IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DETECTING AUTISM IN CHILDREN USING EYE TRACKING

[1]Prof. Amruth V, [2]Adithya S L, [3]Charishma L, [4]Srusti C K, [5]Vaishnavi YV

[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student
Department of Information Science and Engineering,
Maharaja Institute of Technology Mysore, Karnataka, India

Abstract: Autism spectrum disorder (ASD) is a neurodevelopmental disorder that affects social interaction, communication, and behavior. Eye tracking technology has been widely used in autism research because individuals with ASD often exhibit distinctive patterns of eye movements. Deep learning models like ResNet50 and VGG16 are powerful convolutional neural networks that can extract meaningful features from images. These models have been successful in various computer vision tasks, including object recognition. The combination of eye tracking and deep learning techniques can potentially provide valuable insights into early autism detection. By analyzing eye movement patterns captured by eye tracking devices, and leveraging the capabilities of deep learning models like ResNet50 and VGG16, researchers can develop algorithms to detect specific patterns or abnormalities that are indicative of autism in children. These algorithms can learn from large datasets of eye tracking data collected from both typically developing children and those diagnosed with autism. By training the deep learning models on these datasets, they can learn to recognize patterns that are characteristic of ASD and distinguish them from typical eye movement patterns. The ultimate goal of such research is to create a reliable and accurate tool for early autism detection, which could aid in early intervention and improve outcomes for children with ASD. The results of our experiments demonstrate outstanding performance with accuracy scores of 97%, and 98.7% for the ResNet50, and VGG16 models respectively.

***Index Terms* - Autism spectrum disorder (ASD), Eye tracking, Deep learning, Convolutional neural networks (CNNs), ResNet50, VGG16**

I. INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder that affects social interaction, communication, and behavior. Early detection and intervention are crucial for improving outcomes and the quality of life for individuals with ASD. In recent years, there has been growing interest in utilizing eye tracking technology coupled with deep learning models, such as ResNet50 and VGG16, to aid in the detection and diagnosis of autism in children. Eye tracking is a technique that involves measuring and analyzing eye movement patterns and gaze behavior. Children with ASD often exhibit distinct patterns of eye movements, such as reduced eye contact, atypical scanning patterns, and gaze fixation preferences. These eye movement patterns can provide valuable insights into the underlying cognitive and perceptual processes associated with ASD.

Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated remarkable success in various computer vision tasks, including image recognition and object detection. Models like ResNet50 and VGG16 have proven to be effective in extracting meaningful features from images and have achieved state-of-the-art performance in many visual recognition tasks. The combination of eye tracking and deep learning techniques holds promise for developing robust and accurate algorithms to detect autism in children. By leveraging the power of deep learning models like ResNet50 and VGG16, researchers aim to automatically analyze and interpret eye movement data captured through eye tracking devices. These models can learn to identify specific eye movement patterns or abnormalities that are indicative of ASD, distinguishing them from typical eye movement patterns observed in typically developing children.

The objective of this research is to develop a reliable and objective tool for early autism detection using eye tracking technology and deep learning models. Such a tool could potentially assist clinicians and researchers in identifying children at risk of ASD and facilitate early intervention strategies. Furthermore, the insights gained from studying eye movement patterns in children with ASD could contribute to a better understanding of the underlying cognitive and perceptual mechanisms associated with the disorder.

II. Literature Survey

In this paper [1], they proposed a new multimodal feature learning and fusion framework for identifying ASD in children. Its core idea is a two-step multimodal learning model, where at the first step the high-level EEG and ET feature representations are learned via an EEG-SDAE and an ET-SDAE respectively from initial high-dimensional features with information redundancy, and then the learned EEG and ET feature representations are further fused for final classification via a Fusion-SDAE at the second stage. The proposed model realizes the joint modelling and analysis of EEG and ET data and can learn complementary information between two different modalities and enhance identification performance. Experimental results have demonstrated that the proposed method achieved better identification performance with an overall accuracy of 95.56% than unimodal methods and a simple feature-level fusion method. It should be noted that our proposed framework is data-driven and can automatically learn and fuse useful information from neurophysiological and behavioural modalities to identify ASD without the need of much more diagnostic expert experience. It provides a new tool for an easier and more objective diagnosis of ASD in children, which can assist clinicians to make a precise diagnosis decision and improve diagnosis efficiency, suggesting its great potential in clinical applications.

