Learning to Predict Autism Spectrum Disorder based on the Visual Patterns of Eye-tracking Scanpaths

Romuald Carette[1], Mahmoud Elbattah[1], Federica Cilia[2], Gilles Dequen[1], Jean-Luc Guérin[1] and Jérôme Bosche[1]

[1]*Laboratoire MIS, Université de Picardie Jules Verne, Amiens, France*
[2]*Laboratoire CRP-CPO, Université de Picardie Jules Verne, Amiens, France*

Keywords: Autism Spectrum Disorder, Machine Learning, Eye-tracking, Scanpath.

Abstract: Autism spectrum disorder (ASD) is a lifelong condition generally characterized by social and communication impairments. The early diagnosis of ASD is highly desirable, and there is a need for developing assistive tools to support the diagnosis process in this regard. This paper presents an approach to help with the ASD diagnosis with a particular focus on children at early stages of development. Using Machine Learning, our approach aims to learn the eye-tracking patterns of ASD. The key idea is to transform eye-tracking scanpaths into a visual representation, and hence the diagnosis can be approached as an image classification task. Our experimental results evidently demonstrated that such visual representations could simplify the prediction problem, and attained a high accuracy as well. With simple neural network models and a relatively limited dataset, our approach could realize a quite promising accuracy of classification (AUC > 0.9).

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is described as a pervasive developmental disorder characterized by a set of impairments including social communication problems, difficulties with reciprocal social interactions, and unusual patterns of repetitive behavior (Wing and Gould, 1979). According to the U.S. Department of Health and Human Services, ASD has been considered to affect about 1% of the world's population (DOH, 2018). Individuals diagnosed with ASD typically suffer from deficits in social communication and interaction across multiple contexts. Particularly, they could be incapable of making and maintaining eye contact, or keeping their focus on specific tasks. Such troubling symptoms can unfortunately place a considerable strain on their lives and their families.

The diagnosis of ASD is highly desirable at early stages in terms of benefits for both child and the family. The diagnosis process usually involves a set of cognitive tests that could require hours of clinical examinations. In addition, the variation of symptoms makes the identification of ASD more complicated to decide. In this respect, computer-aided technologies have come into prominence in order to provide guidance through the course of examination and assessment. Examples include Magnetic Resonance Imaging (MRI), Electroencephalography, and eye-tracking that is the focus of this study.

Eye-tracking is the process of capturing, tracking and measuring eye movements or the absolute point of gaze (POG), which refers to the point where the eye gaze is focused in the visual scene (Majaranta and Bulling, 2014). The eye-tracking technology received particular attention in the ASD context since abnormalities of gaze have been consistently recognized as the hallmark of autism in general. The Psychology literature is replete with studies that analyzed eye movements in response to verbal or visual cues as signs of ASD (e.g. Coonrod and Stone, 2004; Jones et al., 2014; Sepeta et al., 2012; Wallace et al., 2012).

Furthermore, the coupling of eye-tracking instruments with modern AI techniques is leveraging further capabilities for advancing the diagnosis and its applications. Data-driven techniques, such as Machine Learning (ML), are increasingly embraced to provide a second opinion that is considered to be less biased and reproducible. This study follows on the path of integrating the eye-tracking technology

Carette R., Elbattah M., Cilia F., Dequen G., Guérin J. and Bosche J.
Learning to Predict Autism Spectrum Disorder based on the Visual Patterns of Eye-tracking Scanpaths.
DOI: 10.5220/0007402601030112
In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 12th International Conference on Health Informatics)*, pages 103-112
ISBN: 978-989-758-353-7

in tandem with ML to support the diagnosis of ASD. The study is part of interdisciplinary collaboration between research units of Psychology and AI.

Our approach is distinctively based on the premise that visual representations of eye-tracking scanpaths can discriminate the gaze beahviour of autism. At its core, the key idea is to compactly render eye movements into an image-based format while maintaining the dynamic characteristics of eye motion (e.g. velocity, acceleration) using color gradients. In this manner, the prediction problem can be approached as an image classification task. The potential of our approach is evidently demonstrated in terms of promising classification accuracy though using largely simple ML models.

2 RELATED WORK

Plentiful studies sought to take advantage of eye-tracking for the study and analysis of ASD. For instance, Vabalas and Freeth (2016) demonstrated interesting physiological elements based on eye-tracking experiments. In face-to-face interactions, eye movements were different among individuals who fell on different positions on the spectrum of autism. Specifically, persons with high autistic traits were observed to have shorter and less frequent saccades. In another study, eye-tracking was used to identify ASD-diagnosed toddlers based on the duration of fixations and the number of saccades (Pierce et al., 2011). Their results showed that participants with ASD spent significantly more time fixating on dynamic geometric images compared to other groups.

