A situation-aware system for the detection of motion disorders of patients with Autism Spectrum Disorders

Antonio Coronato, Giuseppe De Pietro, Giovanni Paragliola *

National Research Council (CNR), Institute for High-Performance Computing and Networking (ICAR), Italy

ABSTRACT

Patients with Autism Spectrum Disorders (ASD) show symptoms that in general fall into three areas: (1) social impairment; (2) communication difficulties; and, (3) repetitive and stereotyped behaviors. This paper presents a method and an infrastructure for the detection of the stereotyped motion disorders of patients with ASD. The method adopts Artificial Intelligence techniques for the identification of stereotyped motion disorders and the Situation-Awareness paradigm for the reduction of misclassifications and the extraction of further information useful for clinicians. Signals caught by accelerometers are pre-processed to obtain features that, in turn, are passed to the classifier that classifies the current temporal frame in order to detect stereotyped motions. Once a stereotyped motion is detected, events are generated for intelligent situation-aware components, which collect information related to the frequency and period of the day of the disorders, and help also to reduce false positives and misclassifications by verifying spatio-temporal constraints. Quantitatively, the off-line classifier has shown an accuracy of over 99%; whereas the on-line classifier has an accuracy of 92%. The research activity has been conducted in cooperation with clinicians of the Department of Child Psychiatry at the Children's Hospital Santobono-Pausilipon in Naples, a prototype is deployed and tested at the hospital.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The Autism Spectrum Disorders (ASD) is a group of variable neurodevelopmental disorders that first arise during childhood, and generally follow a fixed progress without remission. Manifest symptoms gradually begin after the age of six months, become established by an age of two or three years and tend to continue through childhood and adulthood. They are distinguished not by a single symptom but by a characteristic triad of symptoms: impairments in social interaction; impairments in communication; and restricted interests and repetitive behavior (Filipek et al., 1999).

Thirty years ago autism was considered to be a rare childhood disorder most often associated with severe intellectual disabilities, a lack of social awareness and the absence of meaningful expressive language (Lotter, 1967). Today, the spectrum of autism disorders is now recognized as a set of common developmental disorders, with an estimated prevalence of, for example, about in every 110 children in the U.S. (Autism, 2006). The symptoms of ASD vary from one child to the next, but in general, they fall into three areas: (1) social impairment; (2) communication difficulties; and, (3) repetitive and stereotyped behaviors. Autistic individuals show many forms of stereotyped behaviors which the Repetitive Behavior Scale-Revised (RBS-R) Lam and Aman (2007) classifies in several categories, one of which being stereotyped movements such as hand flapping, head rolling, or body rocking.

This paper presents a method for the detection of the stereotyped motion disorders and it focuses on the automatic recognition of repetitive and stereotyped behaviors, with a specific emphasis on those that are symptoms of a status of anxiety or of isolation from the surrounding environment. The proposed system uses 3D-axis accelerometer data, whose waveforms, in the case of motion disorders, show clearly identifiable patterns. The method adopts Artificial Intelligence techniques for the detection of motions disorders, thus, a specifically designed Artificial Neural Network (ANN) classifies the temporal frames of patient gestures against such patterns and generates an event whenever a temporal frame is classified as a disorder. Such events are then processed by situation-aware intelligent agents with the aim both of obtaining temporal information useful for clinicians and of reducing misclassifications. The information purged of the misclassifications helps the clinicians to automatically collect reports useful for the tracking of the patient's history of gestures and anomalous behaviors as well. The off-line classifier has shown an accuracy of over 99%; the

* Corresponding author. Tel.: +39 3394814539.
E-mail addresses: antonio.coronato@na.icar.cnr.it (A. Coronato), giuseppe.depietro@na.icar.cnr.it (G. De Pietro), giovanni.paragliola@na.icar.cnr.it (G. Paragliola).

on-line an accuracy of 92%. The system is currently deployed at the Department of Child Psychiatry at the Children's Hospital Santobono-Pausilipon in Naples, in order to validate it against a variety of patients.

2. Theoretical background

2.1. Human behavior understanding

Understanding the nature of human activities is in itself a significant research question for many disciplinary traditions, such as psychology, sociology, and ergonomics. The variety of different perspectives creates problems, since each discipline may exploit different tactics to uncover the nature of human action. Understanding how to represent human achieving activities for the purpose of intelligent environments draws upon these different traditions, and represents a significant multidisciplinary challenge (Chalmers et al., 2006). The need to understand human behavior arose in video surveillance with the aim of detecting crimes. In video surveillance, the analysis is performed in two steps: the segmentation and tracking of human beings in video sequences, and the recognition and understanding of human behavior (Xu, Tang, Liu, & Zhang, 2010). Concerning the understanding of human behavior, two kinds of approach are available: template matching, which aims to extract motion features from the given image sequence and to match such features with predefined activity patterns, and state space that aims to formulate statistical models to recognize several human behaviors. Many research efforts to understand abnormal behavior have been made in the last few years. Some approaches use an internal list of anomalous patterns against which the current behavior is compared.

The effectiveness of such approaches, however, is limited to the set of abnormal behaviors statically defined and does not scale to other unforeseen possible abnormal behaviors. Other approaches inspired by the state space model are (Xiang & Gong, 2008) and (Loy, Xiang, & Gong, 2009). In such cases, all deviations from the normal state space are considered abnormal behaviors. The major drawback of such approaches is the reduced ability to correctly classify anomalous activities and behaviors. An emerging field of application for human behavior representation and recognition techniques is Ambient Intelligence (Cook, Augusto, & Jakkula, 2009). Nater, Grabner, and Gool (2010) propose a data-driven hierarchical approach for the analysis (by means of visual scenes) of human actions in ambient assisted living scenarios. In particular, the authors propose a novel representation and analysis approach, which relies on the learning of a model of normal human behavior in an unsupervised manner. A publicly available action recognition dataset is adopted to determine the micro-actions and to identify actions. This approach, however, focuses exclusively on human motion. Indeed, no kind of interaction with objects within the environment is taken into account. Moreover, anomalous behaviors are identified with deviations from the observed sequence of micro-actions (body motions) with respect to normal models. In Monekosso and Remagnino (2010), a model-based behavior analysis system for assisted living is proposed. The monitoring of human behavior is achieved with unsupervised learning algorithms. Behavior is defined as a recognizable pattern in a sequence of events or activities and is represented by means of Hidden Markov Models (HMM). The system is able to detect and classify human behavior against a predefined subset of normal behaviors. Deviations from normal behaviors are classified as anomalies, but such a situation requires further investigation by a human in order to correctly establish the nature of the anomaly. As far as stereotyped motion disorders are concerned, a few approaches based on template matching techniques have been proposed. In Min and

Tewfik (2010), an approach based on Linear Predictive Coding is defined for the detection and characterization of behavioral patterns. In Goncalves, Rodrigues, Costa, and Soares (2012), the authors adopt the Microsoft Kinect device and gesture recognition algorithms to identify stereotyped motion disorders. On one hand, this method does not require any wearable device for the monitored user; on the other, motion disorders can be detected only when the user is located in a specific area opposite to the device. Albinali, Goodwin, and Intille (2009) and Albinali, Goodwin, and Intille (2012), have adopted wearable sensors and machine learning algorithms to identify the stereotyped motion disorders of children with autism. They have focused on two motion disorders: hand flapping and body rocking. As detailed later in this paper, our method differs from previous approaches because the results obtained by a classic classifier based on a machine learning algorithm (i.e. a neural network) are successively processed by intelligent agents that reduce misclassifications and infer further clinical information.

