Application of Deep Learning with Brain Computer Interfaces

Michael Fairbanks
Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA 56267
fairb205@morris.umn.edu

ABSTRACT

A Brain-Computer Interface is the culmination of neuroscience, computer science, and engineering all in one. A Brain-Computer Interface can potentially restore many functions people have lost through brain trauma, potentially curing disabilities like Alzheimer's and paralysis. After decades of technological progression this technology has also found its way into the entertainment industry hoping to further improve upon current Virtual Reality technology. This paper will explore methods researchers are utilizing Deep Learning algorithms to solve common problems associated with the brain.

1. INTRODUCTION

The Brain-Computer Interface (BCI) has been an up and coming technology for the past couple of decades, and only in recent years is it seeing a growth in popularity. This technology has a lot of potential as it allows the user to control external devices with just a thought. BCI research and development primarily continues to advance current prosthetic technology, allowing more dynamic functions and restoring senses like sight and hearing, but has been theorized to be a potential solution to other disabilities such as paralysis and Alzheimer's.

This technology recently is also being developed for use in the entertainment industry through the use of current Virtual Reality technology. Brain-Computer Interfaces alongside Virtual Reality can bring about a much more immersive experience with a variety of experiences depending on the usage of active or passive BCI's. An active BCI would allow the user to explicitly issue commands versus a passive BCI that just monitors the user's state and is reflected through the virtual interface.

The main issue with the technology's advancement is the variability in brain signals from person to person as well as the numerous artifact issues plaguing EEG data, so engineers and neuro-scientists have been utilizing deep learning algorithms in an attempt to automate the analysis of varying electroencephalogram data.

In this paper I describe and analyze the many facets of Brain-Computer Interfaces, and how this technology is used to predict mental states, and in virtual reality systems for neuro-rehabilitation.

UMM CSci Senior Seminar Conference, November 2021 Morris, MN.

2. BACKGROUND

Due to the general complexity of Brain-Computer Interface technology, the following section will cover key concepts and definitions relating to Electroencephalograms (EEG), the human brain, Deep Learning, for a more in depth understanding on the relationship between these aspects and Brain-Computer Interfaces.

2.1 The Brain

The human brain consists of three parts: cerebrum, cerebellum, and the stem, with the cerebrum being the most important for learning due to its information processing capabilities. The cerebrum can then be divided into four lobes: frontal, temporal, parietal, and occipital. The frontal lobe is responsible for executive functions such as planning or reasoning. The Temporal lobe processes sensory input for appropriate retention of visual memory, language comprehension and emotion association. The parietal lobe as a whole just handles language processing just as the occipital lobe handles only vision. Every part of the brain is made of billions of neurons connected by fiber pathways constantly sending information back and forth creating memories and establishing connections. It is also these sections of the cerebrum that help produce the most important data for electroencephalograms. [3]

2.2 Electroencephalograms

The brain commands the body by continuously sending ionic currents to motor neurons located in the spinal cord, causing muscles to contract. Each individual action produces a different current, whether it's just thinking, or only moving a finger to type.

Electroencephalograms (EEG) are a medical technology used to monitor and record the voltage fluctuations in those ionic currents. EEG diagnostics generally focus on either event-related potentials or its spectral content. Event-related potentials refers to fluctuations due to stimulus onset of something as simple as pressing a button, while spectral content refers to the type of neural oscillations in terms of frequency. The frequency of oscillations can be categorized into Delta, Theta, Alpha, Beta, Gamma, and Mu bands, each capable of communicating what action is being performed. Figure 1 shows relevant categories. For example, Theta waves are most commonly found when a person is actively trying to repress a response or action, while Gamma waves are found during cross-modal sensory processing, which is monitoring when a subject is using two or more senses at the same time such as sound and sight.

Band	Frequency (Hz)	Location	Activity
Delta	0.5–4 Hz	Frontal	Deep sleep
Theta	4–8 Hz	Various	Drowsiness, light sleep
Alpha	8–13 Hz	Posterior region of head	Relaxed
Beta	13–30 Hz	Left and right side, symmetrical distribution, most evident frontally	Active thinking, alert
Gamma	>30 Hz	Somatosensory cortex	Hyperactivity

Figure 1: Frequency Categorization [2]

