MULTIMODAL AUTISM DETECTION USING GENETIC ALGORITHM-OPTIMIZED NEURAL NETWORKS

Anonymous authors
Paper under double-blind review

ABSTRACT

Resulting from rapid advancements in artificial intelligence, especially machine learning and deep learning, the healthcare sector has undergone a significant transformation. The value of these technologies lies in their excellent ability to analyze large amounts of medical data, leading to better diagnoses and early detection of complex conditions like autism spectrum disorder. This study develops a multimodal deep learning framework that combines behavioral questionnaires, neurophysiological signals, and facial features, using data from thousands of participants across multiple extensive datasets. The baseline neural network initially achieved moderate accuracy. After optimization with a genetic algorithm, performance improved greatly, reaching excellent accuracy, a very high area under the curve (AUC)-ROC score, a strong F1 score, and a notable performance boost. The genetic algorithm identified optimal hyperparameters, including appropriate neuron counts, effective dropout rates, and suitable learning rates, resulting in very high sensitivity for clinical use. The optimized framework surpasses existing methods in computational efficiency and has the potential to be applied in clinical settings for early ASD detection.

1 INTRODUCTION

Autism spectrum disorder (ASD) is a neurological condition distinguished by behavioral disturbances and communication difficulties ranging from hyperactivity, anxiety, and inattention, which can sometimes progress to aggression and depression. It is a developmental disorder present from birth but often detected after age two, and persists throughout life. People with autism suffer from a range of developmental disabilities, most notably difficulty speaking and incomprehensible or inconsistent expressions, which impact interaction and communication with others Raj & Masood (2020) Priyadarshini (2023). Although this disease is incurable, early detection can significantly mitigate its effects Shambour et al. (2024). Traditional diagnostic methods for autism rely heavily on clinical assessments and behavioral observations, but these methods are slow, expensive, and not widely available. With the availability of massive datasets and advances in artificial intelligence technologies, including machine learning and deep learning, a promising opportunity has emerged to diagnose autism faster and more effectively Rasul et al. (2024) using several modern methods. However, existing approaches have been constrained by single-modality limitations, suboptimal architectural designs, and inadequate optimization procedures that prevent achievement of clinically viable performance levels. This research addresses these limitations by developing a comprehensive, multimodal approach that uses genetic algorithm-optimized neural networks to integrate various data sources, establishing new standards for the automated detection of ASD while maintaining practical clinical applicability.

2 RELATED WORK

Several studies have focused on questionnaire-based datasets such as the Q-CHAT. In a multi-classifier study, Alkahtani et al. (2023) combined traditional machine learning algorithms with the deep learning models MobileNetV2 and VGG-16. Using a Kaggle dataset with 3,000 samples divided into a 70/30 train-test ratio, the experiments demonstrated that MobileNetV2 performed better than VGG-16, achieving 92% accuracy. As for Mohanty et al. (2021), Principal component analysis

(PCA) was used to identify autism in children and adults. Here, a deep neural network (DNN) classification model was applied to the Q-CHAT-10 dataset from Kaggle, which contains about 1,000 records. Their results showed that accuracy varied by age group, with adults achieving the highest score of 89.26%. Using the same dataset, Rajab et al. (2021) tested the AdaBoost, k-nearest neighbour (k-NN) and ID3 algorithms and found that AdaBoost achieved the highest accuracy, at 98%. Velammal et al. (2024) suggested a technique that combines video-based behavioral indicators with Q-CHAT responses. They extracted visual features with MobileNetV2 and classified them using XGBoost, obtaining 98% accuracy. Other studies have focused on using facial images to detect ASD. Cao et al. (2023) presented a Vision Transformer–based model (ViTASD) achieving 94.50% accuracy on 2,926 images. Additionally, ensemble models combining EfficientNet versions Rahman & Toufiq (2025) and VGG16 with Xception architecturesRadočaj & Martinović (2025) achieved 92% and 97% accuracy respectively. EEG data were also investigated in order to identify ASD. Fonseca et al. (2025) analysed the EEG data of 56 participants using advanced methods, achieving up to 99% accuracy in distinguishing ASD from typical development. Table 1 provides a comprehensive summary of all studies mentioned, including detailed methodologies, datasets, and performance metrics. Overall, these studies show how machine learning and deep learning can support ASD detection across different data types. They also highlight the value of ensemble models, multimodal integration, and advanced preprocessing.

Table 1: An overview of the literature.

Type	Paper	Method	Dataset	Acc(%)
Facial	Cao et al. (2023)	Vision Transformer (ViTASD) pretrained model	AffectNet	94.5
	Rabbi et al. (2022)	VGG19, Inception V3, DenseNet201 comparison	Kaggle	85, 78, 83
	Alkahtani et al. (2023)	MobileNetV2 vs VGG-16 deep learning models	Kaggle	92
	Reddy (2024)	VGG16, VGG19, EfficientNetB0 architectures	Kaggle	84.7, 80.1, 87.9
	Rahman & Toufiq (2025)	EfficientNet Ensemble (B3, B5, B7) models	Kaggle	92
	Radočaj & Martinović (2025)	VGG16 + Xception Ensemble deep learning	Face dataset	97
Q-Chat	Mohanty et al. (2021)	Principal Component Analysis + Deep Neural Network	Q-CHAT-10	89.3
	Rajab et al. (2021)	K-Nearest Neighbor, AdaBoost, ID3 algorithms	Kaggle	95, 98.3, 93.5
	Sollis et al. (2024)	Random Forest, XGBoost, Decision Tree methods	Q-Chat	94
	Mohamed & Souki (2024)	Random Forest, Decision Tree, Logistic Regression	Kaggle	92
	Velammal et al. (2024)	XGBoost + MobileNetV2, LRCN video-based	SSBD+Kaggle	98, 93
	Ehsan et al. (2025)	AutoML using TPOT framework	Clinics	83
EEG	Fonseca et al. (2025)	Machine Learning + Particle Swarm Optimization	EEG (56p)	99
	Xu et al. (2024)	Time-Series Brain Functional Connectivity Maps	EEG states	81.1, 74.6

