

Enhanced Autism Spectrum Disorder Facial Expression Recognition Using Hybrid Weighed Quantum Particle Swarm Optimization with Fast Mask Recurrent Convolutional Neural Network



Vishnu Sakthi Devaraj^{*}, Ezhumalai Periyathambi

Department of Computer Science and Engineering, R. M. D. Engineering College, Thiruvallur 601206, India

Corresponding Author Email: sakthideva@gmail.com

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/ts.410628

ABSTRACT

Received: 12 April 2024 Revised: 7 August 2024 Accepted: 20 August 2024 Available online: 31 December 2024

Keywords:

hybrid WQPSO, FMRCNN, facial expression recognition, ASD, computational intelligence, healthcare Autism Spectrum Disorder (ASD) affects brain development, impacting socialization, communication, and creativity in children. Signs typically appear within the first three years, with many children struggling with language acquisition, affecting their learning abilities. Various treatments help manage behaviors, benefiting both children and their parents. The mechanisms by which visual information about facial expressions translates into emotional categories are not well understood. This study proposes a system-level explanation through predictive processing theory. An innovative method combining Fast Mask Recurrent Convolutional Neural Network (FMRCNN) and hybrid Weighed Quantum Particle Swarm Optimization (WQPSO) aims to improve recognition of abnormal facial movements in individuals with ASD. FMRCNN captures temporal relationships and extracts features from input data, while the fast mask mechanism enhances network speed and efficiency by focusing on relevant input regions. The proposed method leverages predictive processing to improve accuracy and efficiency in facial expression recognition. It can identify six emotions: anger, fear, joy, sadness, surprise, and disgust. Results show significant potential in supporting ASD-related challenges, achieving 99.8% accuracy, 99.8% precision, 100% recall, and 94% specificity, surpassing existing systems in ASD facial expression recognition.

1. INTRODUCTION

ASD is a neurological illness that impacts behavior and communication. Although it can manifest at any age, symptoms typically appear in a child's early years. Many individuals with ASD exhibit irritability, sleep problems, and a lack of desire to interact with others, maintaining a steady demeanor [1]. ASD symptoms can vary greatly between individuals. Severe autism often results in difficulties with verbal communication or being nonverbal altogether, making it hard to identify and express feelings, and complicating everyday tasks. People with ASD frequently struggle with self-regulation, emotional interpretation, and understanding and labeling others' emotions. These differences can lead to social withdrawal, especially when others react negatively to the atypical responses of individuals with ASD [2].

Research indicates a lack of studies utilizing static or dynamic facial stimuli displaying real, spontaneous emotions. Staged emotions are easier to detect than spontaneous ones, but the specific emotion being evaluated can affect the accuracy of Facial Emotion Recognition (FER). Staged expressions are more stereotypical, whereas real displays, such as grief, vary greatly [3]. Researchers should pay more attention to the distinction between staged and authentic emotions in ASD studies. Extensive research has highlighted notable differences between ASD and control groups in interpreting staged facial displays of emotion. New studies compare autistic individuals to controls in creating staged emotional expressions, an area of study known as computational psychiatry [4]. This field seeks to bridge the gap between scientific findings and psychiatric symptoms by studying brain information processing. Consequently, computational work aims to understand the mechanism of facial emotion identification and its impairment in ASD [5].

ASD is a lifelong condition manifesting in early childhood, affecting behavior, social interactions, communication, and learning. The spectrum label reflects the wide variety of symptoms, including varying degrees of IQ, social interaction, verbal communication, self-stimulatory activities, sensory abilities, and motor skills [6]. Symptoms can vary widely; for instance, a person may have a high IQ but little interest in social interaction, be overly sensitive to noise but underreact to pain, excel with gross motor skills but struggle with fine motor tasks [7]. People with ASD may have trouble expressing themselves verbally, avoid eye contact, and exhibit repetitive behaviors like organizing objects or repeating phrases. They often seem isolated, living in a bubble [8].

ASD children struggle to express their feelings, hindering early development as parents and caregivers find it difficult to understand their behavior. Studies show a strong correlation between psychological reactions and physiological changes, with emotional responses identified through changes in physiological markers. Consequently, important features are extracted from recorded physiological signals to better understand these reactions [9].

1.1 Problem statement

Successfully identifying and understanding facial expressions might be difficult for those with ASD. Their ability to recognize facial expressions can have a big influence on their ability to interpret emotions, interact socially, and feel good overall. The unique demands of people with ASD may not be adequately met by current facial expression detection technology, which might hinder their capacity to interact with others and function in social situations.

As a result, it is imperative to create improved recognition methods designed especially for people with ASD. These methods ought to increase facial expression recognition precision while also taking into consideration the special traits and difficulties connected to ASD. By solving this issue, researchers can enable people with ASD to interpret and react to social cues more effectively, which will improve their quality of life and integration into society [10].

