GMR

The Original

GRAPH CONVOLUTION MODEL FOR DIAGNOSING AUTISM SPECTRUM DISORDER USING DIVERSE GENETIC DATA

Suresh Jaganathan[1], D Venkata vara Prasad[2], Kiruthigha M[3]

[1,2]Department of Computer Science and Engineering, Sri Sivasubramaniya Nadar College of Engineering, Chennai, India.
[3]Department of Information Technology, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India.
Emails: sureshj@ssn.edu.in[1], dvvprasad@ssn.edu.in[2], kirthi.r@gmail.com[3]

ABSTRACT
Autism Spectrum Disorder (ASD) affects multiple domains of development, including social interaction, repetitive behavior, restricted interests, sensory processing, motor function, cognition, and emotional regulation. Early intervention can improve outcomes by addressing these challenges and supporting overall development. In this work, we developed a graph convolution model to distinguish children with ASD from non-autistic children using behavioral video recordings during psychiatrist-led intervention sessions and scores from the ISAA questionnaire completed by parents or caretakers. Behavioral videos were processed using a multi-channel 3D CNN to extract node features, while ISAA scores were used to compute similarity-based edges in the graph. We further integrated this behavioral graph with a previously constructed neuroimaging graph to obtain a more comprehensive representation of ASD. The proposed model achieved an accuracy of 86.34% using behavioral data alone and 87.56% after integrating graphs.

KEYWORDS: Autism Spectrum Disorder, Graph Convolutional Networks, Deep Learning, ISAA

1. INTRODUCTION

Behavioral video analysis during psychiatrist-child interaction sessions can provide valuable insight into social communication, repetitive behavior, and sensory responses in children with ASD. Observations of eye contact, gestures, and overall social engagement help identify patterns of communication difficulty. At the same time, verbal and non-verbal responses can reveal differences in language use, interaction style, and responsiveness. Video also enables the analysis of repetitive movements such as hand-flapping or rocking, which can provide evidence of sensory processing differences [1]. In addition, it supports the assessment of play behavior, adaptability to routine changes, facial expression, emotional response, and joint attention. Because recorded sessions can be reviewed across time, video offers a more comprehensive and objective basis for behavioral assessment than live observation alone [2].

The ISAA is a questionnaire (Figure 1) used to assess the severity of autistic symptoms in individuals. It supports decisions about the type and level of intervention required and also contributes to disability certification, which can help individuals with ASD access benefits and support services. Although the ISAA is not intended to be used as a stand-alone diagnostic tool, it can provide valuable supporting evidence when interpreted alongside clinical assessments and direct observations by qualified professionals.

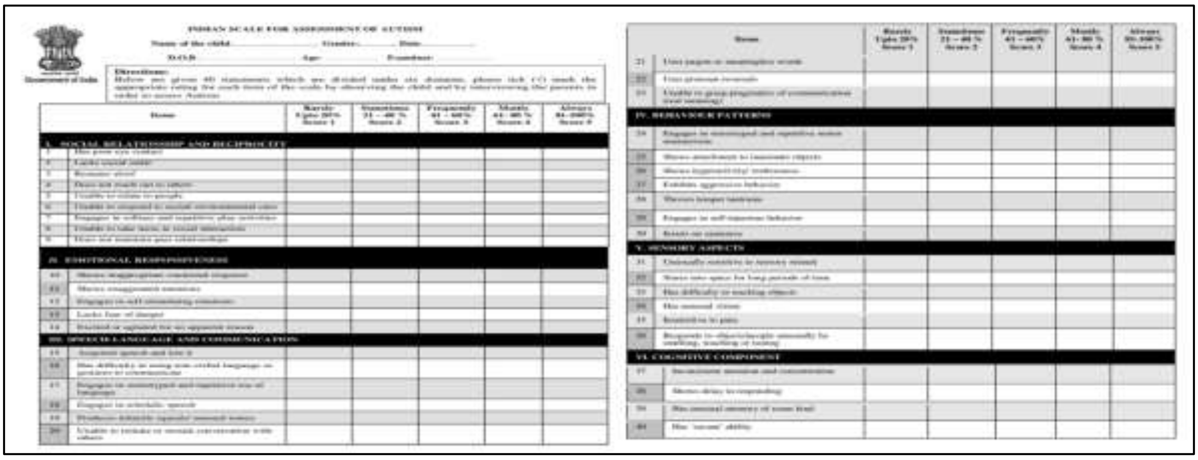

Figure 1: ISAA Questionnaire

Despite its value, the ISAA has limitations. Prior studies suggest that it may be less effective in younger children and may show limited sensitivity and specificity. It should therefore not be used as a stand-alone diagnostic instrument [3]. Accurate diagnosis requires a comprehensive assessment by a qualified healthcare professional using standardized diagnostic criteria. Overall, the ISAA contributes to:

- Assessing the severity of ASD in Indian individuals
- Facilitating disability certification and access to support services
- Supporting the diagnostic process, albeit not as a sole diagnostic tool