2.2. Situation Awareness

Situation-Awareness enhances the concept of the context (Lee, Lunney, Curran, & Santos, 2009) by including rich temporal and other structural aspects: time of day, a situation may only happen at a particular time of the day; duration, it may only last a certain length of time; frequency, it may only happen a certain number of times per week; and sequence, different situations may occur in a certain sequence (Ye, Dobson, & McKeever, 2011). As suggested by Ye et al. (2011), we should distinguish between Context and Situation-Awareness. Indeed, the authors propose the following definitions: the Primary Context is the full set of data caught by real and virtual sensors; the Secondary Context focuses on semantic information inferred and/or derived from the fusion of several data streams (primary contexts) an important kind of secondary context consisting in activities performed within the environment; a situation is, instead, an abstract state of affairs of interest for designers and applications, which is derived from the context and a hypothesis about how the observed context relates to factors of interest. In this paper we use the Artificial Neural Network to identify a piece of the secondary context, that related to special physical activities identified as motion disorders, and intelligent Prolog agents that exploit the Situation-Awareness paradigm to reason on temporal constraints in order to reduce misclassifications and obtain further information useful for the clinician.

The situation model has been described by using Situation Calculus (SC). The basic SC was designed by McCarthy (1963) and has been adopted to model dynamically changing worlds. Three basic kinds of SC are: *Actions*, which can be performed in the world and can be quantified; *Fluents*, that describe the state of the world (these are predicates and functions whose value may change depending on the situation); and *Situations*, which represent a history of action occurrences. A dynamic world is modeled through a series of situations as a result of various actions being performed within the world. The constant S0 denotes the initial situation; whereas, $do(a, S)$ indicates the situation resulting from the execution of the action a in situation S. The dynamic world is axiomatized mainly by adding initial world axioms, effect axioms, and successor state axioms (Reiter, 2001). The initial world axioms describe the initial status of the environment, its objects, their position, their properties, etc. An effect axiom, instead, describes the effect on a fluent caused by the execution of an action in a specific situation. There must be specified for each fluent also the non-effect of the other actions. Intuitively, it is possible to state that a fluent's truth value is true after executing an action a if, and only if, the action has the effect of making the fluent true or, the fluent was already true before executing a and the action does not have the effect of making

it false. In such a case, a successor state axiom must be formalized. Our approach requires that a basic action theory be defined to represent and reason on motion disorders. As an example, a fluent such as *isHandFlapping*, may be defined of representing the situation related to the motion disorder hand flapping. This takes place as long as the temporal frames under observations are classified by the ANN according to the pattern defined for such a kind of motion disorder. It is important to note that, while the ANN classifies a series of temporal frames as different occurrences of the same motion disorder (e.g. Hand Flapping), the intelligent agent is able to recognize that such a sequence represents a unique occurrence of a motion disorder; that is, a unique clinical event. Another approach based on the Situation-Awareness mechanism has been adopted in Coronato, De Pietro, and Esposito (2006) to detect and handle abnormal behaviors in a monitored environment.

3. Process

Succinctly, we have defined an approach and developed tools to support the process shown in Fig. 1. A software component called a *Data Acquirer* collects primary context information from the patient by means of accelerometers. After that, a *Pre-processor* extracts statistical features from the sampled signals. The *ANN Classifier* classifies the temporal frames using statistical features and generates events for four abnormal statuses: Hand hitting against the Ear (HE), Arm Flapping (AF), Hand Rotation Up (HRU) and Hand Rotation Down (HRD). The *ANN Classifier* provides secondary context information by means of such events that indicate time-by-time the current gesture. Recognizing correctly each anomalous status is the main feature of our system in order to provide clinicians with accurate daily reports about all the anomalous statues performed by the patients, as well as information such as frequency and duration which is useful for monitoring their behaviors. Such reports also help clinicians to understand the patient's progress by highlighting and noting any changes in their behaviors since the start of their treatment. From this point of view false positives or misclassification errors are significant problems because they involve uncorrected reports which show a possible anomalous status which in fact never happened. Unfortunately, we have experimentally verified the occurrence of different kinds of problem with the sequences of events generated by the *ANN Classifier*. The most frequent are:

- M1 – spurious *misclassifications*, that are false positives related to one singular temporal frame misclassified as an abnormal status when the gesture is normal;
- M2 – *misclassifications* between Hand Rotation Up and Hand Rotation Down which are gestures with similar statistical features;
- M3 – misclassifications due to short pauses in the motion of the patient. Indeed, in the case of a short pause during the abnormal gesture, a regular classifier such as the *ANN Classifier* or others presented in the related work would treat this as two distinct disorders, but from a clinical point of view it is related to the same event.

The Situation-Awareness paradigm may be useful in reducing such misclassifications giving the possibility of expressing and verifying spatio-temporal constraints on the sequence of actions (events). Therefore, we have defined and realized two subsystems called the Real-Time Advanced Classifier (RTAC) and the Near Real-Time Advanced Classifier (NRTAC). It is worth clarifying that by real-time classifier we mean the system that is able to filter spurious events (M1) and that introduces the minimum delay in the detection. This subsystem is required to provide the earliest detection of abnormal behaviors, whatever they are. The NRTAC, instead, is required to provide the most correct and precise clinical information for the illness assessment; thus, it must be able to correct errors related to spurious misclassifications (M1), cases of misclassifications between HRU and HRD (M2) and short motion pause (M3) although this requires reasoning on a longer sequence of events generated by the ANN Classifier and then causes a slightly longer delay in the classification process. In addition to this, we have assigned to the NRTAC subsystem the responsibility of providing further clinical information about the duration of the motion disorder (Duration Agent) and its frequency in different periods of the day (periodOfDay Agent).

