Stereotyped Gesture Recognition: An Analysis between HMM and SVM

Marcos Y. O. Camada
Instituto Federal Baiano
IFBaiano – Campus Catu
Bahia, Brazil
Email: marcos.camada@catu.ifbaiano.edu.br

Jés J. F. Cerqueira
Eletrical Engineering Department
Federal University of Bahia
Bahia, Brazil
Email: jes@ufba.br

Antonio Marcus N. Lima
Eletrical Engineering Department
Federal University of Campina Grande
Paraíba, Brazil
Email: amnlima@dee.ufcg.edu.br

***Abstract*—Stereotypic behaviours are present in both human and nonhuman primates. Usually, these behaviours are a welfare indicator. However, the stereotypic behaviours may be also a symptom of some mental disorder in the humans. A specific case is Autism Spectrum Disorder (ASD). The individuals with ASD may exhibit stereotypic behaviours through some gestures. The classic stereotyped gestures of autism are: (i) Body Rocking; (ii) Hand Flapping; and (iii) Top Spinning. This paper study the performance between two machine learning algorithms to recognition the stereotyped gestures typical of autism: (i) Hidden Markov Model [HMM]; and (ii) Support Vector Machine [SVM]. Sequence of orientations data from some joints obtained through a RGB-D (Red Green Blue -Depth) camera [Kinect®] are used for analysis. The results of these two machine learning algorithms are compared with state-of-the-art. The HMM approach proposed in this paper have shown 98.89% average recognition rate and 98.9% recall. This value is higher compared to the SVM approach and the others of art method presented.**

***Index Terms*—Stereotypic Behaviours, HMM, SVM, autism.**

I. Introduction

Stereotypic behaviours are abnormal gesture patterns of body that have no obvious function [1]. These behaviours may be present in both humans and nonhuman primates [2]. There are different reasons for stereotypic behaviours. In general, they may be related to the conditions of poor welfare due to either frustration, stress or fear experiences [1], [3]. In particular for humans, the stereotypic behaviours may be also related to neural development disorders as the autism.

Medical researches about autism have pointed out that the stereotypic behaviours may be related to either defense [3] or self-stimulation [4] mechanism due to their hypersensitivity to the environment. These behaviours increase the activation level of the autistic. An autistic can present some classic stereotyped gestures, such as: (i) Body Rocking; (ii) Hand Flapping; (iii) Top Spinning and (iv) Head-Banging [3].

Stereotypic behaviours may interfere directly with learning and socialization of children with autism. In general, traditional treatments rely on techniques of positive/negative reinforcement to mitigate these stereotypies [4]. Recently, both computer softwares [5] and robots [6] have been explored for therapeutic purposes and to helping the autistic. These technologies are part of intervention strategies to both capture their attention and encourage their social and cognitive development.

Computer softwares capable of recognizing or detecting autistic's stereotyped gestures have been useful to reveal either the level of autism severity or their engagement with therapeutic activity. Goncalves et al. [7] have proposed the use of both the Kinect® device and the Dynamic Time Warping [DTW] algorithm to detect stereotyped gesture of the hand flapping. Shamsuddin et al. [8] have analyzed the interaction between children with autism and humanoid robot aiming for the reduction of stereotyped behaviour percentage during the interactive process. Albanali et al. [9] have addressed recognition of stereotyped gesture type Body Rocking and Hand Flapping using wireless accelerometer sensors and used decision tree algorithm to classify it. They have compared also results from both laboratory and classroom environment. Rad et al. [10] have used also accelerometer on limbs of autistic, but they applied Convolutional Neural Network [CNN] to classify the stereotyped gestures [body rocking and hand flapping].

In this paper, it is presented a study about performance of machine learning algorithms for recognition of stereotyped gestures typical of autistic (hand flapping, body rocking and top spinning). Sequence of orientations data from some joints obtained through a RGB-D [Red Green Blue - Depth] camera [Kinect®] are used for analysis. The machine learning chosen are: (i) Hidden Markov Model [HMM]; and (ii) Support Vector Machine [SVM].

