COMBINING GAZE AND DEMOGRAPHIC FEATURE DESCRIPTORS FOR AUTISM CLASSIFICATION

Shaun Canavan, Melanie Chen, Song Chen, Robert Valdez, Miles Yaeger, Huiyi Lin, and Lijun Yin

Binghamton University

ABSTRACT

People with autism suffer from social challenges and communication difficulties, which may prevent them from leading a fruitful and enjoyable life. It is imperative to diagnose and start treatments for autism as early as possible and, in order to do so, accurate methods of identifying the disorder are vital. We propose a novel method for classifying autism through the use of eye gaze and demographic feature descriptors that include a subject's age and gender. We construct feature descriptors that incorporate the subject's age and gender, as well as features based on eye gaze data. Using eye gaze information from the National Database for Autism Research, we tested our constructed feature descriptors on three different classifiers; random regression forests, C4.5 decision tree, and PART. Our proposed method for classifying autism resulted in a top classification rate of 96.2%.

***Index Terms*—** Gaze, autism, classification

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by difficulties with social communication and interaction [13]. People diagnosed with autism also may suffer from repetitive thoughts and obsessive behaviors. These behaviors and communication issues may lead to difficulties learning and forming relationships, as well as bullying [5]. Early detection of ASD is critical as it can help children overcome these disorder-related obstacles. The salient association between eye movements and cognitive processes and abilities in general makes study of the relationship between eye gaze and ASD significant [8]. Previous studies show that impairments in visual attention and gaze patterns in children with ASD serve as important markers for diagnosis and can be the basis for disturbances in communication and interaction [1].

An accurate algorithm that incorporates various factors like age and gender as well as patterns in eye gaze behavior can make measurements that may not be possible with the human eye, thus providing a more reliable and quicker diagnosis. With 25 to 50% of children receiving early intervention beginning general education by kindergarten, it is imperative to screen at-risk individuals (often identified by familial history) and potentially diagnose ASD as early as possible in order to begin treatment and pursue a high quality of life.

Previous studies have used eye tracking devices to measure and categorize eye gaze patterns in children with and without autism. Pierce et al. [12] found that, when given a choice between social images and geometric images, children with ASD prefer to look at geometric patterns, whereas typically developing (TD) children prefer to look at social images. Sasson et al. [16] discovered, when presented with arrays of social and nonsocial objects, the visual attention of children with autism was more circumscribed (fewer images were explored), more perseverative (more time was spent on each image) and more detail-oriented (greater amount of fixation on those images that were explored) than TD children. Bekele et al. [2] found, through the use of different regions in the brain and compared to TD children, children with ASD focus more on the forehead (an information irrelevant area) than on the mouth (information relevant).

Although these studies provide a wealth of information regarding characteristic eye gaze patterns of children with autism when presented with various stimuli, they rarely focus on detection or diagnosis of ASD gaze patterns, instead emphasizing observed differences in individuals known to be diagnosed with autism. Alie et al. [1] performed one of the few studies attempting to classify autism from examining eye gaze of infants with ASD. They used Markov models to assign infant subjects to either a TD or ASD group based on sequences of eye gaze data obtained through video. The study achieved a 93.75% accuracy rate, indicating eye gaze alone is a significant determinant in detecting autism.

In this paper we propose a novel method for autism classification through the use of eye gaze and subject demographic information. Feature descriptors, created using subject's age, gender, and eye gaze data were tested on three different classifiers; random regression forest [4], C4.5 decision tree [14], and PART [7]. Our proposed method resulted in an autism classification rate of 96.2% on the National Database for Autism Research (NDAR) [11]. See figure 1 for an overview of our proposed approach. A summary of the main contributions of this work follows:

(1) We propose novel features descriptors, for autism classification, based on gaze and demographic information such as age and gender.
(2) We propose the use of 3 different machine learning-based classification algorithms for helping to classify autism. These include random regressions forests [4], C4.5 decision trees [14], and PART [7].
(3) We test our new proposed features and classification scheme on data obtained from the National Database for Autism Research [11].