In this pilot study [2], they computed correlations of acoustic, video, and handwriting time- series derived from five children with ASD and five children with neurotypical development during speech and handwriting tasks. These correlations and eigenvalues derived from the correlations act as a proxy for motor coordination across articulatory laryngeal, and respiratory speech production systems and for fine motor skills. They utilized features derived from these correlations to discriminate between children with and without ASD. Eigenvalues derived from these correlations highlighted differences in complexity of coordination across speech subsystems and during handwriting, and helped discriminate between the two subject groups. These results suggest differences in coupling within speech production and fine motor skill systems in children with ASD. Overall, the differences in discriminative ability of features across different tasks, as well as the differences in the patterns witnessed for different features, suggests that the effect of limited motor coordination on speech production is highly nuanced and subsystem or task dependent. This agrees with studies that indicate that coordination during speech is task dependent, such as differences in coupling between upper and lower-lip movement in 2-year-olds during bilabial and ‘‘nonlabial’’ speech tasks . The differences seen across features and tasks may also be influenced by the emotional content present in the task. As they extend the use of these correlation features to function as objective measures of progress related to speech interventions, it will be important to quantify the relationship between the eigenvalue patterns and changes in the underlying signals.

In this study [3], the proposed model achieves comparable results with other models, yet it has the best AUC results which mean it works well regardless of the class distribution (i.e. when data are new/unknown). Overall, our research provides promising findings in regard to early identification through a computer-aid diagnostic tool. Today, the accuracy of the diagnosis is heavily depending on the clinical expertise and their background experience, hence applying an intelligent model for diagnoses will improve early detection by providing a more objective, cost and effort effective approach. To the best of our knowledge, our study is one of the limited number of researches to address these challenges, especially for Arab populations. The proposed system is intended to positively impact Autism field locally as well globally. In terms of local impact, there is a lack in the systems that are available to assist therapists in Saudi Arabia. Our system can serve them as a diagnostic tool for ASD. It will be designed to provide an Arabic interface and it is planned to be used in Autism centers or hospitals. Globally, this research provides a novel contribution to the HCI research by designing assistive technologies for Arabic populations within healthcare domains.

In this study [4], the aim of this work is to present our automated computer aided diagnostic (CAD) system for accurate identification of autism spectrum disorder based on the connectivity of the white matter (WM) tracts. To achieve this goal, two levels of analysis are provided for local and global scores using diffusion tensor imaging (DTI) data. A local analysis using the Johns Hopkins WM atlas is exploited for DTI atlas-based segmentation. Furthermore, WM integrity is examined by extracting the most notable features representing WM connectivity from DTI. Interactions of WM features between different areas in the brain, demonstrating correlations between WM areas were used, and feature selection among those associations were made. The proposed system was tested on a large dataset of 263 subjects from the National Database of Autism Research (NDAR) with their Autism Diagnostic Observation Schedule (ADOS) scores and diagnosis (139 typically developed: 66 males, and 73 females, and 124 autistics: 66 males, and 58 females), with ages ranging from 96 to 215 months, achieving an overall accuracy of 73%. In addition to this achieved global accuracy, diagnostically-important brain areas were identified, allowing for a better understanding of ASD-related brain abnormalities, which is considered as an essential step towards developing early personalized treatment plans for children with autism spectrum disorder.

In this study [5], the aim of this work is to investigate eye gazing images and identify autism applying various machine learning techniques. Therefore, they collected eye-tracking data from the Figshare data repository. But, these scanpath images were almost similar for normal and autistic children. To obtain similar groups, k-means clustering method was used and generated four clusters. Further, several classifiers were applied into primary data and these clusters and evaluated the performance of them using various metrics. After the assessment of overall results, MLP shows the highest 87% accuracy in cluster 1. In addition, it shows the best area under curve, f-measure, g-mean, sensitivity, specificity, fall out and miss rate respectively. They preprocessed eye-tracking scan paths image set by converting them into a gray scale images, increased their brightness and resized them into 128 x 128 format. Nevertheless, K-means clustering is an unsupervised and iterative algorithm that endeavors to split this dataset into dissimilar clusters. In this method, it identifies k number of centroids and distributes each sample into adjacent clusters and updates the centroid's position by averaging its location of all samples in each cluster. In this work, several classifiers such as DT, GB, KNN, LR, MLP, NB, RF, SVM and XGB were implemented with 10 fold cross-validation technique into four eye-tracking clusters and primary dataset using scikit learn library in python. All works had been implemented into Google Collaboratory in the cloud server. Many evaluation metrics such as accuracy, AUC, f-measure, gmean, sensitivity, specificity, fall out and miss rate are used to justify the performance of different classifiers.

