Predicting Autism Spectrum Disorder (ASD) using Eye Tracking Data and Deep Learning

Suhas Nagaraj suhas99@umd.edu
Swaraj Mundruppady Rao swarajmr@umd.edu
FNU Koustubh koustubh@umd.edu
Arya Sunil Gangadkar agangadk@umd.edu

UNIVERSITY OF MARYLAND
FEARLESSLY FORWARD

Problems Statement

Problem Statement

- Detecting Autism Spectrum Disorder(ASD) is a complex process. There is no definitive medical test to diagnose ASD.
- Traditionally, Childhood Autism Rating Scale (CARS) is calculated using behavioural tests to diagnose autism.
- Studies have shown that fixation time percentage of children with ASD showed significant decrease compared with that of TD children.
- The Eye tracking dataset, tries to leverage this diagnostic marker to address the classification problem.
- We have used the eye tracking dataset to come up with a model that can easily help predict ASD in children at an early age.

Technical Challenges

Technical Challenges

- **Limited Dataset size:** The dataset only consists of data from 59 individuals which is very small for generalizing the model.
- **Noise in Dataset:** Variability and outliers, such as image data from individuals diagnosed as ASD-positive but displaying non-ASD characteristics (and vice versa)
- **Difficulty in collecting new data**: Due to special equipment used in creating the dataset.
- Difficulty in re-implementation of baseline described in the reference papers due to unavailability of critical parameters

Related Works

Related Work

- Techniques involving functional Magnetic Resonance Imaging (fMRI) scans have also been explored, with models like ASD-DiagNet [1] achieving reported accuracies of up to 82%.
- Behavioral patterns, such as self-injurious behaviors [2], along with facial expression analysis using deep learning [3], have demonstrated significant success.
- Using these models can be difficult because the person has to take many tests such as fMRI scans and behavioral tests that can take too long and might cause discomfort and skew the results.
- Eye tracking is relatively easy and quick to setup!

Related Work

- **Baseline**: Machine learning model investigating eye-tracking datasets proposed by Akter et al. [4], using MLP classifiers and clustering got a maximum accuracy of **74.2%** on complete dataset.
- Arora et al.[5] proposed a predictive tool for ASD using eye gaze data, achieving an accuracy of 74.57% with ANN and 85.28% with CNN models.
- Carette et al. [6], presented a method for visualizing eye-tracking patterns in ASD, utilizing a pre trained Xception model for feature extraction and a stacking ensemble framework.

Dataset and EDA

Dataset and EDA

- The dataset involves recording eye tracking data of 59 participants out of which 29 were diagnosed with ASD and 30 participants were classified as neurotypical.
- The eye tracking data was collected using SMI RED-m eye tracker paired with a display monitor showing videos.
- Eye movement data collected over time, was compressed into RGBA images, creating a dataset encapsulating both spatial and temporal gaze dynamics

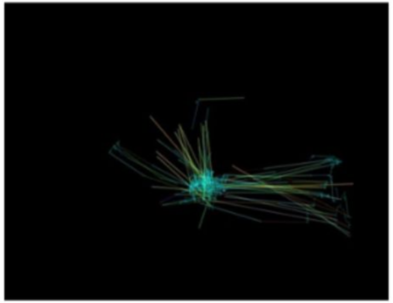


Figure 1 : non-ASD diagnosed participant

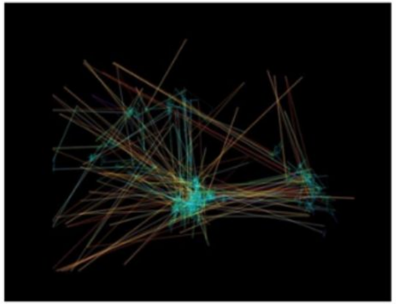
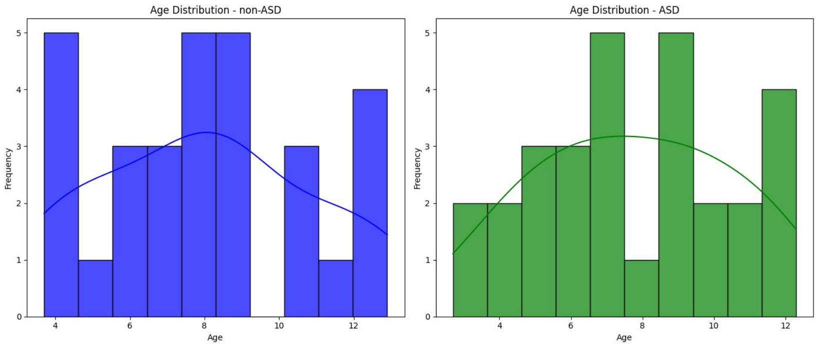