4. Methods

4.1. Data acquisition

In order to acquire motion data, we apply an eZ430-Chronos accelerometer (Fig. 2) to the patient's wrist. The eZ430-Chronos is an integrated, wireless development system that provides a
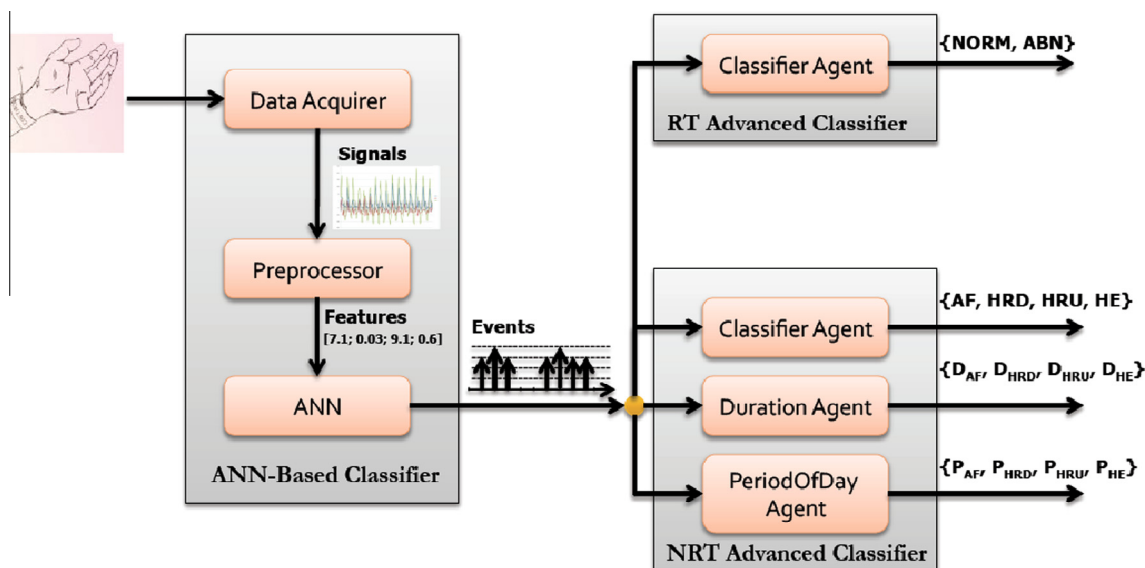

Fig. 1. Workflow of the recognition process.

complete reference design for developers creating wireless smart watch, personal displays for personal area networks, wireless sensor nodes for remote data collection and other applications. It includes an integrated pressure sensor and a 3D-axis accelerometer for motion sensitive control. The device also offers temperature and battery voltage measurements (estimated battery life ranging from 2 days up to several months depending on the operating mode) and is complete with a USB-based CC1111 wireless interface to a PC (wireless range up to 10 m). Such a device comes with a DLL control driver; although in fact, we used the RXTX library (Rxtx library). RXTX is a native library providing serial and parallel communication for the Java Development Toolkit (JDK).

In order to better highlight the repetitive trend of anomalous gestures, we plot 3-D (*X*, *Y*, *Z*) values extracted from the accelerometer sensors at the same time as the anomalous gesture is taking place. The 3-D values are plotted in a *X–Y* graphic where on the horizontal axis (*X*-axis) the time *t* (seconds) is recorded and on the vertical axis (*Y*-axis) is recorded the value of (*X*, *Y*, *Z*).

The values measured with the accelerometer as shown in Figs. 3(a) and 6(a), are the accelerations along the 3 axes (*x*, *y*, *z*) in the range of [−2g, +2g]. These values were coded in integer values in the range [−2500, +2500] to allow us to apply the proposed statistical features for the gesture detection.

Due to the technological limit fixed by our own hardware we could not measure other values such as velocity and position.

The **blue** line is the *Y*-value, the **red** line is the *Z*-value and the **green** line is the *X*-value. As demonstrated in Figs. 3–6(a), in the case of motion disorders, signals show distinctive statistical features and repetitive patterns. It is clear, indeed, how the repetitiveness of the waveform grants specific peculiarities from a statistical point of view, in terms of high frequency, zero crossing, etc. Such examples represent the patterns of the motion disorders.

In Figs. 3 and 6(b), a screen-shoot of the motion patterns of each gesture is shown in order to help the reader to understand better both the monitored gesture and the procedure described in the text.

The sensor has been circled in order to highlight its position on the patient's body. It is worth to note that although the screenshoots show the gestures mimicked with the right hand, all the gestures can be performed with both hands. The system works in both cases.

4.2. Feature extraction

The signals are pre-processed in order to extract statistical features useful for classification. We decided to adopt this approach because it is widely used for the detection of human motions. The literature proposes some works which prove the effectiveness of the method (Borges, 2013), Chen, AlRegib, & Juang, 2013). We also used statistical features such as the average and the variance to compare different kinds of Artificial Intelligence algorithms in order to select the ones most compliant with our needs. In particular, the following features are calculated:

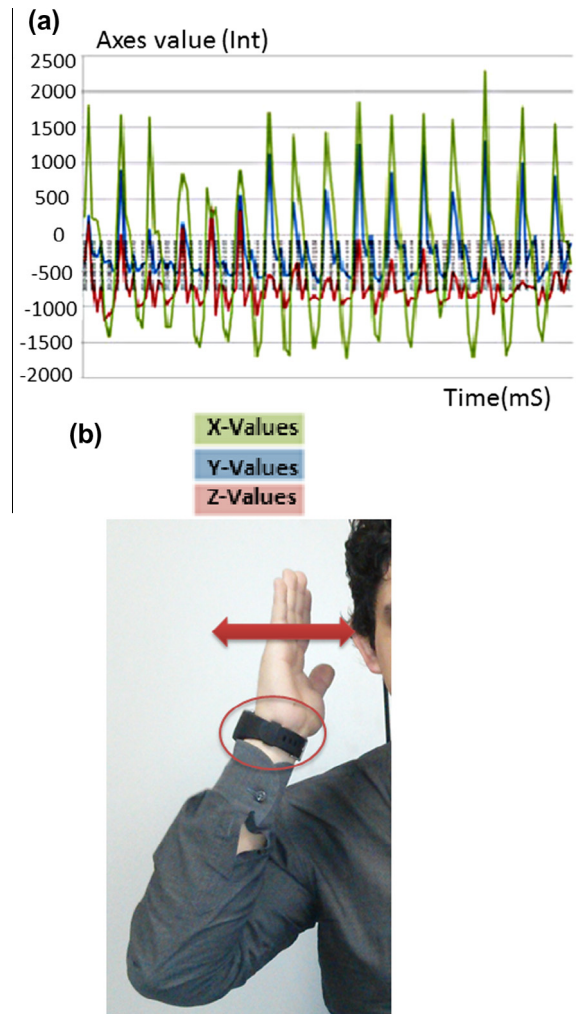**Fig. 2.** eZ430-Chronos accelerometer.

Fig. 3. (a) Hand Against Ears wave form and (b) gesture screen-shoot.

- The *Mean Absolute Value* (MAS), which is obtained from the average of the absolute value of each signal.
- The *Root Mean Square* (RMS), also known as the quadratic mean, which is a statistical measure of the magnitude of a varying quantity.
- The *Variance* (VAR), which is a measure of how far a set of numbers is spread out. It is one of several descriptors of a probability distribution, describing how far the numbers lie from the mean (expected value).
- The *Standard Deviation* (SD), which shows how much variation or dispersion exists from the average (mean, or expected value). A low standard deviation indicates that the data points tend to be very close to the mean, whereas a high standard deviation indicates that the data points are spread out over a large range of values.
- The *Waveform Length* (WL), which is the cumulative length of the waveform over the time fragment.

The WL is related to the waveform amplitude, frequency and time

- The *Zero Crossing* (WZC), which is the number of times that the amplitude value of the signal crosses the zero *y*-axis.

