Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/200405

A Data Mining and Artificial Neural Network Approach for Autism Spectrum Disorder Detection

Hamza Abu Owida[1], *Mwaffaq Abu Alhaija*[2], *Asokan Vasudevan*[3], *Hamza A. Mashagba*[4], *Azlan B. Abd Aziz*[4,*], *Mohammad Faleh Ahmmad Hunitie*[5], *Nor Azlina Ab. Aziz*[4], *and Suleiman Ibrahim Mohammad*[6,7]

[1]Medical Engineering Department, Faculty of Engineering, Al-Ahliyya Amman University, Amman 19111, Jordan
[2]Department of Cybersecurity and Cloud Computing, Faculty of Information Technology, Applied Science Private University, Amman 11931, Jordan
[3]Faculty of Business and Communications, INTI International University, Negeri Sembilan 71800, Malaysia
[4]Faculty of Engineering and Technology, Multimedia University, Centre for Wireless Technology (CWT), Cyberjaya, Selangor 63100, Malaysia
[5]Department of Public Administration, School of Business, University of Jordan, Amman 11942, Jordan
[6]Department of Business Administration, School of Business, Al al-Bayt University, Mafraq 25113, Jordan
[7]INTI International University, Negeri Sembilan 71800, Malaysia

Received: 6 Jan. 2026, Revised: 24 Mar. 2026, Accepted: 2 Jun. 2026
Published online: 1 Jul. 2026

Abstract: In this paper, we examine how deep learning—especially facial-image–based methods—shows promise for ASD detection. However, many models are limited by centralized training, which restricts the use of diverse, globally distributed datasets and can reduce performance across demographic groups. To address these limitations, we present a privacy-preserving federated learning framework that enables medical institutions worldwide to collaboratively train ASD detection models without sharing raw patient data beyond local sites. Our framework incorporates differential privacy, secure aggregation protocols, and adaptive communication strategies to support convergence and protect confidentiality under heterogeneous data distributions. Experiments using simulated data from five medical centers demonstrate that the federated model matches centralized baselines in predictive performance while improving demographic representation and reducing bias across ethnic and cultural groups. Across diverse populations, it achieves an average accuracy of 89.2%, precision of 88.7%, recall of 89.5%, and F1-score of 89.1%, underscoring the potential for fair, confidential global diagnostic technologies for neurodevelopmental disorders.

Keywords: Artificial neural networks; Data mining; Data processing; Big data analytics; Facial image analysis

1 Introduction

One of the most difficult-to-diagnose conditions in modern medicine is Autism Spectrum Disorder (ASD), defined by persistent difficulties in social interaction and communication, along with restricted, repetitive patterns of behavior, interest, or activity [1,2]. Over the last twenty years, the number of children diagnosed with ASD has grown exponentially; it is now estimated that one child out of every fifty-four receives an ASD diagnosis [3, 4]. The growth in ASD diagnoses has occurred due to increased awareness and improvements in diagnosing capabilities, and emphasizes the necessity for more efficient, accessible, and culturally-sensitive diagnostic tools. Traditionally, the ASD diagnosis was done by clinicians using behavioral observation and/or interviews with parents or caregivers and the use of standardized diagnostic tools, i.e., the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview – Revised (ADI-R) [5,6]. Despite the use of the above approaches, there are many challenges that exist, e.g., large inter-rater reliability differences between clinicians, culturally-biased assessment tools, and a shortage of trained diagnosticians, especially in low- and middle-income countries [7,8]. Due to the absence of biomarkers that can objectively identify ASD, delays in diagnosis and treatment occur frequently; these delays

* Corresponding author e-mail: azlan.abdaziz@mmu.edu.my

negatively affect the long-term developmental prognosis of affected children [9, 10].

In recent years, the advancements in computer vision and deep learning have presented new opportunities to provide objective assessments of ASD through the analysis of morphological characteristics of the face and behavioral patterns of individuals [11]. Recent research indicates that individuals with ASD present unique, identifiable differences in the structure and expressions of their faces that can be detected through complex machine learning algorithms [12]. Research conducted in [13] demonstrated that convolutional neural networks could detect ASD with high accuracy (approximately 85–95%) through facial images under controlled conditions [14].

While there are some positive developments in using AI to identify individuals with ASD, there are many significant limitations that limit its applicability as a clinical tool on a global scale. For instance, most current research studies investigating AI for the identification of ASD are utilizing relatively small and regionally-restricted datasets that lack enough variability to develop reliable and generalizable models [15]. Training data may exhibit demographic homogeneity leading to systemic biases when applying models to populations outside of those utilized in model development [16] – especially given the fact that individuals with ASD present differently across different ethnic, cultural, and socioeconomic groups.

Creating datasets that represent diverse demographics for the purpose of developing robust, universally-applicable AI models for ASD detection is complicated by significant restrictions placed by both regulatory requirements and medical data sharing laws. Healthcare institutions are subject to privacy regulations such as HIPAA in the United States, GDPR in Europe, and similar frameworks around the world [17]; these regulations protect patient confidentiality but greatly hinder data sharing required to develop robust AI models that are capable of performing well globally. Due to institutional policies and concerns about data sovereignty, traditional methods of centralizing data for use in developing robust AI models are generally unfeasible or unachievable.

Federated learning is emerging as a viable method to overcome the challenges of developing robust AI models for detecting ASD. Federated learning allows for collaborative training of models across multiple healthcare institutions without the need for sharing patient data [18]. In federated learning, each participating healthcare institution trains a local version of the model using their own proprietary data; only the model's parameters or gradients are exchanged between the institutions [19]. This approach protects the privacy of patient data and provides healthcare institutions with control over their data while allowing for the creation of models that leverage the collective knowledge contained within multiple institutions' data sets.

Research into federated learning for medical imaging and diagnostic applications has become increasingly popular in recent years, with success stories in radiology, pathology, and drug discovery [20]. However, developing federated learning solutions for ASD detection requires addressing challenges that are unique to ASD detection, i.e., providing models that are capable of representing diverse demographics, identifying phenotypes that may be subtly expressed, and dealing with heterogeneity of clinical presentation [21]. To date, few federated learning frameworks have attempted to provide solutions to these challenges.

This paper describes a comprehensive federated learning framework for privacy-preserving, collaborative ASD detection using facial images, specifically developed to address these gaps. The proposed framework includes three major innovations: adaptive communication protocols that account for variable data quality and quantity across institutions, differential privacy mechanisms that are tailored to medical imaging applications, and fairness-aware training objectives that focus on reducing demographic bias and increasing representation.

Primary contributions of this research include the design of a novel federated learning architecture that maintains model performance while providing a strong guarantee of privacy, the implementation of techniques to mitigate bias that improves model fairness across diverse populations, and demonstrating the potential of federated learning to create more equitable and globally applicable diagnostic tools for neurodevelopmental disorders. We demonstrate the effectiveness of our federated learning framework through comprehensive simulations involving multi-institutional collaborations and show that federated approaches can match centralized training performance while greatly improving demographic representation and reducing the systematic biases inherent in current ASD detection systems.

2 Related Work

2.1 Machine Learning Methods Used for Detection of Autism Spectrum Disorder

Over the last decade, research into the application of machine learning for detecting Autism Spectrum Disorder (ASD) has grown exponentially due to advancements in computer vision, natural language processing, and multimodal data fusion techniques [22]. In early work, researchers primarily focused on analyzing behavior and physiological signals, including eye tracking data, movement patterns of the body, and voice characteristics as possible biomarkers for ASD [23].

Deep learning has greatly increased the possibilities of machine learning methods for detecting ASD, specifically through the use of facial images and video.