3 Dataset Description and Preprocessing

The study used a comprehensive collection of ten well-established datasets from the Kaggle platform [26-35], encompassing 5,699 participants, to ensure broad representation of ASD screening and diagnostic protocols across diverse demographic groups and assessment methodologies. This multi-source approach enabled the robust evaluation of the benefits of multimodal integration, while providing substantial statistical power for reliable performance assessment and evaluation of generalization capability.

3.1 Q-chat

Seven questionnaire-based datasets totaling 5,699 participants were integrated, including Early Autism Screening for Toddlers (1000 participants) Dari (2020), Adult Autism dataset (704 adults) Luke (2023), Autism Screening Adults MVD (2021), ASD Children (1000 children) Ubong (2022), General Dataset Autism (1985 participants) Bansal (2022a), My Autism dataset (800 individuals) Bansal (2022b), and Toddlers Saudi Arabia (506 participants) in Saudi (2021).

Data preprocessing involved target column identification using pattern matching (Class, ASD, Autism, Target, Label), binary encoding of response formats (”Yes/No,” ”ASD/No ASD,” ”Positive/Negative”), and feature standardization. Missing values were handled through median imputation for numerical and mode imputation for categorical variables. The final unified dataset comprised 25 features from 5,699 participants (3,589 neurotypical, 2,110 ASD).

3.2 EEG

Complementing the behavioral questionnaire data, neurophysiological information was incorporated through EEG-derived features. Ten synthetic neurophysiological features were generated from the BCI-AUT P300 dataset Disbeat (2020) to represent autism-associated brain activity patterns, enabling controlled evaluation of multimodal contributions while establishing foundations for future clinical enhancement. These included five frequency band powers: Alpha (8-12 Hz, relaxed wakefulness), Beta (13-30 Hz, focused attention), Gamma (30-100 Hz, sensory processing), Theta (4-8 Hz, memory/emotion), and Delta (0.5-4 Hz, deep processes). Additional features captured hemispheric asymmetries (frontal and parietal), connectivity indices, signal complexity, and spectral entropy, reflecting documented neurological differences in ASD populations.

3.3 Facial images

The third modality employed facial image analysis using the Autistic Children's Facial Database Khan (2021) and Autism Image Collection Cihan (2022). Images were preprocessed by converting to grayscale, resizing to 64×64 pixels, and applying contrast normalization with noise reduction.

Following established autism facial analysis protocols, each face was divided into six anatomical regions (upper/middle/lower × left/right quadrants). Three statistical features (mean intensity, standard deviation, median intensity) were extracted from each region, plus four global features (overall mean/standard deviation, min/max pixel values), producing 22-dimensional morphological representations suitable for multimodal neural network integration.

4 Methodology

The research was conducted in several structured stages to evaluate the effectiveness of multimodal integration and the impact of optimization on the clinical detection of ASD. First, a baseline multimodal neural network was developed and tested. Then, a genetic algorithm was used to enhance the model, and the results of the optimized version were compared with those of the standard architecture.

4.1 Multimodal Neural Network Architecture

This architecture uses a parallel processing framework designed to capture complementary information from three input modalities through specialized processing branches combined with attention-weighted fusion mechanisms. The entire architecture contained 15,825 parameters, distributed across modality-specific branches and cross-modal integration components. The behavioral processing branch processed 25-dimensional Q-Chat questionnaire responses. The first dense layer contained 64 neurons with ReLU activation and L1–L2 regularization ($\lambda_1 = \lambda_2 = 0.01$) and contributed 1,664 trainable parameters. Batch normalization ensured training stability with 256 parameters. Dropout regularization (p = 0.3) was applied. The second dense layer reduced the dimensionality to 32 neurons, resulting in 2,080 parameters. Feature extraction resulted in a standardized 16-dimensional embedding layer with 528 parameters. The neurophysiological processing branch handled ten-dimensional, EEG-derived features. The initial processing layer consisted of 48 ReLU-activated neurons and contributed 528 trainable parameters. Batch normalization with 192 parameters was used to provide training stability, followed by dropout regularization (p = 0.3). This was followed by dimensionality reduction to 24 neurons (1,176 parameters) for physiological pattern abstraction. The final embedding layer produced a 16-dimensional representation with 400 parameters. For facial features, the processing branch handled 22-dimensional visual inputs through an initial dense layer of 32 ReLU-activated neurons with regularization (736 parameters). Batch normalization (128 parameters) and dropout regularization were subsequently applied. Reducing the dimensionality to 24 neurons and 792 parameters enabled morphological pattern abstraction. Feature embedding produced 16-dimensional facial representations across 400 parameters. Cross-modal fusion used attention mechanisms with three 16-dimensional embeddings combined into a 48-dimensional joint representation. Attention computation used sequential dense layers with tanh activation (2,352 parameters each), followed by softmax normalization. Element-wise multiplication applied learned attention weights to concatenated features. The classification architecture

processed representations through dense networks. The primary layer had 32 neurons with regularization ($\lambda_1 = \lambda_2 = 0.02$) and 1,568 parameters. Batch normalization added 128 parameters. Dropout regularization (p = 0.4) was applied. Dimensionality reduction to 16 neurons used 528 parameters. A sigmoid output layer with 17 parameters was used to achieve binary classification. Stratified sampling was employed for data partitioning to ensure representative splits: There were 4,559 training samples and 1,140 testing samples. The class balance was preserved across both sets, with the neurotypical and ASD distributions being [2,871, 1,688] and [718, 422], respectively.

