7 Deep Learning Techniques for Thought-Controlled IoT Devices

Adithya Sanyal, Aditya Thaker, Akshath Mahajan, Deap Daru, and Ruhina Karani

7.1 INTRODUCTION

7.1.1 Introduction to Brain Waves and Electroencephalogram

In the most profound way, neurons are electrical components. The nuclear envelope of a neuron contains numerous passageways allowing ions to enter and exit the membrane. The resting membrane potential of the cell is –70 mV, according to neuroscientists. The membrane's potential varies on a frequent basis, due to stimuli from other axons. Some impulses augment the neuron's membrane permeability, while others diminish it. Excitatory or inhibitory impulses are different ways to define them respectively since they simulate or hinder the production of action potentials. When an action potential occurs at the synapse between two neurons, it causes the neuron to release a chemical neurotransmitter. The neurotransmitter has the potential to either stimulate or suppress the following neuron's action potential firing. Electrical impulses empower the brain to engage with one another by transmitting electric pulses. Neurons exploit electrical impulse to acquire and communicate effectively. Analysing signals or pictures from the brain can be used to better understand cognitive behaviour. Eye movement, lip movement, recollection, attentiveness, hand gripping and other motor and sensory states can be used to picture human activity. These states are linked to a certain signal frequency, which aids in the comprehension of the functional activity of a complicated brain structure. An electroencephalogram (EEG) is one technique for identifying electrical impulses in the brain.

EEG is a technique that identifies brainwave activity while using tiny metal plaques (anode and cathode) affixed to the head. The electrodes monitor miniscule electrical impulses generated via neural synapse's activity. The impulses are boosted and displayed on a computer display as a plot or as a recording which can be reproduced in print. An EEG is used to uncover aberrations in activity in the brain that can assist in diagnosing a variety of brain maladies like schizophrenia and sleep disturbances. The EEG is represented as a sequence of trendlines. Regardless of whether one is awake or asleep during the test, the lines will vary in appearance, but each

DOI: 10.1201/9781003436911-7

phase has a standard pattern of brain activity. It could be a symptom of epileptic seizures if the basic pattern of brain waves has been altered (Johns Hopkins Medicine 2021; Mayo Foundation for Medical Education and Research 2022; WebMD n.d.).

The core frequencies affiliated with the human body are Delta, Theta, Alpha, Beta and Gamma, which are outlined below.

The Delta brainwaves have a frequency range of 0.1–3 Hz. These types of waves are emitted in conditions of deep, dreamless, non-rapid eye movement sleep or in the unconscious state. The Theta brainwaves range from 4 to 7 Hz and occur in situations where a person is recalling something or dreaming. The Beta waves are subdivided into low, midrange and high having frequency ranges of 13–15, 16–20 and 21–30 Hz. The low beta waves are emitted in relaxed and focused conditions. The midrange beta is observed when one is thinking or is aware of their respective surroundings whereas high beta is emitted in cases of alertness or agitation. Lastly the Gamma waves with frequency of

31 Hz and above are seen in situations of information processing and cognition (Hinton et al. 2006).

7.1.2 Machine Learning and Deep Learning in Electroencephalogram

Machine learning (ML) can be defined as computer algorithms that can improve over time through experience and by using data. ML systems use large amounts of data, known as training data, to make predictions without being explicitly programmed to do so (Wikimedia Foundation 2022). This finds applications in various computing fields including – feature extraction, pattern recognition, data interpretation, classification and prediction. These systems are also capable of automating EEG analysis, and their results can be used to make decisions. They can be categorised into feature-based (with handcrafted features), and end-to-end approaches (with learned features) (Gemein et al. 2020).

This diverse range of applications lead us to explore the use of ML techniques to infer simple commands by analysing brain wave data, hence, creating systems that are based on the analysis of EEG brain waves.

Before deep learning gained popularity, the standard EEG techniques combined different signal processing and ML methods for enhancing the signal-to-noise ratio, handling of EEG artefacts, feature extraction and interpretation of signals (Bitbrain 2021).

As will be seen further in this chapter, the decoded signals can be fed into an Internet of Things (IoT) device, which can in turn take different actions based on the inputs.

7.1.3 Thought-Controlled Internet of Things Devices

The IoT refers to physical devices that are capable of computing and are equipped with sensors, software and technologies that can be used to exchange data with external devices over the Internet or other communication networks.

A brain–computer interface (BCI) is a computer program that takes brain signals as input and translates them into commands that are passed on to various devices

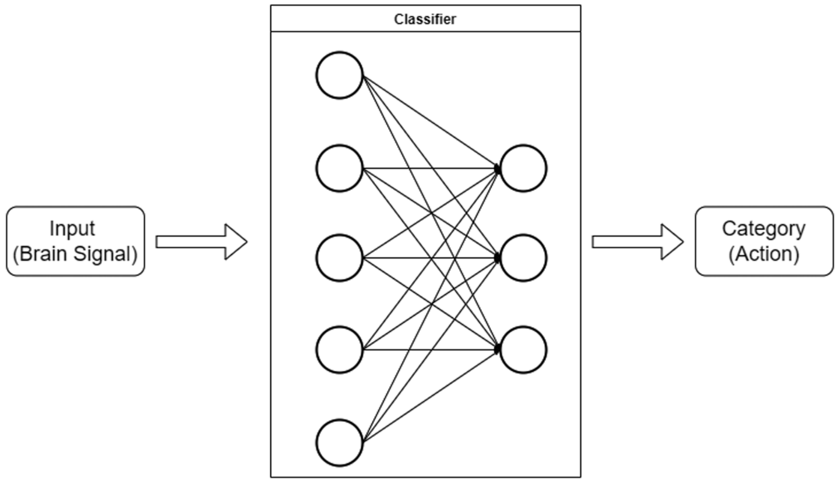

FIGURE 7.1 Simplified representation of a brain-computer interface model.

that carry out desired actions. In general, all types of brain signals are capable of controlling a BCI system. The most popular ones are electrical signals from the brain measured from electrodes that are attached to the scalp, on the cortical surface, or in the cortex (Shih et al. 2012).

When working with brain–computer interfaces, one generally uses ML or deep learning techniques to develop classifiers, which accept the signal data and predict the action to be performed by extracting useful information. Figure 7.1 illustrates a simplified representation of a brain–computer interface model.

7.2 ELECTROENCEPHALOGRAM FEATURE EXTRACTION

EEG devices record a wide range of information about human cognition, behaviour and emotions. The technology has applications in improving healthcare, emotional EEG analysis, as well as brain research. EEG data, on the other hand, is difficult to understand since it differs widely between persons, contains a lot of noise, and changes significantly over time, even for the same person. Visual inspection is a time-consuming, costly and inconvenient technique that does not scale well, and it cannot be used in BCI applications (Bitbrain 2021).

Therefore, in this section, we propose multiple state-of-the-art deep learning techniques to extract information from EEG signals, ranging from Filter Bank Common Spatial Pattern (FBCSP) to hybrid architectures made up of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM).

7.2.1 Filter Bank Common Spatial Pattern

The Common Spatial Pattern (CSP) approach is instrumental in developing appropriate filters occupying space in motor imagery (MI)-based BCI that differentiate among the two categories of EEG data. The variance of two signal matrices of different classes can be enhanced with this technique. It is based on the diagonalisation of both classes' covariance matrices at the same time. However, the efficacy of this spatial filter is determined by the frequency band in which it operates.

The CSP strategy is excellent for computing spatial filters for detecting Event-related Desynchronisation (ERD) and Event-related Synchronisation (ERS).

The CSP strategy is excellent for computing spatial filters for the recognition of ERD and ERS. The goal of spatial filtering in BCI using the CSP method is to construct features with the best variances for distinguishing two classes of EEG readings. The CSP algorithm uses a method that involves diagonalising two covariance matrices at the same time.

$$\mathrm{R} = \mathrm{PQ} \tag{1.1}$$

In this equation, Q represents an M×N matrix containing data of a single trial of the raw EEG measurement; M is defined by the number of channels; N is the number of measurement samples per channel. The CSP projection matrix is P. The horizontal entries of P are the filters occupying space that do not move and the vertical entries of P-1 are the CSPs.

