I.J. Image, Graphics and Signal Processing, **2025, 1, 1-16**
Published Online on February 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2025.01.01

Disorder Facial Emotion Recognition with Grey Wolf Optimized Gefficient-Net for Autism Spectrum

Ramachandran Vedantham*
Department of CSE, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, 522508, India
Email: ramachandranvedantham4@gmail.com
ORCID iD: https://orcid.org/0000-0002-3857-8145
*Corresponding Author

Ravisankar Malladi
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, 522502, India
ORCID iD: https://orcid.org/0000-0002-8250-6595

Sivaiah Bellamkonda
Department of CSE, Indian Institute of Information Technology, Kottayam, Kerala, 686635, India
ORCID iD: https://orcid.org/0000-0001-6948-3483

Edara Sreenivasa Reddy
ANU College of Engineering & Technology, Acharya Nagarjuna University, Nambur, Andhra Pradesh, 522510, India
ORCID iD: https://orcid.org/0000-0001-6948-3483

Received: 29 May, 2024; Revised: 26 June, 2024; Accepted: 15 August, 2024; Published: 08 February, 2025

Abstract: Autism spectrum disorder (ASD) is a neurological issue that impacts brain function at an earlier stage. The autistic person realizes several complexities in communication or social interaction. ASD detection from face images is complicated in the field of computer vision. In this paper, a hybrid GEfficient-Net with a Gray-Wolf (GWO) optimization algorithm for detecting ASD from facial images is proposed. The proposed approach combines the advantages of both EfficientNet and GoogleNet. Initially, the face image from the dataset is pre-processed, and the facial features are extracted with the VGG-16 feature extraction technique. It extracts the most discriminative features by learning the representation of each network layer. The hyperparameters of GoogleNet are optimally selected with the GWO algorithm. The proposed approach is uniformly scaled in all directions to enhance performance. The proposed approach is implemented with the Autistic children's face image dataset, and the performance is computed in terms of accuracy, sensitivity, specificity, G-mean, etc. Moreover, the proposed approach improves the accuracy to 0.9654 and minimizes the error rate to 0.0512. The experimental outcomes demonstrate the proposed ASD diagnosis has achieved better performance.

Index Terms: Autism Spectrum, Disorder, Facial Expression Recognition, Gefficient-Net, Grey Wolf Optimization, Feature Extraction, VGG-16

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects people from birth, with symptoms appearing earlier. The diagnosis of ASD is based on various sessions, including manual coding behavior, exhaustive screening, and behavioral observation. Natural behavior observation of ASD is enhanced with technological related tools for the automatic detection of behavior [1,2]. Nonverbal communications, including voice tone, appearance, facial expression, and gestures, represent an individual's feelings, thinking, and intention as a part of communication [3].

The face represents important trends in social communication, as facial expressions help to identify a person's mental or emotional mood. ASD persons have disrupted communication, and others' facial expressions are not recognized [4].

The ASD features are insufficient for detecting the facial expression. It is required to identify the relationship between social skills and facial expression, which leads to the development of facial expression recognition [5].When examining the facial image with a very intense expression, no impairment could be detected. The linguistic approaches or compensatory cognitive strategies are utilized with automatic processing for recognizing the expression [6]. In recent research, parameters such as gender, allocation time, and kind of expression are considered for detecting ASD [7]. It measures the degree of intensity required to improve the accuracy. The intensity based approaches are processed with a lower number of samples. Hence, the number of samples increased to get adequate feature information for ASD diagnosis. Hence, the capacity to differentiate facial expressions has certain limitations [8].

The training set is labeled with little intensity, and ASD recognition is highly suitable for emotional expressions [9]. With the rapid development of technologies, ASD is diagnosed with Electroencephalogram(EEG), Galvanic Skin Response (GSR), Inertial Measurement Unit (IMU), Magnetic Resonance Imaging (MRI), Electrocardiograph (ECG), and Computer Vision (CV) [10]. These techniques are applied to differentiate ASD patients from TD. The primary assessment approaches include screening and diagnosing based on a symptomatic approach. It provides feasible solutions with computer vision technologies for standardizing clinical assessments [11,12]. In natural data processing, the development of Artificial Intelligence (AI) based technologies provides solutions for image analysis. Machine learning or deep learning analyzes the images in an effective way using traditional image processing approaches [13,14].

Artificial Neural network (ANN) is introduced in ASD detection to improve performance by making the learning process easier. Deep learning approaches such as Deep Neural Networks (DNN) provide accurate prediction with minimal computational time. High level spatial information is obtained by fusing the features of higher and lower level [15,16]. With DNN, the features are learned at varying levels and interconnected during the interaction. The semantic features, including eye movement, pseudo sequence features, and sequence of scanning points from the image, are obtained with Fully Convolutional Networks (FCN) and Convolutional Long Short Term Memory Networks (CLSTMs) [17,18,19]. By integrating these features, salient mapping is produced. The DL architectures are widely used in facial expression analysis and medical image analysis for automated disease diagnosis [20].

The proposed approach is based on the following contribution.

- The face images are pre-processed with cropping, normalization, and redundancy removal. The data ranges are scaled, and the face image part is extracted with pre-processing.
- Extracting the features with the VGG-16 feature extraction technique provides several patterns of face images. The facial features are essential for deciding the final classification of ASD recognition.
- Extracted features are classified based on GEfficient-Net with the GWO algorithm. The proposed classification algorithm makes network training easier by scaling the network in all directions.

The paper has the following organization. The works related to the proposed approach are described in section 2. The details of the proposed GWO-Efficient-Net approach are given in section 3. The evaluation of proposed autism detection is provided in section 4, and the conclusion is provided in section 5.

2. Related Work

To recognize ASD from facial expressions, Sadik et al. [21] utilized a DL model, MobileNet, with transfer learning. Transfer learning was introduced to enhance ASD recognition. It was achieved with the validation and testing accuracy of 89% and 87%.

Early detection of autism changes the behavior of adults. With the advantage of artificial intelligence approaches, A.KALAISELVI et al. [22] developed transfer learning based approaches for ASD diagnosis. The DL architecture model, including VGG19, InceptionV3, ResNet50, and NASNetLarge models, was used to detect facial expressions. Both the expression of autism and non-autism face images were provided as testing, training, and validation data. An accuracy of 87.50% with a 0.372 loss was attained with the NASANetLarge DL model.

The ASD diagnosis requires generalizable, long-lasting, and beneficial results for facial emotion recognition. Zhang et al. [23] suggested meta-analytic approaches for evaluating the benefits of facial emotional training with maintenance and generalization. It was analyzed with the study of 595 ASD patients, and there is no evidence related to improving social skills. To enhance the performance of emotion recognition, a facial expression-based training procedure is involved.

More ASD patients are diagnosed every year based on social communication. Husna et al. [24] developed a DL approach based on CNN for detecting patients with ASD or not. The neural patterns were classified for manipulating the accuracy performance. With ResNet-50 and VGG-16, the accuracy results obtained with the proposed approach are 87% and 63.4 %, respectively.

By analyzing ASD based on facial features, the enhanced autism detection based framework based on transfer learning was developed by Akter et al. [25]. The essential markers were encoded in the human face for ASD identification with respect to eye contact, facial features, etc. The face images were collected from the facial repository, and the pretrained DL models were applied to the images. The performance is improved with MibileNet-v1, and the classification is accomplished with the k-means clustering approach. The accuracy rate of 90.67% and miss rate of 9.33% were achieved with the MObileNet-V1 approach.

The facial emotions are encoded with neural signals due to their FER behavior. Juan Manuel Mayor Torres et al. [26] developed a DCNN facial expression classification approach for ASD and non-ASD groups. Higher accuracy is obtained with independent participant samples. Based on the behavior performance, higher accuracy is obtained with the ASD group. The facial emotion classification of ASD individuals was obtained with feature based analysis. The comparison of existing approaches is given in Table 1.