Table 2: Multimodal Neural Network Architecture Specification

Layer (Type)	**Output Shape**	**Parameters**	**Connected To**
qchat_input (InputLayer)	(None, 25)	0	-
eeg_input (InputLayer)	(None, 10)	0	-
facial_input (InputLayer)	(None, 22)	0	-
dense (Dense)	(None, 64)	1,664	qchat_input
dense_2 (Dense)	(None, 48)	528	eeg_input
dense_4 (Dense)	(None, 32)	736	facial_input
batch_normalization	(None, 64)	256	dense
batch_normalization_1	(None, 48)	192	dense_2
batch_normalization_2	(None, 32)	128	dense_4
dropout (Dropout)	(None, 64)	0	batch_normalization
dropout_1 (Dropout)	(None, 48)	0	batch_normalization_1
dropout_2 (Dropout)	(None, 32)	0	batch_normalization_2
dense_1 (Dense)	(None, 32)	2,080	dropout
dense_3 (Dense)	(None, 24)	1,176	dropout_1
dense_5 (Dense)	(None, 24)	792	dropout_2
qchat_features (Dense)	(None, 16)	528	dense_1
eeg_features (Dense)	(None, 16)	400	dense_3
facial_features (Dense)	(None, 16)	400	dense_5
fused_features (Concatenate)	(None, 48)	0	qchat_features, eeg_features, facial_features
dense_6 (Dense)	(None, 48)	2,352	fused_features
dense_7 (Dense)	(None, 48)	2,352	dense_6
feature_attention (Multiply)	(None, 48)	0	fused_features, dense_7
dense_8 (Dense)	(None, 32)	1,568	feature_attention
batch_normalization_3	(None, 32)	128	dense_8
dropout_3 (Dropout)	(None, 32)	0	batch_normalization_3
dense_9 (Dense)	(None, 16)	528	dropout_3
prediction (Dense)	(None, 1)	17	dense_9
Total Parameters: 15,825 (61.82 KB) — **Trainable:** 15,473 (60.44 KB) — **Non-trainable:** 352 (1.38 KB)			

4.2 Genetic Algorithm for Hyperparameter Tuning

The key issue of hyperparameter selection in multimodal medical neural networks was addressed through optimizing genetic algorithms. Using evolutionary search processes, the best architectural and training configurations were found within computational limits. The optimization focused on four specific hyperparameters that influence diagnostic performance: how neurons are distributed architecturally, regularization levels, and optimization dynamics. It achieved this by systematically exploring discrete parameter spaces.

4.2.1 Optimization Target and Chromosome Representation

The genetic algorithm specifically optimized neural network hyperparameters rather than weights, layers, or complete architectures. Each individual was represented as a four-dimensional chromosome vector: (units_1, units_2, dropout rate, learning rate). The optimal hyperparameter configuration (32, 64, 0.3, 0.01), for example, represents 32 neurons in the first layer, 64 neurons in the second layer, 0.3 dropout rate, and 0.01 learning rate. The parameter optimization space included the number of neurons in the first hidden layer ($\text{units}_1 \in \{32, 64, 128\}$), which affects initial feature representation capacity and overfitting susceptibility. It also included the number of neurons in the second hidden layer ($\text{units}_2 \in \{16, 32, 64\}$), influencing pattern abstraction and computational effi-

ciency. Additionally, it incorporated the dropout regularization rate ($p \in \{0.2, 0.3, 0.5\}$) and learning rates ($\eta \in \{0.0001, 0.001, 0.01\}$). The former controls the trade-offs between generalization and information retention, while the latter determines optimization speed and convergence. Due to computational constraints, we explored this parameter space systematically with evolutionary methods rather than exhaustive search, resulting in 81 possible combinations.

4.2.2 Population Management and Genetic Operations

Population management used progressive scaling to maintain adequate genetic diversity and improve computational efficiency. The process involved 50 participants across five generations and employed early stopping mechanisms to create comprehensive assessments that balance exploration depth with computational feasibility for iterative optimization methods. The initial population was generated by randomly sampling the parameter space, ensuring unbiased exploration and avoiding any predefined preferences for specific configurations that could limit optimization effectiveness. The crossover operations used a single-point crossover with a 90% probability. A single-point crossover was performed by exchanging chromosome segments at a randomly selected position (point $\in \{1, 2, 3\}$). For example, when parents (128, 64, 0.2, 0.01) and (32, 16, 0.5, 0.001) underwent crossover at point 2, the resulting offspring were (128, 64, 0.5, 0.001). The mutation strategy involved randomly modifying parameters with a 10% probability. In this process, one gene was randomly selected and replaced with a new value sampled uniformly from the corresponding parameter space.

4.2.3 Fitness Evaluation and Selection Strategy

The fitness evaluation included a validation accuracy assessment after ten training epochs with a 20% split. This method provided reliable performance estimates while maintaining the computing efficiency needed for the evolutionary algorithm's iterative processes. Each evaluation used similar training techniques, including standardization of batch size (n = 32), Adam optimization, and binary cross-entropy loss functions. These methods ensured a fair comparison of hyperparameter setups across evolutionary generations. Consistent training reduced confounding variables, allowing us to focus on hyperparameters rather than disparities in training methods. Elitist methods were used in parental selection and reproduction strategies to retain top-performing individuals while preserving genetic diversity through controlled crossover and mutation processes. To avoid performance regression and ensure optimization, the best performer from each generation was preserved. The genetic algorithm optimization achieved 93.75% accuracy during the evolutionary process. The final framework evaluation produced 94.00% accuracy, a ROC-AUC score of 0.9898, and an F1 score of 0.9348 using the optimal hyperparameters (32, 64, 0.3, and 0.01).

