

Vision-Based Approach for Autism Diagnosis Using Transfer Learning and Eye-Tracking

Mahmoud Elbattah, Jean-Luc Guérin, Romuald Carette, Federica Cilia, Gilles Dequen
Laboratoire MIS

Université de Picardie Jules Verne (UPJV), France

mahmoud.elbattah@u-picardie.fr

Background: Autism Spectrum Disorder

- Autism Spectrum Disorder (ASD) is a pervasive developmental disorder characterised by a set of impairments including social communication problems.¹
- ASD has been considered to affect about 1% of the world’s population (US Dep. of Health, 2018). ²
- The hallmark of autism is an impairment of the ability to make and maintain eye contact. ³

¹ L. Wing, and J. Gould, “Severe Impairments of Social Interaction and Associated Abnormalities in Children: Epidemiology and Classification”. Journal of Autism and Developmental Disorders, 9(1), pp.11-29, 1979.

² U.S. Department of Health & Human Services. Data and statistics | autism spectrum disorder (asd) | ncbddd | cdc, 2018. URL: <https://www.cdc.gov/ncbddd/autism/data.html>.

³ 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.

Background: Eye-Tracking Technology

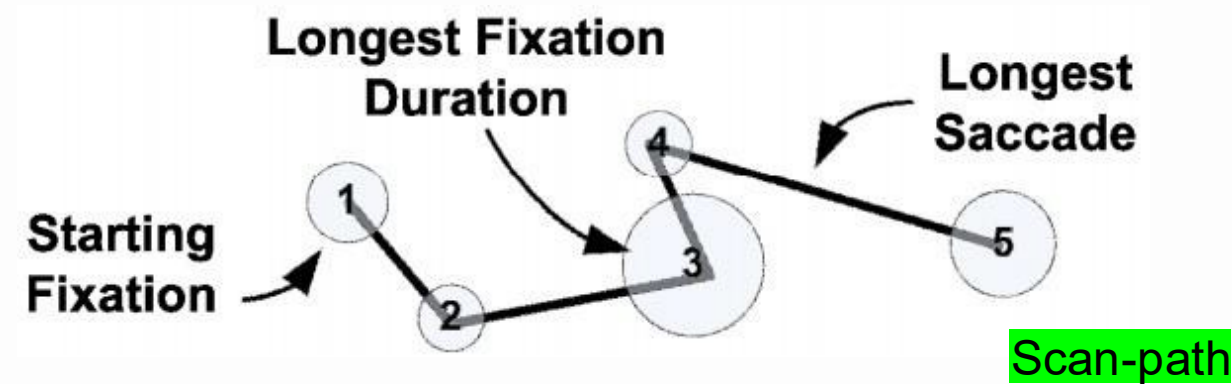
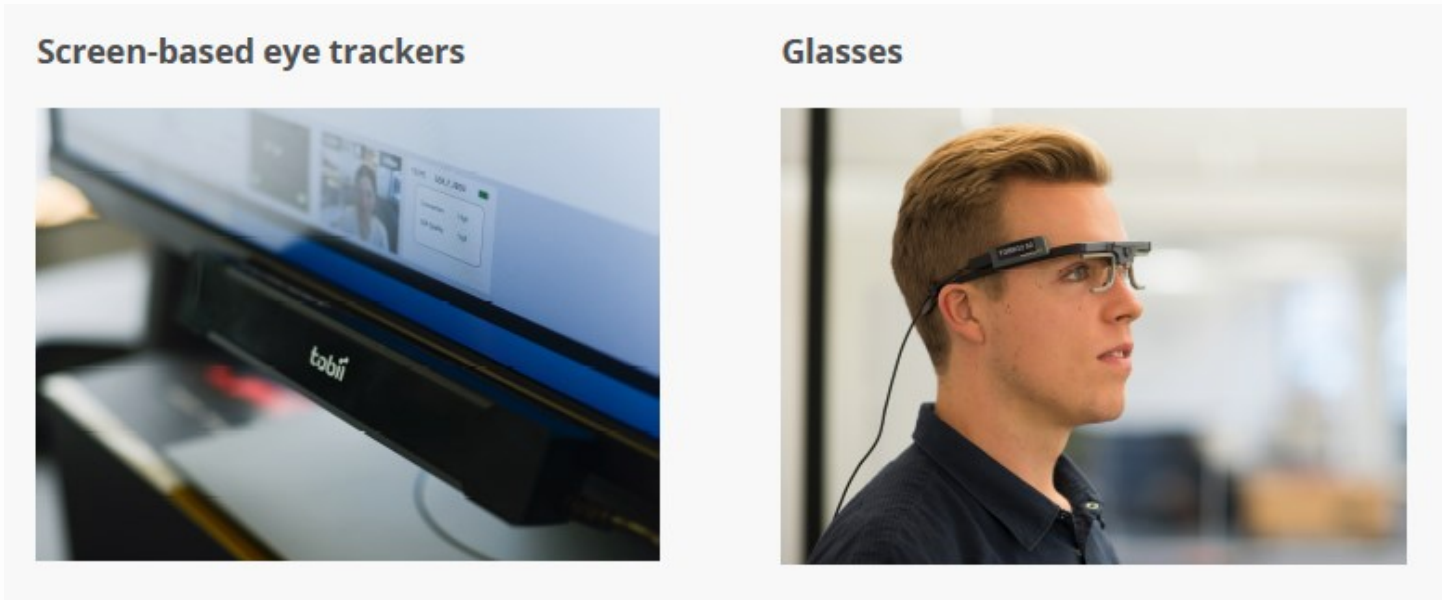


