Depth of Anesthesia Prediction using Convolutional Neural Network and Ensemble Empirical Mode Decomposition

By

Farhan Rahman

Internal Report

July 2020

Department of Mechanical Engineering

Yuan Ze University

Depth of Anesthesia Prediction using Convolutional Neural Network and Ensemble Empirical Mode Decomposition

Student: Farhan Rahman
Advisor: Prof. Jiann-Shing Shieh

Internal Report
Submitted to Department of Mechanical Engineering
College of Engineering
Yuan Ze University
July 2020
Taoyuan, Taiwan, Republic of China.

Acknowledgements

I would like to express my gratitude to all those who gave me the possibility to complete this research.I wish to thank Prof. J.S. Shieh for giving me this opportunity for this internship at Department of Mechanical Engineering of Yuan Ze University, Taiwan and work on this project from which I learnt a lot. I would also like to thank Mr Ravichandra Mandanu for his constant guidance. I would express my gratitude to Dr. Maysam Abbod (Brunel University, London) for his valuable advice.

I would extend my thanks to management at Vellore Institute of Technology (VIT), Chennai, for providing me with such an avenue to help realize how interesting it is to work in today's industry. Last but not the least, I would also like to thank my parents, for being my motivation to take up this internship.

Depth of Anesthesia Prediction using Convolutional Neural Network and Ensemble Empirical Mode Decomposition

Student: Farhan Rahman Advisor: Prof. Jiann-Shing Shieh

Submitted to Institute of Mechanical Engineering Yuan Ze University

ABSTRACT

According to a survey conducted recently the mortality rate due to anesthesia overdose in the case of surgical complication is close to 47%. This data shows us that there is a need to moderate the level of anesthesia. Recently deep learning (DL) methods have played a major role is estimating the depth of Anesthesia (DoA) of patients and has played an essential role to control the anesthesia overdose. In this paper we have used Electroencephalography (EEG) signals for the prediction of DoA. EEG signals are very complex signals which may require months of training and advanced signal processing features. It’s a point of debate whether DL methods are an improvement over the already existing traditional EEG processing approaches. The convolutional neural network (CNN) is very popular for object recognition and is widely growing its application in processing hierarchy in human visual system. In this paper we have used various decomposition methods for extracting the required EEG signal frequency. After acquiring the necessary signals values in image format we have deployed several CNN model for classification of DoA depending upon their Bispectral Index (BIS) and the signal quality indicator (SQI). The EEG signals were converted into the frequency domain using Ensemble Empirical mode decomposition (EEMD) and empirical mode decomposition (EMD).However, because of the inter mode mixing observed in EMD method, we have implemented EEMD for our study. The CNN models were then used for prediction using the EEG spectrum images without the use of handcrafted features which provides intuitive mapping with high efficiency and reliability. The best trained model gives an accuracy of 83.2%. So, this provides further scope and research which can be carried out in the domain of visual mapping of DoA using EEG signals and deep learning methods.

INDEX TERMS: Depth of anesthesia, convolutional neural network, electroencephalography, empirical mode decomposition, ensemble empirical mode decomposition.

Contents

List of Tables

List of Figures

Chapter 1 Introduction

One of the most vital part of surgical procedures is anesthesia and it is very essential to monitor the level of anesthesia. So, measuring the DoA and monitoring of DoA still posses as a challenge for doctors and researches. Accurate analysis and prediction of anesthesia levels in a patient during a surgery facilitates drug administration, preventing awareness and anesthesia overdose thus improving patient outcome. Resorting to traditional anesthesia monitoring requires experienced anesthesiologists through patient's physiological response. These methods might provide inaccurate results as they are highly oriented towards the experience of the doctor or the anesthesiologists and are not considering the external factors like noise interference with the actual signal values. As per the study conducted by World Health Organization in [1], which shows that the mortality rate due to anesthesia overdose is significantly high in surgical complications. This supports our study that there is a need for improved monitoring system for surgical procedures for patient improvement. There has been considerable amount of research work being carried out for establishing relation between DoA and various features which describe the level of anesthesia in a patient. As proposed by the authors of [2] , that the spontaneous change in the brain's electrical activity during the transition of the different level of anesthesia can be recorded using electrodes placed on the scalp(i.e., electroencephalography (EEG)).

In the modern era of the classification of DoA based on EEG spectrum has gained momentum and various feature extraction method has been carried out. Although we cannot commensurate that EEG based DoA classification is the optimum method, the research work in this field seems to be quite promising. Since EEG spectrum was observed to provide substantial information about the anesthesia level, different analysis methods has been adapted and deployed. These include time-frequency domain and the wavelet transform (WT) in [3]. A comparative between Short Time Frequency Transform (STFT) and continuous wavelet transform (CWT) study was carried out in [4], which shows that the STFT was more efficient in real time process while CWT produced high resolution and high performance which can be used for clinical settings. Authors of [5], used the nonlinear property of the EEG signals and used nonlinear chaotic parameters to identify the anesthetic depth levels. It is observed from the results that Elman network yields an overall accuracy of 99% in detecting the anesthetic depth levels. Hutt [6], deployed a linear neural population model which predicts the concentration of anesthetic propofol using the power spectrum of EEG signals. Zhang et al. [7] adopted the spatio-temporal patterns in the electroencephalogram (EEG) using Lempel-Ziv analysis. Various pattern recognition methods for different cognitive task classification was carried out in [8] with an accuracy of 93% using machine learning algorithm. However, these work utilizes high performance GPU and are carried out over a limited data set using specific task oriented features for classification. This results in accuracy compromise and also inefficiency to resolve the internal differences for individual patient's characteristics. Getting better performance has a tradeoff between time and convoluted methods for feature extraction. Therefore for real time processing and monitoring we require simple feature extraction methods and smaller computational time. This will enhance the patient experience and also result in better accuracy for a larger data set using minimal amount of work. In our proposed work we use simple feature extraction for EEG signals and CNN based classifier for DOA.

Previous research has been carried out in biomedical engineering focused on epilepsy [9], emotion recognition [10], sleep [11] and motor imagining [12]. To bring more light into the field of anesthesia level analysis we have proposed CNN-based DOA monitoring. It is important for patient DOA monitoring during general anesthesia surgery. If the level of anesthesia is too low during the surgical process, the patient will have a slight awareness or feel a slight pain resulting in some postoperative memory impairment [13]. Moreover, long-term maintenance of deep anesthesia can lead to other complications in patients, so anesthesia management is very important [14]. With the already research done in [15], it is evident that there is a correlation between the brain wave activity at different frequency component of the EEG signals and different phenomenon. With the induction of anesthetics there is a significant drop in the activity of the high frequency beta and alpha bands while there is an increased response observed in low frequency band during deep anesthesia level [17].

This in turn creates a feature comparison between EEG and activity intensity during anesthetics induction period, the maintenance period and the recovery period in the time-frequency domain image. With the recent advances in artificial intelligence, computer vision and computer hardware, CNN models are preferred over traditional machine learning algorithms. Various studies from [16], shows that CNN based classification model surpasses the traditional classification models. With the large data set and adequate hardware setup it becomes easier to implement CNN models and can solidify the research in measuring DOA.

A

EEG are in themselves not sufficient to provide much information about the brain activity of a patient therefore there is a need to extract the characteristics of the EEG signals which can be helpful for the DOA. Although processing the raw EEG signals is quite challenging task as they have low signal to noise ratio(SNR) and often the brain activity measurement is often buried under multiple resources of hidden information, environmental, physiological and activity-specific artifacts. Various noise reduction methods and filtering methods have been discussed earlier to extract the true brain activity. EEG signals are also non-stationary signals and have their statistics varying across time. As a result poor accuracy maybe observed for smaller training data and user data might get different results at different time for the same patient so it quite essential to gather sufficient data to overcome this discrepancy.

A lot of works have been put in to handle inter subject variability of EEG signals. For generating a time-frequency domain analysis, short-time Fourier transforms (STFT)[18] has been used for visualizing the non-stationary property of the EEG signals in different cognitive states. Authors of [19] used STFT and auto regressive modeling to effectively detect the burst suppression cause by different anesthetics level. Various other works in [20] was carried out to generate the power spectrogram conversion on the EEG signals. All these reflect the success of STFT to determine the practical fluctuations in the brain activity with the changes in anesthetic features. However, another simple method for feature extraction is Ensemble Empirical mode decomposition (EEMD) which represents a signal at different frequency band. The success of this method has been presented in [21], where they have used Empirical Mode Decomposition (EMD) for EEG feature extraction. Ji et al. in [22], have implemented both DWT and EMD for EEG feature extraction. Authors of [23] suggested the use of Multivariate empirical mode decomposition(MEMD) provides a more robust approach to noise. As a result in our proposed work we have used EEMD method for analyzing the different features of the EEG during different stages of anesthesia. Use of EEMD method solves the shortcoming of EMD method by reducing the inter mode mixing and wide frequency band coverage of EMD signals.

