Predicting the Standard and Deviant Patterns In EEG Signals Based On Deep Learning Model

Mhd Saeed Sharif
School of Architecture, Computing and Engineering, College of Arts, Technology and Innovation, UEL. University Way, Dockland Campus London, E16 2RD, UK.
s.sharif@uel.ac.uk

Ali H. Al-Bayatti
Facultiy of Computing, Engineering and Media
De montfort University, The Gateway, LE1 9BH
Leicester, UK

Ahmed S. Alfakeeh
Faculty of Computing and Information Technology
King Abdulaziz University
Jeddah, KSA

Abstract—**In the recent years, there has been a significant growth in the area of brain computer interference. The main aim of such area is to read the brain activities, formulate a specific/desired output and power a specific approach using such output. Electroencephalography (EEG) may provide an insight into the analysis procedure of the human behavior and the level of the attention. Using the deep learning based neural network has a great success in different applications recently, such as making a decision, classifying a pattern and predicting an outcome by learning from a set of data and build the right weight matrices to represent the prediction outcome or the learning patterns. This research work proposes a novel model based on long short-term memory network to predict the standard and the deviant cases within EEG data sets. The EEG signals are acquired utilizing all the 128 electrodes that represent the 128 channels from infants aged between 5 and 7 months. Statistical approaches, principal component analysis (PCA) and autoregressive (AR) power spectral density estimate have been employed to extract the features from the EEG data sets. The proposed deep learning based model has shown great robustness dealing with different types of features extracted from the processed data sets. Very promising results have been achieved in predicting the standard and deviant cases. The standard case was presented with frequent, repetitive stimulus and the deviant case was presented with infrequent sounds.**

Keywords— *Electroencephalography Signals, Artificial Neural Network, Deep Learning, EEG Features, EEG Channels, Long short-term memory.*

I. INTRODUCTION

Electroencephalography (EEG) can be defined as the process of collecting the electrical brain activities utilizing a specific number of electrodes placed on the scalp. There are different types of the signals obtained from our brain which can be classified based on their frequency as follows: the delta signal has less than 4Hz, the theta signal has 4-7Hz, the alpha signal has a range of 8-13Hz, the beta signal has 14-30Hz, while the gamma signal has frequency above 30 Hz [1, 2]. The Figure 1 illustrates examples of the different EEG signals.

Deep Learning approach has been used in many applications. On of which is to reconstruct images based on what people can think of by deploying their EEG signals. The researches state that their technique can support people who are unable to communicate through speech and help people with law matters such as obtaining eyewitness information instead of using a sketch artist. Their statistics from the data show that the human brain approximately takes 0.17 seconds to produce a reasonable representation of the face. The researchers believe that for the future, images other than faces can be represented through EEG and the work for this is currently being carried out, it also can be used for clinical applications [3].

EEG signals from a premature infant can be used to calculate an estimate of the infant brain's functional maturity. This developed software is able to input massive amounts of EEG data and perform hundreds of computational features without the need of a doctor. The software uses a vector machine algorithm so the features can be merged together to produce a reliable approximation of the maturational age. The results were compared with the software and the clinically known age of the infant and the success rate was over 80% as the software results were within two weeks of both results. It was very reliable so the brain could be analysed in terms of functional development by every few weeks using this software [4].

For epilepsy diagnosis, EEG is very useful as there are EEG bands that are able to detect many brain disorders which includes epilepsy. Hence, a cloud-based system which uses the machine learning approach to identify brain disorders in developing countries has been developed by IBM. This is automated approach assists the decision support for non-specialist physicians when they want to conduct EEG analysis. An experienced neurologist is needed to analyse EEG signals since it is very simple to misinterpret. This setback can be tackled by tele-EEG which makes the use of remote EEG analysis but these alternatives are expensive which is relatively insufficient in developing countries [5].