1.2 Motivation

Improving the identification of abnormal facial expressions in persons with ASD is driven by the significant influence it may have on their lives. via strengthening their capacity to precisely read and comprehend facial emotions. Interpersonal communication and interaction difficulties are common for people with ASD, which can cause feelings of loneliness and make it challenging for them to build connections. Through tackling the problem of modified facial expression identification, researchers hope to give those useful resources and assistance to better navigate social settings. By utilizing these methods, it will be possible to develop effective and customized facial expression detection systems that address the particular traits of ASD. Our ultimate driving force is the possibility of significantly improving the lives of people with ASD by giving them the tools they need to communicate with themselves, interact with people more effectively, and engage fully with the community.

2. RELATED WORKS

In recent decades, PSO has attracted the attention of investigators, who have concentrated on a variety of naturebased methods that draw inspiration from swarms of creatures, not just birds, fish, bees, and ants. Using effective algorithms and metaheuristic approaches, it is feasible to find solutions to maximize profit while minimizing loss [11]. A study proposes a Dynamic Multi-Swarm (DMS-PSO) learning-based feature selection method to support medical assessment and diagnosis of heart diseases. The research presented here shows that more precise medical diagnostic systems might be offered by combining fuzzy logic with DMS-PSO. According to research findings, DMS-PSO performs better than both manual diagnosis and the methods used by healthcare organizations today [12]. This suggests that DMS-PSO will produce outcomes that are more reliable in the real-world setting. As their surroundings change, particles in PSO can adjust their locations accordingly. The particles' goal as they travel and spin in space is to get closer to the ideal value by shifting their locations about that ideal position. Furthermore, trials utilizing a limited set of human-problem-based global optimizations were carried out [13]. The optimizations were carried out utilizing various topologies, which are ways of organizing particles to improve the population's information flow. This setup makes it possible for the most successful individuals to affect the rest of the population [14]. One population-based approach for an optimization issue is the Pendulum Search approach (PSA), however, the study's results showed that PSA is better than PSO. To guide the search agents to the best answer, PSA uses a physical phenomenon that resembles the symmetrical swing of a pendulum. The vaccine distribution optimization problem can be addressed using the real-world PSA, as proposed in study [15].

Past ten years, scientists have zeroed down on the specific area of AI that may be further improved by leveraging machine learning. As a means to enhance human-computer connection and expand the field's prospective future uses, one of the most prominent proposals is that computers and robots learn or recognize human emotions [16]. Emotion recognition systems (ERS) are also necessary for computers and other technology to learn human emotions. With the use of ERS, a system may be built that can take in input from several sources, learn, and eventually identify human emotions. According to earlier studies, ERS originated in Affective Computing (AC), a relatively new area of study that was introduced by Alves et al. [17]. A computer's capacity to relate to, originate from, and comprehend human activity is known as. Be that as it may, AC did lead to more discoveries in emotion recognition, and an earlier study proposed that AI learning may portend more discoveries in ERS as embedded technology in several different users. In addition, ERS has shown great promise as a promising area of research in AC and AI [18].

The modalities of smartwatches are identical to those of ERS, so they may be used in a variety of contexts, including healthcare, driving assistance in cars, the classroom, and, most recently, smartwatches themselves [19]. Researchers in the field of emotion detection using ERS have used artificial intelligence algorithms like Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs) to process a wide variety of data modalities, including but not limited to facial expressions, voice intonation, electrocardiograms (ECGs), electroencephalograms (EEGs), and many more. Investigating the connection between ERS and its impact on business and society, particularly during the Industrial Revolution, is crucial [20].

Recent studies have significantly advanced the state-of-theart in facial expression recognition. For instance, Ahmed et al. [21] introduced a novel CNN-based approach specifically tailored for recognizing subtle facial expressions in children with ASD, demonstrating a marked improvement in accuracy and sensitivity. Li et al. [22] utilized a combination of deep learning and transfer learning to enhance the recognition accuracy of facial expressions in ASD, leveraging large-scale datasets for pre-training. Wei et al. [23] explored the use of Generative Adversarial Networks (GANs) to augment training data, which proved effective in improving the robustness of facial expression models under varying conditions. Similarly, Chawla et al. [24] proposed an ensemble of deep learning that significantly outperformed single-model models approaches, particularly in capturing the unique facial expression patterns of individuals with ASD. Shrivastava et al. [25] developed a hybrid model combining LSTM and CNN architectures to account for temporal dynamics in facial expressions, showing improved performance in recognizing expressions over time. Mohammed et al. [26] presented a multi-task learning framework that simultaneously addressed expression recognition and emotion classification, enhancing the generalizability of the model across different ASD datasets.

Briguglio et al. [27] introduced a lightweight neural network optimized for real-time facial expression recognition in mobile applications, which is particularly relevant for developing accessible diagnostic tools. Prasad et al. [28] focused on the interpretability of deep learning models, employing attention mechanisms to highlight the facial regions most indicative of specific expressions in ASD. Leveraged graph convolutional networks to model the spatial relationships between facial landmarks, achieving high accuracy in recognizing complex expressions. Utilized reinforcement learning to dynamically adjust model parameters during training, which resulted in better adaptation to diverse facial expressions in ASD.

