Characterizing Autistic Disorder based on Principle Component Analysis

[1]Wafaa Khazal Shams, [2]Abdul Wahab

Dept of Computer Science, Faculty of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia.

Abstract: Autism is often diagnosed during preschool or toddled age. This diagnosis often depends on behavioral test. It is known that individuals with autism have abnormal brain signals different from typical persons yet this difference in signals is slight that it is often difficult to distinguish from the normal. However, Electroencephalogram (EEG) signals have a lot of information which reflect the behavior of brain functions which therefore captures the marker for autism, help to early diagnose and speed the treatment. This work investigates and compares classification process for autism in open-eyed tasks and motor movement by using Principle Component Analysis (PCA) for feature extracted in Time-frequency domain to reduce data dimension. The results show that the proposed method gives accuracy in the range 90-100% for autism and normal children in motor task and around 90% to detect normal in open-eyed tasks though difficult to detect autism in this task.

Key words: *EEG, PCA, Autism.*

INTRODUCTION

Autistic and Pervasive Development disorders are among the common brain development disorder in this decade. It is also known as Autistic Spectrum Disorder (ASD). Autism can be observed in the early stage of childhood, and characterized by impairment in social interaction shown by eyes contact avoidance, facial expression and difficulty to understand others, deficits in communication, poor in language interpretation, restricted and repetitive (American Psychiatric Association, 2000) These groups of individuals cannot interact with the environment which affects them negatively. Sensitive intervention from their parents is required and different therapy treatments must be provided to ensure they can interact with the community. Unfortunately the cost of therapy is high and is not affordable for most families hence they are not able to pay for the service for long.

Autism was first noticed by Kanner in 1943, as cited in (Bernier, 2007; Williams, 2004; Teitelbaum, 1998; Ozonoff, 2008) and reported that they have poor motor and imitation tasks and this has been established by many researches (Bernier, 2007; Williams, 2004; Vanvuchelen, 2007). Several researches have been done to explain the symptoms of autism. However, it is still unclear and debated among many. The main problem since decades is how to detect autism in early age, Social impairment appears around age 3 years which is the main symptom of the autistic disorders and can be diagnosed however nonsocial impairment could be more effective in detecting autism in earlier ages (American Psychiatric Association, 2000). However; it is difficult to detect autism in infants based on psychology test in motor movement (Teitelbaum, 1998; Ozonoff, 2008) while researchers revealed that autistic individual between the ages of 9 to 10 have deficits in motor task (Green, 2009). Even though it may be possible to detect poor motor movement for autism, it is not usually adopted in most cases to diagnose autism; most clinics depend on social defects rather than other behavior which are not shown in the early age. The human brain is known to control thoughts, emotion and movement, Electroencephalogram (EEG) has been known to be particularly suitable tool to study the brain signal and its functional behavior.

EEG signals have a lot of information which reflect the behavior of brain functions. However, these signals are non-stationary, non-linear, vary with time and represents stochastic system (Sanei, 2007) besides, it is affected by noise from electrode and child movement which cannot be controlled or restricted for long time, Thus, action detection during less than one minute is a challenge. Further, features extraction for complex systems is time dependent. Then, quantized EEG (QEEG) signal to time-frequency domain limit the complexity of the system by decomposition of the wave form. EEGs are multi-dimensional system. Features extraction maps data into samples with fewer dimensions. This is often considered as dimension reduction. EEG is multi-variety, cross correlated data, which records the potentials for enormous number of nerves; therefore orthogonal transformation in Principle Component Analysis (PCA) (Jolliffe, 2004) will be suitable to solve this problem.

EEG records the brain signals characterized by theta, beta, alpha, and gamma waves (Sanei, 2007). Activation of those waves are different according to activities of the nerves, therefore investigating the detection within limited bands may give good accuracy. This work investigates the application of Principle Component Analysis (PCA) - widely known in EEG studies to reduce dimension and detect features and it has shown high

Corresponding Author: Wafaa Khazal Shams, Dept of Computer Science, Faculty of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia.
E-mail: Wafaa_dth@yahoo.com

accuracy in classification -to detection of autistic child using open-eyes and motor tasks. The results of this work can be considered for further work to detect autism in early age.