Likewise, a longitudinal study examined the patterns of eye fixation for infants aged 2 to 6 months (Jones and Klin, 2013). They notably indicated that ASD-diagnosed infants exhibited a mean decline in eye fixation, which was not observed for those who did not develop ASD afterwards. Moreover, another cohort study suggested the strong potential of eye-tracking as an objective tool for quantifying autism risk and estimating the severity of symptoms (Frazier et al., 2016).

More recent studies attempted to makes use of eye-tracking to build predictive ML models. To name a few, Pusiol et al. (2016) worked on the analysis of the eye focus on the face during conversations. This was specifically applied to developmental disorder (DD) children or Fragile X Syndrome (FXS) children. A set of classification models were experimented including Recurrent Neural Networks (RNN), Support Vector Machines (SVM), Naive Bayes, and Hidden Markov Model. With RNN, they were able to reach a high prediction accuracy of 86% and 91% for the classification of female and male FXS patients respectively. Another recent study applied ML on eye-tracking output in order to predict ASD (Carette et al., 2017). The ML model included features related to the saccade eye movement (e.g. amplitude, duration, acceleration). Their experiments aimed at detecting ASD among a set of 17 children aged 8 to 10 years. Despite using a limited dataset and a relatively simple model, they demonstrated promising potentials of ML application in this regard.

Compared to earleir efforts, the main distinction of the present work is that it is purely reliant on the visual representation of eye-tracking recordings. We produce scanpath visualizations that represent the spatial coordinates of the eye movement along with its dyanmics. The approach allowed for simplifying the model training, and attained high accuracy as well. It is claimed that such approach has not been applied before in the context of ASD, to the best of our knowledge.

3 DATA ACQUISTION

3.1 Participants

A group of 59 children took part in this study. It was highly desirable to have participants at an early stage of development, as the principal goal was towards supporting the early detection and diagnosis of ASD. Specifically, all participants were school-aged children of average age about 8 years. Figure 1 provides a box plot showing the distribution of age.

A group of typically developing (TD) children was recruited from a number of French schools in the region of Hauts-de-France. Parental reports of any possible concerns were carefully considered. Both informed consent from parents and assent from subjects were confirmed for all cases. Moreover, all procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

The participants were broadly organized into two groups as: i) ASD-Diagnosed, or ii) Non-ASD. ASD-diagnosed children were examined in multidisciplinary ASD specialty clinics. The intensity of autism was estimated by psychologists

based on the Childhood Autism Rating Scale (CARS) (Schopler et al., 1980).

Figure 2 gives the distribution of CARS score among ASD-diagnosed participants using a box plot. Further, Table 1 gives summary statistics of the participants (e.g. gender distribution, age mean).

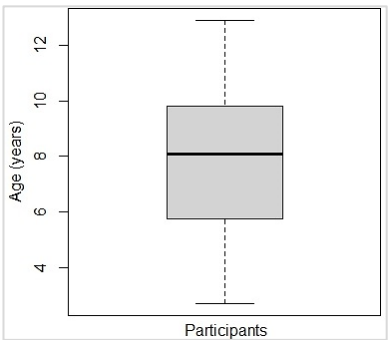

Figure 1: The age distribution in all participants.

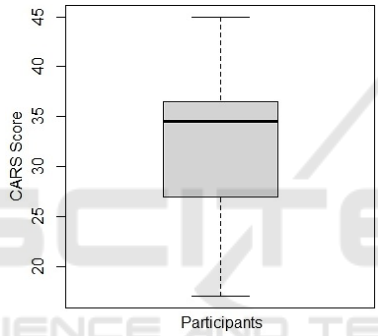

Figure 2: The distribution of CARS in ASD-diagnosed participants.

Table 1: Summary statistics of participants.

Number of Participants	59
Gender Distribution (M / F)	38 (≈ 64%) / 21 (≈ 36%)
Number of Non-ASD	30
Number of ASD-Diagnosed	29
Age (Mean / Median) years	7.88 / 8.1
CARS (Mean / Median)	32.97 / 34.50

3.2 Experimental Protocol

The participants were invited to watch a set of videos, which included scenarios tailored specifically to stimulate the eye movement across the screen area. Participants could be seated on their own or on their parents' lap at approximately 60-cm distance from the display screen. The experiments were conducted in a quiet room at the University premises. Physical white barriers were also used to reduce visual distractions.

The scenarios varied in content and length in order to allow for analyzing the ocular activity of participants from different perspectives. In general, videos were designed to include visual elements that can be especially attractive to children (e.g. colorful balloons, cartoons). The position of elements can also change throughout the course of experiment. In addition, some videos could include a human presenter who speaks and attempts to turn the child's attention to visible or invisible elements around. All conversations were performed in French as the native language of participants. Figure 3 presents a screenshot captured from one of the videos used in eye-tracking experiments.

