Study On Early Detection Of Autism Using Genetic And Kinematic Biomarker

Deepthi V¹ and Vineetha S²
¹ Student of Computer Science &Engineering, RIT, Kottayam.
² Faculty of Computer Science &Engineering, RIT, Kottayam.

E-mail: deepthithulasy@gmail.com

Abstract. Over these years prevalence of autism has shown a drastic increase. A great deal of progress has been made in recognizing autism but still a lot to do. The impact of autism on the individual differ due to difference in the degree of severity and the wide range of symptoms. Early detection can in turn ameliorate overall development of child and help to shorten the diagnostic odyssey that many families experience. One of the great things about machine learning is the capability of computers to work on the actual data to find the complex relationships through different algorithms which in turn result in accurate predictions. Thus by examining the various risk factors that contribute to autism, machine learning approach seems to be beneficial for identifying the potential markers for autism. This study covers machine learning approaches that are used for identification of autism which facilitate early detection and better diagnosis of autistic individuals using different biomarkers.

1. Introduction

Autism seems to be a neurodevelopmental disorder that become visible during the early childhood. It is a wide-spectrum disorder, which indicates that its degree of severity and symptoms may vary from one individual to another. With the numbers increasing annually, the Centre for Disease Control found out that one in 68 has been diagnosed with an autism spectrum disorder.

Real cause of autism cannot be confined to a single known symptom. Even though the symptoms vary but all reveal the deficit in social communication, motor- impairments and restricted behaviour or interest [1]. Presently there is no medical test for autism diagnosis. Doctors diagnose ASD by looking at a child’s etiquette and progress. Young children with ASD can reliably diagnosed by age two. Assessment of adults and children for ASD is based on the behavioural variations in communication, social interactions as reported by parents or teachers.

Detection of autism in adults is a cumbersome procedure because in adults, many symptoms can blend with some of other mental health, motor impairment disorders thus misinterpretation of actual diseases can in turn result in a dreadful life without proper diagnosis and effective treatment mechanisms. However, providing effective and early diagnosis of autism could be useful for overcoming
the difficulties that would in turn persist throughout life of an autistic person and to identify their strengths for leading normal life.

Researchers found that different arguments ponder regarding the cause of autism but a unique consensus which suggest that early intervention can in turn result in better outcomes. Instead of focusing on the scores from the behavioural exams and questionnaires for diagnosis of autism, researcher’s interest lies on using different machine learning approaches which distinguishes autistic individuals from others.

Machine learning is a powerful computational tool that support different application domains by learning complex relationships or patterns from large datasets to make accurate conclusions [10]. Better disease assessment can be made with predictive analysis of health data and offering more appropriate treatment mechanisms which is a hot research area nowadays. Supervised learning is an important part of machine learning [11] which employs a rule based approach by examining empirical datasets for building accurate predictive models for example patients chance of having a disease based on risk factors like genetic information. Movement reflects the way people interact with and sense their environment. Thus by focusing on unique ways in which the autistic brain processes different movements, could be helpful for diagnosis. It was found that untreated motor impairments can strongly affect the overall development of autistic individuals [3].

Only few studies put in machine learning to identify autistic individuals using kinematic data. A proof of concept study called reach to drop task [2] was conducted based on the upper limb movement analysis. It was found that a goal orientated movement can be pivotal in distinguishing autistic individuals. Many studies revealed strong evidence of genetic basis of autism which showed how different gene mutations contribute to the risk of autism. Various twin studies [4, 5] analogized the concordance rate which was found to be higher in monozygotic twins. Such studies provide a clear understanding regarding influence of genetic factors as an indicator for autism detection. Better treatment can be imparted by understanding the changes in the biological pathway and the abnormal proteins produced.

This paper presents a survey on machine learning approach that uses motion and genetic factors as biomarkers for autism identification. Thus early intervention of autism using machine learning technique opens up a new way for autistic individuals to develop a potential for leading a better life by improving their behavioural and emotional skills.

2. Kinematic Indicators
In the past decades, more consistent and troubling, features of autism, such as social problems and difficulties with communication were explored for diagnosis of autism. From 1980s, standardized tests for motor skills were began to be examined. The motion impairments can vary from one individual to other with different degree of severity. Most people with autism were found to lack effective sensory motor responses such as movement coordination towards achieving a certain goal, unstable balance and unusual gait patterns [6, 7].