The spatial filtered signal R optimises the variance differences between the two groups of EEG observations. Nonetheless, only a tiny minority m of the spatial filtered signal's anomalies is typically utilised as classification features. The *n* initial and hindmost rows of Z, that is, Zp, pϵ {1 to 2n} result in Xp that is the feature vector provided as classifier's input.

$$Xp = \log\left(\frac{var(Zp)}{\sum_{i=1}^{2m} var(Zp)}\right) \tag{1.2}$$

When EEG calculations are raw or refined with an inadequately specified range of frequencies, categorisation based on CSP features generally yields low accuracies. As a result, the CSP approach is typically used when working with huge frequency ranges or manually setting a frequency range for a certain issue. Several solutions have been suggested to meet the obstacles of arbitrarily defining the effective particular aspect frequency range for the CSP algorithm.

On each of these sub-bands, spatial filters relying on the CSP methodology are applied.

The Sub-Band Common Spatial Pattern (SBCSP) algorithm was developed to alleviate the complication of arbitrarily designating the CSP algorithm's operating frequency band for a given subject, and it has been shown to enhance classification accuracy. Using a Gabor filter bank, SBCSP breaks down EEG signals into numerous sub-bands. On each of these sub-bands, spatial filters relying on the CSP methodology are applied. The score of the sub-band is fused using recursive band removal or a classification approach once it is obtained. The fused sub-band score then is evaluated using this procedure.

The FBCSP is another innovative ML approach for interpreting EEG data in motor imagery-based BCI (FBCSP). The FBCSP has the following steps: filtering based on frequency, spatial filtering, selection of features and classifying the features. SBCSP uses spatial filters, whereas FBCSP makes use of the potent filters with the designated CSP characteristics coupling. As a result, FBCSP only uses a small set of successful spatial filters. In comparison to employing the whole set of spatial filters, this minimises the computing complexity.

The EEG data are bandpass segmented into distinct frequency bands over the first round. A zero-phase Chebyshev Type II Infinite Impulse Response filter bank is used by FBCSP. This method eliminates the non-linear phase shift imposed by the IIR filter via using a zero-phase filtering. The characteristics are extracted from these bands in the second phase. The attribute selection methodology is adopted in the third step to uncover discriminative frequency band couplings and their corresponding CSP properties.

In pattern classification, feature selection is stated as extracting a group of size k characteristics from a set of d features that yields in the minimal classification errors. The feature selection algorithms utilised here are Mutual Information-based Best Individual Feature (MIBIF), Mutual Information-based Naïve Bayesian Parzen Window, Mutual Information-based feature selection, Mutual Information-based Rough Set Reduction (MIRSR), Fuzzy-Rough set-based Feature Selection and MIRSR.

Following the use of feature selection methods, CSP characteristics are classified using a classification method in the fourth step. The NBPW classifier, Fisher Linear Discriminant (FLD), Support Vector Machine (SVM), Classification And Regression Tree, k-nearest neighbour, Rough set-based Neuro-Fuzzy System and Dynamic Evolving Neural-Fuzzy Inference System are among the classifiers used here.

The architecture of the Filter Bank Common Spatial Pattern is displayed in Figure 7.2.

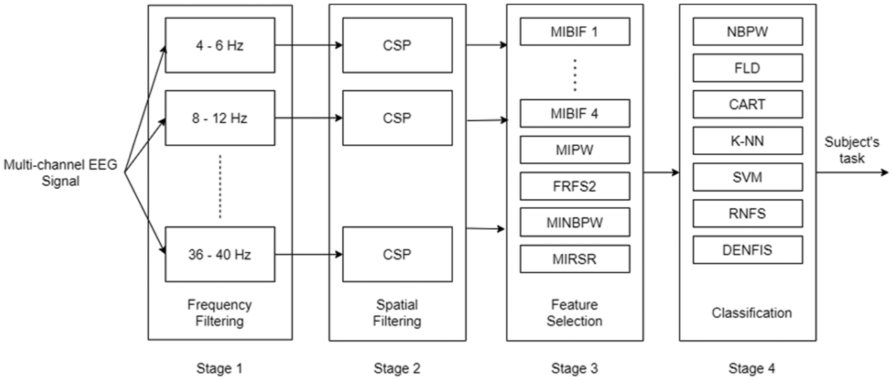

FIGURE 7.2 Architecture of the filter bank common spatial pattern.

The MIBIF selection algorithm, which designates four sets of CSP attributes, as well as the NBPW, FLD or SVM classification algorithms, are strongly proposed for use with FBCSP in MI-based BCIs, experimentally.

In MI-based BCI, the FBCSP is employed to interpret EEG signals. The dilemma of choosing the most appropriate functional frequency range for extracting distinguishing CSP traits has been tackled by FBCSP. FBCSP can learn subject-specific patterns from high-dimensional EEG recordings without the need for operator interaction, and it has reasonably good classification accuracies.

These deep learning techniques mentioned above are used to classify the EEG signals into indicator actions that tell the IoT or robotic devices what action to perform. This method can hence be used to develop EEG-based BCI systems that have the potential to not only help people with disabilities but also provide convenience to everyone (Ang et al. 2008).

7.2.2 Deep and Shallow ConvNets

ConvNets are artificial neural networks that use convolutions as a major component to discover local patterns in data. The network architecture can be as simple as having just one convolutional layer, or as complex as having several thousand layers. They start with extracting low-level features from raw data and work their way up to high-level features in the deeper layers (Schirrmeister et al. 2017).

Deep ConvNets have had a lot of success in recent years, especially in domains like computer imagery and speech recognition, where they are known to frequently outperform earlier state-of-the-art methods (LeCun et al. 2015). Through end-to-end learning, deep learning with CNNs has greatly advanced the field of computer vision.

An EEG signal has properties that distinguish it from the most common CNN inputs, which are primarily images. The EEG signal, unlike static pictures, is essentially time-series data derived from electrode measurements. Moreover, it has a low signal-to-noise ratio, which means that it contains information that is irrelevant to the task, which often affects the EEG signal more strongly than the relevant information (Bast et al. 2006). These characteristics may make learning features for EEG signals more difficult than for standard images. As a result, existing image-based ConvNets architectures must be altered for EEG inputs, and the resulting decoding accuracies must be thoroughly compared to traditional EEG feature extraction techniques (Schirrmeister et al. 2017).

One can describe the dataset as pairs of EEG recording trials and class labels under the assumption that each session of recording (time-segment) focuses on one class. This enables one to treat the task as a supervised learning problem. Such datasets can be created for multiple subjects. The input Xj corresponds to trial j and contains the pre-processed signals of the electrodes. The class label of Xj is denoted by yj and its values correspond to a specific thought-based command.

There are two CNN architectures for extracting information from EEGs proposed in (Schirrmeister et al. 2017), namely – Deep ConvNets and Shallow ConvNets.

The Deep ConvNet architecture is made up of a unique first block built to handle EEG inputs, followed by three generic convolution-max-pooling blocks, and ends with a dense layer that uses the softmax function for classification as shown in Figure 7.3. The first convolutional block is divided into 2 layers to address the huge