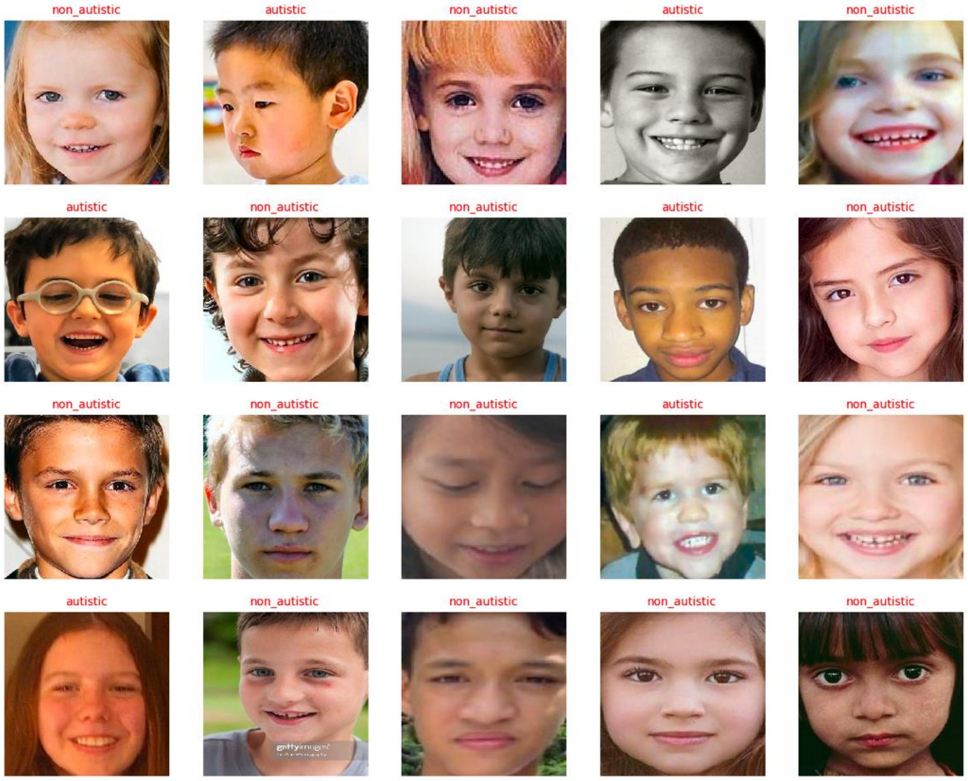
III. METHODOLOGY

The methodology of detecting autism in children using eye tracking is as follows:

- ***Data Collection***

We used a standard data set for children from the Kaggle website, which is for open source data. For training and testing purposes, we used two types of data, which are the types of images.

Fig. 1 Dataset Images



- ***Preprocessing***

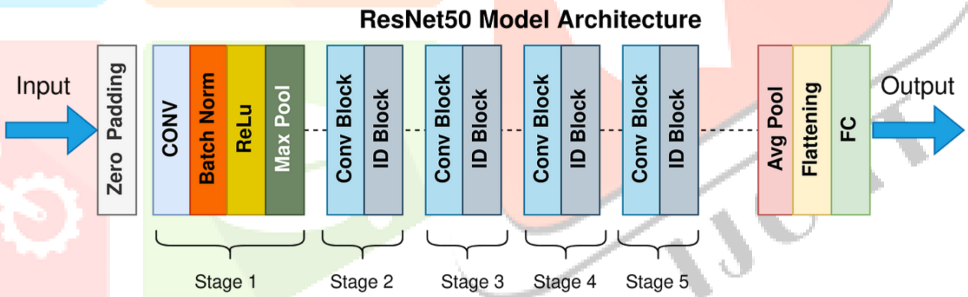
  - Face Alignment: Aligned the extracted face images to a standardized orientation to ensure consistency and reduce variability in subsequent analysis.
  - Resize and Crop: Resized the aligned face images to a consistent size, e.g., 224x224 pixels, as required by ResNet50 and VGG16. Consider cropping the images to remove any irrelevant background or non-facial regions.
  - Image Enhancement: Enhanced the contrast, brightness, and sharpness of the face images to improve their quality and aid feature extraction.
  - Normalization: Normalized the pixel values of the preprocessed face images to a common range, such as [0, 1] or [-1, 1], to ensure consistent representation across the dataset.