This paper is organized as follows. In Section II, the autism and the machine learning algorithms [HMM and SVM] are introduce. The system architecture is Summarized in Section III. The experiments with the machine learning algorithms and their results are discussed in Section IV. Finally, discussions about the contributions of this paper and future works are addressed in Section V.

II. Background

Autism Spectrum Disorder [ASD] is a neuropsychiatric disorder characterized by severe damage for the socialization and communication processes. An autistic may also have either a unusual behaviour pattern or stereotypic behaviours [11]. Novel researches have been indicated that several factors can

978-1-5090-5795-5/17/$31.00 © 2017 IEEE

be associated with the autism. Some of them are genetic, neurological anomalies and psychosocial risks [11].

Stereotypic behaviours in autistic may be related to some defense mechanism due to their hypersensitivity [3]. This type of behaviour increases the level activation of autistic [1]. The stereotyped gestures may occur independently or in pairs. Here, the gestures are considered apart in the experiments. Some stereotyped gestures that can be noted are: (i) Body Rocking [BR] – repetitive movement to forward and backward of the upper torso; (ii) Top Spinning [TS] – walk in a circle; (iii) Hand Flapping [HF] – swing motion of the hands up and down; (iv) Head Banging [HB] – hitting head on the floor or wall. The Head Banging was not considered in this papers specially because the trajectory of their movement is similar to Body Rocking.

A. Hidden Markov Model

Hidden Markov Models (HMM) are stochastic models which present two types of chains: (i) a underlying Markov chain [invisible]; and (ii) an stochastic states symbols [visible]. Since the symbol output probability distribution of a continuous HMM is given by a mixture of Gaussian, a HMM can be expressed as $\lambda = (A, c, \mu, U)$, where: A is a matrix of transition probabilities; c is a set of coefficients [weights for each Gaussian]; μ represents the averages of each Gaussian in the mixture; and U represents the covariance matrix of the Gaussian [12].

The HMM can be applied for supervised learning pattern recognition tasks. The training process of the HMM consists of the presentation of sequences of outputs [training sequences] from a particular system.

A training algorithm can adjust the HMM's parameters such that when a new observation sequence [from the system being modeled] is given as input to the HMM, the resolving of output is based on the probability of the model generated.

B. Support Vector Machines

Support Vector Machines [SVMs] are based on principles of statistical learning theory and of convex optimization, such that two-class can be classified from a set of training examples. Hence, the algorithm aims to find a suitable boundary into a data space and to separate two classes of elements [13].

Extensions of SVM have include methods for regression, clustering, factor analysis. The use of linear SVM for classification is Computationally simpler and to search a 'good' [based on some measure of performance] separating hyperplane in a high-dimensional feature space. The optimal hyperplane is the one with the maximal margin of separation between the two classes. Nonlinear SVMs are able to represent more complex functions to deal with real-world data. The SVMs have been used in various domains of application such as, text categorization, handwriting recognition, face detection, and bioinformatics [13].

III. System Overview

A Kinect® camera is used to capture frames with the stereotyped gestures. This device has a set of resources such as infrared [IR] emitter, IR depth receptor, color sensor, microphone array, and tilt motor. The first three resources were used here to get frames RGB [Red, Green and Blue] with depth information [RGB-D] about a scenery and elements contained into it [people and objects].

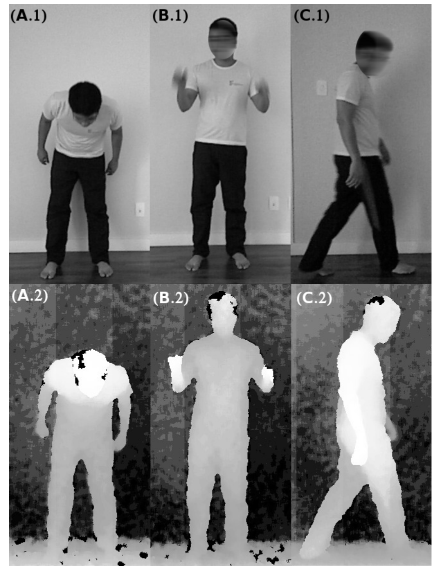


Fig. 1. Stereotyped gestures typical of autism obtained by Kinect® camera: Body Rocking (A.1 and A.2), Hand Flapping (B.1 and B.2) and Top Spinning (C.1 and C.2).

