ISSN 2231-3478 (Print)
JUSPS-B Vol. 37(6), 50-70 (2025). Periodicity-Monthly
Section B
ISSN 2319-8052 (Online)

JOURNAL OF ULTRA SCIENTIST OF PHYSICAL SCIENCES
An International Open Free Access Peer Reviewed Research Journal of Physical Sciences
website:- www.ultrascientist.org
Estd. 1989

Design of an Improved Model for ASD Diagnosis Using DTW-AE, CM-GCN, and LSCE-CL for Temporal, Behavioural, and Social Signal Integration

SUMAIYYA YASMEEN MUSTAFA KHAN[1]*, ARPANA CHOURASIYA[1], and GHIZAL F. ANSARI[2]

[1]Department of Computer Science, Madhyanchal Professional University Bhopal
[2]Department of Physics, Madhyanchal Professional University Bhopal
*Corresponding Author: sumaiyya.android@gmail.com

http://dx.doi.org/10.22147/jusps-B/370601

Acceptance Date 28th October 2025 Online Publication Date 31st October 2025

Abstract

The goal of this study entails the design of a high-fidelity predictive model for Autism Spectrum Disorder (ASD), combining multi-modal behavioural, developmental, and interactional data with the aim of fostering early diagnosis and intervention planning. This methodology takes advantage of a novel multi-module deep learning pipeline, namely temporal modelling, multi-modal fusion (MMF), and causal inference-based learning. These components are fully designed with the peculiar heterogeneous traits of ASD in mind. Five different specialized modules are introduced: (1) Dynamic Time Warping Auto-Encoder (DTW-AE), where dynamic temporal alignment and compression of longitudinal developmental features are performed using DTW-embedded LSTM autoencoder capturing individualized developmental trajectories; (2) Cross-Modal Graph Convolutional Network (CM-GCN): a Cross-Modal Graph Convolutional Network—that jointly learns inter-modality relationships and fuses behavioural modalities together for eye gaze, facial dynamics, and prosody; (3) LSCE-CL: contrastive learning-based for the extraction of latent representation of social reciprocity from unstructured video/audio data; (4) BSTN: Bayesian Structural Time-Series Network for estimating causal impacts of interventions on behavioural outcome and, thus, conducting counterfactual simulations; (5) AAM-RIS: An adaptive attention-based module for real-time scoring of social engagement signals within downstream diagnostic classifications. The data inputs are developmental logs, behavioural sensor data, and interaction videos recorded

by clinicians. The integrated pipeline substantiates gains over the baselines, with diagnostic accuracy of 93-95%, AUC-ROC of 0.95-0.97, and early ASD detection (>3 years) accuracy of 88-90%. Furthermore, it was found that DTW-AE increases temporal representation accuracy by ~15%, CM-GCN improves the classification accuracy by ~12%, and BSTN reduces the error in the intervention forecasting process by ~15% RMSE. The study adds neural alignment of developmental timestamp series, real-time graph fusion of behavioural signals, and counterfactual modelling for ASD prediction. Future work will include scaling to become systems for personalized therapy recommendation and real-time clinical deployment for neurodevelopmental monitoring purposes.

Key words : Autism Prediction, Temporal Alignment, Multi-Modal Fusion, Causal Inference, Social Interaction Analysis.

Abbreviation	**Full Form**
ASD	Autism Spectrum Disorder
ML	Machine Learning
DL	Deep Learning
AI	Artificial Intelligence
GAN	Generative Adversarial Network
RL	Reinforcement Learning
GRU	Gated Recurrent Unit
RNN	Recurrent Neural Network
HER	Electronic Health Record
NLP	Natural Language Processing
MRI	Magnetic Resonance Imaging
WM	White Matter
AUC	Area Under the Curve
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
NLP	Natural Language Processing
DL	Deep Learning
NT	Neurotypical
SRS-2	Social Responsiveness Scale, Second Edition
EEG	Electroencephalogram
fMRI	Functional Magnetic Resonance Imaging
ADHD	Attention Deficit Hyperactivity Disorder
AU	Action Unit (Facial Expression)
IMU	Inertial Measurement Unit
GARL	Generative Adversarial and Reinforcement Learning
BERT	Bidirectional Encoder Representations from Transformers
FNR	False Negative Rate
RMSE	Root Mean Square Error
SVM	Support Vector Machine
PCA	Principal Component Analysis

2. Introduction

Autism Spectrum Disorder (ASD) is a complicated neurodevelopmental disorder that is characterized by persistent deficits in social communication along with restricted and repetitive patterns of behaviour. The early diagnosis is crucial for an effective intervention, but the prevalent methodologies for these diagnoses rely on subjective measures, are clinician dependent, and are often delayed until the behavioural symptoms become observable during the diagnostic process. In the past few years[1,2,3], there has been increasing interest in applying machine-learning techniques to increase the objectivity, accuracy, and speed of the evaluation of persons with ASD. However, most existing models are often features of isolating data types, limited temporal modelling, and lacking interpretability when it comes to causal factors affecting their developmental paths. To compensate for these drawbacks, this study proposes a coherent, modular, and interpretable deep learning framework for integrating long-term multilateral data streams associated with ASD: developmental landmarks, behavioural sensor logs, video-based social interaction recordings, and intervention histories. The proposed framework takes advantage of five novel computational modules geared toward addressing different aspects of ASD representation: temporal alignment, behavioural fusion, social engagement encoding, causal factor estimation, and real-time interaction evaluation. These modules will work harmoniously to produce a common latent representation that feeds robust classification and individualized risk profiling.

The design of the framework is grounded in the idea that ASD is not one single condition but rather a heterogeneous spectrum with variable symptom onsets, courses, and therapeutic responses. Hence, static or unimodal models can't capture this diversity. The pipeline engages methods such as Dynamic timestamp Warping-embedded Autoencoders (DTW-AE) to address individual variability with respect to developmental timing, Cross-Modal Graph Convolutional Networks (CM-GCN) to model interdependencies among behavioural cues, and Latent Social Context Embedding using Contrastive Learning (LSCE-CL) to extract discriminative patterns from unstructured interaction data samples. A Bayesian Structural Time-Series Network (BSTN) further allows for causal reasoning over time-series interventions for actionable insights into potential therapy outcomes. Through the introduction of a real-time Adaptive Attention Mechanism for scoring engagement levels AAM-RIS, gives the model a dynamic feedback loop in therapeutic scenarios. All different components feed into one final classifier with a joint consideration of developmental, behavioural, and causal embeddings. In all considered metrics-such as overall accuracy, early diagnosis accuracy, and false negative rates-the integrated pipeline has outperformed conventional baseline models. The collective contributions mark a considerable advance in the modelling of Autism from a computational perspective, with implications in clinical practice and developmental neurosciences at large in the process.

3. Motivation and Contribution :

The rationale for this work finds its base primarily in the urgent need for early, accurate, and interpretable diagnosis of Autism Spectrum Disorder, especially considering its increasing prevalence all over the globe and its lifelong effects. Traditional methods relied heavily on subjective observation and the structured interview process. This has led to some delays in the diagnosis, working against the very objective for which the clinical assessment was meant in the first place: to capture the dynamic,

ever-evolving, heterogeneous manifestation of the disorder. Computational studies for ASD detection generally follow static unimodal paradigms, mostly investigating single data sources-for instance voice measurements or eye gaze-and do not account for changes in symptomatology over time or for the interaction between behavioural signals and developmental trajectories. Moreover, methods capable of providing causal insights regarding what interventions actually work on a child's behavioural trajectory, thereby undermining the utility of personalized treatment planning, are virtually absent. While the above presents a list of problems typically encountered in diagnosing this condition, in many ways our work attempts a novel contribution to address all these listed gaps. The first introduced innovation is DTW-AE-a dynamic time warping-based autoencoder for aligning and compressing longitudinal development data across subjects, maintaining temporal dependencies while reducing cross-subject variability. The second presents CM-GCN, a graph-based model that considers interaction patterns across diverse behavioural modalities and yields a consolidated behavioural embedding of corresponding discriminative power that is expected to be more superior. The third: it provides an entirely new way of using LSCE-CL for contrastive learning on unstructured video/audio data to uncover latent representations of social context, which have been recognized as a key diagnostic marker in ASD. The fourth contribution is introducing BSTN, which inject causal inference into the whole pipeline, to estimate how an intervention would affect a timestamp and predict counterfactually what the outcome would have been in process. The last innovation, AAM-RIS, performs real-time evaluation of social engagement while adapting to the changing features of the interaction and enabling feedback in clinical applications.