Table 1. Comparison of traditional ASD detection approaches

References	Techniques	Advantages	Disadvantages
Sadik et al. [21]	MobileNet, Transfer learning	Efficient approach	The performance is required to be improved with integrated DL approaches.
A.KALAISELVI et al. [22]	VGG19, InceptionV3, ResNet50, and NASNetLarge	Better performance in terms of accuracy and loss.	More advanced techniques are required for accurate detection.
Zhang et al. [23]	Meta-analytic technique	Face emotion recognition is improved with medium size accuracy.	The training procedure is not based on general social skills.
Husna et al. [24]	ResNet-50, VGG-16	Higher accuracy is obtained	It is not evaluated with real-time applications.
Akter et al. [25]	MobileNet-V1, K-means clustering	Simple, cost effective, and requires low resources.	Prediction performance requires further improvement with multi-class classification approaches.
Torres et al. [26]	DCNN	Highly reliable and accurate.	The performance needs to be improved.

3. Proposed Methodology

Face recognition is essential for diagnosing ASD and for providing earlier treatment. In this paper, GefficientNet with GWO has been proposed for face recognition based ASD detection. Initially, the face image from the dataset is pre-processed for redundant removal, cropping of the facial region, and normalization. Then, the pre-processed images are fed to the feature extraction process. The VGG-16 is utilized for efficient feature extraction. The extracted features are classified using the classification algorithm GEfficientNet. The optimal parameters for classification are selected with the GWO algorithm. VGG16 is selected for feature extraction since it is capable of extracting large feature information for accurate classification. It utilizes a small receptive field to make the decision process non-linear. The proposed VGG16 has convolutional, max pooling, batch normalization, and fully connected layers. In GEfficientNet the learning function is compatible and trains faster than other DL models. It also assures enhanced classification with dimensionality reduction. The GWO benefits from fewer parameters, easy implementation, and simplified processing principles. Due to these advantages, these algorithms are effectively utilized in previous research for enhancing performance. Hence, VGG16 and GEfficientNet with GWO optimization are selected for the proposed approach.

The proposed framework of ASD detection is given in Fig. 1.

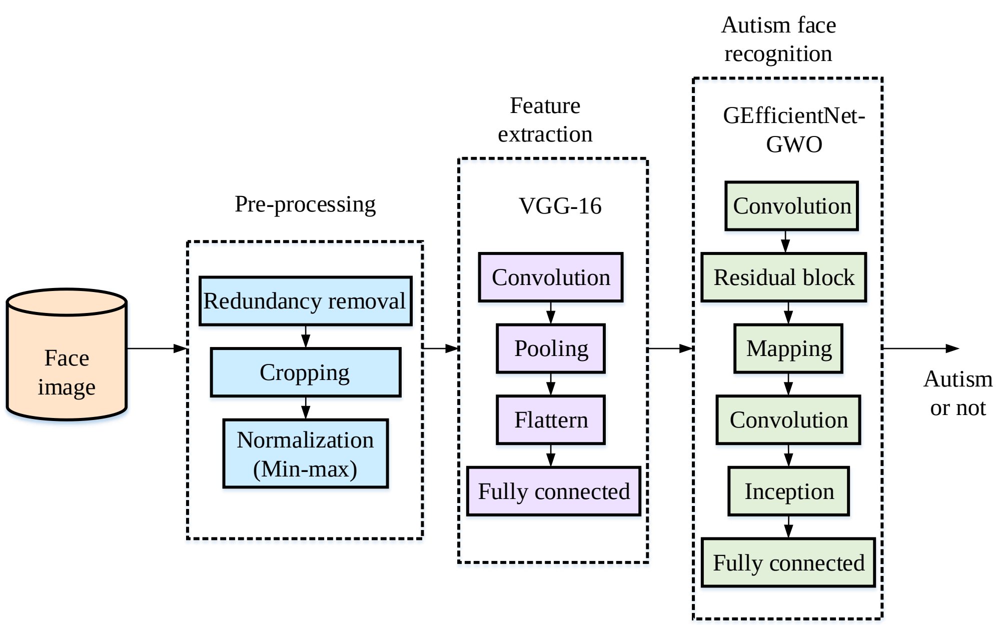

Fig. 1. The proposed ASD recognition architecture based on face emotion recognition

3.1 Pre-processing

The pre-processing improves the image quality for further processing. In pre-processing, the duplicate images are removed from the dataset, and then the images are cropped by considering only the face from the image by eliminating the background portion. After getting the facial portion, normalization is applied to the image.

Normalization: The relationship between the actual data values is determined by min-max normalization to preserve the original facial image shape. The effect of the outlier is eliminated with the standard deviation of the restricted range. It minimizes the variation between image pixels to enhance image quality. The description of min-max normalization is given in equations (1) and (2).

$$Y_{sd} = \frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}} \tag{1}$$

$$Y_{scale} = Y_{sd} \times (\max - \min) + \min \tag{2}$$

Where, $\min$ represents the minimal feature values for the input data Y, and $\max$ represents the feature range of maximum values.

3.2 Feature extraction

In feature representation, the raw data is converted into valid features for processing by preserving the essential information from the dataset. High level feature information can be obtained with VGG-16 to enable precise classification. It is suggested that new feature extraction methods be used to enhance ASD's detection capabilities. Nevertheless, it results in significant computing complexity when the novel approach is included with very little performance improvement. The reasons for selecting the VGG-16 approach are as follows: VGG-16 is a famous feature extraction technique that extracts features from images and performs better classification than other deep learning models. The VGG-16 was used on a collection of MRI scan images. The proposed feature extraction architecture is completed with several levels. VGG-16 is a technique for efficiently reducing data requiring little processing power. Hence, the VGG-16-based feature extraction is used in the proposed ASD diagnosis.

VGG-16 is a convolutional neural network (CNN) with a convolutional pooling and batch normalization layer. The architecture of VGG-16 is shown in Fig. 2. It estimates the outcome of all nodes linked to the local region matrix. In between groups of weight values, dot products related to the local input region are computed. The output from the convolution layer is given to the ReLU activation function. In the pooling layer, the technique of downsampling is included for minimizing the height and width of output data.

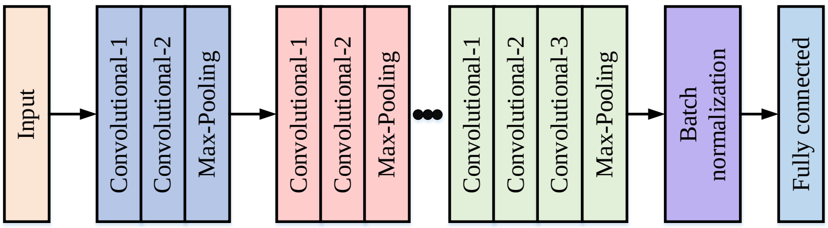

Fig. 2. Architecture of VGG-16 for feature extraction

After several convolution blocks, the channels are increased with reduced height and width. While considering the max pooling operation, the channel height and width are represented with P_i and P_x. The face image is partitioned with the adjacent pixel size of 2×2, producing the resultant image with the size of $P_i / 2$ and $P_x / 2$. The original VGG-16 architecture is slightly varied by eliminating one convolution block. It minimizes the memory load of image size. When the number of layers is increased, deeper features are extracted. However, efficient architecture is required to ensure correct learning probability without data overfitting.