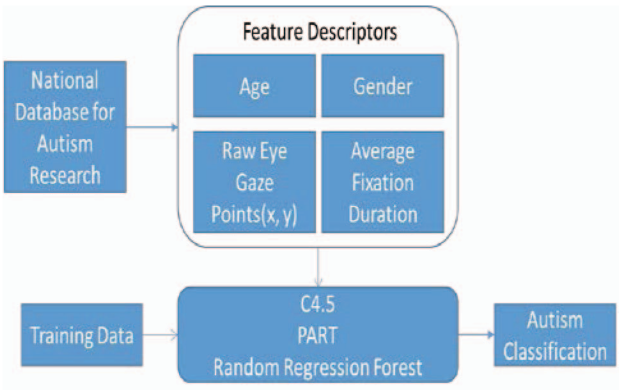


Figure 1. Proposed autism classification overview.

2. FEATURE DESCRIPTORS FOR AUTISM CLASSIFICATION

The proposed classification scheme makes use of eye gaze data, as well as demographic information in order to classify autism. We propose the use four different feature descriptors for help with this classification problem. These include (1) raw eye gaze points (x, y); (2) average fixation duration; (3) age; and (4) gender.

Eye gaze information can detail important features that can be used to classify autism [6]. The raw eye gaze points contain the (x, y) coordinates indicating where the gaze of each test subject was focused. The total number of gaze points for each NDAR test subject is inconsistent and vary in total number. In order to analyze each test subject's gaze over the same period of time and construct a feature vector of consistent size, the first 2580 gaze points were chosen. This number corresponds to the minimum number of gaze points recorded among all test subjects. For this proposed feature descriptor, no pre-processing is performed on the gaze points; raw gaze data was used.

Along with the raw eye gaze data, the NDAR provides eye gaze fixation information. Average fixation was calculated by dividing the total length of fixations over the total number of fixations per subject. Average fixation length was used as subject's with autism tend to fixate on images for a longer amount of time [16]. For example, the average fixation of one test subject classified with a medium ASD risk has an average fixation length of approximately 0.6 seconds, while another subject at high risk has an average fixation length of 0.8 seconds.

The last two feature descriptors are age and gender. Both male and female genders are represented in the NDAR, with ages ranging from 2 to 132. There are a total of 91 female subjects, and 166 male subjects. Section 4.1 and table 1 give detailed statistics on age and gender, as well as our method used to account for large age ranges. Using the four different feature descriptors described, we then created one feature vector, for each subject, to be used as training data to each of the three classification techniques (described in the next section). The total length of each feature vector is 2583 (eye gaze points + age + gender + average fixation length).

3. CLASSIFICATION TECHNIQUES

Using the proposed feature descriptors detailed in section 2, we next propose the use of three different machine learning-based classifiers for help with gaze-based autism classification. These classifiers are (1) Random regression forests [4]; (2) C4.5 decision trees [14]; and (3) PART [7]. Each of the classifiers are detailed in the following subsections.

3.1. Random regression forests

Regression trees [3] are a powerful tool used for classification, which classify by splitting a larger problem into smaller ones that can be solved with simple predictors. Each node of a regression tree represents a question, the answer to which directs towards the left or right child. When they are trained the data is clustered so that simple models can achieve high accuracy. However, it has been shown that regression trees are prone to overfitting. By using a collection of randomly trained trees, Breiman [4] found that this overfitting can be overcome. Random forests work by constructing multiple, randomly trained, regression trees. The mean classification of each of the trees is then taken as the output. Random regression forests have been used in other fields such as gesture recognition [17]. This success in classifying human data, their ability to overcome overfitting, their power for classification, and their speed, makes them a good fit for our proposed autism classification scheme.