This research work performed a series of experiments on the EEG data sets, and developed a deep learning based approach to identify the standard and deviant patterns in the processed EEG signals. Such approach will be the core component of further novel research which will aim at quantifying the distraction levels of the infants during the exposing to specific sounds frequently and then later on measuring the learning capability of these infants. We will also investigate the most suitable features that can elucidate the EEG data. This paper is organized as follows. Section II provides an overview of the state-of-the-art techniques. While Section III presents the methods and data description. Section IV shows the achieved outcomes. Finally, Section V presents the conclusions and future work.

II. RELATED STATE-OF-THE-ART TECHNIQUES

The work presented in [6] compares different features that can be extracted for EEG signals in terms of analysis and methods. One of the approach to extract features is called the Hilbert-Huang transform. It makes the use of an arbitrary continuous signal and outputs a signal which is analytic. The input has to be a real data sequence. FFT is used and calculated within the input sequence and eventually calculates the inverse of FFT of the output. Through this method, the power and 'instantaneous' frequency of the signal is determined. Another approach is called the principal component analysis (PCA) which is commonly used to extract features in EEG. It is a non-parametric method for extracting properties from data which is confuse. The output represents the similarities and differences between the data. PCA makes the use of rotation as it aims to find specific parts of the data with lots of differences. PCA can compress the complex data in order to reveal more information for the data which is hidden within the data. PCA is known for identifying patterns and for graphical representation. For the blind source separation (BSS) problem, the independent component analysis (ICA) comes within context for a method. In this approach the signals could be split into separate, individual components. This includes several sources are split from each other and still be considered as individual components. Therefore, the individual components of these signals can be outputted through ICA [7, 8].
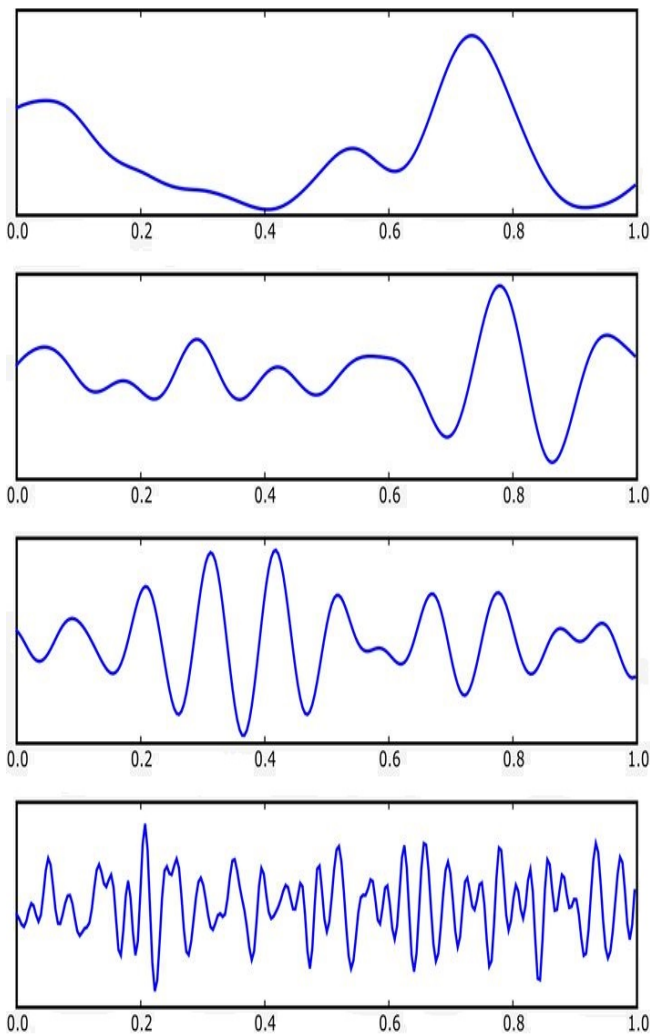

Fig. 1.   Examples of the EEG waves [1].