All these features are computed for each (*X*, *Y*, and *Z*) signal. In addition to this, we also computed the covariance among such
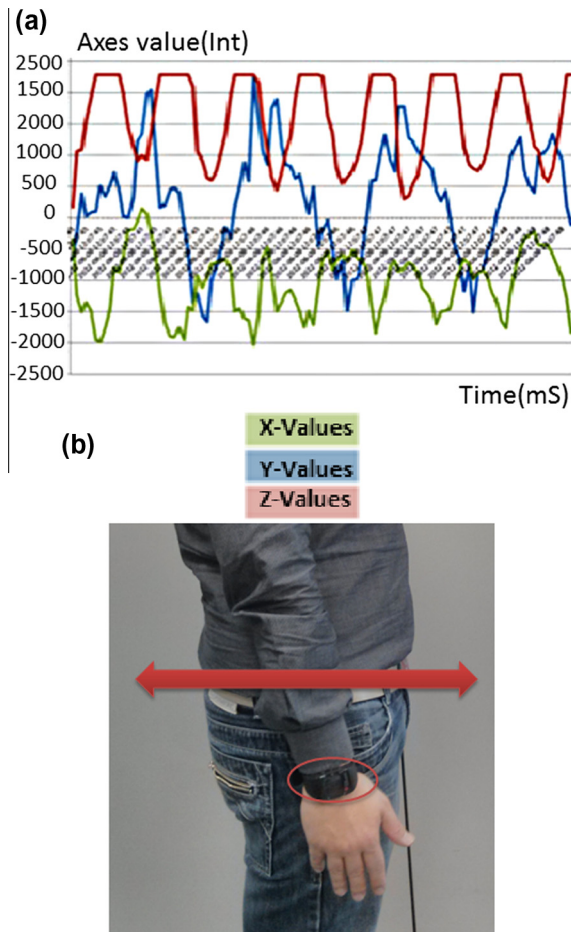
Fig. 4. (a) Hand flapping wave form and (b) gesture screen-shoot.

Fig. 5. (a) Hand Rotation Down wave form and (b) gesture screen-shoot.

signals, which, however, has not proved to be relevant for classification. Further information on the statistical feature extraction may be found in Guyon, Gunn, Nikravesh, and Zadeh (2006).

4.3. ANN Classifier

The next step after the features extraction was to evaluate which Artificial Intelligence algorithms was most suitable for our purposes. In our previous work we performed tests of different classic classical Artificial Intelligence algorithms in order to evaluate their performance with our own ground truth. It is worth explaining how we defined the ground truth for the training and testing. To do that, an expert clinician mimicked one-by-one the four anomalous gestures that we recognize. For each gesture we defined a data-set used for the training.

In this section we introduce two concepts, the off-line classification and the on-line classification the difference between the two being that the off-line classification performed the classification by taking the gesture movement data from the sensors placed on the wrist of the author who had mimicked the gestures, whereas the on-line classification took the data from the sensors placed on a patient's wrist during a medical examination.

In order to identify motion disorders, we evaluated three classifiers based on Artificial Intelligence algorithms: Neural Networks, Naive Bayes and Bayesian Networks. We focused our attention on these three technologies because they have been used successfully in previous works on ambient intelligence (Lin, Wang, & Lin, 2008), (Mozer, 1998),Osborne, Roberts, Rogers, & Jennings, 2012. Experimental results have shown better classification performances
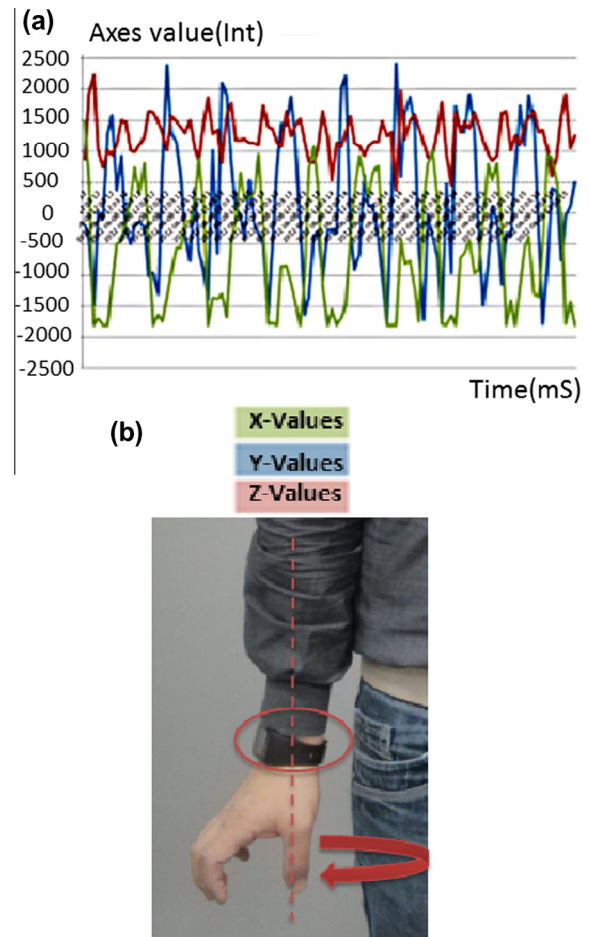
for the neural network, whereas the other classifiers show a faster training. For more information about the results, the experiments, etc. we suggest reading the following work (Coronato & G. D., 2012).

The Table 1 shows the performance of the three methods in terms of accuracy and error. The three methods were trained with the same training data-set with a size equal to 705 temporal frames. As we can see, the NN shows the better classification with an error less than 1%. For this reason we decided to adopt the NN for the classification of the gestures, although it showed a training process slower than the other methods.

The reader can note that all the methods show very good results over 95%. This is due both to a strong statistical characterization of the anomalous gestures and to the fact that the evaluation was performed using a sub-set of the training dataset (20% of whole).

Due to the need for accuracy in the classification process in order to avoid misclassification errors and false positives, we chose to adopt the neural network approach in preference to the others, taking into consideration the performance results obtained in the laboratory. Cross validation was performed with two values of the K-Fold (Stone, 1974).

The off-line experiments allowed us to choose the best Artificial Intelligence algorithm for our purposes and to discover the occurrence of the misclassification errors (M1, M2, M3). Based on the off-line results, we realized an on-line classifier using WEKA libraries (Holmes, Donkin, & Witten, 1994). The training of the on-line classifier followed the same process as the off-line one.

At first, the anomalous gestures were mimicked in order to acquire raw data from the sensor; these mimicked gestures were
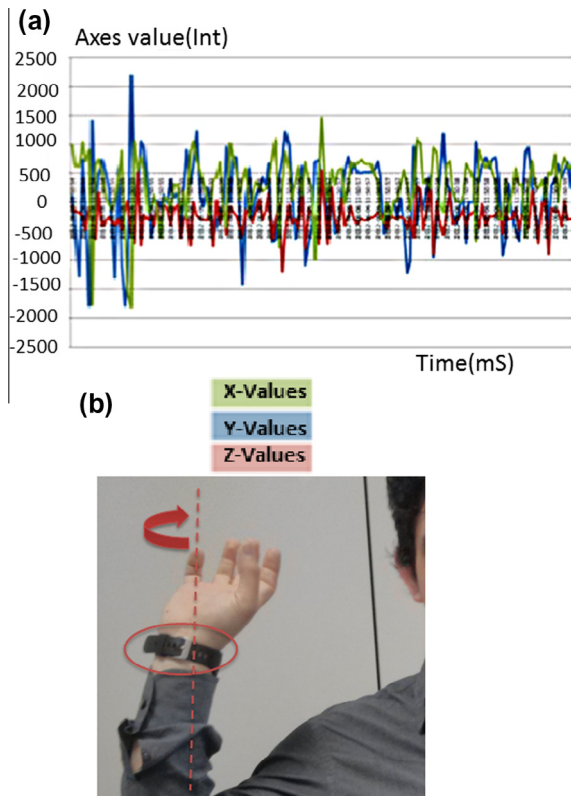
Fig. 6. (a) Hand Rotation Up wave form and (b) gesture screen-shoot.