- ***ResNet50***

ResNet50 is a deep convolutional neural network architecture that consists of 50 layers, including convolutional layers, pooling layers, fully connected layers, and shortcut connections. The key feature of ResNet50 is the use of residual blocks, which enable the network to effectively train very deep networks without suffering from the degradation problem. The architecture of ResNet50 can be summarized as follows:

- Input: RGB images of size 224x224 pixels.
- Convolutional Layers: The initial layers perform convolutional operations on the input images, gradually capturing low-level features and increasing the complexity of the learned representations.
- Residual Blocks: ResNet50 contains multiple residual blocks, each consisting of a series of convolutional layers with shortcut connections. These connections allow the gradient to propagate directly through the network, addressing the vanishing gradient problem.
- Pooling Layers: Max pooling operations are applied periodically to reduce the spatial dimensions of the feature maps while preserving the most salient features.
- Fully Connected Layers: The output of the convolutional layers is flattened and passed through fully connected layers for classification.
- Output: The final layer is a softmax layer that produces the predicted probabilities for the classes.
- The next step is to design the architecture of the deep learning models. In this case, the FCN and RESNET18 models will be used. Each model will need to be tailored to the specific task of kidney lesion detection, with appropriate input and output layers.

Fig. 3 ResNet50 Architecture

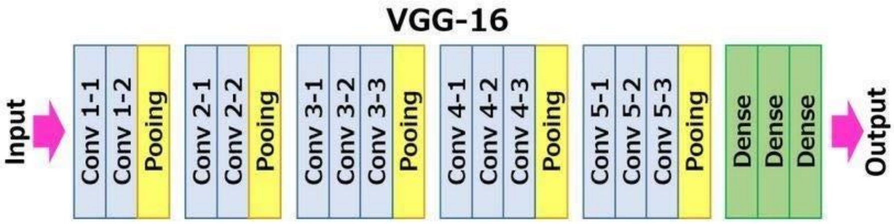


- ***VGG16***

VGG16 is another popular convolutional neural network architecture known for its simplicity and effectiveness. It consists of 16 layers, including convolutional layers, pooling layers, and fully connected layers. The primary characteristic of VGG16 is its use of small-sized filters (3x3) throughout the network, which allows for deeper networks while maintaining a smaller number of parameters. The architecture of VGG16 can be summarized as follows:

- Input: RGB images of size 224x224 pixels.
- Convolutional Layers: The initial layers consist of multiple convolutional layers with small-sized filters (3x3) and a stride of 1. These layers learn hierarchical features, gradually capturing more complex patterns.
- Pooling Layers: Max pooling operations with a stride of 2 are applied periodically to reduce the spatial dimensions of the feature maps.
- Fully Connected Layers: The output of the convolutional layers is flattened and passed through fully connected layers, gradually reducing the dimensions and learning high-level representations.
- Output: The final layer is a softmax layer that produces the predicted probabilities for the classes.

Fig. 4 VGG16 Architecture



- ***Training***

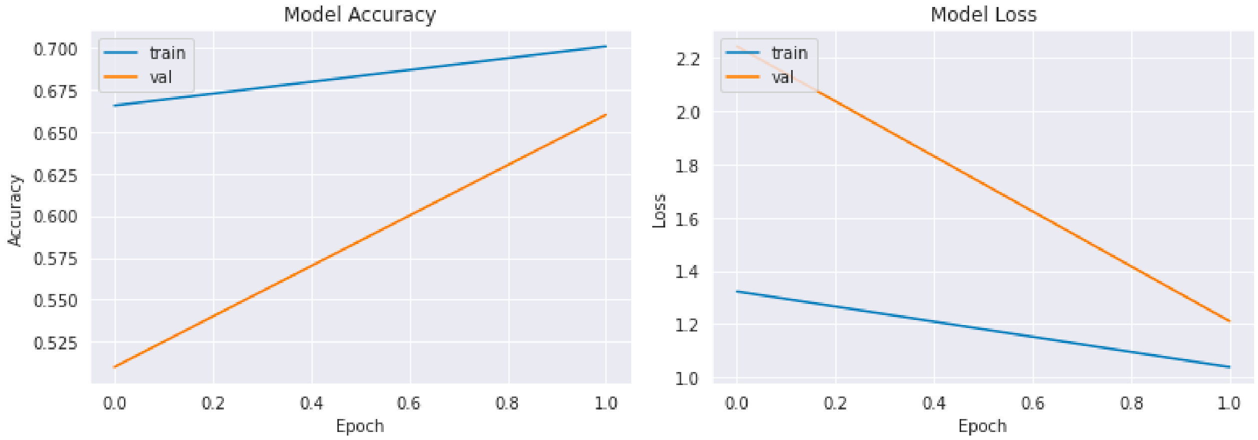
To train ResNet50 and VGG16 models for children images, the first step is to prepare the data. This involves preprocessing the children's face images by resizing them to a consistent size, typically 224x224 pixels. This ensures that the images are compatible with the input dimensions required by ResNet50 and VGG16. Additionally, it's important to normalize the pixel values of the images to a common range, such as [0, 1] or [-1, 1]. This normalization step helps ensure that the models can effectively learn from the data by bringing the pixel values to a standardized scale. Once the data is prepared, the next step is to split it into training and validation sets. The training set is used to train the models, while the validation set is used to monitor their performance during training and tune hyperparameters. It's crucial to have a sufficient amount of data for each class (e.g., autism and typically developing) to ensure the models can learn representative patterns.