The work proposed in [9] explains the advantages and disadvantages of certain methods through its performance and its purpose for suitability. The wavelet transform method is used within the recognition and diagnostic field. The input signal is a biomedical signal which contains several data points and converts it into smaller parameters which makes up the signal. The wavelet transform method is a good as the EEG signal is nonstationary, the input data is raw data and this technique is a spectral estimation technique which in turn can be represented as wavelets in an infinite series. Wavelet transform is used in the time, frequency and linear domain. The signal can be transient and stationary. This method extends the FFT method and fixes issues for nonstationary EEG signals. The eigenvector method (EM) is used to determine the power and frequency from the signals from artifact dominated measurements. The EM is used within the frequency domain and makes the use of signal which is buried with noise. Another method is called the autoregressive method. It approximates the power spectrum density (PSD). This method makes the use of a parametric approach. The Yule-Walker method uses the parameters of the Autoregressive method. The autocorrelation data function is used to approximate. It is used in the frequency domain with a signal that has sharp spectral features.

The work in [10] explains and compares different techniques for learning 'discriminative features' from EEG data with the use of deep learning methods. EEG data is usually obtainable in little amounts with the ratio being not good signal-to-noise. In the data, there is unevenness between the EEG data from the patients and the actual recording sessions. The methods tackle these variables for feature learning. Eventually, models can change for each EEG patient specifically depending on the spatial patterns.

Target images in an image rapid serial visual presentation (RSVP) from EEG data using deep learning were detected in [11]. The features were linked to time-frequency events, the deep learning technique was manipulated with the use of input feature clusters. Several trials and subjects were used for this task. A deep belief net (DBN) classifier was used to classify target and non-target images using a clustering method. The paper also explains how the DBN was used to identify certain parts of brain activities. While the work presented in [12] shows that it is difficult to determine what emotions people can feel at a given moment through their facial expression (for example, functional autism patient). The EEG signals can give indications and information in which specific emotions they are feeling during the recording of the disabled patients. For 21 subjects, each database was built for training involving 180 epochs. Four emotions were classified during the experiment. These four emotions were scared, sad, happy and calm. 45 epochs were used for each emotion in each subject. To fine tune the deep learning model for each subject, every individual epoch was used for pre-training. To learn features and characterise emotions at the same time of the EEG data, the paper talks about a deep learning algorithm to tackle this. They trained the emotion model for each individual subject because the EEG data is individual based and there are not many epochs obtainable in this case. Their algorithm/solution uses three layers of 'restricted Boltzmann machines' making the use of pre-training. Therefore, the epochs for each subject can be used to pre-train the deep network model. The results showed that their solution gained recognition stats than conventional algorithms.

Clinicians use EEG signals/data during sleep to assess different types of neurological disorders. Certain algorithms have been implemented to use upon certain types of events but in [13], it is an event-specific. In a EEG time series, a deep learning method was employed where it anticipates the type of event, locations and durations. This deep learning method makes the use of a convolutional neural network which includes feature representations from EEG data. Compared to conventional algorithms/methods, this deep learning method has proved to be efficient. In this experiment, 19 people/subjects were used; 11 are female and 8 are male. A five-split cross validation was used for training: 10 for training the data, 2 for validation and 3 for spindle detection. Over the whole night of EEG recording, this research introduced a new deep learning method to conduct any type of event detection.

To diagnose brain-related disorders, physicians make the use of EEG analysis in hospitals. The drawback is that EEG analysis requires highly qualified/trained clinicians. In [14], a process priority by recognising whether an EEG recording/session is normal or abnormal is proposed. A deep neural networks that learns important features in an automated procedure was used. It used different types of pre-processing techniques and machine learning approaches. TUH Abnormal EEG Corpus dataset was used to assess the performances of the proposed methods. It was found that modern deep gated recurrent neural networks gained 3.47% better performance than other performances from other statistics.

An architecture of multi-view deep learning was used to record brain abnormalities linked with seizures in [15]. A seizure detection module is linked to the multi-view structure in order to pay more attention to the relevant EEG channels in the dataset to obtain relevant information. Deep learning procedures and broad experiments were carried out to validate the efficiency of the architecture with the EEG data. Five-fold cross validation was used to train the model. This proved to be an efficient method to detect seizures with EEG signals. The dataset was called CHB-MIT which is from the Children's Hospital Boston. It is openly available and can be downloaded online. The subjects' data they used involved 23 patients with 18 females and 5 males. This ranged from the ages of 2 to 22.