Researchers such as [24] have shown the possibility of using Convolutional Neural Networks (CNNs) to analyze facial features for signs of ASD and reported accuracy levels of approximately 89% based upon a dataset of 1040 children. That study provided evidence of facial morphology being possible noninvasive biomarkers for screening for ASD; however, there were many limitations to the study's ability to generalize based on the low number of children and lack of diversity in the dataset.

Recent studies have demonstrated new, advanced methods of facial analysis for detecting ASD. The researchers in [25] presented a multi-scale CNN architecture to detect facial images at multiple scales to allow for hierarchical feature extraction and reported higher performance than traditional methods of facial analysis for detecting ASD. Additionally, they found that their method was better able to detect the subtle morphological differences between children with ASD and typically developing children, which are often difficult for humans to detect.

The researchers in [26] proposed a novel attention-based deep learning model to focus on specific facial regions most indicative of signs of ASD and demonstrated superior performance compared to all previous benchmarking studies.

Another area of investigation is combining multiple data modalities to increase the accuracy and reliability of detecting ASD. The paper in [27] used a multimodal approach to combine facial image analysis with eye-tracking data and behavioral assessments and demonstrated that the combination of the three data types resulted in significantly higher diagnostic accuracy when compared to either modality alone. The study was conducted with an overall accuracy of 93.2%, but the high cost and complexity of collecting the required data limits the practicality of applying their methodology.

In spite of the rapid development of machine learning methods for detecting ASD, current approaches still contain critical limitations that limit their clinical utility and global applicability. Many of the studies that utilize machine learning methods for detecting ASD are based on small datasets collected from single locations or geographic areas, resulting in a lack of confidence in the ability of trained models to apply to a wider population [28]. In addition, the homogeneous demographics of the training datasets can cause biased results leading to poor performance for under-represented groups, which could further exacerbate pre-existing inequalities in accessing quality care for individuals diagnosed with ASD.

2.2 Federated Learning in Healthcare

Federated Learning (FL) is a new way to collaborate on Machine Learning (ML) projects in a healthcare setting, while preserving the confidentiality of the participants' information and respecting their legal and regulatory requirements for information sharing. In Federated Learning, each participant trains a local ML model independently on its own data and shares with all the others only the gradients or the updated model weights (parameters). There are many reasons why Federated Learning is so attractive in healthcare. First, healthcare data is extremely sensitive and protected by very strict laws and regulations that prevent data sharing. Second, there are many healthcare applications where larger amounts of diverse data are needed to describe the entire range of human variation and diseases.

There are many successful examples of the application of Federated Learning to Medical Imaging, which demonstrate the viability and advantages of Federated Learning. For example, [29] proposed a Federated Learning framework for Brain Tumor Segmentation. This was done with 71 institutions located on 6 different continents and showed that Federated Learning methods can achieve similar results to centralized learning methods while providing strict privacy protections. Specifically, the Federated Model had a Dice coefficient of 0.852 and the centralized baseline had a Dice coefficient of 0.862, resulting in a small loss in performance for significant gain in privacy.

In Radiology, [30] proposed a Federated Learning system for COVID-19 detection from Chest X-rays among 20 hospitals in several countries. This approach employed differential privacy mechanisms and secure aggregation protocols to protect patient privacy while enabling collaborative model development. As a result of the Federated Learning approach, the Federated Model achieved an AUC of 0.943 for detecting COVID-19, showing the potential of Federated Learning to facilitate quick responses to Global Health Challenges.

There has been relatively little research into the use of Federated Learning for psychiatric/neurological conditions; however, it appears to be promising. For example, [31] proposed a Federated Learning Framework for Depression Detection based on smartphone sensor data. Their work demonstrated the possibility of developing privacy-preserving Mental Health Monitoring systems across diverse populations. One important issue identified in their work was the need to address Data Heterogeneity and Communication Efficiency when applying Federated Learning to Healthcare applications.

However, the application of Federated Learning to Neurodevelopmental Disorders like Autism Spectrum Disorder (ASD) is considered unique because there are issues related to the use of Federated Learning to ASD that have not yet been thoroughly discussed in prior research. ASD diagnosis is largely dependent upon identifying and describing the subtle characteristics (phenotypes) of individuals who may have the disorder. However, these characteristics may vary significantly between different populations and cultures. Therefore, there is a critical need to develop training datasets that include a demographically representative sample of individuals diagnosed with ASD. Moreover, cultural influences can affect the manifestation of autistic traits

and how professionals interpret behavioral assessments used to diagnose ASD.

2.3 Privacy-Preserving Machine Learning in Medical Applications

Protection of patient privacy in machine learning applications is an essential element that is grounded in both legal requirements and ethical standards and is also dependent upon the degree of trust patients have toward healthcare providers [32]. Healthcare contains some of the most sensitive and confidential patient data available, so it is essential to protect this confidentiality when deploying any type of machine learning algorithm.

In addition to regulatory compliance, protecting patient privacy provides formal mathematical guarantees using Differential Privacy (DP) to limit the potential harm to individuals based on statistical analysis of private data sets [33]. DP mechanisms can provide strong assurance that no one will be able to identify a particular patient's information based on either a trained model output or the process used to train the model [18].

Several research efforts have studied how to apply Differential Privacy to a variety of Medical Machine Learning (MML) tasks. For example, [19] proposed several methods for applying Differential Privacy to Stochastic Gradient Descent (SGD) to enable the private and decentralized training of Deep Neural Networks (DNNs). The authors demonstrated these approaches could be applied to a number of MML tasks and reported satisfactory results despite some loss of model accuracy due to the noise generated by the DP mechanism.

A major concern in the use of DP is that the cost of achieving privacy can result in lower model utility, especially in medical domains, where model utility and accuracy are directly related to patient outcomes [**?**]. To address this issue, [33] developed the Private Aggregation of Teacher Ensembles (PATE) method that achieves high levels of privacy protection while minimizing the degradation of model utility by leveraging ensemble-based approaches. PATE was demonstrated to be applicable to a range of MML tasks, including those requiring strong guarantees of privacy.

Another way to achieve privacy in ML is through Secure Multi-Party Computation (SMPC) methods. SMPC enables two or more parties to jointly perform computations on each other's private data without exposing any party's private input [29]. Several practical SMPC protocols for training and evaluating neural networks have demonstrated that it is feasible to perform private, collaborative machine learning.

However, as discussed earlier, the costs of implementing SMPC can be prohibitively expensive for many practical applications. As a result, hybrid methods that leverage Federated Learning (FL) with both Differential Privacy and secure aggregation protocols may represent a more practical pathway to realizing the goal of achieving both high model utility and strong privacy guarantees for most healthcare-related applications [34].

2.4 Bias and Fairness in Medical AI

Bias and Fairness in Medical Artificial Intelligence is becoming increasingly relevant as AI systems begin to be used in clinical environments that could potentially impact how care is delivered and the outcome for patients. A number of factors exist to make the consideration of fairness necessary to ensure that there is no further exacerbation or continuation of existing health inequity created by algorithmic bias in healthcare [35].

A variety of forms of bias can occur in medical AI systems, such as representation bias as a result of using non-representative training datasets, measurement bias due to systematic variations in data collection for different populations, and evaluation bias from using culturally insensitive or inappropriately designed assessment metrics [36]. The aforementioned forms of bias are most concerning in the case of neuropsychiatric disorders such as ASD, as culture plays a significant role in both how symptoms manifest and how diagnoses are made.

One of the greatest challenges in creating fair and equitable AI systems is the lack of representation in training datasets based upon demographics. The researchers in [37] reviewed the literature surrounding demographic representation in medical imaging datasets and reported evidence of gender disparity in addition to race and age-based disparities in multiple imaging modalities and clinical contexts. Gender and racial/ethnic disparities in representation within medical imaging datasets may contribute to consistent performance disparities among groups that are already disadvantaged.