1.1 Graph Convolution Network (GCN)

Graph Convolution Networks (GCNs) [4] are a class of neural network architecture designed to operate on graph-structured data. They are beneficial for tasks that involve analyzing relationships or interactions between elements in a graph, such as social networks, citation networks, or biological networks. Graph Convolution Networks can effectively capture and learn from the relational information in graph-structured data by combining local graph structure, feature aggregation, and non-linear transformations. The core operation in a GCN is the graph convolutional layer. Each layer takes the node features and their corresponding adjacency matrix as input and applies a convolutional operation to aggregate information from neighboring nodes. The key idea is to combine a node's features with those of its neighboring nodes to capture the local graph structure. The core concept of GCNs is message passing, where information is exchanged between connected nodes in the graph [5].

Autism assessment typically combines input from a pediatrician, child psychologist, neurologist, and therapist. Because ASD diagnosis requires a comprehensive review of development, behavior, and functional ability, each discipline contributes a distinct perspective to the evaluation process. Multiple sources, including parental reports, direct observation, and formal assessment results, inform diagnostic decisions. Accordingly, our proposed model integrates neuroimaging data, behavioral videos, and observations from parents or caregivers to improve classification performance.

The rest of the paper is organized as follows: Section 2 discusses work on GCNs for processing behavioral videos and on deep learning methods for classifying ASD from TC. Section 3 elaborates on the proposed method; Section 4 explores the experimental results and details the research discussion; and Section 5 concludes the paper with possible future work.

2. LITERATURE SURVEY

Research on ASD increasingly uses deep learning and behavioral video analysis to identify clinically relevant patterns from recorded interactions. Existing work generally falls into three categories: eye-gaze analysis, pose estimation, and motion or gesture analysis. Together, these approaches show that video can support objective behavioral assessment, but they also reveal limitations in generalizability, context sensitivity, and multimodal integration.

2.1 Eye Gaze Pattern

Eye-gaze pattern analysis [6,7] examines how individuals with ASD allocate visual attention differently from typically developing children. These studies use eye-tracking or vision-based systems to measure gaze location, fixation duration, and saccadic behavior during structured tasks or visual observation.

Kong Xue-Jun [8] highlighted the importance of collecting diverse parameters, including gaze location, fixation duration, saccadic eye movements, and other relevant factors. In research conducted by Edmonds [2], children in the age range of 3 to 6 who were diagnosed with autism spectrum disorder (ASD) were observed alongside a control group comprising typically developing children. This observation was facilitated using head-mounted cameras. The cameras recorded the children's visual perspective during social interactions, play, and gestures. However, the study's small sample size raises concerns about the comprehensiveness of the results. Point-of-view cameras have limitations, including a restricted field of view and increased resistance, which should be acknowledged when using similar devices.

Another study by Angelina [9] employed eye-tracking technology to analyze social cues in infants with and without ASD. Children faced an eye-shadowing device displaying facial expressions and social stimuli based on their gaze and facial gestures. Eye-tracking systems monitored eye movements in real time, enabling adjustments to the visual stimuli presented. Research suggested that infants without ASD tended to be more attentive and had longer gaze duration toward social stimuli and delighted faces compared to infants with ASD. The revealed conclusions implied that infants with ASD may have reduced sensitivity to social cues. Mariano [10] created a virtual environment to address limitations in earlier studies. This digital setting incorporates various social signals, such as facial expressions and movements, to stimulate active engagement in social interactions. The VR environment captures and analyzes participants' eye movements, gaze duration, and focal points, offering a more sophisticated approach to studying social behaviors.

2.2 Pose Estimation

Pose-estimation methods derive 2D skeletal frames from video and model joints as graph nodes connected by spatial relationships [11]. Spatio-temporal GCNs such as ST-GCN [12] learn both spatial structure and temporal dynamics from these sequences, making them effective for action recognition. However, their performance depends on the assumed graph topology, which may not transfer reliably across behaviors, body configurations, or recording conditions.

2.3 Motion Analysis through Video

Video-based motion and gesture analysis for ASD diagnosis uses computer vision and machine learning to detect behavioral signals such as hand movements, posture, facial expression, and repetitive actions [13,14]. These methods aim to extract discriminative visual patterns associated with ASD from observable behavior during recorded interactions. Farhood [1] proposed a video-assisted system for early detection of ASD based on abnormal and repetitive behaviors. The method uses computer vision to estimate body position and movement and to analyze spatial relationships among joints, enabling the extraction of motion, posture, facial, and behavioural cues from image sequences.

Xuemei [15] analyzed motion patterns in videos collected during Autism Diagnostic Observation Schedule (ADOS) assessments. The study extracted features such as motion force, speed, acceleration, and spatial distribution, showing that structured video analysis can capture movement characteristics relevant to ASD-related behavior.