The work presented in [16] introduced a framework involving signal processing and analysis for EEG. The output from statistical features was used to represent wavelet coefficients. Independent components analysis (ICA) and principal components analysis (PCA) were used to decrease the data in terms of its dimensions. Linear discriminant analysis (LDA) was used as well. The output was used for input into a support vector machine (SVM) to determine if there is epileptic seizure or not. EEG data was obtained from five healthy volunteers. They were awake with two states: eyes open and eyes closed. PCA is a popular and common feature which deals with the reduction of dimensions. It reduces the time complexities and space within the data. It represents the sum-squared error in a variation. PCA is beneficial for splitting signals from various sources.

Independent component analysis (ICA) and principal component analysis (PCA) are used to produce a conclusion in EEG data as there are components that holds significant information of the brain activity. To extract this information, the compression feature of PCA is used with the statistical independence feature of ICA. With cases of cerebral death diagnosis when there is very less brain activity, a small amount of number of components are enough to represent the brain activity. The result shows that using a few number of important components, cerebral activity analysis can be conducted and patients with neurological damages can be considered with this method [17].

Fourier transform and principal component analysis (PCA) was applied to the EEG data before separating it into two parts: deterministic and stochastic in [18]. The Time Varying Vector Auto-regressive (TVVAR) model was used for the stochastic part for output. Experiments were taken place on the EEG signals of the stochastic parts and the model's output. The results of the recognition rates were high in percentages with all being above 90% shows that the TVVAR model is beneficial to study the process of auto-covariance nonstationary vectors.

If the EEG data is corrupted with eye movements and blinks, it is challenging to study the signals. The work presented in [19] provides different strategies to remove artifacts in EEG data. Principal Component Analysis (PCA) and wavelet transforms are methods to remove artifacts which are ocular. PCA was found to be the better method as it outputs the better PSNR value and the process finished quicker when compared to the wavelet method. Electrical activity is recorded over a period through electrodes and the data lies within EEG.

For a brain-computer interface (BCI) system, high values of accuracy are valuable for categorisation or classification. Multi-scale filters and PCA are combined in [20] which introduces a new approach to improve the performance of EEG classification. To extract different components of different frequency bands, a multi-scale filter is needed. This adds the support for frequency resolutions. PCA was used to condense data dimension and improves the accuracy of classification. Data of motor imagery using their own method showed an accuracy of 91%. This value shows that the proposed method may improve the performance of a BCI system.

The research presented in [21] investigated the difference between visual learners and non-visual learners while learning when the participants had no background knowledge. Brain patterns are used to calculate the difference of the learners. 34 university students took part in this study for learning and memory tasks. 128 channels were recorded for EEG data. The students watched an animated learning content as an eight minutes long video. Wavelet features were extracted from the EEG data. The output was filtered into beta and alpha bands. The output was then inputted into a Long-Short Term Memory (LSTM) recurrent neural network (RNN). The training for the network resulted in values over 85% which was very good. The wavelet function has advantages over the Fourier transform so it is more common and mostly used in analysis of EEG data. The localisation in the wavelet function is very accurate in both features of frequency and time. LSTM networks understands information for a long amount of time. It has features to accurately label sequences. It links past inputs and future inputs. The network does classify between two categories and resulted in over 85% accuracy.

Deep learning based neural networks are becoming successful in the computing area as they can learn data efficiently and represent features well. This involves natural language processing, speech recognition and more. It can be used to classify EEG. [22] used 1x3 kernels than standard 3x3 kernels for convolutional filters. They found out that 1x3

kernels produce better performance than the latter. They concluded their technique has good robustness. This study combines two networks: convolutional neural network and Long Short-Term memory. They make use of the center loss function. These networks will categorise the EEG signals. The center loss function is used to increase the classification accuracy as it learns more about the discriminative features.