3.2. C4.5 decision trees

The C4.5 algorithm was developed by Quinlan [14], which is an extension of ID3 algorithm [15]. It is a statistical classifier that builds decision trees based on information entropy. At each node of the tree, the algorithm looks to split the subsets based on the most information gain. From all of the available attributes, the one that has the highest information gain makes the decision. The algorithm then recursively does this for each of the available subsets that are left. The C4.5 algorithm has multiple stopping criteria including (1) All samples belong to the same class; (2) there is no information gain from the attributes at the current node; and (3) an instance of a class type that has not been seen before occurs. The C4.5 algorithm is able to handle both discrete and continuous data, and once the tree is created it attempts to remove branches that don't help (prunes the tree). Because of this they are also a good choice for autism classification.

3.3. PART

The PART algorithm [7] makes use of partial decision trees and a separate and conquer strategy for rule induction. It performs this by first building a rule based on the training set, then removing all of the instances that the rule covers. It then looks at the remaining instances and recursively creates rules and removes the instances associated with each rule until there are no instances left. In order for PART to make a rule it builds a pruned decision tree with the current set of instances, which haven't been removed, and the leaf that has the largest amount of coverage is turned into a rule. The rest of the tree is then discarded. In doing this the algorithm can avoid over-pruning by only generalizing once all of the subtrees have been expanded. These algorithm features help make for a third promising choice for autism classification.

4. EXPERIMENTAL DESIGN AND RESULTS

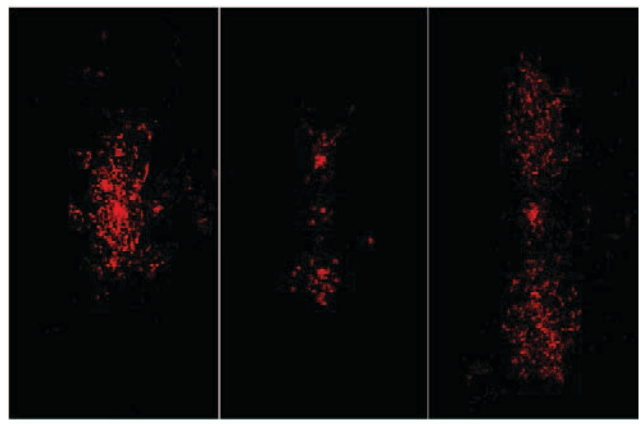
To evaluate our proposed approach for classifying autism using gaze data and demographic information, we make use of the National Database for Autism Research (NDAR) [11]. Using the features detailed in section 2 we use each of the classification techniques detailed in section 3 to conduct the experiments. The database, experimental design, and results are detailed in the following subsections.

4.1. Database

Test subject data was obtained from the National Database for Autism Research (NDAR), an extensive collection of measurements published by the National Institute of Health [11]. This database contains information on each test subject, the stimulus the subject viewed, and the subject's eye gaze while viewing the stimulus. Metrics include the age and gender of each test subject, details about calibration and configuration of the gaze monitoring system, a brief description of the stimulus, gaze point x and y coordinates, and gaze fixation classification (e.g. how long a fixation was recorded). Data also includes a class indicating whether the test subject is diagnosed with autism or, the subject's relative risk (labeled as low, medium, high, or ASD) for being diagnosed. We tested our proposed classification scheme on a total of 257 subjects from the NDAR with a total of 91 subjects being female and 166 being male. Table 1 list details on the age ranges, risk types in those ranges, and the total number of subjects for each. See figure 2 for gaze plots of low, medium, and high risk subjects.

Table 1. Subject information from [11].

Age Ranges	Gender	Risk Types	Total Number of Subjects
2-10	F	low; high	45
	M		49
11-20	F	low; medium; high	17
	M		54
21-30	F	medium	27
	M		41
31-40	F	medium	2
	M		2
60+	F	NA	NA
	M	ASD	20



(a) (b) (c)

Figure 2. Eye gaze for (a) low risk; (b) Medium Risk; and (c) high risk subjects from the NDAR [11].