Batch normalization is applied in the output layer to avoid overfitting. It can be applied to any network layer with the objective of distributing constant activation value. It avoids overfitting issues by minimizing the internal covariate shift. For the input data $y = (y^{(1)} ... y^{(e)})$ with dimension e, the normalized form is represented with the following equation.

$$\hat{y}^{(l)} = \frac{y^{(l)} E[y^{(l)}]}{\sqrt{Var[y^{(l)}]}} \tag{3}$$

Where, $y^{(l)}$ represents each activation, δ represents the constant value for numerical stability. The mean and variance are described with the following equation.

$$E\left[y^{(l)}\right] = \tfrac{1}{n}\sum_{j=1}^{n} y_j^{(l)} \tag{4}$$

$$Var\left[y^{(l)}\right] = \frac{1}{n}\sum_{j=1}^{n}\left(y_j^{(l)} - E\left[y^{(l)}\right]\right)^2 + \delta \tag{5}$$

With the scaling factor, $\lambda^{(l)}$, the normalized value is multiplied to ensure the data is not limited to small. With the offset range $\alpha^{(l)}$, the output is given with the following equation.

$$z^{(l)} = \lambda^{(l)} y^{(l)} + \alpha^{(l)} \tag{6}$$

3.3 Face Recognition based autism Detection

Face recognition is accomplished through a hybrid GEfficient-Net shown in Fig. 3. To address the training challenges associated with EfficientNet, the proposed model incorporates GoogleNet into its architecture. GoogleNet reduces the number of parameters with the Inception module, which enhances efficiency. Although employing fewer parameters, it operates other deep learning methods. Furthermore, hyperparameter adjustment combined with GWO optimization yields the best results. It also uses simple ideas and fewer parameters to accelerate the model. Efficient-Net enhances the network performance by varying dimensions. It is the base model of a convolutional network, which has the dimensions including depth (number of layers), width (number of channels), and resolution of the image. In a convolutional network, scaling is accomplished in a single direction. To improve accuracy and efficiency, the network dimension is balanced in a consistent direction. The compound scaling is accomplished with the constant ratio of scaling. It is utilized with the baseline architecture P by varying the network width D, length M, and resolution (X, I). The network accuracy is improved with the constraints of memory and FLOPs.

$$\max_{e,x,s} Ac(P(e,x,s)); M(P) \leq T_m, F(P) \leq T_f \tag{7}$$

Where, A_c represents the accuracy, T_m represents the target memory, T_f represents the target FLOPs, $e, x, and\ s$ represents the depth, width, and resolution of the baseline network. To minimize the search space $< W, D, X, I >$ the layers are uniformly scaled with a constant ratio. Therefore, the network dimension is represented by the depth, width, and resolution of the base network.

$$e = Al^{\phi}, x = Bt^{\phi}, s = Gm^{\phi} \tag{8}$$

Where, Al, Bt, and Gm are the constants detected from the grid search, ϕ represents the constant coefficient. These coefficients are utilized for controlling the amount of variable resources. The FLOPs used for normal convolution are proportional to e^2, x^2, s^2. The convolution operation increases the computation cost; thus, the compound scaling increases the number of FLOPs.

In Efficient-Net architecture, MBConv represents the inverted residual block with Squeeze and Excitation (SE) block. The residual block is responsible for connecting the starting and ending layers of convolutional layers with a skip connection. In inverted residual blocks, the inverses of residual layers are utilized. Initially, the channels are widened with point-wise convolution, and then the numbers of parameters are minimized with depth-wise convolution. At the beginning and end of the block 1×1 convolution is added to minimize the number of channels. In between channels, the interdependency is improved with channel recalibration instead of weighting the channel. The network itself dynamically assigns higher channel weight to the more important channel.

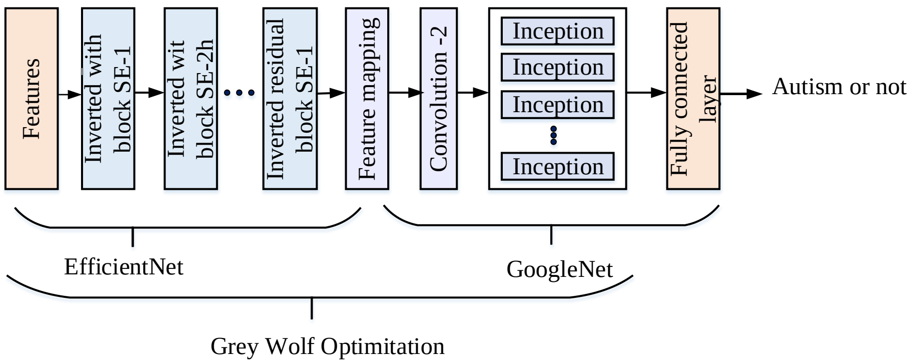

Fig. 3. Proposed architecture of GEfficient-Net with GWO for face image classification

GoogleNet is used with inception architecture, which repeatedly uses the optimal local sparse structure. Inception is used in various situations to minimize expensive convolution. High quality architecture is generated with identification capability.GoogleNet is a robust approach to providing better pattern rendering performance by extracting deep representative details. It has convolutional, fully connected inception and non-linear activation layers with a gradient decent back propagation technique. From high to low dimensional feature maps, complex and non-linear mappings are learned for better classification.

The images are represented with $y \in S^{i \times x \times d}$, and the output of classification is generated with $z = S^{i \times x \times d}$ using $z = I(y, \Theta)$. Where, the parameter set is represented with Θ. By using a convolutional kernel, the input y with bias β is estimated with the following equation.

$$z_{j,k,l} = I\left(\beta_l + \sum_{j=1}^{i}\sum_{k=1}^{x}\sum_{e=1}^{d}\Theta_{j,k,c,l} \times y_{j'+j,k'+k,e}\right) \tag{9}$$

Where, the hypothesis is computed with the non-linear function $I(.)$. The ReLU is used as an activation function.

$$z = \max\{0, y_{j,k,l}\} \tag{10}$$

The subsampling is enabled using the max-pooling layer, which estimates the channel response with the following calculation.

$$z_{j,k,l} = \max_{1<j<I,1<k<x} y_{j'+j,k'+k,l} \tag{11}$$

The final classification is enabled with the softmax classification function described as follows.

$$z_{j,k,l} = \frac{\exp(y_{j,k,l})}{\sum_{e=1}^{d}\exp(y_{j,k,l})} \tag{12}$$

3.3.1 GWO algorithm

In GoogleNet, the number of sequences and layers are optimized with the GWO algorithm. It depends on the group hunting behavior of wolves, with the main phases being tracking, pursuing, and attacking. When GWO is developed, the fittest direction is provided with the parameter β, and the second and third solution is provided with γ and η. The remaining solution is assumed as τ. The hunting behavior is operated with β, γ and η. The hunting behavior of the three wolves is followed by τ.

During the haunting behavior, the grey wolves encircle the prey, and the mathematical modeling of encircling is represented with the following equation.

$$\vec{E} = \left|\vec{D}.\vec{Y}_q(u) - \vec{Y}(u)\right| \tag{13}$$

$$\vec{Y}(u+1) = \vec{Y}_q(u) - \vec{B}.\vec{E} \tag{14}$$

Where, the current iteration is represented with $\vec{B}$, grey wolf position is represented with $\vec{Y}$, $\vec{D}$ is the coefficient, and $\vec{Y}_q$ represents the position of prey.

The values $\vec{B}$ and $\vec{D}$ are estimated with the following equation.

$$\vec{B} = 2\vec{b}.\vec{s}_1.\vec{b} \tag{15}$$

$$\vec{D} = 2.\vec{s}_2 \tag{16}$$

Where, $\vec{b}$ represents the component which is decreased from 2 to 0during each iteration. The values s_1 and s_2 denotes the random vectors within the range of $[0,1]$.