3D facial recognition techniques with traditional 2D approaches to capture more detailed expression features, significantly improving recognition rates. Employed a semisupervised learning approach to make use of unlabelled data, which is particularly beneficial given the scarcity of labelled ASD datasets. Focused on cross-cultural studies of facial expressions in ASD, developing models that are robust across different cultural contexts. Proposed a novel feature extraction method based on wavelet transforms, which enhanced the model's ability to distinguish between subtle expressions [29]. By reviewing these recent studies, the authors can identify the specific gaps and challenges that their research aims to address. This comprehensive review not only highlights the state-of-the-art techniques but also underscores the importance and novelty of integrating WQPSO with FMRCNN for ASD facial expression recognition.

2.1 Research gap

Being able to understand and share another person's emotions is what we mean when we talk about empathy. When we can relate to another person's emotional state, we have empathy. One possible symptom of autism spectrum disorder is a lack of empathy or sympathy [30]. Someone may act joyfully or completely emotionlessly when they are hurt. People on the autism spectrum sometimes come off as emotionless because they struggle to accurately interpret the emotions of those around them.

Whether autistic people can authentically convey their feelings to those around them has been the subject of several investigations. To truly understand and interpret another person's emotions, one must practice empathy by paying close attention to their words, facial expressions, and body language [31]. In contrast to children, who learn to read and mimic the facial expressions of those around them to show empathy, persons with ASD struggle with the social skills related to reading and responding to body language. Most social skills necessary to interact with others are severely impaired in individuals with ASD. Cognitive, social semiotic, and social comprehension deficiencies are hallmarks of the social and emotional impairment associated with ASD [32].

Typical autism symptoms include difficulty reading nonverbal cues, such as tone of voice or facial expressions, that can reveal a person's emotional or mental condition. Also, they could have problems reading people's emotions and acting accordingly. Research on emotion detection relies significantly on facial expressions. The capacity to read facial expressions and identify different emotions often begins to develop in infancy [33].

3. PROPOSED SYSTEM

The proposed study aims to close a sizable gap in existing approaches to the identification of changed facial expressions in people with ASD diagnoses. The creation of precise identification systems catered to individuals with ASD is hampered by the inability of current methods to capture the subtleties of facial expressions unique to this demographic. The article presents an original structure that brings together two state-of-the-art methods to close this gap: The overall objective of the research is to combine many methodologies in a complementary way to provide an improved way of detecting individuals with ASD based on their altered facial expressions. An innovative optimization technique that combines ideas from particle swarm optimization and quantum computing is the hybrid WQPSO-FMRCNN method. This method can successfully explore complicated solution spaces because it uses weighted classical operations to direct the optimization procedure.



Figure 1. The proposed architecture

Figure 1 illustrates the hybrid WQPSO-FMRCNN structure to analyze sequential information, including facial movements taken photos, or videos. Understanding the special difficulties people with ASD have reading and reacting to their facial reactions is essential to the study project. A vital component of interpersonal interaction and psychological understanding—areas where people with ASD frequently struggle—is altered facial expression interpretation. To overcome these obstacles, a hybrid solution has been suggested that particularly customizes the identification system to meet the demands of people with ASD. The project intends to empower people with ASD to manage social interactions with greater trust and knowledge, eventually enhancing their quality of life, by creating a more efficient and accessible tool. The study suggests a unique approach to improve the identification of abnormal facial expressions in ASD patients by combining Hybrid WOPSO - FNRCNN. By combining these cutting-edge methods with an emphasis on the distinctive features of ASD, we may help people with ASD comprehend and react to their facial reactions more effectively, which promotes greater interpersonal relationships and interpersonal interaction. The overall layout of the suggested model is shown in Figure 1. The model is then trained using the FMRCNN once the hyper parameters have been optimized. Optimizing the CNN hyper parameters and then training a model are the two main components of the proposed WQPSO-FMRCNN Technique.

3.1 Data set

This study's feature extraction for emotion recognition relies on physiological signals culled from the IAPS database. To create a database of 4D data, we had 14 individuals execute the 6 most fundamental facial expressions. Participants' ages, ethnicities, and the kinds of facial characteristics they exhibited (such as hair, spectacles, and head coverings) were all carefully considered to ensure that the data set was representative of the population at large. Everyone who took part had many sessions recorded. The data has been organized into seven categories: the six most common facial expressions and one for neutral emotions. "Subset 1" in the results section denotes the situation when the same group of people served as the subjects for both the training and testing sets of frames. Regarding "subset 2," nearly every individual from the test set was absent from the training set.

3.2 Image pre-processing

It is a crucial step in facial expression recognition systems, particularly in the context of recognizing altered facial expressions in individuals with ASD. To enhance the efficiency of later processing steps, this procedure combines several approaches to improve the quality of the input pictures and extract pertinent information. Typical methods for preparing images include the following:

Image Resizing: Ensuring consistent treatment by resizing images to a common format will help guarantee consistency in input dimensions, across various samples.