Video provides an objective and reviewable record of child behavior, while parental observations add context about everyday functioning outside the clinical setting [16]. Prior work typically focuses on a single modality, such as gaze, pose, or gesture. In contrast, our approach combines behavioral videos from psychiatrist-child interaction sessions with ISAA questionnaire scores completed by parents or caretakers, enabling a more contextualized graph-based representation for ASD classification. Table 1 summarizes the works related to autism spectrum disorder.

Table 1: Summary of Related Works

Authors, Paper Title (Year)	Contribution	Drawback
Brief report: using a point-of-view camera to measure eye gaze in young children with autism spectrum disorder during naturalistic social interactions: a pilot study ([2]).	In the recorded scenes, specific Areas of Interest (AOI) were delineated, encompassing the subject’s face and other relevant objects. Software designed to monitor eye movements analyzes gaze patterns, enabling measurement of fixations, saccades, and gaze shifts. Nevertheless, it’s crucial to remember that employing a wearable camera comes with inherent constraints, including a restricted field of vision and the potential for encountering obstacles during the recording process.	There are certain inherent drawbacks in using a wearable camera, most notably a restricted field of view and the potential to encounter obstacles.
Simulating interaction: Using gaze-contingent eye-tracking to measure the reward value of social signals in toddlers with and without autism ([9])	Metrics such as gaze duration on social stimuli, the percentage of fixations on social stimuli, and overall attention patterns were used to compare the ASD and control groups. A statistical analysis was carried out to identify significant differences and correlations among these variables.	Ensuring a comprehensive assessment of the reward value of social cues can be challenging, and relying solely on gaze and exploratory activity may not precisely capture the social reward processing of young infants
Eye gaze as a biomarker in recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis. ([10])	In the application of virtual reality and machine learning, gaze serves as a biomarker for detecting ASD. Children’s gaze patterns may vary due to factors such as nervousness or discomfort when using the device, as well as heightened engagement with virtual reality.	Children’s gaze patterns may change due to their anxiety or discomfort when using the device, as well as their enjoyment of virtual reality.
A 3D graph convolutional networks model for 2D skeleton-based human action recognition. ([11])	The skeleton sequence’s temporal development is illustrated by charting each frame as a node function. This involves making assumptions about connections and spatial relationships within the frame data. These presumptions might not hold in particular	When plotting frame data, assumptions about connections and spatial relationships are inherent. However, in certain situations, these assumptions may not hold, leading to suboptimal or inaccurate graph presentations.

	circumstances, though, which could result in erroneous or inadequate graph displays.	
Vision-assisted recognition of stereotype behaviors for early diagnosis of autism spectrum disorders ([1]).	The pre-processed ADOS videos are used to extract motion features using computer vision techniques. These features capture the movement patterns of children in correlation with their activity levels.	Children with ASD often have concurrent conditions, such as intellectual disabilities or motor impairments. These additional conditions can significantly affect their movement patterns.

3. MATERIAL AND METHODS

The population graph is constructed using behavioral data from the ISAA questionnaire and videos recorded during interactive sessions with a qualified expert. Figure 2 illustrates the overall workflow of the proposed method, which builds a unified graph from two inputs: behavioral video data and ISAA questionnaire scores.