The grey wolf has the capacity to identify the prey and encircle it. The hunting behavior is guided with β and τ which has the details of the prey location. Hence, the first three solutions are obtained, and the position is updated based on the best search agent. Based on these considerations, the following equations are used.

$$\vec{E}_\beta = |\vec{D}_1.\vec{Y}_\beta - \vec{Y}|, \vec{E}_\beta = |\vec{D}_2.\vec{Y}_\gamma - \vec{Y}|, \vec{E}_\tau = |\vec{D}_3.\vec{Y}_\tau - \vec{Y}| \tag{17}$$

$$\vec{Y}_1 = \vec{Y}_\beta - \vec{B}_1.(\vec{E}_\beta), \vec{Y}_2 = \vec{Y}_\gamma - \vec{B}_2.(\vec{E}_\gamma), \vec{Y}_3 = \vec{Y}_\tau - \vec{B}_3.(\vec{E}_\tau) \tag{18}$$

$$\vec{Y}(u+1) = \frac{\vec{Y}_1 + \vec{Y}_2 + \vec{Y}_3}{3} \tag{19}$$

The grey wolf completes the hunting behavior by attacking the prey while moving. The preying technique is mathematically modeled with decreasing the value of $\vec{b}$. The range of fluctuation $\vec{B}$ also reduced with $\vec{b}$, and $\vec{B}$ is the random value within the range of $[-1,1]$. For the search agent, the next position is any place between the current and initial prey location. When $|B|<1$, the wolf attacks the prey. The searching position of the grey wolf is based on the position of β, γ and η. Based on each divergence, the prey is attacked by using $\vec{B}$. The step-by-step procedure of the GWO algorithm is given in Algorithm 1.

Algorithm 1: Optimal weight updating with GWO algorithm

Step 1. The position of Grey wolfs is initialized as $Y_j\,(j = 1,2,..p)$

Step 2. The values of b, B and D are initialized.

Step 3: The best search agent is estimated with the following computation.

Step 3.1: The first search agent Y_β is computed

Step 3.2: The second best search agent $Y\gamma$ is estimated.

Step 3.3: The third best search agent Y_τ is evaluated.

Step 4: If the number of iterations is greater than the value of u, then the following steps are executed.

Step 4.1: For each search agent, the current position is updated with equation 3.7.

Step 4.2: Then the values The values of b, B and D are updated.

Step 4.3: Then, the fitness of all search agents is estimated.

Step 4.4: The search agent Y_β, $Y\gamma$, Y_τ are updated.

Step 5: The search agent with the best fitness value is returned.

The D value is considered to be the effect of approaching the prey. Based on the wolf's position, the prey weight is randomly selected to reach the wolf, which makes the process complicated. The random population is initialized, and the search process is integrated during the prediction of possible positions. Based on prey distance, the candidate solution is updated. To ensure the process of exploration and exploitation, the parameter b is minimized from 2 to 0. The candidate solution is convergent to prey when $|\vec{B}|<1$, and the candidate solution is divergent when $|\vec{B}|>1$. The GWO approach is concluded when the condition is satisfied.

3.4 Ethical considerations

When building ASD using AI technologies, there are various ethical concerns, including privacy consent and ramifications. Artificial intelligence approaches take into account the reproduction of societal biases. If the intrinsic data is related to biases and is not representative, it limits the different needs of individuals with autism. The ethical details revolve around privacy and verified consent. Personal data customization requires the full consent of the individual with ASD. Furthermore, the decision-making process is critical for those persons and difficult to comprehend. Equal access to interventions is not achieved with AI integration of the accessibility issue. It leads to an ethical evaluation of existing imbalances that must be addressed in order to ensure the accessibility and inclusion of AI advantages. In terms of cyber security, sensitive data is a critical element for the autism support system. The impact of cyberattacks on behavioral data, communication patterns, and personal information reduces the importance of strong data security measures. It protects the privacy of people with autistic conditions.

Additionally, harmful flaws in the AI system must be addressed. Individuals with autism are affected by the use of artificial intelligence techniques, data manipulation, and tampering. It is necessary to make proactive efforts to avoid the exploitation of AI technologies. Furthermore, cyber security challenges relate to ASD and AI technologies. The agreement with one system has an impact on other people's privacy. To address the essential cyber security challenges, a resilient and secure infrastructure must be established. The examination includes ethical considerations such as reducing bias, ensuring effective security measures, and improving the lives of people with autism. Ethical considerations and advances are balanced to prevent complex implications.

4. Experimental Evaluation

The proposed approach is evaluated with the working environment of Python, and the result is considered with several performance metrics, including accuracy, specificity, sensitivity, G-mean, etc. VGG16 has three fully connected layers and 13 convolutional layers. The network width is formulated with 16 layers, and it does not have an issue related to the exploding gradient or vanishing gradient. Rather than considering a large number of hyperparameters, it uses the max pool layer with 2×2 filter of stride 2 and the convolutional layer with 3×3 filter of stride 1. VGG-16 is modeled with 64-2048 nodes per layer, 0.0-0.5 dropout between each FC layer, learning rate of 0.01-0.00001, and Adam, SGD optimizers. The GoogleNet is modeled with a number of layers of 22, a learning rate of 0.002, mini batch size of 100, epochs of 20, iterations of 100, and a momentum constant of 0.9. The EfficientNet is modeled with 10 layers, batch size 64, dropout rate 0.4, and epochs 40. GWO is implemented with dimension 3, number of grey wolves 50, iterations 100, upper bound 10, and lower bound -10.

4.1 Dataset description

Autism Children's facial image dataset is used in this paper for the proposed ASD detection. From the Kaggle repository, the facial images of autistic and non-autistic images are gathered. There are 2936 face images collected, of which 1468 are normal, and the remaining 1468 are autistic. In the Kaggle repository, most images are obtained from Facebook pages and other websites. Facial alignment and specific image size are obtained with consistency and better quality. The images are cropped with the input dimension of $224\times 224\times 3$. The dataset images are from children in the United States and Europe. There are equal numbers of autism and non-autism images, with both males and girls represented. The image size in the dataset varies thus, it is transformed to a consistent size during the training method. To recognize face expressions, the image size is reduced to 44x44. A total of 79.5% of the dataset, or 2340 images, are used for training, comprising 1170 autism and 1170 normal images. The remaining 20.5% of the data is used for training, including 300 autism and 300 normal images.

ASD patients' emotions are recognized using real-time images. It is based on the publicly accessible Kaggle dataset (https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset). The dataset contains 35,887 images, separated into training and testing folders. It also contains seven facial expressions: sad, surprised, neutral, joyful, angry, and disgusted. In this work, images are relabelled as positive, negative, and neutral. After pre-processing, 21,513 samples are used for training and testing.

4.2 Performance metrics

The efficiency of the proposed approach is evaluated through various performance metrics. The parameters utilized are accuracy, F1-measure, sensitivity, specificity, G-mean, fall-out, miss-rate, and loss.

Accuracy: For the classification model, the efficiency is measured with accuracy. It detects the ratio of true classification from the total amount of estimation provided with the dataset. Higher accuracy produced with the classification is denoted as an efficient classifier. The result of accuracy is estimated based on the confusion matrix.

$$Accuracy=\frac{Tp+Tn}{Tp+Tn+Fp+Fn} \tag{20}$$

Where, Tp denotes the true positives, Fp represents false positives, Tn indicates true negatives, and Fn represents the false negatives.

F1-score: The F1-score metric combines the recall and precision of a proposed classifier into a unique metric by computing their harmonic mean. The higher f1-score is based on the positive values of precision and recall. The f1-score representation is denoted with,

$$F_1 - score = 2 \times \frac{\Pr ecision \times recall}{\Pr ecision + recall} \tag{21}$$

Receiver operating characteristic (ROC): It is considered a visualization tool for comparing several classifiers. ROC indicates the tradeoff between FPR and TPR for the given model. The given model's ROC curves are estimated regarding the area under the curve (AUC). The False positive rate (FPR), False rejection rate (FRR), and True positive rate (TPR) are also estimated from the confusion matrix.