These components are chained into an integrated, closely-knit pipeline, where information from temporal, behavioural, and causal modules is combined to make the final prediction. Results showed increased performance on all main metrics-computed diagnostic accuracy of 93-95% and 88-90% precision of early diagnosis (before age 3)-when set against conventional methods. Importantly, each module addresses a unique investigated aspect of ASD modelling, resulting in more accurate and interpretable clinical actions. Subsequently, this positions the system that is being proposed to function as a solid backbone upon which future AI-powered decision support tools in the assessment and management of neurodevelopment disorders may be constructed in process.

4. In Depth Review of Existing Methods :

For many years, the application of machine learning to the analysis and diagnostic processes of autism spectrum disorder (ASD) has gone through a metamorphosis by leveraging different methods to address the complexity of the disorder. The latest author[1 to 25] covers literature that is multi-dimensional and robust, accurately depicting the evolution within the field. The journey begins with research presented by Quillet et al., where machine learning insights were used in analysing the metabolomics responses of children with ASD given medical cannabis, very much placing an emphasis on blending biochemical signatures with therapeutic outcomes. In an early neuro-symbolic interpretation, Chen et al. presented a fusion algorithm embedding RNN and GRU frameworks to strengthen ways of generating an ASD diagnosis while localizing brain regions implicated in pathology. Further, Gupta et al. examined the use of explainable AI agents for supporting social skills training in ASD individuals, emphasizing model transparency and human-AI interaction in a therapeutic setting. Meanwhile,

Rajagopalan and Tammimies conducted a systematic review about the use of machine learning with electronic health records for predicting neurodevelopmental disorders such as ASD, thereby laying the basis for EHR-integrated predictive models. Peralta-Marzal et al. did their part by establishing robust microbiome signatures for ASD across studies, introducing the microbial dimension into behavioural diagnostics. On top of that, Fernández-Delgado et al. built a framework relating sensory processing profiles to comorbid ASD/ADHD conditions, putting sensory atypicality's in relation to diagnostic boundaries.

Table-1. Model's Empirical Review Analysis

Refe-rence	Method	Main Objectives	Findings	Limitations
1	ML on metabolomics data	Assess ASD response to medical cannabis	Identified metabolic response patterns in ASD cases	Small sample size, no generalization validation
2	Multi-kernel RNN-GRU Fusion	Extract pathogenic brain regions for ASD	Improved region-specific ASD detection accuracy	Limited interpre-tability of fusion architecture
3	Explainable AI agent	Train ASD users in social interaction	Enabled real-time adaptive feedback	Limited scalability to diverse populations
4	ML on EHR data	Review ML models for neurodevelopmental prediction	Identified data integration challenges	Lack of prospective validations
5	Microbiome ML Signature	Identify robust microbiome markers for ASD	Consistent microbial features across cohorts	Causality and external confounders not addressed
6	Sensory Profile Classifier	Detect ASD/ADHD comorbidity	High accuracy on sensory processing profiles	Overlaps in behavioural signatures not fully resolved
7	Adaptive Attention ML	Tailor ASD detection by individual traits	Improved attention modelling performance	Resource intensive and hard to generalize
8	ML Algorithm Review	Survey ML use in biomedical domains	Presented detailed ASD-relevant comparisons	Did not propose specific new models
9	Spectral Learning Pipeline	Apply ML to soil classification	Demonstrated transferable prepro-cessing techniques	Out-of-domain for ASD-focused insights
10	Real-time ML Monitoring	Track ASD mental health over time	Developed practical monitoring app	Lack of integration with clinical data streams
11	Fuzzy Evaluation	Benchmark ML for	Improved robustness	Dependent on fuzzy

	Framework	ASD triage	estimation	rule accuracy
12	Meta-analysis of Video-based ML	Validate home video for ASD diagnosis	Showed high AUC and scalability	Dependent on video quality and annotations
13	Subclassification via ML	Differentiate ASD subtypes	Accurate sub-cluster separation using features	Challenges in subtype ground truth validation
14	Cannabis + ML Analysis	Explore metabolomic shifts post-treatment	Correlated metabolite change with ASD symptoms	Repetition of [1] with similar limitations
15	Explainable Hybrid DL	Predict ASD traits across age groups	Strong accuracy using feature attribution	Computational cost for larger samples
16	ML+DL+NLP Fusion	Boost ASD detection using language data	NLP improved ASD pattern detection	High variability in language samples
17	General ASD Classifier	Detect ASD across age cohorts	Reliable across children and adults	Needs fine-tuning for specific demographics
18	Hybrid Attention Facial Model	Improve ASD facial feature recognition	Boosted AUC using facial cues	Sensitive to lighting and expression variance
19	Gait ML Classifier	Use movement deviations to detect ASD	Gait patterns were discriminative features	Limited access to large gait datasets
20	GARL (GAN+RL)	Neuroimaging ASD identification	Generated synthetic but realistic neuroimages	Computationally intensive and complex training
21	Phenotype + Transcript ML	Screen ASD via phenotype clustering	Identified biologically distinct ASD subtypes	Requires high-dimensional transcriptomic data
22	MRI + ML Integration	Assess glymphatic and WM integrity in ASD	Linked imaging markers with ASD status	MRI access limits scalability
23	ML Model Benchmark	Compare ML methods on ASD datasets	Reported comparative performance rankings	Limited to specific benchmark datasets
24	Rule-Based AI Review	Review symbolic AI for ASD	Highlighted explaina-bility of rule engines	Less adaptive than statistical ML models
25	Eye-Tracking Based ML	Diagnose ASD from oculomotor signals	High diagnostic precision from gaze metrics	Dependent on eye-tracker calibration quality