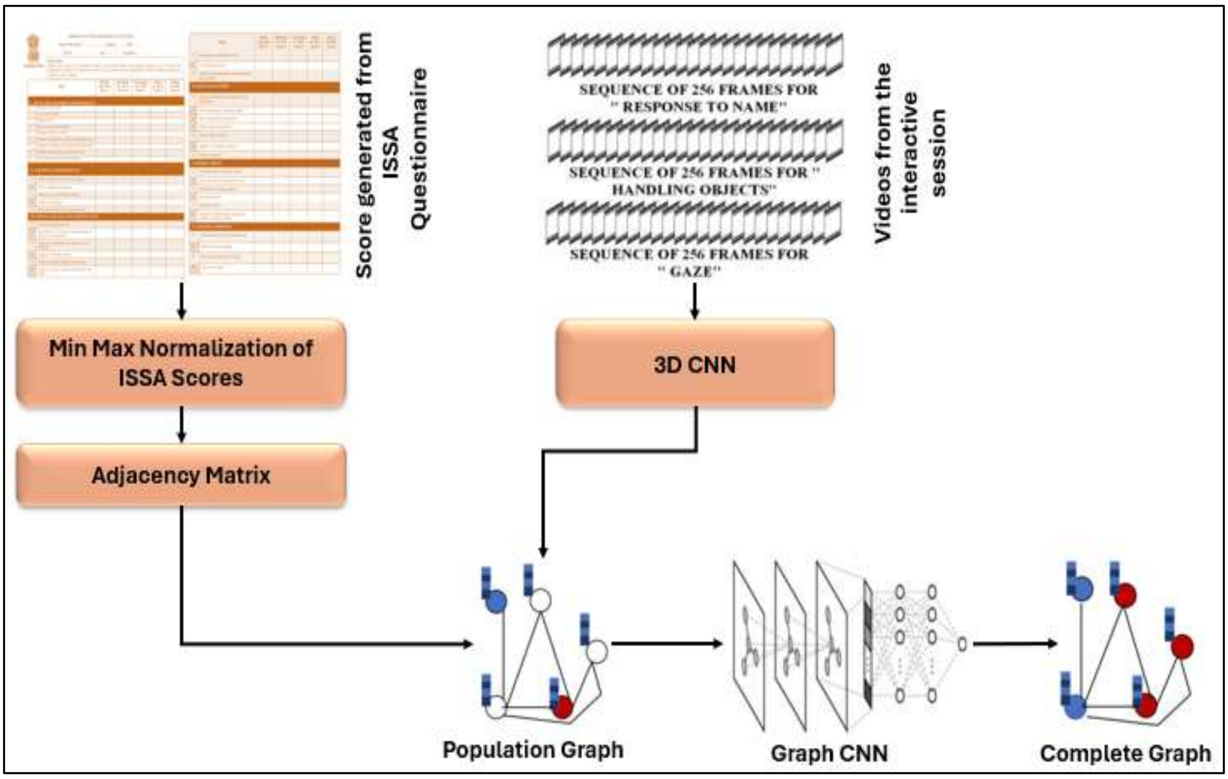

Figure 2: Creation of a Graph based on Behavioral Video

3.1 ISAA Scoring

Behavioral data collected during psychiatrist-led interactive sessions, together with ISAA questionnaire scores completed by parents or caretakers, are used to build the population graph. The ISAA questionnaire contains 40 items that assess the severity of autistic symptoms. These scores are normalized using min-max normalization to generate similarity scores between children.

INCLEN is a semi-structured tool used in ASD assessment. Children suspected of having autism are evaluated by trained professionals using questionnaires and play-based observation. Video footage is recorded during interactive sessions, with cameras positioned to capture relevant facial expressions and behavioral responses. Data were collected for 153 children, of whom 87 were diagnosed with autism, and 66 were not. Based on communication and social interaction criteria, the recordings were divided into sub-videos reflecting three key benchmarks: response to name, eye contact, and change in routine during play [17,18]. Qualified psychologists manually reviewed and labeled the videos, and children diagnosed with autism were reassessed two or three times over one to two months to confirm the diagnosis.

The resulting sub-videos are segmented into 256 frames, each resized to 112x112. These consecutive frames form a 3D volume. For each child, videos are organized into three channels representing response to name, object handling, and eye contact. Two experienced professionals independently scored the videos, and any disagreements were re-evaluated to ensure an appropriate final classification.

3.2 GCN Model for Behavioral Data

ISAA scores are normalized to compute similarity between children [19]. Video clips recorded during participant interactions are pre-processed and passed through a multi-channel CNN to extract subject-level features. These features define the graph nodes, while edges are established from similarity scores derived from the ISAA data, resulting in a population graph for classification.

3.2.1 Construction of Adjacency Matrix

According to [20], an ISAA score below 70 indicates typical development, scores from 70 to 106 indicate mild autism, and scores above 106 indicate more severe autism. To place all questionnaire values on a common scale, we applied min-max normalization, as shown in the equation below.

$$Normalized = \frac{value - min}{maxvalue - minvalue}$$

Because the ISAA contains 40 items, the observed score range extends from 40 to 153. These normalized scores were then used to construct the adjacency matrix. An edge is created between two children when their normalized scores indicate sufficient similarity.

3.2.2 Feature Selection

The 3D CNN takes sequences of video frames as input. In this study, each sequence contains 256 frames, and each frame includes three channels representing response to name, object handling, and gaze. If a video contains L frames, the 3D CNN performs both temporal and spatial down-sampling, following the approach of [21]. The architecture of the proposed 3D CNN is shown in Figure 3. To construct the population graph, 1024 features extracted by the 3D CNN are assigned to each node, while graph edges are defined using the adjacency matrix derived from ISAA scores.

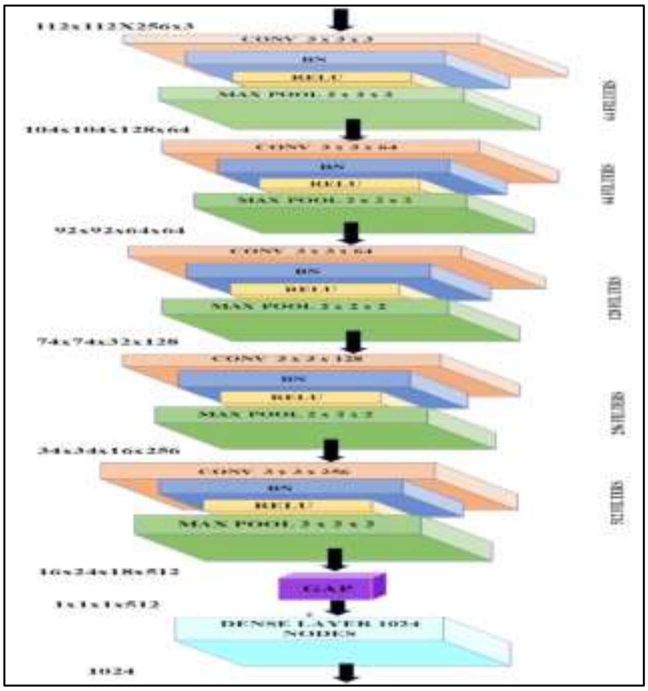

Figure 3: 3D CNN Architecture

The 3D CNN performs joint spatial and temporal down-sampling across the sequence of video frames [22]. Each node in the population graph is represented by 1024 features extracted from the 3D CNN, and graph edges are defined by the adjacency matrix derived from ISAA scores.