The downside to electroencephalograms is how sensitive they are to external stimuli. Alongside the neural oscillations on the report are artifacts, which are electrical signals detected along the scalp that don't originate from the brain. Biological artifacts are the most common, and are caused by eye blinks and movement, muscle activation, and skull defects. Artifacts will show up on EEGs similarly to neural oscillations and their amplitude is often dependent upon the signals of interest. Because EEGs are so commonly plagued by artifacts, it takes a very experienced professional to interpret EEGs clinically, and because clinical professionals can be expensive and inaccurate to interpret EEG data, to utilize machine learning algorithms to interpret the data instead. [2]

Electroencephalograms have multiple reliable modalities being used to measure brain signals, namely Steady-State Evoked Potentials (SSEP), P300, and Motor Imagery (MI), all of which collect specific important oscillations from the brain. SSEPs appear when the user encounters a periodic stimulus like a blinking light or modulated sound. The stimulation of fixed frequency evokes SSEP by yielding EEG activity of the identical frequency as the stimulation is generated. It means that if you were to respond to a stimuli that flickered at a frequency of 10hz, the frequency of the SSEP would also be 10hz, allowing the BCI to determine your target, through SSEP features.

P300 based Brain-Computer Interfaces refers to a specific brain signal called "the P300 wave" that appears approximately 300ms after a target stimulus has been recognized. To trigger this wave, you must use a flickering stimuli that can include colors, sounds or symbols, and the goal for the user is to identify the "oddball", with the wave produced approximately 300ms later is what's important.

BCIs that utilize Motor Imagery, focus on the oscillations in your brain that prepares for or imagines movement called sensorimotor rhythms. The most influential frequency bands for motor imagery are alpha and beta waves. Based on the cortical area, an EEG utilizing Motor Imagery can pick up whether you are thinking about moving your feet, hands, tongue etc. This technology, using fast and accurate readings of motor imagery EEG data, allows for the possibility of movement of these limbs by thought, controlled by a BCI. [12]

2.3 Artificial Intelligence

Artificial Intelligence (AI) is a branch of computer science that leverages a computer's processing power to simulate human intelligence to solve real world problems that usually require human expertise. Companies that use artificial intelligence today include Google for its web search engine, and media platforms such as YouTube and Netflix for its recommendation systems, as well as applications that need to process and understand human language like Siri and Alexa are several forms of Artificial Intelligence. Artificial Intelligence has many branches including Expert Systems, Robotics, Neural Network, Natural Language Processing, Fuzzy Logic, and most importantly Machine Learning.

Machine Learning as a branch of artificial intelligence, uses algorithms and statistical models to learn and adapt without explicit human instruction. Machine Learning has three different approaches to learning, Supervised, Unsupervised, and Reinforcement. Supervised Learning is when the system is given a data set with example inputs and their outputs, and is then required to find the general rule or function that produces the outputs. Unsupervised Learning on the other hand takes a data set with only inputs then tries to find commonalities and produce an output. Finally, Reinforcement Learning is when a program is placed in a dynamic environment and is then asked to perform a goal and as it progresses is provided feedback, maximizing positive actions. Deep learning uses all of these learning types, but with more depth. [11]

Deep learning is a branch of Machine Learning methods based on Artificial Neural Networks. Artificial Neural networks are structured after biological neural networks found in animal brains. It uses a system of connected nodes similar to neurons that processes a signal, represented by a number. The connections between the nodes are called edges, which use weights to adjust how the network learns, increasing or decreasing the signal at said connection. Neurons are usually separated into layers, with each layer performing a different transformation on its inputs. "Deep" in deep learning refers to the use of multiple of these layers in its network. A class of ANN that Deep Learning tends to use is called a Convolutional Neural Network, which most modern Brain-Computer Interfaces use for classification.

Convolutional Neural Networks provide specialized performance in image, speech, and audio classification. All CNN's have at least three types of layers they are composed of: Convolutional Layer, Pooling Layer, and Fully-Connected Layer as shown in Figure 2.

The purpose of the convolutional layer is feature extraction, with each successive convolutional layer increasing the specificity. This layer takes a kernel, small array of numbers, and applies it across the given input which is another array of numbers called a tensor. A product is then calculated between the kernel and the tensor, known as a convolution, creates an output value in an output tensor called a feature map. This process is then repeated with multiple kernels, creating multiple feature maps, representing a different distinct feature of the corresponding input tensor. The convolutional layer also has two hyperparameters, the size and number of kernels. The size of the kernel is usually a 3x3 but can also sometimes be a 5x5 or 7x7. The number of kernels determines the depth of the produced feature maps. Feature maps are the products of a linear operation, which are then processed through a nonlinear activation function, most commonly a Rectified Linear Unit(ReLU) activation function. [13]