Chapter 2 Experiment

A. SIGNAL ACQUISITION

The research consists of four major areas which can be divided into Signal Acquisition, Pre processing, Feature Extraction and Prediction. Figure 2.1.1, shows us the proposed methodology carried out for our study. This work collects the dataset for the complete surgery of general anesthesia raw EEG signals and anesthesia record sheets collected from the National Taiwan University Hospital (NTUH) anesthesiology department. The data was collected for 50 patients ranging between the age of 23 to 72 years who underwent ENT surgery at NTUH as shown in [24]. The data set comprises an average of 2.5 hours of EEG signal. For our analysis we used a processing interval of 5 seconds which generates about 15400 samples which is sufficient for our experimentation. The data is collected with the help of three electrodes, the positive , negative and the ground electrode. As the signals measured are in the units of microvolt, suitable amplifiers were used without changing the true nature of the EEG signal.

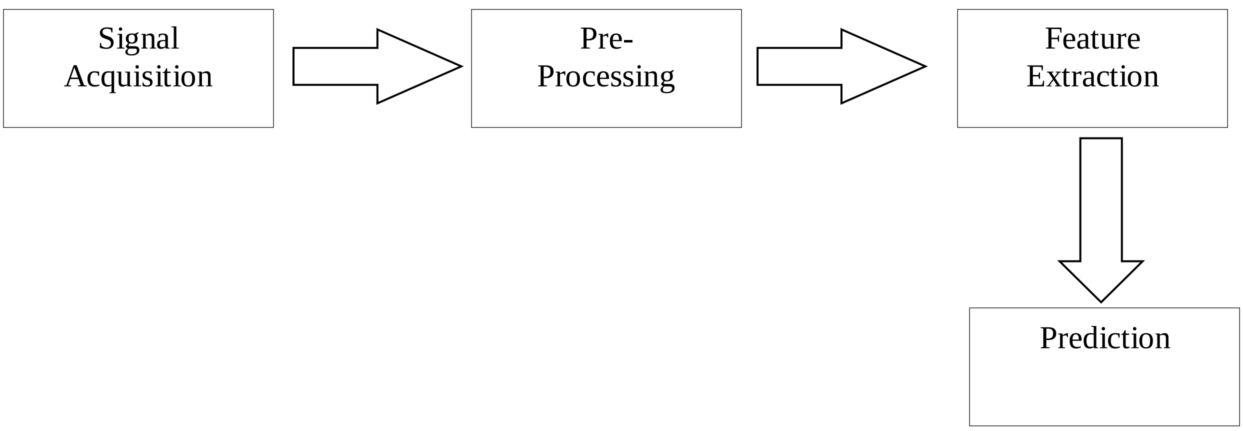


Figure 2.1.1 Block Diagram of Proposed Framework

For the signal acquisition advanced equipments were used like the physiological monitor Phillips IntelliVue MP60, BIS Quatro Sensor module and a portable computer for data-logging in [29]. For the data collection measurements of vital signs like heart rate, blood pressure, SPO2 is collected through an MP60 monitor. Raw EEG, ECG and PPG signals are also collected. The BIS Quatro Sensor is used to measure the BIS index. The monitor is serially connected to NPort through UART (Serial Communication) and uses TCP/IP protocol for data transmission. The NPORT transmits the data received wirelessly to the repeater. The data received from the repeater is then transmitted to the PC. The connection is verified using ping and handshake; when the connection is released, the transmission stops. Fig 2.1.2 shows the block diagram of the signal acquisition process.

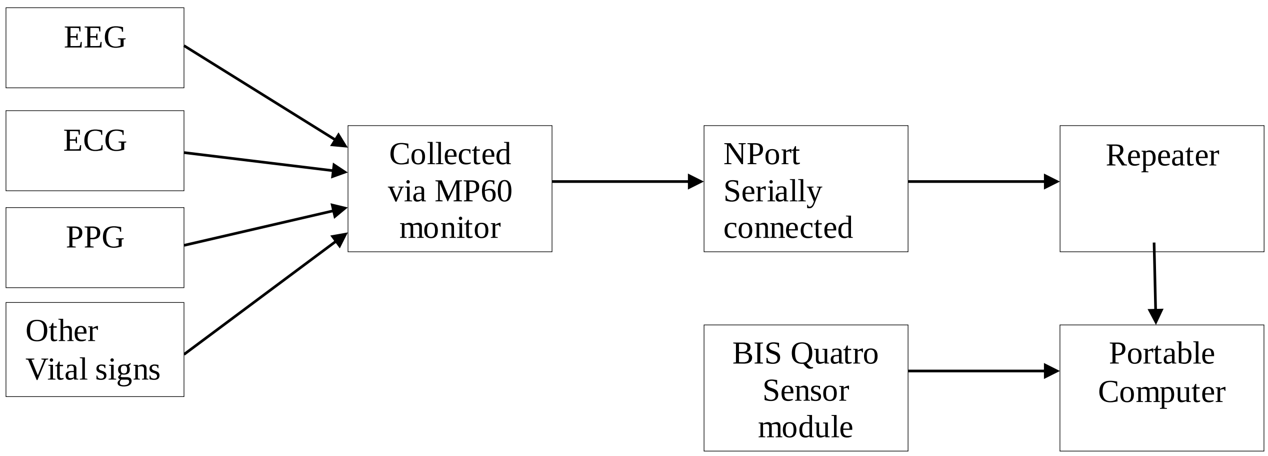


Figure 2.1.2 Signal Acquisition of the EEG signals and other vital signs.

On processing of the EEG signal we observed that there is non-uniformity in the data for each DOA level hence we often say that medical data sets are unbalanced or biased. It was important that we treat each level equally which was overcome by the use of much larger samples for our analysis as compared to the conventional instances. Another way to minimize the unbalanced we have used same number of images for each levels of anesthesia. For classifying the EEG signals into their respective anesthesia levels we have used the average BIS value for categorization which is shown in Table 2.a.

The average BIS value between 40 and 60 is classified as anesthetic OK (AO) thereby this can be considered ‘‘suitable for surgery’’, the value which is less than 40 is considered as anesthetic deep (AD) indicating that the DOA value is low, and the value ranging between 60 and 100 is anesthetic light (AL) indicating that the DOA value is light and may only be preferable for certain types of operational procedures. It is quite natural that the EEG signal will be under the influence of convoluted environmental factors (i.e., electrode off, external frequency interference, etc.), the signal can then be classified as signal polluted (SP) or noise.

Table 2.a . DOA categorization according to BIS Value

BIS Range	Anesthesia Level
0-40	Anesthesia Deep
40-60	Anesthesia Okay
60-100	Anesthesia Light

B. PRE-PROCESSING

Using CNN model for training, the input requires being in the form of a 2D array or in image format so one of the first task before training was to collect the images from the raw EEG data which will be able to provide results using deep learning methods. It was a vital point of our study to process the raw EEG signals into matrix like format with its associated DOA level (i.e. label data), which was crucial as there are many complications associated with the EEG processing signal as discussed earlier. The raw EEG needs to be processed and divide in to three categories which are AL, AO, and AD. As we know that the EEG spectrum shows change in features whenever there is a change in the brain activity for our analysis we have used EEMD method. This method proves to very helpful in extracting the constituent components of a signal. EEMD proves to be very robust and reliable for feature extraction of non-stationary signals which in our case is EEG signal. Since the EEG signal comprises of high to low frequency waves, from Beta (β)-waves ranging from 12-35 Hz, alpha (α)-waves ranging from 8–12 Hz Theta(θ) waves ranging from 4-8 Hz, Delta (δ) waves ranging from 0.5-4 Hz and the frequency beyond 40 Hz can be classified as noise.

Each of the different constituent frequency component of the EEG signal shows variation with DOA levels, so we used these as the classification characteristic. Now, to generate the spectrum plots for EEG signals we have used EEMD which gives us the spectral images of all the four constituent frequency component of the brain wave and then these images we divided according to the DOA levels which includes AO, AL, AD. The raw EEG signals were sampled at an interval of 5 seconds and with a sampling frequency of 125 Hz. With the appropriate window size the signals were processed to generate EEG signal with respect to time. As we know that the EEG signals are non-stationary waves and its characteristics are better observed in the frequency domain, so the signals were decomposed using EEMD method to get the characteristics of the EEG signal at different frequency values.

We often notice that the signal might be compromised because of noise induced in which occurs mainly because of external factors like loss of signal during collection phase or there might be cross connection of hardware setup. Using EEMD method we are able to fragment out the noise signal and use the necessary frequency bands accordingly in our analysis. Previously implemented methods for spectral analysis show that there are chances of feature distortion of the EEG signal. Using STFT for spectral analysis also has the problem of selection of window type, but using EEMD seems to overcome this as it uses the original signal to decompose into the different constituent signals which make up the original signal. Accordingly, EEG data is filtered and decomposed and segmented as mentioned above. For all 50 patients the data is processed and a new DOA index reflecting the three consciousness level is obtained using the CNN method.