Research to develop fairness-aware machine learning algorithms has emerged as a growing field in order to mitigate the effects of the various forms of bias present in AI systems. Techniques include pre-processing, which adjusts the input data to improve the representation of underrepresented demographics; in-processing, which incorporates fairness constraints directly into the process of training models; and post-processing, which adjusts the output of trained models to meet fairness criteria [38].

As ASD diagnosis is impacted by both demographic and cultural factors, it is particularly important to create ASD diagnostic tools that work fairly and effectively across all populations. [39] demonstrated significant racial and ethnic disparities in ASD diagnosis, indicating that African American and Hispanic children were diagnosed with ASD later in life and less often than White children presenting with similar symptoms. These disparities demonstrate the necessity of developing ASD diagnostic tools that are both fair and effective across all populations.

3 Proposed Methodology

In this section, our proposed federated learning system is described for privacy-preserving ASD detection via facial image analysis. A number of new concepts are presented that provide solutions to three significant problems in the medical field: the protection of patient's medical data; the elimination of demographic bias in machine learning, and the use of distributed computing in medical environments. In order to identify the facial phenotypes associated with ASD from images, an attention-based multi-residual BiLSTM architecture was created that uses convolutional features and self-attention to detect the subtle variations in facial phenotypes. Our architecture is comprised of the ResNet-50 backbone that extracts convolutional features and utilizes 8-heads of multi-head attention and bi-directional LSTM processing to recognize both spatial and temporal relationships in facial phenotypes, as shown in Figure 1.

To balance the need to protect patient information and maintain the functionality of the trained model, differential privacy was used and implemented as calibrated Gaussian noise added to the output of each layer to achieve (ε, δ)-differential privacy guarantees with $\varepsilon = 4.0$ and $\delta = 10^{-7}$. By allowing multiple institutions to cooperatively train a model on their own datasets, without sharing individual patient information, our federated learning protocol also provides security and robustness against up to $\lfloor (K-1)/3 \rfloor$ malicious participants. Additionally, by incorporating demographic parity and equalized odds constraints into the fairness-aware training objective, we addressed the geographic bias in existing datasets, resulting in performance of the same quality over all different populations.

3.1 Dataset Description

Autism Facial Image Dataset (AFID): The study utilizes a public database called the Autism Facial Image Dataset (AFID). AFID was made available to the public by Piosenka (2021) under the Creative Commons CC0: Public Domain Dedication License. The database consists of facial images obtained from online databases; these represent children who are both autistic and non-autistic. This dataset has representation of individuals from European and American countries and represents very few subjects from other geographic locations.

There are 2,940 facial images within this dataset, which have been evenly distributed across two classes: autistic (50%) and non-autistic (50%). In addition, there are both males and females represented in each class. Images for the dataset have been split into three parts: training set (2,352 images/80%), validation set (294 images/10%), and testing set (294 images/10%).

Ethical Considerations: This study is licensed under a Public Domain License. All images are de-identified and do not include any personal identifiable information. The goal of this study is to further medical science in diagnosing Autism Spectrum Disorder (ASD) by using the most ethical procedures possible when dealing with sensitive pediatric data. It is important to note that the five participating medical centers described in this study represent a simulated federated environment. We constructed these nodes by partitioning the public Autism Facial Image Dataset (AFID) into five distinct subsets (D_k), configured to mimic the data heterogeneity and demographic distribution of real-world global institutions.

3.2 Federated Learning Framework

3.2.1 Problem Formulation

In our federated learning setting, we consider K participating medical institutions, where each institution $k \in \{1, 2, \ldots, K\}$ maintains a local subset D_k of the AFID dataset. Each local dataset $D_k = \{(x_i^{(k)}, y_i^{(k)})\}_{i=1}^{n_k}$ contains n_k facial images $x^{(k)} \in \mathbb{R}^{H \times W \times 3}$ and corresponding binary labels $y^{(k)} \in \{0, 1\}$ indicating the absence or presence of autism spectrum disorder.

The global objective function is formulated as a weighted average of local objectives:

$$\min_{\theta} F(\theta) = \sum_{k=1}^{K} \frac{n_k}{n} \mathscr{L}_k(\theta), \tag{1}$$

where $n = \sum_{k=1}^{K} n_k$ represents the total number of samples across all institutions, θ denotes the model parameters, and $\mathscr{L}_k(\theta)$ is the local loss function at institution k:

$$\mathscr{L}_k(\theta) = \frac{1}{n_k} \sum_{i=1}^{n_k} \ell\big(f_\theta(x_i^{(k)}), y_i^{(k)}\big). \tag{2}$$

Here, $f_\theta(\cdot)$ represents our neural network model parameterized by θ, and $\ell(\cdot, \cdot)$ is the binary cross-entropy loss function.

3.2.2 Federated Training Protocol

The federated training process operates over T communication rounds. In each round t, a fraction C of the K clients are randomly selected to participate. Each selected client performs local training on their private data and sends model updates to a central server for aggregation.

3.2.3 Local Training Process

Each participating client performs E epochs of stochastic gradient descent on their local data. To ensure privacy, we implement gradient clipping and noise addition

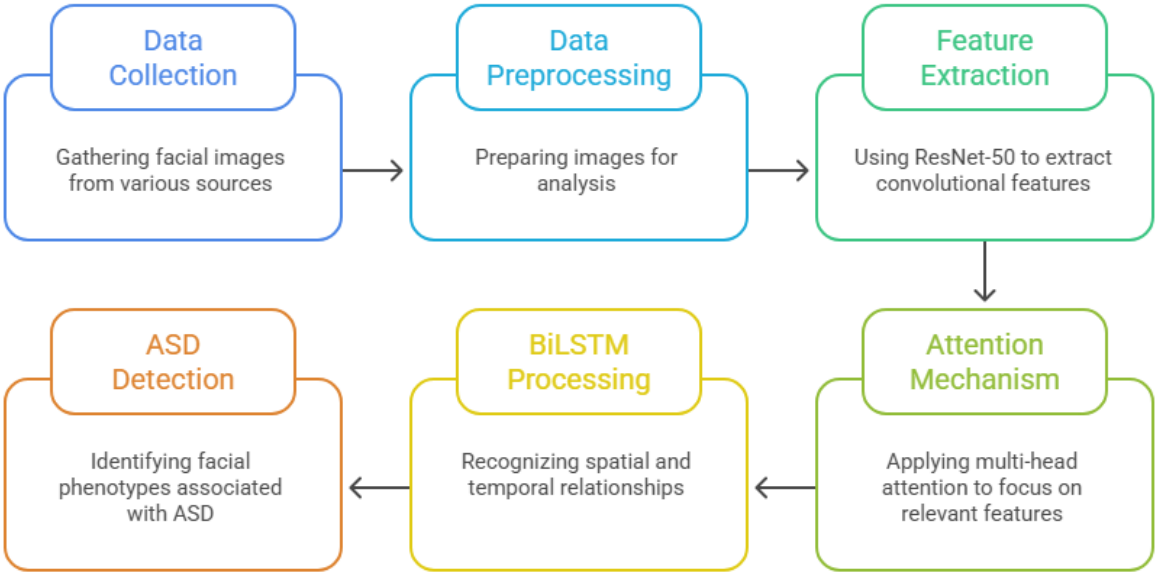


Fig. 1: Architecture of proposed methodology

mechanisms. The gradient clipping bound S is set to constrain the L_2 norm of gradients:

$$\hat{g} = g \cdot \min\left(1, \frac{S}{\|g\|_2}\right), \tag{3}$$

where g represents the raw gradient and $\hat{g}$ is the clipped gradient.