The research in [23] proposed a method of predicting gender and age through EEG data. Cerebral activities were recorded of 60 participants. Participants were either male or female, and were relaxed with their eyes closed. A hybrid learning framework was used through a deep BLSTM-LSTM model. High accuracy results were obtained of over 90% for both age and gender. The results were much better than normal, standard methods. In terms of frequency bands, the beta band worked out to be the best band for predicting age and gender when compared to other frequency bands.

Deep learning based neural networks were used in [24] for the raw EEG data and the power output classification. 56% accuracy was achieved for EEG raw data and 44.75% accuracy for power data. 63% accuracy was achieved using an attention network. Depending on the feature linked with the EEG signals, we can compute a classification system for it. For people suffering from severe motor impairment, signals can be categorised to recognise which EEG signals relate to the feature. The signals can be used for a particular direction for input: left and right. Also, this includes up and down. Since the attention layer is implemented, it is known as an attention network. A recurrent neural network has the weights between the nodes produce a directed graph. It uses memory to process inputs which are sequences. LSTM networks evade the vanishing gradient problem. In another work different range of features with artificial neural network were employed to classify epilepsy cases from the EEG data sets [25].

Brain signals were found as a viable solution for authentication and validation in [26]. For the analysis purpose, an LSTM network is used with features such as event-related potential (ERP) and steady-state visual evoked potential (SSVEP). To conduct this process, the method is split into three stages. Raw EEG signals were recorded from 20 participants (10 males and 10 females) who were presented with ERP stimulus. Participants watched an LED display for around half an hour. Specific channels were recorded from F3 to 02 and sampled at 250Hz. The EEG signals were saved into a private database, filters were applied such as band-pass to pre-process the raw data. The structure of the LSTM network is used to work on the participants signals. After this is done, the performance of the network is evaluated through the use of false metrics such as the false acceptance rates. The proposed method and network showed high accuracy.

III. Methods and Data Description

The deployed dataset contains hundreds of texts files from 23 patients and every text file is standardised with the same matrix: 550x129. There are 550 samples for every channel up to 129. Every channel represents an electrode that records the brain activity in that region. The last five channels are removed as the last five channels only recorded eye movements, so the input data is now 550x124. For each patient, there are two types of the text files one representing the standard and one for the deviant. The EEG signals were acquired using 128 electrodes from the patients. Each of these electrodes represent channels. 0.1-30Hz band pass was used to filter the raw data. The file was then segmented containing 100ms before stimulus presentation and 1000ms after according to the category: standard and deviant. Standard cases were presented with frequent, repetitive stimulus and deviant was presented with infrequent, different sounds. The location and numbers of all the employed channels are illustrated in Figure 2.

In this research, the mean, standard deviation, skewness, kurtosis, energy, principal component analysis (PCA) and autoregressive (AR) power spectral density estimate using the Yule-Walker methods have been employed to extract features from the data sets. These features have proved to be useful in analysis for EEG data. The mean calculates the average of every channel/signal in the EEG data. The standard deviation shows the distance between the signal from the mean. It shows variation from the quantity. The skewness makes the use of the probability density function by representing its symmetry. Also, Kurtosis makes the use of the probability density function by determining the peakedness.
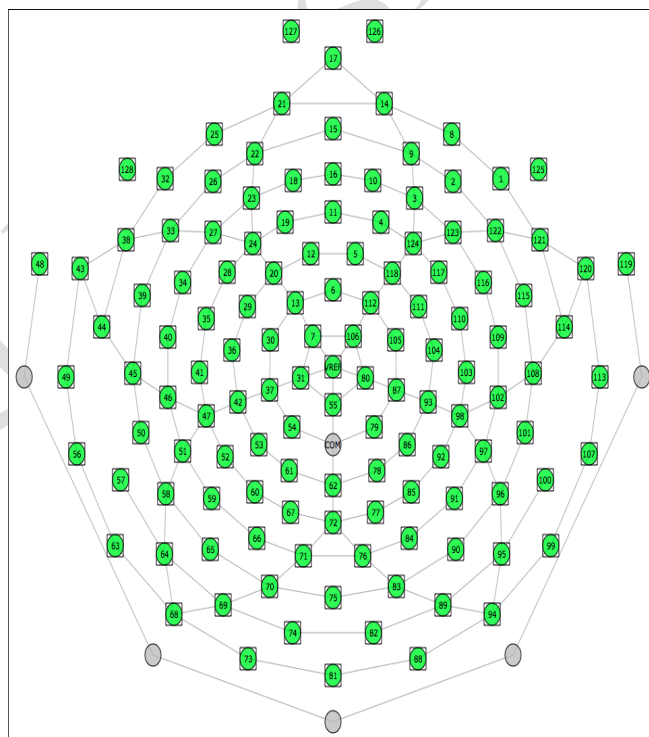

Fig. 2. EEG Channels layout: the location and numbers of all deployed channels.

IV. Results and Discussion

The first step of the developed system is to prepare the acquired data sets and extract the statistics features: mean, standard deviation, skewness, kurtosis and energy. The last 5 channels are removed from the raw EEG data files as the last 5 channels represents eye movement which is not important, hence the data point was 550x124 in each file. The developed approach also creates a categorical variable of targets so 1 represents deviant case and 2 represents standard case. After that sequence classification (long short-term memory network) using deep learning has been employed. All the patient's matrices are combined together, creating one big matrix so it can be used as input for the deep learning network. This network classifies sequences based on the deep learning. It takes a time-series sequence as input, the input size is defined as 550 as there are 550 samples in the input data.

There are two classes as we only have deviant and standard cases. 100 is the default number of hidden units. The layers and options are set by default to define the network. four separate LSTM deep learning networks were created to handle the following types of data: raw EEG data, PCA output, AR output and statistical features of the EEG data respectively.

Different number of training epochs was used to achieve the most appropriate epochs which can lead to the best accuracy. Figure 3 shows the training progress and classification accuracy of the developed network with 20 epochs. While Figure 4 illustrates the training progress and classification accuracy of the developed network with 20 epochs. 75% of the data was used for the training and 25% for the testing and validation.

The peak of the training accuracy for the first network, when the raw EEG data was used, was 95% and maintained an average of 75%. The training uses a smoothing algorithm, making it easier to recognise trends which is the dark blue line displayed in the graph. Different number of patients data files which were not used for training, were used as a testing data to see if the network can recognise if the given test input data is standard or deviant based on the trained network. This resulted a classification accuracy of 79.02% which means that the network predicted 79.02% of the standard and deviant files correctly when the prediction target data is compared to the actual target data of the test data. Given the complication levels of the processed EEG data sets, the achieved accuracy was considered very promising.
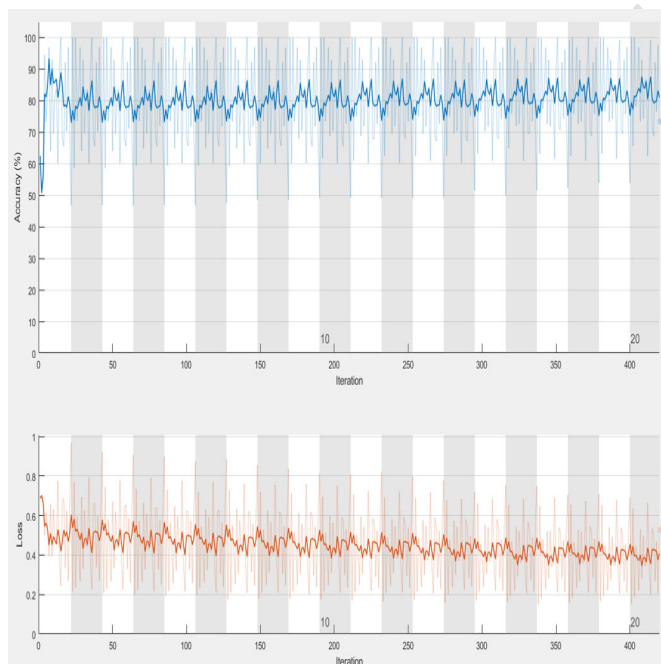

Fig. 3.   Training progress and classification accuracy of the developed network with 20 epoch.

In the second network, based on PCA data, the peak of the training accuracy achieved 95% and maintained an average of above 70% as well. The classification accuracy resulted a value of 79.44% which is similar to the performance of the previous network. This is an indication that the networks are producing consistent values which means that the networks have a good ability to recognise the deviant and standard cases.

The third network was using the AR features. In this network the peak almost achieved 100% of the training accuracy and maintained an average of 75% which is also very similar to previous networks. The classification accuracy resulted in 79.9% which is the highest among these three networks so far when compared to the other networks. Beside the achieved accuracy shows the network consistency of recognising deviant and standard cases.

Fig. 4.   Training progress and classification accuracy of the developed network with 10 epoch.

Another set of features have been deployed, and the network is trained using the statistical features. The training of the network produced similar values to previous networks, which is an average of 70% with the peak near 100%. Testing the developed network produced an accuracy of 79.28%. This shows that the developed network adapted well with the exposed input and had the ability to recognise 79% of standard and deviant cases correctly.

V.   CONCLUSIONS AND FUTURE WORK

This research paper has proposed a novel model based on long short-term memory network to predict the standard and the deviant cases within EEG data sets. Different network structures have been developed, tested and validated. Statistical approaches of the mean, standard deviation, skewness, kurtosis, energy, beside principal component analysis, and autoregressive (AR) power spectral density estimate have been employed to extract the features from the EEG data sets. The developed deep learning based model has shown great robustness dealing with these different types of features extracted from the processed data sets. The developed networks have shown very good consistency of recognising deviant and standard cases in the processed data sets for different patients. The developed model has achieved a classification accuracy of 79.9%, which was considered a very promising result, due to the complication levels and the difficulties associated with the processed EEG data sets.

The developed approach will have a significant contribution to the field of brain computer interference which has grown rapidly with the purpose of reading a human's mind, generating a certain output and controlling objects using brain signals. The acquired EEG data sets are from infants aged between 5 and 7 months. These data sets will be deployed to build an objectively evaluated framework that has the ability to provide better understanding for the infants' behaviour and their attention for stimuli from different environments.

The future work of this research will involve a wider series of experiments, which are conducted to measure the levels of a distraction of the infants while exposed frequently to some specific sounds or distractions within different environments. The main motivation of the whole research is to develop an objectively evaluated complete system, which has the ability to address the research question of whether an infant can detect a change in a stream of sounds when frequently presented standard changes to deviant. The response to such environment can be suggested to be a predictor of further language development for the participants in question. As the employment of electroencephalography data, representatives of the brain activities, may provide an insight into infant behaviour and attention within the living environment. The acquired infant EEG data will be analysed to build an objectively evaluated approach that has the ability to provide an automated understanding of the infants' behaviour, underpin the infant specialists in analysing the infant attention for stimuli within different environments.

REFERENCES

[1] J. Ward. The Students Guide to Cognitive Neuroscience. Psychology Press, 2010.

[2] M. S. Sharif, M. H. Alsibai and E. Kushnerenko. Evaluating Bad and Good EEG Segments Based on Extracted Features: Towards an Automated Understanding of Infant Behavior and Attention. 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2017.

[3] Health Innovations. Artificial intelligence uses EEG Data to reconstruct images based on what participants perceive. 2019.

[4] Science Daily. Artificial intelligence to evaluate brain maturity of preterm infants. 2019.

[5] IBM. Machine Learning Treats Brain Disorders Where They Most Often Occur. IBM Research Blog, 2019.

[6] W.A.W Azlan, Y. F. Low. Feature extraction of electroencephalogram (EEG) signal-A review. 2014 IEEE Conf. on Biomedical Engineering and Sciences (IECBES), pp. 801-806, 2014.