Figure 2: ASD-diagnosed participant

UNIVERSITY OF MARYLAND FEARLESSLY FORWARD

Dataset and EDA

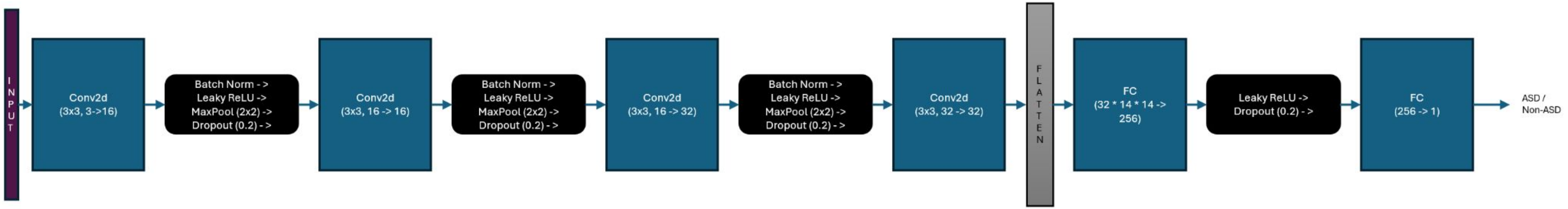
- A total of 547 images were generated, with 219 images representing children diagnosed with ASD and 328 images corresponding to neurotypical participants.
- The participants ranged from approximately 3 to 13 years of age, with a mean age of 7.8 years
- The dataset comprised 38 male and 21 female participants. Among ASD-labeled participants, 76% were male, reflecting the higher prevalence of ASD in males.