Related Work:
The Electorencephalogram (Eeg):

Electroencephalogram (EEG) signals measure potential at the scalp of brain which contains the activity of millions of nerves. Electroencephalogram compared to other methods is inexpensive, completely safe for all ages, including infants (Sanei, 2007). EEG records the brain signal characterized by theta, beta, alpha, and gamma waves; recently, researchers consider using it in clinics to diagnose and detect neurological disease by detecting abnormal wave form of signal (Sanei, 2007; Adiel, H., 2010). However, EEG signal is a dynamic system that changes with time (Sanei, 2007). In the last two decades, many mathematical and statistical models have been proposed to extract features from EEG signals. Principle component analysis is one of the techniques that are widely used in EEG researches and it has shown to be effective and useful in brain computer interface, in particular and neurological disease studies (Sanei, 2007; Ke, L., 2009; Wolpaw, 2007).

Motor Impairment in Autism:

As mentioned above, many researches have been done to study the motor impairment for autism. In literature, so far research done in (Teitelbaum, 1998), try to detect autism during motor in early age with infants by analyzing video recording of infants' movements. Even though, the research findings mentioned that it could be possible to detect autism depending on movement but it has not been confirmed in infants. In (Ozonoff, 2008), it was found that defects in motor is shown clearly in children with Delay Disorders (DD) by analyzing infants' movements while there have not been clear signs for motor deficits for autism in infants. Researchers in (Green, 2009) confirmed in their study that autism within ages 7 to 9 years had definite movement impairments on the Movement Assessment Battery for Children (M-ABC). Others studies have considered brain activities; it is found that there is increase activity in cerebral region in brain and supplementary motor region during motor task (Allen, 2003) and high cerebella motor activation for autism in motor task(Mostofsky, 2009).

Recent Studies to Detect Autism Based on EEG:

The complex mental for individuals with autism is associated with abnormal EEG signals, capturing the early marker for autism in early age help to treat this case earlier to avoid further complication. In recent years, researches have paid more attention to using EEG for detection of individuals with autism among control subjects. In (Sheikani, 2007) attempt to detect autisms was by using the Time-frequency domain (Short Time Fourier Transform - STFT) of the EEG signals. The mean and variance values are computed within all bands and within limited bands (STFT-BW) and using ANOVA with K-Nearest Neighbour classification.

The approach gives a result of 81% accuracy suggesting that it is difficult to discriminate a significant marker for ASD in STFT while there is discriminate in STFT-BW. Further, (Ahmadlou, 2010) investigated about children with ASD, by using wavelet chaos theory and KFD to extract EEG signal features, the study focused on the difference between autistic child and normal in close-eye active, reporting an accuracy of 90%. Another study (Sudirman, 2010) depends on the visual task for autism and normal child and reveals that there is difference in alpha wave; which suggests that autistic child has disorder in visual from the earlier age. While (Bosl, 2011) investigated about finding biomarker for autism in infants, by computing multi-scale entropy (mMSE) of EEG signals in resting state as a feature vector and using multiclass Support Vector Machine (SVM) algorithm for classification of each age group from 6 to 24 months. Classification accuracy for boys was close to 100% at age 9 months and between 70% to 90% at ages 12 and 18 months. For girls, classification accuracy was the highest at age 6 months, but declines thereafter. Overall the accuracy was 80% for control and ASD in 9 months.

From the foregoing, the approach used in this work is to identify autistic child during motor and open-eyes tasks by applying Principle Components Analysis (PCA) to Short Time Fourier Transform (STFT) of Electroencephalogram (EEG) signals.

