DIAGNOSIS OF AUTISM IN CHILDREN USING DEEP NEURAL NETWORKS

N.V. Kousik
School of Computing Science and Engineering, Galgotia University, India

Abstract

The autism spectrum disorder is a common term for a group of complex brain and neurodevelopment disorder. The EEG medical imaging technique is a perfect tool for the brain signal analysis. In this study, we identify the variations in EEG signals on the auto-regressive features for classifying the normal and autistic features using Artificial Neural Networks. The simulation result shows that the proposed DNN in classifying the autism features achieves a classification rate of 95.23%.

Keywords:

Autism Disorder, Auto Regressive Features, Electroencephalography, ANN

1. INTRODUCTION

Autism is a childhood-related neurodevelopmental defect. Language deficiency and lack of social skills are a couple of major concerns that are related to autism. Kanner and Asperger explained at first that while the children appear slightly ordinary, they could not relate with others and their surroundings since the beginning of their lives on a regular basis. In all children with autism, repetitive behavior and strict interest were typical [1]. ASD is generally regarded as a life-long disability, without a single accurate laboratory examination and without any precise or curative treatment yet established [2] - [4].

Autism is reported to occur around 1% globally and yet, in comparison with other similar diseases, this disorder is given much less attention. The newly diagnosed ASDs are DSM-V criteria include three conditions: impairment of the social interaction, deformity of communication, and restricted repetitive and stereotypical patterns [6]. DSM-V is a diagnosis of the disease. Electroencephalography (EEG) is a system used to study brain activity and function by recording brain electrical activity. In order to capture the signals generated from the scalp surface, metal electrodes and conduction media are used [7].

EEG signals are generated by electrical voltages within the brain structure in human scalp. The electrical activities are responsible for both types of neuronal activity, namely action potential and post synaptic potential [5]. When there are series of fast electrochemical differences between the axon and axon terminals in the cellular body, potential for action is created. At the membrane receptors, postsynaptic potentials are observed. Neurotransmitters are available for the post-synaptic cell, which causes ion channels to open and close and cause the electrical potential across the membrane to be gradually shifted [8]. As a result, the activities in the brain of the patient can be used to be known in detail [9].

EEG is mostly non-invasive and cost-effective compared with other laboratory tests. DSM-V criteria for diagnosis of autism are currently used. It is a difficult task for doctors to manually measure the degree of a disorder by examining the patient and his reports, without the assistance of a tool. In the automated autism diagnosis using EEG signals, further innovation and development are therefore promoted globally.

For finding a biomarker of autism in babies, Bosl et al. [10] used multi-scale entropy (mMSE) and multiclass Support Vector Machine (SVM) algorithms. For control and ASD children, they were given an overall precision of 80%.

Seven EEGs from 17 ASD kids and 11 control kids in the age group of 6 to 11 years were assessed by Sheikhani et al. [11] assessed EEG signals using spectrograms and consistency values. Statistical analysis showed that the best differentiation level of 96.4% was found in the alpha frequency band in relaxed eye-open conditions.

Multilayer Perception Neural network (MLP) used by Shams et al. [12] for Principle Component Analysis (PCA). The autism signals were classified into two tasks, namely the open eye and engine movement tasks. Routing between autism and control objects has been found to be around 90-100% accurate. For normal subjects in open-eyed tasks 90% accuracy was found.

2. PROPOSED MODEL

In this paper, we develop a model that consists of data acquisition of EEG signal, feature extraction of autism behavior and classification using a machine learning model i.e. Artificial Neural Network (ANN) with Genetic Algorithm (GA).