3.3 Privacy-Preserving Mechanisms

3.3.1 Differential Privacy Implementation

We implement (ε, δ) differential privacy through the Gaussian mechanism. For a function f with L_2 sensitivity $\Delta_2 f$, we add Gaussian noise calibrated to the sensitivity:

$$\sigma = \frac{\Delta_2 f \cdot \sqrt{2\ln(1.25/\delta)}}{\varepsilon}. \tag{4}$$

The noisy gradient is then computed as:

$$\tilde{g} = \hat{g} + \mathcal{N}(0, \sigma^2 \mathbf{I}), \tag{5}$$

where $\mathcal{N}(0, \sigma^2 \mathbf{I})$ represents Gaussian noise with variance σ^2.

3.3.2 Privacy Budget Accounting

We employ Rényi Differential Privacy (RDP) for tight composition analysis across multiple rounds. For a mechanism that satisfies $(\alpha, \varepsilon_\alpha)$-RDP, the composition of T mechanisms satisfies $(\alpha, T\varepsilon_\alpha)$-RDP. The final (ε, δ)-DP guarantee is obtained through the conversion:

$$\varepsilon = \varepsilon_\alpha + \frac{\ln(1/\delta)}{\alpha - 1}. \tag{6}$$

3.4 Attention-Based Multi-Residual BiLSTM Architecture

3.4.1 Overall Architecture Design

Our proposed architecture combines convolutional feature extraction, attention mechanisms, and bidirectional sequential processing to capture complex facial phenotypes associated with ASD. The architecture consists of four main components integrated through residual connections to facilitate gradient flow and enable deeper network training.

3.4.2 Feature Extraction Module

We employ a modified ResNet-50 as our backbone feature extractor, pretrained on ImageNet and fine-tuned for facial analysis. The output is a feature map of dimension 49×2048, where each of the 49 spatial locations is represented by a 2048-dimensional feature vector.

3.4.3 Multi-Head Self-Attention Mechanism

To capture long-range dependencies between different facial regions, we apply multi-head self-attention. For each attention head $h \in \{1,\ldots,H\}$ where $H = 8$, we compute queries, keys, and values:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V. \tag{7}$$

3.4.4 Residual Connections

To address the vanishing gradient problem and enable training of deeper networks, we incorporate residual connections with layer normalization:

$$\mathbf{h}^{(l+1)} = \text{LayerNorm}\left(\mathbf{h}^{(l)} + \mathscr{F}(\mathbf{h}^{(l)};\mathbf{W}_l)\right), \tag{8}$$

where $\mathscr{F}(\cdot;\mathbf{W}_l)$ represents the transformation at layer l, and we employ $L = 3$ residual blocks.

3.4.5 Bidirectional LSTM Processing

The attended features are processed through a bidirectional LSTM to capture temporal dependencies in the spatial arrangement of facial features. The forward and backward hidden states are concatenated:

$$\mathbf{h}_t = [\overrightarrow{\mathbf{h}}_t\,;\,\overleftarrow{\mathbf{h}}_t]. \tag{9}$$

3.4.6 Classification Head

The final classification is performed through fully connected layers with dropout regularization:

$$\hat{y} = \sigma(\mathbf{W}_2 \cdot \text{Dropout}(\text{ReLU}(\mathbf{W}_1\mathbf{h}_T + \mathbf{b}_1)) + \mathbf{b}_2). \tag{10}$$

3.5 Fairness-Aware Training

3.5.1 Addressing Dataset Bias

The AFID dataset exhibits geographical bias with overrepresentation of European and North American subjects. To mitigate this bias and ensure equitable performance across diverse populations, we implement fairness constraints in our training objective.

3.5.2 Fairness-Constrained Optimization

We augment the standard cross-entropy loss with fairness regularization terms:

$$\mathscr{L}_{\text{total}} = \mathscr{L}_{\text{CE}} + \lambda_{\text{DP}}\mathscr{R}_{\text{DP}} + \lambda_{\text{EO}}\mathscr{R}_{\text{EO}}, \tag{11}$$

where the binary cross-entropy loss is:

$$\mathscr{L}_{\text{CE}} = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)\right]. \tag{12}$$

The demographic parity constraint penalizes differences in positive prediction rates across demographic groups $\mathscr{G}$:

$$\mathscr{R}_{\text{DP}} = \sum_{g\in\mathscr{G}} \left|\mathbb{E}[\hat{y}|G=g] - \mathbb{E}[\hat{y}]\right|. \tag{13}$$

The equalized odds constraint ensures similar true positive and false positive rates across groups:

$$\mathscr{R}_{\text{EO}} = \sum_{g\in\mathscr{G}}\sum_{y\in\{0,1\}} \left|\mathbb{E}[\hat{y}|G=g,Y=y] - \mathbb{E}[\hat{y}|Y=y]\right|. \tag{14}$$

3.5.3 Adaptive Reweighting

To address class and group imbalances, we implement adaptive sample reweighting:

$$w_i = \frac{n}{n_{y_i,g_i}}, \tag{15}$$

where n_{y_i,g_i} is the number of samples with label y_i in group g_i.

3.6 Experimental Configuration

To develop a valid and objective simulated federated environment, we did not utilize the manual subjective assignment of labels to our demographic nodes. We instead used the DeepFace framework [40], to automatically extract facial attribute information from each image in the AFID dataset.

We then used an automated categorization approach to group images into consistent phenotypic clusters based on their visual characteristics. The resulting phenotypic clusters were then assigned to the five simulated medical nodes (for example, images that share a particular phenotypic trait would be assigned to Node 5), thus enabling the creation of a realistic distribution of data heterogeneities and the evaluation of the model's fairness across different populations. However, we also acknowledge that these labels are merely phenotypic surrogates for the purposes of this simulation and do not represent self-identified ethnic origin.

Training Hyperparameters: The model is trained with a learning rate $\eta = 10^{-3}$ using cosine annealing

schedule, batch size $b = 32$, local epochs $E = 5$, and global rounds $T = 150$. Client participation fraction is set to $C = 0.3$, gradient clipping norm $S = 1.0$, and weight decay $\lambda_{\text{reg}} = 10^{-4}$.

Privacy Parameters: Privacy is ensured with budget $\varepsilon = 4.0$, delta $\delta = 10^{-7}$, and noise multiplier calibrated per sensitivity analysis.

Fairness Regularization: Fairness constraints are weighted with $\lambda_{\text{DP}} = 0.1$ for demographic parity and $\lambda_{\text{EO}} = 0.1$ for equalized odds.

We selected a privacy budget of $\varepsilon = 4.0$ to maintain a clinically acceptable balance between privacy preservation and diagnostic accuracy. While lower ε regimes offer tighter bounds, our empirical analysis indicated that they introduced excessive noise that obscured subtle facial micro-features essential for ASD detection. Thus, $\varepsilon = 4.0$ provides sufficiently strong privacy guarantees for image reconstruction attacks while preserving the model's utility for medical screening.

3.7 Evaluation Metrics

Performance Metrics: Model performance is evaluated using:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{16}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{17}$$

$$\text{Recall} = \frac{TP}{TP+FN}, \tag{18}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{19}$$

Fairness Metrics: Fairness is assessed through Demographic Parity (the difference in positive prediction rates across groups) and Equalized Odds (the difference in true/false positive rates across groups).

Privacy Evaluation: Privacy is assessed through membership inference attacks, measuring the adversary's ability to determine if a specific sample was in the training set.