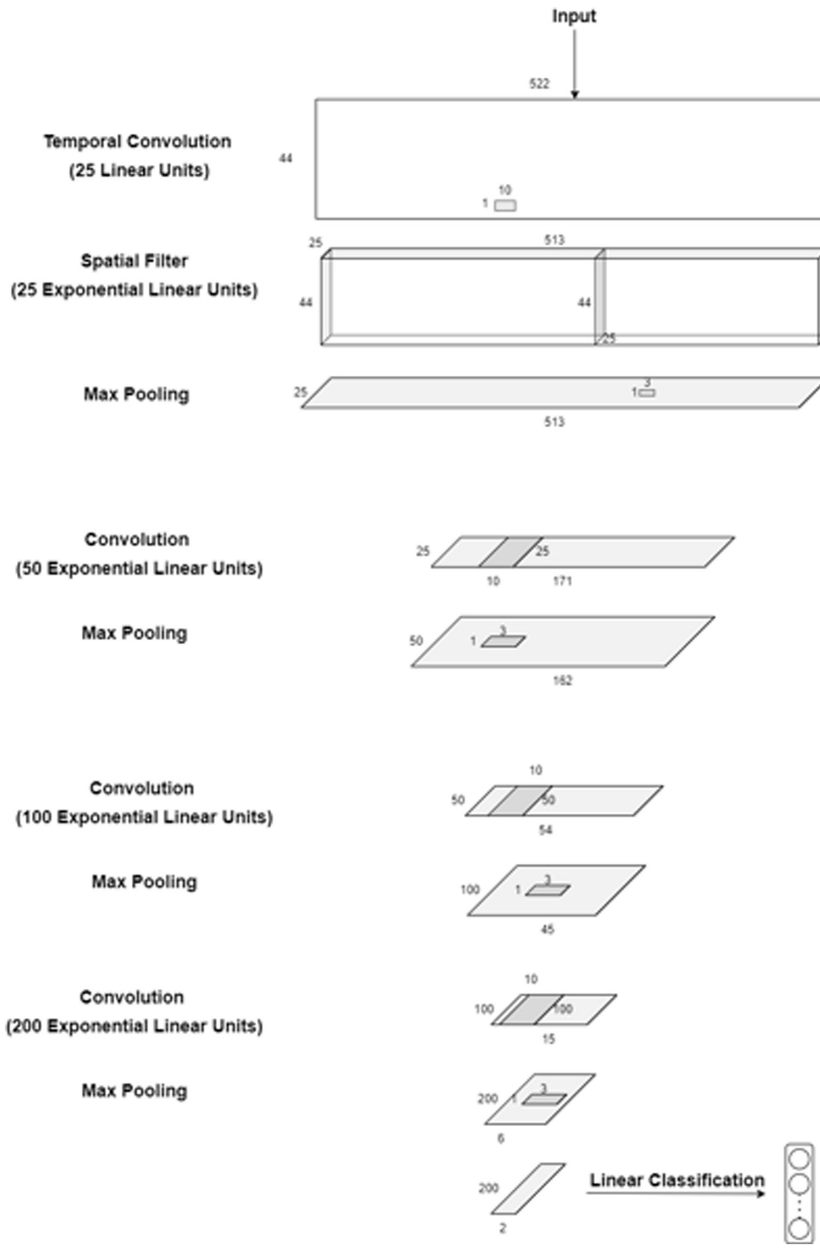

FIGURE 7.3 The deep ConvNet architecture.

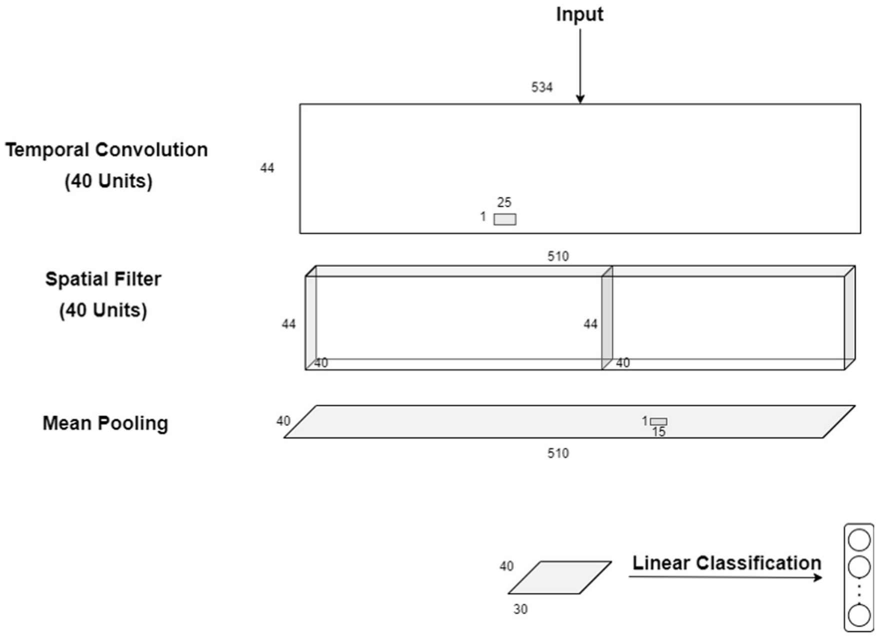

FIGURE 7.4 Architecture of shallow ConvNet.

amount of input channels. Each filter performs a temporal convolution in the first layer, followed by spatial filtering with weights for all pairs of electrodes using filters from the earlier temporal convolution in the second layer.

Similarly, a shallow architecture designed based on the FBCSP pipeline is also found to perform well at decoding raw EEG signals.

The Shallow ConvNet architecture is designed such that its first two layers are similar to the Deep ConvNets, and conduct a temporal convolution followed by a spatial filter. These steps are similar to FBCSP's bandpass and CSP spatial filter phases. These layers are followed by a squaring non-linearity, a mean pooling layer and a logarithmic activation to replicate FBCSP's trial log-variance computation as shown in Figure 7.4. The Shallow ConvNet encapsulates all computations in a single network, unlike FBCSP, and allows the simultaneous optimisation of all steps.

To use ConvNets for classification, the outputs are generally converted to conditional probabilities P(lk | Xj) using the softmax function, where lk represents a class label. This is a standard linear classification task.

The Deep and Shallow ConvNets show a better decoding accuracy as compared to the traditional FBCSP approach. In particular, the Deep ConvNet shows a 0.5 to 2.9% increase on different EEG datasets. The Shallow ConvNet is a bit more extreme, showing a 1.9% decrease to a 5.7% increase in the decoding accuracy on different datasets (Schirrmeister et al. 2017).

7.2.3 EEGNet

Another approach based on the use of ConvNets is proposed in (Lawhern et al. 2018), which describes a CNN-based architecture that is designed to be compact and can classify EEG signals from various BCI paradigms. The model, which is referred to as EEGNet, is designed to be as compact as possible and introduces the use of depthwise and separable convolutions to encapsulate well-known EEG feature extraction concepts.

In essence, EEGNet architecture is proposed to generalise across multiple BCI paradigms without the need for large volumes of data and for extracting neurophysiologically interpretable features. The model is made up of two convolutional blocks followed by a dense classifier that uses softmax activation.

The first convolutional block consists of two sub-blocks. The first sub-block is a standard convolutional layer with F1 filters followed by batch normalisation, and the second sub-block is made up of a depthwise convolutional layer having D*F1 filters (where D is the depth of the layer), followed by batch normalisation, average pooling and a dropout of either 25% or 50% for regularisation. This is done not only to learn a spatial filter but also to reduce the number of trainable weights in the model. When utilised in EEG-specific applications, this process enables the efficient extraction of frequency-specific spatial filters by providing a straightforward means to learn spatial filters for each temporal filter. This two-step convolution technique is in part inspired by a traditional EEG extraction technique – FBCSP.

The second block implements a separable convolutional layer with F2 filters followed by batch normalisation, average pooling and dropout (25–50%) for regularisation. The separable convolution is primarily responsible for clearly separating relationships inside and across the feature maps by learning a kernel to first summarise each feature map separately and then combining the results ideally. It also reduces the number of trainable parameters, and thus increases the compactness of the model. This procedure, when utilised for EEG-based tasks, separates learning the summarisation process for individual feature maps in time (known as depthwise convolution) from learning the merging process for the feature maps optimally (the pointwise convolution). This method is especially beneficial for EEG signals because distinct feature maps are capable of representing data at different time scales.

Finally, the classification block flattens the resulting output of the third block and directly uses a Softmax layer having N units (where N is the number of class labels). The use of a dense layer is avoided to increase the compactness of the model. The architecture of EEGNet is depicted in Figure 7.5.

All convolution layers in the model use the linear activation function, and the batch normalisation layers of the depthwise and separable convolution block use the ELU activation. The filter sizes in EEGNet are chosen based on the sampling rate (frequency) of the data. For example, in the original paper on EEGNet (Lawhern et al. 2018), the authors propose the filters of the first convolutional block of size (1, 64) so that the filter length is half the data sampling rate.