Image Source: <https://imotions.com/blog/eye-tracking/>

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

Motivation

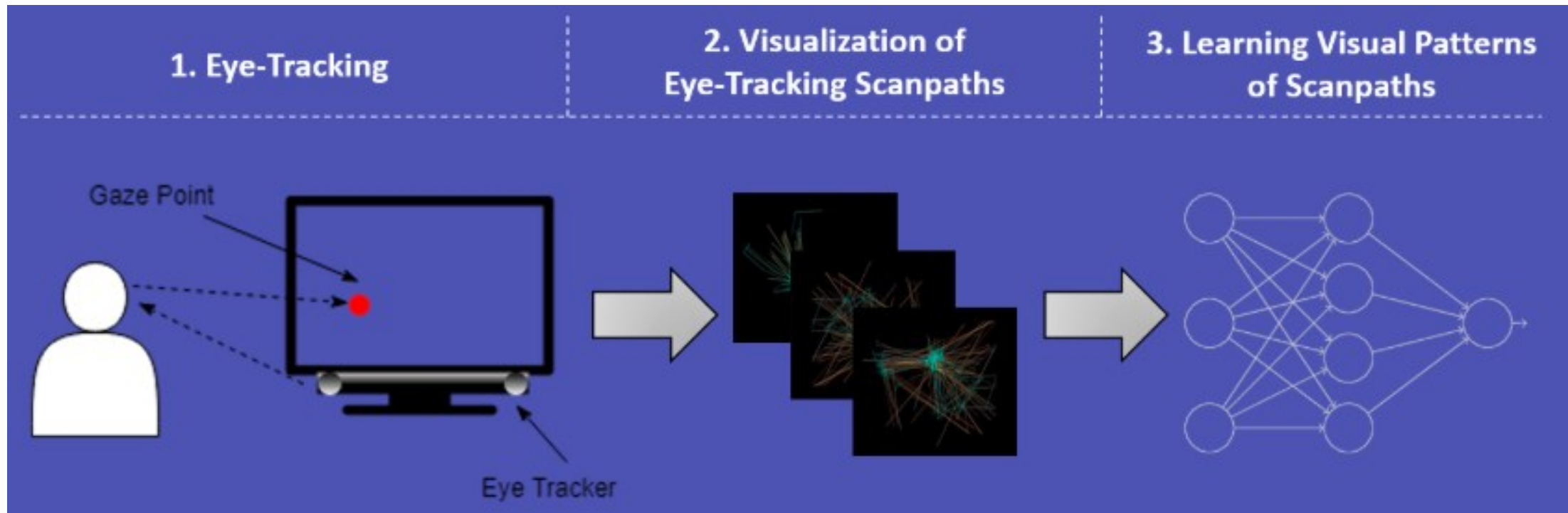
Our Goal:

- Detecting ASD-diagnosed individual in eye-tracking data.