It was perceived that children with adaptive behavioural skills are those with good motor skills. Thus the motor impairments are highly correlated with practical difficulties in habitual life skills like writing, walking etc. [8, 9]. Untreated motor problems can aggravate the impact of psychological and behavioural problems of autistic individuals. Motor abnormalities are found to be an important endophenotype for autism thus contributing to remarkable exploration in appropriate treatment mechanisms [12].
Proper characterization of kinematic impairments that gathers both spatial and temporal nature of movements could result in understanding the aetiology of motor impairments and contribute to diagnostic classification of ASD [13]. Here kinematic parameters of imitation of hand movement were taken for extracting the most desired features using machine learning which may classify autistic individuals and prove validity of motor bio signatures for preliminary diagnosis. Their result included two imitation condition and nine motion features extraction with a classification accuracy of 93%.

Movement ABC [14] was one among the frequently used evaluation method for many previous work in which four types of item assessment was done such as ball skill, static and dynamic balance etc. This method serves as a gauge in the recognition of motor impairments in children but motion patterns are poorly designated without considering the spatial temporal nature of movements. Poor motor coordination [9] can often limit the self-perception of an individual in their capabilities and handling of their environment and situation. However this lack of perceived competence can lead to anxiety and low self-worth.

Facial expressions [16] of children were recorded by direct motion capture and for these expressions pairwise causal dependencies between facial regions were studied. The complexity inspection of facial expressions were carried out using multiple scale entropy method. Six facial conditions were tested and found that highest complexity is observed in cheek regions for autistic children than in eye region for emotions like anger, joy, sadness and surprise. Dynamic time wrapping method and local complexity based similarity analysis showed that the major facial expression differences come from eye region in autistic individuals.

A number of studies applied the machine learning techniques for differentiating individuals with autism from others. Multivariate pattern analysis (MVPA) [19] conducted using two different fMRI experiments with social stimuli where support vector machines and recursive feature elimination method try to find out the neural markers for autism diagnosis. Machine learning algorithms were employed for extracting relevant information from Autism Diagnostic Observation Schedule-Generic (ADOS) [20, 21]. Here behavioural biomarkers were explored for autism detection. Stahl et al. [15] examined the infant event related potential and discussed the disadvantages of the commonly used methods such as mean group comparison. Discriminant function analysis, regularized function analysis and support vector machines were employed that avoid the multiple testing problem and errors of group mean comparisons. These classification measures were capable of increasing the distinct power of event related potential measurements. Here the purpose of machine learning methods to discriminate infant groups with high or low risk for detection of autism were discussed.

In Wilson et al. [17] cognitive skills were evaluated for discriminating autistic individuals deviating from behavioural diagnosis. Cognitive abilities are those voluntary responses of a person such as perception, motor skills, attention which contribute to a bank of knowledge based on combining the extracted information. Different range of cognitive skills were sampled. Tests valuating the symptom and intellectual capabilities along with neuropsychological tests were carried out. Support vector machines were used for classification using these twelve cognitive scores. Five cognitive task offers a significant difference and the important thing was that motor task fall under this category. However the cognitive classification models along with exploration of different empirical datasets were identified to be vital for autism diagnosis.

With the help of data from Autism Diagnostic Interview-Revised (ADI-R) [18] and Social Responsiveness Scale (SRS) an attempt was made to design more effective algorithms using machine learning with support vector machine was proposed to classify masses of verbal individuals with and without ASD. Greedy forward feature selection method provides information about most critical codes.
utilised for evolving a screener algorithm and provides an insight for improving the screening and diagnosis tools for ASD.

Artificial neural networks are adaptive systems influenced by the information processing mechanism of human brain. Neural networks are well-adapted to identifying non-linear patterns, in which there isn’t a direct, one-to-one relationship between the input and the output and could be employed in many application domains. Hence it can form a predictive model with high noise tolerance with high performance. Grossi et al. [22] assessed the risk factors during pregnancy in autism development. The frequency of potential risk factors in autistic and control groups were compared. This work demonstrate the deployment of artificial neural networks predictive capacity that differentiate different risk factors with an accuracy of 80% and could be used as diagnosis tool.

Artificial neural networks usage for analysing the effect of treatment in autism spectrum disorder was proposed in the work of Narzisi et al. [23]. The Auto Contractive Map (Auto-CM) a kind of ANN with a three-layer architecture. It include an Input layer, with inputs captured from the environment, a Hidden layer in which modulation is carried out, and an Output layer, which make use of a feeds back mechanism based on the previously received and processed stimuli, here the association among variables were created as semantic connectivity map. ADOS computed severity score (ADOS-CSS) which distinguished children who actively respond to treatments versus non responders. ADOS-CSS indicate the extremity of autism symptoms.