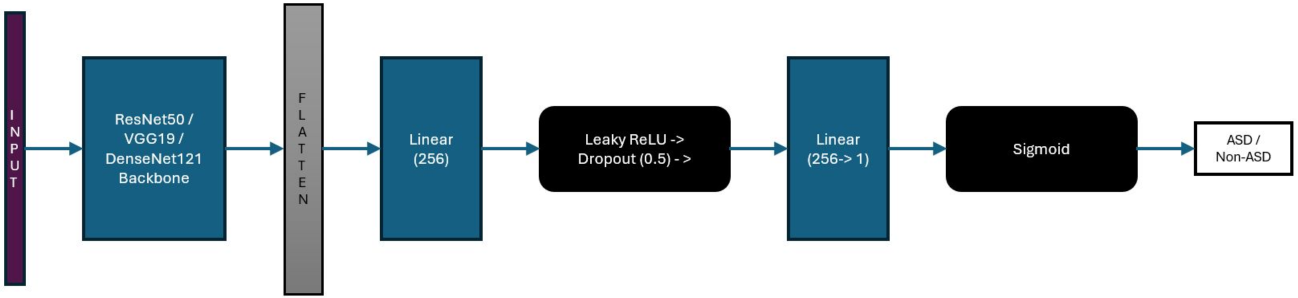
Methodology

Image Level Classification Model

Custom CNN model

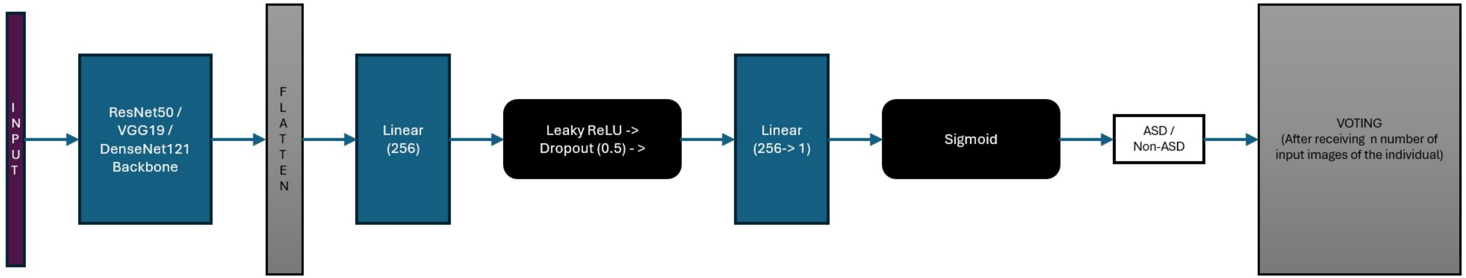


Transfer Learning Models

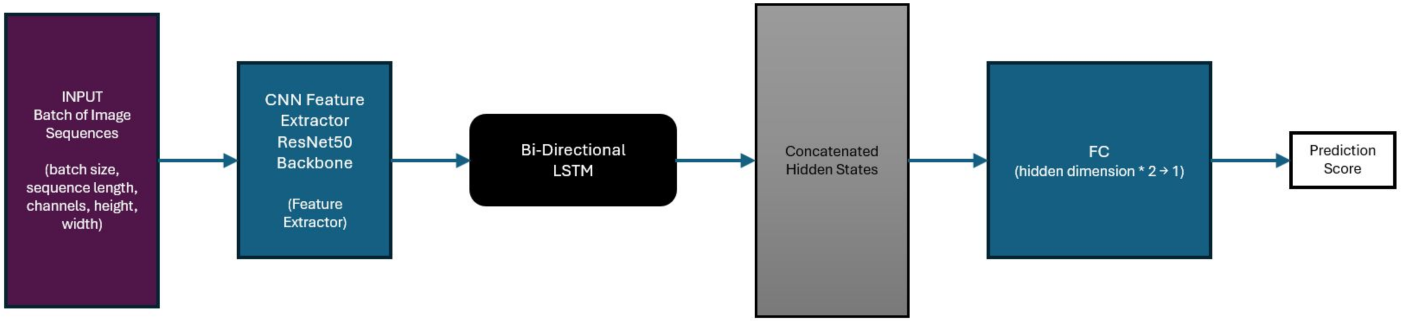


Individual Level Classification

Feature Extractor(ResNet) + Voting Model

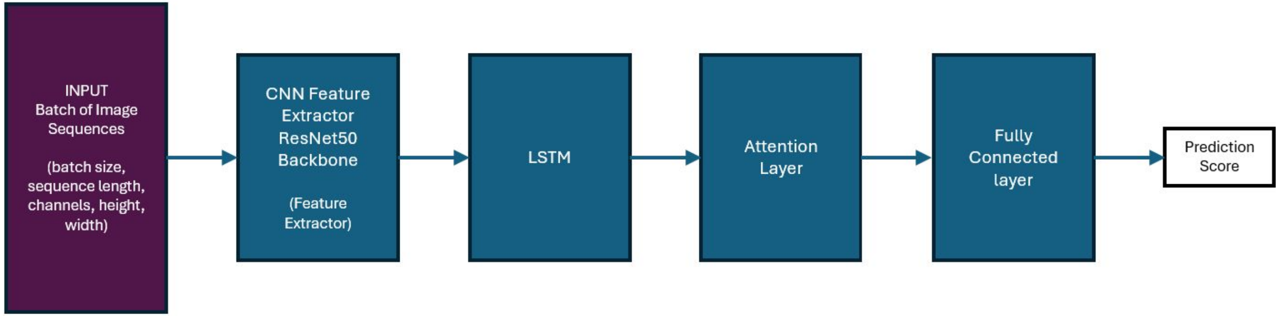


Sequence modelling with LSTM

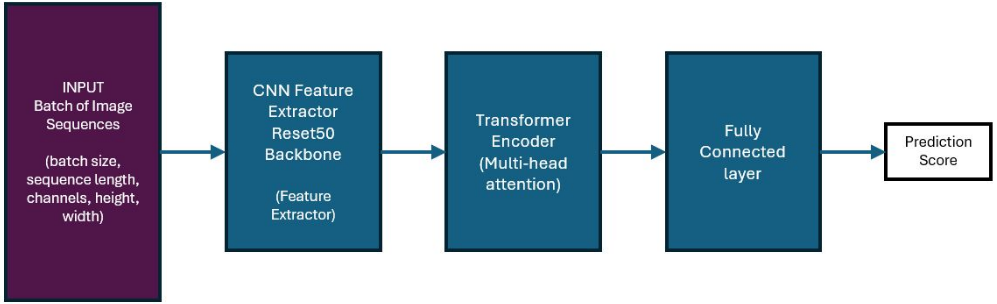


Individual Level Classification Models

Attention Augmented LSTM



Transfer Learning CNN layers + Transformers



Experiments

Experiments

Experimental Setup

- To prevent data leak, the dataset was split at individual level ensuring no similar images from one person were shared by training, validation and test sets.
- Participants were divided into 70% training, 15% validation, and 15% testing subsets, maintaining an equal ratio of positive and negative data points in each subset.
- Data augmentation was applied exclusively to the training set to enhance generalization.
- Preprocessing involved resizing images to a uniform size of 224 x 224 and normalizing pixel values using the mean and standard deviation computed from the training set.