Next, the pooling layer is aiming to downsample the feature maps, reducing the dimensions of the maps in order by summarizing the features within each feature map. The two most common forms of pooling operations are max pooling and average pooling. Max pooling would take a filter, most commonly 2x2, and take the max value when comparing the

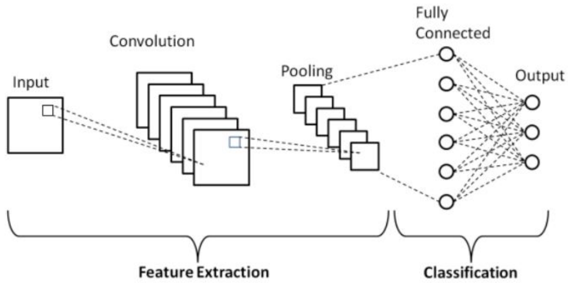

Figure 2: Example of Convolutional Neural Network

filter to the input tensor. Average pooling on the other hand takes the average value of a feature map and downsamples it into a 1x1 array. Because of average poolings downsampling power, it is generally only used once before the fully connected layers.[13]

It is called the fully connected layer because every input is connected to every output through a learnable weight. This layer(s) maps the features extracted and downsampled by the convolutional layer(s) and pooling layer(s) to the final outputs of the network, usually the probability of a class in classification tasks. This probability of classification is calculated by an activation function on the last fully connected layer, usually the softmax function. Because the last fully connected layer will have the same number of output nodes as classes, the softmax function will normalize these outputs, applying a probability to each class with values between 1 and 0 that sum to 1.[13]

Because of the technical difficulty that comes with interpreting electroencephalograms, we can use the strengths of Convolutional Neural Networks to interpret EEG data in real-time, and a lot more consistently. Making the interpretation of the EEG data cheaper and almost instant, can bring many new technologies and usages of EEG-based BCIs to the table like Virtual Reality and improved brain monitoring systems in clinics.

2.4 Brain-Computer Interfaces

A Brain-Computer Interface is the usage of a person's brain waves to control an external device, which was mainly developed for but not limited to the medical field with increasing uses every year, even reaching into the entertainment industry with Virtual Reality. Every Brain-Computer Interface system consists of 5 components: Brain activity measurement, Pre-processing, Feature Extraction, Classification then Translation into a command. Brain activity is measured with EEG electrodes, feeding the device data. This data then goes through pre-processing which removes all artifacts from the EEG data, so that the BCI can interpret clean data. From the clean data, features are extracted meaning making connections between oscillations and external stimuli. Then through a series of machine learning algorithms, the connections are classified as actions. The BCI then translates that action into an actual command and performs it. With this complicated series of calculations to perform a simple action, come its challenges, limiting its true potential. [8]

3. DECODING MENTAL STATE

Table 2: Comparisons of the Prediction Accuracy with other Deep Learning Models.

Methods	Prediction Accuracy
EEGNet ([17])	51.01%
FBCSPShallowNet ([18])	49.17%
DeepConvNet ([18])	52.91%
Our proposed	53.22%

Figure 3: Decoding Results [14]

In 2019 an experiment [14] was conducted by Dongdong Zhang, Dong Cao, and Haibo Chen using Brain-Computer Interfaces and their proposed Deep Convolutional Neural Network using 6 convolution blocks to accurately predict mental states. The catalyst for this research was the knowledge of how common traffic accidents are when the driver is mentally fatigued. Direct changes in EEG activity is usually manifested by fatigue, so they used an EEG and electrooculogram (EOG), which monitors eye movement, as a bio-marker to achieve high accuracy prediction of fatigue.

3.1 Method

This data was gathered from 5 different participants with over 25 hours of EEG recordings, around 6 sessions each. Each participant was put in a virtual environment simulating driving a train for 30 to 50 minutes. For the first 10 minutes the participants were asked to be super focused and in control of the train. The next 10 minutes the participants were asked to be in a relaxed state without controlling the train. The last 10 minutes the participants were allowed to fully relax, closing their eyes or sleep as they wanted.

The EEG data was collected using the Emotiv Epoc+ EEG headset, which has 14 electrodes at a 128hz sampling rate with a bandwidth of 0.2-43hz. A filter was then used to cut off all frequencies under 1hz and over 40hz. The data was then staged for each 10 minute segment and labeled as: Alert, Normal, and Drowsy. They then had to normalize their data due to the 14 separate electrodes in their headset, using total average and deviation to preserve the correlations between the electrodes. $x_{i,j}$ is data point x_i in channel j, y is the mean of all the channel data, s is the standard deviation of the data.

$$x^n_{i,j} = \frac{x_{i,j} - \bar{y}}{s}$$

The clean data was then processed through their own Deep Neural Network using an input layer, 6 convolution blocks and an output layer.