Table 1
Performance comparison of the three methods.

	Neural network	Bayesian networks	Naive Bayes
Accuracy	99.127%	97.3%	98.15%
Error	0.873%	2.7%	1.85%

performed by a clinician who is an expert about the motion disorders associated to autism. This step produced thousands of raw data samples in the form of (day-time, x-axis, y-axis, z-axis).

Next, we aggregated the raw data samples in temporal frames each of which lasted two seconds. The statistical features described in the paper were extracted frame-by-frame. The result of this step is a tuple like this (MAXx, MASy, MASz, RMSx, RMSy, RMSz, VARx, VARy, VARz, SDx, SDy, SDz, WLx, WLy, WLz, WZCx, WZCy, WZCz).

All tuples were labeled as AF, HE, HRU, HRD, NORM according to the anomalous stereotyped gestures presented in Section 4.3.

All tuples define the training dataset of the on-line classifier with a total size of 705 occurrences.

For the classification we set the software module that acquires data from the sensors to collect 32 samples per second and considered temporal frames of two seconds for the classification, with one second of overlap frame-by-frame. Each second the classifier produces one event of one of the following types:

- AF – (repeated) Arm Flapping;
- HE – (repeated) Hand against Ear;
- HRU – (repeated) Hand Rotation Up;
- HRD – (repeated) Hand Rotation Down;
- NORM – a Normal gesture

In order to evaluate the performance of the neural network we performed an experiment in our lab. Table 1 reports the performance and accuracy of the neural network; such results were

achieved during off-line experiments by monitoring one of the authors who performed both normal (NORM) and anomalous gestures (ANOM). In particular, it is possible to see that the neural network reports seven false negatives and two false positives out of one thousand samples. The resulting accuracy is over 99%. This result supports the choice of the neural network. The entire designing process of the ANN network was made using the KNIME tool (Silipo & Mazanetz, 2012), an open frame-work for the analysis process, data transformation and reporting; the Fig. 8 shows the ANN parameters. It is worth note that the parameter *HiddenLayers* set to a means that the number of hidden layers is $\frac{number\ of\ attributes + number\ of\ classes}{2}$, tributes is 18 (RMSx, RMSy, etc.) and the number of classes is 5 (NORM, HRU, HRD, AF, HE), the Fig. 7 shows the ANN's structure. The other parameters such as LearninggRate, Momentum were used to calibrate the BackPropagation algorithm for the training of the neural network (Riedmiller & Braun, 1993).

5. Intelligent Prolog agents

The reason for the choice of SC over other methods of planning such as the Planning Domain Definition Language (PDDL) McDermott, 1998 lies in the consideration that SC is an event-driven model for reasoning over a sequence of events. This feature is suitable for the sequence of events provided by the ANN. The sequence of events is used to model changes in patient behaviors and recognize motion disorders; in SC such changes are described in terms of fluents and properties whereas in PDDL and other methods of planning, they are described in terms of the effects of the events themselves, an approach more complex than SC (Eyerich, Nebel, Lakemeyer, & Classen, 2006).

As said before, we have defined two sub-systems, RTAC and NRTAC, which both model changes in patient behaviors and reduce the misclassifications errors. The first one works as fast as possible without providing any kind of temporal information, while the latter, NRTAC, also provides temporal information such as duration and period of day. However, NRTAC recognizes the anomalous gestures slower than RTAC.

We would remind the reader that we define such agents in order to remove the misclassification errors (M1, M2, M3) that we have experimentally verified with the ANN.

In this section, we describe the intelligent agents which the RTAC and the NRTAC components include. The specifics of each agent is implemented using SC formalism. The input of both components is the sequence of events generated by the ANN by classifying the patient's movements. Each event of the sequence is sent to the agents which reason according to their specifics in SC.

5.1. Primitive actions

As introduce before, our agents reason over a sequence of events generated by the ANN so first we have to define such kinds of events in SC formalism first. Axiom 1 introduces the set of actions (events) of interest. Among such actions, *resetCount* and *updateGesture* are endogenous and are executed for technical reasons (i.e. to reset a counter and update the total duration of the motion disorder). The others (HE, AF, HRD, HRU, and NORM) are those performed by the patient and detected by the ANN Classifier as motion disorders; the names are the acronyms of the type of motion disorder.

$$a \in \{\mathbf{HE(t)}, \mathbf{HRU(t)}, \mathbf{AF(t)}, \mathbf{HRD(t)}, \mathbf{NORM(t)}, \mathbf{resetCount(t)}, \mathbf{AN(t)}, \mathbf{updateGesture(t)}\} \quad (1)$$

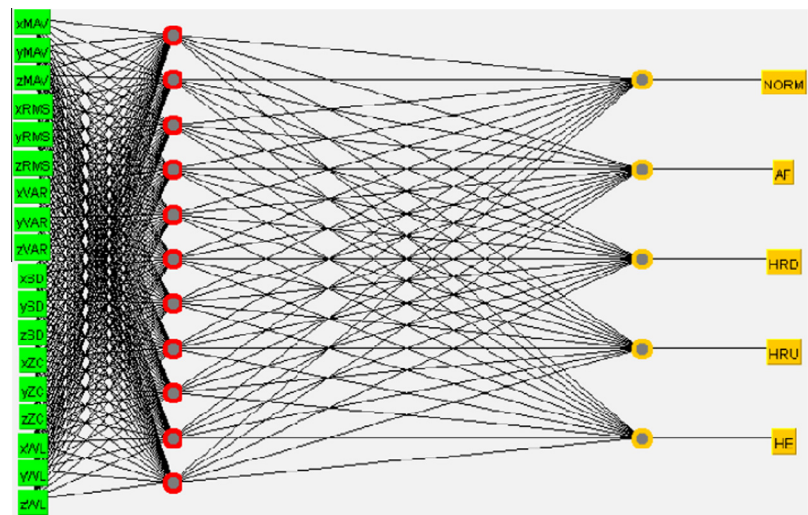where the t variable means the time of the occurrence of the event.
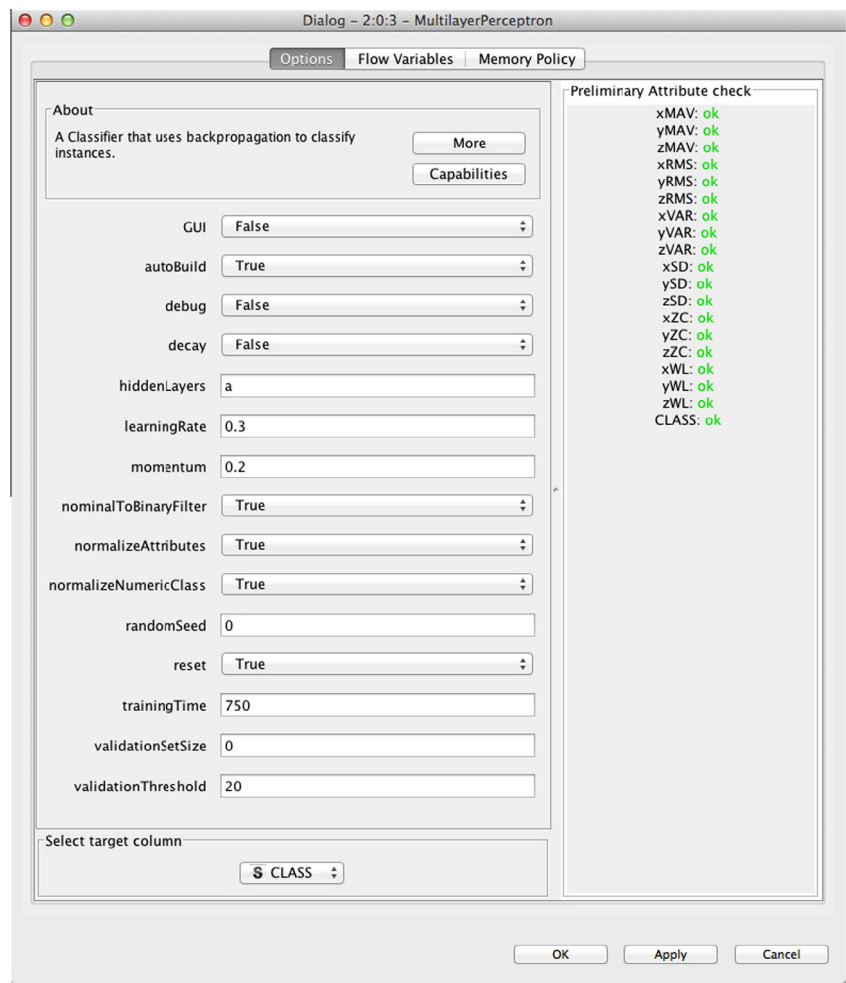
Fig. 7. ANN networks.