Experiments

Evaluation Metrics and Hyperparameter tuning

- The performance of the models were evaluated using metrics such as accuracy, AUC, F1 score and confusion matrices.
- Hyperparameter tuning was performed using grid search to maximize validation accuracy and model generalization.
- Parameters tuned included the number of layers, layer dimensions, kernel sizes, learning rate, weight decay, dropout, batch size, and optimizer type.
- Techniques like Learning rate decay and early stopping were used to get better and generalized training performance.

Results

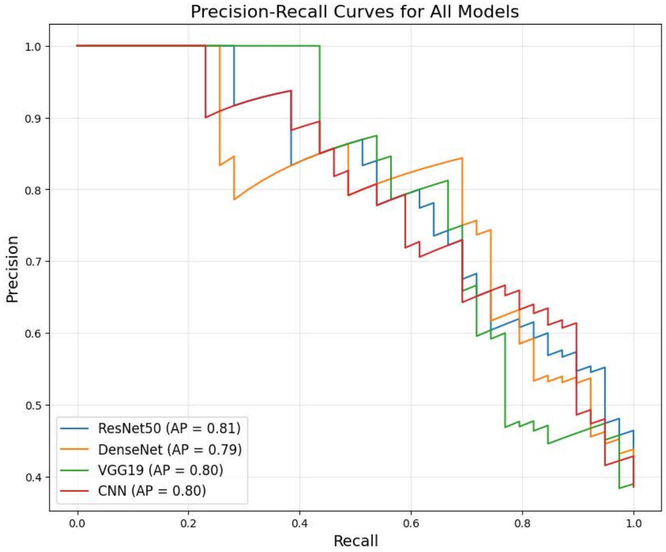
Results

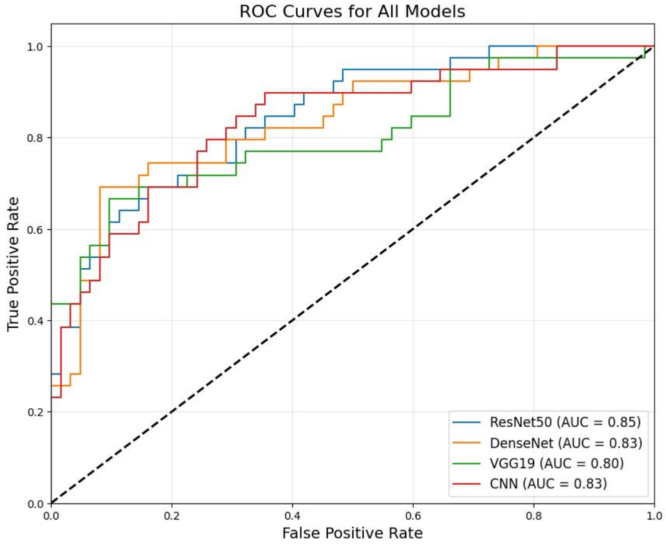
Image Level Classification Models

Method/ Model	Baseline	Custom CNN	ResNet50	DenseNet 121	VGG19
Accuracy	0.742	0.7624	0.7822	0.7921	0.7228
AUC	0.715	0.8346	0.8465	0.8341	0.8032
F1 Score	0.736	0.49056	0.6764	0.72	0.6666