Further stimuli were presented from the SMI Experiment Center Software. Stimuli represented multiple distinct types used in the eye gaze literature. Examples included static and dynamic naturalistic scenes with and without receptive language, joint attention stimuli, static face or objects and cartoons stimuli. The average duration of eye-tracking experiments was about 5 minutes.

Participants were mainly examined for the quality of eye contact with the presenter, and the level of focus on other elements. A five-point calibration scheme was used. The calibration routine was followed by a set of verification procedures.

Figure 3: Screenshot captured from one of the videos.

3.3 Eye-tracking Records

The SMI remote eye-tracker (Red-m 250Hz) was the main instrument used to perform the eye-tracking function. The device belongs to the category of screen-based eye-trackers. It can be conveniently placed at the bottom of the screen of a desktop PC or laptop. In our case, a 17-inch monitor of 1280x1024 resolution was used.

Three basic categories of eye movements are aimed to be captured by eye-trackers including: i) Fixation, ii) Saccade, and iii) Blink. A fixation is the brief moment that happens while pausing the gaze on an object in order that the brain can perform

Table 2: A snapshot of eye-tracking records.

Recording Timestamp [ms]	Category Right	Category Left	Point of Regard Right X [px]	Point of Regard Right Y [px]	Point of Regard Right Y [px]	Point of Regard Left Y [px]
8005654.069	Fixation	Fixation	1033.9115	834.0902	1033.9115	834.0902
8005673.953	Fixation	Fixation	1030.3754	826.0894	1030.3754	826.0894
8005693.85	Saccade	Saccade	1027.337	826.3127	1027.337	826.3127
8005713.7	Saccade	Saccade	1015.0085	849.2188	1015.0085	849.2188
8005733.589	Saccade	Saccade	613.7673	418.1735	613.7673	418.1735

the perception process. The average duration of fixation was estimated to be 330 milliseconds (Henderson, 2003). Further, the accurate perception requires constant scanning of the object with rapid eye movements, which are so-called saccades. Saccades include quick, ballistic jumps of 2º or longer that take about 30–120 milliseconds each (Jacob, 1995). On the other hand, a blink would often be a sign that the system has lost track of the eye gaze.

Likewise, the initial records of our eye-tracking experiments essentially included the features described above. In addition, the eye-tracker provided the (x,y) coordinates that represented the participant's gaze direction into the screen. The coordinates were of special significance to implement our approach for drawing the virtual path of the viewer's POG and the dynamics of movement as well (e.g. velocity, acceleration).

Table 2 provides a simplified snapshot of the eye-tracking records. The records describe the category of movement and the POG for the left and right eyes over time. The table lists five records of eye movements including two fixations and three saccades. Due to limited space, many variables are not included in the table (e.g. pupil size, pupil diameter, pupil position).

3.4 Visualization of Eye-tracking Scanpaths

A scanpath represents the sequence of consecutive fixations and saccades as a trace through time and space that may overlap itself (Goldberg and Helfman, 2010). The premise of the study is based on learning the visual patterns of scanpaths. Specifically, the core idea was to compactly render the eye movements into a visual representation that can simplify and describe the path and dynamics of eye movement.

To achieve this, we availed of the coordinates included in eye-tracking records, which represented the change in participant's gaze direction into the screen with respect to time. Based on the change in position along associated time, we were able to calculate the velocity of gaze movement. The acceleration of movement could be computed based on the change in velocity, and the jerk of movement could be accordingly computed based on the change in acceleration. As such, the variation of velocity, acceleration and jerk could describe properly the dynamics of eye motion.

Subsequently, the path of eye motion along with computed dynamics were transformed into images. For every participant, a set of images could describe the visual patterns of gaze behavior. Specifically, an image is constructed as below:

- A line is drawn for each transition from position (x_t, y_t) to (x_{t+1}, y_{t+1}), where t is a defined point of time during the experiment.
- The change in color across the line represented the movement dynamics. The values of RGB components were tuned based on velocity, acceleration and jerk with respect to time. For instance, the values of velocity range from black (i.e. low) to red (i.e. high). In this manner, higher values of velocity shift gradually towards deeper red values. Likewise, the acceleration and jerk were set using color gradients of green and blue respectively.
- The images produced were vertically mirrored since the origin was located at the bottom of the screen.

All color values were capped to one-quarter of the diagonal length of the display screen. With this cap, images represented the eye movements including saccades (yellow or white, white representing very fast movements, exceeding the cap), and fixations as cyan.