In Figure 1, it is shown some RGB-D frames of the stereotyped gestures: Body Rocking [A.1 – A.2)], Hand Flapping [B.1 – B.2], and Top Spinning [C.1 – C.2)].

In order to extract from these frames, the information about joints of one subject [actor] the OpenNI/NITE package was used. The orientation of joints is quaternion notation and after it is converted to euler notation [right side of (1), where T (defined in [14, eq.(6.3)]) is a transformation matrix]. In this paper, the joints considered are head, shoulders, elbows, trunk, hips, and knees. Hence, ten joints were used [$j = 10$].

$$\begin{bmatrix} c_1^1 & c_2^1 & c_3^1 & c_4^1 \\ c_1^2 & c_2^2 & c_3^2 & c_4^2 \\ \vdots & \vdots & \vdots & \vdots \\ c_1^j & c_2^j & c_3^j & c_4^j \end{bmatrix} T = \begin{bmatrix} x^1 & y^1 & z^1 \\ x^2 & y^2 & z^2 \\ \vdots & \vdots & \vdots \\ x^j & y^j & z^j \end{bmatrix} \tag{1}$$

IV. Experiments and Results

The Waikato Environment for Knowledge Analysis [WEKA] was used as a support tool for evaluate the performance of different machine learning algorithms from joint orientation data [15]. The WEKA provides different interfaces for the users to test and evaluate machine learning algorithms available by itself or third-party. It allows also to build workbench for the main data mining problems: regression, classification, clustering, association rule mining, and attribute selection. The files for test and training are usually either Attribute-Relation File Format [ARFF] or Comma-Separated

Values [CSV]. The performance of machine learning algorithms to recognize stereotyped gestures typically autistic are compared through experiments.

A. Methodology

Several RGB-D frames of each scenario of the stereotyped gestures were recorded repeatedly. The joint [orientation] data of an actor were extracted using the OpenNI/NITE frameworks and stored with their corresponding RGB-D frames. After that, the samples were manually extract. For each stereotyped gesture, two scenarios were simulated: high and low activation. Each scenario had 120 records with varying amounts of joint samples and also varying amounts of joint samples represented by (2), where e_i^j is euler notation from (1), j is index of a joint and n is size of a sample. Therefore, each stereotyped gesture has 240 records. A stereotyped gesture is recognized regardless of the activation level.

$$\begin{bmatrix} e_1^1 & e_1^2 & \cdots & e_1^j \\ e_2^1 & e_2^2 & \cdots & e_2^j \\ \vdots & \vdots & \ddots & \vdots \\ e_n^1 & e_n^2 & \cdots & e_n^j \end{bmatrix} \tag{2}$$

The 10-folds cross validation technique was used. In this one, the training is randomly split into 10 sets, such that nine sets are used in training and the reaming set is used for validations then another nine sets is picked and so forth.

In this paper, for each configuration of machine learning algorithms [Diagonal–HMM, Spherical–HMM, Radial–SVM, and Polynomial–SVM] is applied 25 experiments. In order to compare the performance of the machine learning, it uses two relevant statistical qualifiers: Correctly Classified Instances [CCI], and recall (measures the classifiers completeness).

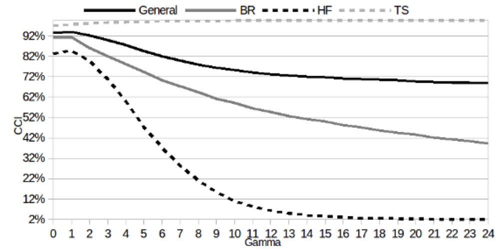
B. Experiments with SVM

The experiments with SVM have had 1437 instances. Each instance is a row of (2). The definition these experiments was organized in relation to two kernel type: (i) radial [depends on gamma variable]; and (ii) polynomial [depends on degree and gamma variables].