EEGNet is found to have considerably less number of parameters as compared to Shallow ConvNets and Deep ConvNets, while not only maintaining the performance

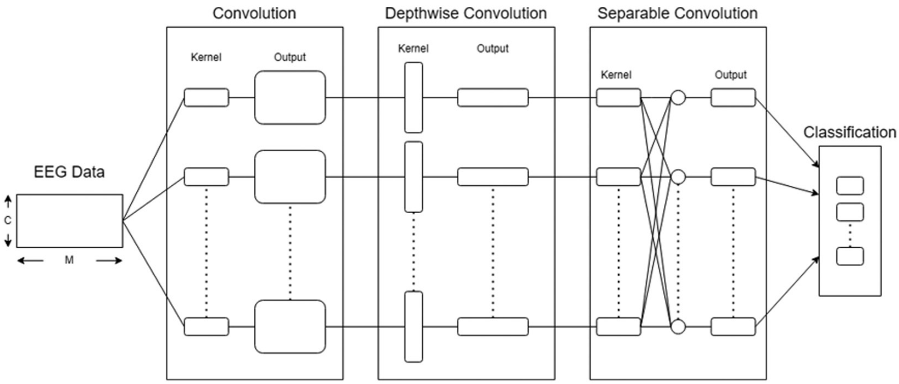

FIGURE 7.5 Architecture of the EEGNet.

standard similar to Deep ConvNet across cross-subject analyses, but also outperforming it in nearly all within-subject analyses.

ERP (event-related potential) is a recorded brain response that is the direct outcome of a specific sensory, cognitive, or motor event (Luck 2014). Shallow ConvNet is observed to perform poorer on event-related prediction BCI datasets in both within-subject and cross-subject assessments.

7.2.4 Hybrid Deep Neural Networks Using Transfer Learning

A hybrid Deep Neural Network employs CNN and LSTM that is capable of extracting spatial and temporal features of the Motor Imagery (MI) signal of the EEG. One major hindrance in BCI is the difference arising between individuals in MI patterns. This leads to a needless time consumption where the classification model undergoes training from scratch on data from a new candidate. Transfer Learning (TL) effectively addresses this issue by fine-tuning the Fully Connected (FC) layer that catches upon the new candidate's features with minimal training data, consuming less time. The proposed technique of Hybrid Deep Neural Networks (HDNN)-TL (Zhang et al. 2021) deals with the individual differences in the 4-class MI signals.

The deep learning approach for classification of enormous and complex data are unprecedented, as they are successfully able to extract the non-linear features better than any other conventional classification method. Various methods for EEG classification are applied to MI signals ranging from traditional classifiers, like CSP and GNB (Sreeja et al. 2017), LDA (Vidaurre et al. 2010), to deep neural network architectures including RBM with fast Fourier transformation (Lu et al. 2016) for 2-class MI classification, new CNN architecture to extract temporal representation in MI signals (Sakhavi et al. 2018) and a deep CNN and LSTM model (Zhang et al. 2019). However, these do not emphasise solving the individual differences of candidates. TL is a method that focuses on imbibing knowledge gained from training on one task and is flexible enough to apply to a different but relevant task. Influenced by the above factors, this strategy has evolved into an HDNN-TL model that when applied to the MI signals of EEG gives state-of-the-art accuracy. The dataset used is the 2008 BCI Competition IV dataset 2a, consisting of a 4-class MI task for tongue, feet, right-hand, and left-hand recorded on 9 healthy candidates (Tangermann et al. 2012).

The in-depth architecture of HDNN-TL is shown in Figure 7.6. Feature extraction is done by OVR-FBCSP and HDNN with the feature-subject correlation handled by FC with TL technique.

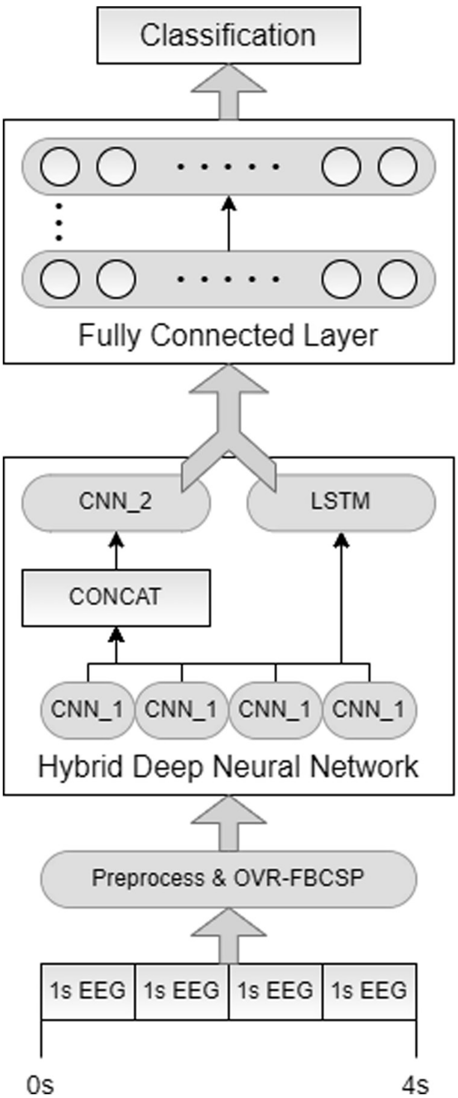

FIGURE 7.6 Architecture of the hybrid deep neural network with transfer learning.

OVR-FBCSP deals with multi-class MI tasks by combining four one-versus-rest CSP filters for the 4-class data to compute output for each filter bank. Spatially transformed signal Z is obtained through (Zhang et al. 2021)

$$Z = W^{T}X \tag{1.3}$$

where X is the originating signal, with W as the projection matrix.

The proposed architecture utilises two sub CNNs as mentioned in Figure 7.6. First has a monolayer involving one convolution with 5 × 5 size kernel and an FC layer. Second has three hidden layers with each convolution layer having a 3 × 3 size convolution kernel. ReLU activation is exercised to extract features of the MI signal. A max-pooling layer follows every convolution layer that diminishes the feature matrix size. Moreover, zero-padding technique is applied to ensure consistent output size matching input size. LSTM, an extension of RNN, is vital in revealing the temporal correlations of time series EEG signals. Therefore, a parallel LSTM with a cell state of 32 is employed with CNN_2 to process the input signal of each time step fed by the CNN_1 layer sequentially.

FC follows the architecture after the CNN and LSTM components extract spatial and temporal features of the MI signal. FC has a three-layer feedforward neural network with the number of nodes in the respective layers as 512, 512 and 32 in each hidden layer, with a dropout network in the second layer to avoid overfitting. ReLU as the activation function for every hidden layer, the following Softmax function is selected,

$$y_{p,m} = \frac{e^{y}m}{\sum_{m}^{T} e^{y}m} \tag{1.4}$$

where m is the index of each class and T denotes the total number of classes (Zhang et al. 2021). We can utilise this through TL by fine tuning the parameters for the FC to train a classification model on new candidates. The learning rate of FC is 0.001. For this consolidated architecture, cross-entropy function is used as the loss function with adaptive method estimation used as the optimiser for the neural network training.

HDNN and HDNN-TL give a huge improvement over the traditional method giving a Cohen kappa value of 0.78 and 0.81 respectively (Zhang et al. 2021). HDNN-TL performs better due to fewer training data for a new subject which takes less computational time as well, whereas traditional methods would involve larger training data with a particular number of subject samples that need to be classified. Therefore, this proposed method avoids the loss of spatial and temporal features and addresses the individual differences in the EEG signals of diverse candidates. This technique can yield fruitful results when applied to imagery-based BCI or other types of EEG-based BCIs.

7.2.5 Siamese Neural Networks

Learning marvellous feature representations for ML applications may have high computational cost, and it can be especially onerous in situations when data are scarce. Moreover, when the number of classes representing the cerebral tasks for classification is increased, a fundamental hurdle in the development of pragmatic and user-friendly BCI systems is their restricted performance. The one-shot learning scenario is a definitive example of this, where one must make accurate predictions based on a couple samples of every class. Here, we will look at a learning technique for Siamese neural networks that use a unique structure to prioritise similarity (Koch et al. 2015).