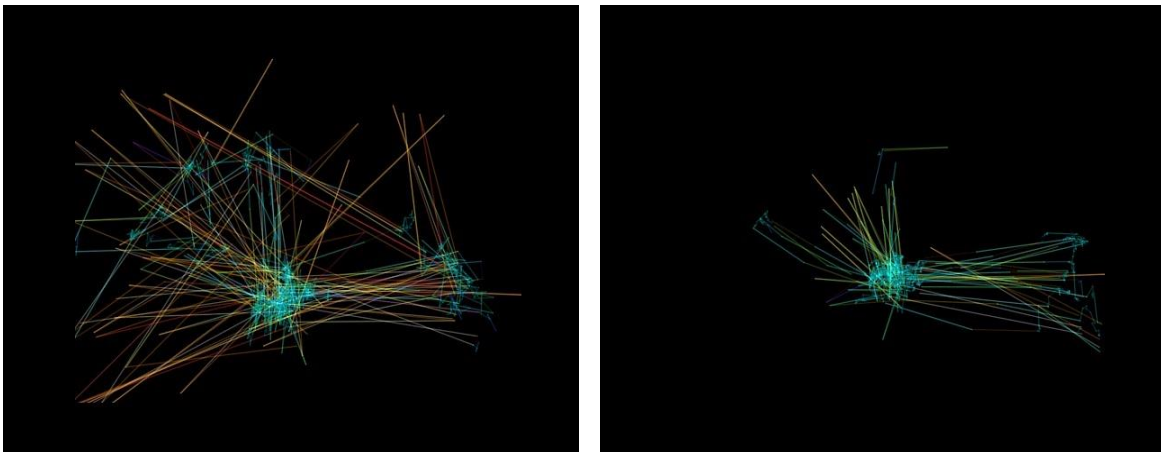

Figure 4: Visualization of eye-tracking scanpaths. The left-sided image represents an ASD-diagnosed participant, while the other is for non-ASD.

In addition, images were constrained to contain the same level of information approximately. Specifically, a threshold was applied to limit the number of points to be drawn. The threshold was aimed to be high enough to sufficiently describe the pattern of gaze behavior. However, overly high values could increase the possibility of producing cluttered visualizations. Therefore, several tests were conducted to choose an appropriate value for the threshold. With a threshold ranging from 100 to 150, images seemed to include fewer lines, which turned out to poorly discriminate the two classes of participants. Eventually, it was decided to set the threshold at 200, which largely struck an adequate balance, and could capture the key features of eye motion. The visualizations were produced using Python along with the popular Matplotlib library (Hunter, 2007).

Figure 4 presents examples of scanpaths visualizations corresponding to ASD and non-ASD participants. As it appears, the center of both images includes areas of high density, which probably represent one of the main points of focus in the video scenario. The visualizations may also highlight other points of focus x at the right side of the screen. These focus points are drawn as cyan (i.e. low velocity, but high acceleration and jerk), while other widely diffused lines seem yellow (i.e. high velocity, medium acceleration and jerk).

The figures can also describe the distinction of the gaze movement in both cases. For example, it can be noticed that the ASD-diagnosed participant had a tendency to look at the bottom of the screen, where the eye-tracking device was placed.

The visualizations resulted in an image dataset containing 547 images. Specifically, 328 images for the non-ASD participants, and 219 images for others (i.e. ASD-diagnosed). The default image dimension was set as 640x480. A more comprehensive presentation of the process of data acquisition and transformation was elaborated in an earlier publication (Carette et al., 2018). Further, the dataset along with metadata files were recently published and made publicly available on the Figshare data repository (Figshare, 2018). It is conceived that the dataset itself could be useful for developing further applications or discovering interesting insights using data mining or other AI techniques.

4 DATA AUGMENTATION AND PRE-PROCESSING

4.1 Image Augmentation

Image augmentation is a common technique to enlarge datasets, and help models generalize better and reduce the possibility of overfitting. The basic idea of augmentation is to produce synthetic samples using a random set of image transformations (e.g. rotation, shearing). Augmentation was recognized to improve the prediction accuracy in image classification applications (e.g. Xu et al., 2016; Perez and Wang, 2017).

Similarly, we applied augmentation to produce variations of the eye- scanpath visualizations. The dataset was augmented with additional 2,735 samples, where five synthetic images were generated for each visualization. The data augmentation process was greatly simplified thanks to the Keras library (Chollet, 2015), which includes an easy-to-use API for that purpose.

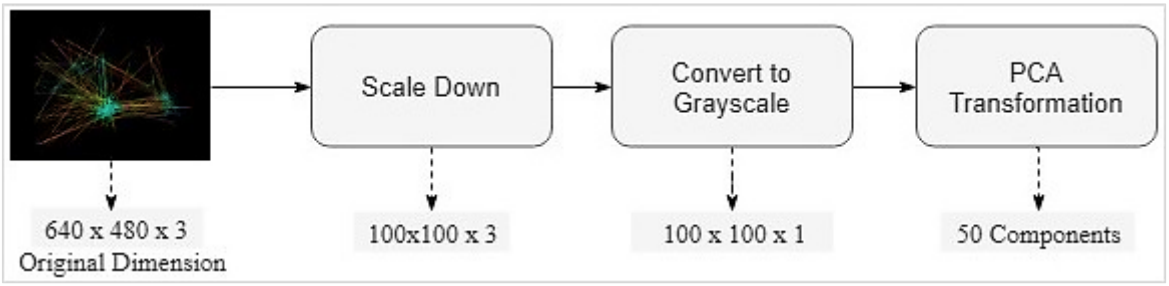

Figure 5: The procedures of dimensionality reduction.