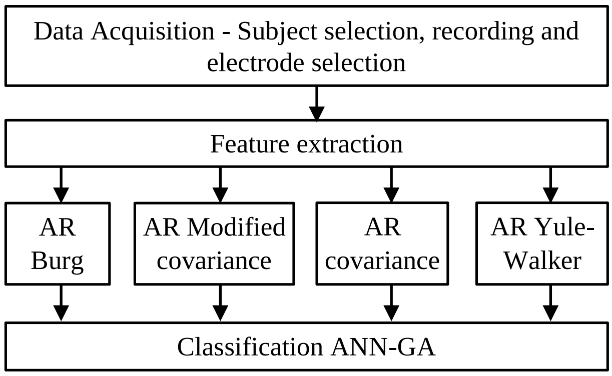

Fig.1. Architecture of Proposed Model

3. DATA ACQUISITION

The major components for data acquisition stage are:

- *Subjects*: Sample size included 4 (3 boys and 1 girls) normal children and 6 (4 boys and 2 girls) children with autism. Age group of 6-12 years was chosen. The normal children group consisted of children with no present or past neurological disorders.
- *Recording*: 4 tasks namely relax, flashcards read and spell, video read and spell and video hand movement imitation were used for the study. During the recording, band pass filter with frequency band (0.1-60) Hz was used to filter the

data and digitization took place at 256Hz. 50 Hz notch filter was also used.

- *Selection of Electrode*: Since the EEG signals of children are only taken into consideration, pediatric montage is chosen for the study. The pediatric montage consists of electrodes A1, A2, O1, O2, T3, T4, C3, C4, Fp1 and Fp2 according to 10-20 International System.

4. FEATURE EXTRACTION

The estimate of the parametric spectrum depends on the previous system information. The Auto Regressive (AR) method is commonly used as a parametric method. The AR method is used in the study where the signal coefficient is derived in a given instance, by inserting the sample's previous coefficient and adding the error estimate. p^{th} model order of Autoregressive (AR) process is given by

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + e(n) \tag{1}$$

where

a_k indicates AR coefficients,

p indicates the model order,

$x(n)$ represents EEG signal at the sampled point n and

$e(n)$ indicates the error term independent of previous samples.

Thus, four extraction features, such as AR Burg, AR Modified Covariance, AR Covariance and AR Yule Walker, have been employed to obtain the estimates of the AR coefficient.

- *AR Burg Method*: The least quadratic sense method is employed in this method to minimize forward and backward forecast errors by matching AR models to EEG indicators. High frequency resolution, stability and highly effective calculation are the main advantages of the AR Burg assessment. Without the interference of autocorrelation function, the AR burg method generates the reflection coefficient automatically.
- *AR Modified Covariance Method*: This method also uses less square sensory techniques to minimize forward and reverse prediction errors. By fitting the AR model to the EEG signals, the AR coefficients can be identified. In addition, the autocorrelation matrix window is not necessary for calculating this method. Spectral line-splitting is not involved, but it can produce unstable models.
- *AR Covariance Method*: This method identifies AR coefficients through the use of lowest squares sensory techniques to minimize future prediction errors. These coefficients are then adapted to EEG signals' AR models. The AR covariance estimate produces a greater resolution spectrum for short data records compared to the Yule-Walker AR estimate. The autocorrelation matrix window calculation is not required in this method.
- *AR Yule-Walker Method*: This method identifies AR coefficients through the use of lowest squares sensory techniques to minimize future prediction errors. These coefficients are then adapted to EEG signal. The AR covariance estimate produces a greater resolution spectrum for short data records compared to the Yule-Walker AR estimate. The autocorrelation matrix window calculation is not required in this method.

5. ANN CLASSIFICATION

Classification is one of the most dynamic fields of research and application for ANN. The major disadvantage when using ANN is that the training, learning and transmission functions for classifying the data sets with an increasing number of functionalities and classified sets are most suitable. The various combinations of functions and their effect are examined and the correctness is analysed for various types of datasets while using ANN as a classifier.

The Back Propagation Neural Network (BPNN) trains the neural network and the Gradient Descent Method (GDM) method was used for reducing the average squared error between the output of the network and the actual error rate. In order to measure the network efficiency, convergence rate, network convergence times, and the calculated mean square error (MSE), the following parameters are considered. The classification of the data set uses the most successful tool called the back propagation neural network in the appropriate combination of training, learning and transmission functions.