For the pre-processing of signal, we considered a window of 5 seconds and as we are aware that the sampling frequency of the EEG signal is 125 Hz so we obtain 625 sample points which represents the row whereas the BIS value corresponding to each window size of the raw EEG signal represents the column of the matrix. The plots of the matrix are generated corresponding to its BIS value. As these raw EEG signals are not adequate to make a prediction model therefore further feature extraction process has to be carried out. For getting differentiable features we have using EEMD method to get the characteristics of the non-stationary signal at different frequency values. After the generation of the respective Intrinsic Mode Functions using EEMD, we analyzed the respective IMFs in their time-frequency domain using power spectrogram. These spectrograms were converted into jpeg format and saved for CNN model training. Figure 2.2.1, shows us the flow chart of the steps taken starting from pre-processing to the feature extraction and the prediction process of our analysis.

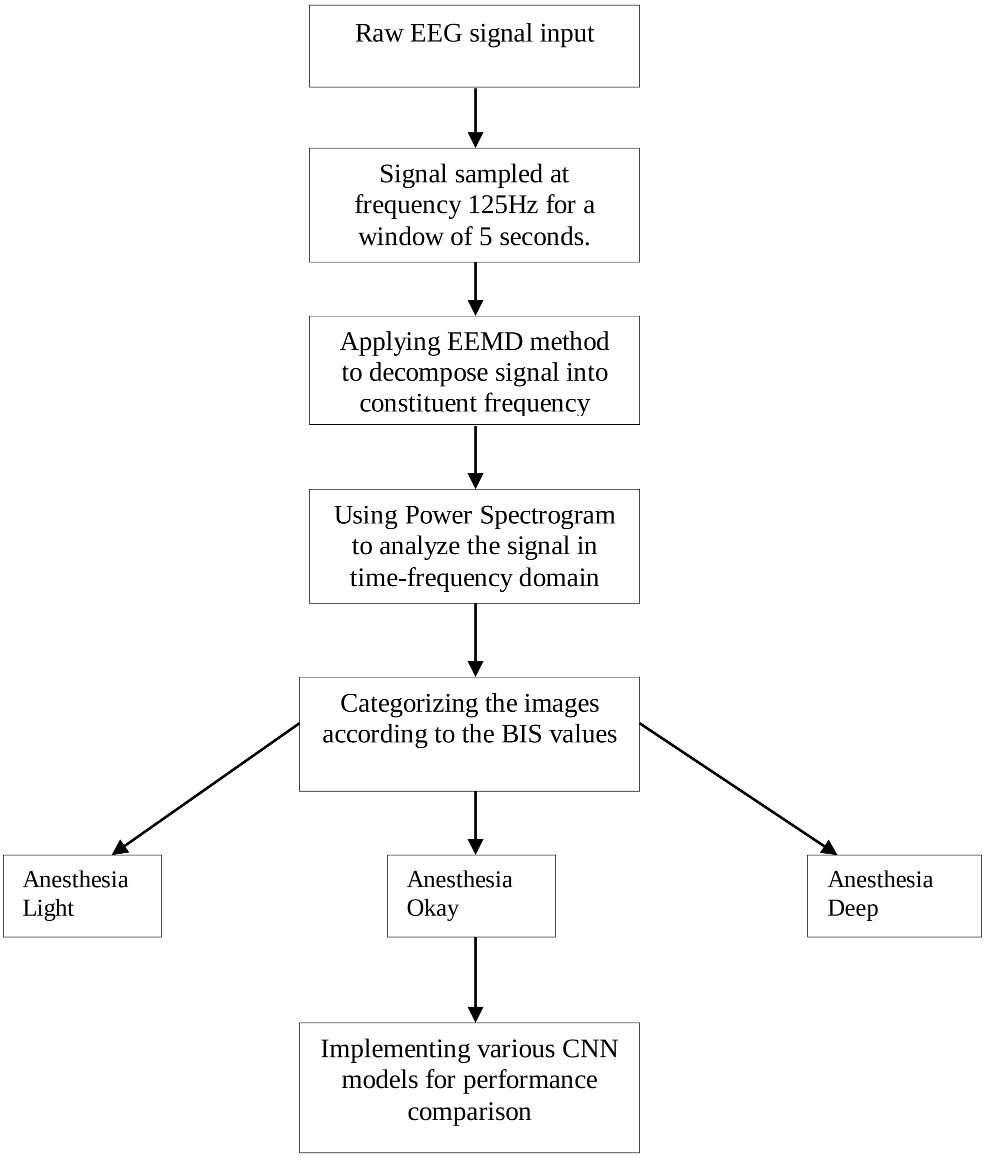
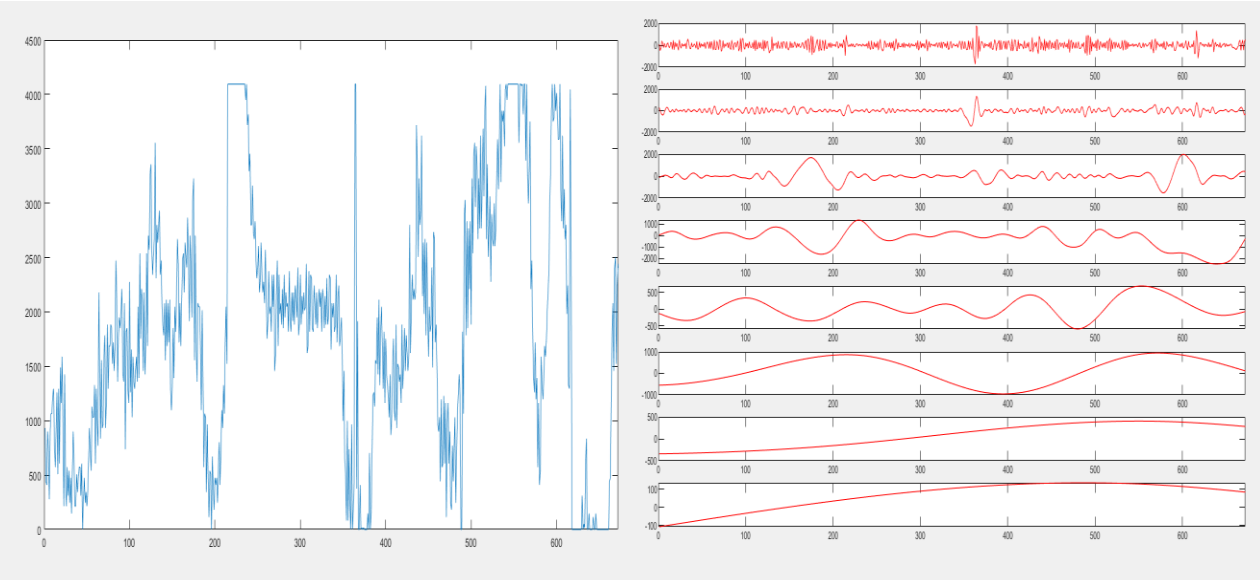


Figure 2.2.1 Flow chart for pre-processing, feature extraction and Prediction

C. FEATURE EXTRACTION

The EEMD method used for our study derives the simple intrinsic characteristics dynamically without the any prior knowledge of the system. The first step of EEMD comprises the addition of an independent uniformly distributed and zero mean white noise with matching intensity of the noise in the signal and then EMD is applied to generate a set of IMFs. The above step is iterated for N times to generate the ensemble of the IMF sets and then the ensemble is averaged to receive a set of IMFs.



(a). Raw EEG Signal **(b). IMFs of EEG via EEMD method**

Figure 2.3.1 The raw EEG and IMFs via EEMD method

The basic working of the EEMD can be concluded as follows:

(1) The original signal is assigned to $x(t)$.

(2) The local maxima and minima of $x(t)$ is calculated.

(3) The upper and lower envelope is generated using cubic spline interpolation between the local maxima and minima: $f_{max}(t)$ and $f_{min}(t)$.

(4) Mean value of the envelope is subtracted from *x(t)*.

$$r(t) = x(t) - \frac{f_{max}(t) - f_{min}(t)}{2}$$

(5)The shift relative tolerance (rel_{tol}) which is the stop criterion of IMF which is set to 0.3 for our study.

$$rel_{tol} = \frac{(r_{i-1}(t) - r_i(t))^2}{r_i(t)^2}$$

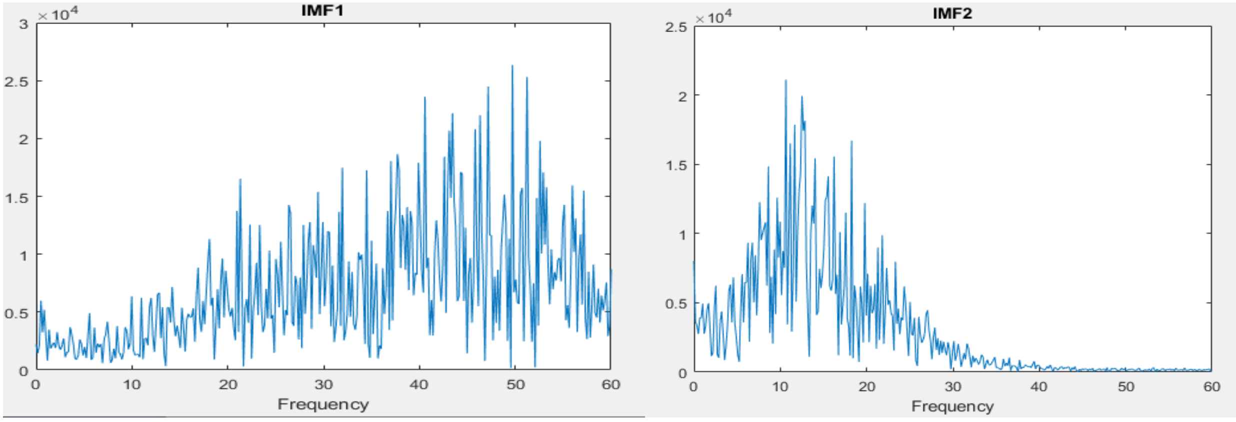
(6)Check if $rel_{tol}$is less than 0.3, if it is, stop the loop and consider the current *r(t)* as an IMF; otherwise assign *r(t)* to *x(t)* and loop over the steps from (2) to(6) .