4.2 Dimensionality Reduction

The reduction of problem dimensionality is a vital part in the development of ML models. The dimensionality refers to the number of variables (i.e. features) under consideration. In our case, the initial dimensionality was 640*480*3 (i.e. image size* RGB components). This translates into more than 900K features to be considered, which can complicate the model training and largely increases the possibility of overfitting as well.

To simplify the learning process, a set of image processing techniques was applied in sequence as follows. First, all images were consistently scaled down to 100x100 dimensions. It was expected that such resizing would not lead to much loss of information since most images contained a large blank space of unused pixels. The impact of new dimensions was also examined in the initial ML experiments.

Second, the images were converted into a grayscale format for further simplification. The grayscale transformation reduces the image representation by eliminating the hue and saturation information while retaining the luminance. Specifically, the grayscale values were computed by forming a weighted sum of the R, G, and B components as in the equation below. This contributed to reducing the visual representation from 100x100x3 to 100x100x1. It turned out that the grayscale spectrum was mostly sufficient to discriminate the eye-tracking patterns in terms of velocity, acceleration, and jerk.

*Luminance=0.299*R + 0.587*G + 0.114*B*

Where R, G and B represent the values of the Red, Green and Blue components, and the coefficients are used to calculate luminance (ITU, 2017).

Eventually, Principal Component Analysis (PCA) was implemented to transform grayscale images into a more compressed format. Using orthogonal transformations, PCA attempts to convert a possibly correlated set of data (e.g. signals or images) into a linearly uncorrelated set of reduced dimensions. PCA is one of the most popular techniques for dimensionality reduction that has been widely applied in problems dealing with data of high dimensionality such as image compression (Du and Fowler, 2007), and face recognition (Draper et al., 2003). In our case, the 10K feature set was transformed into 50 components. This significantly reduced the dimensionally into less than 1% of the original dataset. The number of components was empirically decided based on the model accuracy. Figure 5 summarizes the pre-processing procedures along with the dimensions output from each step.

5 EXPERIMENTAL RESULTS

The experimental results are divided into three sections as follows. Initially, we aimed to develop a binary classifier that can basically predict the two categories of participants. Subsequently, the accuracy of a multi-label classification model was experimented. Further, a simple web-based tool is presented as a practical demo that can be used during the diagnosis process.

5.1 Binary Classifier

We conducted our experiments using several ML models. Initially, non-neural network approaches were tested including: Naive Bayes, Logistic Regression, SVM, and Random Forests. Those models were implemented using the Scikit-Learn library (Pedregosa et al., 2011). Generally, the accuracy realized by that category of models was relatively fair (AUC ≈ 0.7).

Subsequently, the model was experimented using various Artificial Neural Network (ANN) structures as follows. Initially, the simplest model structure included a single hidden layer of 50 neurons. The complexity of the model was gradually increased by adding more neurons (e.g. 200, 500).