Fig. 8. ANN parameters.

5.2. Fluents and situation-independent predicates

In this subsection we describe the specifics of the agents to model the changes in the patient's behaviors and the anomalous gestures.

It is possible to define the periods of the day of interest by means of a predicate such as the one formalized by functional axiom 2. In this case, we have defined three periods: morning, afternoon, and evening with actual values, the variables which define them being: tS and tF two thresholds that define a temporal

window for classifying periods of the day and p, a variable that returns the period of day according to the temporal window.

$$periodOfDay(\mathbf{tS}, \mathbf{tF}) \equiv \tag{2}$$
$$\{(P = Morning \land tS = 6:00 \land tF = 12:00)$$
$$\lor (P = Afternoon \land tS = 12:01 \land tF = 18:00)$$
$$\lor (P = Evening \land tS = 18:01 \land tF = 24:00)\}$$

For the sake of brevity, we report only a subset of the situation-dependent fluents including those for the HRU disorder. However, it is quite easy to figure out that similar fluents have been defined also for the other disorders.

$$isHRU(\mathbf{S}); \tag{3}$$
This fluent denotes a motion disorder of type HRU

$$isGesture(\mathbf{S}) \equiv p; \tag{4}$$
This denotes the period of day within which the gesture is performed

$$HRUDuration(\mathbf{S}) \equiv d; \tag{5}$$
This functional fluent denotes the duration of the current disorder HRU

$$isEndDay(\mathbf{S}); \tag{6}$$
This fluent resets all counters at the end of the day

$$start(\mathbf{S}); \tag{7}$$
This fluent denotes the starting time of situation S

The variable S, that is common to all the fluents, describes the current situation, the p variable of the functional fluent $isGesture(S)$ describes the period of day within which the patient performs an anomalous gesture, while the d variable of $HRUDuration(S)$ describes the duration time of the anomalous gesture.

A specification of successor state axioms for some of the previous fluents is now presented, which allows us to explain the functioning of some mechanisms. In particular, let's suppose that the ANN Classifier generates the sequence of events **{NORM, NORM, HRU, HRU, HRU, HRU, NORM, NORM}**, which corresponds to a motion disorder of type HRU and a duration of four seconds. In this case, both fluents *isHRU* and *isGestureHRU* change their truth values to true.

Both *isGestureHRU* and *isHRU* are relation fluents so they return Boolean values, the difference between them being the following: the first fluent triggers whenever an event such as an HRU is generated from the ANN so that for each coming HRU event, its value is true whereas at the first occurrence of a non-HRU event (e.g. a NORM event or a pause between two HRU events), its value changes to false. It basically maps one-by-one the occurrence of each HRU event generated by the ANN in an agents's truth values; *isHRU* triggers when a sequence of HRU events is recognized instead of one event at time so it is also able to filter spurious misclassifications which *isGestureHRU* cannot do. However, the fluent *isHRU* keeps the value true until the first event not equal to an HRU and indicates a situation of motion disorder HRU that is taking place. The fluent *HRUDuration*, instead, computes the duration of the current HRU disorder.

$$isHRU(\mathbf{do}(\mathbf{a}, \mathbf{S})) \equiv (a = HRU(t)$$
$$\lor isHRU(S)) \land$$
$$\neg\{a = NORM(t) \lor a = HRD(t) \lor a = HE(t) \tag{8}$$
$$\lor a = AF(t)\}$$

$$isGestureHRU(\mathbf{do}(\mathbf{a}, \mathbf{S})) \equiv$$
$$a = HRU(t) \land t \in periodOfDay(tS, tF, p) \tag{9}$$
$$\land \neg(isHRU(S))$$

$$isHRUDuration(\mathbf{do}(\mathbf{a}, \mathbf{S})) \equiv$$
$$a = HRU(t) \land start(S) = tStart \land \neg(isHRU(S))$$
$$\lor isHRUDuration(S) \land (a = HRD(t) \land a = HE(t) \tag{10}$$
$$\lor a = AF(t) \lor a = NORM(t)) \lor start(S) = tFinish \lor isHRU(S)$$
$$D = \|tFinish - tStart\|$$

In addition to this, let's now consider the following sequence: **{NORM, NORM, HRU, HRU, HRU, HRU, NORM, HRU, HRU, NORM, NORM, NORM}**, which corresponds to an HRU disorder with a short interruption in the middle. In this case, the prolog agent computes two different clinical events related to the two subsequences. From the clinicians point of view, this is an error because two or more sequences with a short pause in the middle should be consider as a single clinical event *(M3)*. To avoid this misclassification, we have modified the fluent *isHRU* to let it check a temporal constraint between two or more consecutive HRU disorders. In particular, if the temporal distance is less than a certain threshold, the two sequences are considered part of the same clinical event.

$$isHRU(\mathbf{do}(\mathbf{a}, \mathbf{S})) \equiv (a = HRU(t) \land start(S) = LastGesture$$
$$\lor (isHRU(S)) \land$$
$$\neg\{(a = NORM(t) \lor a = HRD(t) \lor a = HE(t) \tag{11}$$
$$\lor a = AF(t)) \land start(S) = tCurrent$$
$$\land \|LastGesture - tCurrent\| > \Delta\}$$

The Agents have been realized within ECLiPSe, an environment for the development and deployment of the high-level modeling and control programming of Prolog agents. Interactions between the Java-based ANN Classifier and Prolog agents have been realized by means of the Java-Eclipse (JE) interface. JE is a set of API embedded in the ECLiPSe kernel library for the Java Development Toolkit, which provides methods for the iteration with the agents.