(7)The original is then subtracted from *r (t)* and steps from (1) to (7) are repeated until *x (t)* can't be decomposed. The original signal is then expressed as:

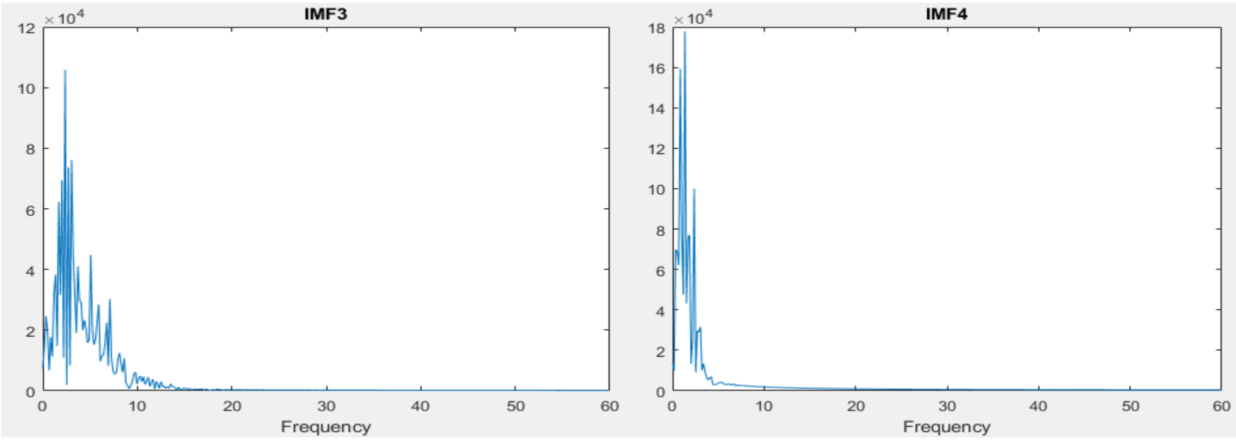
$$x(t) = \sum_{i=1}^{n} IMF_i(t) + res(t)$$

Where *n* is the total number of IMFs and *res (t)* is the residual component.

After the application of the EEMD method over the raw EEG signals for the 50 patient's data we get the desired IMF plots for the different brain waves and the noise signal shown in Figure 2.3.1.Figure 2.3.2, is the representation of the Fourier transform of different Intrinsic Mode Functions (IMF) at different frequencies. For our analysis we have used the first four IMF values. The IMF value ranging from 0-52 Hz is considered as the polluted signal or noise, the IMF value in the interval of 10-33 Hz is considered is the Beta (β)-wave while the signal varying from 7-11 Hz is the alpha (α)-wave and the signal with the interval of 3-7 Hz is the Theta(θ) wave. The remaining IMFs can be neglected as their contribution to the original signal is minimal and therefore only certain frequency signals contribute to the generation of the original signal.



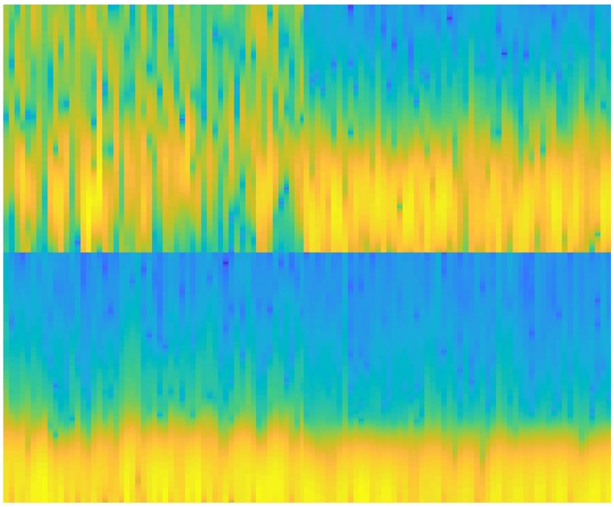
(a) Signal Polluted **(b) Beta (β)-wave**



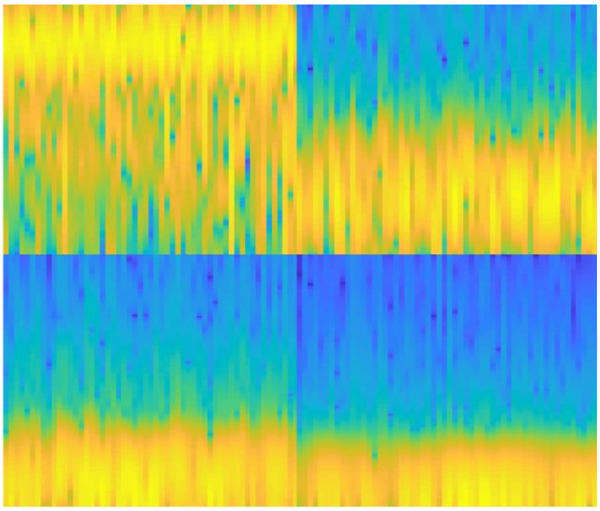
(c) Alpha (α)-wave **(d) Theta (θ)-wave**

Figure 2.3.2 Decomposed Fourier transform plots of IMF of EEG signal.

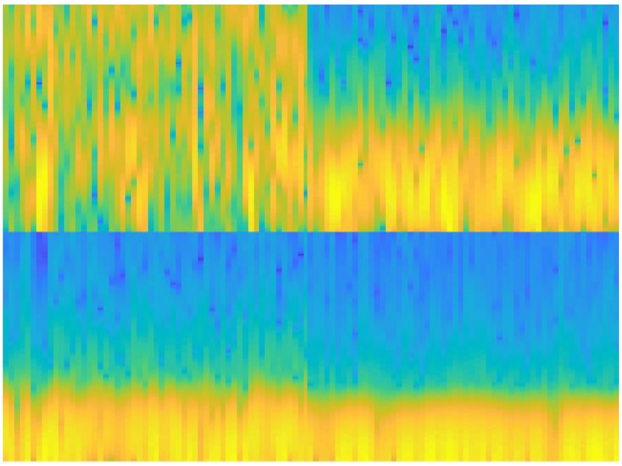
For the analysis in the frequency-time domain, we use power spectrogram which gives us the relation in time-frequency domain of the respective IMF plots. As a result the raw EEG signals were converted into two -dimensional matrix which gives us the sample points on the vertical axis and the horizontal axis gives the frequency corresponding to the sample points. The raw signal was processed for an interval of 5 seconds and each plot generated by EEMD was classified as AO, AL, AD according to the average BIS value as mentioned above. Figure 2.3.3, shows the EEMD spectrum according to the DOA. These plots result in visual mapping of the brain activity without the actual use of any external setup and also reduce the cost and time of computation. It is worth noticing that by taking the above steps we are digitizing the raw values of the input in accordance with the BIS values suggested by experienced anesthesiologists.