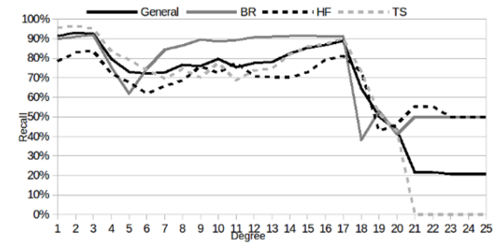
The results of SVM experiments using radial kernel are observed in Figure 2. The high recognition rate for BR [94.52% CCI and 94.5% recall] occurs with gamma equal 1. The high recognition rate for HF [85.34% CCI and 85.3% recall] occurs with 1 gamma. The best recognition rate for TS [100% CCI] occurs with value of the gamma from 12 to 24. The highest average recognition rate for SVM radial [94.52% CCI and 94.5% recall] has gamma equal 1.

In order to find the degree of polynomial SVM with the best CCI, the gamma value was set to 1 (Fig. 3a and 3b). After that, it is fixed the value of degree with the value defined in the previous step. Thus, the value of gamma is varied to determine its best value (Fig. 3c and 3d).

In the Figure 3 (3a and 3b), it can observe that the best general performance (93.11% CCI and 93.1% recall) is achieved with 2 degree (and 1 gamma). These values of degree and gamma are also the best values for general, BR and HF in variation of gamma (Figure 3c and 3d). For TS, the high recognition rate (96.84% CCI and 96.8% recall) is 6 gamma (and 2 degree).



(a)



(b)

Fig. 2. Result of radial–SVM experiments, CCI (2a) and recall (2b), for each stereotyped gestures BR, HF and TS

TABLE I
CONFIGURATIONS OF SVM FOR STEREOTYPED GESTURES WITH GAMMA/DEGREE TYPE KERNEL AND MAXIMUM CCI/RECALL

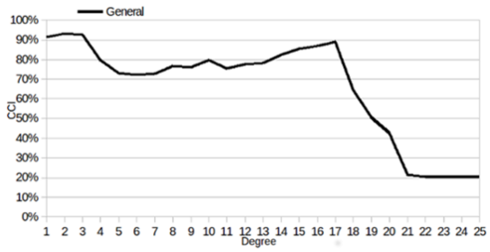
Stereotyped Gesture	**Kernel Type**	**Gamma(G)/ Degree(D)**	**Max CCI/Recall**
BR	Radial	G=1	94.52%/94.5%
HF	Polynomial	G=1;D=2	93.11%/93.1%
TS	Radial	G=12	100%
General	Radial	G=1	94.52%/94.5%

The experiment results of SVM with the best performance for each stereotyped gesture are summarized in Table I. This table categorizes these results in relation to stereotyped gesture, kernel type, (gamma/degree), and maximum CCI/recall.

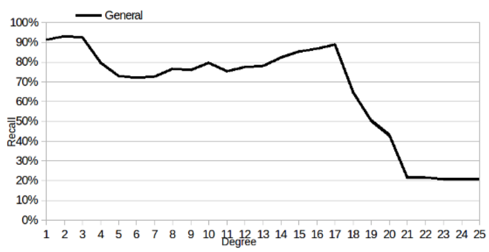
In the SVM experiments, the general and BR had the best recognition rate with radial type kernel [94.52% CCI and 94.5% recall], and 1 gamma. The HF had high recognition rate [93.11% CCI and 93.1% recall] with polynomial [1 gamma and 2 degree] type kernel. The TS had the best performance for recognition rate [100%] with radial type kernel and 12 gamma.