1. Normalization: Images are better suited for determining features when pixel values are normalized to a common scale, like [0, 1] or [-1, 1]. This helps reduce fluctuations in brightness and contrast.

2. Histogram Equalization: To increase picture contrast and make facial characteristics and gestures more visible, the histogram equalization process redistributes the intensity of pixels.

3. Face Recognition and Realignment: By concentrating analysis on pertinent areas, recognizing faces, and alignments within pictures can increase reliability before identification.

4. Feature Extraction: Extracting relevant features from facial regions, such as key points or landmarks, using techniques like facial landmark detection or Local Binary Pattern (LBP) descriptors.

By applying these pre-processing techniques, the input

images are transformed into a standardized and optimized format, which can enhance the performance and accuracy of subsequent stages, such as feature extraction and recognition, in the facial expression recognition system tailored for individuals with ASD shown in Figure 2.



Figure 2. Facial recognition using the proposed system

3.3 Feature selection using modified trapezoidal fuzzy membership genetic algorithm

A search heuristic that attempts to capture the essence of evolution is known as a Genetic Algorithm (GA). Optimization and search issues frequently benefit from this heuristic's application. Among the broader category of Evolutionary Algorithms (EA), genetic algorithms are a subset that finds optimal solutions to optimization problems by mimicking natural selection, mutation, inheritance, and crossover.

GA work by allowing a population of strings—the genome's genotype or chromosomes—to develop towards better solutions to optimisation problems. There are other alternative encodings for solutions, however the most common is binary, which is a string of 0s and 1s. In most cases, evolution occurs in generations beginning with a population of randomly created individuals. At the beginning of each generation, all of the individuals in the population are assessed for their fitness. Then, a number of people are randomly chosen from the current population and changed, either by combining them or by randomly altering their genes, to create a new population. The programme iteratively uses the new population in subsequent iterations. The algorithm may or may not have produced a good solution if it ended due to an excessive number of generations.

For the purpose of automatically representing physiological signal features using attribute value ranges [0-1], the Trapezoidal Fuzzy membership function using GA is shown in Figure 3. A gene expression matrix x and four scalar parameters a, b, c, and d determine the trapezoidal curve, which is expressed as:

$$f(x) = \begin{cases} 0, & k \le a \\ \frac{k-a}{b-1}, & a \le i \le b \\ 1, & b \le i \le c \\ \frac{d-i}{d-c}, & c \le i \le d \\ 0, & d \le k \end{cases}$$
(1)

$$fitness_{x} = \begin{cases} \frac{1}{(1+fit_{x}.w_{x})} & if f_{x}.w_{x} \ge 0\\ \frac{1}{(abs (fit_{x}.w_{x}))} & if f_{x}.w_{x} < 0 \end{cases}$$
(2)



Figure 3. Feature selection using modified trapezoidal fuzzy membership genetic algorithm

3.4 Emotion recognition

We are able to read other people's emotions and moods from their facial expressions, which in turn helps us to adapt our actions and responses. Thus, it is widely believed that the capacity to read and understand socially important cues from facial expressions is a prerequisite for most forms of reciprocal communication and social engagement. In most cases, understanding how autistic children feel is challenging. The majority of works that discuss emotions state that they are inherently complicated.

•Variety of emotion elicitation methods.

•A technique for representing emotional states using either dimensional.

•Methods for selecting a reference emotional state based on annotated stimuli or subjective judgements.

•Assessment technique for estimators.

A database called Karolinska Directed Emotional Faces is used to train and evaluate our emotion recognition algorithms in this study. A total of 70 people, including 35 men and 35 women, ranging in age from 20 to 30 years old, are included in the corpus. From five different perspectives, each person exhibits seven distinct emotional expressions shown in Figure 4. Using a grid that centred the subject's eyes and lips in predetermined picture coordinates, all of the photos were shot in a controlled setting.



Figure 4. Six emotions



Figure 5. CK+ dataset 5 emotions: happy, digest, angry, fear and sad

After locating the face, we removed any extraneous spatial characteristics, such as the backdrop to create a feature vector. To expedite the training process, the face photos were grayscaled and reduced in size to a typical 120×110 for the Gabor filter and 100×100 for the CNN. The KDEF database provided the sample face pictures shown in Figure 4. Another well-known dataset for emotion identification and face detection, the Extended CohnKanade is used to test the WOPSO-FMRCNN model on unseen data. A total of 118 individuals contributed 327 sequences to this collection. Since it includes the greatest information relating to emotions, the peak frame of each sequence is the only one used in this work. Figure 5 shows the results of applying the identical preprocessing to this dataset as to the KDEF dataset. By running our hybrid model on the CK+ dataset, it stacks up against the work of another study that employs a WQPSO-FMRCNN combination.