3.3 Graph Union

In graph theory, a graph union combines two graphs into a single graph (Figure 4). This operation consolidates information from multiple sources and can provide a richer representation of the relationships within the data. In our work, the population graph derived from neuroimaging data is merged with the graph derived from behavioral data to form a unified network with more connections. The resulting graph supports improved message passing and more accurate classification of unlabeled nodes (Figure 5).

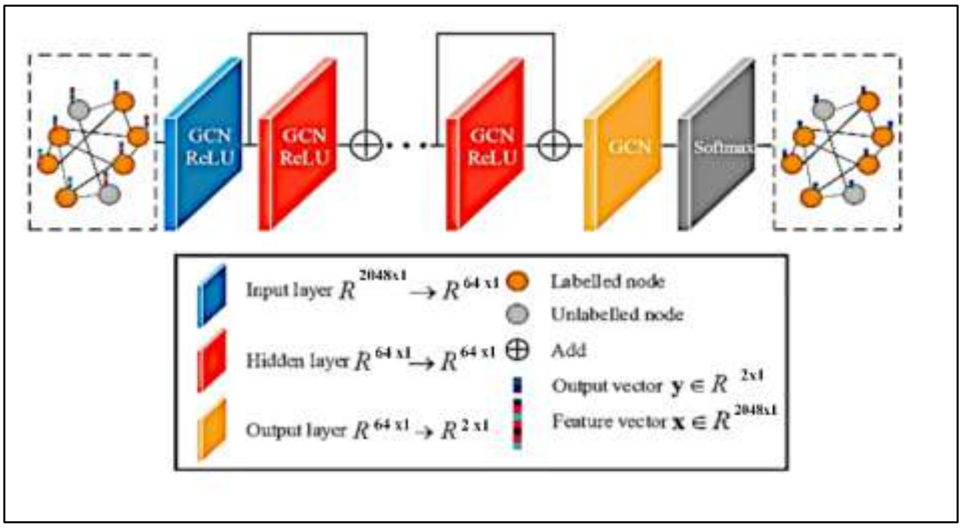

Figure 4: Graph Union

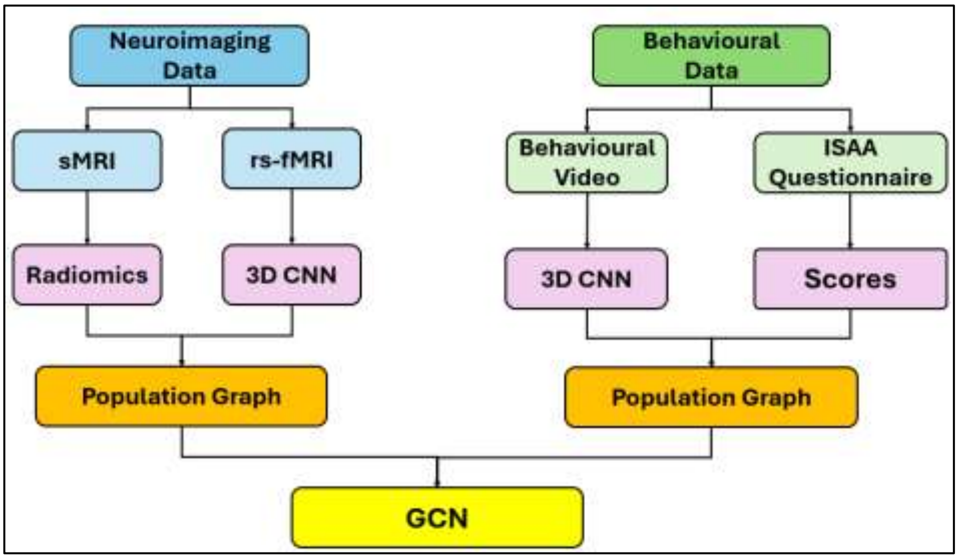

Figure 5: Combined Neuroimaging and Behavioral Data

3.4 Graph constructed from Neuroimaging Data

For the neuro-imaging graph, regions of interest extracted from sMRI brain images were processed using radiomics to derive quantitative structural features. Feature dimensionality was reduced using stacked autoencoders, and improved square-root cosine similarity was used to determine similarity between nodes. In rs-fMRI analysis, brain summary methods were used to reduce the temporal dimension of the 4D data by deriving measures that reflect functional connectivity among brain regions. These summaries produced 19 3D image volumes [23,24]. A multi-channel 3D CNN was then used to extract subject-level features, with each subject represented as a node and inter-node edges defined by the similarity metric.