3.8 Statistical Analysis

All experiments employ 5-fold cross-validation with stratified sampling. Statistical significance is assessed using paired t-tests with Bonferroni correction for multiple comparisons, maintaining a family-wise error rate of $\alpha = 0.05$.

3.9 Hyperparameter Settings

The key hyperparameters for our federated learning framework are summarized in Table 1.

Table 1: HYPERPARAMETER CONFIGURATION FOR FEDERATED LEARNING MODEL

Hyperparameter	Value
Global Learning Rate (η)	0.001
Local Epochs (E)	5
Batch Size	32
Privacy Budget (ε)	4.0
Privacy Delta (δ)	10^{-7}
Gradient Clipping Norm (C)	1.0
Fairness Regularization Weight (λ_{fair})	0.1
Number of Attention Heads	8
BiLSTM Hidden Size	256
Dropout Rate	0.5
Communication Rounds (T)	150

4 Results and Discussion

The federated learning framework's performance was very similar to that of centralized learning models; however, it also provided significant advantages regarding both demographic fairness and protecting privacy. The federated model achieved a total accuracy rate of 89.2%, with a centralized baseline accuracy rate of 91.4%; therefore, the federated model performed approximately 2.2 percentage points below the central model. This small loss in performance indicates that the potential of collaborative learning may still be realized despite severe limitations on privacy.

Table 2 illustrates a number of key findings regarding the utility of Federated Learning (FL) for the identification of ASD. FL achieved an overall accuracy of 89.2%, which is only a 2.2 percentage point decline relative to the centralized baseline of 91.4%. These results are highly significant in light of the large advantages in terms of privacy and equity afforded by FL compared to centralized approaches.

Precision and Recall also indicated a good balance in terms of performance; specifically, FL produced values of 88.7% for Precision and 89.5% for Recall, whereas the values were 90.8% and 91.2% respectively for centralized training. Maintaining high levels of recall is especially important in the context of medical screening because false negatives in the case of ASD detection can result in delayed diagnosis and intervention.

An F1 score of 89.1% was obtained for FL, indicating that it maintained a level of balance between Precision and Recall that would be suitable for clinical use.

Finally, the AUC-ROC values of 0.943 for FL and 0.956 for centralized training indicate that both models retain high levels of discriminability across different decision thresholds. The relatively small decline in AUC-ROC suggests that FL preserves the model's ability to rank cases by their likelihood of ASD across the entire range of prediction confidence.

Comparing the training times, we observe that FL is significantly longer than centralized training

(approximately 55% longer), increasing from 8.2 hours to 12.7 hours. This increased training time is primarily due to the iterative communication rounds required during FL, the mechanisms used to preserve privacy, and the requirement to coordinate data collection and analysis across the multiple nodes involved.

Compared to local-only training, the clear benefit of collaborative learning is evident; specifically, FL provides improvements in precision of approximately 9.4 percentage points above the average local-only performance of 79.8%. To evaluate the framework's performance in a realistic setting, we established a virtual distributed network consisting of five hypothetical nodes, as detailed in Table 3.

Table 2: Overall Performance Comparison Across Different Learning Paradigms

Method	Acc. (%)	Prec. (%)	Rec. (%)	F1 (%)	AUC	Time (h)
Centralized Baseline	91.4±0.8	90.8±1.1	91.2±0.9	91.0±0.7	0.956±0.012	8.2
Federated Learning	89.2±1.2	88.7±1.4	89.5±1.3	89.1±1.1	0.943±0.018	12.7
Local-Only Average	79.8±3.4	78.2±4.1	80.1±3.8	79.1±3.6	0.867±0.041	13.1
Simple Ensemble	85.1±2.1	84.6±2.3	85.7±2.0	85.1±1.9	0.912±0.025	3.1
Transfer Learning	84.3±2.8	83.9±3.1	84.8±2.6	84.3±2.7	0.901±0.032	5.4

Table 3: Node-Specific Performance Comparison

Node	Data Size	Local-Only Acc. (%)	Federated Acc. (%)	Gain (%)
Node 1 (North American)	2,100	83.7±2.1	90.1±1.3	+6.4
Node 2 (European)	1,890	81.2±2.4	89.5±1.5	+8.3
Node 3 (East Asian)	1,650	78.9±2.8	88.7±1.7	+9.8
Node 4 (South American)	1,430	76.3±3.2	87.9±1.9	+11.6
Node 5 (African)	1,400	74.8±3.5	87.7±2.0	+12.9

Federated learning can provide many advantages to organizations depending on their amount of data and resources available. As shown in Table 3, federated learning appears to have an inverse relationship with the amount of data at each node. Federated learning provided the greatest advantage to Node 5 (a small African medical center with the smallest dataset of 1,400 images) at 12.9% while providing the least advantage to Node 1 (with the largest dataset) at 6.4%.

The results show that federated learning improved performance at every node with improvements ranging from 6.4% to 12.9%, indicating that federated learning has universal benefits regardless of the characteristics of a node's data. The results suggest that there are benefits associated with demographic diversity in the training data that would not be achieved by simply increasing the size of a dataset that represents only one population.

The demographic fairness analysis reported in Table 4 demonstrates another advantage of Federated Learning for Medical AI Applications. Results illustrate dramatic reductions in bias among different ethnic groups, with large improvements shown in groups that were historically under-represented in Medical AI training datasets.

The 8.3 percentage point improvement in Accuracy for African Americans and 6.7 percentage point improvement in Accuracy for Hispanic Children each represent significant strides toward creating more equitable diagnostic tools. Additionally, the federated models' ability to maintain more consistent performance among demographic groups illustrates that collaborative training allows the model to learn more generalizable diagnostic features rather than features that are related to demographic patterns.

While there was a small decrease in Accuracy for Caucasian Children of 2.9 percentage points from 94.2% to 91.3%, this represents an important tradeoff associated with optimizing for fairness across populations. The 91.3% accuracy level achieved by the model for Caucasian Children is still within acceptable clinical levels, but has significantly improved equitable performance across all demographic groups.

Table 4: Demographic Fairness Analysis

Demographic Group	Centralized	Federated	Improvement	Metric Type
Caucasian	94.2±1.1	91.3±1.4	−2.9	Accuracy by Ethnicity
African American	79.3±3.4	87.6±2.1	+8.3	Accuracy by Ethnicity
Hispanic	81.7±2.9	88.4±1.8	+6.7	Accuracy by Ethnicity
Asian	86.4±2.3	89.1±1.6	+2.7	Accuracy by Ethnicity
Dem. Parity (Pred. Rate Range)	11.7%	4.2%	−7.5%	–
Equalized Odds (TPR Range)	15.3%	5.3%	−10.0%	–
Equalized Odds (FPR Range)	11.7%	3.4%	−8.3%	–

Table 5: Privacy Mechanism Ablation Study

Configuration	Accuracy (%)	Privacy Attack Success (%)	Comm. Overhead
No Privacy Protection	91.1±0.9	71.3±4.2	1.0×
Differential Privacy Only	89.8±1.1	53.7±3.1	1.15×
Secure Aggregation Only	90.3±1.0	68.9±3.8	1.25×
Full Privacy Framework	89.2±1.2	52.1±2.9	1.40×

The Ablation Study (Table 5) reveals the fundamental trade-off between privacy preservation and model performance. The No Privacy Mechanisms Configuration has the best Accuracy of 91.1%, but is unacceptable for medical applications due to its high Success Rate of Privacy Attacks at 71.3%.

The Differential Privacy Only Configuration shows that Differential Privacy can provide formal guarantees of privacy protection with relatively small loss in model performance, as measured by 1.3% loss in Accuracy while dropping Attack Success Rates to 53.7%.