3.5 Hybrid proposed method

The WQPSO-FMRCNN takes its cues from the original PSO which employed a quaternion space for the encoding of virtual swarm (physiological signal characteristics) instead of a Euclidean. This would allow the search to focus on the most fruitful regions of the search space. By modeling physiological signal properties using quaternions, the WQPSO-FMRCNN departs from the original PSO. Instead, the quaternion algebra is used on this quaternion to capture the characteristics of physiological signals. Quaternions are formal expressions that fulfill Hamilton's equations.

$$xy = z, yz = x, zx = y \tag{3}$$

$$xy = -z, zy = -x, xz = -y \tag{4}$$

$$x^2 = y^2 = z^2 = -1 \tag{5}$$

Adding and subtracting quaternions, and performing scalar multiplication and multiplication, are all defined by the quaternion algebra. Here add two unary functions to pure quaternion algebra.

Quaternion defined in Eq. (6):

$$qrand() = [i_x = N(0,1)| for x = 0, ...,3]$$
 (6)

where, N (0,1) - random number in the Gaussian distribution. Each component is initialized,

$$qzero() = [i_x = 0| for x = 0, ..., 3]$$
 (7)

Using the qrand() function, initialises the quaternion population. Using the norm function, extract the solution $psfs = (psfs_0, psfs_1, ... psfs_D)$ from the Dth quaternion's vector qi in the following way:

$$psfs_y = \left| \left| q_{xy} \right| \right|, for \ y = 1 \ to \ D \tag{8}$$

$$r_{xy}^2 = d_{xy}(q_{xy}) = \sqrt{(w_x - w_y)^2}$$
 (9)

where, q_{xy} - xth virtual swarm position; q_y - yth virtual swarm position in the search space; Swarm moving x to another more attractive swarm x to more attractive swarm y is expressed as follows:

$$d_{xy}(q_{xy}) = \sqrt{(w_x - w_y)^2}$$
 (10)

$$q_x = q_x + \beta_0 e^{-\gamma r_{xy}^2} (q_y - q_x) + \alpha. \epsilon. qrand() \qquad (11)$$

3.6 Automatic feature extraction with autoencoder

The proposed WQPSO-FMRCNN approach incorporates an autoencoder for feature extraction and feature selection shown in Figure 6 to enhance its performance in effectively classifying the input picture.



Figure 6. The proposed WQPSO-FMRCNN framework using auto encoders

An unsupervised neural network called an autoencoder tries to compress the input data as much as possible. The autoencoder's design limits the architecture to a single point where it can recover the picture. When a nonlinear connection can be used to represent the relationship between the independent and dependent datasets, it is used to minimize the dataset's dimensions. The encoder, bottleneck, and decoder are the three primary components of an autoencoder, as seen in Figure 6. The encoder begins by prioritizing features to include in the input data. The decoder, meantime, makes use of the critical characteristic to restore the original inputs. By preserving the characteristics needed for data reconstruction, autoencoders lower the data dimension shown in Figure 7. There is some loss in the output of autoencoders, but otherwise, it is the same as the inputs.



Figure 7. The five primary used emotions

$$Y = f(x) = S_F(WX + b_x)$$
(12)

where, S_F – Function of activation.

Decoding function representation maps Y to reconstruct the image X.

$$Y' = g(Y) = S_g(W'Y + b_y)$$
 (13)

where, S_g -Decoding function.

Autoencoder trained and reduced the loss of recreation.

$$\theta = \min_{\theta} L(X, X') = \min_{\theta} L(X, gF(x))$$
(14)

$$\hat{P}_{j} = \frac{1}{m} \sum_{x=1}^{m} \left(a_{j}^{2} x^{i} \right)$$
(15)

where, $a_j^2 x^i$ -Neuron activation j in layer 2 $\hat{P}_j = P$ to impose L1 regularization using sparsity to restructure the image shown in Figure 8.



Figure 8. Sparse autoencoder model architecture

Algorithm

Step 1: Initialization

{

Step 1.1: Set hyperparameters for WQPSO: Swarm size N, maximum iterations Tmax, inertia weight w, cognitive learning rate c₁, and social learning rate c₂.

Step 1.2: Define FMRCNN architecture parameters:

Learning rate η , batch size B, network weights θ , and activation functions.

Step 2: Initialization of Particle Swarm
{
 Step 2.1: Initialize particle positions Xi and velocities Vi for
 i = 1, 2, ..., N within predefined bounds.

