DeepConvNet, and EffNet with 95% CI on NVIDIA Jetson Xavier NX.
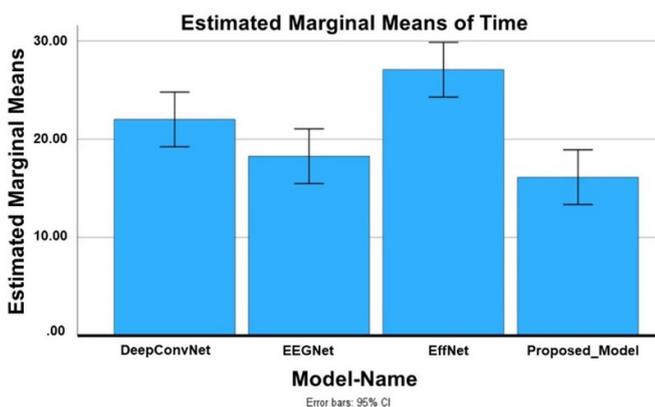

Figure 10. Average inference time across 15 subjects of the proposed CNN model along with EEGNet, DeepConvNet, and EffNet with 95% CI on NVIDIA Jetson Nano.

Jetson Nano proposed model and EEGNet performs better as shown in figures 8 and 10. So overall on jetson Nano the proposed model performs better as compared to the other 3 DL models.

On jetson Xavier EEGNet uses less power as compared to the proposed model as shown in figure 7 but in terms of the

inference time proposed model performs better as compared to the other 3 DL models shown in figure 9. Overall, in terms of accuracy, power, and time model proposed in this work shows the best result on 3 parameters on Jetson Nano and on 2 parameters (accuracy, inference time) on Jetson Xavier.

6. Conclusion

In this research, the dataset of BNCI horizon 2020 is utilized to classify hand motions using ML and DL models. The primary focus of this research is the development of an efficient and lightweight DL model and evaluating its performance on the embedded platforms. Additionally, we compared our proposed models with ML models (LDA, SVM, KNN) and three previously used DL models (EEGNet, DeepConvNet, EffNet).

To make the model lightweight, we leverage the separable convolution layers which decomposes convolution operation in depth-wise and point wise convolutions. Separable convolution layer significantly reduces the number of computations and enhances the representational capacity of the network. The results show that our proposed models provide better results as compared to ML models and from DL models our models perform better than EEGNet and EffNet models. The classification results of the Proposed model and DeepConvNet models are approximately the same but because of the higher model size, the DeepConvNet model uses more computational power and inference time on both embedded platforms (Nvidia Jetson Nano & Nvidia Jetson Xavier).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

ORCID iD

Jamshed Iqbal ⬤ https://orcid.org/0000-0002-0795-0282

References

[1] World Health Organization 2022 WHO disability (available at: www.who.int/news-room/fact-sheets/detail/disability-and-health) (Accessed 9 Jan 2023)
[2] Bank Group W 2022 Disability inclusion and accountability framework (available at: www.worldbank.org)
[3] Berns G S and Moore S E 2012 A neural predictor of cultural popularity *J. Consum. Psychol.* **22** 154–60
[4] Kaiser J *et al* Self-initiation of EEG-based communication in paralyzed patients (available at: www.elsevier.com/locate/clinph)
[5] Ramadan R A and Refat S 2015 Basics of Brain Computer Interface *Brain-Computer Interfaces: Current Trends and*

Applications eds A E Hassanien and A T Azar (Springer) pp 31–50

[6] Mahajan R and Bansal D 2017 Real time EEG based cognitive brain computer interface for control applications via arduino interfacing *Proc. Comput. Sci.* **115** 812–20

[7] Bin Nasir T, Lalin M A M, Niaz K, Karim M R and Rahman A 2021 EEG based human assistance rover for domestic application *Int. Conf. on Robotics, Electrical and Signal Processing Techniques* pp 461–6

[8] Fujiyoshi H, Hirakawa T and Yamashita T 2019 Deep learning-based image recognition for autonomous driving *IATSS Research* **43** 244–52

[9] Roy P P, Kumar P and Chang V 2020 A hybrid classifier combination for home automation using EEG signals *Neural Comput. Appl.* **32** 16135–47

[10] Ieracitano C, Mammone N, Hussain A and Morabito F C 2022 A novel explainable machine learning approach for EEG-based brain-computer interface systems *Neural Comput. Appl.* **34** 11347–60

[11] Fu B, Li F, Niu Y, Wu H, Li Y and Shi G 2021 Conditional generative adversarial network for EEG-based emotion fine-grained estimation and visualization *J. Vis. Commun. Image Represent.* **74** 102982

[12] Ramadan R A, Refat S, Elshahed M A and Ali R A 2015 Basics of brain computer interface *Intell. Syst. Ref. Libr.* **74** 31–50