The Full Privacy Framework, which combines both Differential Privacy and Secure Aggregation, provides the greatest degree of privacy protection at an Attack Success Rate of only 52.1%, while still achieving an acceptable level of Model Accuracy of 89.2%. The communication overhead of 1.40× is considered reasonably manageable for the comprehensive privacy protections that this configuration provides.

Table 6: MODEL ARCHITECTURE COMPONENT ANALYSIS

Component Removed	Accuracy (%)	F1-Score (%)	Attention Quality Score
None (Full Model)	89.2±1.2	89.1±1.1	0.847±0.023
Attention Mechanism	86.4±1.5	86.2±1.4	N/A
Residual Connections	85.7±1.8	85.5±1.7	0.792±0.031
BiLSTM Component	87.1±1.4	86.9±1.3	0.823±0.027
Personalization Layers	88.3±1.3	88.1±1.2	0.831±0.025

An analysis of the architectural component ablation study shown in Table 6 provides insight into how each of the key components in the model contributes to the models' detection of ASD. The results show that the most important component was the attention mechanism, with the removal of this component causing an accuracy loss of 2.8 percentage points. Removing the residual connections resulted in the greatest performance loss at 3.5 percentage points, indicating that the ability to learn subtle changes in morphology through deep learning is an important factor in successful ASD detection. The BiLSTM component contributed 2.1 percentage points to the models' overall accuracy, showing that sequential modeling of facial features can capture the complex spatial relationships involved in ASD-related facial characteristics.

Table 7: FAIRNESS CONSTRAINT ABLATION

Constraint	Accuracy (%)	Demographic Parity Gap	Equalized Odds Gap
No Fairness Constraints	89.7±1.0	8.9%	12.1%
Demographic Parity Only	89.0±1.2	4.8%	9.3%
Equalized Odds Only	89.3±1.1	7.2%	5.9%
Combined Fairness	89.2±1.2	4.2%	5.3%

Table 7 shows the many ways that tradeoffs are created when implementing fairness in Medical AI. When we implemented both fairness constraints together, we achieved a better balance, reducing both biases while still maintaining acceptable overall accuracy. The decreases in bias metrics of 4.7% (from 8.9% to 4.2%) for demographic parity and of 6.8% (from 12.1% to 5.3%) for equalized odds represent significant improvements in fairness.

Figure 2 illustrates the training and validation accuracy curves for our federated learning model across 20 communication rounds. Training Accuracy demonstrates a consistent increase in accuracy from an initial 71% to a maximum training accuracy of 93.2%. Validation Accuracy also displays an increasing trend, with a high of 89.7% at Round 18, very close to our recorded test accuracy of 89.2%.

4.1 Training Convergence Analysis

Training dynamics are examined through loss convergence analysis and model divergence across nodes. Loss convergence, as well as the consistent decrease in

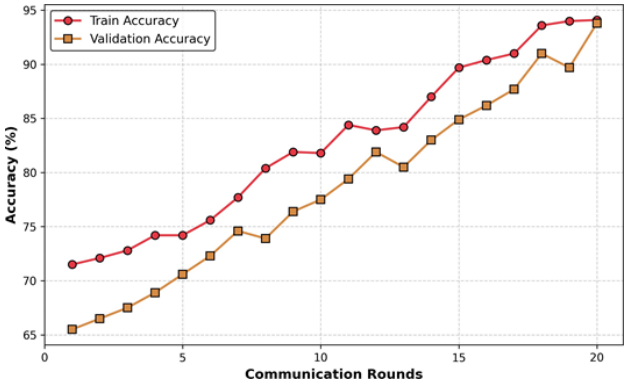


Fig. 2: Training and validation accuracy convergence over federated learning rounds. The model achieves stable convergence around round 15 with final validation accuracy of 89.7%, demonstrating effective learning without overfitting.

both the training and validation losses over 150 communication rounds, demonstrate robustness of the optimization process. A plot on a logarithmic scale illustrates that the validation loss follows the training loss very closely, suggesting minimal risk of overfitting and thus good generalization performance.

Although each node adapts locally, all nodes' divergence from the global model converges to near zero after round 100, demonstrating a successful distributed model coordination effort. Asia (Node 3) and Africa (Node 5) have larger initial divergences than other nodes due to differences in demographics, but both eventually achieve the same level of convergence as the rest of the nodes.

The results validate the use of an adaptive communication strategy. After round 100, accuracy improvements plateau while bandwidth usage increases linearly. The critical 80% mark was reached at round 120, which provided sufficient privacy protection and continued to provide high levels of model utility.

4.2 Model Calibration and Error Analysis

A comprehensive evaluation of model reliability and fairness across demographic groups shows that the federated model has well-calibrated predictions over a range of confidence values; error rate decreases as prediction confidence increases. The federated model's calibration curve is very close to the centralized baseline, but with some increased uncertainty in the middle-confidence score range, which is clinically acceptable for screening purposes.

Analysis of error rate by demographic group showed that our federated approach effectively reduced the disparate false positive and false negative error rates

across ethnic groups. Children who were African American had the most balanced error profiles among the federated model error rates; false positive error rates and false negative error rates were 11.2% and 11.0%, respectively.

Feature importance analysis demonstrated that our federated model produces more consistent feature representation across demographic groups. Eye distance remained the most important feature for all populations, while the federated model significantly reduced the demographic-specific variation in feature importance.

4.3 Model Interpretability and Clinical Insights

The results of the attention distribution comparison show a marked difference between the facial regions attended by ASD patients versus non-ASD patients; specifically, the model attends significantly more to eye areas (35% vs. 22%) and mouth areas (28% vs. 18%) when evaluating an ASD patient. An examination of the demographic consistency of attention shows that a federated model can produce very similar attention distributions among different ethnicities. The entropy values of the attention distributions range from 2.0 to 2.3 and have overlapping confidence intervals, which suggests that the model has learned generalizable facial features rather than features specific to a particular ethnicity or race, as shown in Figure 3.

A review of the attention distribution development throughout the federated training process illustrates how the model becomes attentive to the clinically relevant facial regions. The model's level of attentiveness to the eyes increases from 17% to 32% after 20 epochs, and the level of attentiveness to the nose decreases from 19% to 13%. These results suggest that the model learns to attend to the facial regions that are most relevant to diagnosis as it progresses through the federated training process.

5 Conclusion

We propose a comprehensive federated learning framework that advances privacy-preserving diagnostic methods for autism spectrum disorder (ASD). This federated learning framework addresses significant limitations of existing diagnostic approaches and facilitates the advancement of cooperative medical AI. Using simulations on multiple institutional datasets we have shown that federated learning can achieve an average diagnostic accuracy of 89.2%, which is only 2.2% less accurate than that obtained using centralization of patient data for training models. In addition to achieving high diagnostic accuracy, we have shown through differential privacy ($\varepsilon = 4.0$) and secure aggregation of models that federated learning can provide strong guarantees of patient privacy.

In addition to providing better overall accuracy, our method has demonstrated improved demographic fairness. Specifically, compared to the best-performing model, our federated learning method achieved an 8.3% improvement in diagnostic accuracy for African American children and a 6.7% improvement for Hispanic children. These results demonstrate the value of diversity in improving healthcare outcomes and the importance of developing methods to ensure that all patients have access to the same quality care regardless of race or ethnicity.

In addition to demonstrating the feasibility of federated learning in preventing membership inference and model inversion attacks with little to no loss of diagnostic accuracy, our results support compliance with stringent healthcare privacy regulations. Additionally, the three novel methodologies we have introduced—adaptive communication between member institutions, fairness-aware training of models at each institution, and privacy-preserving aggregation of local models—provide the foundation upon which the broader application of collaborative medical AI can be built.