METHODS AND RESULTS

A. Description of Eeg Data Recording:

The data was collected from six autistic children from the National Autistic Society of Malaysia (NASOM), and six typical subjects from preschool, both group around (7 to 9) years old. The EEG signal was measured and recorded from 8 channels based on the EEG International (10-20) Standard System by using BIMEC EEG machine with sampling frequency of 250 hertz. The 8 channels represent C3, C4, F3, F4, T3, T4, P3, P4 with Cz as reference, these channels collected data from motor cortex, frontal and partial regions of brain shown below in Figure 1. The subjects were subjected to two task for experimentation. The first task contained motor movement where subject sit in 75 cm away from monitor screen and asked to follow the right and left hand

movement movie while the other task is done by asking the subject to sit in rest condition with open eyes looking on black screen. The time for both tasks was 1 min but in this work, 35 sec of the collected time is considered which showed less affected by noise. This work interests on detect autism in motor and open eyes actives.
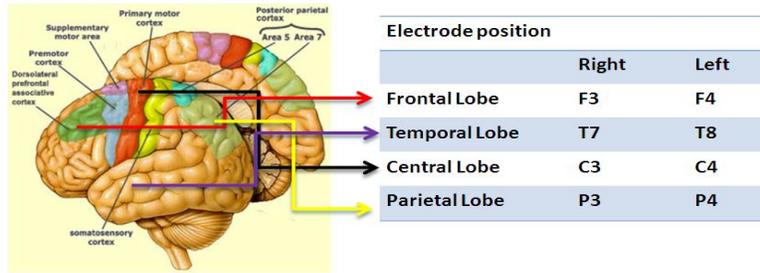

Electrode position		
	Right	**Left**
Frontal Lobe	F3	F4
Temporal Lobe	T7	T8
Central Lobe	C3	C4
Parietal Lobe	P3	P4

Fig. 1: The position of EEG electrodes (Brain diagram adapted from Buckeye psychaitry, LLC(Adam, 2009).

B. Data Processing:

EEG signal for each channel is normalised and processed to reduce the size and noise from the data. Down sampling was carried on the data to converting from 250 Hz to 83.3 Hz for all channels, and subtracting the mean values. Low quality EEG segments (artefacts) were manually excluded from the analysis. The EEG data was filtered to the interested bands of 0.5-30 Hz, using Butterworth band pass filter that assume to be sufficient as shown in figure 2. Then average filter is used to smooth the signal in figure 3. Similarly the average filter was done for 3 samples to reduce loss information.
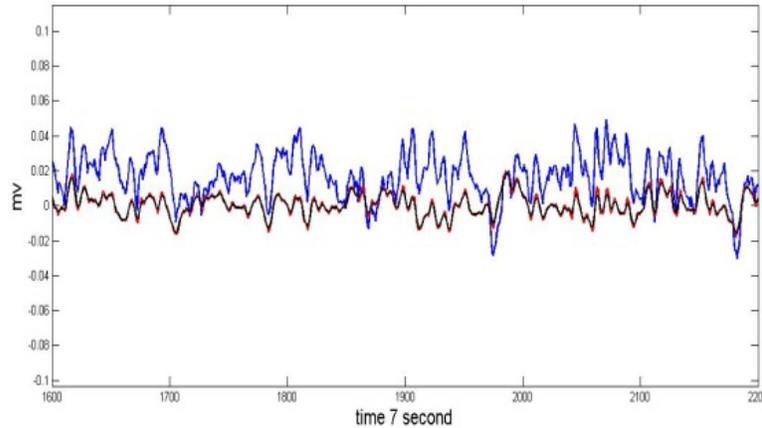

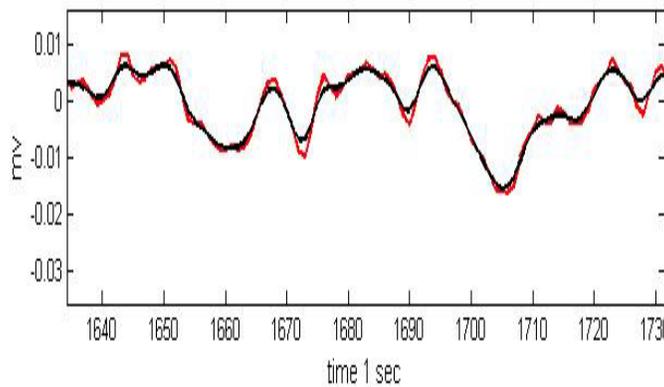**Fig. 2:** The original EEG signal (blue) and the filter signal(red).