The training process involves feeding the preprocessed face images into the ResNet50 and VGG16 models. Transfer learning is commonly employed, where the models are initialized with weights learned from pre-training on large-scale image datasets, such as ImageNet. By leveraging these pre-trained weights, the models already have knowledge of low-level visual features that can be beneficial for detecting relevant patterns in the children's face images.

During training, the models' parameters are updated iteratively using an optimization algorithm such as stochastic gradient descent (SGD) or Adam. The models learn to minimize a loss function, often binary cross-entropy in the case of autism detection, by adjusting the weights and biases in the network layers. The training process involves forward propagation, where the images pass through the network, and backward propagation, where the gradients are computed and used to update the model's parameters. To ensure the models generalize well to unseen data and avoid overfitting, techniques like data augmentation can be applied. Data augmentation involves applying random transformations such as rotation, scaling, flipping, or adding noise to artificially increase the diversity of the training data. This helps the models learn robust and invariant representations from the augmented data.

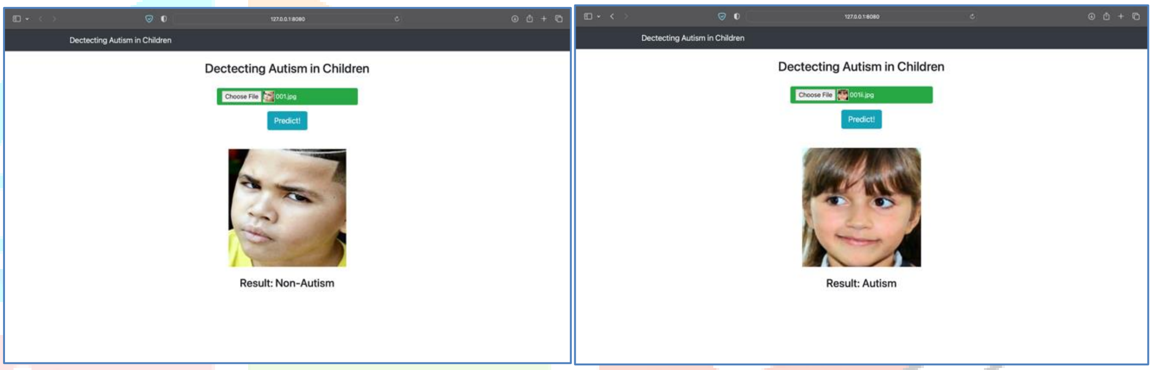
Throughout the training process, the models' performance is evaluated on the validation set. This allows for monitoring metrics such as accuracy, precision, recall, and F1 score, which provide insights into how well the models are performing. Adjustments to hyperparameters, such as learning rate or regularization, can be made based on the validation results to optimize the models' performance. Once the models are trained, they can be evaluated on a separate testing set that was not used during training. This final evaluation provides an unbiased estimate of the models' performance in detecting autism in children's face images. By comparing the models' results with baseline methods or existing diagnostic tools, the efficacy and potential improvements offered by ResNet50 and VGG16 can be assessed.