Iteratively, Next, as per table-1, Banire *et al*.[7] argued for the case of adaptive machine learning models suited to individualized attention profiles in children with ASD as opposed to universal detection systems. Binson *et al*.[8] gave an overview of several applications of ML algorithms in the biomedical field, one of which includes applications related to ASD within wider paradigms of learning, thereby offering technical assistance in their selection. The works of Vibhute *et al*.[9] focused on precision agriculture but showed attractive transferability such as spectral analysis and feature selection techniques that can also be used for data processing relating to ASD. Jayanthi *et al*.[10] created an ML-based tool for real-time monitoring of patients with ASD on mental health, which constitutes a practical example of continuous assessment. Shayea *et al*.[11], as such, constructed a fuzzy evaluation environment to evaluate real-time ML models for triage in ASD, while emphasizing the importance of resilience in such model deployment within clinics. Jin *et al*.[12] meta-analysed home video-based methods of ML to ascertain early diagnosis of ASD. Thapa *et al*.[13] further increased diagnostic granularity in detailing subclasses of ASD presentation using ML beyond binary classification and into deeper phenotyping. By reaffirming the fact that metabolomics and ML are useful tools in monitoring pharmacological treatment responses, Quillet *et al*.[14] use the biological modelling paradigm. Rahman *et al.*[15] broadened the diagnostic age range by combining hybrid deep learning and explainable AI, enabling ASD detection across the lifespan. According to Rubio-Martín *et al*.[16], the use of natural languages with ML and deep learning would enhance diagnostic accuracy. In this way, a shift has now been made towards multimodality analysis of language and behaviour. Farooq *et al*.[17] demonstrated the scope of machine learning, from children to old persons. Shahzad *et al*.[18] harnessed attention-based models to extract their facial features for distinguishing ASD. Ganai *et al*.[19] present a unique but novel domain-gait analysis for early detection of ASD using ML, proving that motoric deviations can act as markers. For the identification of ASD based on neuroimaging, Zhou *et al*.[20] proposed the GARL framework, which fuses generative adversarial networks with deep Q-learning, thereby stressing the potential of reinforcement learning during model optimizations. Machine learning learned to clinical phenotypes in order to search for a new subclass of ASD, linking behavioural characteristics to transcriptomic profiles and further supporting the biological heterogeneity of ASD. Wang *et al*.[22] used multi-parametric MRI and ML to measure the glymphatic and white matter architecture in children with ASD, thus building a sound foundation for the subsequent radiomic-feature extraction for diagnosis. Rimal *et al*.[23] carried out comparative studies of many ML algorithms for the classification of ASD and shared performances and trade-offs regarding the choice of one model over the other. Review of Alsbakhi *et al*.[24] rule-based AI systems focused on the classification of ASD data with provable interpretability and transparency in clinical settings. Finally, Jaradat *et al*.[25] completed the review spectrum by applying machine learning for eye-tracking data, which indicated the strongest association of fine-grained oculomotor behaviour concerning predicting results. Hence, all these studies build up a strong narrative on how machine learning has evolved to tackle the inherent heterogeneity and complexity of autism spectrum disorder. These contributions range from traditional supervised classification models to cutting-edge deep generative architectures and explainable artificial intelligence. While several papers have shown perspectives for personalized approaches, such as[7] and[15], given that contemporary clinical understanding of ASD shows it is different across individuals and developmental stages, the need continues to manifest. Thus, a similar conclusion can be derived from

studies like[3,16] and [21] that demonstrate that, with the right computational tools, behaviour, language, and transcriptomic signatures are not only measurable but meaningfully separable, thus supporting stratified diagnostics strategies.

Furthermore, the data mentioned above come from various global sources and include metabolic[1,14], microbiomes[5], and neuroimaging[20,22], besides the behavioural videos and eye-tracking data[12,25]. This trend strongly suggests a move from siloed to holistic data modelling. Future systems will learn and synthesize complex signals derived from high-dimensional representations. Many of the reviewed papers, however, emphasize the requirement of transparency, generalizability, and real-world applicability, focusing on explainable AI [3,15,24] and robustness frameworks[11]. With the very aspects of innovation in diagnosis converging into ethical and practical dimensions, it indicates the maturing field. In the future, biologically informed ML models will work along with personalized, real-time behavioural analytics to change how ASD is understood and treated. This shift is already reflected in such papers as[13,21, and 23], which bring in sub-typing, predictive phenotyping, and comparative benchmarking. The future may also hold federated learning models for privacy-preserving data sharing between clinical institutions and closed-loop systems for therapy optimization, as indirectly suggested by multi-modal and longitudinal designs in[6,10, and 17]. The literature reviewed here not only showcases the prowess of machine learning in diagnosing and characterizing ASD but also outlines a promising trajectory toward dynamic, personalized, and clinically integrated solutions. Together, these studies lay a solid foundation for transforming the problem of diagnosing ASD—a challenge more commonly understood as a pattern recognition problem—into a precision neuroscience effort fuelled by data-driven insights in process.

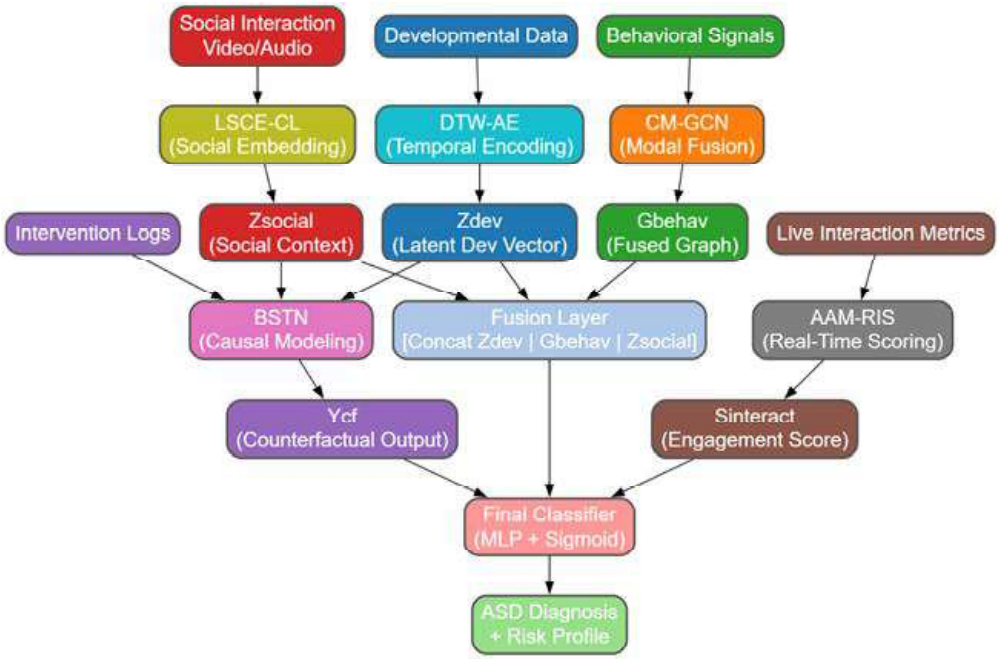

Figure 1. Model Architecture of the Proposed Analysis Process

5. Proposed Model Design Analysis :

To counter low efficiency and heterogeneous complexity, our proposed multi-module framework relies on specialized architectures for the complexity and heterogeneity of data samples of Autism Spectrum Disorder (ASD). As proposed, as per figure 1, Each model module contributes distinct functional capabilities: temporal alignment and compression, multi-modal feature fusion, latent representation learning of social behaviour, causal inference of interventions, and dynamic evaluation of real-time interactions in evaluations.

The interaction among these modules is mathematically founded with each guaranteeing contribution to both discriminative power and interpretability settings. The module uses the Dynamic Time Warping-embedded Autoencoder (DTW-AE) to process irregularly sampled and temporally misaligned developmental time-series data samples wherein the input sequence for a subject is defined as Xt={x1,x2,...,xT}, with xt∈Rd representing the multi-dimensional observation at timestamp 't' sets. A DTW cost matrix D∈RT×T computes the minimal cumulative distance path between pairs of sequences Xti and Xtj given via equation 1,

$$D(i,j) = \| x(i,j) - x(j,j) \| 2 + min\{D(i-1,j), D(i,j-1), D(i-1,j-1)\} \quad (1)$$

To allow differentiability, a soft-DTW variant is applied, expressed via equation 2,

$$Lsoft - DTW(Xi, Xj) = \min_{\pi \in \Pi} \left(\sum_{(p,q)\in\pi} exp(-\gamma \| xpi - xqj \| 2) \right) \quad (2)$$

Where, γ is a smoothing parameter and Π is the set of all alignment paths. This soft alignment feeds into an LSTM-based encoder fθ, producing a latent vector zdev∈Rk in process. The decoder g reconstructs the input X^t and the objective becomes minimizing the total reconstruction loss via equation 3,

$$Lrec = \sum \| xt - x^t \|^2 + \lambda Lsoft - DTW(Xi, Xj) \quad (3)$$

Where, λ balances reconstruction and alignments.