Results

Image Level Classification Models





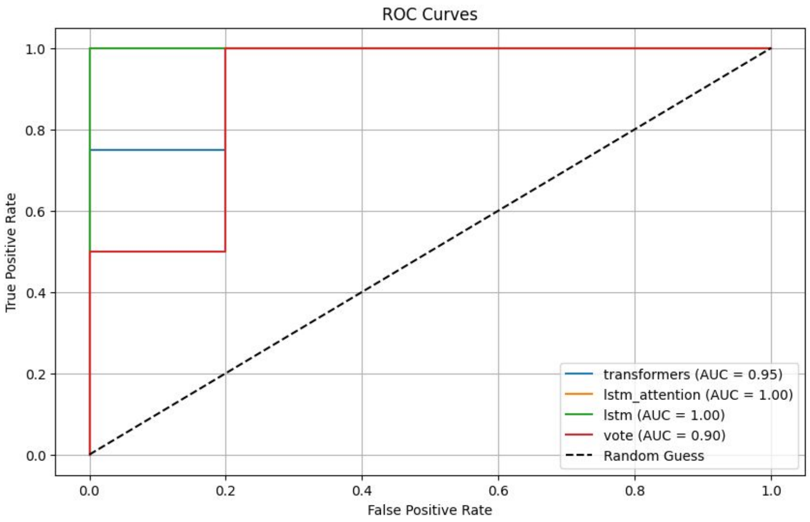
Results

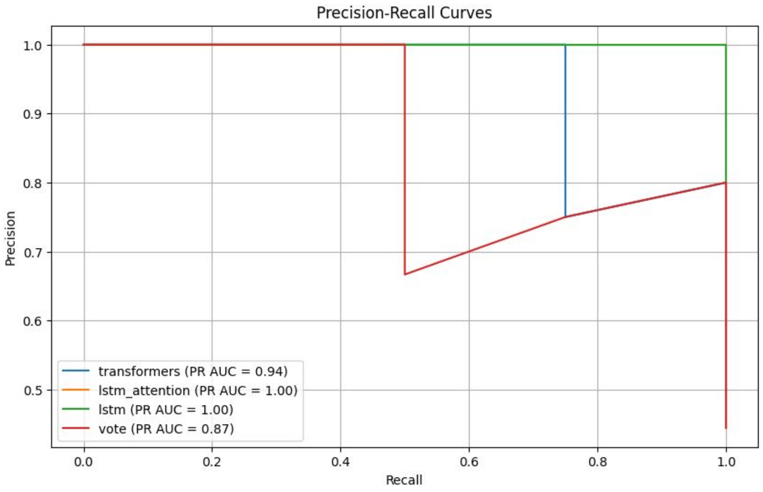
Individual Level Classification Models

Method/ Model	CNN + Voting	CNN + LSTM	CNN + LSTM + Attention	CNN + Transformer
Accuracy	0.7778	0.7778	0.7778	0.8889
AUC	0.9000	1.000	1.000	0.95
F1 Score	0.6153	0.8000	0.8000	0.888

Results

ROC and PR curves **individual level classification** using voting, LSTM, LSTM with attention and transformers based models





Results

Individual Level Classification Method Validation

Broader Impact, Conclusion, Limitations

Conclusion, Limitations and Broader Impact

Conclusion

- Individual level classification is a better approach compared to image level classification as we diagnosing individuals and not the images.
- The CNN + transformers show the most promising results.

Limitations

- The dataset is small and might result in overfitting, transformers require a greater amount of data to generalise better.

Broader Impact

- In our study we have shown how the transfer learning and Attention based hybrid models can be very powerful tool in ASD detection.
- The same tools can be used in other medical imaging based problems.

References

[1] Eslami T, Mirjalili V, Fong A, Laird AR, Saeed F. ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data. Front Neuroinform. 2019;13:70. Published 2019 Nov 27. doi:10.3389/fninf.2019.00070

[2] Cantin-Garside, K.D., Kong, Z., White, S.W. et al. Detecting and Classifying Self-injurious Behavior in Autism Spectrum Disorder Using Machine Learning Techniques. J Autism Dev Disord 50, 4039–4052 (2020). https://doi.org/10.1007/s10803-020-04463-x

[3] Beary, M., Hadsell, A., Messersmith, R., & Hosseini, M. (2020). Diagnosis of Autism in Children using Facial Analysis and Deep Learning. ArXiv, abs/2008.02890.

[4] T. Akter, M. H. Ali, M. I. Khan, M. S. Satu, and M. A. Moni, ”Machine Learning Model To Predict Autism Investigating Eye-Tracking Dataset,” 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 2021, pp. 383-387, doi:10.1109/ICREST51555.2021.9331152\\\

[5] E. Arora, H. S. Jolly and A. Rehalia, ”Prediction of Autism Spectrum Disorder Using ANN and CNN,” 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2023, pp. 457- 461, doi: 10.1109/ICTACS59847.2023.10390530.

[6] Carette et al. [1] presented a method for visualizing eye-tracking patterns in Autism Spectrum Disorder (ASD), utilizing a pretrained Xception model for feature extraction and a stacking ensemble framework

Thank You!