Table 2.1: Summarizing the impact of motion parameters for autism identification.

| Selection Criteria | Importance |
|--------------------|------------|
| Kinematic factors  | Highly Quantifiable and measurable |
|                    | Motor-abnormalities act as endophenotype for autism, exploring better treatment mechanisms. |
|                    | Better characterization of motor functions were critical for improving the behavioural and functional characteristics. |
|                    | Provides a method for examining the variations in neural circuits leading to autism |
|                    | Predicting the severity by analysing the repetitive behaviour of autistic individuals. |

3. Genetic Indicators

The gene susceptibility to autism has shown a remarkable development with the ongoing researches. Several genomic studies showed that many genes and chromosomal locations were associated to ASD. Advanced researches with new methods identified some significant mutations and the chromosomal instability which were analysed as an important risk factors for ASD. Here the main focus is to find whether the genetic factors provide the basis for autism biomarkers in machine learning scenarios.

One way of knowledge about the genetic influence in autism is through twin studies [4,5]. Researchers compared twin pairs both identical and fraternal pairs to see if the pattern of a disease or trait’s inheritance differs between them. Twin studies illustrated the phenotypic variation due to genetic factors to be about 90%.

Family studies inspected the siblings of individuals known to have autism. Family history data analysis [24, 25] showed that the some of the problems experienced by the relatives have a close resemblance with behavioural characteristic commonly found in autistic persons. Reduced intrafamily variance provides an evidence related to hereditary influence in autism. Multiple predisposing factors
were involved in autism which resulted in the variation of severity of disease and the phenotype [32]. The large degree of convergence from the causal factors obstructs the identification of predisposing factors and their development pathways.

The variation in the gene expressions were evaluated from blood samples collected which were used to differentiate between ASD individuals and unaffected controls. Glatt et. al [29] proposed a transcriptome-wide biomarker classifiers with high accuracy from the blood samples of different children having or not having autism. Gene expression profiling of mRNA samples collected from blood were examined by microarray to indicate that the abnormalities observed in these expression profiles mark the probability of autism occurrence. Another work focused on a class prediction algorithm [30] applied to find genetic mutations in lymphoblastoid cell lines (LCL). A rule based machine learning methods captured the relevant mutations that may result in autism by identifying the significant variation in genetic expression. ADI-R questionnaire scores were obtained by microarray analysis. Prediction was based on observed gene mutations with an accuracy of 94%. This work suggests the genes may be useful biomarkers for diagnosis of subtypes of idiopathic autism.

Kong et. al [26] conducted a largest transcriptome study that distinguished autistic individuals from the age/sex-matched controls. Repeated leave-group out cross-validation (LGOCV) strategy was used to build the predictive model. It was found that a favourable comparison occurs with DNA-based tests currently proposed for ASD diagnosis and ASD55 predictor’s classification ability. Age of the individuals can significantly influence gene expression but ASD55 genes were not significantly correlated with age except two. Fisher's exact test was used to find the biological pathway of differentially expressed genes. The outcome indicated feasibility of blood based transcriptomic approach for ASD detection and hence it serves as a diagnosis tool.

In toddlers early identification of individuals at risk for autism spectrum disorder (ASD) was difficulty due to the impact of overlapping of symptoms with other disorders along with genetic diversity. Pramparo et. al [27] analysed illumina microarrays which included the leucocytes collected from different sets of children below four years. Differentially coexpressed genes were identified and the symptomatic categorization of children with ASD were obtained with 83% accuracy. Logical regression and curve analysis evaluated the classifiers performance. It showed the possibility of using blood based clinical test in pediatric diagnostic settings and the use of genetic biomarkers in machine learning.

Genetic cause [31] including rare de novo and inherited mutations or chromosomal rearrangements were investigated by conducting whole-exome sequencing (WES) that is a genomic technique for sequencing all of the protein-coding genes in a genome and a transcriptomic approach. This strategy identified the rarely inherited mutations that result in greater probability autism emergence and monogenetic forms of ASD. Dong Hoon [28] employed the transcriptomic approach which identifies the gene expressions suitable for diagnostic prediction of young adults with ASD. Both supervised and unsupervised machine learning methods were taken in to consideration. Supervised learning used three different machine earning algorithms, such as the support vector machine (SVM) K-nearest neighbours (KNN) and linear discriminant analysis (LDA) and other by clustering. Thus the results of this and four previous studies suggest that gene expression profiles in the peripheral blood samples serves as biomarker to predict the ASD risk in all type of age groups.