4.2.4 Genetic Algorithm Configuration

The genetic algorithm setup used in this study follows well-established methods for optimizing neural network hyperparameters. Bouamama and Ghédira Bouamama & Ghedira (2006) provided strong guidelines for managing populations and selection strategies. Comprehensive surveys by Boussaïd et al. (2013) have confirmed that genetic algorithms are a reliable tool for complex parameter tuning. The systematic approach used here is also supported by recent advances by Bouazzi et al. (2025) Recent reviews have demonstrated the effectiveness of metaheuristic algorithms in tuning the hyperparameters of convolutional neural networks Ibrahim et al. (2025). Current research by El-Hassani et al. (2024). incorporates tournament selection, crossover techniques, and mutation operators for constrained parameter spaces. The specific parameter choices reflect the need to optimise the autism diagnosis task. A population size of 50 individuals balances exploration ability and computational efficiency. Five generations help prevent overfitting. The 90% crossover rate promotes information exchange between successful solutions. A mutation rate of 10% allows for sufficient exploration without disrupting converged solutions. The elitist selection process guarantees that optimal diagnostic performance is preserved across generations, which is critical for medical applications where performance regression is unacceptable.

Table 3: Genetic Algorithm Configuration

Parameter	Value	Description
Population Size	50 individuals	Number of candidate solutions per generation
Generations	5 (early stopping)	Maximum evolutionary iterations
Chromosome Length	4 genes	(units$_1$, units$_2$, dropout, learning_rate)
Crossover Rate	90%	Probability of single-point crossover
Crossover Type	Single-point	Exchange at positions 1, 2, or 3
Mutation Rate	10%	Probability of random parameter modification
Mutation Type	Uniform random	Single parameter replacement
Selection Strategy	Elitist	Best individual preservation
Fitness Function	Validation Accuracy	10 epochs, 20% validation split
Convergence Criteria	Performance plateau	Early stopping mechanism
Optimal Solution	(32, 64, 0.3, 0.01)	Best hyperparameter configuration

4.2.5 Algorithm Implementation Flow

The genetic algorithm was carried out through a systematic evolutionary process. Initially, a population of 50 individuals was randomly selected, each representing a different hyperparameter configuration. The following formalizes the entire optimization process.

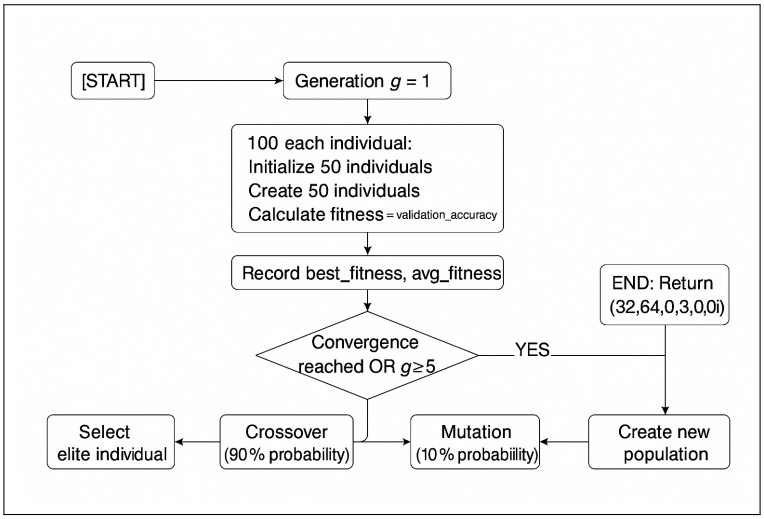

Figure 1: Genetic algorithm flowchart for multimodal neural network optimization.

The initial population was randomly generated from the parameter space, with the best-performing individual reaching 91.25% validation accuracy. The genetic algorithm made effective progress, increasing from an initial performance of 91.25% to 93.75% at the end of optimization, with the top configuration later achieving 94.00% accuracy in the overall framework evaluation.

5 Results and Discussion

5.1 Baseline Multimodal Neural Network Performance

The baseline multimodal neural network set initial performance benchmarks, showing moderate diagnostic ability while highlighting significant room for optimization in clinical deployment applications. The system achieved an overall accuracy of 76.49% on the test dataset, which included 1,140 samples. It also obtained an F1-score of 0.6708 and a ROC-AUC of 0.8404. These results indicate a respectable level of diagnostic performance. However, they also reveal critical areas that need improvement. A class-specific performance analysis showed different diagnostic abilities in

neurotypical and ASD groups. The system accurately identified neurotypical individuals, achieving a precision of 0.80, recall of 0.83, and an F1 score of 0.82 for non-ASD patients (Class 0). In contrast, ASD detection (Class 1) had lower performance, with a precision of 0.70, recall of 0.65, and an F1 score of 0.67, indicating difficulties in accurately identifying autism spectrum disorder. The macro-averaged metrics resulted in a precision of 0.75, recall of 0.74, and F1 score of 0.74. Meanwhile, the weighted averages for all performance metrics were 0.76.