Sensitivity: The ratio of real positives against positively predicted values is measured with sensitivity S_n measure.

$$S_n = \frac{Tp}{Tp + Fn} \tag{22}$$

***Specificity*:** It represents the number of negative values against those predicted as negative.

$$Sp = \frac{Tn}{Tn + Fp} \tag{23}$$

G-mean: The geometric mean is estimated with the product squared root of specificity and sensitivity. It is also referred to as a product root for class-wise sensitivity.

$$Gm = \sqrt{\frac{Tp}{Tp + Fn} \times \frac{Tn}{Tn + Fp}} \tag{24}$$

Miss-rate: The miss rate is also termed a false negative ratio. Miss rate estimates the proportion of false negatives and the summation of false negatives and true positives.

$$Mr = \frac{Fn}{Fn + Tp} \tag{25}$$

***Fault-out*:** The fall-out ratio is also termed a false positive rate. It is computed with the true positive ratio and the addition of false positives and true negatives.

$$Fo = \frac{Fp}{Fp + Tn} \tag{26}$$

Loss: It estimates the variation between the original data sample and the output obtained with the model. It is utilized to direct the model towards best convergence, and for the training data, it obtains model fitting.

4.3 Experimental results

The proposed model of ASD diagnosis is analyzed with the classification related metrics and compared with existing deep learning approaches. The proposed approach's performance is compared to CNN, ResNet, DenseNet, mobileNet, Multi-Layer Perceptron (MLP), K-Nearest Neighbor (KNN), Decision Tree (DT), Logical Regression (LR), Gradient Boost (GB), Random Forest (RF), Extended Gradient Boost (XGB), and Support Vector Machine (SVM) [25]. In previous approaches, pre-processing and feature transformation are performed before being supplied to the classifier. The missing values in the dataset are replaced with mean values. As a result, the association among the values is examined. Then, the feature transformation is used, in which several categorical variables are converted into a suitable format for improving the classifier performance.

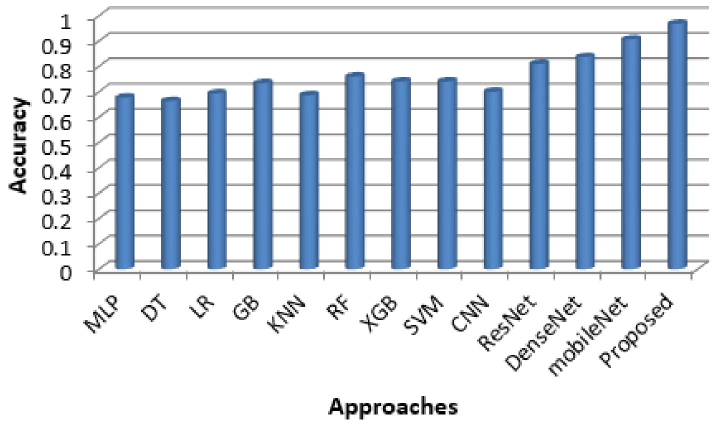

Fig. 4. Comparison of accuracy performance with machine learning models

When compared with machine learning approaches, better performance is obtained with deep learning approaches such as CNN, ResNet, DenseNet, Mobilenet, and the proposed approach in Fig. 4. The performance obtained with MLP, KNN, DT, LR, GB, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet and proposed approaches is0.6767, 0.6633, 0.6933, 0.7333, 0.6867, 0.7600, 0.7400, 0.7400, 0.700, 0.8100, 0.8367, 0.9067, and 0.9654 respectively. However, the higher accuracy is obtained with the proposed approach. The accuracy metric is lower for other existing approaches.

The f-measure performance of the proposed approach is compared with the existing deep learning and machine learning approaches in Fig. 5. The F-measure performance obtained with MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet and proposed approaches are 0.6646, 0.6631, 0.6920, 0.7331, 0.6627, 0.7599, 0.7400, 0.7399, 0.6998, 0.8082, 0.8365, 0.9067, and 0.9635. The higher value indicates the efficiency of the proposed approach.

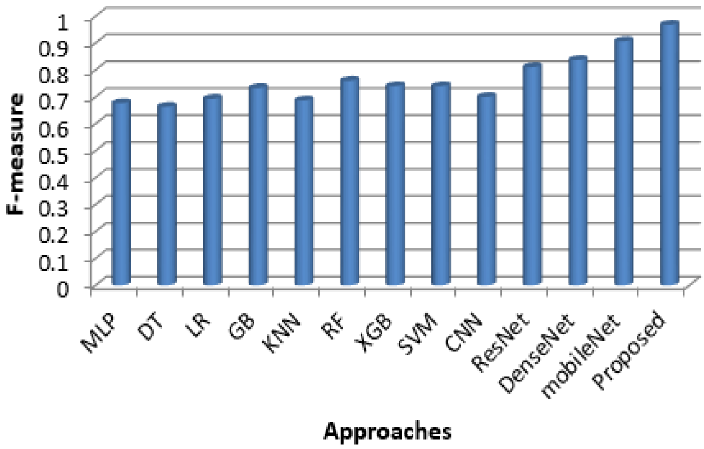

Fig. 5. F-measure performance comparison with other techniques

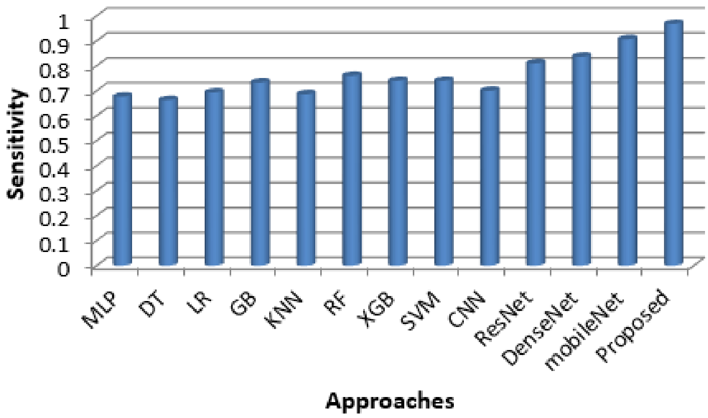

Fig. 6. Sensitivity performance comparison with other techniques

The sensitivity result is analyzed with the existing machine learning and deep learning approaches in Fig. 6. The sensitivity performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches are 0.6767, 0.6633, 0.6933, 0.7333, 0.6867, 0.7600, 0.7400, 0.7400, 0.7000, 0.8100, 0.8367, 0.9067, and 0.9615. A higher sensitivity performance is obtained with the proposed ASD detection than with traditional deep learning models. Higher sensitivity is obtained due to hybrid classification techniques and the proposed optimization algorithm for hyperparameter optimization.

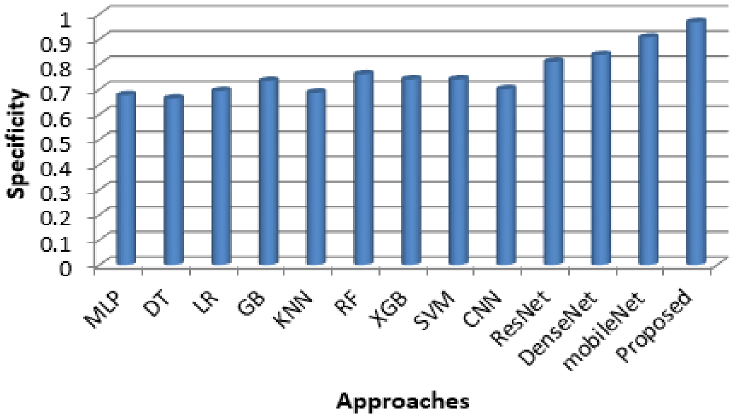

Fig. 7. Comparison of Specificity performance with existing approaches

The specificity performances are given in Fig. 7. The specificity performance is higher for the proposed approach and lower for other deep learning approaches. The increased specificity of the model is due to the influence of optimal parameter selection and the hybrid deep learning approach. The specificity performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches are 0.6767, 0.6633, 0.6933, 0.7333, 0.6867, 0.7600, 0.7400, 0.7400, 0.7000, 0.8100, 0.8367, 0.9067, and 0.9615 respectively. The hybrid deep learning and other deep learning approaches provide better performance than the machine learning based approaches.

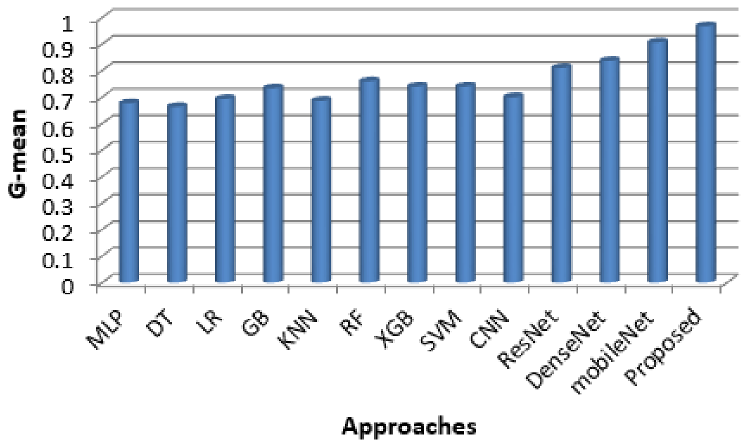

Fig. 8. Comparison of G-mean performance with existing approaches

The G-mean comparisons of the proposed model with existing machine learning techniques are given in Fig. 8. Due to the learning capability of the DNN, better performance is obtained than the baseline techniques of machine learning. The weighted mean is computed for all approaches based on the weights of the approaches. The g-mean performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches are 0.6767, 0.6633, 0.6933, 0.7333, 0.6867, 0.7600, 0.7400, 0.7400, 0.7000, 0.8100, 0.8367, 0.9067, and 0.9632 respectively. The higher result obtained with the proposed approach is due to the efficient learning procedure.

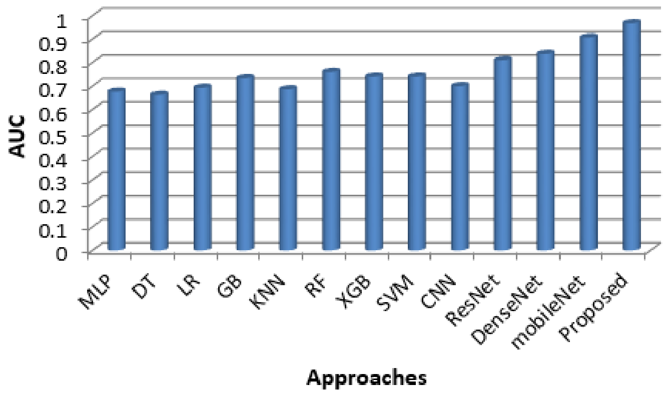

Fig. 9. Comparison of AUC performance with other classification models

The AUC performance of the proposed classification model is shown in Fig. 9. Comparing with traditional face recognition approaches, efficient recognition for autism detection is obtained with the proposed techniques. With the proposed technique, the optimal level of performance is obtained with the GWO algorithm. The AUC performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches are 0.6767, 0.6633, 0.6933, 0.7333, 0.6867, 0.7600, 0.7400, 0.7400, 0.7000, 0.8100, 0.8367, 0.9067, and 0.9678, respectively. The probability of severability was measured with this curve by classifying different classes. It is working based on varying discrimination thresholds. AUC is computed on a general function with points on the curve.

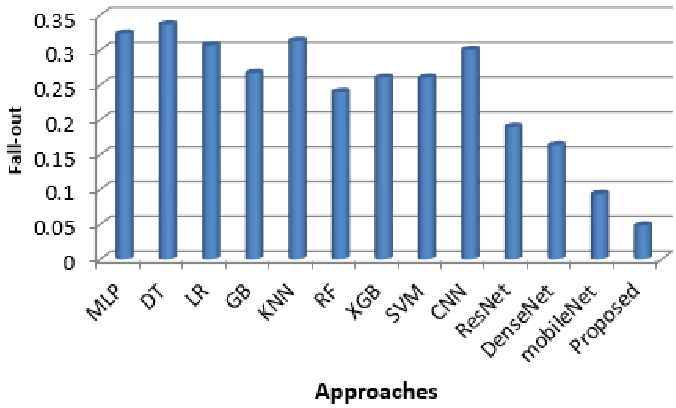

Fig. 10. Comparison of Fall-out performance with other approaches

The fall-out performance is compared with the existing machine learning as well as deep learning approaches given in Fig. 10. The fall out rate is lower for the proposed approach, but it is higher for other existing increases with other techniques. The fall out rate is higher than 0.3 for the existing approaches such as DT, MLP, and KNN. For the remaining approaches, it is lower than 0.3. The Fall-out rate performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches are 0.3233, 0.3367, 0.3067, 0.2667, 0.3133, 0.2400, 0.2600, 0.2600, 0.3000, 0.1900, 0.1633, 0.0933, and 0.0512 respectively. Improper prediction or the system's failure to predict certain aspects is termed a fall-out rate.

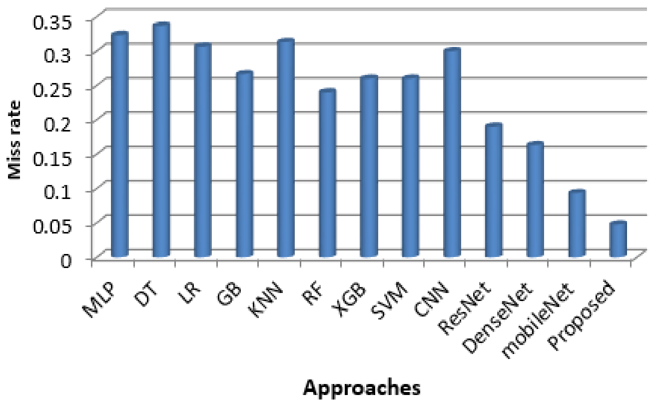

Fig. 11. Comparison of Miss-Rate performance with deep learning techniques

The miss rate performance of the proposed approach is shown in Fig. 11. When compared with standard baseline models, the performance is higher with the proposed approach. The miss-rate performance obtained with the MLP, DT, LR, GB, KNN, RF, XGB, SVM, CNN, ResNet, DenseNet, mobileNet, and proposed approaches is 0.3233, 0.3367, 0.3067, 0.2667, 0.3133, 0.2400, 0.2600, 0.2600, 0.3000, 0.1900, 0.1633, 0.0933, and 0.0478 respectively. The memory access that is not detected with the upper detection level is estimated with the miss rate.

Fig. 12 indicates the loss and accuracy of the proposed classifier during the stage of training and testing. At the training and testing stages of the children's autistic dataset with the proposed GWO-GEfficient-Net architecture, the loss and accuracy values are analyzed. With different epochs from 0 to 100, the accuracy and loss are estimated. When the epoch size is 0 to 20, the accuracy of the testing data is greater than that of the training data. From the epoch of 20 to 100, the values maintain the same level. In addition, the loss achieved with several phases is approximately the same from epoch 0 to 100. The value obtained with proposed and existing approaches of face recognition for detecting autism disorder is represented in Table 2. The performance of the proposed approach is evaluated with real–time data, as given in Table 3. The performance of the proposed approach is slightly dropped with real-time data. However, the performance is not lower than the existing algorithms.

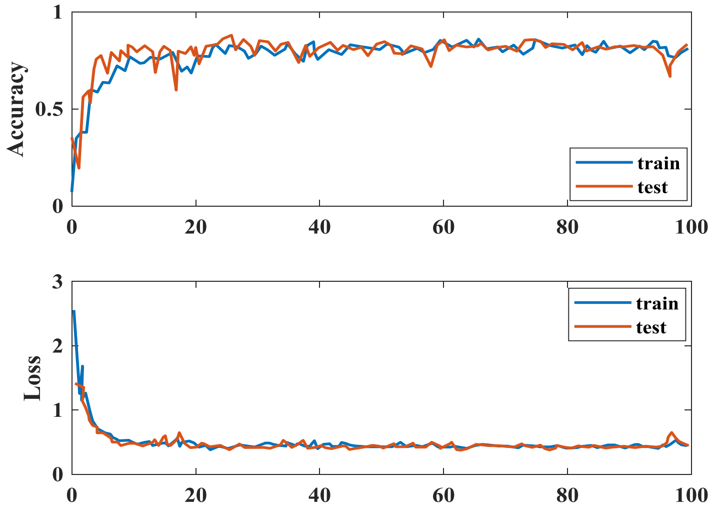

Fig. 12. Number of epochs is varied for Accuracy and loss of performance

To achieve realistic comparison, the same structures are employed with the neural network model. The ResNet, DenseNet, InceptionNet, U-net, and MobileNet models use the same pre-processing and feature extraction methodologies. The accuracy performance of these models varies between 0.8954, 0.8965, 0.8791, 0.8974, and 0.8765. However, the accuracy achieved using the identical setup of the proposed technique is 0.9654. The proposed strategy improves performance, indicating that the model is efficient. The proposed approach performance differs slightly from the configuration employed in the present model.

Table 2. ASD detection performance with different approaches (Autism Children's facial image dataset)

Classifier	Accuracy	F-Measure	Sensitivity	Fall-Out	Specificity	G-Mean	Miss Rate	AUC
MLP	0.6767	0.6646	0.6767	0.3233	0.6767	0.6767	0.3233	0.6767
DT	0.6633	0.6631	0.6633	0.3367	0.6633	0.6633	0.3367	0.6633
SVM	0.7400	0.7399	0.7400	0.2600	0.7400	0.7400	0.2600	0.7400
GB	0.7333	0.7331	0.7333	0.2667	0.7333	0.7333	0.2667	0.7333
KNN	0.6867	0.6627	0.6867	0.3133	0.6867	0.6867	0.3133	0.6867
XGB	0.7400	0.7400	0.7400	0.2600	0.7400	0.7400	0.2600	0.7400
RF	0.7600	0.7599	0.7600	0.2400	0.7600	0.7600	0.2400	0.7600
CNN	0.700	0.6998	0.7000	0.3000	0.7000	0.7000	0.3000	0.7000
ResNet	0.8100	0.8082	0.8100	0.1900	0.8100	0.8100	0.1900	0.8100
DenseNet	0.8367	0.8365	0.8367	0.1633	0.8367	0.8367	0.1633	0.8367
mobileNet	0.9067	0.9067	0.9067	0.0933	0.9067	0.9067	0.0933	0.9067
LR	0.6933	0.6920	0.6933	0.3067	0.6933	0.6933	0.3067	0.6933
Proposed	0.9654	0.9635	0.9615	0.0512	0.9648	0.9632	0.0478	0.9678

Table 3. ASD detection performance of proposed approach with real-time data

Classifier	Accuracy	F-Measure	Sensitivity	Fall-Out	Specificity	G-Mean	Miss Rate	AUC
Proposed	0.9624	0.9610	0.9601	0.0712	0.9621	0.9732	0.0878	0.9634

4.4. Limits of model generalization to ASD children

The limitations of model generalization are explored in this group to demonstrate model training for the ASD group. It can be used by altering the number of participants in the control and ASD groups (ASD=300, TD=300). To undertake model generalization, the number of people from the ASD group is limited to 200. The model's performance with imbalanced data is evaluated using the F1-score and AUC metrics. When the binary categorization class is used to deal with AUC and the F1-score, only one direction is changed. Due to its improved recognition, the proposed technique is analyzed for generalized and participant-specific models. Two-way splitting patterns are taken into account while conducting cross-model evaluations. Table 4 shows the model generalization between the children with ASD and TD.

Table 4. Model generalization between the children with ASD and TD

Techniques	**Train on TD test on TD**	**Train on TD test on ASD**	**Train on ASD test on ASD**	**Train on ASD test on TD**
ROC-AUC	0.8954	0.8574	0.8947	0.8365
F1-score	0.8910	0.8365	.0.8769	0.8452

The F1-score reveals that the model generalizes better in TD (F1-score = 0.8910) than in ASD (F1-score = 0.8769). The F1-score is reduced from the same group to the cross-group model for TD and ASD to 0.8365 and 0.8452, respectively. The deviation in the group indicates that the model generalization from TD to ASD is inefficient. Similarly, ROC data show that TD (ROC=0.8954) has more generalization than the ASD group (ROC=0.8947). In the cross-group model, the ROC results are reduced to 0.8574 and 0.8365. Furthermore, the same-group model outperforms the cross-group model. The performance variation implies a loss in generalization skills, which can be addressed by making the number of data in each class equal.

5. Conclusion

The hybrid deep learning approach with an optimization algorithm is used in the proposed approach to improve face recognition performance in ASD. The proposed technique provides efficient outcomes in terms of accuracy, specificity, sensitivity, G-mean, Fall-out rate, miss rate, and loss. Compared with other machine learning approaches, the proposed approach provides efficient detection of ASD. With the proposed approach, the accuracy, sensitivity, and specificity obtained are 0.9654, 0.9615, and 0.9648. These values are obtained with the fall-out and miss-rate of 0.0512 and 0.0478. The higher performance is obtained with the efficient feature extraction framework VGG-16. Also, the performance is enhanced with hybrid learning procedures for effective learning and optimal classification. The proposed approach has achieved accurate classification with large feature information. Also, the dimensionality reduction of the proposed model enables a compatible learning process and enhances classification. The higher performance obtained with the experimental evaluation shows the efficiency of the proposed model. Moreover, the parameters are optimized and selected in all directions for efficient processing. In the future, the performance can be improved with a large dataset of real-time applications.

References

[1] J. Manfredonia, A. Bangerter, N.V. Manyakov, S. Ness, D. Lewin, A. Skalkin, M. Boice et al. "Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder." Journal of autism and developmental disorders vol. 49, pp. 279-293, 2019.

[2] A.A. Ramirez-Duque, A. Frizera-Neto, and T.F. Bastos. "Robot-assisted autism spectrum disorder diagnostic based on artificial reasoning." Journal of Intelligent & Robotic Systems vol. 96, pp. 267-281, 2019.

[3] S.V. der Donck, M. Dzhelyova, S. Vettori, S.S. Mahdi, P. Claes, J. Steyaert, and B. Boets, "Rapid neural categorization of angry and fearful faces is specifically impaired in boys with autism spectrum disorder." Journal of Child Psychology and Psychiatry vol. 61, no. 9, pp. 1019-1029, 2020.

[4] E.M. Weiss, C. Rominger, E. Hofer, A. Fink, and I. Papousek, "Less differentiated facial responses to naturalistic films of another person's emotional expressions in adolescents and adults with High-Functioning Autism Spectrum Disorder." Progress in Neuro-Psychopharmacology and Biological Psychiatry vol. 89, pp. 341-346, 2019.

[5] S. Griffiths, C. Jarrold, I.S. Penton-Voak, A.T. Woods, A.L. Skinner, and M.R. Munafò, "Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity." Journal of autism and developmental disorders vol. 49, pp. 2768-2778, 2019.

[6] T.-L. Liu, P.-W. Wang, Y.-H.C. Yang, R.C. Hsiao, Y.-Y. Su, G.C.-W. Shyi, and C.-F. Yen. "Deficits in facial emotion recognition and implicit attitudes toward emotion among adolescents with high functioning autism spectrum disorder." Comprehensive psychiatry vol. 90, pp. 7-13, 2019.

[7] N.A. Shanok, N.A. Jones, and N.N. Lucas. "The nature of facial emotion recognition impairments in children on the autism spectrum." Child Psychiatry & Human Development vol. 50, pp. 661-667, 2019.