[13] Aldayel M, Ykhlef M and Al-Nafjan A 2020 Deep learning for EEG-based preference classification in neuromarketing *Appl. Sci.* **10** 1525

[14] Yang J, Yao S and Wang J 2018 Deep fusion feature learning network for MI-EEG classification *IEEE Access* **6** 79050–9

[15] Huong N TM, Linh H Q and Khai L Q 2018 Classification of left/right hand movement EEG signals using event related potentials and advanced features *6th Int. Conf. on the Development of Biomedical Engineering in Vietnam (BME6)* pp 209–15

[16] Ramírez-Quintana J, Macias-Macias J, Corral-Saenz A and Chacon-Murguia M 2019 Novel SSVEP processing method based on correlation and feedforward neural network for embedded brain computer interface *Pattern Recognition: 11th Mexican Conference, MCPR 2019* (*Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*) vol 11524 pp 248–58

[17] Wang X, Hersche M, Tömekce B, Kaya B, Magno M and Benini L 2020 An accurate EEGNet-based motor-imagery brain-computer interface for low-power edge computing (available at: http://arxiv.org/abs/2004.00077)

[18] ACM Special Interest Group on Design Automation 2020 IEEE circuits and systems society, IEEE Council on Electronic Design Automation, and Institute of Electrical and Electronics Engineers *ASP-DAC 2020: 25th Asia and South Pacific Design Automation Conf.: Proc. (13–16 January)* (National Convention Center)

[19] Na R, Hu C, Sun Y, Wang S, Zhang S, Han M, Yin W, Zhang J, Chen X and Zheng D 2021 An embedded lightweight SSVEP-BCI electric wheelchair with hybrid stimulator *Digit. Signal Process.* **116** 103101

[20] Kanuparthi B, Turlapaty A C and Gokaraju B 2021 An ensemble approach for classification of reach and grasp movements based on EEG signals *Proc.—Applied Imagery Pattern Recognition Workshop* vol 2021

[21] Deng X, Zhang B, Yu N, Liu K and Sun K 2021 Advanced TSGL-EEGNet for motor imagery EEG-based brain-computer interfaces *IEEE Access* **9** 25118–30

[22] Mwata-Velu T, Ruiz-Pinales J, Rostro-Gonzalez H, Ibarra-Manzano M A, Cruz-Duarte J M and Avina-Cervantes J G 2021 Motor imagery classification based on a recurrent-convolutional architecture to control a hexapod robot *Mathematics* **9** 606

[23] Musallam Y K, AlFassam N I, Muhammad G, Amin S U, Alsulaiman M, Abdul W, Altaheri H, Bencherif M A and Algabri M 2021 Electroencephalography-based motor imagery classification using temporal convolutional network fusion *Biomed. Signal Process. Control* **69** 102826

[24] Altuwaijri G A and Muhammad G 2022 Electroencephalogram-based motor imagery signals classification using a multi-branch convolutional neural network model with attention blocks *Bioengineering* **9** 323

[25] Yu Z, Chen W and Zhang T 2022 Motor imagery EEG classification algorithm based on improved lightweight feature fusion network *Biomed. Signal Process. Control* **75** 103618

[26] Ma W, Gong Y, Xue H, Liu Y, Lin X, Zhou G and Li Y 2022 A lightweight and accurate double-branch neural network for four-class motor imagery classification *Biomed. Signal Process. Control* **75** 103582

[27] Rajesh Kumar D T, Mahalaxmi U S B K, Ramakrishna M M and Bhatt D D 2023 Optimization enabled deep residual neural network for motor imagery EEG signal classification *Biomed. Signal Process. Control* **80** 104317

[28] BNCI Horizon 2020 BNCI Horizon 2020 (available at: http://bnci-horizon-2020.eu/)

[29] Oostenveld R and Praamstra P The ®ve percent electrode system for high-resolution EEG and ERP measurements (available at: www.elsevier.com/locate/clinph)

[30] Breitwieser C, Kreilinger A and Neuper C The TOBI hybrid BCI-the data acquisition module (available at: www.researchgate.net/publication/267801490)

[31] Schwarz A, Pereira J, Kobler R and Müller−putz G R 2020 Unimanual and bimanual reach-and-grasp actions can be decoded from human EEG *IEEE Trans. Biomed. Eng.* **67** 1684–95

[32] Lawhern V J, Solon A J, Waytowich N R, Gordon S M, Hung C P and Lance B J 2018 EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces *J. Neural Eng.* **15** 056013

[33] Freeman I, Roese-Koerner L and Kummert A 2018 EffNet: an efficient structure for convolutional neural networks (available at: http://arxiv.org/abs/1801.06434)

[34] Schirrmeister R T *et al* 2017 Deep learning with convolutional neural networks for EEG decoding and visualization *Hum. Brain Mapp.* **38** 5391–420