Fitness for the classification of remote images based on 3 steps is proposed by the BPNN. Initially, the characteristics are extracted from first-order histograms. The second stage is a BPNN-based characteristic classification, and in the third stage the results are compared with the MLC method. The statistical characteristics of this article depend on the distribution measure of the first order, for example, medium, Standard deviation, skew, kurtosis, energy and entropy. There are 3 layers in the network.

In the field of EEG signal classification, ANN is an important element of artificial intelligence. Due to the complexity of wetland areas, the wetland remote sensing classification based on ANN is complex. On the training samples is carried out the remote sensing image supervised classification. The clarity is investigated and it has been found that it is difficult to guarantee, as it will impact the results of the classification.

The nonlinear mapping function of the BP ANN results in better classification for complex areas. First, the theory of statistical analysis in training samples for the eradication of noise is used. In order to train the BP ANN individually, the original samples and purified samples were then used. They created two classification maps based on two BP ANN trained. The statistical method of analysis for the purification of BP ANN samples is carried out.

The study uses multi-dimensional datasets, which are important to classify data sets. The dataset is divided into training set and test set and is not used during training. These data sets are used to produce the results and to test them. The training set consists of 2/3 of the data set and the rest was tested. This is achieved by evaluating the accuracy achieved by testing these data sets. The network will then be simulated with the same information.

5.1 ANN FOR FEATURE EXTRACTION

BPNN trained with back propagation was used in this paper. The study used three topologies to classify stress. A GA was used for selecting stress characteristics for two ANNs that formed ANN inputs. As a result, only the number of inputs used differed from the three ANNs.

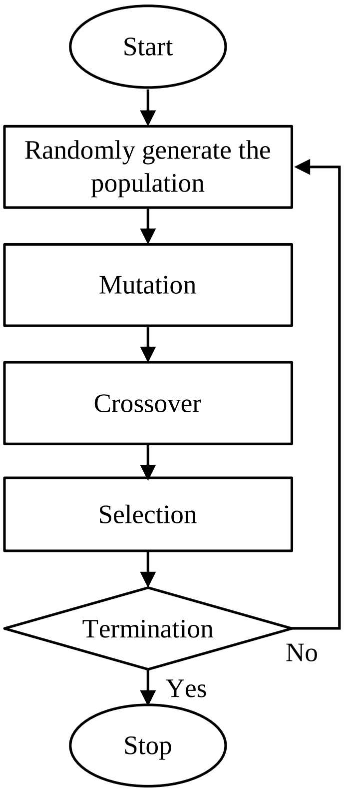

Fig.2. Genetic Algorithm

For each participant, a feed-forward ANN was generated. The ANN contains 215 inputs, which reflect the number of characteristics derived from the main stress signals. The ANN entries were similar but the number of entries depended on the functionality selected by a GA. The GA selected the classification rates for stressed and unstressed countries to be improved. ANN-3Seg-GAInputs: For each participant, an ANN was developed with a maximum of 10 features with 3 sequential time frames each as the ANN inputs. The GA selected the classification rates for stressed and unstressed countries to be improved.

In addition to the selection of relevant functions, GA was used to improve classification, but also to exclude corrupted features. For example, when the ECG sensors acquired ECG signals for the participant, the corresponding features were corrupted. This information is also visible through the observation of graphs of raw signals during data acquisition. The GA gives the ANN classification a greater chance of not using corrupted features to develop relationships from stress classification features.

One reason for choosing features less than the overall ANN GA feature was to decide whether fewer features can be used to represent the area with a smaller network successfully. Furthermore, if all features were inputs then the number of inputs would be 3 times higher than the number of ANN inputs. ANN GA input results show that the 215 functions for the ANN classification have not been required. A smaller subset of features can produce a more accurate classification using ANN GA inputs. This also led to features that are less important for ANN GA inputs than for inputs.

For each participant, the data sets have been divided into 3 subsets – training, validation and testing in which 50% of the data samples were used to train ANN and the rest of the data sets were equally divided to validate and test the ANN.