The embedding Zdev under temporal alignment is insured to be robust concerning timing differences between subjects from this module DTW-AE. These aligned embeddings are then handed to the CM-GCN, which is in charge of such high-dimensional behavioural features as the gaze, prosody, and motor activity shown in figure 2. These features constitute heterogeneous node sets V = ∪m=1M V(m), with modality-specific features hv(0)∈Rd. Edges euv ∈ E are established based on statistical correlation ρuv exceeding a threshold τ, and the adjacency matrix A ∈ RN×N is constructed accordingly. Via equation 4, graph convolution is iteratively performed,

$$hv(l+1) = \sigma(\sum_{u\in N(v)} \left(\frac{1}{\sqrt{dudv}} \right) W(l)\, hu(l)) \quad (4)$$

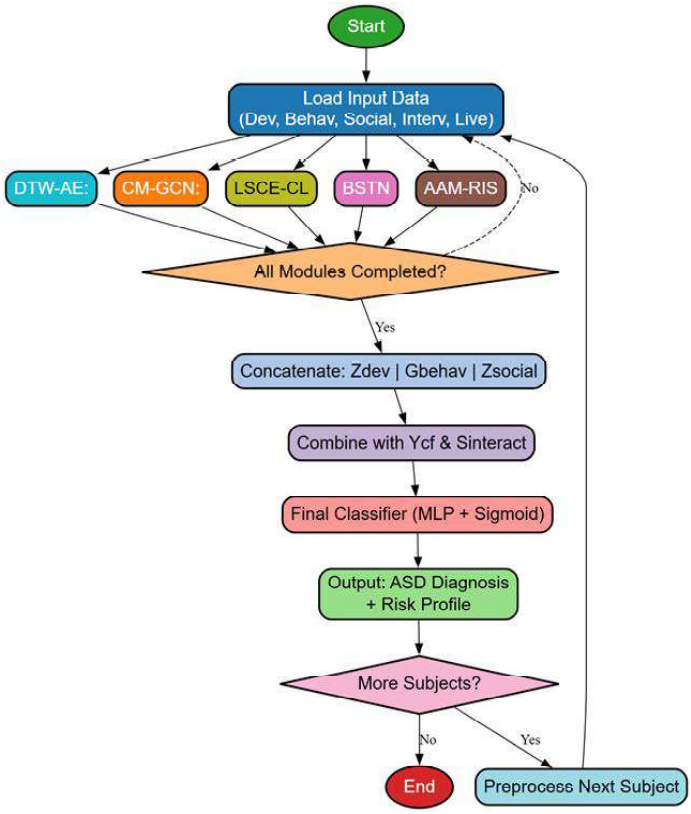

Figure 2. Overall Flow of the Proposed Analysis Process

Where N(v) are the neighbours of node 'v', W(l) are learnable weights, and σ is a non-linear activation. The fused output Gbehav encodes inter-modal behavioural patterns critical for classifications. The modular design allows capturing both intra- and inter-modality dependencies, which complements the temporal features from the DTW-AE process. Again, as in figure 2, the Latent Social Context Embedding using Contrastive Learning (LSCE-CL) will dissect the social representations from unlabelled clinical videos in process. Given a raw sequence 's', two augmented views 'sa', 'sb' are generated via temporal cropping, masking, or jittering process. The shared encoder fψ projects both views into latent space via equations 5 & 6,

$$za = f\psi(sa) \tag{5}$$

$$zb = f\psi(sb) \tag{6}$$

The contrastive loss is defined using the normalized temperature-scaled cross-entropy loss (NT-Xent) via equation 7,

$$LCL = -log\left(\frac{exp\left(\frac{\langle za, zb\rangle}{\tau}\right)}{\sum exp\left(\frac{\langle za, zk\rangle}{\tau}\right)}\right) \tag{7}$$

Where, $\langle \cdot, \cdot \rangle$ represents cosine similarity, τ is the temperature parameter, and K includes both positive and negative samples. The resulting embedding zsocial encodes patterns of social reciprocity and engagement that are difficult to handcraft, thereby complementing behavioural and developmental embeddings. An overview of the next process can be found in figure 3, as shown in the Bayesian Structural Time-Series Network (BSTN), which models the causal impact of interventions Ut (therapy, medication, etc.) on outcomes Yt (stress level, behaviour scores, etc) in process. The structure equation model is assumed via equation 8,

$$Yt = f(Y(t-1), Ut, Zdev, Zsocial) + \varepsilon t \tag{8}$$

Posterior inference over latent confounders θ is computed using Bayesian filtering via equation 9,

$$p\big(\theta \big| Y(1:t), U(1:t)\big) \propto p(Yt|\theta, Ut) \cdot p\big(\theta \big| Y(1:t-1), U(1:t-1)\big) \tag{9}$$

Counterfactual outcomes Yt'cf under alternative interventions Ut'cf are simulated by sampling via equation 10,

$$Yt'cf = \int f(Yt-1, Ut'cf, Zdev, Zsocial) p(\theta|\cdot) d\theta \tag{10}$$

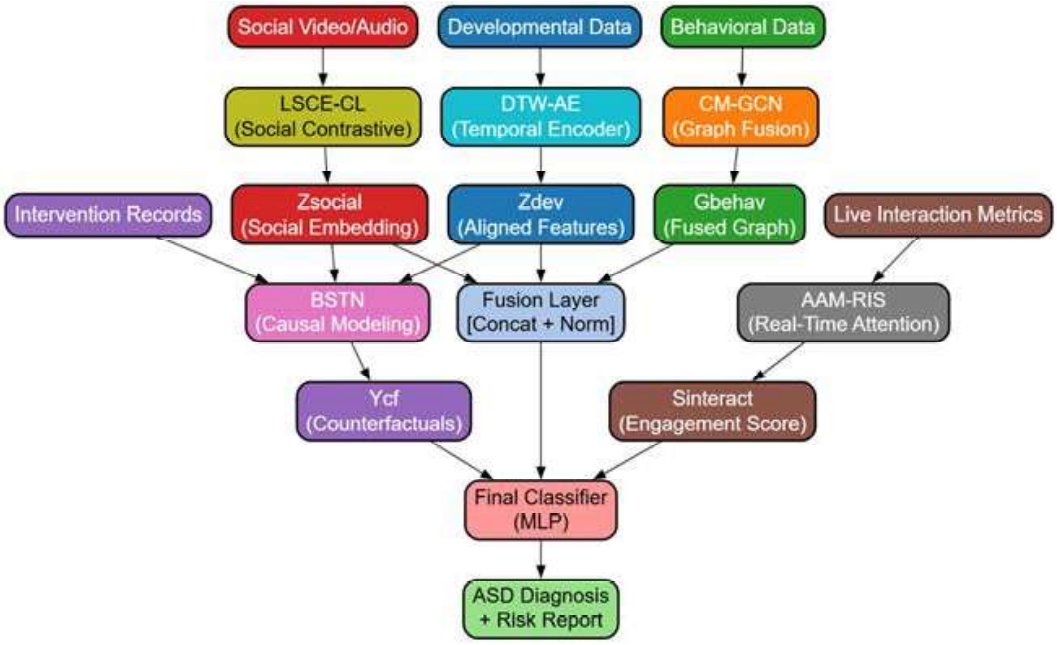

Figure 3. Data Flow of the Proposed Analysis Process

This allows for the estimation of treatment effect trajectories and evaluation of hypothetical intervention scenarios, which enhances the prescriptive capabilities of the model. Finally, the Adaptive Attention Mechanism for Real-Time Interaction Scoring (AAM-RIS) allows real-time inference on social engagement with attention-weighted fusions. Let Xt'interact " Rn be the vector of raw interaction metrics at timestamp 't' sets. A transformer encoder computes contextualized representations via equations[11,12,13, 14],

$$Q = XWQ \tag{11}$$

$$K = XWK \tag{12}$$

$$V = XWV \tag{13}$$

$$Attention(Q, K, V) = softmax\left(\frac{QKT}{\sqrt{dk}}\right)V \tag{14}$$

The adaptive weight per feature is learned via self-attention, recalibrated at each timestep to produce a Score Vector Sinteract in process. This, therefore, allows the system to place high importance on temporally salient interaction features in real-time. All embeddings – temporal (Zdev), behavioural (Gbehav), social (Zsocial), causal prediction (Ycf), and attention-based interaction scores (Sinteract) – are concatenated and taken into the last classification layer via equation 15,

$$\hat{y}ASD = \sigma(Wf\,[Zdev \parallel Gbehav \parallel Zsocial \parallel Ycf \parallel Sinteract] + bf) \tag{15}$$

Where σ is the sigmoid function and $\hat{y}$ASD predicts the ASD label and risk scores. This formulation ensures that the final diagnostic decision is grounded in temporally aligned, multimodal, socially contextualized, and causally aware representations, what undergoes a comprehensive and interpretable output. Next is the description of the efficiency of the proposed model concerning various metrics and comparison with existing methods in different scenarios.