In order to explain how the agents are used and how they can be helpful for clinicians, we provide a case study scenario as shown in Fig. 9. A sequence of AF events is generated by the neural network, which means that an anomalous gesture is occurring. The short pause in the middle of the sequence, the red event NORM, causes the neural network to perform a misclassification because two distinct anomalous gestures (G1 and G2) are recognized instead of a single clinical event (an anomalous behavior). This kind of misclassification is avoided by submitting the sequence to the Prolog agents. The *isAF* agent detects only one gesture, the *AFDuration* agent returns the duration time ($\Delta_T = 15$ s), and the *PeriodOfDay* agent returns the period of day in which the gesture happens. Each AF event is sent to the agents which change their truth values according to their specifics. When the sequence finishes an opportune report about the anomalous gesture is generated and made available to clinicians.

6. Results

A prototype of this system is currently deployed in a hospital setting; clinicians have made available a living room at the hospital with a workstation that hosts all the components of the system such as the database and agents, etc. Currently the system is able to capture the data stream from one patient at a time so clinicians and caregivers make daily observations of one child at a time. Whenever a medical examination finishes, the accelerometer is removed from the child's wrist and the report about the observation is saved for future consultations. If a new observation begins, clinicians will place a sensor on the patient's wrist and re-start the system.

In order to evaluate the accuracy of the ANN Classifier, we performed an experiment involving one patient. Table 2, reports
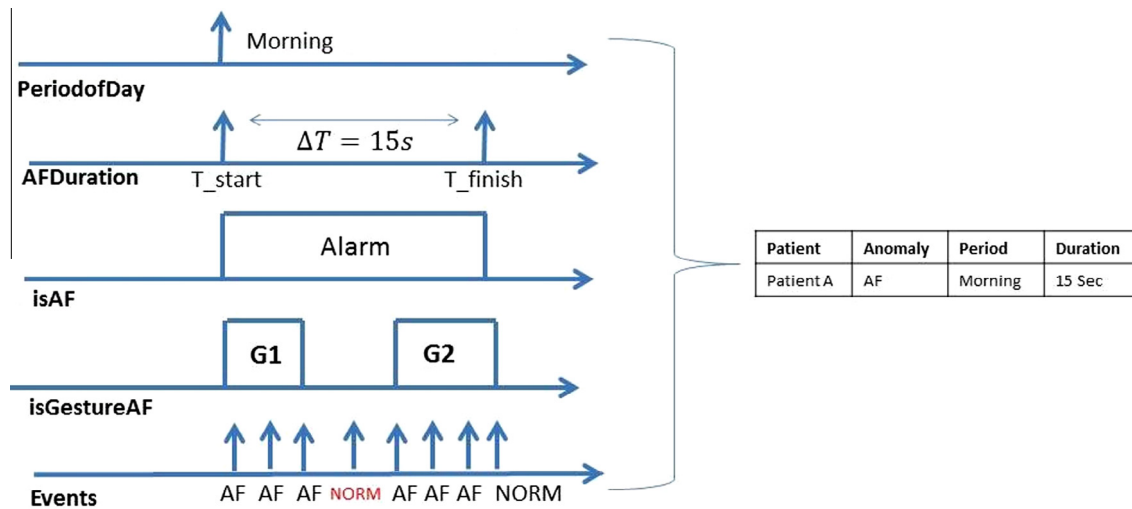
Fig. 9. Use case scenario.

the performance of the on-line ANN Classifier over more than 700 temporal frames classified.

The data belong to the patient who performed the anomalous gestures during a daily observation. In detail, we had 11 false positives and two other misclassifications out of 705 temporal frames. It is important to note that such a result is related to the on-line classifier that which has been trained with gestures mimicked by clinicians. In the hospital setting we have observed for the on-line ANN Classifier a worse performance, of course, with an accuracy near to 92% (see Table 3). This result has been obtained by the ANN Classifier matching real patient's gestures with patterns of motion disorders mimicked by clinicians. We are now collecting data from other real patients, which are continuously being validated by clinicians, in order to build a larger training data-set that will improve the performance of the online ANN Classifier. In addition to this, however, it is important to report the corrections and other clinical in-formation due to the intelligent agents (RTAC and NRTAC).

In order to evaluate the performance of the agents we performed a second experiment.

Tables 4 and 5, indeed, reports the results of the monitoring in the hospital of four different patients in four daily observations. The intelligent Prolog agents have identified two motion disorders of type HE for one patient (Table 4) and two of type HE and AF, respectively, for another patient (Fig. 5). The other patients did not show any motion disorders during the observation. Let us focus on the disorder of type AF. This was the result of a sequence of seven temporal frames classified as AF by the Artificial Neural Network. The disorder lasted for seven seconds. It is finally useful to note that the ANN Classifier detected one spurious event classified as an HDR for the first patient. Such a misclassification was purged by the Advanced Classifiers.

Table 2
Off-line performance and accuracy of the neural network.

	NORM	ANOM
Off-line confusion matrix		
NORM	879	2
ANOM	7	143
Off-line results		
Accuracy	99.127%	
Error	0.873%	
Wrong classified	9	

Table 3
On-line performance and accuracy of the neural network.

	AF	HRD	HRU	HE	NORM
On-line confusion matrix					
AF	15	0	0	0	0
HRD	0	8	0	1	0
HRU	0	0	16	0	0
HE	0	0	1	9	0
NORM	7	1	1	2	644
On-line results					
Accuracy					92.156%
Error					7.844%
Wrong classified					13

7. Discussion

In this paper we have presented an approach to detect stereotyped motion disorders in children with Autism Spectrum Disorders. We have adopted a template matching technique and compared the performance of different classifiers. Artificial Neural Networks have shown a slightly better performance during off-line motion disorder recognition. We have realized an on-line recogni-

Table 4
Results of four daily observation – Patient A.

	ANN Classifier	Advanced Classifier							
		Day		Morning		Afternoon		Evening	
	N	N	Δ	N	Δ	N	Δ	N	Δ
HE	15	2	15 s	2	15 s	0	0	0	0
HRU	0	0	0	0	0	0	0	0	0
HRD	1	0	0	0	0	0	0	0	0
AF	0	0	0	0	0	0	0	0	0

Table 5
Results of four daily observation – Patient B.

	ANN Classifier	Advanced Classifier							
		Day		Morning		Afternoon		Evening	
	N	N	Δ	N	Δ	N	Δ	N	Δ
HE	9	1	9 s	0	0	1	9 s	0	0
HRU	0	0	0	0	0	0	0	0	0
HRD	0	0	0	0	0	0	0	0	0
AF	7	1	7 s	0	0	1	7 s	0	0

tion system, which classifies temporal frames against four types of motion disorders. Such classification results have been next processed by intelligent Prolog agents that are responsible for (i) identifying episodes of motion disorders; (ii) constructing temporal statistics and other information useful for the clinician; and (iii) reducing misclassifications by applying temporal constraints, the results show as the Agents remove misclassification errors and improve the quality of the detection.