Fig. 3: The effect of average filters to reduce noise and smooth the signal: original signal (red) and filtered signal (black).

C. Short Time Fourier Transform (Stft):

STFT is an appropriate technique useful to map the EEGs signal to its constituent frequencies. The limitation of STFT lies in the size of segment windows. Each filter signal is divided into finer band by using STFT, therefore feature extraction can be more efficiently done.

D. Feature Extraction By Principle Component Analysis:

Principle Component Analysis (PCA) is a widely used technique to reduce data dimension and to reconstruct data with less square errors. Its principle is based on rotating the data towards the longest variance and less residual error. Thus its projection represents the variance of the original data. Let B be two dimensional matrix (n, m) representing the time-frequency domain of the data after applying STFT, where m=0,1,2,....m-1 the time sample and n=0,1,2,3...n-1, the frequency sample.

The covariance matrix will be

$$C = \sum_{i=1}^{n}(B(i) - M(i))(B(i) - M(i))^T \tag{1}$$

Where M is the mean of each frequency sample with time (m).

Then performing Singular Value Decomposition (SVD) on the matrix C to get the Eigen values $\lambda(i = 1,2,3,....,N)$ and the corresponding Eigen vectors, $v(i = 1,2,3,...,N)$, are descending from the largest value to small one.

This then maps the spectrum matrix to new features domain, given as

$$PC = V * (B - M) \tag{2}$$

where PC represents the features coefficient and V denotes the Eigen vectors. Thus, the spectrum of EEG signal for both groups will distribute in principal component space. As shows in Figure 4, it can be seen that the data distribution has some few overlap.
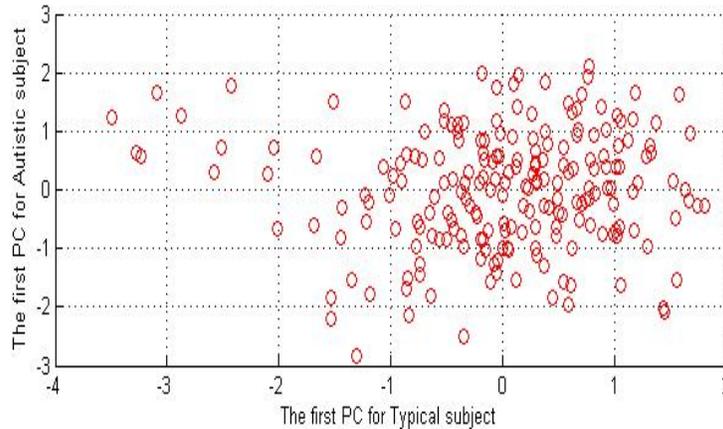

Fig. 4: The first PC for both Autistic and typical subject.

PC represents the feature space in (n, m) dimension, therefore to reduce features space, variances of PC were extracted. This will reduce the dimension of the feature space (n, m) for each channel. Further, all PC are used, considering that the low Eigen values also have important information. Thus, the feature space for each subject will be (R_m, nu) , where nu is the number of channels, R_m is the variance of PC features with time.

E. Classification Process:

For each group subject, features from PC subspace are computed from the transform matrix, for each channel. Then Neural network, Multilayer Perception, (MLP), is used to test the similarity and the difference of the features for two classes – autistic and typical subjects. The two groups are labeled as class 0 for autism and class 1 for typical. In the first step, both typical and autism groups are mixed together and then randomly divided for training and testing. The result is shown in Table I. In motor task, autistic feature is detected strongly and the accuracy is around 100% in MLP while typical group has 85.3% accuracy. In open eyes task, the ratio of detection in both autistic and normal is less than in motor task.

Table 1: The Accuracy Of Features Classification In Motor And Eyes Open Task.