3.2 Results

Figure 3 displays the prediction accuracies of the respective method or deep learning model. Prediction Accuracy refers to the model or methods ability to classify a brain wave from EEG data as: Alert, Normal, or Drowsy.

In Table 2 of Figure 3, they compared their proposed Convolutional Neural Network shown in Figure 4, with other deep learning architecture solutions like DeepConvNet, Filter Bank Common Spacial Patterns ShallowNet (FBCSP) and EEGNet. These other three CNN models are well known modern classification solutions, with EEGNet generally considered to have the best classification accuracy. Despite how well known these other CNN solutions are for their high accuracy classification, the researchers newer proposed CNN model scored a higher classification accuracy using the same training and evaluation for all four models with a 53.22 percent.

Their proposed CNN in Figure 4 is constructed using an input layer, 6 convolutional blocks and an output layer. Each convolutional block performs 2 1-D convolutions with batch normalization(BN) and a ReLU activation function. The dimensionality of a convolution defines how many directions the kernel will move, so in the case of the researchers 1-D convolutional layers, the kernel only moves in one direction. These types of convolutions are mainly used for things like text and 1D signals, like EEG data. Batch Normalization is used to standardize the inputs to the network and often accelerates the training process of these networks. A dropout layer was also added to reduce overfitting. Overfitting occurs when the network reflects the training data too well to the extent that its performance is negatively impacted on new data sets. [10]

Because a models classification accuracy in this context is how accurately classify a persons mental state based on EEG data, the accuracy of their proposed model shown in Figure 4 only being 53.22 percent, is still too low for it to be used clinically or commercially. However at the same time, we can see that this technology can still be refined and yield greater classification accuracy. The researchers for this experiment reported that a lot of the inaccuracy stems from prediction of a normal mental state because it's easy to be misclassified. While an improvement was made through these researchers new algorithm, this experiment also highlights another current problem of EEGs, their accuracy.

Electroencephalograms experience major accuracy issues due to artifacts, and artifacting can originate from any movement, meaning even with the most advanced Deep Learning algorithms, chances are there are still going to be inaccuracy issues, currently limiting this technology's growth. [14]

4. VR FOR NEURO-REHABILITATION

Brain-Computer Interfaces for a while have been used to assist in recovery from cognitive and motor impairments such as Apraxia common in stroke patients. Motor functions in the brain are built and strengthened through repetitive usage, but conditions like a stroke may destroy the connectivity between your neurons causing movement to be difficult or even impossible. Two common approaches to recovery using Brain-Computer Interfaces is through continuous usage of robotic prosthetic to simulate movement to rebuild those neural pathways. The other approach is through the facilitation of neuroplasticity and motor learning to increase motor function recovery. However a recent study from 2019 has developed the use of an electroencephalogram based Motor Imagery-Brain-Computer Interface (MI-BCI) with a Virtual Reality game for stroke rehabilitation.

Motor Imagery practice has been shown to activate and improve motor pathways in patients regardless of the the state of their motor functions, more so if combined with conventional physical therapy. The main pillar behind motor rehabilitation is progressive and skilled motor practice; however the problem with standard physical therapy procedures, is the fact that it gets too repetitive and may influence the motivation and engagement of patients leading to less effective recovery, thus Virtual Reality Brain-Computer Interfaces have begun to be researched. Another benefit to MI-BCIs is active measurement of brain activity, which can be monitored to tune the patients experience to make sure they are getting the most out of the therapy.

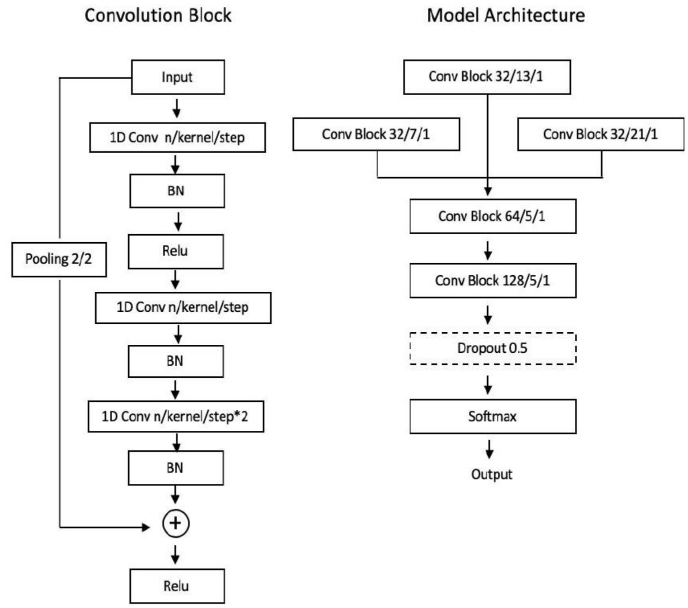

Figure 4: Proposed Deep Neural Network [14]