The preliminary laboratory results have encouraged the adoption of such an approach in real case studies. For this reason, we have deployed the monitoring system at the Department of Child Neuropsychiatry at the Santobono-Pausillipon Children's Hospital, where we are collecting data from a large set of patients. The final objective, however, is to realize a full context-aware environment (Coronato et al., 2006), which should be able to detect patient's anomalous behaviors and their causes (e.g. actions performed within the environment by the patient or other users).

References

Albinali, F., Goodwin, M. S., & Intille, S. S. (2009). *Recognizing stereotypical motor movements in the laboratory and classroom a case study with children on the autism spectrum*.

Albinali, F., Goodwin, M. S., & Intille, S. S. (2012). Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms. *Pervasive and Mobile Computing*, 103–114.

Autism and developmental disabilities monitoring network surveillance year 2006. In *Prevalence of autism spectrum disorders-Autism and Developmental Disabilities Monitoring Network*, MMWR Surveillance Summary, Centers for Disease Control and Prevention, 2006.

Borges, P. (2013). Pedestrian detection based on blob motion statistics. *IEEE Transactions on Circuits and Systems for Video Technology, 23*(2), 224–235. http://dx.doi.org/10.1109/TCSVT.2012.2203217.

Chalmers, D., Chalmers, M., Crowcroft, J., Kwiatkowska, M., Milner, R., ONeill, E., Rodden, T., Sassone, V., Sloman, M. 2006. Ubiquitous computing: Experience, design and science.

Chen, M., AlRegib, G., & Juang, B.-H. (2013). Feature processing and modeling for 6d motion gesture recognition. *IEEE Transactions on Multimedia, 15*(3), 561–571. http://dx.doi.org/10.1109/TMM.2012.2237024.

Cook, D. J., Augusto, J. C., & Jakkula, V. R. (2009). Review: Ambient intelligence: Technologies, applications, and opportunities. *Pervasive Mobile computing, 5*(4), 277–298. http://dx.doi.org/10.1016/j.pmcj.2009.04.001.

Coronato, A. Pietro, G. D. 2012. Detection of motion disorders of patients with autism spectrum disorders. In *IWAAL* (pp. 415–422).

Coronato, A., De Pietro, G., Esposito, M. 2006. A semantic context service for smart offices. In *Proceedings of the 2006 International Conference on Hybrid Information Technology – Vol. 02* (pp. 391–399). IEEE Computer Society.

Eyerich, P., Nebel, B., Lakemeyer, G., Classen, J. 2006. GOLOG and PDDL: what is the relative expressiveness? In *Proceedings of the 2006 international symposium on Practical cognitive agents and robots, PCAR '06, ACM* (pp. 93–104). New York, NY, USA. doi:10.1145/1232425.1232440. URL: http://doi.acm.org/10.1145/1232425.1232440.

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Jr., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*(6), 439–484.

Goncalves, N., Rodrigues, J., Costa, S., Soares, F. 2012. Preliminary study on determining stereotypical motor movements. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1598–1601). doi:10.1109/EMBC.2012.6346250.

Guyon, I., Gunn, S., Nikravesh, M., & Zadeh, L. A. (2006). *Feature extraction: Foundations and applications (studies in fuzziness and soft computing)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc..

Holmes, G., Donkin, A., & Witten, I. H. (1994). *WEKA: A machine learning workbench*, 357–361.

Lam, K. S., & Aman, M. G. (2007). The repetitive behavior scale-revised: Independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(5), 855–866.

Lee, K., Lunney, T., Curran, K., & Santos, J. (2009). *Ambient Middleware for Context-Awareness*. doi:10.4018/jaci.2009070105.

Lin, T., Wang, C., & Lin, P.-C. (2008). A neural-network-based context-aware handoff algorithm for multimedia computing. *ACM Transactions on Multimedia Computing, Communications, and Applications, 4*(3), 17:1–17:23. http://dx.doi.org/10.1145/1386109.1386110. http://doi.acm.org/10.1145/1386109.1386110.

Lotter, V. (1967). Epidemiology of autistic conditions in young children. *Social Psychiatry and Psychiatric Epidemiology, 1*, 163–173. http://dx.doi.org/10.1007/BF00578950.

Loy, C. C., Xiang, T., & Gong, S. (2009). *Surveillance video behaviour profiling and anomaly detection* (pp. 74860E–74860E-11). doi:10.1117/12.832188.

McCarthy, J. 1963. Situations, actions, and causal laws, Tech. Rep. Memo 2, Stanford Artificial Intelligence Project, Stanford University.

McDermott, D. 1998. The planning domain definition language, version 1.2. Technical report, Yale center for computational vision and control.

Min, C.-H., Tewfik, A. 2010. Automatic characterization and detection of behavioral patterns using linear predictive coding of accelerometer sensor data. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 220–223). doi:10.1109/IEMBS.2010.5627850.

Monekosso, D., & Remagnino, P. (2010). Behavior analysis for assisted living. *IEEE Transactions on Automation Science and Engineering, 7*(4), 879–886. http://dx.doi.org/10.1109/TASE.2010.2049840.

Mozer, M. C., 1998. The neural network house: An environment that adapts to its inhabitants.

Nater, F., Grabner, H., Gool, L. J. V. 2010. Exploiting simple hierarchies for unsupervised human behavior analysis. In Computer Vision and Pattern Recognition. doi:10.1109/CVPR.2010.5539877.

Osborne, M. A., Roberts, S. J., Rogers, A., & Jennings, N. R. (2012). Real-time information processing of environmental sensor network data using Bayesian Gaussian processes. *ACM Transactions on Sensor Networks, 9*(1), 1:1–1:32. http://dx.doi.org/10.1145/2379799.2379800. http://doi.acm.org/10.1145/2379799.2379800.

Reiter, R. (2001). *Knowledge in action: Logical foundations for specifying and implementing dynamical systems*. Massachusetts, MA: The MIT Press.

Riedmiller, M. Braun, H. 1993. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *1993 IEEE International Conference on Neural Networks, Vol. 1* (pp. 586–591). doi:10.1109/ICNN.1993.298623.

Rxtx library. URL http://rxtx.qbang.org/wiki/index.php/Main Page.

Silipo, R., & Mazanetz, M. P. (2012). *The knime cookbook: Recipes for the advanced user*.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society Series B (Methodological), 36*(2), 111–147. http://dx.doi.org/10.2307/2984809. http://dx.doi.org/10.2307/2984809.

Xiang, T., & Gong, S. (2008). Video behavior profiling for anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*(5), 893–908. http://dx.doi.org/10.1109/TPAMI.2007.70731.

Xu, X., Tang, J., Liu, X., Zhang, X. 2010. Human behavior understanding for video surveillance: Recent advance. In *2010 IEEE International Conference on Systems Man and Cybernetics (SMC)* (pp. 3867–3873). doi:10.1109/ICSMC.2010.5641773.

Ye, J., Dobson, S., & McKeever, S. (2011). Situation identification techniques in pervasive computing: A review. *Pervasive and Mobile Computing*. doi: 10.1016/j.pmcj.2011.01.004. In Press, Corrected Proof.