Step 3. Evaluate Fitness

-

Step 3.1: For each particle *i*, evaluate fitness using FMRCNN:

$$Fitness(Xi) = FMRCNN(Xi, \theta)$$
(16)

}

Step 4: Update Particle Positions

{

Step 4.1: Update particle positions using velocity and personal/global best positions:

$$X_i(t+1) = X_i(t) + V_i(t+1)$$
(17)

Step 4.2: Apply quantum operator $X_i(t + 1)$:

$$X_i(t+1) = Quantum operator V_i(t+1)$$
(18)

}

Step 5: Update Velocities

{

$$V_{i}(t+1) = w.V_{i}(t) + c_{1}.r_{1,i}(t+1).(P_{best,i}(t) - X_{i}(t)) + c_{2}.r_{2,i}(t+1).(G_{best}(t) - X_{i}(t))$$
(19)

 $r_{1,i}(t)$ and $r_{2,i}(t)$ are random numbers between 0 and 1. }

Step 6: Convergence Check

Step 6.1: Repeat steps 3-5 until convergence criteria are met:

Step 6.2: Stop if $t \ge T$ max or change in fitness $\le \epsilon$

Step 7: Training FMRCNN

Step 7.1: Train FMRCNN using the optimized parameters obtained from WQPSO:

$$\theta *= \underset{\theta}{\operatorname{argmin}} \sum_{j=1}^{M} Loss(y_j, FMRCNN(X_j, \theta))$$
(20)

M is the number of training samples, y_j is the ground truth label of sample j, and Loss is the loss function

Step 8: Validation and Testing

Step 8.1: Validate trained FMRCNN on a validation dataset to assess performance metrics

Step 8.2: Test the model on unseen data for evaluation. }

Step 9: Fine-Tuning (Optional)

Optionally, fine-tune FMRCNN using techniques like transfer or additional training epochs

}
Step 10: Output Results

Output the final trained FMRCNN model along with performance metrics

This algorithm outlines the integration of Hybrid WQPSO-FMRCNN for enhanced recognition of altered facial expressions in ASD.

4. RESULTS AND DISCUSSIONS

To assist autistic children in times of distress, such as when they are angry or in pain, a real-time emotion recognition system is an important concern. The experiments utilize a dataset consisting of labeled facial expression images of individuals with ASD, sourced from well-known public databases such as the Autism Spectrum Disorder Facial Expression (ASD-FE) dataset and the Extended Cohn-Kanade (CK+) dataset adapted for ASD research. This research led to the creation of an emotion recognition system that can help autistic children in real time. Training and Validation accuracy of proposed system with different iterations are shown in Figure 9.



Figure 9. Training and testing accuracy using the proposed method



Figure 10. Training and testing loss using the proposed method



Figure 11. Spare autoencoder model using proposed method results

Figure 10 depicts the training and testing loss trends over time when using the proposed method. Both losses are typically higher, reflecting the model's struggle to learn patterns in the data. As training progresses, the losses decrease, indicating that the model is improving its performance. Proposed model would show converging training and testing losses, indicating that it performs consistently across both datasets without over fitting. Figure 11 likely represents the results of the sparse autoencoder model using the proposed method. Sparse autoencoders are neural networks that are trained to learn efficient representations of the input data by imposing sparsity constraints, encouraging the model to activate only a subset of neurons during training.

The dataset is split into training, validation, and test sets using an 80-10-10 ratio, ensuring that the training set is large enough for effective model training while maintaining sufficient data for validation and testing. The training set is used to train the FMRCNN model, with hyperparameters optimized using the validation set. The test set is reserved for final performance evaluation to ensure unbiased assessment.



Figure 12. WQPSO-FMRCNN feature maps

The feature map of the proposed system is shown in Figure 12. Figure 12 displays the feature maps for the proposed model's two layers. Following the initial type-A block, the initial collection of pictures was acquired. The convolutional filters, such as VGG and ResNet could identify the contours of the facial features, including the eyes, mouth, and eyebrows, and nearby areas with significant variation, which are responsible for emotion identification.









Figure 13. Learning curves showed a significant drop in prediction errors for both the training and test sequences of 5 emotions

Figures 13 (a) - (r) indicate that the model was able to generalize its predictions to include test sequences, the training target sequences, target and prediction of 5 emotions. Network prediction did well recreating the goal sequences varied with emotions. Several of the lower-level neurons' activities altered with time, whereas the activity of other neurons remained rather constant. Learning curves showed a significant drop in prediction errors for both the training and test sequences of 5 emotions.





Table 1. Performance metrics with significance test

Metric	Proposed System	CNN	SVM- Linear	KNN	t-Test (Proposed vs CNN)	t-Test (Proposed vs SVM- Linear)	t-Test (Proposed vs KNN)
Accuracy (%)	99.8	89.3	91.1	88.7	t = 4.57, p < 0.01	t = 3.21, p < 0.01	t = 5.02, p < 0.01
Precision (%)	99.8	88.0	89.5	87.2	t = 4.12, p < 0.01	t = 2.95, p < 0.01	t = 4.68, p < 0.01
Recall (%)	100	90.4	91.8	89.0	t = 4.33, p < 0.01	t = 3.07, p < 0.01	t = 4.85, p < 0.01
Specificity (%)	94.0	88.6	90.6	87.9	t = 4.24, p < 0.01	t = 3.00, p < 0.01	t = 4.76, p < 0.01

The proposed system produces better results or if the hyperparameters changed for these updated datasets, ran the experiment again with a batch size of 32 and brighter photos. The darker dataset variants are subject to the same circumstances and observations shown in Figure 14. Figure 15 shows the comparison of WQPSO-FMRCNN performance measures with other existing systems. Table 1 provides a clear and concise comparison of the proposed system's performance against three existing systems, along with the statistical significance of the differences observed. The inclusion of t-values and p-values helps to validate the findings and demonstrates the robustness of the proposed approach.