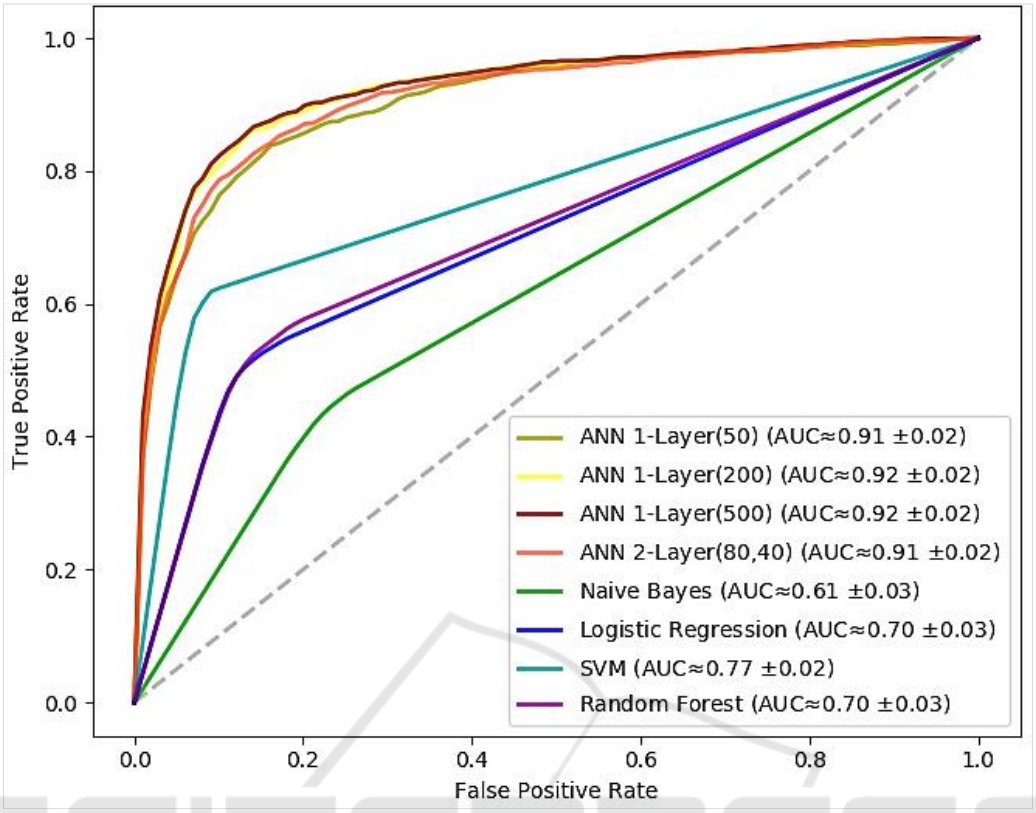

Figure 6: ROC curves of the binary classification models.

Subsequently, another hidden layer was included in the model. The two layers consisted of 80 and 40 neurons respectively. It turned out the there was no need to introduce further complexity in the model based on the prediction accuracy as given in the next section. Table 3 lists the ANN architectures included in our experiments. The ML experiments were implemented using the Keras library (Chollet, 2015) with Python. The models were trained based on 10 rounds of cross-validation including 100 epochs and 20% dropout.

The classification accuracy is analyzed based on the Receiver Operating Characteristics (ROC) curve. The ROC curve plots the relationship between the true positive rate and the false positive rate across a full range of possible thresholds. Figure 6 plots the ROC curves for the set of ANNs experimented. The figure also shows the approximate value of the area under the curve along with its standard deviation over the 10-fold cross-validation.

At it appears, the neural network models obviously outperform other approaches. All neural networks provided a notable prediction accuracy that went beyond 90%. Specifically, the single-layer model of 200 could yield the best performance (≈92.0%).

However, it is noteworthy that there was no substantial improvement by growing the model complexity either by increasing the number of neurons or stacking more hidden layers. Thus, we can say that a single-layer neural network was sufficient in our case, which is a promising outcome using a relatively limited dataset.

Table 3: ANN architectures.

	Model Architecture	
	Hidden Layers	Number of Neurons
Experiment #1	Single-Layer	50
Experiment #2		200
Experiment #3		500
Experiment #4	Two-Layer	(80, 40)

5.2 Multi-Label Classifier

A finer classification of ASD-diagnosed participants was applied to allow for training a multi-label model. We followed the basic grouping that describe

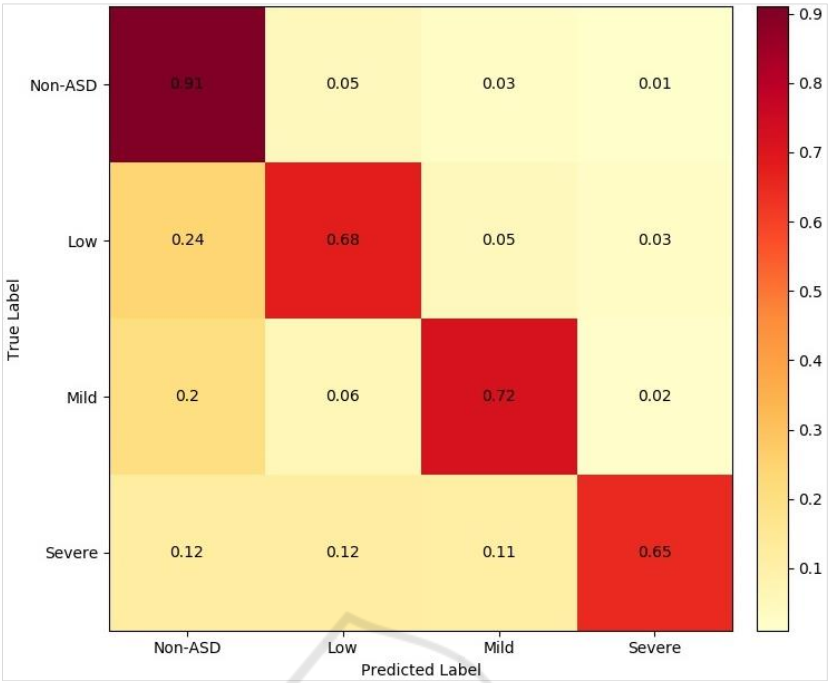

Figure 7: Confusion matrix of the multi-label classifier.

the severity of autism symptoms based on the CARS score (Schopler et al., 1985). Specifically, ASD participants were organized into smaller groups as follows: i) Low, ii) Mild, and iii) Severe. Table 4 gives the specific categoreis of ASD participants based on the CARS score.