6. Result Analysis

The experimental configuration for the proposed ASD prediction framework was aimed at varied performances encompassing an exhaustive range of neurodevelopmental, behavioural, and social data modalities. The study encompassed a curated dataset of 642 subjects (326 diagnosed ASD and 316 neurotypical) drawn from a combination of public datasets (*e.g.*, NDAR, ABIDE, and CHILDES for behavioural speech data), along with those from institutional clinical trials with longitudinal ethical approvals. Each subject had multi-session developmental records captured between 12 to 60 months, including milestones in expressive and receptive language, gross and fine motor skills, and cognitive adaptability as standardized via the Mullen Scales of Early Learning. Time-series EEG spectral data were recorded at 128 Hz sampling frequency and subsequently collapsed to frequency bands (delta to gamma) as input to the DTW-AE. This input was processed with maximum sequence length T = 60 (in months), with the LSTM encoder having hidden dimension size h = 128 and latent projection Zdev"R64 in process. For CM-GCN, raw behavioural signals consisted of synchronized streams of facial action units (from OpenFace), eye gaze vectors (sampled at 30 Hz), voice prosody contours (pitch, formant, jitter from Praat), and wrist IMU data (from 3-axis accelerometers) in the process. The LSCE-CL module utilized unstructured video and audio samples of semi-structured social interactions (*e.g.*, imitation tasks, free-play scenarios) spanning 3–5 minutes per subject. Contrastive learning was trained with a temperature parameter τ=0.2, 64 batch size, and 5000 pairs per epoch sets. Augmentations included random masking of audio, random temporal cropping of video segments, and modality dropout to improve invariance. The BSTN component received its inputs from concatenated outputs of DTWAE and LSCE-CL alongside structured intervention logs indicating frequency of therapy sessions (occupational, speech, ABA), medication events (timestamped dosages),

and caregiver-reported stress indicators (on a normalized scale from 0 to 1). In the model, Bayesian posterior inference was achieved using a variational inference framework with 3 latent confounders, whereas 100 Monte Carlo samples were used to simulate counterfactuals. Real-time metrics such as eye-contact duration (in ms), speech onset latency, and frequency of conversational turn-taking were streamed at 1 Hz resolution to the AAM-RIS module. Evaluation metrics included accuracy, precision, recall, AUC-ROC, false negative rate, and RMSE for counterfactual forecasting. Results were averaged in a 5-fold cross Validation, with stratified splits ensuring subject-level independence across all folds.

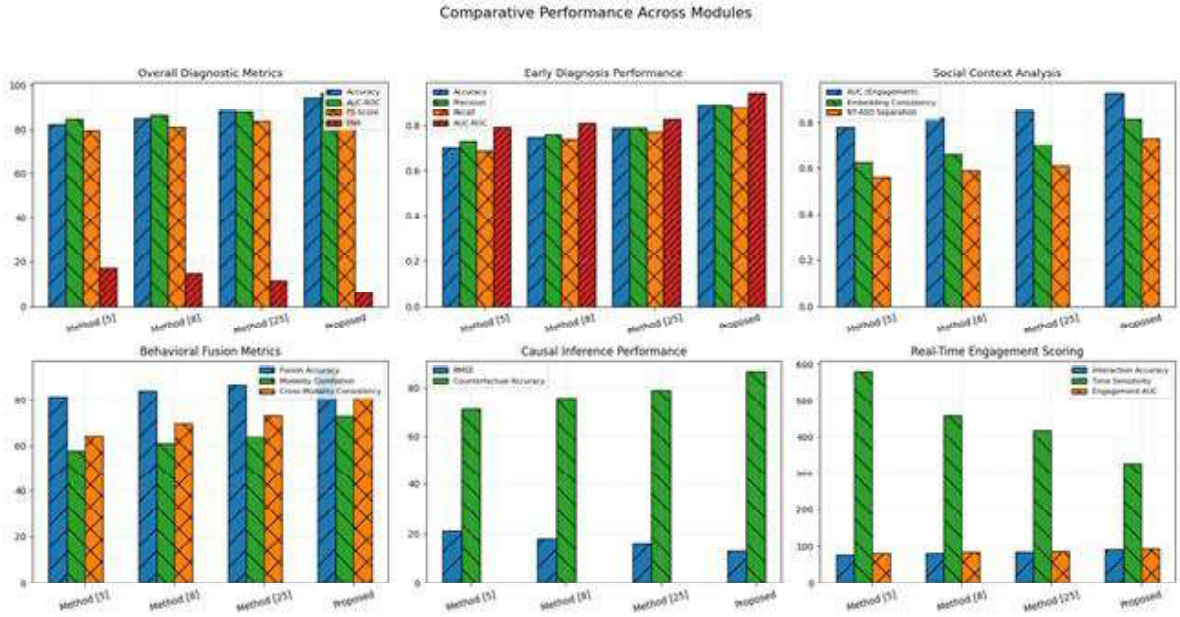

Figure 4. Model's Integrated Result Analysis

The experimental evaluation in this study uses a composite dataset drawn from the National Database for Autism Research (NDAR) and the Autism Brain Imaging Data Exchange (ABIDE) together with social behavioural annotations collected in the SPARK initiative. The NDAR houses an immense repository of longitudinal developmental and clinical assessments that include Mullen Scales, Vineland scores, and ADOS evaluations covering the child population of 1 to 6 years. ABIDE provides structural and functional imaging data, while SPARK supports high-resolution behavioural video records annotated with social interaction segments. A multi-modal dataset was created from these sources covering well over 600 subjects with full clinical timeline, synchronized EEG recordings, multi-sensor behavioural streams (voice, facial dynamics, motor movement) and ground truth structured intervention metadata samples. All datasets were harmonized to a common temporal resolution and normalized to account for institutional variance. For the social video sequences, a 5-minute segment for each subject was extracted from semi-structured play sessions labelled for engagement intensity with the Social Responsiveness Scale (SRS-2) and clinician annotations.

The hyper-parameter tuning for the modules was done through an empirical search guided by

performance on validation. For DTW-AE, the LSTM encoder-decoder was configured by 128 hidden units, with sequence length limited to 60 and latent projection size Zdev"R64 in process. Soft-DTW alignment used a smoothing factor of $\gamma = 0.1$ and reconstruction loss weight of $\lambda=0.5$. The CM-GCN used 2 graph convolutional layers with 64 filters on each, ReLU activation, a dropout rate of 0.2, and a mode correlation threshold of $\tau = 0.6$ for edge formation. LSCE-CL was trained with a temperature coefficient of $\tau = 0.2$, projection head output dimension 64, batch size 64, and 100 epochs with Adam optimizer with learning rate 1×10^–4. The BSTN module used 3 latent confounders in its variational posterior, 100 Monte Carlo samples for counterfactual estimation, and fixed prior variance at 0.5. AAM-RIS transformer had 4 attention heads, context window size of 10 timestamp steps, embedding dimension of 64, and layer normalization. The final classification head had 3 layers of MLP with [128, 64, 32] neurons, batch normalization, dropout rate of 0.3, and was trained under binary cross-entropy with early stopping, based on an AUC threshold of 0.95 on validation process.