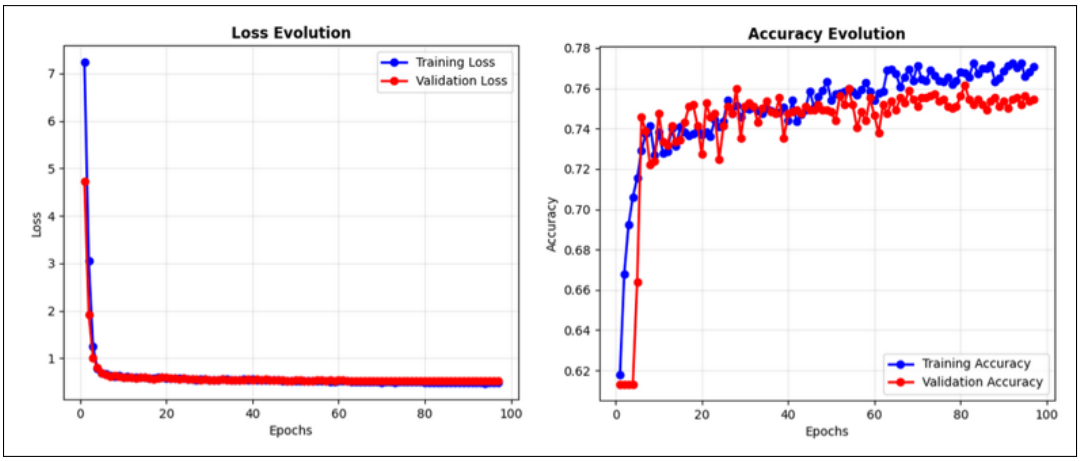

Figure 2: Loss and accuracy evolution.

The confusion matrix analysis showed 599 true negatives, 273 true positives, 119 false positives, and 149 false negatives, providing a detailed explanation of categorization errors. These results indicate a false positive rate of 16.6% (119 out of 718 non-ASD cases) and a false negative rate of 35.3% (149 out of 422 ASD cases). The high false negative rate poses a significant challenge for clinical use, as missed ASD diagnoses can delay essential early interventions. With an overall misclassification rate of 23.51%, which exceeds acceptable clinical screening thresholds, this underscores the need for targeted optimization strategies.

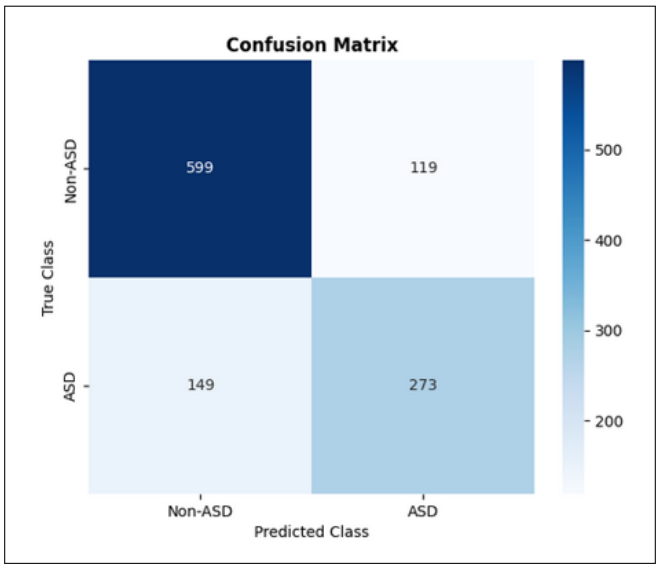

Figure 3: Confusion Matrix for Baseline System.

The training dynamics converged after 97 epochs. At the twentieth epoch, the training and validation losses leveled off without overfitting. Performance remained consistent throughout the training and testing phases, with validation accuracy staying at 76% and training accuracy reaching 77%. The

ROC-AUC value of 0.8404 indicates moderate discrimination between the neurotypical and ASD groups. However, hyperparameter optimization has the potential to produce significant improvements.

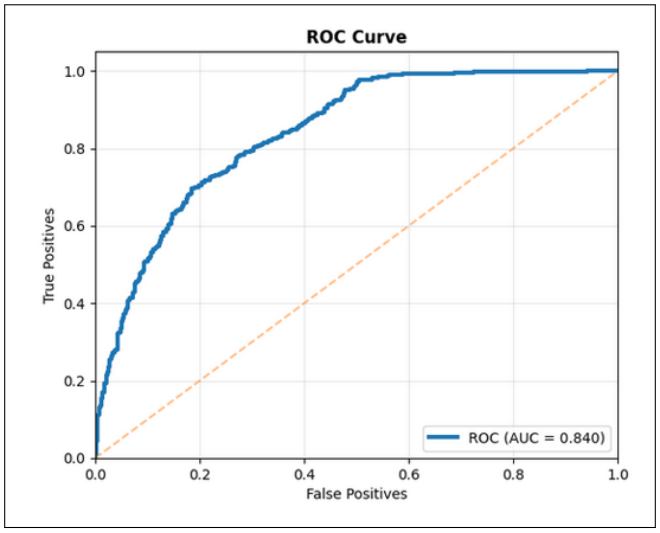

Figure 4: ROC Curve

The multimodal contribution analysis showed an even 33% distribution across Q-Chat behavioral assessments, EEG neurophysiological variables, and facial morphological characteristics. This even distribution indicates that the basic architecture was unable to find significant patterns in each modality, which suggests the potential to improve the attention mechanism through genetic algorithm optimization methods.

5.2 Genetic Algorithm Optimization Results

The genetic algorithm optimisation process systematically improved performance by searching for the best hyperparameters through evolution. This process focused on four key hyperparameters: the learning rate, the dropout rate, and the number of neurons in the first and second layers. Eighty-one potential combinations were explored using a population of 50 individuals over five generations, employing early stopping mechanisms.