4.2 Experimental design and classification results

In order to evaluate the efficacy of our proposed features we conducted experiments using three separate classifiers; PART [7], C4.5 [14], and random regression forests [4]. It should be noted that there are some abnormalities in the NDAR data including subjects of 129 and 132 years of age, and all of the subjects older than 60 were diagnosed with ASD. In order to acknowledge these abnormalities, tests were conducted on two separate sets of data. The first contains all 257 subjects and the second contains all subjects ages 2-40 (237 total subjects). For both of these experiments, all gaze and demographic feature descriptors were concatenated into one feature vector for each subject. Using 10-fold cross validation, data is randomly split into 10 subsets. One was used for testing and the other nine are used for training. This is done for each of the subsets being used as the testing data with the average error used.

Using this classification scheme, the training data where outliers were removed (ages 60+), yielded a max classification rate of 96.2% using the PART classifier. Using all of the training data, we obtained a max classification rate of 94.16% using the PART and C4.5 classifiers. The equal classification rate of these two algorithms may be attributed to their use of decision trees, however, they did not achieve the same results when the outliers were removed. This leaves an open question of which classification scheme can best be used to help classify autism. See table 2 for a comparison each classifier and tables 3-8 for the confusion matrices of each classifier.

Table 2. Classification rates of 3 tested classifiers on [11].

Classifier	PART	C4.5	Random Regression Forest
Original Data	94.16%	94.16%	91.05%
Outliers Removed	96.2%	94.94%	93.25%

Table 3. Confusion matrix, from PART [7], of subjects of ages 1-40 from [11].

	low	medium	high
low	41	0	6
medium	0	131	0
high	3	0	56

Table 4. Confusion matrix, from PART [7], of all subjects tested from [11].

	low	medium	high	ASD
low	39	0	8	0
medium	0	130	0	1
high	6	0	53	0
ASD	0	0	0	20

Table 5. Confusion matrix, from C4.5 [14], of subjects of ages 1-40 from [11].

	low	medium	high
low	40	0	7
medium	0	131	0
high	5	0	54

Table 6. Confusion matrix, (C4.5 [14], of all subjects tested [11].

	low	medium	high	ASD
low	39	0	8	0
medium	0	130	0	1
high	6	0	53	0
ASD	0	0	0	20

Table 7. Confusion matrix, from random regression forest [4][3], of subjects of ages 1-40 from [11][10].

	low	medium	high
low	38	4	5
medium	0	131	0
high	2	5	52

Table 8. Confusion matrix, from random regression forest [4], of all subjects tested from [11].

	low	medium	high	ASD
low	37	2	8	0
medium	0	131	0	0
high	3	4	52	0
ASD	0	6	0	14

Tables 3-8, show that for each of the classifiers tested, medium (out of low, medium, and high risk) has the highest classification rate, with only one misclassification. For the PART and C4.5 classifiers, the ASD diagnosis was successfully classified 100% of the time, with random regression forests successfully classifying 70% of the testing data. These results are encouraging as it agrees with various studies [12][6][2] that gaze can be used as a marker for ASD diagnoses. It also points to age being a significant factor in being able to diagnose ASD, as all subjects with ASD, in the NDAR, were over 60 years of age.

5. DISCUSSION AND FUTURE WORK

We have presented a novel approach to classifying autism by constructing features from a subject's raw eye gaze points (x,y), average fixation length, and demographic information such as age, gender. Our results on the NDAR [11] are encouraging with a max classification rate of 96.2%. We have the discussed the potential use of three different machine learning classifiers, (1) Random regression forests [4]; (2) C4.5 decision trees [14]; and (3) PART [7], to help with classifying autism. While the results are encouraging, we are interested in testing on larger datasets, especially those including predominantly children, as early intervention in ASD is crucial.