C. Experiments with HMM

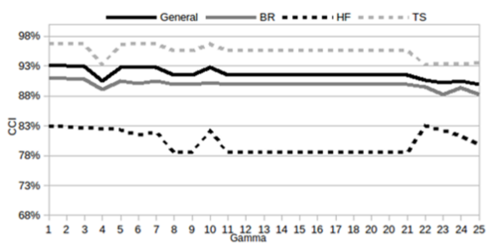
Experiments were accomplished with HMM of the type right-left and 0.01 interaction cut off. In these experiments, 720 instances have been used, which were several sequences of joint orientation extracted manually from eq. (2). The performance of HMM was analyzed for each gesture over number of HMM states and covariance type [diagonal and spherical matrices].



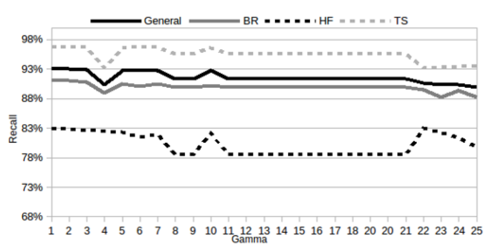
(a)



(b)

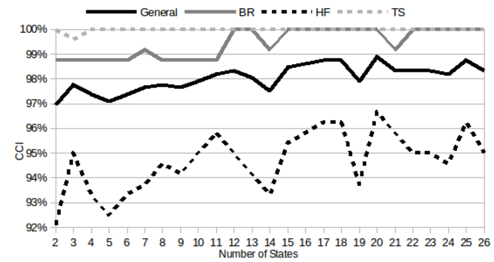


(c)

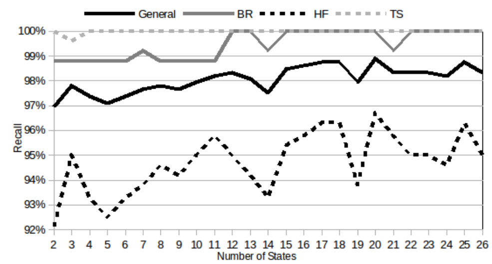


(d)

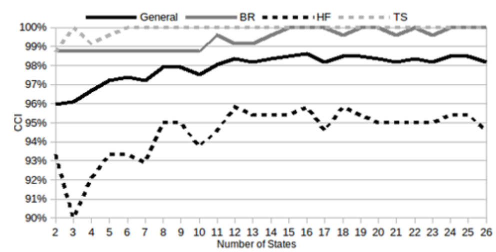
Fig. 3. Result of polynomial–SVM experiments varying in the degree, CCI (3a) and recall (3b), and in the gamma, CCI (3c) and recall (3d), for each stereotyped gestures BR, HF and TS.



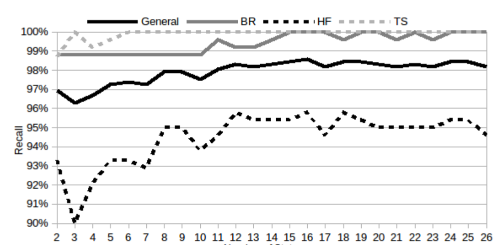
(a)



(b)



(c)



(d)

Fig. 4. Results of diagonal–HMM experiments, CCI (4a) and recall (4b), and spherical covariance (CCI (4c) and recall (4d)) types in relation to number of states for each stereotyped gestures BR, HF and TS.

Results of experiments (CCI and recall) of HMMs with diagonal and spherical type covariance varying the amount of the states are shown in Figure 4.

In experiments with diagonal covariance HMM (see Figure 4a and 4b), the BR had 100% recognition rate (both CCI and recall) with 12 and 13, from 15 to 20, and from 22 to 26 states. The highest recognition rate for HF [96.67% CCI and 96.7% recall] occurred with 20 states and TS [both CCI and recall were 100%] except for 3 states. The highest average recognition rate for HMM diagonal type covariance [98.89%] had 20 states.

Experiments with spherical covariance HMM present CCI (Fig. 4c) and recall (Fig. 4d) results. The highest recognition rate for BR [both CCI and recall with 100%] occurred from 15 to 17 states, 19 and 20 states, and from 24 to 26 states. For HF, 12 and 18 states had the highest average recognition rate [95.83%CCI and 95.8% recall]. The TS had 100% recognition rate except for 2, 4 and 5 states. The best general average recognition rate [98.61% CCI and 9.6% recall] was with 16 states.