	Motor task			Open eyes	
	AUTISM	*TYPICAL*		*AUTISM*	*TYPICAL*
AUTISM	100%	0%		86%	13%
TYPICAL	14.5%	85.3%		21.1%	78.9%
MLP classification					

Further classification was done to detect autistic and typical by testing all subject individually. This is done by putting all subjects in training test with excluding the tested ones.

The results are shown in Table II, and further illustrated in Figure 5, Figure 6, Figure 7 and Figure 8, for motor task and the accuracy was 100% for autism and 99% for normal. In contrast, the result is less in the open eyes task, it is around 90% to detect typical subject but its variance between (80% -40%) for autism. However its under the acceptance range of classification in typical subjects but it is ineffective for autism. For clinic purpose, it is necessary to get high accuracy.

Table 2: The Accuracy Of Detect Typical And Autism Subject.

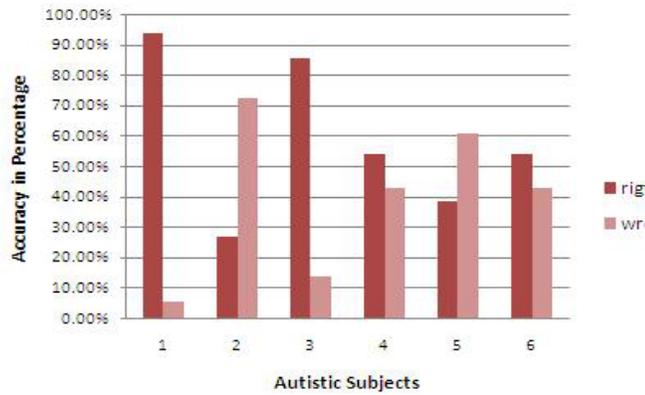| Open eyes		Motor		Autism	Typical
	Autism	*typical*		*Autism*	*Typical*
Autism(1)	94.3%	5.7%		100%	0%
Autism(2)	27.2%	72.8%		100%	0%
Autism(3)	85.7%	14.3%		100%	0%
Autism(4)	54.3%	43.3%		100%	0%
Autism(5)	38.7%	61.3%		79.5%	20.5%
Autism(6)	54.3%	43.3%		100%	0%
Typical(1)	10%	90%		0%	100%
Typical(2)	10%	90%		0.95%	99.05%
Typical (3)	8.1%	91.9%		0%	100%
Typical (4)	3.1%	96.9%		0.95%	99.05%
Typical (5)	8.1%	91.9%		1%	99%
Typical (6)	8.1%	91.9%		0%	100%

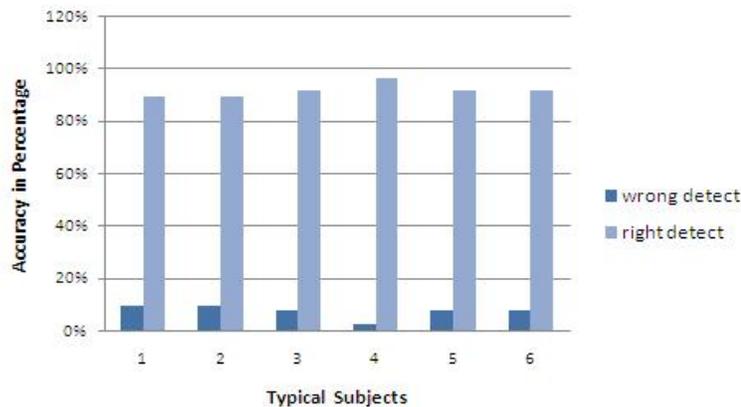**Fig. 5:** The right and wrong rate classification for autism group in open eyes task.

