Upper Limb Motor Coordination based Early Diagnosis in High Risk Subjects for Autism

Mohammad Wedyan
Faculty of Engineering and Information Technology
University of Technology Sydney
Sydney, Australia
MohammadOmar.M.Wedyan@student.uts.edu.au

Adel Al-Jumaily
Faculty of Engineering and Information Technology
University of Technology Sydney
Sydney, Australia
Adel.Al-Jumaily@uts.edu.au

***Abstract*—Autism is a lifelong condition present from early childhood. Medical specialists' diagnosis autism based on observation is of great difficulty in communicating, difficulties for forming relationships with other people, and delayed speech. The scientists tried to discover other early signs to reach the early detection of Autism Spectrum Disorders (ASD). Early diagnosing is very important to initiate and improve treatment results. One of these signs is based on examination of upper limb motor movements. This study aims to determine whether a simple upper limb motor movement could be useful to classify High Risk (HR) infants for autism and comparison infants with Low Risk (LR) for autism. Also, this paper presents a computational intelligence method that uses HR and LR subjects between the ages of 12 and 36 months to make an early autism diagnosing. The paper examined one task which asks to insert an object into a box. It analyzed the data by using Support Vector Machine (SVM) and Extreme Learning Machine (ELM). The results show engorging results in comparison to other state or art methods.**

Keywords—Spectrum Disorders Children; Upper Limb Motor;Autism; SVM; ELM.

I. INTRODUCTION

Autism is an organic brain syndrome of childhood shows a spectrum of developmental disability[1]. It is a lifelong neuro developmental condition, with difficult ability of the person to communicate and relate to others [2]. The term 'autism spectrum disorders' (ASDs) covers conditions such as autism, childhood disintegrative disorder, and Asperger syndrome [3].

Many statistics shows that autism prevalence is rising, for example, according to the National Institute of Mental Health (NIMH), in the 1970's the ASD rate was 1 in 10,000. In 1995 the rate increased to 1 in 1,000. In 1999 the rate was 1 in 500. In 2001, the rate was 1 in 250. In 2005, the rate was 1 in 166. In 2007, the rate was revised to 1 in 150. In 2009, the rate rise to 1 in 90 and in 2012 the rate increased to 1 in 68[4, 5]. Fig.1 shows how autism diagnosis have climbed steadily since 1970s that displays the previous statistic.

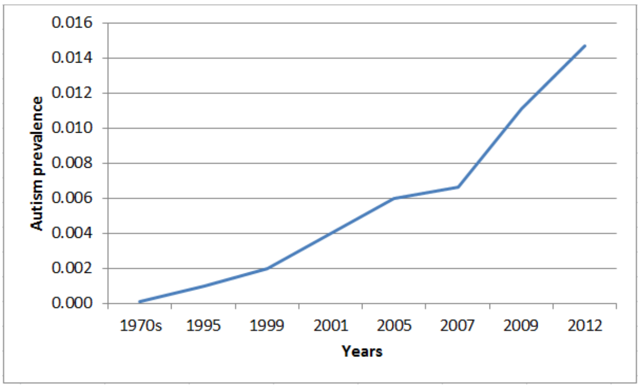

Fig. 1. Shows some counts, autism diagnosis have climbed steadily since the 1970s

Autism has lifetime negative effects, that impact on the health, happiness, education, social integration and goodness of life of a person, families, and society. Many of those impacts are economic that is different from society or country to another [6-14].

There are three classes of autism costs; the first class is a direct cost, for example; health care, social services, education; the second class Indirect tangible costs, for instance; reducing in income from lost employment; the third class of cost is informal intangible cost for example; care for adults with autism child or infant [15-18].

Autism is difficult to diagnose by medical doctors with support from physical, occupational and speech therapists since there is no medical test. Doctors look at the child's behavior and development to make a diagnosis [19-21]. The diagnosis of autism is based on the child's failure to develop good social interaction, impairments in communication, poor eye contact, delayed speech restricted repetitive and stereotyped pattern [1, 19, 22-26].

Scientist found that many autistic children have atypical motor patterns [27-30]. such as asymmetric gait or toe walking [31].The defect of gross motor function in kids with autism are well known to clinicians [32-36] but have not received much experimental documentation with the exception of repetition, are not among its diagnostic criteria [34-36].

Human motion was divided by specialist into different levels. For example, researchers divided human motion into “action” and “activity” that are frequently used reciprocally, but what is the different between these terms by “actions”. The study refer to simple motion patterns usually executed by one human and typically lasting for brief durations of time, on the order of several of seconds. Examples of actions include bending, walking, and swimming [37, 38]. Furthermore, by “activities” the studies refer to the complex sequence of actions performed by a few persons who could associate with each other in a constrained manner. They are commonly described by much longer temporal durations, e.g., two persons shaking hands, a football team scoring a goal [37, 38].

Motion analysis in autism is based on upper limb data and kinematic gait data [28]. Autism specific motor signs which, being obtained before the development of language, and being assessed quantitatively [27, 28].