The configurations of HMM for each stereotyped gesture with both the minimum number of states and the maximum CCI/recall are summarized in Table II. Thus, the best configuration for BR is HMM diagonal type covariance with 12 states that reaches 100% recognition rate. The HF has the highest recognition rate [96.67% CCI and 96.7% recall] with

TABLE II
CONFIGURATIONS OF HMM FOR STEREOTYPED GESTURES WITH MAXIMUM CCI/RECALL AND MINIMUM NUMBER OF STATES

Stereotyped Gesture	Covariance Type	Min. Number of States	Max. CCI/Recall
BR	Diagonal	12	100%
HF	Diagonal	20	96.67%/96.7%
TS	Diagonal	2	100%
General	Diagonal	20	98.89%/98.9%

TABLE III
PERFORMANCE FOR RECOGNITION OF STEREOTYPED GESTURES AMONG MACHINE LEARNING.

Stereotyped Gesture	HMM	SVM
BR	**100% CCI**	94.52% CCI and 94.5% recall
HF	**96.67% and 96.7% recall**	93.11% CCI 93.1% recall
TS	**100% CCI**	**100% CCI**
General	**98.89% CCI and 98.9% recall**	94.52% CCI 94.5% recall

the diagonal–HMM with 20 states. The diagonal–HMM with 2 states reaches the best recognition rate [100%] for TS. The general average recognition rate [98.89% CCI and 98.9% recall] is with the HMM spherical type covariance with 16 states.

D. Discussions

The best results for recognition of the stereotyped gestures [BR, HF, and TS] using machine learning algorithms [HMM and SVM)] are summarized in Table III. The HMM have presented better or equal (for TS) results for all stereotyped gestures than the SVM. The highest difference between them have occurred with BR [5.48% CCI]. The general performance of HMM [98.89% CCI and 98.9% recall] was also higher than SVM [94.52% CCI and 94.5% recall], with 4.37% CCI difference.

Currently, the works related to recognition of stereotyped gestures for autism are still rare and, Table IV summarize some these papers comparing all of them the target stereotyped gestures, type of sensor, machine learning algorithm used, the body feature considered and the format for data generation [actors or non-actors].

A study on the performance of stereotyped gesture [body rocking, hand flapping] recognition into different environments [between classroom and laboratory] was presented by [9].

In this paper, accelerometers were used in wrists and torso of 6 autistic children, jointly with Decision Tree for classification. The final results were not organized by stereotyped gesture, but by environment. The performance for recognition of each stereotyped gesture class was not determined. The average performance stereotyped gesture recognition for laboratory [89.5%] was better than classroom [88.6%].

A comparative study about performance between RGB-D [Kinect®] and accelerometer to detect stereotypical motor movements [hand flapping] in real time was proposed in [16] with two approaches. In their first approach, the authors have used Dynamic Time Warping [DTW] to recognize the gesture through joint coordinates data. In the second one, they have used a statistical methods [standard variation, root mean square and peaks and valleys]. Four autistic children are used in both approaches. The accuracy of the first and second approaches were respectively 51% and 76%. An extension of this previous works was also presented [7]. In this one, the work have aimed to determining ground truth of the first approach using Kinect® with regard to video analysis. The average performance is 65% ([disregarding the experiments with false-positive values].

An approach using Convolutional Neural Network [CNN] was proposed in [10]. In this one, data from accelerometer were used to classify stereotyped gesture [BR and TS] of 6 autistic. Two approaches were proposed: (i) to use cross trained dataset [transferred]; and (ii) to use only trained dataset. The accuracy of first and second approach were respectively 78% and 74%.