[8] Z. Wang, J. Liu, K. He, Q. Xu, X. Xu, and H. Liu. "Screening early children with autism spectrum disorder via response-to-name protocol." IEEE Transactions on Industrial Informatics vol. 17, no. 1, pp. 587-595, 2019.

[9] A.A. Al-Jubouri, and I.H. Ali. "A Survey On Movement Analysis (Hand, Eye, Body) And Facial Expressions-Based Diagnosis Autism Disorders Using Microsoft Kinect V2." Compusoft Int. J. Adv. Comput. Technol vol. 9, pp. 3566-3577, 2020.

[10] E. Shephard, B. Milosavljevic, L. Mason, M. Elsabbagh, C. Tye, T. Gliga, E.J.H. Jones et al. "Neural and behavioural indices of face processing in siblings of children with autism spectrum disorder (ASD): A longitudinal study from infancy to mid-childhood." cortex vol. 127, pp. 162-179, 2020.

[11] Y. Fang, H. Zhang, Y. Zuo, W. Jiang, H. Huang, and J. Yan. "Visual attention prediction for Autism Spectrum Disorder with hierarchical semantic fusion." Signal Processing: Image Communication vol. 93, pp. 116186, 2021.

[12] K. Safar, M.M. Vandewouw, and M.J. Taylor. "Atypical development of emotional face processing networks in autism spectrum disorder from childhood through to adulthood." Developmental cognitive neuroscience vol. 51, pp. 101003, 2021.

[13] K. Wang, B. Zhang, and Y. Cho. "Using mobile augmented reality to improve attention in adults with autism spectrum disorder." In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1-9, 2020.

[14] A. Kuno‐Fujita, T. Iwabuchi, K. Wakusawa, H. Ito, K. Suzuki, A. Shigetomi, H. Kosaka, M. Tsujii, and K.J. Tsuchiya. "Sensory processing patterns and fusiform activity during face processing in autism spectrum disorder." Autism Research vol. 13, no. 5, pp. 741-750, 2020.

[15] D. Bakheet, and K. Maharatna. "Linear and non-linear analysis of intrinsic mode function after facial stimuli presentation in children with autism spectrum disorder." Computers in Biology and Medicine vol. 133, pp. 104376, 2021.

[16] B. Helfer, S. Boxhoorn, J. Songa, C. Steel, S. Maltezos, and P. Asherson. "Emotion recognition and mind wandering in adults with attention deficit hyperactivity disorder or autism spectrum disorder." Journal of psychiatric research vol. 134, pp. 89-96, 2021.

[17] J.R. Sweigert, T.S. John, K.K. Begay, G.E. Davis, J. Munson, E. Shankland, A. Estes, S.R. Dager, and N.M. Kleinhans. "Characterizing olfactory function in children with autism spectrum disorder and children with sensory processing dysfunction." Brain Sciences vol. 10, no. 6, pp. 362, 2020.

[18] F. Yasuno, M. Makinodan, M. Takahashi, K. Matsuoka, H. Yoshikawa, S. Kitamura, R. Ishida et al. "Microstructural anomalies evaluated by neurite orientation dispersion and density imaging are related to deficits in facial emotional recognition via perceptual‐binding difficulties in autism spectrum disorder." Autism Research vol. 13, no. 5, pp. 729-740, 2020.

[19] F. Shic, Q. Wang, S.L. Macari, and K. Chawarska. "The role of limited salience of speech in selective attention to faces in toddlers with autism spectrum disorders." Journal of Child Psychology and Psychiatry vol. 61, no. 4, pp. 459-469, 2020.

[20] R. Sadik, S. Anwar, and M.L. Reza. "Autismnet: Recognition of autism spectrum disorder from facial expressions using mobilenet architecture." International Journal vol. 10, no. 1, pp. 327-334, 2021.

[21] A. Kalaiselvi, S. Nagarathinam, T.D. Paul, and M. Alagumeenaakshi. "Detection of autism spectrum disorder using transfer learning." Turkish Journal of Physiotherapy and Rehabilitation vol. 32, no. 2, pp. 926-933, 2021.

[22] Q. Zhang, R. Wu, S. Zhu, J. Le, Y. Chen, C. Lan, S. Yao, W. Zhao, and K.M. Kendrick. "Facial emotion training as an intervention in autism spectrum disorder: A meta‐analysis of randomized controlled trials." Autism Research vol. 14, no. 10, pp. 2169-2182, 2021.

[23] R.N.S. Husna, A.R. Syafeeza, N.A. Hamid, Y.C. Wong, and R.A. Raihan. "Functional magnetic resonance imaging for autism spectrum disorder detection using deep learning." Jurnal Teknologi vol. 83, no. 3, pp. 45-52, 2021.

[24] K. Kovarski, F. Caetta, M. Mermillod, C. Peyrin, C. Perez, L. Granjon, R. Delorme, A. Cartigny, T. Zalla, and S. Chokron. "Emotional face recognition in autism and in cerebral visual impairments: In search for specificity." Journal of neuropsychology vol. 15, no. 2, pp. 235-252, 2021.

[25] T. Akter, M.H. Ali, M.I. Khan, M.S. Satu, M.J. Uddin, S.A. Alyami, S. Ali, A.K.M. Azad, and M.A. Moni. "Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage." Brain Sciences vol. 11, no. 6, pp. 734, 2021.

[26] J.M.M. Torres, T. Clarkson, K.M. Hauschild, C.C. Luhmann, M.D. Lerner, and G. Riccardi. "Facial emotions are accurately encoded in the neural signal of those with autism spectrum disorder: A deep learning approach." Biological Psychiatry: Cognitive Neuroscience and Neuroimaging vol. 7, no. 7, pp. 688-695, 2022.

Authors' Profiles

Ramachandran Vedantham is an engineering educator since two decades and a proven researcher in Computer Science engineering in different domains. He did his B.Tech from Andhra University and M.Tech from JNTUK and PhD from Acharya Nagarjuna University, in Computer Science and Engineering. His research areas include Computer Vision, Medical Image Retrieval as major.

M. Ravisankar, working as an Associate Professor in Department of Computer Science and Engineering, K.L. deemed to be University, Guntur dist., Andhra Pradesh, India. He has 24 years of teaching experience. Dr. Ravi has received Excellence in Research Award and Best Senior Faculty Award. He has Published 5 Patents and he has published more than 20 articles in Scopus, SCI, WOS and other prominent International Journals., His areas of specializations are Data Mining and Artificial Intelligence.

Sivaiah Bellamkonda received doctoral degree from National Institute of Technology, Tiruchirappalli in December 2020. He received Batchelor and master's degrees from Jawaharlal Nehru Technological University, Kakinada in 2007 and 2010 respectively. Currently he is working as assistant professor in computer science & engineering department at Indian Institute of Information Technology Kottayam. His research interests spawn around the domains like computer vision, machine learning and image processing.

E. Srinivasa Reddy is a Professor & Dean R&D of Dr.YSR University College of Engineering of Acharya Nagarjuna University. He did Master of Technology in Computer Science Engineering from Sir Mokhsagundam Visweswariah University, Bangalore. He also did Mater of Science from Birla Institute of Technological Sciences, Pilani. He did Philosophical Doctorate in Computer Science Engineering from Acharya Nagarjuna University, Guntur. Prof.E.S.Reddy, guided successfully more than twenty five research scholars for PhD degree. His research areas of interest are in various domains including Digital Image Processing.

How to cite this paper: Ramachandran Vedantham, Ravisankar Malladi, Sivaiah Bellamkonda, Edara Sreenivasa Reddy, "Disorder Facial Emotion Recognition with Grey Wolf Optimized Gefficient-Net for Autism Spectrum", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.17, No.1, pp. 1-16, 2025. DOI:10.5815/ijigsp.2025.01.01