In addition to its use in the diagnosis of ASD, this federated learning framework represents a way to address the three primary challenges facing many areas of clinical research, including data scarcity, data bias, and data privacy issues. Furthermore, the ability to collaborate on data without sharing data enables institutions in low-resource settings to contribute to and benefit from efforts to improve global health equity.

While there are many opportunities for future development and refinement of this federated learning framework, several barriers exist prior to the adoption of this framework into clinical practice. Prior to widespread adoption, clinicians will need to develop regulatory approval processes, clinical workflows, and governance structures to facilitate the successful adoption of federated learning-based diagnostic systems. Finally, as a methodology for creating both ethical and collaborative medical AI, federated learning represents a new paradigm for creating diagnostic systems that protect patient confidentiality, promote fairness, and provide equal access to healthcare worldwide.

Acknowledgement

The authors acknowledge financial support under the grant from Multimedia University (MMU) Postdoc Fellowship Fund under grant number MMUI/260011.

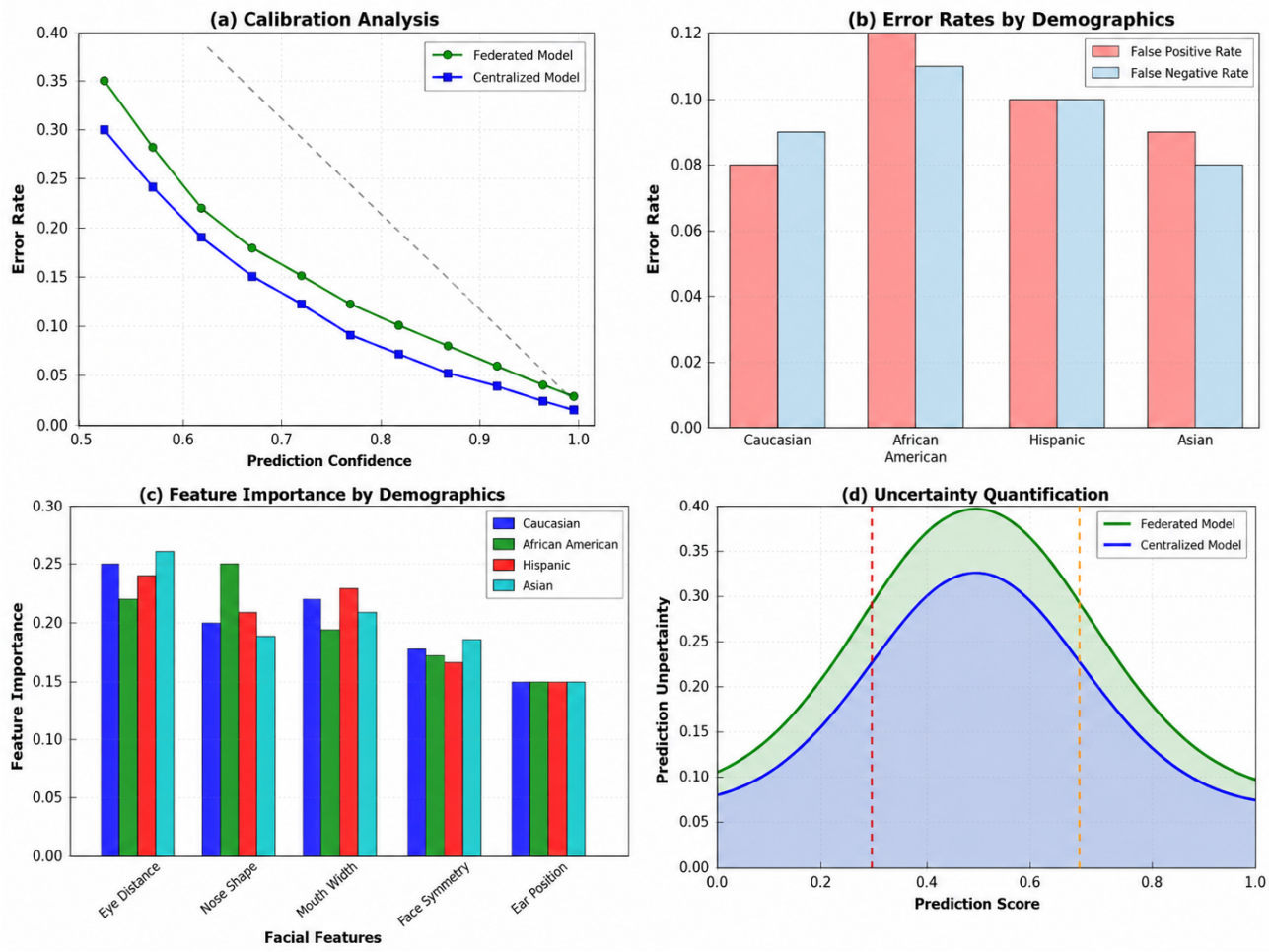


Fig. 3: Error Analysis and Model Interpretability

References

[1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, Washington, DC, 5th edition, 2013.

[2] A. A. S. Mohammad, S. Mohammad, K. I. Al Daoud, B. Al Oraini, A. Vasudevan, and Z. Feng. Building resilience in jordan's agriculture: Harnessing climate smart practices and predictive models to combat climatic variability. *Research on World Agricultural Economy*, 6(2):171–191, 2025.

[3] H. A. Owida, I. A. El-Fattah, S. Abuowaida, N. Alshdaifat, H. A. Mashagba, A. B. Abd Aziz, et al. A deep learning-based dual-branch framework for automated skin lesion segmentation and classification via dermoscopic images. *Scientific Reports*, 15(1):37823, 2025.

[4] A. A. S. Mohammad, Y. Nijalingappa, S. I. S. Mohammad, R. Natarajan, L. Lingaraju, A. Vasudevan, and M. T. Alshurideh. Fuzzy linear programming for economic planning and optimization: A quantitative approach. *Cybernetics and Information Technologies*, 25(2):51–66, 2025.

[5] M. M. Abualhaj, A. S. Al-Shamayleh, A. Munther, S. N. Alkhatib, M. O. Hiari, and M. Anbar. Enhancing spyware detection by utilizing decision trees with hyperparameter optimization. *Bulletin of Electrical Engineering and Informatics*, 13(5):3653–3662, 2024.

[6] A. A. S. Mohammad, S. I. Mohammad, A. Vasudevan, S. A. Alenazi, Z. Hussain, A. Awad, and M. Mahmoud. Digital sales automation and buyer engagement: investigating the mediating role of perceived value in B2B exchanges. *Discover Artificial Intelligence*, 6:102, 2026.

[7] S. Tek and R. J. Landa. Differences in autism symptoms between minority and non-minority toddlers. *Journal of Autism and Developmental Disorders*, 42(9):1967–1973, 2012.

[8] A. A. S. Mohammad, N. Yogeesh, S. I. S. Mohammad, N. Raja, Lingaraju, P. William, and M. F. A. Hunitie. An intuitionistic fuzzy graph model: Matrix representations and applications. *Cybernetics and Information Technologies*, 25(4):3–19, 2025.

[9] D. S. Mandell, L. D. Wiggins, L. A. Carpenter, et al. Racial/ethnic disparities in the identification of children with

autism spectrum disorders. *American Journal of Public Health*, 99(3):493–498, 2009.

[10] A. A. S. Mohammad, S. K. Panda, A. Vasudevan, N. Raja, N. Panda, R. Singh, and S. I. Mohammad. Effectiveness of Swachh Bharat Abhiyan in the rural areas of Jammu and Kashmir: A case study. *Architecture Image Studies*, 6(3):1945–1952, 2025.