[7] N. Huang. Introduction to the Hilbert Huang Transform and Its Related Mathematical Problems in Hilbert-Huang Transform and Its Applications. Singapore:World Scientific Publishing, Co. Pte. Ltd. pp. 1-26, 2005.

[8] P. Comon "Independent component analysis a new concept?" Signal Processing vol. 36 no. 3 pp. 287-314 1994.

[9] A. S. Al-Fahoum and A. A. Al-Fraihat. Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. ISRN Neuroscience, vol. 2014, Article ID 730218, 7 pages, 2014.

[10] S. Stober, A. Sternin, A. M. Owen and J. A. Grahn. Deep Feature Learning for EEG Recordings. CoRR abs/1511.04306 ,2015.

[11] S. Ahmed, L. Mauricio Merino, Z. Mao, J. Meng, K. Robbins and Y. Huang, "A Deep Learning method for classification of images RSVP

[12] Y. Gao, H. J. Lee, R. M. Mehmood, "Deep learning of EEG signals for emotion recognition", Multimedia & Expo Workshops (ICMEW) 2015 IEEE International Conference on, pp. 1-5, 2015.

[13] S. Chambon, V. Thorey, P. Arnal, E. Mignot, A. Gramfort. A deep learning architecture to detect events in EEG signals during sleep. IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2018), Sep 2018.

[14] S. Roy, I. Kiral-Kornek, and S. Harrer, "Deep Learning Enabled Automatic Abnormal EEG Identification," 39th Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018

[15] Y. Yuan, G. Xun, K. Jia and A. Zhang, "A Multi-View Deep Learning Framework for EEG Seizure Detection," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 1, pp. 83-94, Jan. 2019.

[16] A. Subasi, M. I. Gursoy. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Systems with Applications: An International Journal, v.37 n.12, p.8659-8666, December, 2010.

[17] F. La Foresta, F. C. Morabito, B. Azzerboni, M. Ipsale. PCA and ICA for the extraction of EEG components in cerebral death assessment. IEEE International Joint Conference on Neural Networks, vol.4, pp.2532-2537, July 2005.

[18] H. Wang, L. Bai, J. Xu, W. Fei, "EEG recognition through Time-varying Vector Autogressive Model" August 2015.

[19] P. A. Babu and K. V. S. V. R. Prasad. Removal of Ocular Artifacts from EEG Signals Using Adaptive Threshold PCA and Wavelet Transforms. IEEE International Conference on Communication Systems and Network Technologies (CSNT), pp.572-575, 3-5, 2011.

[20] L. Ke, R. Li. Classification of EEG signals by multi-scale filtering and PCA. IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS, vol1, pp 362-366, 2009.

[21] S. Jawed, H. Amin, A. Malik, I. Faye. EEG Visual and Non-Visual Learner Classification Using LSTM Recurrent Neural Networks. IEEE-EMBS Conference, 2018.

[22] W. Zhang, Q. Liu. Using the Center Loss Function to Improve Deep Learning Performance for EEG Signal Classification. Tenth International Conference on Advanced Computational Intelligence. 2018.

[23] P. Kaushik, A. Gupta, P. Roy, D. Dogra. EEG-Based Age and Gender Prediction Using Deep BLSTM-LSTM Network Model. IEEE Sensors Journal, Vol. 19, No. 7, 2019.

[24] B. Adhikari, A. Shrestha, S. Mishra, S. Singh, A. Timalsina. EEG based Directional Signal Classification using RNN Variants. IEEE International Conference on Computing, Communication and Security (ICCCS), 2018.

[25] S. Priyanka, D. Dema, T. Jayanthi. Feature selection and classification of Epilepsy from EEG signal. International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017.

[26] S. Puengdang, T. Sattabongkot, S. Tuarob, B. Sakboonyarat. EEG-based Person Authentication Method Using Deep Learning with Visual Stimulation. 11th International Conference on Knowledge and Smart Technology (KST), 2019.