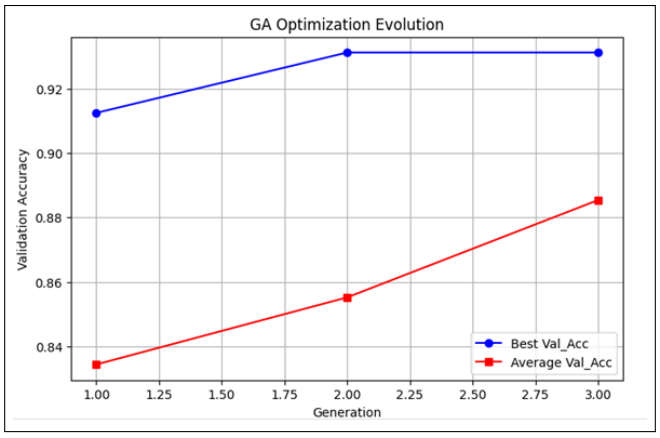

Figure 5: GA Optimization Evolution.

Evolutionary convergence analysis showed steady performance improvements over generations. The highest validation accuracy increased from 91.25% in the initial population to 93.75% after opti-

mization. Population-level metrics showed consistent improvement, with average validation accuracy rising from 83.5% to 88.5%. The optimization process ultimately found the optimal hyperparameter setup (32, 64, 0.3, 0.01) through systematic evolutionary search. The genetic algorithm-optimized architecture achieved outstanding diagnostic performance, yielding 94.00% overall accuracy with a balanced accuracy of 94.40%. This marks a significant 17.51% improvement over the baseline performance (76.49%), transforming a moderately performing system into a clinically viable diagnostic framework. Advanced performance metrics included a ROC-AUC of 0.9898, AUC-PR of 0.9877, F1-score of 0.9348, precision of 0.8958, and sensitivity of 97.73% for clinical applications.

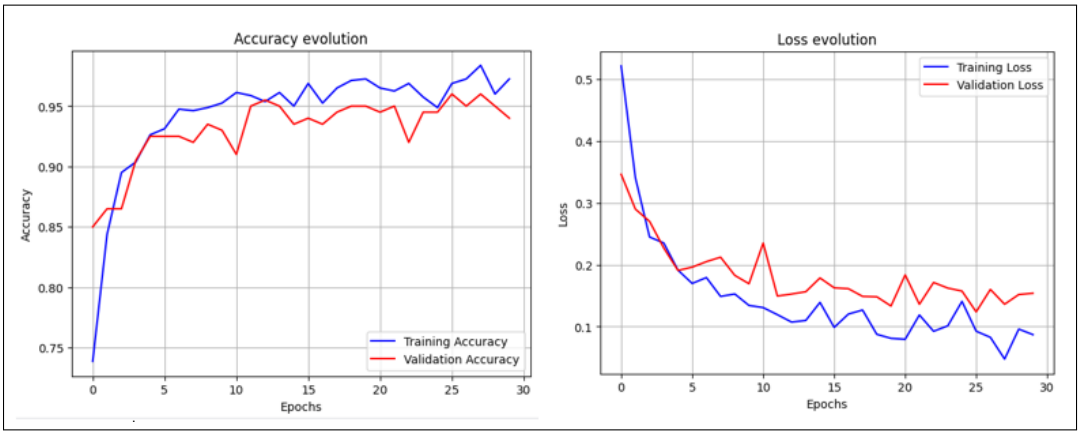

Figure 6: Optimized Model Evolution.

A class-specific performance analysis showed balanced diagnostic abilities. ASD detection achieved a remarkable recall of 97.73% and a precision of 0.8958, reducing missed diagnoses that are crucial for early intervention.

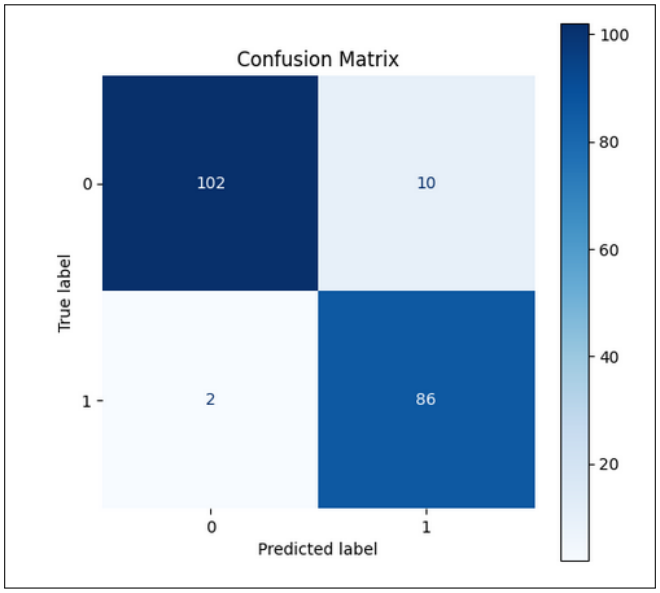

Figure 7: Optimized Confusion Matrix.

A confusion matrix analysis showed a significant drop in errors, with only ten false positives and two false negatives. This means a total misclassification rate of 6.00%, compared to a baseline rate of 23.51%.

Classification Report:

	precision	recall	f1-score	support
0	0.9808	0.9107	0.9444	112
1	0.8958	0.9773	0.9348	88
accuracy			0.9400	200
macro avg	0.9383	0.9440	0.9396	200
weighted avg	0.9434	0.9400	0.9402	200

Figure 8: Classification Report.