In this present paper, from joint orientation data obtained using the Kinect® device, HMM have presented better result for stereotyped gestures recognition [BR, HF, TS] than SVM [see Table III]. However, the others papers considered in Table IV took in account only two stereotyped gestures: BR and HF. Hence, only these two gestures were considered in the analysis between the approaches of this paper [HMM and SVM] and the approaches from others papers presented in Table IV.

HMM reaches high recognition rate 100% and 96.67% for BR and HF, respectively [see Table III]. The average recognition rate with these gesture is 98.89%. SVM reaches also high recognition rates for BR and HF gestures [94.52% and 93.11%, see Table III], and the average recognition rate is 94.52%. The performance of the approach using HMM presented here is greater than those presented in the other papers in Table IV.

The use of accelerometer sensors in autistic can be considered as the disadvantage of [9], [16], and [10]. The use of devices embedded in the body may be intrusive or inconvenient in hypersensitive autistic. Hence, a less intrusive approach [e.g. camera or infrared sensors] may be recommend for most cases of ASD (Autism Spectrum Disorder).

However, the approach presents some restrictions. The use of single non-autistic subject to test and validate the machine learning algorithms from simulated gestures has two disadvantage: (i) subtle gestural features of the autistic may not be accomplished or perceived; and (ii) the influence of both the idiosyncrasy and the artistic quality of actor become more relevant. Anyhow, the use of actors can be justified because allows to obtaining a large amount of samples for training and testing, which in the real conditions of a medical environment would be difficult. In addition, invariant body information [orientation joints] were used for all experiments.

TABLE IV
AUTOMATIC STEREOTYPED GESTURE RECOGNITION SYSTEMS. ML = MACHINE LEARNING, (L)LABORATORY AND (C)CLASSROOM ENVIRONMENT. ACC = ACCELERATION, COORD = COORDINATES, DTW = DYNAMIC TIME WARPING, CNN = CONVOLUTIONAL NEURAL NETWORK, N = NON-TRANSFERRED-CNN, T = TRANSFERRED-CNN

References	**Stereotyped Gesture**	**Sensor**	**ML Algorithm**	**Body Feature**	**Actors**	**Accuracy (%)**
[9]	BR, HF	accelerometer	Decision Tree	wrists and torso acc.	6 non-actors	89.5(L), 88.6(C)
[16]	HF	RGB-D accelerometer	DTW Statistical	joint coord. wrists and torso acc.	4 non-actors	51 76
[7]	HF	RGB-D	DTW	joint coord.	5 non-actors	65
[10]	BR, HF	accelerometer	CNN	wrists and torso acc.	6 non-actors	74(N) 78(T)

V. CONCLUSIONS AND FUTURE WORKS

In this paper, a study about the performance of some machine learning algorithms [HMM and SVM] applied to the recognition of stereotyped gestures typical of autism from joint orientation data was presented. The approach using HMM reaches a general average recognition rate of 98.89% CCI and 98.9% recall and recognition rate of 100% and 96.67% for BR and HF, respectively. For TS, the HMM has recognition rate 100%. These values are higher than all others approaches either presented or cited here. The high values of CCI and recall presented by classifiers studied in this paper (HMM and SVM) reveal a satisfactory performance from the considered samples.

Of course the idiosyncrasy of a single non-autistic subject [actor] to obtaining of samples for training and testing may have influenced the results. However, the use of actor is acceptable in the prototype phase of some projects. For instance, this allowed to collecting a greater amount of stereotyped gesture samples [with invariant feature] than an artistic.

This research is one part of HiBot Project [17], [18]. HiBot is a robotic platform for experiments of Human-Robot Interaction [HRI] in development at Robotic Laboratory of Federal University of Bahia. This research is a prototype for a HiBot's interface. In the context of the HiBot Project, the robot will acts as an assistive device to aid the treatment of autistic patients as well as others types of interaction with humans. Of course the ground truth needs to be improved for better results in the real world. Performances from either different actors or autistic can be used to minimize the influence of idiosyncrasy.

REFERENCES

[1] G. J. Mason, "Stereotypies: a critical review," *Animal Behaviour*, vol. 41, no. 6, pp. 1015–1037, jun 1991.