II. Early Diagnosis

The early detection of autism spectrum disorders is very important to increase child treatment outcomes, allowing early interventions that rise development and treatment results.[30, 39-41]

Early diagnosis help in the creation of an environment in which everybody who interacts on a usual daily basis with the child with autism has new knowledge about the basic communication, theory of mind deficits in autism and how these can, at least partly, be controlled by structured, concrete modes of interaction and education, and informed by increased knowledge about the typical behaviors related with autism [42].

Early diagnosis prevents of side symptoms for example; aggression, and self-injury. These manners are not in the diagnostic criteria for ASD but are secondary symptoms that develop when primary symptoms are not treated. Early intervention techniques to treat core symptoms of ASD may avoid side symptoms and minimize the need for more substantial and expensive interventions later in life [41].

III. Methods

This section introduces description about participants, methods applied for data collection, task, and analysis. It is comprised of different sections for data acquisition, feature extraction, classification, and evaluation.

In this study, validation of the classifier for the HR versus LR comparison was performed by using support vector machine and extreme learning machine. The input is a number of sub-movements (i.e., 8) was used to extract features by using a Linear Discriminant Analysis (LDA). After that data set is divided into N subsets the testing set is only composed of one sample of the original dataset and the training set is made up of the remaining samples of the original dataset performed by using two machine learning approaches; Support Vector Machine (SVM) and Extreme Learning Machine (ELM). A schematic description of the whole procedure is shown in Fig.2.

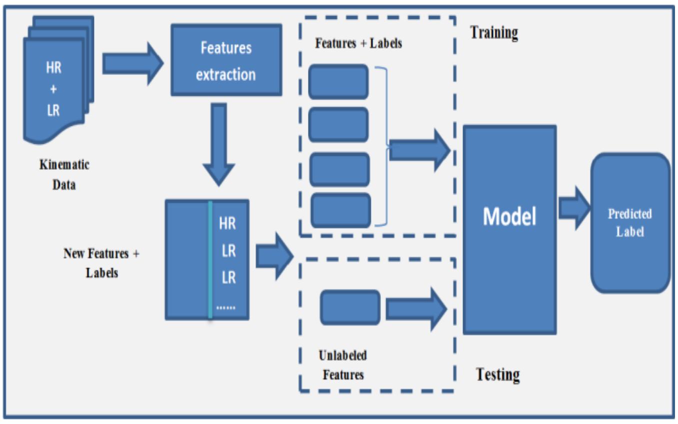

Fig. 2. Flowchart of procedures in this paper

A. Participants

Participants were 17 infants at high risk for autism; where infants high risk due to an older sibling with autism [43] and 15 low-risk infants for autism; where low risk for autism is typically developing older sibling and no family history of autism[43]. All Infants are between the ages of 12 and 36 months[44]. The group characteristics are shown in table 1. This group was obtained from the National Database for Autism Research (NDAR) (https://www.ndar.nih.gov)

B. Data Acquisition

Three sensors were used, two wired sensors (weight: 14 g) that can be worn by infants on their wrists, and a wireless Bluetooth sensor which is embedded into manipulated objects. The sensing core, common both to instrumented toy and wearable devices, is designed to sense samples/sec rate. These data are collected and rearranged in a specific message format by a microcontroller and then retransmitted via (wrist sensors) or Bluetooth device (object sensor). Data transmitted over the Bluetooth are collected by a nearby computer, for later data analysis. This method allows the study of infants in their environment, decreasing their pressure and allowing the collection of extra significant quantitative information. Refer to the technical paper [44], for more details.

C. Task

The data was collected to study and examine fine motor and object manipulation skills in (HR) and comparison of (LR) between the ages of 12 and 36 months. Children’s motor was examined shape sorter task. In this task, the child is asked to insert an object into a box with interchangeable lids (see fig. 3). Each lid has a slot identical, but only slightly larger than the cross-section of the shape to be inserted. Each shape is presented six times: three times in initial vertical direction and three in initial horizontal direction. Refer to the technical paper [44], for more details.

TABLE I. Participant Characteristics

	High Risk (HR)	**Lower Risk (LR)**
number	17	15
male:female	9:8	8:7
age (months)	between the of 12 and 36 months	between the of 12 and 36 months

Fig. 3. Task equipment [44].

D. Feature extraction

The features extracted from the collected sub-movement were more discriminative for the HR versus LR comparison. Table II shows sub-movements values of the two groups of children included in the study (HR vs. LR) and the results of ANCOVA calculated on all kinematic measures. The study found several significant group differences based on the Kinematic variables. Feature extraction was implemented by using a Linear Discriminant Analysis (LDA). LDA techniques are used for data classification and dimensionality reduction[45]. This study used LDA for dimensionality reduction only.

The above aims can be achieved by finding the projection hyperplane that minimizes the inter-class variance and maximizes the distance between the two classes. See fig. 4. This hyperplane can be utilized in many targets for classification, dimensionality reduction and for determination of the significance of the given features [46].