Voineagu et. al [33] distinguished the mRNA molecules expressed from the genes in autistic and normal brain by gene co-expression network analysis and revealed molecular abnormalities in autistic individuals. The experiments were carried out on post-mortem brain tissue samples from 19 autism cases and 17 controls and found that differentially expressed genes of autistic individuals concentrated more
in cerebral cortex. Gene expression changes related to many diseases and hereditary characteristics were modelled using co-expression network.

Chromosomal micro array [34] is more powerful detecting capability than G-banded karyotype and fragile X DNA testing in patients with ASD. The benefits [35] of array comparative genomic hybridization which is a strong disease gene identification mechanism not only pave way for effective diagnosis but also helpful in the determining ASD genes with chromosomal imbalances. Hence this array offers an important aspect for genetic analysis. An attempt to identify the structural variations in ASD using both high-resolution microarray technologies and karyotyping was suggested and potential contribution of gene variation in ASD were identified [36]. This also supported the use of microarray analysis for clinical characterization in the detection of ASD.

Reviews on clinical and research genetics [37] gathered all genes and recurrent genomic imbalances which can become a cause of ASD. This work found that autism is a behavioural demonstration several genes and genomic disorders from the microarray analysis. Whole exome and genome sequencing act as genetic signatures for ASD identification.

Figure 3.1: Importance of genetic factors for autism identification.

The advancement in autistic detection for improved treatment and diagnosis can be made using the genetic biomarkers. Thus the machine learning methods using genetic indicators result in early detection of autism which reduces the stressful diagnostic odyssey.

4. Summary
However it was shown that both genetic and motion parameters are relevant in their own aspect for autism detection. Thus the benefits offered by these two biomarkers can be combined to facilitate better diagnosis and effective treatment mechanisms through a machine learning based approach. As the neurology of autism on a genetic level may be complicated due to the heterogeneity of motor deficits, the kinematic factors which involve the hand movements is helpful and a blood based gene expression analysis would be helpful to find the mutated genes that offer the risk of autism. This combination pave way in early intervention of autism than those studies which involve these biomarkers independently.
5. Conclusion
In this paper a brief review of different machine learning approaches employed in identifying autistic individuals are presented. Early intervention of autism reduces the diagnosis stresses faced by many families and appropriate treatment mechanisms as early as possible which in turn helps the autistic individuals to lead a normal life in the society. Motor impairments and the genetic factors have strong influence in autistic individual. Hence they found to be effective as a diagnosis biomarkers for autism detection using machine learning algorithms with better accuracy and computation speed. Heterogeneity of symptoms along with the lack of knowledge about gene susceptibility paved way for researchers to explore new domains for autism identification. Thus machine learning prediction by analyzing behavioural and hereditary features provide an aid in developing a better diagnostic and screening tool for ASD.

6. References
[1] American Psychiatric Association, Diagnostic and statistical manual of mental disorders (5th Ed.), Arlington, 2013, American Psychiatric Publishing.
[2] Crippa A, Salvatore C, Perego P, Forti S, Nobile M et al., 2015, Use of machine learning to identify children with Autism and their motor abnormalities, Journal of autism and developmental disorders.
[3] Fournier KA, Hass CJ, Naik SK, Lodha N, Cauraugh JH, Motor coordination in Autism Spectrum Disorders: A synthesis and meta-analysis, Journal of autism and developmental disorders.
[4] Steffenburg S, Gillberg C, Hellgren L, Andersson L, Gillberg IC, et al., 1989, A twin study of Autism in Denmark, Finland, Iceland, Norway and Sweden, Child Psychol Psychiatry 30: 405–416.
[5] Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, et al., 1995, Autism as a strongly genetic disorder: evidence from a British twin study, J Psychol Med 25: 63–77.
[6] Gowen E, Hamilton A, 2013, Motor abilities in autism: A review using a computational context, Journal of autism and developmental disorders.
[7] Bhat AN, Landa RJ, Galloway JCC, 2011, Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders, Physical Therapy.
[8] Jasmin E, Couture M, McKinley P, Reid G et al., 2009, Sensory-motor and daily living skills of preschool children with autism spectrum disorders, Journal of autism and developmental disorders.
[9] MacDonald M, Lord C, Ulrich D, 2013, The relationship of motor skills and adaptive behavior skills in young children with autism spectrum disorders, Research in autism spectrum disorders.
[10] Mitchell TM, 1997, Machine learning (McGraw-Hill Boston).
[11] Kotsiantis SB, Zaharakis I, Pintelas P, 2007, Supervised machine learning: A review of classification techniques.
[12] Esposito G, Pasca SP, 2013, Motor abnormalities as a putative endophenotype for Autism Spectrum Disorders, Front Integr Neurosci.
[13] Baihua Li1, Arjun Sharma, James Meng et al., 2016, Applying machine learning to identify autistic adults using imitation: An exploratory study.
[14] Henderson SE, Sugden DA, Barnett AL, Smits-Engelsman C, 1992, Movement assessment battery for children, Psychological Corporation London.
[15] Stahl D, Pickles A, Elsabbagh M, Johnson MH, Team B et al., 2012, Novel machine learning methods for ERP analysis: a validation from research on infants at risk for autism, Developmental neuropsychology.
[16] Guha T, Yang Z, Ramakrishna A, Grossman RB, Hedley D, Lee S, et al., On quantifying facial expression-related atypicality of children with autism spectrum disorder, IEEE Int Conf Acoustics, Speech and Signal Processing (ICASSP).
[17] Wilson CE, Happe F, Wheelwright SJ, Ecker C, Lombardo MV, Johnston P et al., 2014, The Neuropsychology of Male Adults With High-Functioning Autism or Asperger Syndrome, Autism Research.
[18] Bone D, Bishop S, Black MP, Goodwin MS, Lord C, Narayanan SS, 2016, Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion, *Journal of Child Psychology and Psychiatry*.