(a) Anesthesia Light



(b) Anesthesia Okay



(c) Anesthesia Deep

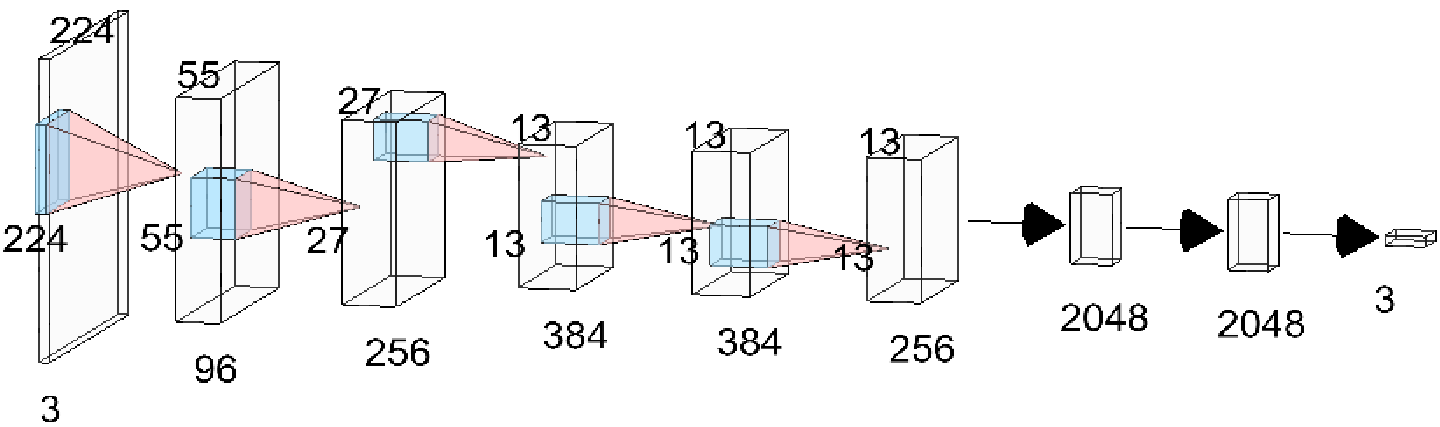
Figure 2.3.3 The power spectral plot of anesthetic states of IMFs 1 to 4

D. CONVOLUTIONAL NEURAL NETWORK MODEL

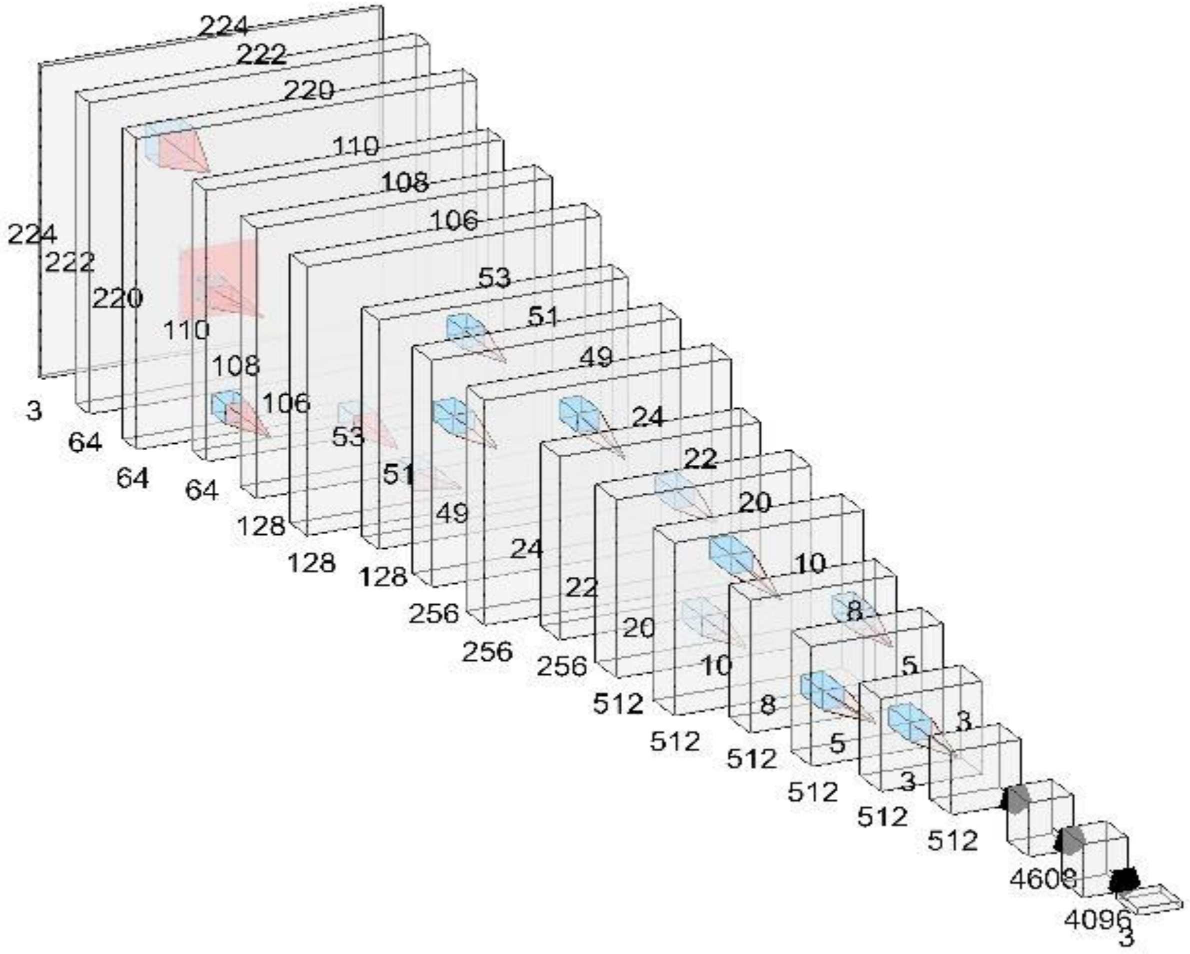
As we are dealing with images belonging to anesthesia spectrum classification, it is reasonable to use CNN model for predicting the DOA levels in the EEG power spectrogram. But, it is very essential that we select the best fit model for our classification because most of the conventional models were trained over data set consisting of general objects. Previous work show that the spectrum analysis of EEG signals using various models like CifarNet, AlexNet and VGGNet model, which effectively proves the advances of in the field of computer vision. For our research work we have deployed AlexNet, VGG16Net, VGG19Net, InceptionRESV2 a 5 layered, 6 layered and finally a 10 layered convolution layer as shown in Figure 2.4. [26-28] shows that better accuracy is attained using complex neural network. However, training complex CNN has high computational time and GPU capacity. After using various model, we inferred that the models which use pre-trained weights of ImageNet data set gives less accurate prediction results. But considering the GPU capacity and the sample data we can use simple convolutional neural network.

We used 5 layers deep CNN for simple analysis with much smaller 3×3 filters in each convolutional layers and combined them as a sequence of convolutions. Further improvement was made using the AlexNet model, 6 layered deep CNN and a 10 layer deep CNN model which uses multiple smaller kernel sized filters stacked up one after the other. VGG16 and VGG19 is 16 layers deep and 19 layer deep respectively with 3x3 sized filters used at different stages of convolutional layer instead of a larger sized filter used at a single point of the model. The back end for the model training was TensorFlow framework and various python libraries were used for our analysis. For the model training of the five layers CNN we used the input shape of 128x128x3 as we used the RGB image format and the output is three classes. The other models were trained with 128x128 size as an input image except for the VGG16 and VGG19Net model. By changing the dimension of the input size and using multiple non-linear layers to increase the depth of the network increases the capacity of the model to differentiate between the complex features of the input EEG spectrum images with lower cost. With a given smaller receptive field of the effective area size of input image where output depends, this multiple non-linear layers can increase the depth of the network which enables it to learn more complex features with a lower cost.VGG16 training the input and output size of the image was set to 224x224x3 while dropout were added in between very dense layers with minor changes in the activation function (i.eRelu,tanh).

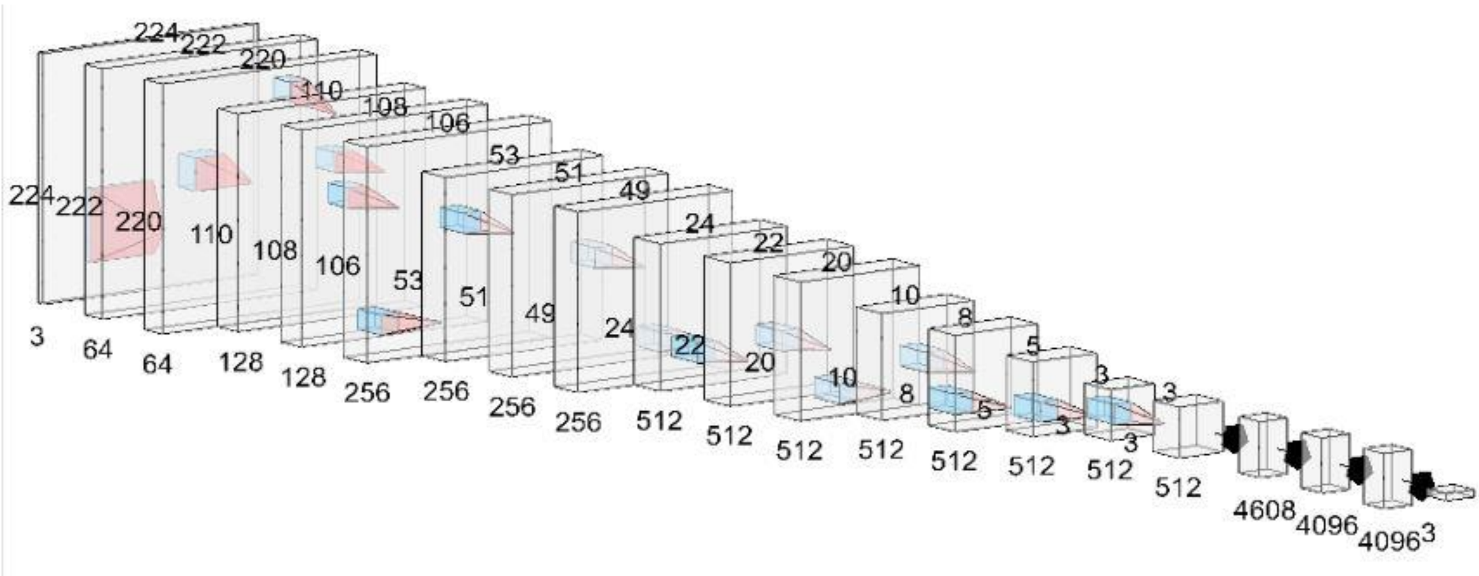
We have used simple processing and minimal optimization, mathematical operations for multi-dimensional array for achieving the best fit. The learning rate for the models was varied from 0.01 to 0.0001 and the various batch sizes of 32, 64, and 128 were used to attain the best fit for the model. Each of the models was trained over different epochs varying from 50 to 200 to get smoother accuracy curves. Our study focuses on comparison being made on the different CNN model and simple feature extraction method for DOA classification. All the models were trained over the same data set and a comparative analysis was carried out.