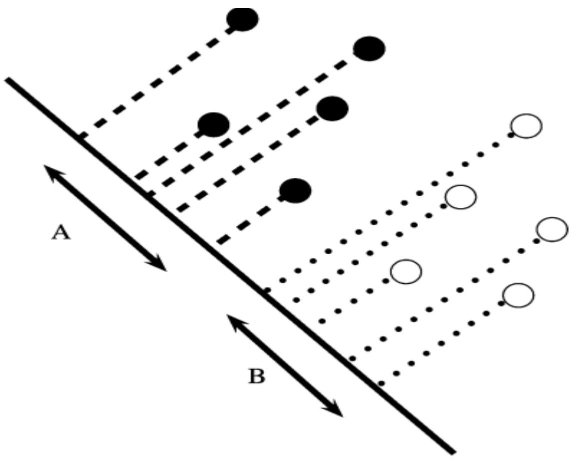

Fig. 4. The line has to be selected so that the projection maximizes the separability of the projected samples [46]

TABLE II. KINEMATIC DATA WERE INITIALLY ANALYZED THROUGH AN ANCOVA WITH GROUP (HR VS.LR)

Sub-movement	HR (mean)	LR (mean)	sig	Partial Eta Square
Reach duration	(0.669)	(0.589)	0.068	0.137
Place duration	(4.671)	(4.561)	0.464	0.024
Latency reach place	(1.380)	(1.134)	0.580	0.013
Score place	(1.875)	(1.896)	0.619	0.011
Hand reach	(0.667)	(0.470)	0.979	0.0
Hand place	(0.571)	(0.359)	0.450	0.025
Shape orientation	(0.399)	(0.393)	0.370	0.035
First contact duration	(28.163)	(3.313)	0.350	0.038

[a]. The alpha level was set to .05 for all data analyses.

Specifically, the LDA score was calculated with base-technique using LDA, for a sample x and a given discriminant function g(x). In this case, Fisher's LDA can be decided. A discriminant function that is a linear collection of the components of x can be written as: [47]

$$g(X) = W^{T} * X + W_0 \tag{1}$$

W is the weight vector

W_0 is the threshold weight.

$$W = \sum_{c}^{-1}(\mu 2 - \mu 1) \tag{2}$$

Where μi is the mean value of class i from C classes and

The covariance (cov) and the mean (μ) are calculated as:

$$\text{cov} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)\,(X_i - \mu)^T \tag{3}$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4}$$

E. Classification Algorithm

Classification of HR and LR subjects was performed by using a two machine learning approaches; Support Vector Machine (SVM) and Extreme Learning Machine (ELM).

The study used SVM classifier because it is easily separable and is a known classifier and ELM we desired of improving the results and used relatively recent classifier; this is what has been achieved. In this study, the whole machine-learning method was implemented on the Matlab platform (Matlab version R2016a). In particular, the study used functions of the toolbox of Matlab to implement the classification algorithm.

Support vector machine (SVM): is a learning algorithm that learns by example to specify labels to objects, so it is supervised learning technique from applicable to both classification and regression. SVM, therefore link the problems they are designed for with a big body of existing work on kernel based methods[48, 49].

Functions that describe Support Vector Machine which is used for binary classification can be written as [50]

Training data is of the form:

$\{x_i, y_i\}$ where i=1….L; $y_i \in \{-1, 1\}$, $x \in \mathbf{R}^D$

L is a training point

D is a dimensional

The hyperplane can be described by

$$w.x+b=0 \tag{5}$$

w is a normal to hyperplane

$\frac{b}{\|w\|}$ is perpendicular distance between the hyperplane to the origin.

Referring to graph in fig.5 the training data can be described by

$$x_i \bullet w+b \geq +1 \quad \text{for } y_i=+1 \tag{6}$$

$$x_i \bullet w+b \leq -1 \quad \text{for } y_i=-1 \tag{7}$$

w and b are training data.

These two equations can be combined into:

$$y_i(x_i \bullet w+b)-1 \geq 0 \ \forall_i \tag{8}$$

The two planes H_1 and H_2 that these points lie on can be described by:

$$x_i \bullet w+b= +1 \quad H_1 \tag{9}$$

$$x_i \bullet w+b = -1 \quad H_2 \tag{10}$$

d_1 is the distance from H_1 to the hyperplane

d_2 is the distance from H_2 to the hyperplane

The hyper plane's equidistance from H_1 and H_2 means that d1 = d2 - a quantity called SVM's margin

Margin is equal to $\frac{1}{\|w\|}$

Maximize the margin means minimize the ||w||

min ||w|| such that

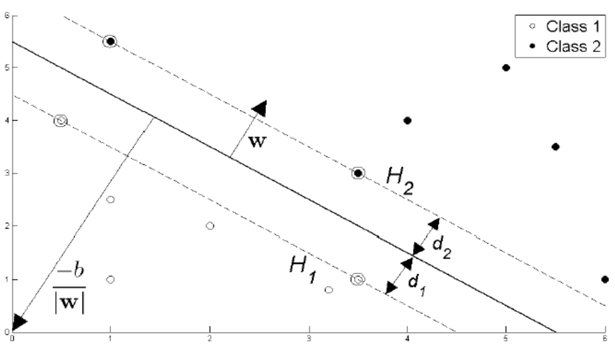

Fig. 5. Hyperplane through two linear separable classes [50]

$$y_i(x_i \bullet w+b)-1 \geq 0 \ \forall_i \tag{11}$$

Minimizing ||w|| is equivalent to minimizing $\frac{1}{2}\|w\|^2$

min $\frac{1}{2}\|w\|^2$ such that

$$y_i(x_i \bullet w+b)-1 \geq 0 \ \forall_i \tag{12}$$

To allocate them Lagrange multipliers α, where $\alpha \geq 0 \ \forall_i$

$$L_p \equiv \frac{1}{2}\|w\|^2 - \sum_{i=1}^{L} \alpha_i y_i (x_i \cdot w + b) + \sum_{i=1}^{L} \alpha_i \tag{13}$$

Now we shall look that when w and b minimizes, the α maximizes "(13),". This shall be achieved by differentiating L_P with respect to w and b and setting the derivatives to zero:

$$\frac{\partial L_P}{\partial W} = 0 \rightarrow w = \sum_{i=1}^{L} \propto_i y_i x_i \tag{14}$$

$$\frac{\partial L_P}{\partial b} = 0 \rightarrow \sum_{i=1}^{L} \propto_i y_i = 0 \tag{15}$$

Substituting the last two equations in "(13),"gives a new equation which, being dependent on $\propto$, so need to maximize

$$max_{\propto}\left[\sum_{i=1}^{L} \alpha_i - \frac{1}{2}\alpha^T H \alpha\right] \ \ s.t\ \alpha \geq 0 \ \ and \sum_{i=1}^{L} \alpha_i y_i = 0 \tag{16}$$

L_D is dual form of primary L_P.

So we moved from minimizing L_P to maximizing L_D to do that we need to find:

$$L_D \equiv \sum_{i=1}^{L} \propto_i - \frac{1}{2} \propto^T H\alpha \ \ s.t \ \ \alpha_i \geq 0 \ \forall_i, \sum_{i=1}^{L} \propto_i y_i = 0 \tag{17}$$

Any data point satisfying "(15)," which is a Support Vector x_2 will have the form:

$$y_s(x_s \cdot W + b) = 1 \tag{18}$$

Substituting in "(14),"

$$y_s\left(\sum_{m \in S} \alpha_m y_m x_m \cdot x_s + b\right) = 1 \tag{19}$$

Where S denotes the set of indices of the support vectors. Multiplying through by y_s and then using $y_s^2 = 1$ from "(6)," and "(7),".

$$y_s^2\left(\sum_{m \in S} \alpha_m y_m X_m \cdot X_s + b\right) = y_s \tag{20}$$

$$b = y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s \tag{21}$$

$$b = \frac{1}{N_s} \sum_{s \in S}\left(y_s - \sum_{m \in S} \alpha_m y_m x_m \cdot x_s\right) \tag{22}$$

After we get the variables w and b that define separating hyper plane's optimal orientation and hence Support Vector Machine.

Extreme learning machine (ELM) is learning algorithm for one hidden layer feed-forward neural networks which randomly selects the input weights and analytically set the output weights of this neural network. In theory, it is extremely fast learning speed. The algorithm can present best generalization efficiency in some cases and can learn much quicker than common learning algorithms for feed forward neural networks [51-53].

Functions that that describe single hidden layer feedforward networks can be written as [52]

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(x_j) = \sum_{i=1}^{\tilde{N}} \beta_i g(W_i \,.\, X_j + b_i) = o_j, \quad (23)$$

Where j=1….N,

N=arbitrary distinct sample (x_i , t_i)

$$x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$$

$$t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$$

g(x) is an activation function

Ñ is hidden nodes

Standard single hidden layer feedforward networks (SLFNs)

Where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector

connecting the i hidden node

The input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is the weight vector connecting the *i* th hidden node and the output nodes, and b_i is the threshold of the *i* th hidden node. $w_i \cdot x_i$ denotes the inner product of $\mathbf{w}_i w_i$ and $\mathbf{x}_j x_j$. The output nodes are chosen linear in this paper.

N samples with zero error means that

$\sum_{j=1}^{\tilde{N}} \| o_j - t_j \| = 0$ i.e., there exist β_i, w_i and b_i such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\, w_i \,.\, X_j + b_i) = t_j, \;\; j = 1, \;\; \cdots, N \quad (24)$$

The above *N* equations can be written compactly as equation

$$H\beta = T$$

$$H = \begin{bmatrix} g(w_1 . X_1 + b_1) & \dots & g(w_{\tilde{N}} . X_1 + b_{\tilde{N}}) \\ \vdots & \dots & \vdots \\ g(w_1 . X_N + b_1) & \cdots & g(w_{\tilde{N}} . X_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \quad (25)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad (26)$$

And

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (27)$$

H is called the hidden layer

the *i* th column of **H** is the *i* th hidden node output with respect to inputs $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$. See fig. 6 to explain the Elm structure.

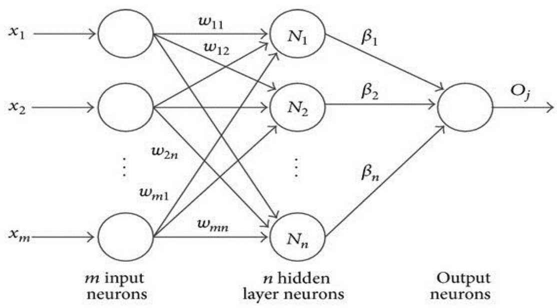

Fig. 6. ELM structure mode [54]

F. Performance of the Classification Algorithm

The study assessed the performance of the classification algorithm by using a cross-validation strategy. Generally, cross-validation includes splitting the dataset into two complementary subsets: a training group and a testing group. The training group is a group of data mapped to a label and used to perform the training of the classifier; the testing group is a group of data not associated with a label and used to perform the validation of the classifier. By considering various segmentations of the data, multiple rounds of cross-validation can then be executed.

In a special case of cross-validation, called leave-one out (LOO) cross-validation, data set is divided into N subsets. The testing set is only composed of one sample of the original dataset and the training set is made up of the remaining samples of the original dataset (N - 1). Therefore, each time, one of the N subsets is used as the test group and the other N-1 subsets are put together to form a training group to test all N samples in the original dataset, then it is sufficient that the number of times to be performed equals the number N of samples in the original dataset. LOO is a commonly used validation way in studies because it is an unbiased estimate of the probability of error [27, 55].

To quantify the performance of the classification algorithm, the sensitivity, specificity and accuracy were computed. sensitivity measure the rate of correctly classified samples in the positive (HR) class, specificity measure the rate of correctly classified samples in the negative (LR) class, and accuracy of classification measures the rate of correctly classified samples in both positive (HR) and negative (LR) classes [56].

Sensitivity = (Number of true classified children HR)/ (Number of all HR)

Specificity = (Number of true classified children LR)/ (Number of all LR)

Accuracy = (Number of correct classified)/ (Number of all children)

G. Evaluation

In this study, accuracy is used as the major index to explain the performance of classification. Accuracy is defined as the rate of correct classification of all data HR and LR in a test set. The accuracy of each session is computed using five-fold cross validation. In each fold, a block is chosen as the test set, and the remainder used as a training set. The accuracy of each session was calculated as the average accuracy of five classifications. In addition, statistical analyses are applied to explain the experimental results. As previously mentioned, the experiments were founded on data collected.

IV. RESULTS

As mentioned before we used cross-validation and SVM. The study used SVM code [57-59] which divides the database randomly into two group's; dataset for training and dataset for testing. The SVM algorithm was repeated five times (N=5), the first four times it selects different 6 infants as test dataset and different 26 infants as training dataset. But in the last time it selects eight infants as a testing dataset and 24 infants as a training dataset. We traced the execution and documented the result each time.

The machine-learning method (SVM) was able to successfully classify participants. The classification accuracy reached a maximum accuracy of 83.3 % (sensitivity 80 % and specificity 100 %). Overall mean accuracy, specificity, and sensitivity rates were also calculated. The overall mean classification was calculated: mean sensitivity = % 76.5, mean specificity = % 66.7 and mean accuracy = % 71.9

In fig. 7, the dependence of the metrics on the number of considered inputs is shown. The resulting data are shown for a number of inputs ranging from one to eight. As can be noticed, accuracy, specificity, and sensitivity rates increase with the number of selected inputs.

Then the study performed the ELM algorithm. This algorithm was published in[60] and we implemented the algorithm on the Matlab platform (Matlab version R2016a). The study accomplished this stage by repeating the previous steps (that performed in SVM) more clearly. The database was divided into two groups; training group and testing group. This step was repeated five times the first four times; training dataset was selected with 26 as a training dataset and different 6 infants as a testing dataset. In the last time, training dataset was selected with 24 as a training dataset and eight children as a testing dataset same as SVM.

The EL was able to successfully classify participants by diagnosis. The classification accuracy reached a maximum accuracy of 100 % (sensitivity 100 % and specificity 100 %) overall mean accuracy, specificity, and sensitivity rates were also calculated. The overall mean classification for sensitivity, specificity and accuracy is %78.5, % 66.7 and % 72.5, respectively.

In fig.8 the graph shows classification sensitivity, specificity and accuracy rates (%) of ELM (Y-axis) in a relation of the number of considered same as SVM features graph, but the sensitivity, specificity, and accuracy are closed to each other and almost similar in behavior.

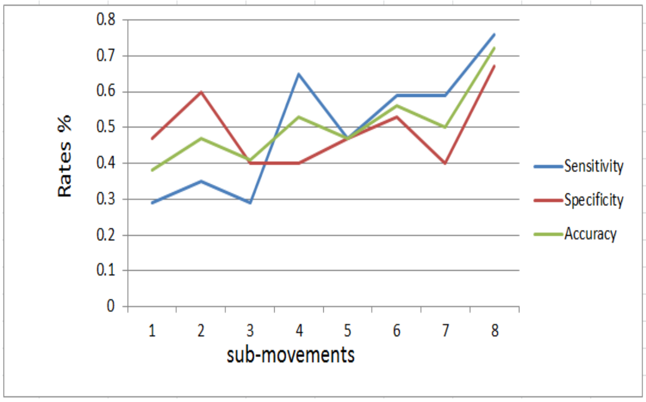

Fig. 7. Graph showing classification sensitivity, specificity and accuracy rates (%) of SVM (Y-axis) in a relation of the number of considered Kinematic data(X-axis). As expected, accuracy, specificity, and sensitivity rates increased with the number of Kinematic data

V. DISCUSSION AND FUTURE WORK

The optimal separating hyper-plane for the (HR) group (LR) group (first, second and third sub-movements) representative shown in fig. 9, is an example of the training phase of the machine-learning method before features extraction phase.

This study aims to examine and analysis the upper limb movements to correctly discriminate between HR infants and LR infants. We are looking to follow-up the study and find out which HR infant will be diagnosed as an autistic child. Someone may look at these results as low outcomes if it is compared with other studies like [27] where it reached an accuracy of 84.9 % and the participants were diagnosed as an autistic child in the mentioned study. However, that work used dataset where all subjects are autistic but our dataset is not because it is not every HR infant is an autistic child. Autism is now diagnosed on the basis of symptoms as qualitatively judged by clinicians and by means of semi-structured observations (ADOS) and standardized interviews or questionnaires (ADI-R) these take too long time and diagnosis age is high, so the study new technique that may help families and clinicians to save time and efforts.

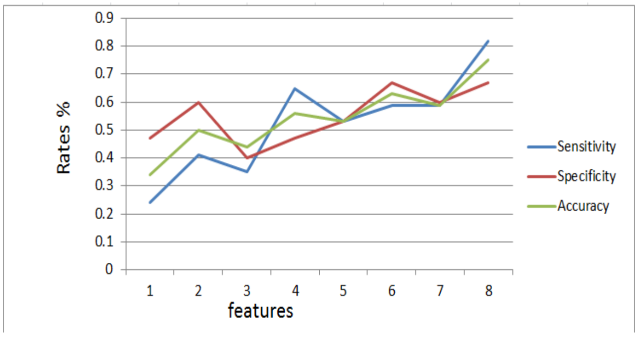

Fig. 8. Graph showing classification sensitivity, specificity and accuracy rates (%) of ELM (Y-axis) in a relation of the number of considered

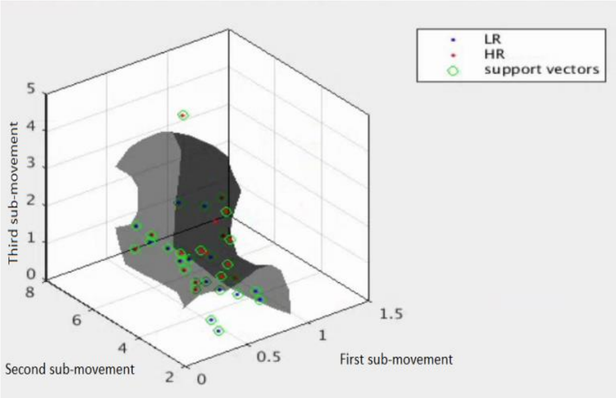

Fig. 9. The optimal separating hyperplane for the autism High-Risk (HR) group versus Low Risk before using LDA

Classification of HR and LR children was performed using a (SVM) approach (Vedaldi's MATLAB wrapper [1]). Because it is easily separable and is a known classifier so the purpose of the considered SVM is to create a model can learn from the selected features of labeled children how to discriminate children of different groups (binary labeled), and correctly classify, by means of the same selected features, new unlabeled children as belonging to one of the two groups (HR or LR).The learning process of the classifier consists of a training stage in which the selected features of the HR and LR children are two training datasets associated to the HR and LR labels, respectively. In our study, if we have training data associated binary label, then SVM uses the principle of structural risk minimization to design an optimal hyperplane that maximizes the distance between the two training groups (HR and LR) that separates training groups. The lower the distance of a training child from the optimal hyperplane, the more important that training child is to define the optimal hyperplane. So, the distance matches the ''weight'' of that training child in the definition of the optimal hyperplane. The optimal hyperplane can then be used as a model to classify new children. ELM we desired of improving the results and used relatively recent classifier; this is what has been achieved.

VI. Conclusion

Autism is a lifelong disorder condition; with difficult ability of the child to communicate to others and develop mental disability. Also, autism prevalence is rising around the world. During autism, it effects an individual person, family, and society during the lifetime. It is difficult to diagnose by medical doctors with support from physical. Studies show that many autistic children have atypical motor patterns but have not received much experimental documentation. Furthermore, we presented that the early detection of autism spectrum disorders is very important to increase child treatment outcomes, allowing early interventions that rise development and treatment results.

Then the study performed kinematic analysis of a simple insert of an object into a box with interchangeable lids performed by infants with High Risk for autism (HR) in comparison to their age-matched, Low Risk (LR) developing peers neuro anatomical networks implicated in ASD using a classification approach employing Support Vector Machine (SVM) and Extreme Learning Machine (ELM). SVM provided good group separation and ELM provided slightly better results, but the speed ELM is much better than SVM. In the future work, we will study children younger than this study sample in different ways, and we are looking to follow-up study, find out which HR infant will be diagnosed as an autistic child and the result will be more accurate.

Acknowledgment

Authors acknowledge for National Database for Autism Research (NDAR) for granting the authors opportunity to access this database.

References

[1]. P. I. Markowitz, "Autism in a child with congenital cytomegalovirus infection," Journal of autism and developmental disorders, vol. 13, pp. 249-253, 1983.

[2]. M. Elsabbagh, G. Divan, Y. J. Koh, Y. S. Kim, S. Kauchali, C. Marcín, C. Montiel-Nava, V. Patel, C. S. Paula, and C. Wang, "Global prevalence of autism and other pervasive developmental disorders," Autism Research, vol. 5, pp. 160-179, 2012.

[3]. W. H. Organization, "Autism spectrum disorders & other developmental disorders. From raising awareness to building capacity," ed: Geneva, Switzerland: Author, 2013.

[4]. C. R. Jensen, "A machine learning approach for identifying subtypes of autism," 2015.

[5]. nimh. (2016, 10-07-2016). http://www.nimh.nih.gov/health/statistics/index.shtml. Available: http://www.nimh.nih.gov/health/statistics/index.shtml

[6]. M. Knapp, R. Romeo, and J. Beecham, "The economic consequences of autism in the UK," 2007.

[7]. M. Knapp, R. Romeo, and J. Beecham, "Economic cost of autism in the UK," Autism, vol. 13, pp. 317-336, 2009.

[8]. K. Järbrink, "The economic consequences of autistic spectrum disorder among children in a Swedish municipality," Autism, vol. 11, pp. 453-463, 2007.

[9]. A. V. Buescher, Z. Cidav, M. Knapp, and D. S. Mandell, "Costs of autism spectrum disorders in the United Kingdom and the United States," JAMA pediatrics, vol. 168, pp. 721-728, 2014.

[10]. Z. Cidav, S. C. Marcus, and D. S. Mandell, "Implications of childhood autism for parental employment and earnings," Pediatrics, vol. 129, pp. 617-623, 2012.

[11]. T. A. Lavelle, M. C. Weinstein, J. P. Newhouse, K. Munir, K. A. Kuhlthau, and L. A. Prosser, "Economic burden of childhood autism spectrum disorders," Pediatrics, vol. 133, pp. e520-e529, 2014.

[12]. R. L. Mendoza, "The economics of autism in Egypt," American Journal of Economics and Business Administration, vol. 2, p. 12, 2010.

[13]. D. L. Sharpe and D. L. Baker, "The financial side of autism: Private and public costs," A comprehensive book on autism spectrum disorders, p. 275, 2011.

[14]. M. L. Ganz, "The lifetime distribution of the incremental societal costs of autism," Archives of pediatrics & adolescent medicine, vol. 161, pp. 343-349, 2007.

[15]. L. Roth, "Autism spectrum disorder, Briefing Paper No. 5/2013," ed, 2014.

[16]. C. Horlin, M. Falkmer, R. Parsons, M. A. Albrecht, and T. Falkmer, "The cost of autism spectrum disorders," PloS one, vol. 9, p. e106552, 2014.

[17]. K. Järbrink, P. McCrone, E. Fombonne, H. Zandén, and M. Knapp, "Cost-impact of young adults with high-functioning autistic spectrum disorder," Research in developmental disabilities, vol. 28, pp. 94-104, 2007.

[18]. C. Dudley and J. Emery, "The value of caregiver time: Costs of support and care for individuals living with autism spectrum disorder," SPP Research Paper, 2014.

[19]. E. A. Alsaggaf and S. S. Baaisharah, "Directions of Autism Diagnosis by Electroencephalogram Based brain Computer Interface: A Review," Life Science Journal, vol. 11, 2014.

[20]. M. D. King and P. S. Bearman, "Socioeconomic status and the increased prevalence of autism in California," American Sociological Review, vol. 76, pp. 320-346, 2011.

[21]. M. King and P. Bearman, "Diagnostic change and the increased prevalence of autism," International journal of epidemiology, vol. 38, pp. 1224-1234, 2009.

[22]. M. Sigman, S. J. Spence, and A. T. Wang, "Autism from Developmental and Neuropsychological Perspectives," Annual Review of Clinical Psychology, vol. 2, pp. 327-355, 2006/04/01 2006.

[23]. R. C. M. Philip, "Emotion processing in autism spectrum disorder," 2009.

[24]. L. Zwaigenbaum, S. Bryson, T. Rogers, W. Roberts, J. Brian, and P. Szatmari, "Behavioral manifestations of autism in the first year of life," International journal of developmental neuroscience, vol. 23, pp. 143-152, 2005.

[25]. G. Dawson, J. Osterling, A. N. Meltzoff, and P. Kuhl, "Case study of the development of an infant with autism from birth to two years of age," Journal of Applied Developmental Psychology, vol. 21, pp. 299-313, 2000.

[26]. P. J. Krantz and L. E. McClannahan, "Social interaction skills for children with autism: A script-fading procedure for beginning readers," Journal of applied behavior analysis, vol. 31, pp. 191-202, 1998.

[27]. A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, and I. Castiglioni, "Use of Machine Learning to Identify Children with Autism and Their Motor Abnormalities," Journal of autism and developmental disorders, vol. 45, pp. 2146-2156, 2015.

[28]. P. Perego, S. Forti, A. Crippa, A. Valli, and G. Reni, "Reach and throw movement analysis with support vector machines in early diagnosis of autism," in Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE, 2009, pp. 2555-2558.

[29]. M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," Biomedical Engineering, IEEE Transactions on, vol. 55, pp. 1956-1965, 2008.

[30]. J. Hashemi, T. V. Spina, M. Tepper, A. Esler, V. Morellas, N. Papanikolopoulos, and G. Sapiro, "A computer vision approach for the assessment of autism-related behavioral markers," in Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on, 2012, pp. 1-7.

[31]. G. Esposito, P. Venuti, F. Apicella, and F. Muratori, "Analysis of unsupported gait in toddlers with autism," Brain and Development, vol. 33, pp. 367-373, 2011.

[32]. J. Bo, C.-M. Lee, A. Colbert, and B. Shen, "Do children with autism spectrum disorders have motor learning difficulties?," Research in Autism Spectrum Disorders, vol. 23, pp. 50-62, 2016.

[33]. R. Downey and M. J. K. Rapport, "Motor activity in children with autism: a review of current literature," Pediatric Physical Therapy, vol. 24, pp. 2-20, 2012.

[34]. M. Shetreat-Klein, S. Shinnar, and I. Rapin, "Abnormalities of joint mobility and gait in children with autism spectrum disorders," Brain and Development, vol. 36, pp. 91-96, 2014.

[35]. G. Dawson and R. Watling, "Interventions to facilitate auditory, visual, and motor integration in autism: A review of the evidence," Journal of autism and developmental disorders, vol. 30, pp. 415-421, 2000.

[36]. G. T. Baranek, "Efficacy of sensory and motor interventions for children with autism," Journal of autism and developmental disorders, vol. 32, pp. 397-422, 2002.

[37]. S. Vishwakarma and A. Agrawal, "A survey on activity recognition and behavior understanding in video surveillance," The Visual Computer, vol. 29, pp. 983-1009, 2013.

[38]. P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 18, pp. 1473-1488, 2008.

[39]. C. Fountain, M. D. King, and P. S. Bearman, "Age of diagnosis for autism: individual and community factors across 10 birth cohorts," Journal of Epidemiology and Community Health, vol. 65, pp. 503-510, 2011.

[40]. C. P. Johnson and S. M. Myers, "Identification and evaluation of children with autism spectrum disorders," Pediatrics, vol. 120, pp. 1183-1215, 2007.

[41]. L. K. Koegel, R. L. Koegel, K. Ashbaugh, and J. Bradshaw, "The importance of early identification and intervention for children with or at risk for autism spectrum disorders," International journal of speech-language pathology, vol. 16, pp. 50-56, 2014.

[42]. E. Fernell, M. A. Eriksson, and C. Gillberg, "Early diagnosis of autism and impact on prognosis: a narrative review," Clinical epidemiology, vol. 5, pp. 33-43, 2013.

[43]. T. St John, A. M. Estes, S. R. Dager, P. Kostopoulos, J. J. Wolff, J. Pandey, J. T. Elison, S. J. Paterson, R. T. Schultz, and K. Botteron, "Emerging Executive Functioning and Motor Development in Infants at High and Low Risk for Autism Spectrum Disorder," Frontiers in Psychology, vol. 7, p. 1016, 2016.

[44]. F. Taffoni, V. Focaroli, D. Formica, E. Gugliemelli, F. Keller, and J. M. Iverson, "Sensor-based technology in the study of motor skills in infants at risk for ASD," in 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), 2012, pp. 1879-1883.

[45]. S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," Institute for Signal and information Processing, vol. 18, 1998.

[46]. P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear discriminant analysis," in Robust Data Mining, ed: Springer, 2013, pp. 27-33.

[47]. W. Dwinnell and D. Sevis, "LDA: Linear Discriminant Analysis," Matlab Central File Exchange, vol. 29673, 2010.

[48]. C. J. Burges, "A tutorial on support vector machines for pattern recognition," Data mining and knowledge discovery, vol. 2, pp. 121-167, 1998.

[49]. W. S. Noble, "What is a support vector machine?," Nature biotechnology, vol. 24, pp. 1565-1567, 2006.

[50]. T. Fletcher, "Support vector machines explained," Online]. http://sutikno. blog. undip. ac. id/files/2011/11/SVM-Explained. pdf.[Accessed 06 06 2013], 2009.

[51]. G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on, 2004, pp. 985-990.

[52]. G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," Neurocomputing, vol. 70, pp. 489-501, 2006.

[53]. Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, "Evolutionary extreme learning machine," Pattern recognition, vol. 38, pp. 1759-1763, 2005.

[54]. Z.-H. You, Y.-K. Lei, L. Zhu, J. Xia, and B. Wang, "Prediction of protein-protein interactions from amino acid sequences with ensemble extreme learning machines and principal component analysis," BMC bioinformatics, vol. 14, p. 1, 2013.

[55]. R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in Ijcai, 1995, pp. 1137-1145.

[56]. W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations," NESUG proceedings: health care and life sciences, Baltimore, Maryland, pp. 1-9, 2010.

[57]. F. Lauer and Y. Guermeur, "MSVMpack: a multi-class support vector machine package," Journal of Machine Learning Research, vol. 12, pp. 2293-2296, 2011.

[58]. J. Weston and C. Watkins, "Multi-class support vector machines," Citeseer1998.

[59]. A. Vedaldi, "A MATLAB wrapper of SVMstruct," ed, 2011.

[60]. ELM, "http://www.ntu.edu.sg/home/egbhuang/elm_codes.html," 2013.