For implementing and testing ANNs, MATLAB was used. For the incremental training of ANN, the MATLAB adaptation function was employed. The Levenberg-Marquardt (LM) algorithm trained every network for 1000 epochs. In the output layer, the network had seven hidden neurons and one neuron. In future work, the ANN topology for autism classification in the reading dataset could be optimized.

ANN GA inputs and ANN GA inputs are used to select features using quality measures to determine stress in ANNs. A GA is an international search technique and has proved useful for problems with optimization. Due to a population of sub functions, the GAs have been developed to identify sets of features that produce a better ANN classification quality through the use of crossover, mutation and selection feature sets.

All features of the initial population for the GAs are created. The number of chromosome characteristics varied but the length of the chromosome was fixed. The chromosome length was equal to the number of characteristics in the space for the feature. A binary string was a chromosome, the index representing a feature for a bit and the bit value showed if the function is used in the ANN. The population of ANN with GA had characteristics that vary equally and the chromosomes had 5 characteristics with feature values for 3 consecutive time segments, on the other hand for ANN GA inputs.

6. RESULTS AND DISCUSSION

For every EEG signal, feature extraction provides 16 features. The input to the neural network is this value. To design an FFNN, seventeen input neurons, three output neurons, and ten hidden neurons are used. In order to choose the values for hidden neurons, the test and error method is used. Network testing and training are conducted with 75% and 100% of the data set respectively. The 0.001 error tolerance rate of training is fixed.

Table.1. AR-Burg Accuracy

Subjects	Hidden Neuron	AR-Burg	
		Standard Deviation	Average Accuracy (%)
1	10	2.41	90.53
2	15	2.36	89.86
3	20	2.28	92.41
4	25	2.31	91.26
5	30	2.04	93.25
6	35	2.44	93.36
7	40	2.19	92.77
8	45	2.38	93.07
9	50	2.23	94.87
10	55	1.89	93.28

Table.2. AR-Modified Covariance Accuracy

Subjects	Hidden Neuron	Standard Deviation	Average Accuracy (%)
1	10	2.81	90.17
2	15	2.97	89.54
3	20	2.78	92.10
4	25	2.46	91.03
5	30	2.50	93.04
6	35	2.36	93.25
7	40	2.60	92.51
8	45	2.51	92.63
9	50	1.95	93.48
10	55	2.06	93.16

Table.3. AR-Covariance Accuracy

Subjects	Hidden Neuron	Standard Deviation	Average Accuracy (%)
1	10	2.92	90.38
2	15	2.55	89.22
3	20	3.24	91.73
4	25	2.93	90.69
5	30	2.69	93.11
6	35	2.77	93.04
7	40	3.24	92.35
8	45	3.11	92.60
9	50	2.72	93.36
10	55	2.87	92.82

Table.4. AR-Yule Walker Accuracy

Subjects	Hidden Neuron	Standard Deviation	Average Accuracy (%)
1	10	2.77	89.98
2	15	2.41	89.13
3	20	3.02	91.56
4	25	2.50	90.42
5	30	3.04	92.70
6	35	2.69	92.61
7	40	3.05	91.83
8	45	3.00	92.26
9	50	2.76	93.13
10	55	2.69	92.61

The testing error tolerance rate is 0.6 and the accuracy of classification of the four AR functions, namely AR-Burg, AR-modified covariance, AR-covariance and AR-Yule Walker is shown in Table.1 – Table.4. Normal subjects are subject to 1 to 4, and autistic subjects are subject to 5 to 10.

The testing error tolerance rate is 0.6. The accuracy of classification of the four AR functions, namely AR-Burg, AR-modified covariance, AR-covariance and AR-Yule Walker is shown in Table.1 – Table.4. Normal subjects are subject to 1 to 4, and autistic subjects are subject to 5 to 10.

In children with autism, higher classification accuracy is found than ordinary children which show a greater distinction between neuronal behaviors in each task. In addition, the highest accuracy in classification is achieved with the AR-Burg feature while AR-Yule Walker obtains the lowest accuracy in classification. The best way to classify autistic children and children using FFNN is AR-Burg method.