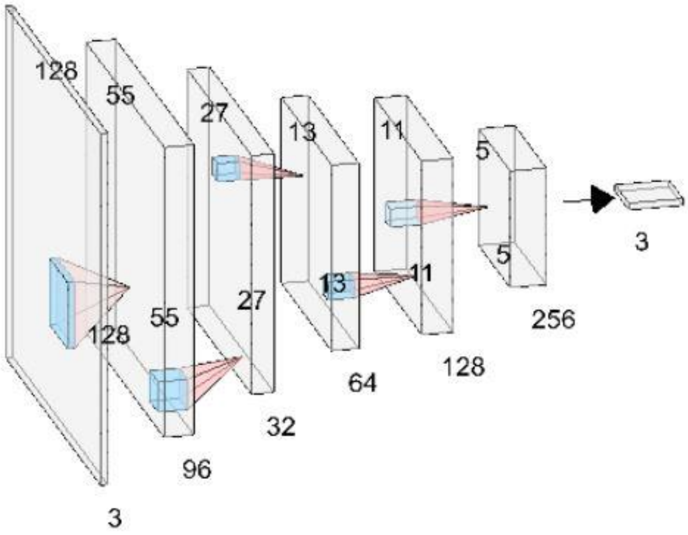
(a). AlexNet model

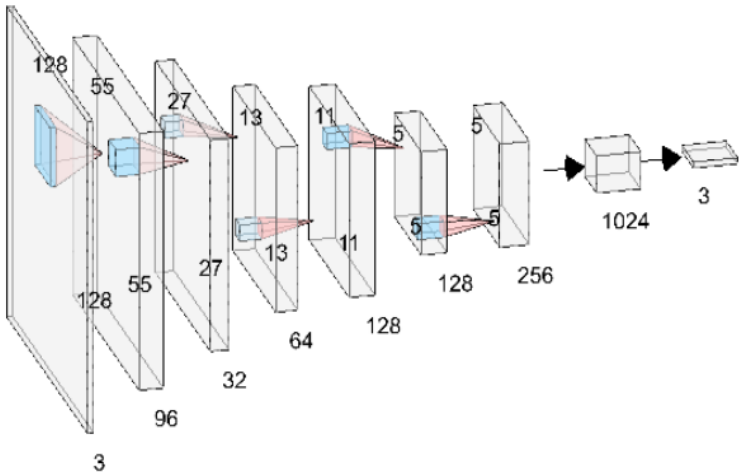


(b). VGG16 Net Model

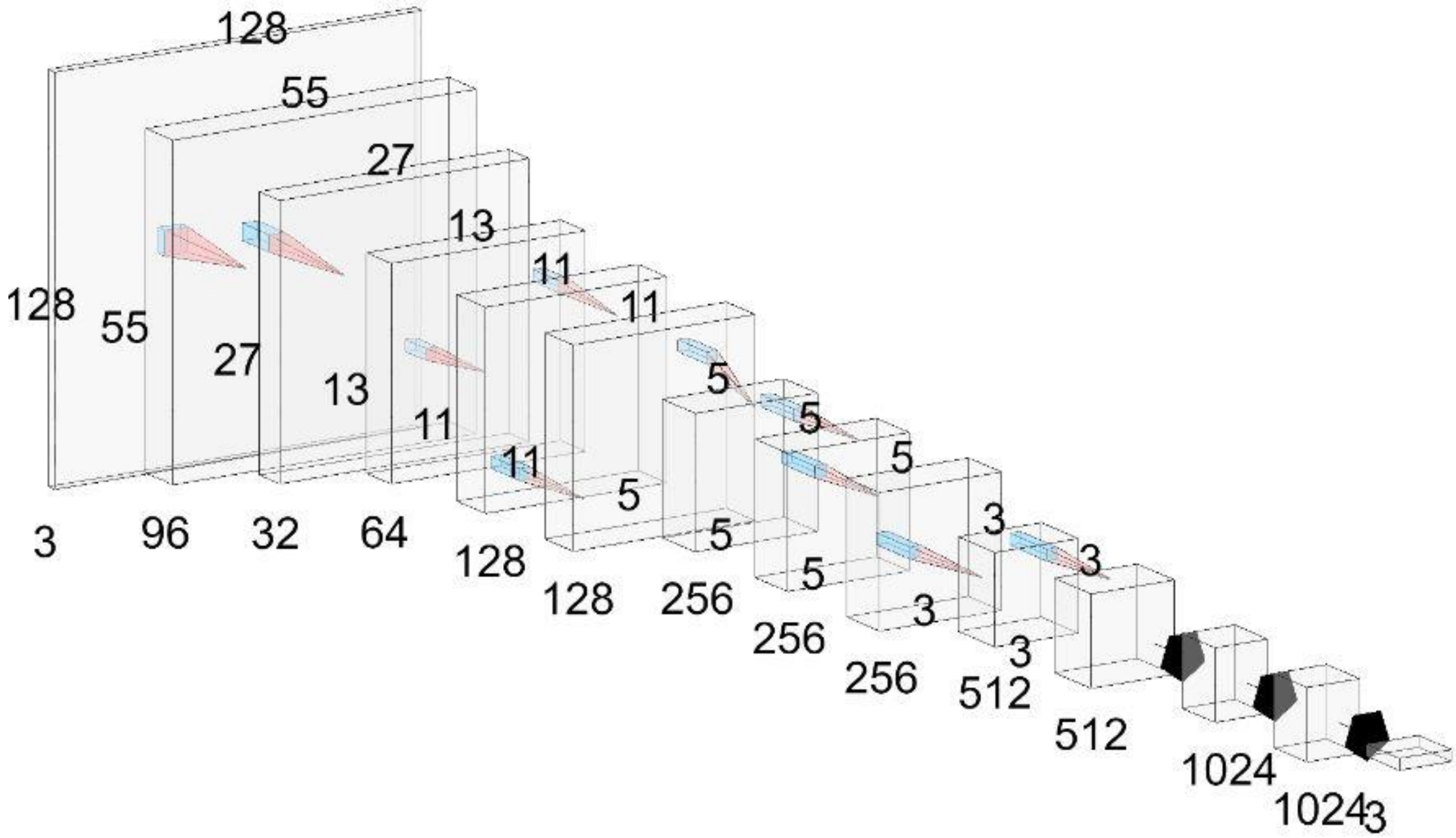


(c). VGG19 Net Model





(d). 5 layer deep CNN Model **(e). 6 layer deep CNN Model**



(f). 10 layer deep CNN Model

Figure 2.4 : The structure of deep learning model with fully connected layers and convolutional layers

Chapter 3 Results

A. PREDICTION PERFORMANCE

Initially a small sample size was used for our analysis and the models were compared. For a data set belonging to 25 patients, from Table 3.a, we get an accuracy of 74% for the 5 layer deep CNN whereas we get an improved accuracy of 81.2% using 6 layer CNN while best accuracy of about 87.8% was observed for a 10 layer deep CNN model. AlexNet model gave an accuracy of about 75.6% while VGG16 and VGG19 gave 76.7% and 74.3% accuracy respectively. While considering the distribution of the image set, we have divided it into three sections namely training set, validation set and test set. 70% of the data was used as the training set while we used 20% of the images for validation set and the test set comprised of 10% of the images. For GPU we have used Nvidia T4s which reduces the computational time by one-tenth. The time taken for complete training of the VGG16Net takes around 2.3 hours while the 5 layered CNN takes 1.2 hours. The images are complex and high-pixel input so having deep convolutional layer provides an advantage to detect the features of the power spectrogram plots. As discussed earlier that the features of the EEG signal are changing with time so we can get better accuracy with larger sample size. At the beginning of training we use training step, batch size, optimizer and epoch according to the small sample size but the parameters have to change constantly to get better accuracy. In this way the CNN models provides the flexibility to change the layer size, the sub-layer size, the kernel size, learning rate and batch size to attain maximum accuracy of the training model.

Table 3.a Accuracy for different model structure for dataset of 25 patients.