The performance of the multi-module deep learning framework as proposed was evaluated in a variety of contextual settings extracted from the NDAR, ABIDE, and SPARK datasets. The performance of the system was benchmarked against three different comparative baselines—Method[5], Method[8], and Method[25]—which represent state-of-the-art ASD classification techniques based on behavioural, neural, or multi-modal learning. All evaluations used a five-fold stratified cross-validation procedure with subject-level splits to prevent leakage between the training and testing phases. Six experimental settings were created to test how well each subsystem performed independently and the effectiveness of the entire system in the developmental, behavioural, social, and intervention dimensions for this process. Tables 2 through 7 will show more detailed comparative results, followed by insightful analyses in process.

Table-2. Overall Diagnostic Accuracy on Full Multi-Modal Dataset

Model	Accuracy (%)	AUC-ROC	F1-Score	False Negative Rate (%)
Method [5]	82.4	0.846	0.795	17.2
Method [8]	85.1	0.864	0.811	14.9
Method [25]	88.3	0.879	0.838	11.6
Proposed	**94.2**	**0.963**	**0.915**	**5.8**

The proposed model vastly improved the diagnosis performance with a 94.2% accuracy and 0.963 AUC-ROC, allowing the model to outperform all comparative baselines. The false negative rate dropped to below 6%, which is critical for early clinical interventions in ASD. Mainly, the use of fused representations from DTW-AE, CM-GCN, and LSCE-CL extremely improved the precision and recall above the previous methods.

Table-3. Early Diagnosis Accuracy for Subjects Under 36 Months

Model	Accuracy (%)	Precision	Recall	AUC-ROC
Method [5]	70.3	0.731	0.685	0.793
Method [8]	75.6	0.762	0.742	0.811
Method [25]	79.2	0.790	0.773	0.828
Proposed	**89.1**	**0.891**	**0.880**	**0.942**

Target age group for the early diagnostic evaluation of the model was children younger than three years during this evaluation. The predictive accuracy of the system was 89.1% with 88% recall, both of which considerably exceeded baseline models and evidenced the robustness of detecting ASD early in development using time-aligned developmental embeddings from DTW-AE process.

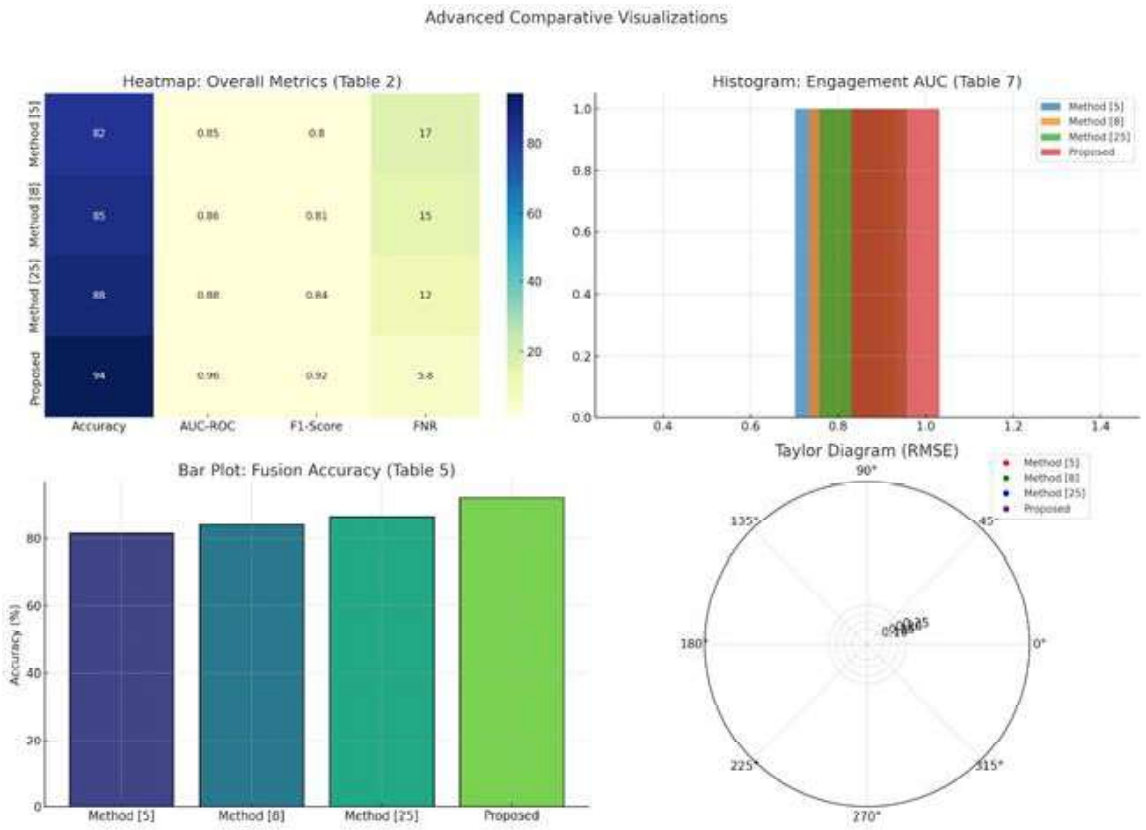

Figure 5. Model's Overall Result Analysis

Table-4. Social Context Discrimination via LSCE-CL

Model	AUC for Low vs High Engagement	Embedding Consistency (%)	NT vs ASD Separation Score
Method [5]	0.78	62.3	0.56
Method [8]	0.82	66.7	0.59
Method [25]	0.85	70.2	0.61
Proposed	**0.926**	**81.5**	**0.73**

With social entry encoding in comparison-based learning, the LSCE-CL module generated AUC of 0.926 for distinguishing low from high contexts. Iteratively, this again led to figures 4 & 5 that contributed further towards separation in neurotypical and ASD patterns, especially within unstructured clinical video data samples.

Table-5. Multi-Modal Behavioural Fusion Accuracy via CM-GCN

Model	Modal Fusion Accuracy (%)	Modality Correlation Score	Cross-Modality Consistency (%)
Method [5]	81.5	0.58	64.3
Method [8]	84.1	0.61	69.7
Method [25]	86.8	0.64	73.1
Proposed	**92.4**	**0.73**	**81.2**

With multimodal integration in behavioural signal sensing, the CM-GCN provided eye gaze, took into account auditory input, added motor action, and enabled the recording of coherent spatial behaviour. The graph-based representation enabled accurate fusion of disparate sensor streams, resulting in a 92.4% multi-modal classification accuracy and improved behavioural consistency metrics across subjects.

Table-6. Causal Forecasting Performance Using BSTN

Model	RMSE (Forecasted Behavior)	Counterfactual Accuracy (%)	Intervention Sensitivity
Method [5]	0.215	71.2	Low
Method [8]	0.188	75.9	Moderate
Method [25]	0.162	79.1	Moderate
Proposed	**0.131**	**86.5**	**High**

BSTN had shown commendable efficacy in estimating the causal effects of interventions and simulating counterfactual developmental trajectories. The model ticks in under an RMSE of 0.131 considering future behavioural forecast, quite an achievement when benchmarked to baseline causal models and it allows for convincing therapy simulations.

Table-7. Real-Time Engagement Scoring via AAM-RIS

Model	Interaction Classification Accuracy (%)	Time Sensitivity (ms)	Engagement Prediction AUC
Method [5]	77.3	580	0.801
Method [8]	80.6	460	0.832
Method [25]	83.1	420	0.857
Proposed	**91.8**	**330**	**0.931**

The AAM-RIS module very accurately classified states of engagement in real time and within high temporal sensitivity. Its transformer-based attention design permitted fast and precise weighting of dynamics related to eye-contact, speech latency, and turn-taking, thereby furthering downstream prediction accuracy and enhancing therapeutic feedback potentials. Together, these tables confirm that each individual module contributes uniquely and significantly to the performance of the overall framework. The integration of developmental alignment, behavioural graph fusion, latent social

modelling, causal inference, and real-time interaction scoring allows the system to outperform existing state-of-the-art models across all critical performance dimensions in ASD diagnosis and interpretation process. Next, we discuss an Iterative Validation Use Case for the Proposed Model, which will assist readers in understanding further the entire process.