The study conducted by professors from the Technical University of Denmark deployed their PhysioNet EEG Motor Imagery Brain-Computer Interfaces. Their data sets totaled over 1500 recordings between one to two minutes of EEG recordings from over 100 participants. The headsets were recorded with 160hz sampling frequency and 64 electrodes. The EEG recordings would display and classify motor execution and Motor Imagery data of opening and closing of both fists, their rest state, and both feet. The recording sessions would proceed with patients sitting relaxed and following the visual cues shown in the VR headset, prompting usage of their hands or feet. The cues were used to create three separate data sets to test the performance of their classifier for complexities. Data set one consisted of the MI from the opening and closing of the left and right fist. Data set two contains data set one as well as their resting state. Data set three was the MI from both hands and feet.

With the current development of EEG classification, there is no well established approach or Neural Network with standard hyperparameters, so they developed their own Deep Neural Network shown in Figure 5. Their Neural Network was made of four 1-D convolutional layers, two max pooling layers and six Fully-Connected Layers. Every FC Layer and Convolutional Layer except the last softmax layer, employs ReLU activation, batch normalization and a dropout layer.

The game played by participants in the Virtual Environment had the goal of catching fruits and kicking footballs, and if successful resulted in a small explosion of the object and their score was recorded on the leaderboard. The participants are able to see their own hands and feet, but the arms and legs were left void to allow the brain to make the

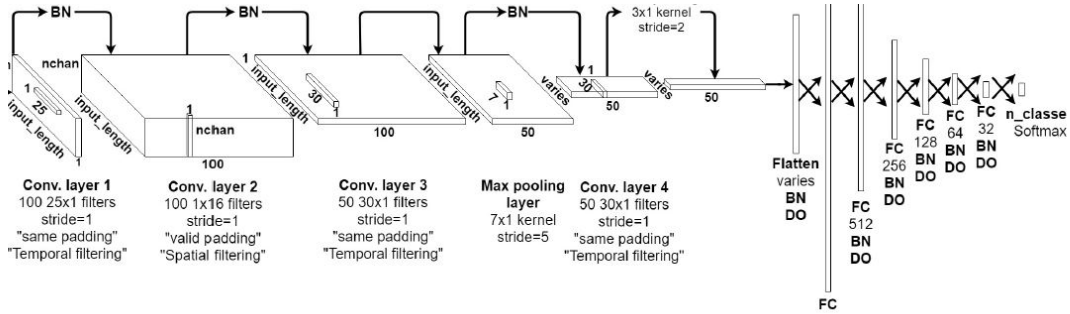

Figure 5: CNN Architecture Used [9]

connections.

In Data set one denoted by global 2cl in Figure 6, all participants achieved 100 percent goal accuracy, but this was partly because of the low activation thresholds producing false positives, meaning some of their misses would be interpreted. Although having false positives, the classifier was still able to accurately distinguish unique hand usage with 80.45 percent accuracy.

Next in Data set two denoted by global 3cl in Figure 6, the highest goal accuracy scored was at 87 percent, and was said to have been enjoyed the most by the participants because they felt like they had more control over the system. Classification accuracy of hand usage increased from data set one to 84.08 percent.

Finally in Data set three denoted by global 4cl in Figure 6, the participants achieved 70 percent goal accuracy with their hands, but 100 percent goal accuracy with their feet. The classifier was able to accurately distinguish feet movement at 72.28 percent, but was reported to have been more challenging to control.

Once finished with the experiment, the subjects were asked to take a questionnaire to evaluate their experience with the Virtual Reality game and headset. All participants found the game to be fun, interesting and very immersive but was also reported to be a little strenuous because of the games demand for their constant concentration. Overall the experiment can be considered a success from the standpoint that it was fun and immersive, yet provided enough of a challenge for the subjects to feel stressed. However the actual systems classification accuracy must be judged as well.