The final graph contains all the vertices and edges from the two input graphs. Let V_1 and V_2 denote the vertex sets and E_1 and E_2 denote the edge sets of the two graphs. Algorithm 1 summarizes the graph-union procedure for combining. G_1 and G_2.

Algorithm 1: Graph Union

Input : $G1 \leftarrow (V1, E1)$ & $G2 \leftarrow (V2, E2)$
Output: Graph Union G
begin
Generate an empty graph
for v in $G1$ do
Get G.addVertex(v)
end
Copy vertices from $G2$ to G, avoiding duplicates
for v in $G2$ do
  if v not in G.vertices then
    G.addVertex(v)

end
end
Copy edges from $G1$ to G
for e in $G1.edges$ do
G.addEdge(e)
end
Copy edges from $G2$ to G, avoiding duplicates
for e in $G2.edges$ do
if e not in $G.edges$ then
G.addEdge(e)
end
end
return G
end

The combined population graph represents each node with 2048 features, obtained by concatenating 1024 features from behavioral videos and 1024 features from rs-fMRI images. Edges are constructed using concordance between ISAA scores and sMRI radiomic properties. The GCN model uses two hidden layers with 64 units each and dropout values of 0.005 and 0.4. The output layer contains two nodes and applies Softmax to classify unlabelled nodes as either normal or ASD [25,26].

4. Experimental Results and Discussion

All experiments were conducted on a system with an Intel i9-12900 CPU, 16 GB RAM, an NVIDIA RTX A2000 GPU, and Ubuntu OS. The hyperparameters for constructing a 3D CNN are the Number of convolution blocks after the first convolution block, the Number of Filters, the Dropout rate, and the number of nodes in the dense layer. We made an extensive analysis to identify the appropriate parameters for both neuroimaging and behavioral data. The values used for analysis are provided in Table 2. Using the previously described 3D CNN configuration layers=4, filters=64, dropout rate=0.5, and dense nodes=1024, the model achieved an accuracy of 69.45%. Figure 5 shows the parallel coordinates plot for the considered hyperparameters.

Table 2: Hyperparameter Search

Hyper Parameters	Values
No. Convolutional blocks after the First layer	2,3,**4**
Number of Filters	32,**64**,128
Dropout Rate	0.4,**0.5**,0.6
No. nodes in the dense layers	128,256,512,**1024**

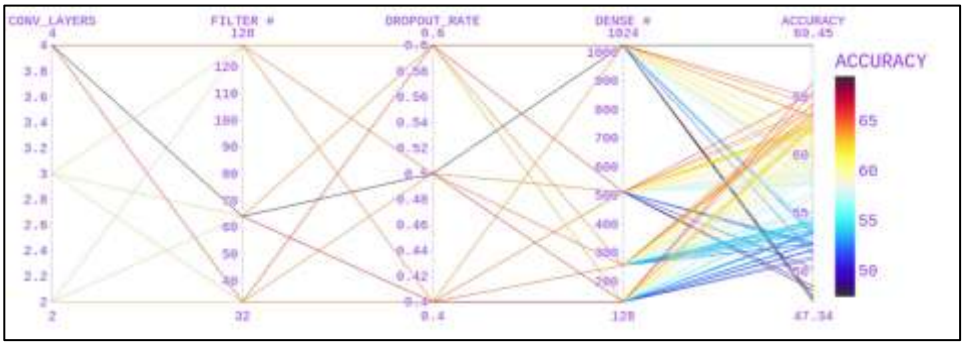

Figure 6: Hyperparameter search for 3D CNN

4.1 Performance Metrics

Performance metrics are quantitative measures used to evaluate a model's effectiveness. In classification tasks, they indicate how well the model distinguishes between classes across several dimensions. The metrics used in this study are:

- Accuracy: The proportion of correctly classified instances out of the total instances.

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- Precision: The proportion of true positive predictions out of all positive predictions. It measures the model's ability to avoid false positives.

$$precision = \frac{TP}{(TP + FP)}$$

- Sensitivity: The proportion of true positive predictions out of all positive instances. It measures the model's ability to identify all positive instances.

$$sensitivity = \frac{TP}{(TP + FN)}$$

- Specificity measures the ability of a model to correctly identify the negative instances out of all actual negative instances. Specificity is complementary to sensitivity.

$$specificity = \frac{TN}{(TN + FP)}$$

4.2 Graph Convolution Network for Behavioral Data

Deep learning approaches for ASD diagnosis from video commonly use inputs such as eye tracking, facial analysis, or kinematic features for action classification. However, many existing classification methods do not incorporate parent-reported observations, even though parents often provide the most consistent view of a child's developmental milestones and day-to-day behavior.