Model Structure	**Accuracy**
5 Layered Model	74
6 Layered Model	81
10 Layered Model	**88**
AlexNet Model	75
Pre-trained VGG16 Model	76
Pre-trained VGG19 Model	73
Pre-trained InceptionRESV2 Model	67

Another parameter that we need to monitor is the maximum epoch which depends on the size of the data set and is determined by the model's ability to reach the steady state. After we have all the parameters of the model figured out we can deploy the model for training the entire data set of 50 patient's data. As we progress with the model training it becomes important to select to monitor the model after each epoch. ModelCheckpoint is used from the callback function is it a major task to get the best results. For our model training we save the best model after every epoch so that the model weights can later be used for testing stage. After including all the images for training, from Table 3.b ,we observe the accuracy for the 5 layered CNN was 72.5% while we get an accuracy of 74.6% with AlexNet whereas VGG16Net and VGG19Net gives an accuracy of 80.1% and 77.4% respectively. The best accuracy of 83.2 % was observed for a model with 10 layers. All the models were trained following early stopping criteria and batch-size of 128 was selected for AlexNet and VGG16Net while 64 batch-size was used for 5 layered CNN model. The datasets in our study maintains balanced distribution in the training and testing procedures to avoid over and under fitting of the classes. As a result we can conclude the proposed method provides robust and reliable benchmark for DOA level classification.

The loss of the models can be observed using the confusion matrix. The confusion matrix is different for each model which is because of the variability of the EEG spectrum features and the CNN model helps to overcome the individualism and offers more precision and consistency. The highest error was observed for the AL level while the AD level gave the best accuracy with least error. This shows us that further fine tuning or better models can be used to minimize the error and help us to analyze the features of the EEG signal efficiently. Although classifying the DOA level according to the BIS values of 40-60 and below 40 is not a standard way of classifying DOA levels. There is no boundary for classification so we can say we sometimes get anomalous behavior for the AL and AD classes.

Table 3.b Accuracy for different model structure for dataset of 50 patients.

Model Structure	**Accuracy (%)**
5 Layered Model	72
6 Layered Model	74
10 Layered Model	**83**
AlexNet Model	74
VGG16 Model	80
VGG19 Model	77
InceptionRESV2 Model	70

We observe that we get better accuracy for simple CNN model while the models with more layers gives us less accuracy as compared to the 5 layer deep CNN. This trend can be explained as a result of the use of pre trained weight of the VGG16Net which uses the ImageNet weights for classification might not to be able to capture the features of the the EEG spectrum images thereby giving less accuracy. Generally we can say that the classification of the models are in accordance with the expected predictions, gives reasonable error rate and the DOA level prediction using CNN and EEMD feature extraction method was successful.

Chapter 4 Discussion and Conclusion

A. DISCUSSION

The above results validate our approach and opens up potential fields of research that was carried out in the time frequency domain. The conversion of the raw EEG signals into spectral plots using EEMD method for DOA evaluation proves to have certain advantages over the already exiting methods. For our study we used simple preprocessing methods without the use of trivial conventional hardware setup. With our proposed framework, we are avoiding convoluted mathematical calculations and segmentation of the DOA level physically. This saves us a lot of time, cost and even the use of experts to some extent. The work still needs to be done to reach a stage where machines or computer vision might be able to replace the anesthesiologists for classification task. With the help of the different CNN models we tried to show case the reliability of our method for DOA level classification. The application field is not unique and limited to a particular area but its uses are widespread. Furthermore, different layers of CNNs are used to explore the effectiveness of the proposed method. Few of the advantages our proposed work are (1) All the models and processing of the EEG signals were done without the use of external hardware and manual support. (2) The use of EEMD method provides more robustness, reliability and helps us to overcome the challenging problem of the inter mode mixing of the EEG signals. Although the use of our proposed work seems to be quite promising, there are certainly few short comings in our findings. Firstly, because of the availability of a biased medical dataset our model might not perform as expected. Secondly, because of the use of different types of anesthetic drugs we cannot differentiate the DOA depending on the type of drugs used and the model can show poor performance. The CNN models tend to show poor performance during the transition the state and especially in the case of AL the model falls short to predict correctly as shown in Fig 8 . The reason being that the transition states are often quite unstable and rapid change of brain activity might be observed which might go unnoticed by our CNN based models.

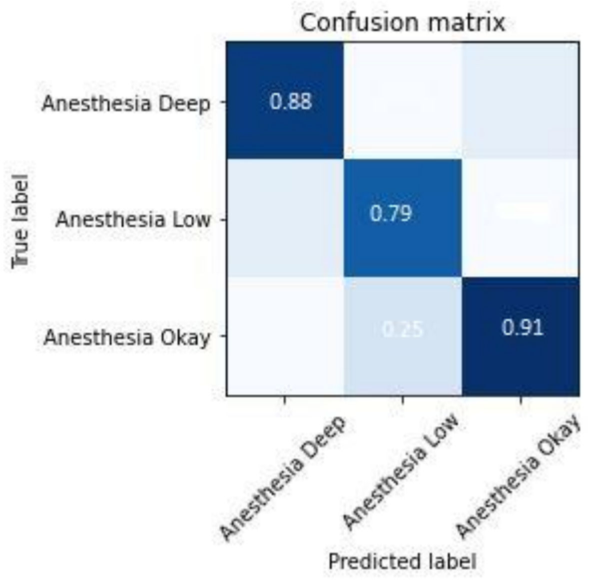


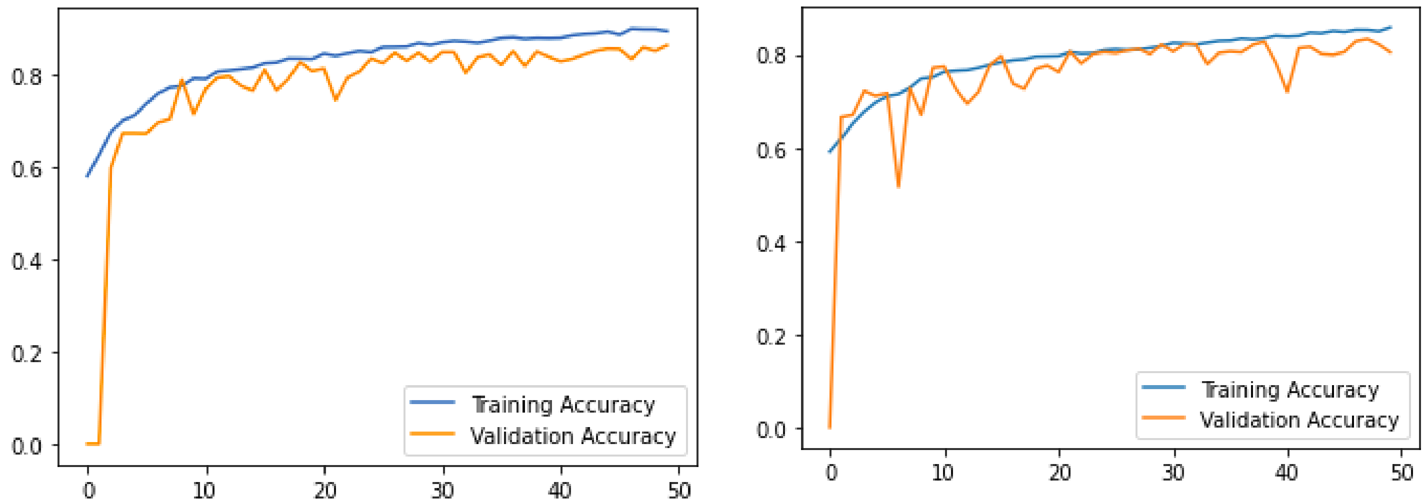
Figure 4.1.1 Confusion Matrix for the best trained model for 50 patients

Since the conventional deep learning models are trained to analyze the object in an image, however, in our study we are using features which do not resemble a particular object and has varying characteristics overtime, so we might expect some researches to be done in the development of deep learning models specific to our image data set. With the help of recent advances in computer vision we were able to decide on the model parameters, neural network layer, optimization technique, activation functions, batch normalization and fine tuning.

Table 4.a Accuracy summary for different models with 25 patients and 50 patients

Model Structure	**Accuracy (%) for 25 Patients**	**Accuracy (%) for 50 Patients**
5 Layered Model	74	72
6 Layered Model	81	74
10 Layered Model	**88**	**83**
AlexNet Model	75	74
VGG16 Model	76	80
VGG19 Model	73	77
InceptionRESV2 Model	67	70

From the training results in Table 4, it can be seen that the deep CNN can accurately identify the DOA features in the EEG signals of the entire anesthesia patient, and the wrong classification results are within acceptable error. The results acquired are support the use of CNN framework for DOA level estimation as many of the models showed accuracy over 80% and the accuracy plot of the best trained model for a dataset of 25 and 50 patients in shown in Fig 9. With the of these kind of framework we can decide on the amount of anesthesia that the patient have to be infused with depending on the type of surgical procedure without any delay. The scope of CNN models is very wide in this field as they are very efficient and has universal application. One of the most interesting attribute if CNN is its ability to learn by itself, even if the data is non-uniform a deep layer CNN will be able to extract the optimal weights for feature training and these weights can later be used to train on similar input images. Even though the advantages of CNN are quite significant but still improvement is required to improve the DOA classification stability and accuracy. In future work, we can explore into the fields of adding more features which can help us categorize the anesthesia levels like the ECG signals and the PPG of a patient and establishing a relation between these features and DOA. With the help of high speed GPU even more deep layered CNN models can be used like VGG19, InceptionNet, GoogleNet, ResNet. Use of additional data will help us to introduce more reliability into the models and help them understand the varying nature of the EEG signals. Another approach can be use of CNN which might be able to provide desired accuracy for a limited patient data set.