Siamese neural networks are two or more neural networks that have a homogeneous structure and are trained simultaneously, to distinguish between classes to which the input may belong. 'Homogenous' here means that they have the same weight matrices and neural architecture; the changes in the weight matrix are also duplicated across the networks. It only uses lesser input data points, as its data-hungry contemporaries, to attain discriminatory power that is equivalent to them. This method could be very useful in cases where there is a need to manufacture a thought-controlled device for a specific individual and only a handful of samples are available for training.

A Siamese neural network can be a pair of any two or more neural networks, however, to classify EEG signals we require that the twin neural networks be CNNs. Siamese networks learn the classification tasks by calculating the distance between inputs, rather than learning the probability distribution of the classes over the training data.

The input data must be partitioned into positive and negative samples where each sample is a triplet that includes a randomly selected image, another image of the same class for a positive sample, or an image from any other class, along with a Boolean denoting the correlation between the two images. These pairs are then passed through the networks in parallel, after which a feature vector is obtained, and let us assume its denoted by Hk, this is then utilised for finding differences between the two images. The network through which these input signals are passed is a convolutional neural network, as shown in the Figure 7.7, it provides us with a feature vector for a given input signal. There are a myriad of options available for the choice of architectures that can be used in this feature extraction step, they range from simple models built with intuition from the dataset, or extensively used pre-trained models like VGGNet, AlexNet, ResNet, etc. However, the best choice for this step must be derived from experimentation of several architectures on different splits of data (Shahtalebi et al. 2020).

For the MI classification problem, we denote the first EEG trial with { X C × Sk }T k=1, where C, S, T are the number of channels, samples and trials respectively. The preprocessing step is the application of a function f on each { X C × Sk } resulting in Hk = f(Xk) as the output of the preprocessing step. Let, Hk1 and Hk2 are the two outputs of the twin models, which are then passed to the distance function. Euclidean distance is one of the most used measurements, but Manhattan distance

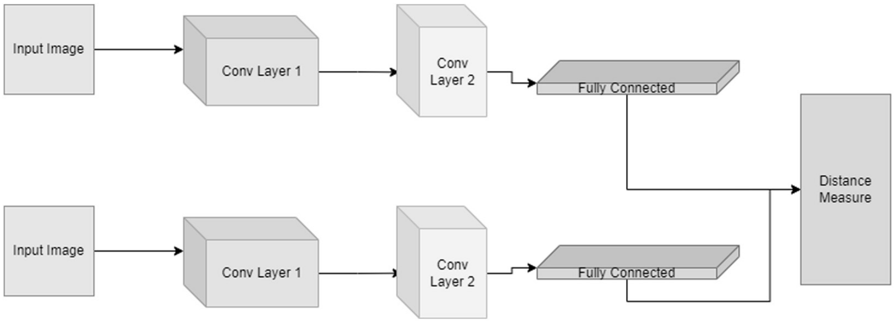

FIGURE 7.7 Architecture of the siamese neural network.

is also a good alternative. Let D be the Euclidean distance metric, then the distance between two signals is given by,

$$D\left(H_{k1}, H_{k2}\right) = \left\|H_{k2} - H_{k1}\right\|_2 \tag{1.5}$$

After calculating the difference between the two inputs, contrastive loss, which is a prominent loss function that is widely employed, is used as a loss function. It is a distance-based loss rather than the more traditional losses. This loss is used to train embeddings to have a low Euclidean distance between two similar points and a large Euclidean distance between two different points. Contrastive loss is calculated as,

$$L\left(W, H_{k2}, H_{k1}, Y\right) = \left(1-Y\right)\frac{1}{2}\left(D\right)^2 + \left(Y\right)\frac{1}{2}\left\{max\left(0, m-D\right)^2\right\} \tag{1.6}$$

Where m > 0 denotes a radius around Hk to decide if a pair of signals are similar or not. The value of m is a hyperparameter that must be tuned to a specific dataset. The model is then trained using Stochastic Gradient Descent, with the weights of the twin models tied together. Siamese Networks offer several advantages as a network architecture. One notable advantage is their resistance to class imbalance, which stems from their capability for one-shot learning. This means that Siamese Networks can effectively recognise and classify classes with only a few instances, making them well-suited for future scenarios where limited data is available for each class. Additionally, Siamese Networks excel at learning from inherent similarity. These networks are designed to focus on finding representations that classify similar classes, allowing them to capture semantic similarity in the data. However, there are some disadvantages associated with Siamese Networks. One of the drawbacks is that they typically require longer training times compared to traditional networks. This extended training duration is primarily a consequence of using quadratic pairings to learn from all available data, which can be slower compared to pointwise learning employed in traditional classification models. Another limitation is that Siamese Networks do not report probabilities. Instead, they rely on paired learning during training, and the output represents the distance from the input to each class. This lack of probabilistic output can be a drawback in situations where probability estimates are essential for decision-making or interpretation.

7.2.6 Restricted Boltzmann Machine and Deep Belief Network

Boltzmann machines are graphs that are a crucial component while studying Statistical Mechanics and are named after the Boltzmann distribution. In this network all the nodes are connected to each other, even the input nodes, unlike conventional neural networks, also have edges between them. They are used to learn the distributions of input data and generate more data by sampling from the learned distribution. Boltzmann machines, on the other hand, have a serious practical problem: they appear to cease learning correctly when the machine is built up beyond a trivial size since the cost of statistical calculation increases exponentially. The Boltzmann machine architecture is shown as in Figure 7.8.

Restricted Boltzmann machines (Salakhutdinov et al. 2007), better known as RBMs, are an alternative to Boltzmann machines with more practical use cases. Their architecture is restricted to a bipartite graph of input layer neurons and the

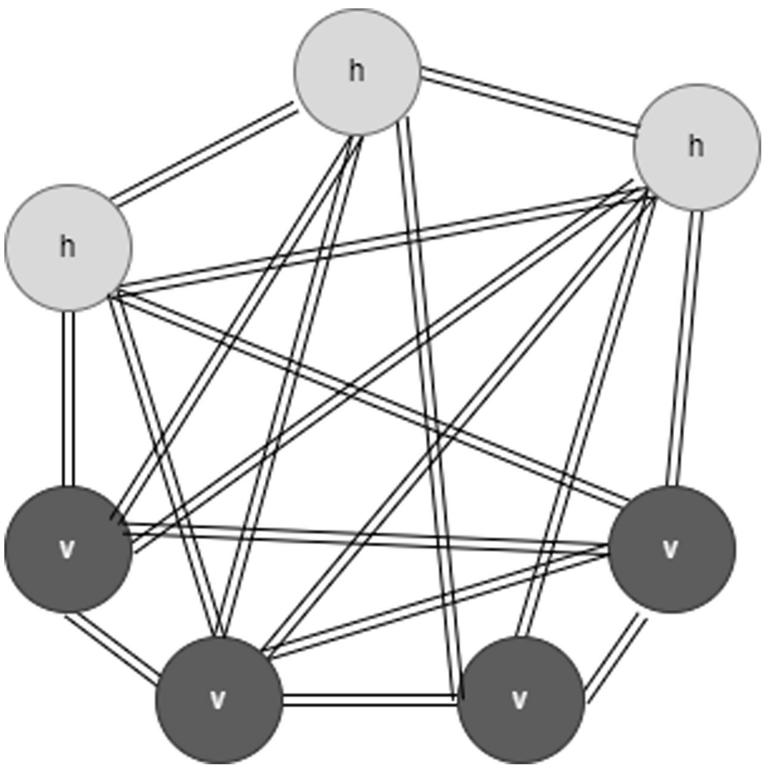

FIGURE 7.8 The Boltzmann machine architecture.

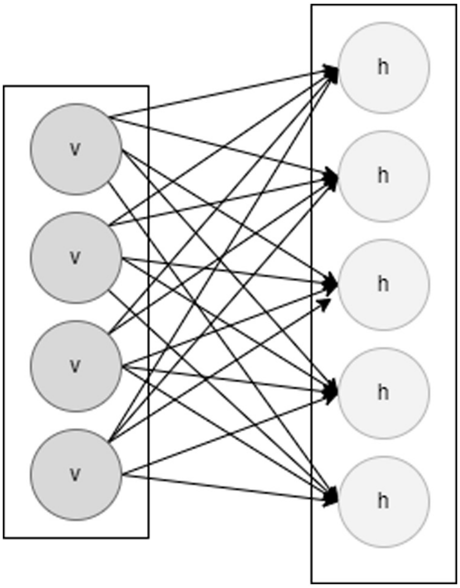

FIGURE 7.9 The restricted Boltzmann machine architecture.

hidden layer neurons. Although, they provide advantages over Boltzmann machines, they too suffer from slower training times. Training algorithms such as contrastive divergence are used to overcome these issues. The Restricted Boltzmann machine architecture is depicted in Figure 7.9.

A deep belief network is a generative artificial neural network that is constructed by stacking several RBMs on top of each other, the number of which is a hyper-parameter that must be tuned while training (Hinton et al. 2006). The training of DBNs takes place by pre-training each of the component RBMs individually and then fine-tuning the entire model. Each RBM is trained with the objective to be able to reconstruct the input provided to the visible layer by tuning the weights of the hidden layer. These RBMs are then stacked on top of each other and a final softmax layer is applied to get a probability distribution over the mental imagery classes.

Before the signals can be passed to the network one needs to ensure that it has a minimum amount of noise to perform band-pass filtering and FFT on the signals. EEG signals have a very low signal-to-noise ratio because they are highly susceptible to noise due to insignificant bodily movements like the blinking of eyes, flinching neck muscles, etc. The band-pass filter allows frequencies in a fixed range to pass and attenuates frequencies that are outside the range (Oon et al. 2018). Research on neurology suggests that the reaction of the brain to MI lies in 5 frequency ranges, namely Alpha, Sigma, low Beta, high Beta and low Gamma, Electrooculography and electromyography introduce noise in the signals useful for motor imagery. Band-pass filtering is applied to attenuate the signals that are outside the 8 and 35 Hz range. The classical architecture of DBNs is shown in Figure 7.10.

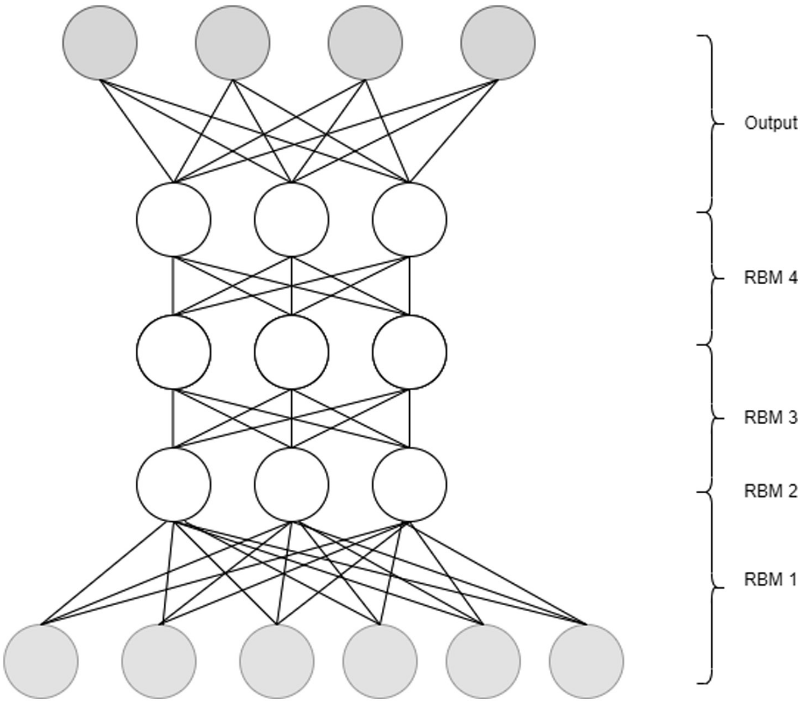

FIGURE 7.10 Architecture of the deep belief network.

7.3 APPLICATIONS

7.3.1 Electroencephalogram-based Brain-Controlled Prosthetic Arm

Myoelectric prosthesis have gained popularity in recent years. However, it has several shortcomings, such as healthy nerves dependency and being quite costly. EEG-based brain-controlled prosthetic arms can be utilised to address the shortcomings of myoelectric prosthetics and to support individuals with neuromuscular ailments. This is a broad spectrum approach that can aid severely disabled people in their daily affairs, notably by allowing them to move their arm autonomously. The four phases of this technology entail acquiring waves emitted by the brain, analysing the signal, systemising them into diverse command signals, and transmitting them to the artificial arm. The power of one's mind to focus and concentrate influences the control of the prosthesis.

It is a BCI device which uses brain waves for command signals to control the prosthetic arm's operations. This system may be separated into four steps in general. Detection of the brain wave signals, collection, propagation, and mapping are all steps in the signal processing process, which are stated as follows:

7.3.1.1 Signal Detection

This step focuses on detecting the signals from the scalp of a person with care. Electroencephalography is a technique for detecting these impulses. The EEG measures voltage fluctuations in the brain's neurons mediated by ionic current. The frequency band of the EEG is commonly used to characterise it. Delta, Theta, Alpha, Beta and Gamma are the primary types of frequencies of human EEG waves.

A Neurosky Mindwave mobile Headset as well as ThinkGear are used for this objective. The Neurosky system is made of dry electrodes and an electrical circuit for the dry electrodes. The system is impervious to noise. A single AAA battery powers the Mindwave headset. It uses Bluetooth to wirelessly transmit the detected signals. It has a ThinkGear chip. The analogue electrical impulses also known as brain waves with help of ThinkGear technology converts brain waves into digital signals using. It is a device that communicates with the wearer's brainwaves.

7.3.1.2 Signal Acquisition

The signals read by the Neurosky Mindwave headset are relayed over Bluetooth to the laptop. The signal is solely detected, interpreted, and converted into digital form by the headset. The laptop is used for signal acquisition. The raw brainwave data from the Neurosky Mindwave Headset is captured using MATLAB® code on a laptop. MATLAB® assists in the extraction of real-time raw brainwave measurements that are used to operate the prosthetic arm. Using the ZigBee transceiver module, the analogue electrical impulses will be sent as command signals to the microcontroller.

7.3.1.3 Signal Transmission

Signal communication between the laptop and the microcontroller is required. With tiny, low-powered digital radios, ZigBee is used to create personal area networks (PANs). The transmission rate of ZigBee is 250 kbit/s. ZigBee is simple to set up and may be used with any type of microcontroller. Wireless connection between the laptop and the microcontroller is accomplished via ZigBee. The EEG signals are sent to the Arduino Uno from MATLAB®.

7.3.1.4 Mapping

In the Arduino Uno microcontroller, the message from the ZigBee transponder must be routed to the artificial hand. The received signal will be used to control the prosthetic arm as a command signal.

The Arduino Uno is a microcontroller platform which is open source built on a basic I/O board. It comes with a development environment that allows you to use the processing language. It is an ATmega328P-based microcontroller board. There are fourteen pins that are digital on it. Six of these pins may be utilised as Pulse Width Modulation outputs, while the other six can be used as analogue inputs. It contains a Universal Serial Bus (USB) connection and a quartz crystal with a frequency of 16 MHz. It may be powered via a USB connection linked to a computer or by simply connecting to a battery.

The framework of the prosthetic arm might be 3D printed or made of metal. The cost of a 3D-printed limb will be less than that of a robotic arm. Each fingertip will be hooked to its servo motor. The flexion, extension, and pinch movements will be regulated by all these linear actuators. The command signal created by Arduino Uno in response to the brainwave values received will control these three actions. As a result, the prosthetic arm may be manipulated in real time utilising the command signal.

The three actions of the prosthetic arm are controlled by the attention level. These numbers may be divided into three categories. A distinct action is assigned to each range. These activities will be carried out by the Arduino in response to the raw EEG readings received. The attention ranges for various actions are shown in Table 7.1.

To regulate these three actions, the user must educate his or her brain to improve control precision and significantly minimise latency.

EEG-based brain-controlled prosthetic arms have sparked considerable interest because they have tremendous potential to raise the standard of living for patients with disabilities. The attention values have been divided into three groups, two of which are used to regulate the two primary hand motions. The brainwaves were used to regulate these three movements with an accuracy of 80%. Using a large number of sensors to collect brain waves from various parts of the brain allows for greater precision. A long period of training would enable the user to manipulate the arm more precisely. Broader EEG sensors would also improve accuracy and allow for more ranges to be explored (Bright et al. 2016).