Table 3: Methods for Analyzing Behavioral Data

S. No	Input	Methodology	Findings
1	Eye Tracking	The temporal component of the gaze pattern is found using the length of the eye-tracking map.	The use of wearable devices and secular surroundings may increase the child's anxiety, causing variations in the eye tracking map.
2	Pose Estimation	It categorizes videos based on 2D skeleton sequences and graph convolution networks (GCN).	Different actions require different topology graphs.
3	Video Gesture Analysis	An attention map identifies the zones in a video of significant interest in space and time.	Identifying kinematic discriminants in gestures for action recognition may cause an overload in analysis.

Table 4: GCN for Behavioral Data

S. No	Model	Accuracy	Sensitivity	Specificity
1	Eye Tracking	82.8	79.3	86.2
2	Pose Estimation	74	89	87
3	Video Gesture Analysis	82	87	91
4	Proposed Method	86.34	91	90.4

Table 3 highlights the methodological trade-offs of prior behavioral approaches, while Table 4 shows that the proposed method achieves the strongest overall balance across accuracy, sensitivity, and specificity. This improvement likely stems from combining subject-level video features with parent-reported ISAA information in a graph-based framework.

4.3 GCN after Graph Union

Graph union improves connectivity and information flow in the merged graph by combining multiple graph structures into a single representation. Because GCNs rely on neighborhood aggregation, merging graphs expands the available neighborhood information, enabling richer message passing and improving the model's ability to capture broader relational patterns.

The neuro-imaging graph contained 7,947 edges, and the behavioral graph contained 8,122 edges. After graph union, the merged graph contained 8,179 edges. Although the number of nodes remained unchanged, the denser connectivity improved neighborhood information for message passing, which in turn improved the classification of unlabelled nodes, as shown in Figure 7.

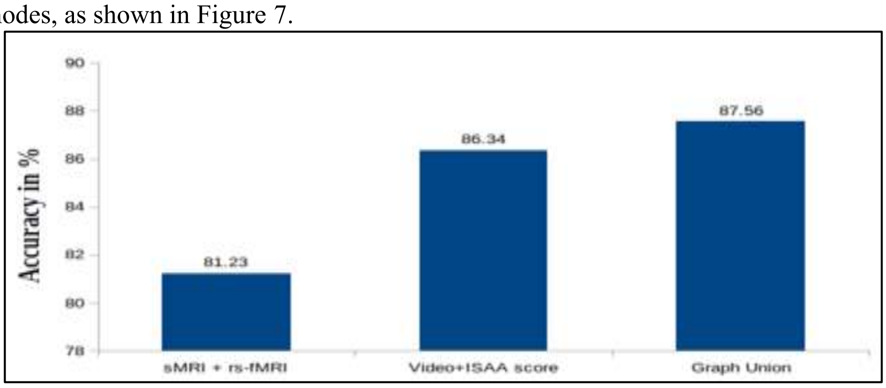

Figure 7: Performance of the GCN Model Using Neuro-imaging Data, Behavioral Data, and the Combined Model

Neuro-imaging data capture high-dimensional brain-related features, whereas behavioral data provide complementary categorical and continuous indicators of observed function. Integrating these modalities allows the GCN to learn more informative representations of the relationship between brain measures and behavior than either modality alone. This multimodal view improves classification performance and supports broader applications, such as brain-based diagnosis, prediction of therapy response, and understanding of ASD-related cognitive patterns.

CONCLUSION

GCNs show strong potential for improving the diagnosis and understanding of ASD. By integrating behavioral video analysis, ISAA questionnaire scores, and neuro-imaging data, the proposed approach captures relationships among behavior, symptoms, and brain-based features more effectively than unimodal methods. The model achieved strong classification performance and demonstrated the value of multimodal graph-based learning for ASD assessment. Future work can further validate this framework on larger datasets and explore its potential for predicting therapy response.

REFERENCES

[1] Farhood Negin, Baris Ozyer, Saeid Agahian, Sibel Kacdioglu, Gulsah Tumuklu Ozyer (2021). Vision-assisted recognition of stereotype behaviors for early diagnosis of autism spectrum disorders. *Neurocomputing*. Volume 446, Pages 145-155, https://doi.org/10.1016/j.neucom.2021.03.004.

[2] Edmunds, S.R., Rozga, A., Li, Y. et al (2017). Brief Report: Using a Point-of-View Camera to Measure Eye Gaze in Young Children with Autism Spectrum Disorder During Naturalistic Social Interactions: A Pilot Study. *Journal of Autism Developmental Disorders*. 47, 898–904. https://doi.org/10.1007/s10803-016-3002-3.

[3] Kanimozhi, S. et al. (2020). Autism spectrum disorder: The dilemma of untimely recognition, intervention, and diagnostic scales available in the Indian subcontinent. *Journal of Psychological Research*. 2(1): 1–11.

[4] Zhang, S., Tong, H., Xu, J. et al. (2019). Graph convolutional networks: a comprehensive review. *Computational Social Networks*. 6(11).