[2] C. K. Lutz, "Stereotypic behavior in nonhuman primates as a model for the human condition," *ILAR Journal*, vol. 55, no. 2, pp. 284–296, 2014.

[3] J. Kuhn, "Stereotypic behavior as a defense mechanism in autism," *Harvard Brain Special*, vol. 6, no. 1, pp. 11–15, 1999.

[4] A. B. Cunningham and L. Schreibman, "Stereotypy in autism: The importance of function," *Research in Autism Spectrum Disorders*, vol. 2, no. 3, pp. 469–479, jul 2008.

[5] S. Parsons, P. Mitchell, and A. Leonard, "The use and understanding of virtual environments by adolescents with autistic spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 34, no. 4, pp. 449–466, aug 2004.

[6] M. A. Goodrich, M. Colton, B. Brinton, M. Fujiki, J. A. Atherton, L. Robinson, D. Ricks, M. H. Maxfield, and A. Acerson, "Incorporating a robot into an autism therapy team," *IEEE Intelligent Systems*, vol. 27, no. 2, pp. 52–59, mar 2012.

[7] N. Gonçalves, S. Costa, J. Rodrigues, and F. Soares, "Detection of stereotyped hand flapping movements in autistic children using the kinect sensor: A case study," in *IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2014*, 2014, pp. 212–216.

[8] S. Shamsuddin, H. Yussof, S. Mohamed, F. A. Hanapiah, and L. I. Ismail, "Stereotyped behavior of autistic children with lower iq level in hri with a humanoid robot," in *Proceedings of IEEE Workshop on Advanced Robotics and its Social Impacts, ARSO*, 2013, pp. 175–180.

[9] F. Albinali, M. S. Goodwin, and S. S. Intille, "Recognizing stereotypical motor movements in the laboratory and classroom," in *Proceedings of the 11th international conference on Ubiquitous computing - Ubicomp '09*. New York, New York, USA: ACM Press, 2009, p. 71.

[10] N. M. Rad, A. Bizzego, S. M. Kia, G. Jurman, P. Venuti, and C. Furlanello, "Convolutional neural network for stereotypical motor movement detection in autism," in *5th NIPS Workshop on Machine Learning and Interpretation in Neuroimaging (MLINI)*, 2015, pp. 1–9.

[11] S. E. Levy, D. S. Mandell, and R. T. Schultz, "Autism," *The Lancet*, vol. 374, no. 9701, pp. 1627–1638, nov 2009.

[12] L. R. Rabiner and B. H. Juang, "An introduction to hidden markov models," *IEEE ASSP Magazine*, vol. 3, no. January, pp. 4–16, 1986.

[13] A. Mammone, M. Turchi, and N. Cristianini, "Support vector machines," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 3, pp. 283–289, 2009.

[14] K. Großekatthöfer and Z. Yoon, "Introduction into quaternions for spacecraft attitude representation," *TU Berlin*, pp. 1–16, 2012.

[15] E. Frank, M. A. Hall, and I. H. Witten, "The weka workbench. online appendix for "data mining: Practical machine learning tools and techniques"," *Morgan Kaufmann, Fourth Edition*, 2016.

[16] N. Gonçalves, J. Rodrigues, S. Costa, and F. Soares, "Automatic detection of stereotyped hand flapping movements: Two different approaches," in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 2012, pp. 392–397.

[17] C. M. Salcedo, C. A. Peña, J. J. F. Cerqueira, and A. M. N. Lima, "Designing a real time artificial vision system for human interaction with an omnidirectional mobile platform," in *2012 Brazilian Robotics Symposium and Latin American Robotics Symposium*, Oct 2012, pp. 21–26.

[18] M. Y. O. Camada, D. Stéfano, J. J. F. Cerqueira, A. M. N. Lima, A. G. S. Conceição, and da A. C. P. L. Costa, "Recognition of affective state for austist from stereotyped gestures," in *Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics*, vol. 1, no. Icinco, 2016, pp. 197–204.