7.4.2 Mind-Controlled Wheelchair

Wheelchairs are used by people that are unable to walk due to illness, injury, problems related to old age, or disability, which can be caused by various mental and physical conditions. Modern electric wheelchairs have completely changed how disabled people go up steep inclines, and ramps and also provide tighter turning radius due to a powerful engine and smaller wheels. However, this requires some amount of proficiency to control the wheelchair as it is essentially a motor vehicle in today's day and age, and not every patient who needs a wheelchair is able to control the modern wheelchair deftly. A mind-controlled wheelchair, that allows the user to calibrate their will to physically control the speed and direction of the wheelchair, proves really beneficial to users who are not able to drive such machines. This allows venerated users to control the wheelchair as an extension of their bodies, empowering them to overcome their physical disabilities.

This section discusses an approach to building mind-controlled wheelchairs, including the data requirement, signal processing tasks, the various ML models employed, and the hardware components required for the same.

The first step is data acquisition, the unfinished data from the brain is collected with the aid of an EEG headset called Emotiv EPOC+, which is set up according to the 10/20 international electrode location scheme. AF3, F7, FC5, FC6, F3, T7, P7, O1, O2, P8, T8, F4, F8, and AF4 are the 14 channel electrodes on the Emotiv headset, including two reference electrodes. The acquired signals are sampled at 128 Hz by the headset. The gathered signals should be sent to the CPU. Since the data collected from all 14 channels is massive, the system automatically should choose the top k maximum power channels utilising power spectral density to reduce processing times and increase system throughput (Swee et al. 2016).

TABLE 7.1
Summary of All the Feature Extraction Techniques Discussed

Section	Key Points
Filter Bank Common Spatial Pattern	- CSP is employed in the development of spatial filters for Event-related Desynchronisation and Event-related Synchronisation. • The intent of CSP spatial filtering is to elevate the volatility distinctions among two groups of EEG observations.SBCSP (Sub-Band Common Spatial Pattern) emerged to solve the problem of establishing frequency ranges for CSP. • It improves the precision of classification by using spatial filters on sub-bands.
Deep and Shallow ConvNets	- ConvNets are employed for EEG classification, but EEG information remains peculiar in that it is data from time series with a low signal-to-noise ratio. - Deep ConvNets and Shallow ConvNets are utilised to classify EEG data. - Deep ConvNets are made up of convolution-max-pooling blocks. - Shallow ConvNets duplicate each stage of FBCSP, incorporating bandpass and spatial filtering. - These ConvNets outperform standard approaches in terms of deciphering accuracy.
EEGNet	- EEGNet is a small CNN-based framework for classifying EEG signals. - It captures EEG feature extraction principles using depthwise and segmented convolutions. - The EEGNet design is intended to be generalisable across many BCI paradigms. - It has far less variables than comparable models while yet performing well.
Hybrid Deep Neural Networks using Transfer Learning	- HDNN-TL is an amalgamated framework for EEG signal categorisation that employs CNN and LSTM. - It handles specific variations in EEG signals by fine-tuning models for classification with little training data using Transfer Learning. - In regard to correctness and speed of computation, HDNN-TL exceeds standard approaches.
Siamese Neural Networks	- Siamese Neural Networks are used for one-shot learning, recognising few signals per class with minimal training data. - They prioritise similarity and focus on learning representations for similar classes. - However, they require longer training time compared to traditional networks and donot provide probabilities but distances.
Restricted Boltzmann Machine and Deep Belief Network	- Restricted Boltzmann Machines (RBMs) are a practical alternative to Boltzmann Machines. - Deep Belief Networks (DBNs) stack multiple RBMs and use pre-training and fine-tuning to classify EEG signals. - DBNs are used for feature extraction and classification. - Band-pass filtering is applied to EEG signals to remove noise.

The second step is preprocessing, raw EEG signals are extremely noisy and provide little information for the classification task at hand, however, there are signal preprocessing methods that can be applied to extract quantifiable features from the signals. Band-pass filtering is one such method, which allows only the signals in a certain frequency range to pass through and filters out or attenuates the signals that lie outside the predefined frequency range. This reduces noise in signal by eliminating signals that are proven to be an impediment to MI tasks, in previous research. Before the signals can be sent for feature extraction, they must be divided into several classes such as Delta, Theta, Alpha, etc.

These are some of the preprocessing techniques used for signal processing that can be used for EEG signals.

After the signals have been rid of the noise by preprocessing, the next step is feature learning. There are several techniques for feature extractions that fall under two main branches namely, signal analysis and representation learning. Signal analysis techniques include Short-Time Fourier Transform, Fast Fourier Transform, Wavelet Transform, etc. EEG signals are non-stationary in nature, that is, changing with time, and hence cannot be represented accurately with FFT. STFT can be used for non-stationary signals. STFT essentially breaks down each signal into equal time segments and applies Fourier Transform on each of the segments, to decompose the signal into sinusoidal components.

Representation Learning can also be used by ML models to learn the feature autonomously, reducing the time allocated for manual feature selection and also works for different kinds of signals. CNNs are widely used for feature extraction of EEG signals, and some of the popular architectures were introduced in the 'Approaches' section. Shallow and Deep CNNs are the simplest form of the model and can be used if the dataset collected is large as it learns the probability distribution of the classes over the input data. On the other hand, if the dataset is small, or specific to the user then one-shot learning techniques like Siamese neural networks can be used for feature extraction that uses a distance measure to classify the inputs.

The hardware system of this project is the one proposed by Sim Kok Swee et al (Swee et al. 2016). Hardware components of this system include the Emotiv EPOC+ headset, an Arduino board, an electric wheelchair, and a Bluetooth module. The entire system is connected as shown in Figure 7.11, the computer is connected for debugging purposes during development and is removed from the system after deployment.

High torque motors are necessary to operate the wheelchair with a person on it. As a result, scooter motors are used to power the wheelchair's back wheels. This motor's maximum voltage is DC 24 V. The motor has a power rating of 250 W. The motor spins at 3000 rpm and has a rated torque of 0.80Nm. The highest current drawn by this motor is 13.4 A.

Since the scooter motors demand a lot of electricity, a motor driver that can keep up is required. The scooter motor requires a lot of current, up to 13.5 A, thus the motor driver's current rating must be higher than the DC motor's maximum current. The motor driver that must be chosen such that it is a high current motor driver which

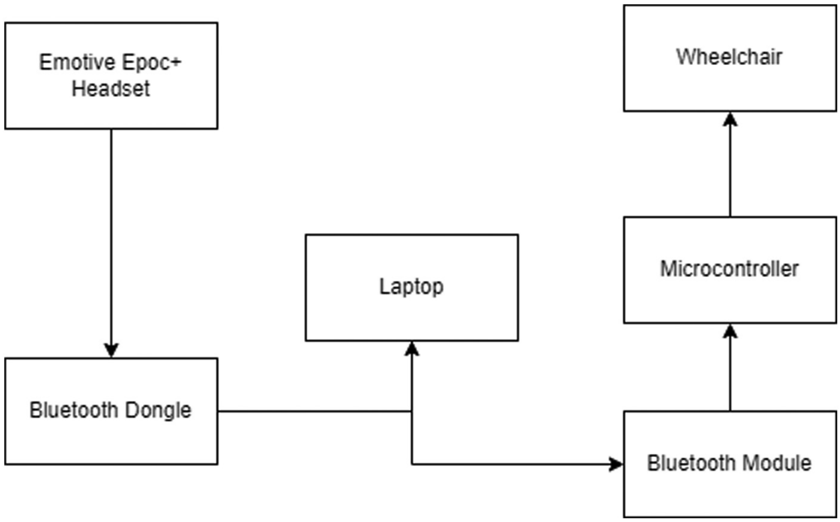

FIGURE 7.11 Information flow diagram of mind controlled wheelchair.

works with DC voltages ranging from 12 to 30 V. It should also be able to withstand a peak current of 160 A and a constant current of 60 A.