Key Idea:

- Compactly render eye movements into an image-based format while maintaining the dynamic characteristics of eye motion.
- As such, the classification task could be approached as image classification.

Our Earlier Work (HEALTHINF 2019)



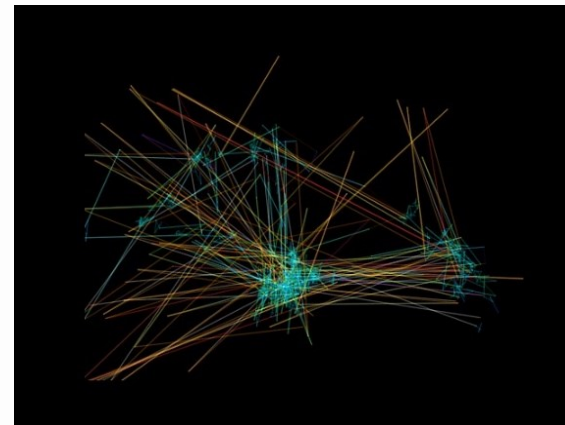
Carette, R., Elbattah, M., Cilia, F., Dequen, G., Guérin, J., & Bosche, J. (2019). Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths. *In Proceedings of the 12th International Conference on Health Informatics (HEALTHINF)*.

Data Description

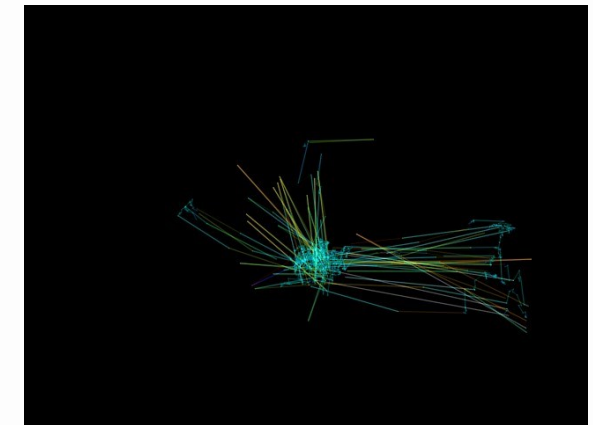
Number of Participants (ASD, TD)	59 (29, 30)
Gender Distribution (M, F)	38 (\approx 64%), 21 (\approx 36%)
Age (Mean, Median) years	7.88, 8.1
CARS Score (Mean, Median)	32.97, 34.50

Image Dataset Description

- 547 images: 328 (Non-ASD), 219 (ASD)
- Image dimensions: 640x480



ASD



Non-ASD

Transfer Learning Models (Feature Extraction)