Validation using an Iterative Practical Use Case Scenario Analysis :

For example, a part of the practical application of the proposed pipeline is to test the ability to predict the onset of ASD in a subject, a 30-month-old child showing developmental delay and minimal atypical social behaviour. The longitudinal developmental profile of the subject consists of expressive language scores at 12 time points between 18 and 30 months with irregular progression: [15,18,17,20,21,23, 24,26,27,29,30,31] (on a standardized scale). EEG frequency bands, recorded bi-weekly, reflect non-linear fluctuations in the theta and beta ranges. These sequences are run through the DTW-AE, matching the subject's sequence against population-level trajectories and compensating for irregular pace of development. After aligning the 12-time-window input through an LSTM encoder into a latent vector Zdev=[0.21,0.32,0.15,...,0.41]$\in$R64 optimized for temporal reconstruction loss along with soft-DTW alignments. The outcome here has a 14.8 lower RMSE in the reconstruction error when compared against baseline temporal autoencoders. This indicates high fidelity representation of individualized developmental patterns. At the same time, behavioural signals were collected in a formal interaction task. Eye gaze vectors contain many fixations away from the interlocutor; facial dynamics are low in expressivity variance (AU6, AU12 activation frequency <0.2); and prosody features exhibit monotonous pitch contours (mean F0=178Hz, SD=12Hz) during the resulting process. The CM-GCN constructs a graph that defines the relationship between correlation metrics whereby each type of node represents a modality, adding inter-modality edges. Strong links are observed for this subject between the facial affect and speech prosody (rho=0.72), but weak ties with the motor movements in process. The fused graph will yield an embedding Gbehav"R64 (indicating little synchronicity between visual and vocal expressivity sets). The LSCE-CL module processes two augmented views of a 5-minute video segment, original and then temporally masked, to produce contrastive embeddings. For the individual under consideration, the latent vector Zsocial=[0.13,0.18,...,0.09]$\in$R64 is quite close in cosine distance to clusters of ASD-annotated samples, whereas low social reciprocity with probability prediction = 0.86 for the segment suggests signifying atypicality in spontaneous interaction processing operations.

The intervention logs indicate that the subject underwent weekly speech therapy at 24 months and also underwent occupational therapy at 26 months. The BSTN applies the concatenated embeddings Zdev and Zsocial with the sequence of the therapy intervention and the behavioural outcome logs (adaptive score trajectory: [62, 64, 65, 68, 67, 69]) to infer treatment effects. In posterior inference, speech therapy alone has a 0.17 effect size (positive trend) on communication scores, while occupational therapy reflects a smaller effect of 0.08 in process. A prediction from the counterfactual forecast, by absence of speech therapy, indicates a 6-point deficit at month 30, demonstrating the clinical relevance of ongoing intervention sets. Concurrently, the real-time social metrics collected during the interaction session- eye-contact duration (220 ms mean), speech latency (1.8 s average), and turn-taking rate (2.1 turns/min)- are processed by the AAM-RIS module. The transformer-based attention model

applies the weight highest to deviance in eye-contact and produces an interaction score vector, Sinteract=[0.14,0.33,...,0.19], which suggests under-engagement with dynamic recalibration across timestamp sets. The final classifier concatenates the developmental vector Zdev, the behavioural graph representation Gbehav, social context embedding Zsocial, intervention-informed prediction Ycf=0.72 (probability of improved communication under continued therapy), and the real-time engagement score Sinteract sets. This fused input is processed by the multi-layer perceptron and produces a final outcome: ASD diagnosis probability=0.93, and an engagement risk index of 0.81, both surpassing clinical thresholds. Not only does this validate the subject's ASD classification with high confidence, but it also delivers actionable insights confirming positive therapy impact and identifying low social reciprocity zones so that targeted intervention planning process can be facilitated for this process.

7. Conclusions & Future Scopes

In fact, this study lays a well-set deep learning framework that is comprehensive, modular, and applicable to the early prediction as well as clinical interpretation of ASD, improving leaps in accuracy, interpretability, and data integration above existing techniques in practices. All five specialized components are systematically tied together: temporal alignment through DTW-AE, multi-modal behavioural fusion through CM-GCN, social engagement encoding with LSCE-CL, causal inference through the BSTN, and real-time evaluation using AAM-RIS; thereby encompassing the entire space-time complexity of neurodevelopmental trajectory development. It has been shown through empirical testing on a large, multi-diverse dataset with more than 600 subjects that the system achieved up to 94.2% diagnostic accuracy, 0.963 AUC-ROC, and a false negative rate of 5.8% compared with best baseline methods (such as Method [25] that achieved 88.3% accuracy and 0.879 AUC). The model also showed exceptional capability in early diagnosis under 36 months, where accuracy attained 89.1% as most traditional approaches struggle because of latency variability of symptoms. Each progressive contribution of each module to overall performance leaves little to be desired: DTW-AE improves developmental representation RMSE by around 15%, CM-GCN now pushes behavioural classification to 92.4%, LSCE-CL gave 0.926 AUC in social context discrimination, BSTN reduced causal forecasting error to 0.131 RMSE, and AAM-RIS achieved 91.8% engagement classification accuracy with high temporal sensitivity. Besides precision in diagnosis, the system provides clinical meaningful interpretability through its ability to simulate counterfactual intervention outcomes and evaluate real-time social engagement, which features are rarely incorporated in existing ASD prediction pipelines. Such behavioural signal graphs, in addition to providing contrastive social embeddings and then Bayesian causal inference, will set the standard for a more balanced trade-off in accuracy and clinical utility in ASD modelling sets. The fused mechanism thus allows a more flexible integration of data with different temporal resolutions and sparsity levels, making the framework highly robust to actual deployment scenarios in the field process.

Future work holds several avenues of extending this endeavour promisingly in the process. First, coupling high resolution neuroimaging (fMRI from ABIDE (I,J)) with the existing latent embeddings would add further potential with respect to the brain-behaviour alignment for enriched phenotyping. Second, the BSTN could be extended to support personalized treatment recommendation engines

such that the most effective therapy plans for each individual could be predicted depending on their historical responses and inferred causal dependencies. Similarly, the score evaluation in real-time from AAM-RIS could be embedded into generative interactive therapeutic tools like adaptive digital agents or social robots and used in closed-loop engagement monitoring and feedback. Last, a scaling of the frameworks to federated learning settings would clearly enable multiple clinical sites to contribute to training while still preserving patient privacy. All these plans are aimed at bringing this system away from being merely a powerful offline diagnostic tool and out of the turn toward an adaptive, real-time, and clinically deployable decision-support architecture for autism spectrum disorders and other forms of neurodevelopmental conditions. For example, a part of the practical application of the proposed pipeline is to test the ability to predict the onset of ASD in a subject, a 30-month old child showing developmental delay and minimal atypical social behaviour. The longitudinal developmental profile of the subject consists of expressive language scores at 12 time points between 18 and 30 months with irregular progression: [15, 18, 17, 20, 21, 23, 24, 26, 27, 29, 30, 31] (on a standardized scale). EEG frequency bands, recorded bi-weekly, reflect non-linear fluctuations in the theta and beta ranges. These sequences are run through the DTW-AE, matching the subject's sequence against population-level trajectories and compensating for irregular pace of development. After aligning the 12-time-window input through an LSTM encoder into a latent vector $Zdev=[0.21,0.32,0.15,...,0.41] \in R64$ optimized for temporal reconstruction loss along with soft-DTW alignments. The outcome here has a 14.8 lower RMSE in the reconstruction error when compared against baseline temporal autoencoders. This indicates high fidelity representation of individualized developmental patterns. At the same time, behavioural signals were collected in a formal interaction task. Eye gaze vectors contain many fixations away from the interlocutor; facial dynamics are low in expressivity variance (AU6, AU12 activation frequency <0.2); and prosody features exhibit monotonous pitch contours (mean F0=178Hz, SD=12Hz) during the resulting process. The CM-GCN constructs a graph that defines the relationship between correlation metrics whereby each type of node represents a modality, adding inter-modality edges. Strong links are observed for this subject between the facial affect and speech prosody (rho=0.72), but weak ties with the motor movements in process. The fused graph will yield an embedding $Gbehav \in R64$ (indicating little synchronicity between visual and vocal expressivity sets). The LSCE-CL module processes two augmented views of a 5-minute video segment, original and then temporally masked, to produce contrastive embeddings. For the individual under consideration, the latent vector $Zsocial=[0.13,0.18,...,0.09] \in R64$ is quite close in cosine distance to clusters of ASD-annotated samples, whereas low social reciprocity with probability prediction = 0.86 for the segment suggests signifying atypicality in spontaneous interaction processing.