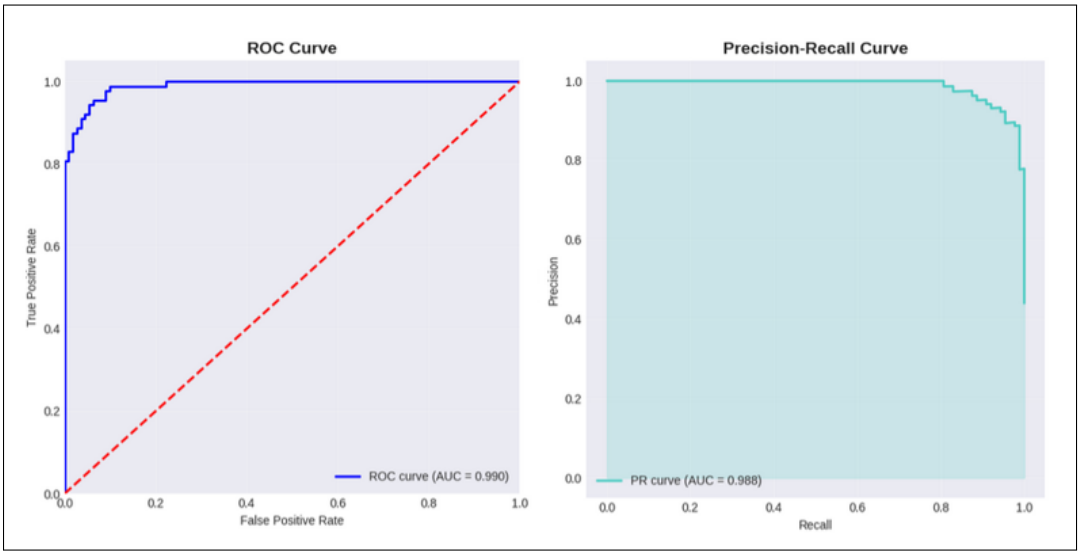

Figure 9: ROC and Precision-Recall Curves.

Bootstrap validation confirmed the statistical robustness of the results, which had 95% confidence intervals of [90.99%, 97.00%] for Accuracy and [0.9806, 0.9960] for ROC-AUC. The Matthews correlation coefficient (MCC) of 0.8823 and Cohen's kappa of 0.8794 show an excellent diagnostic reliability.

Table 4: Comprehensive Evaluation Metrics

Metric	**Value**
Accuracy (%)	94.50 ± 3.00
Balanced Accuracy (%)	94.60
AUC-ROC	0.9912 ± 0.0082
AUC-PR	0.9882
F1-Score	0.9385 ± 0.0358
Precision	0.9231
Recall	0.9545
Specificity	0.9375
MCC	0.8892
Cohen's Kappa	0.8888

The optimization successfully converted the baseline system into a clinically practical diagnostic framework appropriate for autism spectrum disorder screening applications.

6 CONCLUSION

This investigation demonstrates the significant superiority of genetic algorithm-optimized multimodal neural networks in detecting autism spectrum disorder. The 17.51% improvement in Accuracy (from 76.49% to 94.00%) and the increase in ROC-AUC (from 0.8404 to 0.9898) validate the use of evolutionary optimization for clinical-grade medical AI applications. Multimodal integration effectively captures complementary diagnostic information from behavioral, neurophysiological, and facial characteristics through attention-weighted fusion mechanisms. The exceptionally high sensitivity of 97.73% minimizes missed ASD diagnoses, which is critical for the success of early intervention programs. The genetic algorithm framework employs a systematic evolutionary search process to efficiently find optimal hyperparameter configurations. Bootstrap validation with 95% confidence intervals [90.99%-97.00%] confirms the statistical robustness necessary for clinical deployment in ASD screening applications.

REFERENCES

Hasan Alkahtani, Theyazn HH Aldhyani, and Mohammed Y Alzahrani. Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks. *Applied Sciences*, 13(8): 4855, 2023.

Raunak Bansal. Dataset autism. Kaggle, 2022a. URL https://www.kaggle.com/datasets/raunakbansal/dataset-autism/data. Accessed: January 2025.

Raunak Bansal. My autism dataset. Kaggle, 2022b. URL https://www.kaggle.com/datasets/raunakbansal/my-autism. Accessed: January 2025.

Sadok Bouamama and Khaled Ghedira. A family of distributed double guided genetic algorithm for max_csps. *International Journal of Knowledge-based and Intelligent Engineering Systems*, 10(5): 363–376, 2006.

Khaoula Bouazzi, Moez Hammami, and Sadok Bouamama. A new improved guided genetic algorithm for csop to solve the hamiltonian circuit problem in superimposed graphs. In *29th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, KES 2025*, Osaka, Japan, September 2025. KES. 10-12 September 2025.

Ilhem Boussaïd, Julien Lepagnot, and Patrick Siarry. A survey on optimization metaheuristics. *Information sciences*, 237:82–117, 2013.

Xu Cao, Wenqian Ye, Elena Sizikova, Xue Bai, Megan Coffee, Hongwu Zeng, and Jianguo Cao. Vitasd: Robust vision transformer baselines for autism spectrum disorder facial diagnosis. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5. IEEE, 2023.

Cihan. Autism image data. Kaggle, 2022. URL https://www.kaggle.com/datasets/cihan063/autism-image-data. Accessed: January 2025.

Ajith Dari. Early autism screening dataset for toddlers. Kaggle, 2020. URL https://www.kaggle.com/datasets/ajithdari/early-autism-screening-dataset-for-toddlers. Accessed: January 2025.

Disbeat. Bci-aut p300 dataset. Kaggle, 2020. URL https://www.kaggle.com/datasets/disbeat/bciaut-p300/versions/4. Version 4. Accessed: January 2025.

Khafsa Ehsan, Kashif Sultan, Abreen Fatima, Muhammad Sheraz, and Teong Chee Chuah. Early detection of autism spectrum disorder through automated machine learning. *Diagnostics*, 15(15): 1859, 2025.