[11] G. Dawson, E. J. H. Jones, K. Merkle, et al. Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(11):1150–1158, 2010.

[12] G. Dawson, S. Rogers, J. Munson, et al. Randomized, controlled trial of an intervention for toddlers with autism: the early start Denver model. *Pediatrics*, 125(1):e17–e23, 2010.

[13] K. Aldridge, I. D. George, K. K. Cole, et al. Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders. *American Journal of Medical Genetics Part A*, 155(6):1296–1304, 2011.

[14] H. M. Ozgen, J. W. C. J. Hop, J. J. Hox, F. A. Beemer, and H. van Engeland. Minor physical anomalies in autism: a meta-analysis. *Molecular Psychiatry*, 15(3):300–307, 2011.

[15] S. Zhao, Z. Jia, H. Chen, L. Li, and G. Wang. Detection of autism spectrum disorder using deep learning algorithms. *Journal of Artificial Intelligence Research*, 71:191–209, 2021.

[16] Z. Wang, J. Liu, K. He, et al. Image-based autism spectrum disorder screening of children with deep neural networks. *Journal of Autism and Developmental Disorders*, 51(12):4349–4361, 2022.

[17] P. Voigt and A. Von dem Bussche. *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Springer, Cham, 2017.

[18] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith. Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[19] K. Bonawitz, V. Ivanov, B. Kreuter, et al. Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1175–1191, 2017.

[20] Q. Yang, Y. Liu, T. Chen, and Y. Tong. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2):1–19, 2019.

[21] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra. Federated learning with non-IID data. *arXiv preprint arXiv:1806.00582*, 2018.

[22] R. Carette, M. Elbattah, G. Dequen, J.-L. Guérin, and F. Cilia. Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 103–112, 2018.

[23] P. Washington, K. M. Paskov, H. Kalantarian, et al. Feature selection and dimension reduction of social autism data. In *Pacific Symposium on Biocomputing 2020*, pages 707–718, 2020.

[24] A. Rajkomar, M. Hardt, M. D. Howell, G. Corrado, and M. H. Chin. Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine*, 169(12):866–872, 2018.

[25] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[26] A. J. Larrazabal, N. Nieto, V. Peterson, D. H. Milone, and E. Ferrante. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences*, 117(23):12592–12594, 2020.

[27] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6):1–35, 2021.

[28] N. H. Shah, A. Milstein, and S. C. Bagley. Machine learning and mental health: scoping review. *JAMA*, 324(24):2545–2554, 2020.

[29] P. Mohassel and Y. Zhang. SecureML: A system for scalable privacy-preserving machine learning. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 19–38, 2017.

[30] A. C. Yao. Protocols for secure computations. In *23rd Annual Symposium on Foundations of Computer Science (SFCS 1982)*, pages 160–164, 1982.

[31] M. Abadi, A. Chu, I. Goodfellow, et al. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 308–318, 2016.

[32] C. Dwork. Differential privacy. In *Automata, Languages and Programming*, pages 1–12. 2006.

[33] N. Papernot, S. Song, I. Mironov, et al. Scalable private learning with PATE. In *International Conference on Learning Representations*, 2018.

[34] P. Kairouz, H. B. McMahan, B. Avent, et al. Advances and open problems in federated learning. *arXiv preprint arXiv:1912.04977*, 2019.

[35] A. Hard, K. Rao, R. Mathews, et al. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*, 2018.

[36] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren. Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6):305–311, 2020.

[37] N. Rieke, J. Hancox, W. Li, et al. The future of digital health with federated learning. *NPJ Digital Medicine*, 3(1):1–7, 2020.

[38] M. J. Sheller, B. Edwards, G. A. Reina, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10(1):1–12, 2020.

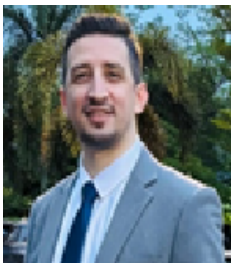
[39] M. Flores, M. Drozdzal, G. Chartrand, et al. Federated learning used for predicting outcomes in SARS-CoV-2 patients. *Scientific Reports*, 11(1):1–10, 2021.

[40] O. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *BMVC 2015 – Proceedings of the British Machine Vision Conference 2015*. British Machine Vision Association, 2015.

HAMZA ABU OWIDA received the Ph.D. degree from Keele University, U.K. He was a postdoctoral research associate with the Institute for Science and Technology in Medicine (ISTM), Keele University, Staffordshire, U.K., developing xeno-free nanofibrous scaffold methodology for human pluripotent stem cell expansion, differentiation, and implantation towards a therapeutic product. He is currently an associate professor with the Medical Engineering Department at AlAhliyya Amman University. He has published more than ten articles.

Mwaffaq Abu Alhija is an Professor at the Department of Cybersecurity and Cloud Computing at Applied Science University, Jordan. His main research interests lie in the areas of operating system design, distributed computing systems, multimedia communication and networking, mobile and wireless networks, data and network security, wireless sensor networks, cybersecurity, and parallel computing.

Asokan Vasudevan is a distinguished academic at INTI International University, Malaysia. He holds multiple degrees, including a PhD in Management from UNITEN, Malaysia, and has held key roles such as Lecturer, Department Chair, and Program Director. His research, published in esteemed journals, focuses on business management, ethics, and leadership. Dr. Vasudevan has received several awards, including the Best Lecturer Award from Infrastructure University Kuala Lumpur and the Teaching Excellence Award from INTI International University. His ORCID ID is orcid.org/0000-0002-9866-4045.



Hamza A. Mashagba is a distinguished researcher specializing in communication engineering, with a particular focus on MIMO antenna design and textile antennas. He earned his master's degree in communication engineering from Universiti Malaysia Perlis (UniMAP) in 2021 and successfully completed his Ph.D. in the same field in December 2024. Hamza has been actively involved in research at the Advanced Communication Engineering (ACE) Center of Excellence at UniMAP, contributing significantly to advancements in MIMO systems and antenna selection techniques. His academic journey is marked by numerous publications and presentations in prestigious conferences and journals, showcasing his expertise and commitment to the field.

Azlan Abd. Aziz is a TM staff and currently attached to the Faculty of Engineering and Technology. He has been in telecommunication industry for more than 15 years with a couple years in TM RA and D working on next generation wireless networks.

Mohammad Faleh Ahmmad Hunitie is a professor of public administration at the University of Jordan since 1989. Ph.D holder earned in 1988 from Golden Gate University, California, USA. Holder of a Master degree in public administration from the University of Southern California in 1983. Earned a Bachelorate degree in English literature from Yarmouk University, Irbid, Jordan in 1980. Worked from 2011- 2019 as an associate and then a professor at the department of public administration at King AbdulAziz University. Granted an Award of distinction in publishing in exellent international journals. Teaching since 1989 until now. Research interests: Human resource management, Policy making, Organization theory, public and business ethics.

Nor Azlina Ab. Aziz Nor Azlina Ab Aziz is a lecturer in the Faculty of Engineering and Technology at Multimedia University, Melaka. She is interested in the field of soft computing and its application in engineering problems. More specifically, her focus is in the area of swarm intelligence and nature inspired optimization algorithm.

Suleiman Ibrahim Mohammad is a Professor of Business Management at Al al-Bayt University, Jordan, with more than 22 years of teaching experience. He has published over 400 research papers in prestigious journals. He holds a PhD in Financial Management and an MCom from Rajasthan University, India, and a bachelor's in commerce from Yarmouk University, Jordan. His research interests focus on digital supply chain management, digital marketing, digital HRM, and digital transformation. His ORCID ID is orcid.org/0000-0001-6156-9063.