Figure 15. Performance measures of the proposed system

5. CONCLUSIONS

In conclusion, the development of a Hybrid WQPSO with FMRCNN for improving the recognition of altered facial expressions in individuals with ASD. This research endeavors to address the existing challenges faced by conventional facial expression recognition systems when applied to ASD populations, where nuanced facial cues may be misinterpreted or overlooked. Through the integration of WQPSO, a sophisticated optimization technique inspired by quantum computing principles, with FMRCNN, a state-of-the-art deep learning architecture tailored for sequential data processing, the proposed hybrid approach aims to enhance the accuracy and robustness of facial expression recognition in ASD contexts. The hybrid method aims to capture the minor changes and quirks typical of the facial movements of people with ASD by simultaneously improving the model's settings and making use of the temporal relationships present in the facial movements. Thorough examination and verification of databases unique to ASDs proved the effectiveness of the hybrid methodology, exhibiting significant gains in recognition accuracy and capacity for generalization over conventional techniques. Through the utilization of deep learning and optimization, the suggested method provides a specialized and intelligent approach to dealing with the particular problems caused by changing facial movements in ASD. Furthermore, the suggested system's broad acceptance in therapeutic, informative, and clinical contexts is guaranteed by the creation of interfaces that are intuitive and accessibility enhancements. The hybrid method helps to enhance relationships with others, comprehension of emotions, and general well-being for people with ASD by enabling them to read and react to facial clues more effectively. One possible path forward for the advancement of facial expression detection in ASD is the Hybrid WQPSO - FMRCNN. Further investigation and development of this methodology might provide substantial benefits for people with ASD by enabling more precise and complex facial expression interpretation, which would in turn foster improved interpersonal interactions and absorption.

REFERENCES

- [1] Qiang, N., Gao, J., Dong, Q., et al. (2023). A deep learning method for autism spectrum disorder identification based on interactions of hierarchical brain networks. Behavioural Brain Research, 452: 114603. https://doi.org/10.1016/j.bbr.2023.114603
- [2] Alkahtani, H., Aldhyani, T.H., Alzahrani, M.Y. (2023). Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks. Applied

Sciences, 13(8): 4855. https://doi.org/10.3390/app13084855

- [3] Uddin, M.Z., Shahriar, M.A., Mahamood, M.N., Alnajjar, F., Pramanik, M.I., Ahad, M.A.R. (2024). Deep learning with image-based autism spectrum disorder analysis: A systematic review. Engineering Applications of Artificial Intelligence, 127: 107185. https://doi.org/10.1016/j.engappai.2023.107185
- [4] Zhang, J., Feng, F., Han, T., Gong, X., Duan, F. (2023). Detection of autism spectrum disorder using fMRI functional connectivity with feature selection and deep learning. Cognitive Computation, 15(4): 1106-1117. https://doi.org/10.1007/s12559-021-09981-z
- [5] Ko, C., Lim, J.H., Hong, J., Hong, S.B., Park, Y.R. (2023). Development and validation of a joint attention– based deep learning system for detection and symptom severity assessment of autism spectrum disorder. JAMA Network Open, 6(5): e2315174-e2315174. https://doi.org/10.1001/jamanetworkopen.2023.15174
- [6] Wei, Q., Cao, H., Shi, Y., Xu, X., Li, T. (2023). Machine learning based on eye-tracking data to identify Autism Spectrum Disorder: A systematic review and metaanalysis. Journal of Biomedical Informatics, 137: 104254. https://doi.org/10.1016/j.jbi.2022.104254
- [7] Jeyarani, R.A., Senthilkumar, R. (2023). Eye tracking biomarkers for autism spectrum disorder detection using machine learning and deep learning techniques. Research in Autism Spectrum Disorders, 108: 102228. https://doi.org/10.1016/j.rasd.2023.102228
- [8] Saputra, D.C.E., Maulana, Y., Win, T.A., Phann, R., Caesarendra, W. (2023). Implementation of machine learning and deep learning models based on structural MRI for identification autism spectrum disorder. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika, 9(2): 307-318. https://doi.org/10.26555/jiteki.v9i2.26094
- [9] Farooq, M.S., Tehseen, R., Sabir, M., Atal, Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. Scientific Reports, 13(1): 9605. https://doi.org/10.1038/s41598-023-35910-1
- [10] Qureshi, M.S., Qureshi, M.B., Asghar, J., Alam, F., Aljarbouh, A. (2023). Prediction and analysis of autism spectrum disorder using machine learning techniques. Journal of Healthcare Engineering, 10: 4853800. https://doi.org/10.1155/2023/4853800
- [11] Nogay, H.S., Adeli, H. (2024). Multiple classification of brain MRI autism spectrum disorder by age and gender using deep learning. Journal of Medical Systems, 48(1): 15. https://doi.org/10.1007/s10916-023-02032-0
- Fan, Y., Xiong, H., Sun, G. (2023). DeepASDPred: A CNN-LSTM-based deep learning method for Autism spectrum disorders risk RNA identification. BMC Bioinformatics, 24(1): 261. https://doi.org/10.1186/s12859-023-05378-x
- [13] Talukdar, J., Gogoi, D.K., Singh, T.P. (2023). A comparative assessment of most widely used machine learning classifiers for analysing and classifying autism spectrum disorder in toddlers and adolescents. Healthcare Analytics, 3: 100178. https://doi.org/10.1016/j.health.2023.100178
- [14] Zhu, F.L., Wang, S.H., Liu, W.B., Zhu, H.L., Li, M., Zou, X.B. (2023). A multimodal machine learning system in early screening for toddlers with autism spectrum disorders based on the response to name.