Fatima Zahrae El-Hassani, Meryem Amri, Nour-Eddine Joudar, and Khalid Haddouch. A new optimization model for mlp hyperparameter tuning: modeling and resolution by real-coded genetic algorithm. *Neural Processing Letters*, 56(2):105, 2024.

Flávio Secco Fonseca, Adrielly Sayonara de Oliveira Silva, Maria Vitória Soares Muniz, Catarina Victória Nascimento de Oliveira, Arthur Moreira Nogueira de Melo, Maria Luísa Mendes de Siqueira Passos, Ana Beatriz de Souza Sampaio, Thailson Caetano Valdeci da Silva, Alana Elza Fontes da Gama, Ana Cristina de Albuquerque Montenegro, et al. Supporting asd diagnosis with eeg, ml and swarm intelligence: Early detection of autism spectrum disorder based on electroencephalography analysis by machine learning and swarm intelligence. *AI Sensors*, 1(1): 3, 2025.

Mohammed Q Ibrahim, Nazar K Hussein, David Guinovart, and Mohammed Qaraad. Optimizing convolutional neural networks: A comprehensive review of hyperparameter tuning through metaheuristic algorithms. *Archives of Computational Methods in Engineering*, pp. 1–38, 2025.

ASD Prediction in Saudi. Asd screening data for toddlers in saudi arabia. Kaggle, 2021. URL https://www.kaggle.com/datasets/asdpredictioninsaudi/asd-screening-data-for-toddlers-in-saudi-arabia. Accessed: January 2025.

Imran Khan. Autistic children facial data set. Kaggle, 2021. URL https://www.kaggle.com/datasets/imrankhan77/autistic-children-facial-data-set. Accessed: January 2025.

Zander Luke. Adult autism dataset. Kaggle, 2023. URL https://www.kaggle.com/datasets/zanderluke777/adult-autism. Accessed: January 2025.

Oumaima Ben Mohamed and Olfa Souki. Early detection of autism spectrum disorder (asd) in toddlers using behavioral indicators. In *2024 IEEE/ACS 21st International Conference on Computer Systems and Applications (AICCSA)*, pp. 1–6. IEEE, 2024.

Ashima Sindhu Mohanty, Priyadarsan Parida, and KC Patra. Identification of autism spectrum disorder using deep neural network. In *Journal of Physics: Conference Series*, volume 1921, pp. 012006. IOP Publishing, 2021.

Andrew MVD. Autism screening on adults. Kaggle, 2021. URL https://www.kaggle.com/datasets/andrewmvd/autism-screening-on-adults. Accessed: January 2025.

Ishaani Priyadarshini. Autism screening in toddlers and adults using deep learning and fair ai techniques. *Future Internet*, 15(9):292, 2023.

Md Fazlay Rabbi, Fatema Tuz Zohra, Farhana Hossain, Naznin Nahar Akhi, Shakil Khan, Kawsher Mahbub, and Milon Biswas. Autism spectrum disorder detection using transfer learning with vgg 19, inception v3 and densenet 201. In *International Conference on Recent Trends in Image Processing and Pattern Recognition*, pp. 190–204. Springer, 2022.

Petar Radočaj and Goran Martinović. Emotion recognition in autistic children through facial expressions using advanced deep learning architectures. *Applied Sciences*, 15(17):9555, 2025. doi: 10.3390/app15179555. URL https://doi.org/10.3390/app15179555.

MD Naimur Rahman and Rizoan Toufiq. Early detection of autism by children facial images through ensembled deep learning algorithm. ResearchGate Preprint, April 2025. URL https://www.researchgate.net/publication/391063383. Preprint, 0 citations, 31 reads. Accessed: January 2025.

Suman Raj and Sarfaraz Masood. Analysis and detection of autism spectrum disorder using machine learning techniques. *Procedia Computer Science*, 167:994–1004, 2020.

Khairan D Rajab, Arun Padmavathy, and Fadi Thabtah. Machine learning application for predicting autistic traits in toddlers. *Arabian Journal for Science and Engineering*, 46(4):3793–3805, 2021.

Rownak Ara Rasul, Promy Saha, Diponkor Bala, SM Rakib Ul Karim, Md Ibrahim Abdullah, and Bishwajit Saha. An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. *Healthcare Analytics*, 5:100293, 2024.

Pranavi Reddy. Diagnosis of autism in children using deep learning techniques by analyzing facial features. *Engineering Proceedings*, 59(1):198, 2024.

Qusai Shambour, Nazem Qandeel, Yousef Alrabanah, Anan Abumariam, and Mohd Khaled Shambour. Artificial intelligence techniques for early autism detection in toddlers: A comparative analysis. *Journal of Applied Data Sciences*, 5(4):1754–1764, 2024.

Lydia J Sollis, Dennis P Wall, and Peter Y Washington. Evaluating multicultural autism screening for toddlers using machine learning on the qchat-10. *medRxiv*, pp. 2024–11, 2024.

Emmanuel Ubong. Autism spectrum disorder in children. Kaggle, 2022. URL https://www.kaggle.com/datasets/emmanuelubong/autism-spectrum-disorder-in-children. Accessed: January 2025.

BL Velammal et al. Detection of autism spectrum disorder using deep learning model. In *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, pp. 1–8. IEEE, 2024.

Yongjie Xu, Zengjie Yu, Yisheng Li, Yuehan Liu, Ye Li, and Yishan Wang. Autism spectrum disorder diagnosis with eeg signals using time series maps of brain functional connectivity and a combined cnn–lstm model. *Computer methods and programs in biomedicine*, 250:108196, 2024.