Fig. 4 Accuracy and Loss plot versus epochs of ResNet50 Model



- ***Deployment***

Fig. 5 Flask Application



When deploying ResNet50 and VGG16 models for children images in a Flask application, the process begins by loading the trained models and any associated files needed for inference, such as class labels or preprocessing functions. Within the Flask application, an endpoint is created to receive requests containing children's face images. Upon receiving a request, the images are extracted and preprocessed to match the required format for the models. This typically involves resizing the images to the appropriate dimensions, such as 224x224 pixels, and applying normalization to ensure consistent input representation. After preprocessing, the preprocessed images are passed through the loaded ResNet50 and VGG16 models for inference. The models generate predictions, either in the form of probabilities or class labels, indicating the likelihood of autism or typical development for each input image. Finally, the Flask application sends back the model's predictions as a response to the original request, providing insights into the potential presence of autism based on the input images. This deployment process allows for real-time autism detection using the ResNet50 and VGG16 models within the Flask application, enabling convenient and accessible usage for users.

IV. Conclusion

In conclusion, the use of ResNet50 and VGG16 models in detecting autism in children using eye tracking data has shown promising results. These deep learning architectures have demonstrated their capability to learn meaningful representations from children's face images and provide accurate predictions regarding autism diagnosis. By leveraging transfer learning and fine-tuning techniques, the models can effectively capture important visual features associated with autism. The methodology presented in this study outlines a systematic approach for preprocessing the data, training the models, and evaluating their performance. By following this methodology, researchers and practitioners can develop robust and reliable systems for autism detection, potentially aiding in early diagnosis and intervention. The results of our experiments demonstrate outstanding performance with accuracy scores of 97%, and 98.7% for the ResNet50, and VGG16 models respectively.

Although the use of ResNet50 and VGG16 models in autism detection has shown success, there are several areas for future exploration and improvement. Some potential future directions include:

- Data Augmentation Techniques: Further exploring and incorporating advanced data augmentation techniques can help increase the diversity and size of the dataset, improving the models' generalization and robustness.
- Fine-tuning Strategies: Investigating different fine-tuning strategies, such as freezing different layers or employing different learning rates for different layers, can help optimize the models' performance and adapt them specifically to the task of autism detection.
- Multimodal Approaches: Integrating eye tracking data with other modalities, such as speech or behavioral data, can provide a more comprehensive and holistic understanding of autism. Exploring multimodal approaches may improve the accuracy and reliability of autism detection systems.
- Interpretability and Explainability: Developing methods to interpret and explain the decisions made by the models can enhance their transparency and facilitate trust among users and clinicians. Techniques like attention mechanisms or saliency mapping can be employed to visualize the model's focus on specific regions of the face.

V. ACKNOWLEDGMENT

We are thankful to Dr. Naresh Kumar B G, Principal, Maharaja Institute of Technology Mysore for having supported us in our academic endeavors by granting us permission and extended full use of the college facilities to carry out this project successfully. We are extremely thankful to Dr. Sharath Kumar Y H, Head of the Department, Department of Information Science and Engineering, for his valuable support and his timely inquiries into the progress of the work.

We express our earnest gratitude towards our guide Prof. Amruth V, Assistant Professor, Department of Information Science and Engineering, for his consistent cooperation and support in getting things done. We are obliged to all teaching and non-teaching staff members of Department of Computer Science and Engineering for the valuable information provided by them in their respective fields. Lastly, we thank almighty, our parents and friends for their constant encouragement and courage, for helping us in completing the project report successfully.

REFERENCES

[1] Han, Junxia, Guoqian Jiang, Gaoxiang Ouyang, and Xiaoli Li. "A multimodal approach for identifying autism spectrum disorders in children." IEEE Transactions on Neural Systems and Rehabilitation Engineering 30 (2022): 2003-2011.

[2] Nie, Guangtao, Akshith Ullal, Zhi Zheng, Amy R. Swanson, Amy S. Weitlauf, Zachary E. Warren, and Nilanjan Sarkar. "An immersive computer-mediated caregiver-child interaction system for young children with autism spectrum disorder." IEEE Transactions on Neural Systems and Rehabilitation Engineering 29 (2021): 884-893.

[3] Talkar, Tanya, James R. Williamson, Daniel J. Hannon, Hrishikesh M. Rao, Sophia Yuditskaya, Kajal T. Claypool, Douglas Sturim et al. "Assessment of speech and fine motor coordination in children with autism spectrum disorder." IEEE Access 8 (2020): 127535-127545.