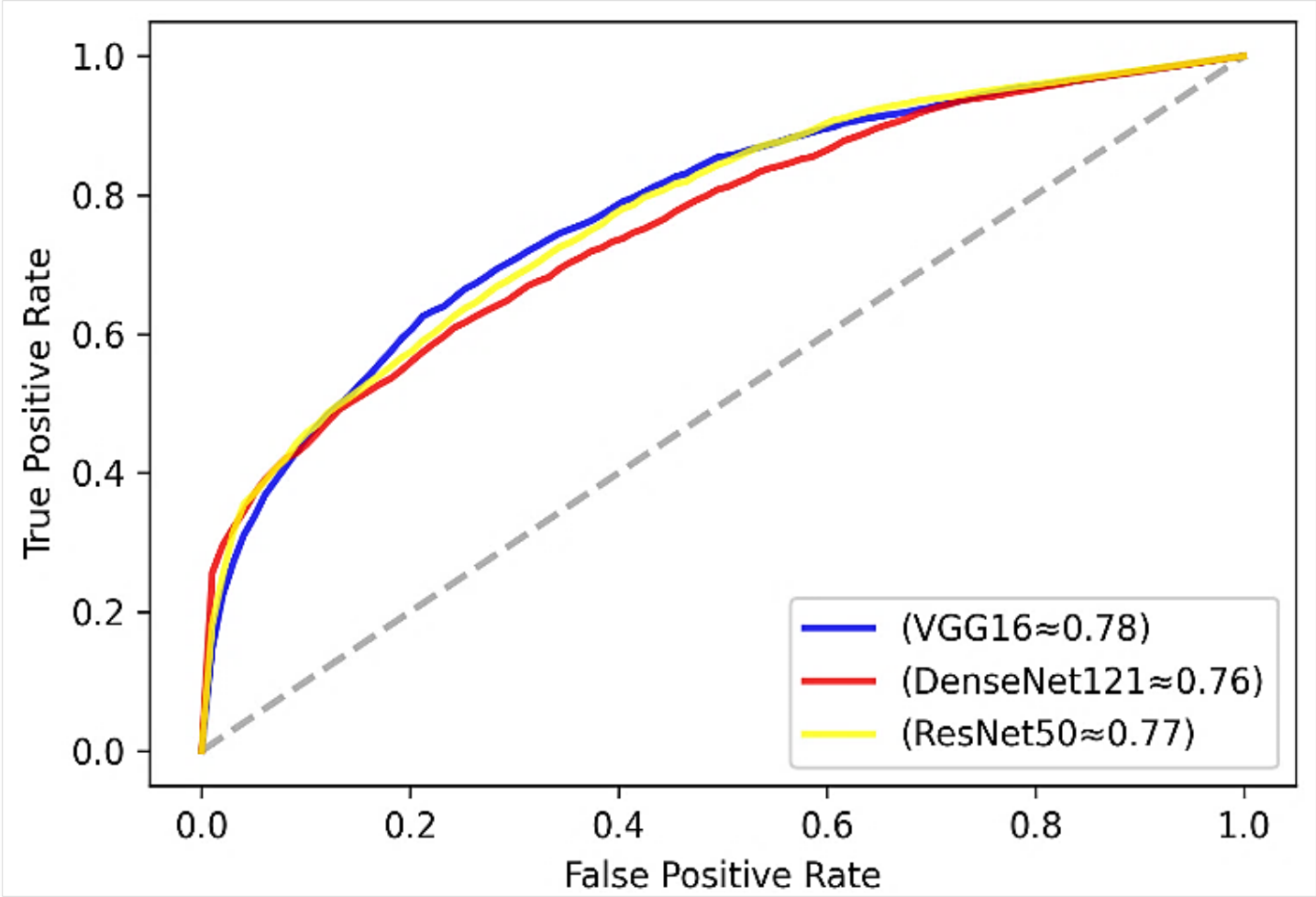
- **VGG-16:** A deep CNN architecture developed by a group of researchers from the University of Oxford (Simonyan, and Zisserman, 2014).
- **ResNet:** Residual Networks (ResNet) (He et al., 2016).
- **DenseNet:** Densely Connected Convolutional Network (DenseNet) architecture (Huang et al., 2017).

Dataset Splitting (3-Fold Cross-Validation)

The dataset was split using the following stepwise procedures:

1. **Split Participants:** Initially, the group of 59 participants was randomly split into two independent sets (i.e. train and test).
2. **Match Images:** Based on the IDs of participants, the images were matched and loaded into the train and test sets.
3. **Repeat:** Step #1 and Step #2 would be repeated for each round of the cross-validation process.

Experimental Results



Experimental Results (cont'd)

AVERAGE PRECISION-RECALL OF MODELS.

Model	Recall (~)	Precision (~)
VGG-16	0.56	0.67
ResNet	0.54	0.65
DenseNet	0.55	0.65

Conclusions and Limitations

- Popular vision models such as VGG-16, ResNet, and DenseNet could achieve a quite promising performance.
- The TL approach was largely applicable, even though the source dataset (i.e., ImageNet) is assumed to have included quite different types of images.
- It is not claimed that the TL approach could provide superior performance.
- However, it is conceived that the scarcity or imbalance of datasets could make such TL approaches attractive for further investigation.

Thank You!