7. CONCLUSION

The study utilized both ANN and AR-features to differentiate children with autism from normal children by using EEG signals. The study used 10 subjects, including 4 normal subjects and 6 autistic subjects. The AR-Burg function attains the maximum classification rate. Experimental data shows that the characteristic AR-Burg is comparatively better than the extraction methods of AR-covariance, AR-modified covariance and AR-Yule Walker. The results also show that children with autism have higher accuracy of classification than ordinary children. In order to practically implement the automated EEG-based autism diagnostic system, improved recognition rates must be achieved in the future.

REFERENCES

[1] A.R. Aslam, T. Iqbal and M. Aftab, “A10. 13uJ/Classification 2-Channel Deep Neural Network-based SoC for Emotion Detection of Autistic Children”, *Proceedings of IEEE International Conference on Custom Integrated Circuits* Conference, pp. 1-4, 2020.

[2] J.K.E. Wei, S.L. Oh, N. Arunkumar and E.W. Abdulhay, “Autism Spectrum Disorder Diagnostic System Using HOS Bispectrum with EEG Signals”, *International Journal of Environmental Research and Public Health*, Vol. 17, No. 3, pp. 971-988, 2020.

[3] L.E. Mash, B. Keehn, A.C. Linke and T.T. Liu, “Atypical Relationships between Spontaneous EEG and fMRI Activity in Autism”, *Brain Connectivity*, Vol. 10, No. 1, pp. 18-28, 2020.

[4] F.A. Alturki, K. AlSharabi, A.M. Abdurraqeeb and M. Aljalal, “EEG Signal Analysis for Diagnosing Neurological Disorders Using Discrete Wavelet Transform and Intelligent Techniques”, *Sensors*, Vol. 20, N0. 9, pp. 2505-2511, 2020.

[5] D. Haputhanthri and G. Brihadiswaran, “An EEG based Channel Optimized Classification Approach for Autism Spectrum Disorder”, *Proceedings of International Conference on Moratuwa Engineering* Research, pp. 123-128, 2019.

[6] D. Haputhanthri and G. Brihadiswaran, “EEG-based Processing and Classification Methodologies for Autism Spectrum Disorder: A Review”, *Journal of Computer Science*, Vol. 15, No. 8, pp. 1-14, 2019.

[7] A.A. Ramzy and B.S. Abdallah, “EEG-Based Screening Tool for School Student’s Brain Disorders Using Machine Learning Algorithms”, *International Journal of Biomedical*

and Biological Engineering, Vol. 13, No. 6, pp. 288-293, 2019.

[8] M. Hashemian and H. Pourghassem, "Decision-Level Fusion-Based Structure of Autism Diagnosis using Interpretation of EEG Signals Related to Facial Expression Modes", *Neurophysiology*, Vol. 49, No. 1, pp. 59-71, 2017.

[9] J. Kang, T. Zhou, J. Han and X. Li, "EEG based Multi-Feature Fusion Assessment for Autism", *Journal of Clinical Neuroscience*, Vol. 56, pp. 101-107, 2018.

[10] S. Ibrahim, R. Djemal and A. Alsuwailem, "Electroencephalography (EEG) Signal Processing for Epilepsy and Autism Spectrum Disorder Diagnosis", *Biocybernetics and Biomedical Engineering*, Vol. 38, No. 1, pp. 16-26, 2018.

[11] E.A. Alsaggaf and S.S. Baaisharah, "Directions of Autism Diagnosis by Electroencephalogram Based brain Computer Interface: A Review", *Life Science Journal*, Vol. 11, No. 6, pp.298-304, 2014.

[12] M. Ahmadlou, H. Adeli and A. Adeli, "Fuzzy Synchronization Likelihood-Wavelet Methodology for Diagnosis of Autism Spectrum Disorder", *Journal of Neuroscience Methods*, Vol. 211, No. 2, pp. 203-209, 2012.