(a). Accuracy plot for 25 Patients

(b). Accuracy plot for 50 Patients

Figure 4.1.2. Accuracy plots for 50 epochs of the best trained model

B. CONCLUSION

A significant research has been done in the estimation of DOA and several breakthroughs has been noticed in this field. It has been observed that a large number of assessments of the DOA was based on conventional manual processing of the EEG and a few works showed the visual mapping of the attributes of the EEG using the time-frequency domain. A lot of research has been seen where the use of raw EEG signals are used as time series input for RNN model training. The recently gained momentum in computer vision facilitates the use of CNN model with the capability to assess the DOA level. Use of EEMD method for feature extraction opens up a new scope for research and this approach is rarely experienced. Using EEMD provides sturdiness and makes the feature extraction easier and without the need of manually selecting the EEG signal characteristics and classifying them. With the use of different classes such as the AL, AO and AD we make the task of the anesthesiologist to monitor and evaluate the state of a patient's brain during different anesthetics drug infusion into a patient and take rudimentary course of action to prevent the disparity caused by the different drug usage and the overdose of anesthesia. This further enhances the patient's condition as there maybe chances of psychological trauma if the level of anesthesia is not monitored properly during the transition state. This research work shows us that there a significant correlation between EEG and DOA. The use of EEMD method introduces a novel approach to extract and analyze the EEG features with nominal feature engineering provides an opportunity to establish safer surgical procedure with the use of simpler DOA predictive device.

REFERENCES

[1]. A. Gottschalk, H.V. Aken, M. Zenz, and T. Standl. “Is anesthesia dangerous? “*Deutsches Arzteblatt international*, vol. 108, no. 27, pp. 469–474, Jul. 2011.

[2]. B. Musizza, S. Ribaric. "Monitoring the Depth of Anaesthesia." *Sensors* , vol. 10, no. 12, pp. 10896-10935, 2010.

[3]. M. G. Frasch, L.D. Durosier,N. Gold,M. Cao, B. Matushewski,L. Keenliside,Y. Louzoun, M.G. Ross and B.S. Richardson. “Adaptive shut-down of EEG activity predicts critical acidemia in the near-term ovine fetus” *Physiological reports*, vol. 3, no. 7, e12435, Jul. 2015.

[4]. M.K. Kiymik, I. Güler , A. Dizibüyük , M. Akin.” Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application.” *Comput Biol Med*. 2005, vol. 35, no. 7, pp. 603-616. doi:10.1016/j.compbiomed.2004.05.001, Oct 2005.

[5]. V. Lalitha, C. Eswaran. “Automated detection of anesthetic depth levels using chaotic features with artificial neural networks”, *J Med Syst*, vol. 31, no. 6, pp. 445-452, Dec 2007.

[6]. A. Hutt. “The anesthetic propofol shifts the frequency of maximum spectral power in EEG during general anesthesia: analytical insights from a linear model” *Front ComputNeurosci*, vol.7, no. 2, Feb 2013.

[7]. X. -. Zhang, R. J. Roy and E. W. Jensen, "EEG complexity as a measure of depth of anesthesia for patients," *in IEEE Transactions on Biomedical Engineering*, vol. 48, no. 12, pp. 1424-1433, Dec. 2001.

[8]. H.U. Amin, W. Mumtaz , A.R. Subhani , M.N.M. Saad , A.S. Malik . “ Classification of EEG Signals Based on Pattern Recognition Approach”, *Front ComputNeurosci*, vol. 11, no. 103, Nov 2017.

[9]. U.R. Acharya, S. L. Oh, Y. Hagiwara, J.H. Tan and H. Adeli, “Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals.” *Comput, Biol. Med.*, vol. 100, pp. 270–278, Sep.2017.

[10]. S. Tripathi, S. Acharya, R.D. Sharma, S. Mittal and S. Bhattacharya,“Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset,” *in Proc*, IAAI, pp. 4746–4752, 2017.

[11]. O. Tsinalis, P.M. Matthews and Y. Guo, “Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders,” *Ann. Biomed. Eng.*, vol. 44, no. 5, pp. 1587–1597, 2016.

[12]. Y.R. Tabar and U. Halici, “A novel deep learning approach for classification of EEG motor imagery signals,” *Journal of Neural Engineering*, vol. 14, no. 1, 016003, 2017.

[13]. G. Kotsovolis and G. Komninos, “Awareness during anesthesia: how sure can we be that the patient is sleeping indeed?” *Hippokratia*, vol. 13, pp. 83–9, 2009.

[14]. A. Petsiti, V. Tassoudis, G. Vretzakit, D. Zacharoulis, K. Tepetes, G.Ganeli, et al. “Depth of anesthesia as a risk factor for perioperative morbidity,” *Anesthesiol Res Pract*, 829151,2015.

[15]. Y.Bengio, A. Courville, and P. Vincent, ‘‘Representation learning:A review and new perspectives,’’ *IEEE Trans. Pattern Anal. Mach. Intell.*,vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[16]. O. Tsinalis, P. M. Matthews, Y. Guo, and S. Zafeiriou. (2016). ‘‘Automatic sleep stage scoring with single-channel EEG using convolutional neural networks.’’ [Online]. Available: https://arxiv.org/abs/1610.01683.

[17]. K. Kuizenga, J.M. Wierda, and C.J. Kalkman, “Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane,” *Br J Anaesth*, vol. 86, pp. 354–360, 2001.

[18]. G. Muhammad, M. Masud, S. U. Amin, R. Alrobaea and M. F. Alhamid, "Automatic Seizure Detection in a Mobile Multimedia Framework," *in IEEE Access*, vol. 6, pp. 45372-45383, 2018.

[19]. M. Särkelä, S. Mustola , T. Seppänen , et al. “ Automatic analysis and monitoring of burst suppression in anesthesia” *J ClinMonitComput.*,vol. 17, no. 2, pp. 125-134, 2002.

[20]. N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, and O. Kavehei. (2017). ‘‘A generalised seizure prediction with convolutional neural networks for intracranial and scalp electroencephalogram data analysis.’’ [Online]. Available: https://arxiv.org/abs/1707.01976 .

[21]. M. Bueno-López, E. Giraldo and M. Molinas, "Analysis of neural activity from EEG data based on EMD frequency bands," 2017 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Batumi, pp. 401-405, 2017.

[22]. N. Ji, L. Ma, H. Dong, X. Zhang ,” EEG Signals Feature Extraction Based on DWT and EMD Combined with Approximate Entropy.”, *Brain Sci*. 2019, vol. 9,no. 8, 201, Aug. 2019,doi:10.3390/brainsci9080201

[23]. Q. Liu, L. Ma, S.Z. Fan, M.F. Abbod , J.S. Shieh. "Sample entropy analysis for the estimating depth of anaesthesia through human EEG signal at different levels of unconsciousness during surgeries". *PeerJ*, vol. 6, e4817, May 2018.

[24]. Q. Wei et al., ''A critical care monitoring system for depth of anaesthesia analysis based on entropy analysis and physiological information database,'' *Australas. Phys. Eng. Sci. Med.*, vol. 37, no. 3, pp. 591–605, 2014.

[25]. H. Ge, G. Chen, H. Yu, H. Chen, F. An, 2018. "Theoretical Analysis of Empirical Mode Decomposition." *Symmetry* , vol. 10, no. 11: 623.

[26]. A. Krizhevsky, ''Learning multiple layers of features from tiny images,''M.S. thesis, Dept. Comput. Sci., Univ. Toronto, Toronto, ON, Canada, 2009.

[27]. X. Liu, D. H. Kim, C. Wu and O. Chen, "Resource and Data Optimization for Hardware Implementation of Deep Neural Networks Targeting FPGA-based Edge Devices," *2018 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)*, San Francisco, CA, pp. 1-8, 2018 ,doi: 10.1145/3225209.3225214.

[28]. S. H. Hasanpour, M. Rouhani, M. Fayyaz , and M. Sabokrou. (2016).''Lets keep it simple, using simple architectures to outperform deeper and more complex architectures.'' [Online]. Available: https://arxiv.org/abs/1608.06037 .

[29]. Q. Liu, L. Ma, S.Z. Fan, M. F. Abbod, C.W. Lu, T.Y. Lin, K.K. Jen, S.J. Wu, and J.S. Shieh. 2018. "Design and Evaluation of a Real Time Physiological Signals Acquisition System Implemented in Multi-Operating Rooms for Anesthesia." *J. Med. Syst*, vol. 42, no. 8, pp. 1–19, (August 2018). DOI:https://doi.org/10.1007/s10916-018-0999-1.

.