[5] Yao Ma, Suhang Wang, Charu C. Aggarwal, Dawei Yin, Jiliang Tang (2018). Multi-dimensional graph convolutional networks. *arXiv:1808.06099*.

[6] Modi, N., Singh, J. (2021). A Review of Various State-of-the-Art Eye Gaze Estimation Techniques. *In Proceedings of Advances in Computational Intelligence and Communication Technology. Advances in Intelligent Systems and Computing*, vol 1086. Springer, Singapore. https://doi.org/10.1007/978-981-15-1275-9_41.

[7] Jeyarani, A. and Senthilkumar, R (2023). Eye tracking biomarkers for autism spectrum disorder detection using machine learning and deep learning techniques: Review. *Research in Autism Spectrum Disorders.* 108: 102–228.

[8] Kong Xue-Jun, W.Z. et al. (2022). Different eye-tracking patterns in autism spectrum disorder among toddler and preschool children. *Frontiers in Psychiatry.* 13: 1–10.

[9] Angelina, V., Atsushi, S., Tony, C., H J.M., and Teodora, G. (2018). Simulating interaction: Using gaze contingent eye-tracking to measure the reward value of social signals in toddlers with and without autism. *Developmental cognitive neuroscience*. 29: 21–29.

[10] Alcaiz Mariano, C.G. et al. (2022). Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis. *Autism Research*. 15(1): 131–145.

[11] Weng Libo, L.W. et al. (2023). A 3d graph convolutional networks model for 2d skeleton-based human action recognition. *IET Image Processing*. 17(3): 773–783.

[12] Sijie Yan, Yuanjun Xiong, Dahua Lin (2018). Spatial-temporal graph convolutional networks for skeleton-based action recognition. *arXiv:1801.07455*.

[13] E., L.D., and H, M.S. (2021). Moving toward understanding autism: Visual-motor integration, imitation, and social skill development. *Pediatric neurology*. 98—105.

[14] Ziyuan Huang, Shiwei Zhang, Liang Pan, Zhiwu Qing, Mingqian Tang, Ziwei Liu, Marcelo H. Ang Jr (2021). Tada! Temporally adaptive convolutions for video understanding. *arXiv:2110.06178*.

[15] Jin Xuemei, Z.H. et al. (2023). Identifying activity-level movement features of children with ASD based on ADOS videos. *Scientific Reports*. 13(1): 1–12.

[16] ID. Berardini, L., Migliorelli, S., Moccia, M., Naldini, G. D., Angelis and E. Frontoni (2020). Evaluating the autonomy of children with autism spectrum disorder in washing hands: a deep-learning approach. *In Proceedings of IEEE Symposium on Computers and Communications (ISCC)*. France, pp. 1-7, doi: 10.1109/ISCC50000.2020.9219648.

[17] Howsmon DP, K.U. et al. (2017). Classification and adaptive behavior prediction of children with autism spectrum disorder based upon multivariate data analysis of markers of oxidative stress and DNA methylation. *PLoS Computational Biology*. 13.

[18] K. Sun, L. Li, L. Li, N. He and J. Zhu (2020). Spatial Attentional Bilinear 3D Convolutional Network for Video-Based Autism Spectrum Disorder Detection. *In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Spain. pp. 3387-3391, doi: 10.1109/ICASSP40776.2020.9054641.

[19] Waselallah, A.F. and Saeed, A.M. (2022). Classification and detection of autism spectrum disorder based on deep learning algorithms. *Computational Intelligence and Neuroscience*. 2022.

[20] Suravi, P. and Kumar, K.S. (2021). Autism spectrum disorder in India: A scoping review. *International Review of Psychiatry*. 33(1): 81–112.
[21] Qureshi, M.N. et al. (2019). Evaluation of functional decline in Alzheimer's dementia using 3d deep learning and group ICA for rs-fMRI measurements. *Frontiers in Aging Neuroscience*. 11.
[22] Roberta, V., Robert, H., Patrik, K., and Peter, S. (2022). Human activity classification using the 3dcnn architecture. *Applied Sciences*. 12(2): 9–31.
[23] Kiruthigha M and Suresh Jaganathan (2021). Graph Convolutional Model to Diagnose Autism Spectrum Disorder Using Rs-Fmri Data. *In Proceedings International Conference on Computer, Communication and Signal Processing (ICCCSP).* Chennai, India, pp. 1-5, doi: 10.1109/ICCCSP52374.2021.9465490.
[24] Kiruthigha, M. and Suresh, J. (2023). A model for diagnosing autism patients using spatial and statistical measures using rs-fmri and smri by adopting graphical neural networks. *Diagnostics*. 13(6): 11–43.
[25] Chu Y, W.G. et al. (2022) Multi-scale graph representation learning for autism identification with functional MRI. *Frontiers in Neuroinformatics*. 15.
[26] Nie W, C.R. et al. (2021) I-gcn: Incremental graph convolution network for conversation emotion detection. *IEEE Transactions on Multimedia*. 24: 4471–4481.