Our future work includes extending our current feature descriptors to include average gaze velocity over time, as well as, the potential for using deep learning approaches for classifying autism. Recently deep learning has been successfully used to increase the performance of eye tracking systems [10]. Future work also includes other non-rule-based classifiers. We also want to know which features are important for classifying autism. We have shown that gaze and demographic information can be used to classify autism. The next step is to determine which of these exact features are the strongest for classification.

We are also interested in a multi-modal approach to classifying autism. It has been shown that children diagnosed with ASD have difficulty with coordination which adversely influences gait [5]. We are currently exploring combining gaze, gait, and demographic feature descriptor to help classify autism. While the focus on this work is to classify autism, we are also investigating the use of gaze to classify patients with schizophrenia, as they have been shown to have problems with gaze perception [9].

6. REFERENCES

[1] D. Alie, M. Mahoor, W. Mattson, D. Anderson and D. Messinger, "Analysis of eye gaze pattern of infants at risk of autism spectrum disorder using Markov models", *2011 IEEE Workshop on Applications of Computer* Vision, 2011.

[2] E. Bekele, Z. Zheng, A. Swanson, J. Crittendon, Z. Warren and N. Sarkar, "Understanding How Adolescents with Autism Respond to Facial Expressions in Virtual Reality Environments", *IEEE Trans. Visual. Comput. Graphics*, 19(4), pp. 711-720, 2013.

[3] L. Breiman, J. Friedman, R. Olshen, and C. Stone, "Classifcation and regression trees," *Wadsworth and Brooks*, Monterey, CA, 1984.

[4] L. Breiman, "Random forests," *Machine Learning*, 45(1), pp.5-32, 2001.

[5] K. Deirdre, L. Gallagher, and J. Gormley, "Gait deviations in children with autism spectrum disorders: a review," *Autism Research and Treatment*, 2015.

[6] T. Falck-Ytter, S. Bölte and G. Gredebäck, "Eye tracking in early autism research", *Journal of Neurodevelopmental Disorders*, 5(1), pp. 28, 2013.

[7] E. Frank, and I. Witten, "Generating accurate rule sets without global optimization", *Machine Learning,* 1998.

[8] M. Hayhoe and D. Ballard, "Eye movements in natural behavior", *Trends in Cognitive Sciences*, vol. 9, no. 4, pp. 188-194, 2005.

[9] C. Hooker, and S. Park, "You must be looking at me: The nature of gaze perception in schizophrenia patients," *Cognitive neuropsychiatry,* 10(5), pp. 327-345, 2005.

[10] K. Krafka, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye Tracking for Everyone," *Computer Vision and Pattern Recogniiton*, 2016.

[11] "National Database for Autism Research," 2016. [Online] Available: https://ndar.nih.gov/, 2016. [Accessed 24-Apr-2016].

[12] K. Pierce, D. Conant, R. Hazin, R. Stoner and J. Desmond, "Preference for Geometric Patterns Early in Life as a Risk Factor for Autism", *Arch Gen Psychiatry*, 68(1), pp. 101, 2011.

[13] G. Powell, S. Wass, J. Erichsen and S. Leekam, "First evidence of the feasibility of gaze-contingent attention training for school children with autism", *Autism*, 2016.

[14] J. Quinlan, "C4.5: Programs for machine learning", *Morgan Kaufmann Publisher*, 1993.

[15] J. Quinlan, "Induction of Decision Trees," *Machine Learning*, 1(1), pp.81-106, 1986.

[16] N. Sasson, L. Turner-Brown, T. Holtzclaw, K. Lam and J. Bodfish, "Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays", *Autism Res*, 1(1), pp. 31-42, 2008.

[17] T. Schmidt, F.P. Araujo, G.L. Pappa, and E.R. Nascimento, "Real-Time hand Gesture Recognition Based on Sparse Positional Data," *Brazalian Workshop on Computer Vision*, 2014.