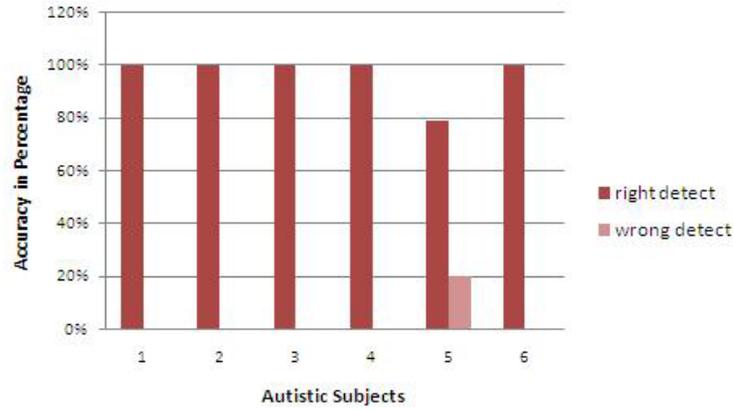**Fig. 6:** The right and wrong rate classification for typical group in open eyes task.

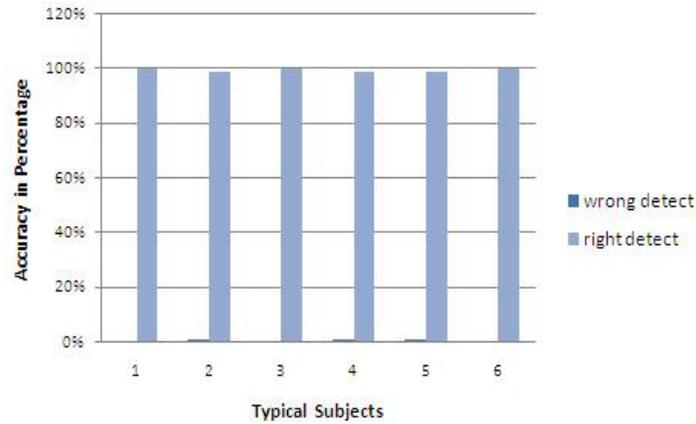**Fig. 7:** The right and wrong rate classification for autism group in motor task.

Fig. 8: The right and wrong rate classification for typical group in motor task.

Observation and Discussion:

Individuals with Autistic Disorder are known to have impairment in motor and motor imitation. This result of work shows that it is more practical to detect autism by depending on motor activities than in open eyes activities. This result suggests that in motor task, the motor cortex region of autistic subjects is active in different ways from normal subjects. Therefore, this can be useful in clinic by paying attention to the changes in EEG signal during motor task for early detection of autism. As mention in [13] and [14], the cerebrum region for autism is active in different way from typical one, regardless of whether it is less or high in activeness. The question that arise here is can this great changes in motor region for autistic individuals, be detected in less than age 7 using quantized EEG? The open eyes task proves more difficult to detect in both classes in rest condition. From six autistic subjects, two are detected as strongly autistic, two as correct autistic but near typical while two were detected wrongly as typical. Furthermore, extracted features from different brain regions such as frontal, cortex, temporal and partial within all band waves make the detection process more accurate. In most researches in autism that consider sub-band, this approach may be more effective in rest condition; but for detection purpose it is necessary to take all bands. Besides, capturing waveforms changes in different regions during task improve the features more significantly.

Principle Component Analysis (PCA) shows an effective method to detect feature from EEG system when it is applied to Time-Frequency domain. However, there is the inherent weakness when it is applied to open eyes activity. In general, QEEG shows good result to detect autism in motor task, which indicates that using QEEG for infants may show important marker for autism in motor movement. Besides, dysfunction in motor cortex regions could be responsible on social impairments for autism if the abnormal activities in cortex region showed the same changes in attention task.

Conclusion And Future Work:

In this paper, a new approach to diagnosis of autism is proposed. The proposed method applies time-frequency domain and principal component analysis (PCA) to extracted features from Electroencephalogram (EEG) signal for autism and typical subjects. A Multilayer Perception Neural network (MLP) method was used to detect the autism subject in two different tasks - motor and open eyes. The result shows that the *PCA* features

improve the classification accuracy. Moreover, its noticeable that autism subjects can be detected by motor more than open eyes activity. This recommends that future work should focus more on motor cortex region for autism and using sub – band analysis to open and close eyes actions.