The provided table in Figure 6 shows an overview of the conducted experiment performing left and right classification, as well as the number of channels used, the type of classifier, its accuracy, and the deep learning method of choice. Two things this table can tell us, is that generally a higher amount of channels used, leads to a higher classification accuracy, as well as Convolutional Neural Networks generally yielding a high classification accuracy. Using their 64 channel system, the researchers were able to achieve 88.50 percent accuracy when classifying between left and right hand usage, but only 76.37 percent when including the feet.

This experiment was extremely impactful for many reasons. They developed a Convolutional Neural Network, displayed in Figure 5, with higher classification accuracy than other researchers, and was also used in a real-time experiment. This as well as the game being enjoyable, provides a realistic potential solution to patients that feel unmotivated or fatigued with their standard physical therapy, so it is

This work			6 [sec]	0.5 [sec]	
	64	global 2cl	85.94%	80.45%	CNN
		global 3cl	88.50%	84.08%	
		global 4cl	76.37%	72.28%	
	16	global 2cl	79.66%	72.81%	CNN
		global 3cl	84.13%	78.62%	
		global 4cl	65.96%	60.37%	

Figure 6: VR Rehab Results [9]

easy to recommend Motor-Imagery Brain-Computer Interface Virtual Reality systems with Deep Learning classifiers for real time applications like video games to assist in stroke rehabilitation. [9]

5. CONCLUSIONS

As Deep Learning and Neural Network Technology progresses, so does Brain-Computer Interfaces and corresponding technologies. We are still in early stages of development, yet the provided studies have shown how effective and powerful this technology can be in the future, clinically and in the entertainment industry.

While other issues do exist such as the major artifacting in EEG data as well as the variety of the patients brain signals affecting Neural Network classification, Deep Learning overall is the correct path to advancing this technology treating more brain disorders like Alzheimer's and provide never before seen immersiveness with Virtual Reality.

Acknowledgments

I would like to thank Elena Machkasova as well as my alumni reviewer Justin Mullin for their valuable feedback .

I would also like to thank Kristin Lamberty for guidance throughout the research process, giving needed feedback.

6. REFERENCES

[1] Deep learning. `https://en.wikipedia.org/wiki/Deep_learning`, Nov 2021.

[2] Electroencephalography. `https://en.wikipedia.org/wiki/Electroencephalography`, Nov 2021.

[3] Human brain. `https://en.wikipedia.org/wiki/Human_brain`, Nov 2021.

[4] J. Brownlee. 14 different types of learning in machine learning. `https://machinelearningmastery.com/`

types-of-learning-in-machine-learning/, Nov 2019.

[5] B. I. C. Education. What are convolutional neural networks? https://www.ibm.com/cloud/learn/convolutional-neural-networks.

[6] R. D. Fields. The brain learns in unexpected ways. https://www.scientificamerican.com/article/the-brain-learns-in-unexpected-ways/, Mar 2020.

[7] D. J. Ford. How the brain learns. https://trainingindustry.com/articles/content-development/how-the-brain-learns/, Jun 2021.

[8] A. Gonfalonieri. Deep learning algorithms and brain-computer interfaces. https://tinyurl.com/deeplearnbci, Nov 2019.

[9] T. Karácsony, J. P. Hansen, H. K. Iversen, and S. Puthusserypady. Brain computer interface for neuro-rehabilitation with deep learning classification and virtual reality feedback. In *Proceedings of the 10th Augmented Human International Conference 2019*, AH2019, New York, NY, USA, 2019. Association for Computing Machinery.

[10] C. Maklin. Dropout neural network layer in keras explained. https://tinyurl.com/dropoutlay, Jun 2019.

[11] M. Middleton. Deep learning vs. machine learning - what's the difference? https://flatironschool.com/blog/deep-learning-vs-machine-learning, Feb 2021.

[12] M. Rashid, N. Sulaiman, A. P. P. Abdul Majeed, R. M. Musa, A. F. Ab. Nasir, B. S. Bari, and S. Khatun. Current status, challenges, and possible solutions of eeg-based brain-computer interface: A comprehensive review. https://www.frontiersin.org/articles/10.3389/fnbot.2020.00025/full, Jan 1AD.

[13] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi. Convolutional neural networks: an overview and application in radiology. https://tinyurl.com/cnnvalid.

[14] D. Zhang, D. Cao, and H. Chen. Deep learning decoding of mental state in non-invasive brain computer interface. In *Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing*, AIIPCC '19, New York, NY, USA, 2019. Association for Computing Machinery.