[19] Chanel G, Pichon S, Conly L, Berthoz S, Chevallier C, Grèzes J, 2016, Classification of autistic individuals and controls using cross-task characterization of fMRI activity, *NeuroImage: Clinical*.

[20] Kosmicki J, Sochat V, Duda M, Wall D, 2015, Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning, *Translational psychiatry*.

[21] Wall D, Kosmicki J, Deluca T, Harstad E, Fusaro V, 2012, Use of machine learning to shorten observation based screening and diagnosis of autism, *Translational psychiatry*.

[22] Grossi E, Veggo F, Narzisi A, Compare A, Muratori F, 2016, Pregnancy risk factors in autism: a pilot study with artificial neural networks, *Pediatric research*.

[23] Narzisi A, Muratori F, Buscema M, Calderoni S, Grossi E, 2015, Outcome predictors in autism spectrum disorders preschoolers undergoing treatment as usual: insights from an observational study using artificial neural networks, *Neuropsychiatric disease and treatment*.

[24] Bolton P, Macdonald H, Pickles A, Rios P, Goode S, et al, 1994, A case-control family history study of autism, *J Child Psychol Psychiatry 35*: 877–900.

[25] Bailey A, Palferman S, Heavey L, Le Couteur A, 1998, Autism: the phenotype in relatives, *J Autism Dev Disord 28*: 369–392.

[26] Kong SW, Collins CD, Shimizu-Motohashi Y, Holm IA, Campbell MG, Lee IH, et al, 2012, Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders, *PLoS One*.

[27] Pramparo T, Pierce K, Lombardo MV, Carter Barnes C, Marinero S, Ahrens-Barbeau C, et al, 2015, Prediction of autism by translation and immune/inflammation coexpressed genes in toddlers from pediatric community practices, *JAMA Psychiatry*.

[28] Dong Hoon Oh, Bin Kim, Seok Hyeon et al, 2017, Predicting Autism Spectrum Disorder Using Blood-based Gene Expression Signatures and Machine Learning.

[29] Glatt SJ, Tsuang MT, Winn M, Chandler SD, Collins M, Lopez L, et al, 2012, Blood-based gene expression signatures of infants and toddlers with autism. *J Am Acad Child Adolesc Psychiatry*.

[30] Hu VW, Lai Y, 2013, Developing a predictive gene classifier for autism spectrum disorders based upon differential gene expression profiles of phenotypic subgroups. *N Am J Med Sci (Boston)*.

[31] Codina-Solà M, Rodríguez-Santiago B, Homs A, Santoyo J, Rigau M, Aznar-Lain G, et al, 2015, Integrated analysis of wholeexome sequencing and transcriptome profiling in males with autism spectrum disorders, *Mol Autism*.

[32] Belmonte MK, Cook EH, Jr., Anderson GM, Rubenstein JL, Greenough WT, et al, 2004, Autism as a disorder of neural information processing: directions for research and targets for therapy.

[33] Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, et al, 2011, Transcriptomic analysis of autistic brain reveals convergent molecular pathology.

[34] Shen Y, Dies KA, Holm IA, Bridgemoohan C, Sobeih MM, et al, 2010, Clinical genetic testing for patients with autism spectrum disorders.

[35] Jacquemont ML, Sanlaville D, et al, 2006, Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders.

[36] Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, et al, 2008, Structural variation of chromosomes in autism spectrum disorder.

[37] Betancur C, 2010, Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting.