31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Reach and throw movement analysis with Support Vector Machines in early diagnosis of autism.

Paolo Perego*, Sara Forti, Alessandro Crippa, Angela Valli and Gianluigi Reni**

Abstract— **Movement disturbances play an intrinsic part in autism. Upper limb movements like reach-and-throw seem to be helpful in early identification of children affected by autism. Nevertheless few works investigate the application of classifying methods to upper limb movements. In this study we used a machine learning approach Support Vector Machine (SVM) for identifying peculiar features in reach-and-throw movements. 10 pre-scholar age children with autism and 10 control subjects performing the same exercises were analyzed. The SVM algorithm proved to be able to separate the two groups: accuracy of 100% was achieved with a soft margin algorithm, and accuracy of 92.5% with a more conservative one. These results were obtained with a radial basis function kernel, suggesting that a non-linear analysis is possibly required.**

I. INTRODUCTION

Autism is a complex developmental disability which primarily affects a person's ability in social interaction and communication. Autism is known as a spectrum disorder, because it affects each individual in different ways and degrees. According to the Center for Disease Control and Prevention (2007) [1], autism is estimated to occur in one out of 150 children in the USA.

Diagnostic criteria for autism are primarily based on impairments in social functioning and communication skills (DSM-IV_TR & ICD-10). However, there is a growing research interest for motor impairment in autism, and, it's even been suggested that most of the social symptoms may actually originate from deficits in motor functioning [2].

In our study, we analyze upper limb kinematics during a reach-and-throw task in a group of children with autism at nursery age, compared to a group of normally developing children of the same age.

Support Vector Machines (SVMs) are learning networks used for two-group classification problems. SVMs calculate the weights of the network by solving a quadratic programming problem with linear constraints, rather than by solving a not-convex, unconstrained minimization problem as in standard neural network training. SVMs are suitable for data a priori not linearly independent. They are normally used in bimolecular analysis but also in EEG and ECG [3] and motion analysis. Many works on the application of SVMs in motion analysis are based on kinematic gait data [4], [5], [6] and only a few deal with upper limb data [7].

With this research, we aim to use SVM to identify some autism-specific motor signs which, being acquired before the development of language, and being evaluated quantitatively, may represent a praecox, and indeed an objective indicator of risk in autism, supporting clinicians towards earlier diagnoses in life, with substantial effects on the therapy outcomes.

Paolo Perego and Gianluigi Reni are with the Bioengineering Lab of I.R.C.C.S. E. Medea

Sara Forti, Alessandro Crippa and Angela Valli are with the Clinical Unit of I.R.C.C.S. E. Medea

* Corresponding author paolo.perego@bp.lnf.it

** IEEE Member

II. METHODS

The reach-and-throw movement was chosen because it is a milestone in the growth of a child and its incomplete or faulty acquisition could prevent the development of successive superior neuro-motor functions. It can be studied in infants, as it requires the application of any internal motor models (already acquired at 6 months) and just a few strategic programming (present at 3 years). In the end it is simple enough to be executed by infants, and it is useful to stress some early cognitive abilities in order to differentiate between normal children and risky ones. Many studies show that children affected by autism do not control this movement completely even at scholar age, being slower and less accurate compared to same age normal children [8], [9].

A. Experimental design

Each child sat in front of a shaped table, whose height was adjusted to align to the base of the child's trunk. The experimental task consisted of grasping a rubber ball (6.0 cm diameter) placed over a support, and throwing it in a see-through squared basket (21 cm high), with a hole large enough (7.0 cm diameter) to require not so fine movements, see "Fig. 1". Ten throwing trials were conducted for each participant, 5 with the right hand and 5 with the left one.

B. Subjects

Two groups of children were studied. The first group included 10 normal subjects (mean age of 41.60 months $\pm$ 9.23, mean IQ 119.02 $\pm$ 16.07). The second group included 10 subjects affected by autism (mean age of 41.44 months $\pm$ 7.13, mean IQ 72.86 $\pm$ 14.10).

C. Instrumentation

Kinematic variables were recorded by means of an optoelectronic system, with 8 infrared cameras working at 60 Hz (SMART-BTS ®). Passive markers were used to reflect the infrared concentric beam of light emitted from around each camera. The markers were positioned on the child (shoulder, elbow, medial and lateral position of the wrist, 4-5 metacarpus for every child), on the basket under the goal area and on the ball, see "Fig. 1"

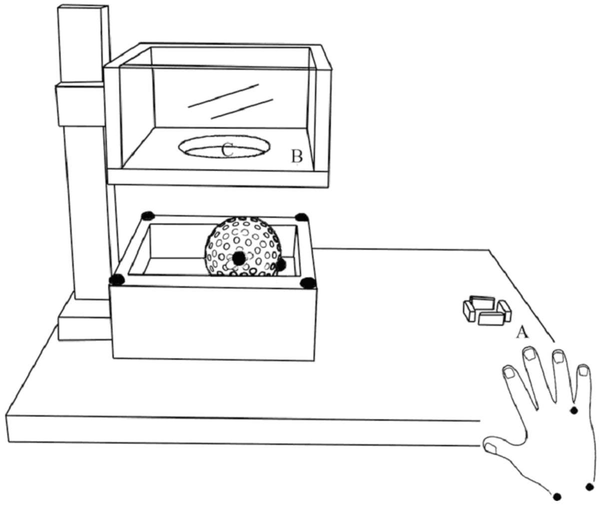

Fig. 1. Experimental setting. The ball is initially positioned on the support (A). The child has to reach the ball, grasp it and do a second reach to throw the ball in the goal area (B) and through the hole (C). The goal area is transparent to allow seeing through. 4 markers (black points) are placed on the basket under the goal area, 2 on the ball and 3 on each hand.

D. Movement analysis

The complete movement from the starting position till the ball release, was split in parts. Criteria for identifying single parts are presented in Table I. Anomalous movements were excluded from the analysis in order to avoid classification errors. An 8-Hz Butterworth low-pass filter was applied to the data [10] and segmentation was then automatically performed with specifically developed software written in Matlab (Matworks ®).
Specific spatio temporal parameters were identified in order to analyze each segment (Table II) according to previous studies on upper limb movements analysis in children [8], [9].

TABLE I
MOVEMENT SEGMENTATION

Segment	Starting	Ending
Reaching 1: *"Reaching the ball"*	T0 - a non-stop movement of the hand towards the ball starts and child wrist reaches a speed greater than 10 mm/sec [11]	T1 - the speed slows under 10 mm/sec, after the hand has touched the ball
Reaching 2: *"Reaching the basket"*	T2 - a non-stop movement towards the goal area starts. Only movements starting within 20 cm from the support are considered valid	T3 - the wrist speed slows under 10 mm/sec, after the hand has entered the goal area
Adjustment movements (if present [a])	T3	T4 - the child releases the ball

[a] When the child programs a complete Reaching 2 movement since its starting, no adjustment movement is observed and the ball is released at the end of Reaching 2 (T3=T4)

TABLE II
SPATIO-TEMPORAL PARAMETERS DEFINED FOR EACH SEGMENT

Reaching 1	**Reaching 2**	**Adjustment movements**
Total Duration [TD1, $msec$]	Total Duration [TD2, $msec$]	Total Duration [TDf, $msec$]
Number of Movement Units [a] [MU1,#]	Number of Movement Units [MU2,#]	Number of Movement Units [MUf,#]
Peak Velocity [PV1, cm/sec]	Peak Velocity [PV2, cm/sec]	
Time of Peak Velocity [TPV1, sec]	Time of Peak Velocity [TPV2, sec]	
Straightness[b] [str, $ratio$]	Wrist angle between the vertical axis and the hand of the subject at time T3 [WA,$degrees$]	

[a] A Movement Unit is determined by an acceleration phase followed by a deceleration one, where the speed module both in acceleration and in deceleration is greater than 10 mm/sec and the module of either acceleration or deceleration is greater then 20 mm/sec^2 [11].
[b] Straightness is calculated as the ratio between Cumulative Distance and minimal distance, where minimal distance is computed by a linear regression between the detected positions of the wrist at time T0 and T1 [12].

E. SVM Classifier

The support-vector network is a learning machine for two-group classification problems. Input vectors are non-linearly mapped to a n-dimension feature space where a linear decision surface is constructed [13], [14].
Lateral mean values, in a normalized z-score form ($z = \frac{x_i - \mu}{\sigma}$), were used to facilitate the training of SVMs [13]. SVM formulation is presented in Appendix. SVMs require the selection of a proper kernel function. Till now neither empirical nor analytical studies proves the superiority of any kernel function. In this study we experimented the use of 3 different kernel functions:

- *linear:* $K(x_i, x_j) = \mathbf{x}_i \cdot \mathbf{x}_j$
- *polynomial:* $K(x_i, x_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^d$
- *radial basis:* $K(x_i, x_j) = \exp(-\sigma|\mathbf{x}_i - \mathbf{x}_j|^2)$

The regularization parameter C, that settles the trade-off between the maximum margin and the minimum classification error [4], was used to tune the machine. Different values of C (from 0.01 to 100) were tested for each kernel solution. LibSVM package of the R project was used for the classification [15].

III. EXPERIMENTAL RESULTS

A. Statistical Analysis

We analyzed the movement parameters with a 1-way anova F(1,38) to test the equality between the two groups (null hypothesis H0). WA, TDf and MUf showed the most significant differences (Table III).

B. SVM Classification

Area Under ROC Curve (AUC) values were computed for each movement feature in the case of linear kernel and C=1 (Table IV), in order to obtain the power of every feature in the basic classification condition. Accuracy[1] was estimated for each considered kernel function in different conditions, as mentioned above. A 10 k-fold cross-validation method was used, where all these data are considered for both training and testing in a recursive procedure [16]. The maximum accuracies observed are reported in Table V. The best classification (accuracy 95%) was obtained with the radial basis kernel (σ = 0.05) and the highest C value (C=100), suggesting that the separation margin between the two groups was small.
The Hill-Climbing feature selection [17] was then performed in two cases: firstly in case of maximum accuracy with this sample of data (rbf kernel, σ=0.05, C=100) and secondly in case of major robustness of the classification in view of new test data (rbf kernel, σ=0.5, C=1). Features were added considering the reverse order of their corresponding AUC values (Table IV). Results are shown in “Fig. 2”. In both cases the accuracy increased when including MU1, decreased and then increased again with MU2.

TABLE III
1-WAY ANOVA F VALUES FOR EACH FEATURE. F(1,38)

TD1	MU1	PV1	TPV1
9.21 **	12.45 **	7.01 *	9.17 **
str	**TD2**	**MU2**	**PV2**
8.26 **	10.52 **	7.93 **	6.88 *
TPV2	**WA**	**TDf**	**MUf**
6.88 *	18.21 ***	27.53 ***	54.71 ***

*** p<0.001 ** p<0.01 * p<0.05

TABLE IV
AUC VALUES FOR EACH MOVEMENT PARAMETER (LINEAR KERNEL, C=1).

MUf	TDf	WA	MU1	str	TD1
0.98	0.98	0.87	0.85	0.78	0.77
TPV1	**TD2**	**PV1**	**MU2**	**PV2**	**TPV2**
0.77	0.76	0.75	0.73	0.71	0.66

TABLE V
MAXIMUM ACCURACY VALUES FOR DIFFERENT KERNEL FUNCTIONS AND PARAMETERS.

Kernel	C	Accuracy
Linear	10	90
Polynomial (d=2)	10	65
Polynomial (d=3)	10	87.5
Rbf (σ=0.5)	1	85
Rbf (σ=0.05)	100	95

[1]Accuracy is defined as $Acc = \frac{TPR+TNR}{TPR+FPR+FNR+TNR}$ where TPR=True Positive Rate, TNR=True Negative Rate, FPR=False Positive Rate, FNR=False Negative Rate

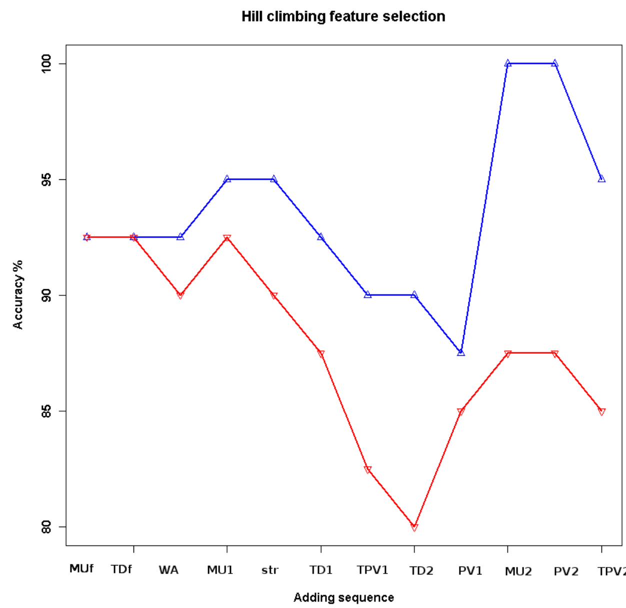

Fig. 2. Hill-Climbing feature selection in the case of maximum accuracy (rbf kernel, σ=0.05, C=100, line with triangle point-up) and in an higher robustness case (rbf kernel, σ=0.5, C=1, line with triangle point-down). In both cases an increase of accuracy can be observed when including MU1 feature as well as when including MU2 feature.

These results suggest that Movement Units (MU1 and MU2) have to be considered as discriminating features in addition to TDf, MUf and WA, because they explain the remaining variance.
Accuracy of 100% was achieved adding these features to the classification with the soft margin separation (line with triangle point-up, “Fig. 2”).

IV. DISCUSSION AND FUTURE WORKS

A. Conclusions

The reach-and-throw movement was segmented and described by means of a set of 12 features.
Classical statistical analysis was able to identify a first group of these features that were useful in differentiating between the autistic children group and the control group. They were related to the final part of the movement.
SVM classification proved its better performance with a radial basis function kernel (non-linear separation) and demonstrated that also the movement units MU1 and MU2 are important features to be considered.
The best accuracy (100%) was obtained with a soft margin classification, where high accuracy is achieved by means of a narrow separation margin. Very good accuracy (92.5%) was obtained even in a more conservative case with a harder margin classification, assuring the power of the classification system in view of new collecting data.
According to these results, signs of autism impairment can be really extracted from the reach-and-throw movement confirming the hypothesis that movement classification can be helpful in supporting early diagnosis, showing high values both for sensitivity and specificity.

B. Future Works

Many methods are available besides the Hill-Climbing feature selection used in this work. For example a data mining or a feature extraction via PCA [5], Wavelet [18] or rough set theory [6] could be performed. Their use could lead to a good classification with a reduced number of features and they should be tested.
Moreover a larger sample of data has to be collected, to allow using part of them for training the classification system and the rest for validating it.

APPENDIX
SVM FORMULATION

Considering a training set $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^{l}$, with $\mathbf{x}_i \epsilon \Re^n$ as input and $y_i \epsilon \{-1, +1\}$ as output, each input $\mathbf{x}$ is firstly mapped into another feature space Υ by $\mathbf{g}= \phi(\mathbf{x})$ with a linear or a nonlinear mapping $\phi : \Re^n \rightarrow \Upsilon$. Υ feature space can both be higher or equal to n upon on which mapping function is used. A linear separation in a higher feature space (Υ) could be viewed as a non-linear separation in the starting space ($\Re^n$). Considering the case where data are linearly separable in Υ, there exists a vector $\mathbf{w} \epsilon \Upsilon$ and a scalar b that define the optimal separating hyperplane (OSM) as $\mathbf{w} \cdot \mathbf{g} + b = 0$ such that

$$y_i = (\mathbf{w} \cdot \mathbf{g}_i + b) \geq 1, \forall i \tag{1}$$

Maximize the margin of separation between the two classes $(2/\|\mathbf{w}\|)$ is minimize $\mathbf{w} \cdot \mathbf{w}/2$ under the constrain of Eq.(1).
For non linearly separable data the above minimization problem is modified in the constrain formula as follows

$$y_i = (\mathbf{w} \cdot \mathbf{g}_i + b) \geq 1 - \xi_i, \forall i \tag{2}$$

where ξ_i are errors of classification. So ξ_i can be regarded as a measure of misclassification. The position of the hyperplane is determined by the weights, w_i and bias, b which are obtained by solving the following SVM quadratic programming problem formulated as a Lagrangian dual problem:

$$\begin{aligned} &\textit{maximize } W(\alpha) = \sum_{i=1}^{L} \alpha_i - \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j \mathbf{g}_i \cdot \mathbf{g}_j \\ &\textit{subject to } \sum_{i=1}^{L} y_i \alpha_i = 0, \text{ and } 0 \leq \alpha_i \leq C, \forall i \end{aligned} \tag{3}$$

where $\alpha = (\alpha_i, \ldots, \alpha_L)$ is the non-negative Lagrangian multiplier. The data points $\mathbf{x}_i$ corresponding $\alpha_i > 0$ lie along the margin of the hyperplane and are named as Support Vectors (SVs). Kernel function **K(.,.)** can be used in Eq. 3 as $\mathbf{g}_i \cdot \mathbf{g}_j$ obtaining $\mathbf{g}_i \cdot \mathbf{g}_j = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$. Determined the optimum Lagrange multipliers, the optimal solution for weight vector $\mathbf{w}$ is

$$\mathbf{x} = \sum_{i \in SVs} \alpha_i y_i \mathbf{g}_i \tag{4}$$

having for any test vector $\mathbf{x} \in \mathbf{Re}^n$ the output given by

$$\begin{aligned} y &= f(\mathbf{x}) = sign(\mathbf{w} \cdot \mathbf{g} + b) \\ &= sign(\sum_{i \in SVs} \alpha_i y_i \mathbf{k}(\mathbf{x}_i, \mathbf{x}) + b) \end{aligned} \tag{5}$$

REFERENCES

[1] http://www.cdc.gov/ncbddd/autism/.
[2] M. R. Leary and D. A. Hill, "Moving on: autism and movement disturbance." *Ment Retard*, vol. 34, no. 1, pp. 39–53, Feb 1996.
[3] A. H. Khandoker, M. Palaniswami, and C. K. Karmakar, "Support vector machines for automated recognition of obstructive sleep apnea syndrome from ecg recordings." *IEEE Trans Inf Technol Biomed*, vol. 13, no. 1, pp. 37–48, Jan 2009.
[4] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data." *J Biomech*, vol. 38, no. 3, pp. 401–408, Mar 2005.
[5] J. Wu, J. Wang, and L. Liu, "Feature extraction via kpca for classification of gait patterns." *Hum Mov Sci*, vol. 26, no. 3, pp. 393–411, Jun 2007.
[6] Z. Yan, Z. Wang, and H. Xie, "Joint application of rough set-based feature reduction and fuzzy ls-svm classifier in motion classification." *Med Biol Eng Comput*, vol. 46, no. 6, pp. 519–527, Jun 2008.
[7] M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb." *IEEE Trans Biomed Eng*, vol. 55, no. 8, pp. 1956–1965, Aug 2008.
[8] P. Teitelbaum, O. Teitelbaum, J. Nye, J. Fryman, and R. G. Maurer, "Movement analysis in infancy may be useful for early diagnosis of autism." *Proc Natl Acad Sci U S A*, vol. 95, no. 23, pp. 13 982–13 987, Nov 1998.
[9] L. Rnnqvist and E. Domellf, "Quantitative assessment of right and left reaching movements in infants: a longitudinal study from 6 to 36 months." *Dev Psychobiol*, vol. 48, no. 6, pp. 444–459, Sep 2006.
[10] E. M. Robin Roithner, Hermann Schwameder, "Determination of optimal filter parameters for filtering kinematic walking data using butterworth low pass filter," in *Proceedings of the XVII International Symposium of Biomechanics in Sport*, 2000.
[11] C. von Hofsten, "Structuring of early reaching movements: a longitudinal study." *J Mot Behav*, vol. 23, no. 4, pp. 280–292, Dec 1991.
[12] E. Thelen, D. Corbetta, and J. P. Spencer, "Development of reaching during the first year: role of movement speed." *J Exp Psychol Hum Percept Perform*, vol. 22, no. 5, pp. 1059–1076, Oct 1996.
[13] V. N. Vapnik, *The Nature of Statistical Learning Theory (Information Science and Statistics)*. Springer, November 1999.
[14] C. Cortes and V. Vapnik, "Support vector networks," in *Machine Learning*, 1995, pp. 273–297.
[15] http://www.csie.ntu.edu.tw/ cjlin/libsvm/.
[16] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification." *IEEE Trans Biomed Eng*, vol. 52, no. 5, pp. 828–838, May 2005.
[17] K. Chan, T.-W. Lee, P. A. Sample, M. H. Goldbaum, R. N. Weinreb, and T. J. Sejnowski, "Comparison of machine learning and traditional classifiers in glaucoma diagnosis." *IEEE Trans Biomed Eng*, vol. 49, no. 9, pp. 963–974, Sep 2002.
[18] A. H. Khandoker, D. T. H. Lai, R. K. Begg, and M. Palaniswami, "Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly." *IEEE Trans Neural Syst Rehabil Eng*, vol. 15, no. 4, pp. 587–597, Dec 2007.