Table 4: Classification of ASD participants.

ASD Category	Range of CARS Score
Low	CARS < 30
Mild	30<= CARS < 36
Severe	CARS >= 36

The multi-label classification model was trained using neural networks only. We experimented single-layer model of 200 neurons and two-layer model as before. The average accuracy of the single-layer model was still higher (≈83%) compared to about 81%.

Though the accuracy relatively declined, the approach still proved very good performance. Figure 7 provides a confusion matrix that visulaizes the normalized classification accuracy (single-layer model). The model turned out to discriminate the non-ASD labels very well compared to others. The prediction accuracy of ASD labels was 20% lower (at least), especially for the severe-ASD examples.

5.3 Demo Application

A demo application was developed to serve as a practical illustration of our approach. The application links the three components altogether including eye-tracking, visualization and ML to support the diagnosis of ASD. The application includes three layers as: i) Presentation, and ii) Web services, and iii) Prediction.

The presentation layer performs the basic user interface functionality and interactivity. The presentation elements were delivered using ASP.NET along with JavaScript. The layers of web services and prediction were fully implemented by the Azure ML Studio. Specifically, Azure ML is employed to host the classification model and the Python implementation used to produce visualizations. The Azure ML platform provides an ideal environment for data analytics with the ability to deploy ML models as web services. In this manner, ML models can be conveniently used through web services using request/response calls. Figure 8 sketches the application architecture.

The application goes through three steps as follows. First, the user is asked to upload the eye-tracking data. The data records should describe the coordinates of the viewer's gaze into the screen along with associated time as shown earlier in Table2. Second, the application produces a

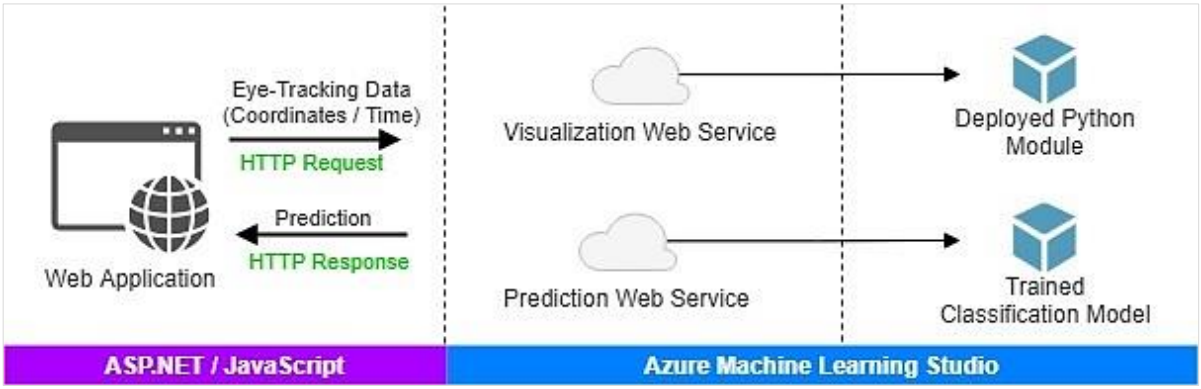

Figure 8: The demo application architecture.

visualization of the eye-tracking records. The visualization is constructed through calling a web service hosted on the Azure ML Studio. The web service executes a Python module deployed to create visualizations of the eye movement and its dynamics.

Eventually, the application calls the prediction web service, which returns the prediction from the trained classification model. All communication with the web services is conducted through standard HTTP requests/response. The application can be accessed online via <https://goo.gl/i4N7Zj>.

6 LIMITATIONS

Even though the results presented in this study are promising, a set of limitations need to be highlighted as follows. The primary limitation could be the relatively small number of participants.

Another relevant issue of concern is the duration of video scenarios, which was fairly short. Perhaps longer scenarios could have allowed for a richer visual representation of the gaze behavior of ASD. Despite limitations, the study is still believed to serve as a kernel for further interesting applications of the proposed approach.

7 CONCLUSIONS

The coupling of eye-tracking, visualization and ML can hold a strong potential for the development of an objective tool to assist the diagnosis of ASD. The ML experiments confirmed the core idea behind our approach, which hinges on the visual representation of eye-tracking scanpaths. The classification accuracy indicated that visualizations were able to successfully pack the information of eye motion and its underlying dynamics.

From a practical standpoint, it is noteworthy that we could reach that high accuracy with largely simple ML models. Using simple ANN Classifiers, the prediction accuracy could go beyond 90%. This should be compared positively to related efforts that used different sets of features and more complex models (e.g. Wan et al., 2018; Carette et al., 2017).