REFERENCES

Adam, B., 2009. "http://buckeyepsych.wordpress.com/2009/05/09/cortical-brain-stimulation-may-improve-symptoms-of-intractable-depression-research-suggests/",.

Adiel, H., 2010. Wavelet-Chaos-Neural Network Models for EEG-ased Diagnosis of Neurological Disorders. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 26485 LNCS, pp: 1-11

Ahmadlou, M., H. Adeli, A. Adeli, 2010. Fractality and a Wavelet-Chaos-Neural Network Methodology for EEG-based Diagnosis of Autistic Spectrum Disorder. Journal of Clinical Neurophysiology, 27(5): 328-333.

Allen, G. and E. Courchesne, 2003. Differential effects of developmental cerebellar abnormality on cognitive and motor functions in the cerebellum: an fMRI study of autism. Am J Psychiatry; 160: 262-73.

American Psychiatric Association, 2000. Diagnose and statistical manual of mental Disorder DSM-IV-IR. Washington DC:American Psychiatric Association.

Bernier, R., G. Dawson, S. Webb and M. Murias, 2007. EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. Brain and Cognition, 64: 228-237.

Bosl, W., A. Tierney, H. Tager-Flusberg and C. Nelson, 2011. EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine*, 9: 18.

Green, D., T. Charman, A. Pickles, S. Chandler, T. Loucas, Simonoff, G. Braid, 2009. Impairment in movement skills of children with autistic spectrum disorders. Developmental medicine and Child Neurology, 52: 311-316.

Jolliffe, 2004. Principal Component Analysis. Springer Science, Inc., New York.

Ke, L., R. Li, 2009. Classification of EEG Signals by Multi-Scale Filtering and PCA. In the Proceedings - IEEE International Conference on Intelligent Computing and Intelligent Systems, art. no. 5357825, pp: 362-366.

Mostofsky, S.H., S.K. Powell, D.J. Simmonds, M.C. Goldberg, B. Caffo and J.J. Pekar, 2009. Decreased connectivity and cerebellar activity in autism during motor task performance. *Brain*, 132(9): 2413-2425.

Ozonoff, S., G.S. Young, S. Goldring, L. Greiss-Hess, A.M. Herrera, J. Steele, S. Macari, S.J. Rogers and Gross, 2008. Motor development, movement abnormalities, and early identification of autism**.** *Journal of Autism and evelopmental Disorders*, 38: 644-656.

Sanei, S. and Chambers, 2007. EEG signal Processing. John Wiley & Sons Ltd, England.

Sheikani, A., H. Behnam, M. Mohammadi, M. noroozian and P. Golabi, 2007. Analysis of quantitative Electroencephalogram background activity in autism disease patients with Lempel-Ziv complexity and Short time Fourier Transfoem measure. Proceedings of the 4th IEEE-EMBS International Summer School and Symposium on Medical Devices and Biosensors,art no. 4338305, pp: 111-114.

Sudirman, S., N. Saidin and Mat Safri, 2010. Study of electroencephalography signal of autism and down syndrome children using FFT. IEEE symposium on Industrial electronics and applications, penang, Malaysia, art., 5679434, pp: 401-406.

Teitelbaum, P., O. Teitelbaum, J. Nye, J. Fryman and R.G. Maurer, 998. Movement analysis in infancy may be useful for early diagnosis of autism. Proceedings of the National Academy of Sciences of the United State, 95: 13982-13987.

Vanvuchelen, M., H. Roeyers and W. Weerdt, 2007. Nature of motor imitation problems in school-aged males with autism: how congruent are the error types?. *Developmental Medicine & Child Neurology*, 49: 6-12.

Williams, J.H., A. Whiten and T. Singh, 2004. A systematic review of action imitation in autistic spectrum disorder", *Journal of Autism and Developmental Disorders.,* 34: 285-299.

Wolpaw, R.J., 2007. Brain-computer interfaces for communication and control. Proceding of the 9th international ACM SIGACCESS conference on computers and accessibility.