The ATmega328P microcontroller IC is used in the Arduino Uno microcontroller which includes 14 digital I/O pins and 6 analogue inputs on this microcontroller. Six of the 14 digital I/O pins can be set as PWM outputs directly. This microcontroller board also features a 16-MHz quartz crystal, a voltage regulator IC that can handle input voltages ranging from 6 to 20 V. The dimensions of the microcontroller board are 68.6 mm × 53.4 mm. There are four input switches throughout the wheelchair.

As mentioned previously, the motor driver requires two PWM pins for control signals and four output pins. The usage of pins is as follows, four analogue pins are used for input switches, four digital output pins are used for output, and two PWM pins are used to operate the motor driver. Moreover, because this microcontroller is miniature, it can be easily embedded inside the wheelchair. Bluetooth HC-06 can be used as the wireless module. The HC-06 is a Bluetooth serial module that utilises a serial port to facilitate a Bluetooth connection by using the Serial Port Protocol. It has a baseband and a 2.4 GHz radio transceiver. This module has a low-power I/O range of 1.8–3.6 V. The HC-06 connects to the Arduino Uno, this enables the wheelchair's microcontroller board to communicate and accept command signals wirelessly from any Bluetooth device.

The Emotiv EPOC headset is a portable EEG device with several channels and excellent resolution for BCI applications. This headgear has 14 channels, the data for which are collected by 14 electrodes to capture the brain's EEG signals. It has the most electrodes of any commercially available EEG gadget. A single ADC with sequential sampling is used in this device. A sampling rate of 128 samples per second is used by the authors. The headset's battery is a rechargeable LiPoly battery that can last for up to 12 hours. Push, pull, left, right and other mental instructions are some examples that should be considered during implementation. The brain command classification model is extremely helpful in reading the user's mental command and controlling the electrical wheelchair's movement.

7.5 CONCLUSION

Brain-controlled interfaces (BCIs) are used to enable users to interact with IoT devices using their thoughts. Such systems are gaining popularity due to their diverse range of potential applications, as well as the continued benefits they provide in the medical field. The most challenging and important component of a BCI is the classifier, which is responsible for interpreting and classifying brain waves into different actions to be executed by the IoT devices. In this chapter, EEG-based BCIs are explored which analyse EEG signals to execute different actions. The section titled 'EEG Extraction Techniques' focused on various ways in which a BCI classifier can be built using modern ML and deep learning techniques. The summary of the methods discussed have been compared and noted down in Table 7.1. Two major applications of these systems in the medical field are also discussed, which helped in easing the lives of people suffering from physical disabilities.

REFERENCES

Ang, K. K., Chin, Z. Y., Zhang, H., & Guan, C. (2008, June). "Filter bank common spatial pattern (FBCSP) in brain-computer interface." In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)* (pp. 2390–2397). IEEE.

Bast, T., Boppel, T., Rupp, A., Harting, I., Hoechstetter, K., Fauser, S., ... Scherg, M. (2006). "Noninvasive source localization of interictal EEG spikes: Effects of signal-to-noise ratio and averaging." *Journal of Clinical Neurophysiology*, 23(6), 487–497.

Bright, D., Nair, A., Salvekar, D., & Bhisikar, S. (2016, June). "EEG-based brain-controlled prosthetic arm." In *2016 Conference on Advances in Signal Processing* (CASP) (pp. 479–483). IEEE.

Gemein, L. A., Schirrmeister, R. T., Chrabąszcz, P., Wilson, D., Boedecker, J., Schulze-Bonhage, A., ... Ball, T. (2020). "Machine-learning-based diagnostics of EEG pathology." *NeuroImage*, 220, 117021.

Hamel, P., & Eck, D. (2010, August). Learning features from music audio with deep belief networks. In *ISMIR* (Vol. 10, pp. 339-344).

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). "A fast learning algorithm for deep belief nets." *Neural Computation*, 18(7), 1527–1554.

How deep learning is changing machine learning AI in EEG Data Processing. Bitbrain. (2021, March 25). Retrieved March 29, 2022, from https://www.bitbrain.com/blog/ai-eeg-data-processing.

Johns Hopkins Medicine. (2021, August 8). "Electroencephalogram (EEG)." Retrieved March 28, 2022, from https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electroencephalogram-eeg.

Koch, G., Zemel, R., & Salakhutdinov, R. (2015, July). Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop* (Vol. 2, No. 1, pp. 1-30).

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces." *Journal of Neural Engineering*, 15(5), 056013.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning." *Nature*, 521(7553), 436–444.

Lu, N., Li, T., Ren, X., & Miao, H. (2016). "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(6), 566–576.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Mayo Foundation for Medical Education and Research. (2022, May 11). "EEG (electroencephalogram)." *Mayo Clinic*. Retrieved March 25, 2022, from https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875.

Oon, H. N., Saidatul, A., & Ibrahim, Z. (2018, August). Analysis on Non-linear features of electroencephalogram (EEG) signal for neuromarketing application. In *2018 international conference on computational approach in smart systems design and applications (ICASSDA)* (pp. 1-8). IEEE.

Sakhavi, S., Guan, C., & Yan, S. (2018). "Learning temporal information for brain-computer interface using convolutional neural networks." *IEEE Transactions on Neural Networks and Learning Systems*, 29(11), 5619–5629.

Salakhutdinov, R., Mnih, A., & Hinton, G. (2007, June). Restricted Boltzmann machines for collaborative filtering. In *Proceedings of the 24th international conference on Machine learning* (pp. 791-798).

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., ... Ball, T. (2017). "Deep learning with convolutional neural networks for EEG decoding and visualization." *Human Brain Mapping*, 38(11), 5391–5420.

Shahtalebi, S., Asif, A., & Mohammadi, A. (2020, July). "Siamese neural networks for EEG-based brain-computer interfaces." In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 442–446). IEEE.

Shih, J. J., Krusienski, D. J., & Wolpaw, J. R. (2012, March). "Brain-computer interfaces in medicine." Mayo Clinic Proceedings. Retrieved March 29, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3497935/

Sreeja, S. R., Rabha, J., Nagarjuna, K. Y., Samanta, D., Mitra, P., & Sarma, M. (2017, October). "Motor imagery EEG signal processing and classification using machine learning approach." In *2017 International Conference on New Trends in Computing Sciences (ICTCS)* (pp. 61–66). IEEE.

Swee, S. K., Kiang, K. D. T., & You, L. Z. (2016). "EEG-controlled wheelchair." In *MATEC Web of Conferences* (Vol. 51, p. 02011). EDP Sciences.

Tangermann, M., Müller, K. R., Aertsen, A., Birbaumer, N., Braun, C., Brunner, C., ... Blankertz, B. (2012). "Review of the BCI competition IV." *Frontiers in Neuroscience*, 55.

Types of boltzmann machines. GeeksforGeeks. (2021, November 20). Retrieved April 3, 2022, pp 6-55from https://www.geeksforgeeks.org/types-of-boltzmann-machines/

Vidaurre, C., Kawanabe, M., von Bünau, P., Blankertz, B., & Müller, K. R. (2010). "Toward unsupervised adaptation of LDA for brain–computer interfaces." *IEEE Transactions on Biomedical Engineering*, 58(3), 587–597.

WebMD. (n.d.). "EEG test (electroencephalogram): Purpose, procedure, & results." *WebMD*. Retrieved March 28, 2022, from https://www.webmd.com/epilepsy/guide/electroencephalogram-eeg.

Wikimedia Foundation. (2022, May 13). "Machine learning." *Wikipedia*. Retrieved April 1, 2022, from https://en.wikipedia.org/wiki/Machine_learning.

Wikimedia Foundation. (2022, May 13). "Restricted Boltzmann machine." *Wikipedia*. Retrieved April 3, 2022, from https://en.wikipedia.org/wiki/Restricted_Boltzmann_machine

Zhang, R., Zong, Q., Dou, L., & Zhao, X. (2019). "A novel hybrid deep learning scheme for four-class motor imagery classification." *Journal of Neural Engineering*, 16(6), 066004.

Zhang, R., Zong, Q., Dou, L., Zhao, X., Tang, Y., & Li, Z. (2021). "Hybrid deep neural network using transfer learning for EEG motor imagery decoding." *Biomedical Signal Processing and Control*, 63, 102144.