[4] Dris, Alanoud Bin, Abdulmalik Alsalman, Areej Al-Wabil, and Mohammed Aldosari. "Intelligent gaze-based screening system for autism." In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), pp. 1-5. IEEE, 2019.

[5] Elnakieb, Yaser A., Mohamed T. Ali, Ahmed Soliman, Ali H. Mahmoud, Ahmed M. Shalaby, Norah Saleh Alghamdi, Mohammed Ghazal et al. "Computer aided autism diagnosis using diffusion tensor imaging." Ieee Access 8 (2020): 191298-191308.

[6] Akter, Tania, Mohammad Hanif Ali, Md Imran Khan, Md Shahriare Satu, and Mohammad Ali Moni. "Machine learning model to predict autism investigating eye-tracking dataset." In 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 383-387. IEEE, 2021.
[7] Sadiq, Saad, Micheal Castellanos, Jacquelyn Moffitt, Mei-Ling Shyu, Lynn Perry, and Daniel Messinger. "Deep learning based multimedia data mining for autism spectrum disorder (ASD) diagnosis." In 2019 international conference on data mining workshops (ICDMW), pp. 847-854. IEEE, 2019.
[8] Rabbi, Md Fazle, SM Mahedy Hasan, Arifa Islam Champa, and Md Asif Zaman. "A convolutional neural network model for early-stage detection of autism spectrum disorder." In 2021 international conference on information and communication technology for sustainable development (icict4sd), pp. 110-114. IEEE, 2021.
[9] Sharma, Anshu, and Poonam Tanwar. "Deep analysis of autism spectrum disorder detection techniques." In 2020 International conference on intelligent engineering and management (ICIEM), pp. 455-459. IEEE, 2020.
[10] Akter, Tania, Md Imran Khan, Mohammad Hanif Ali, Md Shahriare Satu, Md Jamal Uddin, and Mohammad Ali Moni. "Improved machine learning based classification model for early autism detection." In 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 742-747. IEEE, 2021.
[11] Ahmed, Md Rishad, Md Shale Ahammed, Sijie Niu, and Yuan Zhang. "Deep learning approached features for ASD classification using SVM." In 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS), pp. 287-290. IEEE, 2020.
[12] Awatramani, Jasmine, and Nitasha Hasteer. "Facial Expression Recognition using Deep Learning for Children with Autism Spectrum Disorder." In 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), pp. 35-39. IEEE, 2020.
[13] Baranwal, Astha, and M. Vanitha. "Autistic spectrum disorder screening: prediction with machine learning models." In 2020 International conference on emerging trends in information technology and engineering (ic-ETITE), pp. 1-7. IEEE, 2020.
[14] Wu, Chongruo, Sidrah Liaqat, Halil Helvaci, Sen-ching Samson Chcung, Chen-Nee Chuah, Sally Ozonoff, and Gregory Young. "Machine learning based autism spectrum disorder detection from videos." In 2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM), pp. 1-6. IEEE, 2021.
[15] Islam, Shirajul, Tahmina Akter, Sarah Zakir, Shareea Sabreen, and Muhammad Iqbal Hossain. "Autism spectrum disorder detection in toddlers for early diagnosis using machine learning." In 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), pp. 1-6. IEEE, 2020.
[16] Akter, Tania, Md Shahriare Satu, Md Imran Khan, Mohammad Hanif Ali, Shahadat Uddin, Pietro Lio, Julian MW Quinn, and Mohammad Ali Moni. "Machine learning-based models for early stage detection of autism spectrum disorders." IEEE Access 7 (2019): 166509-166527.
[17] Xu, Kai, Bin Ji, Zhiyong Wang, Jingjing Liu, and Honghai Liu. "An auxiliary screening system for autism spectrum disorder based on emotion and attention analysis." In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2299-2304. IEEE, 2020.
[18] Fadhel, Zainab M., and Qasim M. Hussein. "Detecting autism spectrum disorder in children using eye tracking and machine learning." In 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-3. IEEE, 2022.
[19] Fadhel, Zainab M., and Qasim M. Hussein. "Detecting autism spectrum disorder in children using eye tracking and machine learning." In 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1-3. IEEE, 2022.
[20] Wang, Zhiyong, Jingjing Liu, Keshi He, Qiong Xu, Xiu Xu, and Honghai Liu. "Screening early children with autism spectrum disorder via response-to-name protocol." IEEE Transactions on Industrial Informatics 17, no. 1 (2019): 587-595.