It is conceived that our approach might be applicable to comparable diagnostic problems. In a broader sense, the visualization of eye-tracking scanpaths could possibly be utilized for assisting the diagnosis of similar psychological disorders.

ACKNOWLEDGMENTS

This work has been generously supported by the Evolucare Technologies company. Run by a team of medical information technology professionals, Evolucare aspires to bring a broad range of products and services to the whole healthcare community. www.evolucare.com.

REFERENCES

Carette, R., Cilia, F., Dequen, G., Bosche, J., Guerin, J.-L. and Vandromme, L. (2017). Automatic autism spectrum disorder detection thanks to eye-tracking and neural network-based approach. *In Proceedings of the International Conference on IoT Technologies for HealthCare (HealthyIoT)*. Springer.

Carette, R., Elbattah, M., Dequen, G., Guérin J.L. and Cilia F. (2018). Visualization of eye-tracking patterns in autism spectrum disorder: method and dataset. *In Proceedings of the 13th International Conference on Digital Information Management*. IEEE.

Chollet, F. (2015). Keras. https://github.com/fchollet/keras

Coonrod, E. E. and Stone, W. L. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants & Young Children*, 17(3), 258–268.

DOH-U.S. (2018). Data and statistics | Autism Spectrum Disorder (ASD). URL: https://www.cdc.gov/ncbddd/autism/data.html

Draper, B. A., Baek, K., Bartlett, M. S. and Beveridge, J. R. (2003). Recognizing faces with PCA and ICA. *Computer Vision and Image Understanding*, 91(1-2), 115-137.

Du, Q. and Fowler, J. E. (2007). Hyperspectral image compression using JPEG2000 and principal component analysis. *IEEE Geoscience and Remote Sensing Letters*, 4(2), 201-205.

Figshare (2018). URL: https://figshare.com/s/5d4f93395cc49d01e2bd

Frazier, T. W., Klingemier, E. W., Beukemann, M., Speer, L., Markowitz, L., Parikh, S., … and Strauss, M. S. (2016). Development of an objective autism risk index using remote eye tracking. *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(4), 301–309.

Goldberg, J.H. and Helfman, J.I., (2010). Visual scanpath representation. *In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (203-210). ACM.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498-504.

Hunter, J.D., (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90-95.

ITU. (2017). Rec.ITU-R BT.601-7 Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios. International Telecommunication Union-Radiocommunication Sector, Geneva.

Jacob, R. (1995). Eye tracking in advanced interface design. In W. Barfield W and T.A. Furness (eds), *Virtual Environments and Advanced Interface Design* (pp. 258–288). New York: Oxford University Press.

Jones, E., Gliga, T., Bedford, R., Charman, T. and Johnson, M.H. (2014). Developmental pathways to autism: a review of prospective studies of infants at risk. *Neuroscience & Biobehavioral Reviews*, 39, 1-33.

Jones, W. and Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*, 504(7480), 427-431.

Majaranta, P. and Bulling, A. (2014). Eye tracking and eye-based human–computer interaction. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J.P. Hansen, A. Hyrskykari and K.J. Räihä (Eds.), *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies* (pp. 39-65). Hershey: IGI-Gloal.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825-2830.

Perez, L. and Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv:1712.04621*.

Pierce, K., Conant, D., Hazin, R., Stoner, R. and Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry*, 68(1), 101–109.

Pusiol, G., Esteva, A., Hall, S., Frank, M., Milstein, A. and Fei-fei, L. (2016). Vision-based classification of developmental disorders using eye-movements. *In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 317–325). Springer.

Sepeta, L., Tsuchiya, N., Davies, M. S., Sigman, M., Bookheimer, S. Y. and Dapretto, M. (2012). Abnormal social reward processing in autism as indexed by pupillary responses to happy faces. *Journal of Neurodevelopmental Disorders*, 4(17), 1–9.

Schopler, E., Reichler, R.J., DeVellis, R.F. and Daly, K., (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10(1), 91-103.

Vabalas, A. and Freeth, M. (2016). Brief report: patterns of eye movements in face to face conversation are associated with autistic traits: evidence from a student sample. *Journal of Autism and Developmental Disorders*, 46(1), 305–314.

Wallace, S., Coutanche, M. N., Leppa, J. M., Cusack, J., Bailey, A. J. and Hietanen, J. K. (2012). Affective – motivational brain responses to direct gaze in Children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 7(53), 790–797.

Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., ... and Lin, Y. (2018). Applying eye tracking to identify autism spectrum disorder in children. *Journal of Autism and Developmental Disorders*, 1-7.

Wing, L. and Gould, J. (1979). Severe impairments of social interactin and associated abnormalities in children: epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9(1), 11-29.

Xu, Y., Jia, R., Mou, L., Li, G., Chen, Y., Lu, Y. and Jin, Z. (2016). Improved relation classification by deep recurrent neural networks with data augmentation. *arXiv:1601.03651*.