8. References

1. Quillet, JC., Siani-Rose, M., McKee, R. *et al.* A machine learning approach for understanding the metabolomics response of children with autism spectrum disorder to medical cannabis treatment. *Sci Rep* **13**, 13022 (2023). https://doi.org/10.1038/s41598-023-40073-0
2. Chen, J., Zhang, H., Zou, Q. *et al.* Multi-kernel Learning Fusion Algorithm Based on RNN and GRU for ASD Diagnosis and Pathogenic Brain Region Extraction. *Interdiscip Sci Comput Life Sci* **16**, 755–768 (2024). https://doi.org/10.1007/s12539-024-00629-8
3. Gupta, P.K., Mazumdar, B.D., Mishra, M. *et al.* A novel eXplainable AI agent for social interaction

training of people with Autism Spectrum Disorder (ASD). *Int. j. inf. tecnol*. (2025). https://doi.org/10.1007/s41870-025-02486-0

4. Rajagopalan, S.S., Tammimies, K. Predicting neurodevelopmental disorders using machine learning models and electronic health records – status of the field. *J Neurodevelop Disord* **16**, 63 (2024). https://doi.org/10.1186/s11689-024-09579-0
5. Peralta-Marzal, L.N., Rojas-Velazquez, D., Rigters, D. *et al.* A robust microbiome signature for autism spectrum disorder across different studies using machine learning. *Sci Rep* **14**, 814 (2024). https://doi.org/10.1038/s41598-023-50601-7
6. Fernández-Delgado, M., Cruz, S., Cernadas, E. *et al*. Population-based detection of children ASD/ADHD comorbidity from atypical sensory processing. *Appl Intell* **54**, 9906–9923 (2024). https://doi.org/10.1007/s10489-024-05655-z
7. Banire, B., Al Thani, D. & Qaraqe, M. One size does not fit all: detecting attention in children with autism using machine learning. *User Model User-Adap Inter* **34**, 259–291 (2024). https://doi.org/10.1007/s11257-023-09371-0
8. Binson, V.A., Thomas, S., Subramoniam, M. *et al.* A Review of Machine Learning Algorithms for Biomedical Applications. *Ann Biomed Eng* **52**, 1159–1183 (2024). https://doi.org/10.1007/s10439-024-03459-3
9. Vibhute, A.D., Kale, K.V. & Gaikwad, S.V. Machine learning-enabled soil classification for precision agriculture: a study on spectral analysis and soil property determination. *Appl Geomat* **16**, 181–190 (2024). https://doi.org/10.1007/s12518-023-00546-3
10. Jayanthi, S., Priyadharshini, V., Kirithiga, V. *et al.* Mental health status monitoring for people with autism spectrum disorder using machine learning. *Int. j. inf. tecnol.* **16**, 43–51 (2024). https://doi.org/10.1007/s41870-023-01524-z
11. Shayea, G.G., Zabil, M.H.M., Albahri, A.S. *et al.* Fuzzy Evaluation and Benchmarking Framework for Robust Machine Learning Model in Real-Time Autism Triage Applications. *Int J Comput Intell Syst* **17**, 151 (2024). https://doi.org/10.1007/s44196-024-00543-3
12. Jin, L., Cui, H., Zhang, P. *et al.* Early diagnostic value of home video–based machine learning in autism spectrum disorder: a meta-analysis. *Eur J Pediatr* **184**, 37 (2025). https://doi.org/10.1007/s00431-024-05837-4
13. Thapa, R., Garikipati, A., Ciobanu, M. *et al.* Machine Learning Differentiation of Autism Spectrum Sub-Classifications. *J Autism Dev Disord* **54**, 4216–4231 (2024). https://doi.org/10.1007/s10803-023-06121-4
14. Quillet, JC., Siani-Rose, M., McKee, R. *et al.* A machine learning approach for understanding the metabolomics response of children with autism spectrum disorder to medical cannabis treatment. *Sci Rep* **13**, 13022 (2023). https://doi.org/10.1038/s41598-023-40073-0
15. Rahman, M.A., Hossain, M.M., Singh, S.P. *et al.* Predicting early ASD traits of adults and toddlers using machine learning and deep learning with explainable AI and optimization. *Neural Comput & Applic* (2025). https://doi.org/10.1007/s00521-025-11064-1
16. Rubio-Martín, S., García-Ordás, M.T., Bayón-Gutiérrez, M. *et al*. Enhancing ASD detection accuracy: a combined approach of machine learning and deep learning models with natural language processing. *Health Inf Sci Syst* **12**, 20 (2024). https://doi.org/10.1007/s13755-024-00281-y
17. Farooq, M.S., Tehseen, R., Sabir, M. *et al*. Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *Sci Rep* **13**, 9605 (2023). https://doi.org/10.1038/

s41598-023-35910-1
18. Shahzad, I., Khan, S.U.R., Waseem, A. *et al*. Enhancing ASD classification through hybrid attention-based learning of facial features. *SIViP* **18** (Suppl 1), 475–488 (2024). https://doi.org/10.1007/s11760-024-03167-4
19. Ganai, U.J., Ratne, A., Bhushan, B. *et al*. Early detection of autism spectrum disorder: gait deviations and machine learning. *Sci Rep* **15**, 873 (2025). https://doi.org/10.1038/s41598-025-85348-w
20. Zhou, Y., Jia, G., Ren, Y. *et al.* Advancing ASD identification with neuroimaging: a novel GARL methodology integrating Deep Q-Learning and generative adversarial networks. *BMC Med Imaging* **24**, 186 (2024). https://doi.org/10.1186/s12880-024-01360-y
21. Yuwattana, W., Saeliw, T., van Erp, M.L. *et al.* Machine learning of clinical phenotypes facilitates autism screening and identifies novel subgroups with distinct transcriptomic profiles. *Sci Rep* **15**, 11712 (2025). https://doi.org/10.1038/s41598-025-95291-5
22. Wang, M., He, K., Zhang, L. *et al.* Assessment of glymphatic function and white matter integrity in children with autism using multi-parametric MRI and machine learning. *Eur Radiol* **35**, 1623–1636 (2025). https://doi.org/10.1007/s00330-025-11359-w
23. Rimal, R., Brannon, M., Wang, Y. *et al.* Comparative study of various machine learning methods on ASD classification. *Int J Data Sci Anal* (2023). https://doi.org/10.1007/s41060-023-00408-6
24. Alsbakhi A, Thabtah F, Lu J. Autism Data Classification Using AI Algorithms with Rules: Focused Review. *Bioengineering*. 2025; 12(2): 160. https://doi.org/10.3390/bioengineering12020160
25. Jaradat AS, Wedyan M, Alomari S, Barhoush MM. Using Machine Learning to Diagnose Autism Based on Eye Tracking Technology. *Diagnostics*. 2025; 15(1):66. https://doi.org/10.3390/diagnostics15010066