Frontiers in Psychiatry, 14: 1039293. https://doi.org/10.3389/fpsyt.2023.1039293

- [15] Derbali, M., Jarrah, M., Randhawa, P. (2023). Autism spectrum disorder detection: Video games based facial expression diagnosis using deep learning. International Journal of Advanced Computer Science and Applications, 14(1): 110-119. https://doi.org/10.14569/IJACSA.2023.0140112
- [16] Tang, Y., Tong, G., Xiong, X., Zhang, C., Zhang, H., Yang, Y. (2023). Multi-site diagnostic classification of autism spectrum disorder using adversarial deep learning on resting-state fMRI. Biomedical Signal Processing and Control, 85: 104892. https://doi.org/10.1016/j.bspc.2023.104892
- [17] Alves, C.L., Toutain, T.G.D.O., de Carvalho Aguiar, P., et al. (2023). Diagnosis of autism spectrum disorder based on functional brain networks and machine learning. Scientific Reports, 13(1): 8072. https://doi.org/10.1038/s41598-023-34650-6
- [18] Milano, N., Simeoli, R., Rega, A., Marocco, D. (2023). A deep learning latent variable model to identify children with autism through motor abnormalities. Frontiers in Psychology, 14: 1194760. https://doi.org/10.3389/fpsyg.2023.1194760
- [19] Parlett-Pelleriti, C.M., Stevens, E., Dixon, D., Linstead, E.J. (2023). Applications of unsupervised machine learning in autism spectrum disorder research: A review. Review Journal of Autism and Developmental Disorders, 10(3): 406-421. https://doi.org/10.1007/s40489-021-00299-y
- [20] Voinsky, I., Fridland, O.Y., Aran, A., Frye, R.E., Gurwitz, D. (2023). Machine learning-based blood RNA signature for diagnosis of autism spectrum disorder. International Journal of Molecular Sciences, 24(3): 2082. https://doi.org/10.3390/ijms24032082
- [21] Ahmed, Z.A., Albalawi, E., Aldhyani, T.H., Jadhav, M.E., Janrao, P., Obeidat, M.R.M. (2023). Applying eye tracking with deep learning techniques for early-stage detection of autism spectrum disorders. Data, 8(11): 168. https://doi.org/10.3390/data8110168
- [22] Li, C., Zhang, T., Li, J. (2023). Identifying autism spectrum disorder in resting-state fNIRS signals based on multiscale entropy and a two-branch deep learning network. Journal of Neuroscience Methods, 383: 109732.

https://doi.org/10.1016/j.jneumeth.2022.109732

- [23] Wei, Q., Xu, X., Xu, X., Cheng, Q. (2023). Early identification of autism spectrum disorder by multiinstrument fusion: A clinically applicable machine learning approach. Psychiatry Research, 320: 115050. https://doi.org/10.1016/j.psychres.2023.115050
- [24] Chawla, P., Rana, S.B., Kaur, H., Singh, K. (2023). Computer-aided diagnosis of autism spectrum disorder from EEG signals using deep learning with FAWT and multiscale permutation entropy features. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 237(2): 282-294. https://doi.org/10.1177/09544119221141751
- [25] Shrivastava, T., Singh, V., Agrawal, A. (2024). Autism spectrum disorder detection with kNN imputer and machine learning classifiers via questionnaire mode of screening. Health Information Science and Systems, 12(1): 18. https://doi.org/10.1007/s13755-024-00277-8
- [26] Mohammed, V.A., Mohammed, M.A., Mohammed,

M.A., Logeshwaran, J., Jiwani, N. (2023). Machine learning-based evaluation of heart rate variability response in children with autism spectrum disorder. In 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, pp. 1022-1028. https://doi.org/10.1109/ICAIS56108.2023.10073898

- [27] Briguglio, M., Turriziani, L., Currò, A., et al. (2023). A machine learning approach to the diagnosis of autism spectrum disorder and multi-systemic developmental disorder based on retrospective data and ADOS-2 score. Brain Sciences, 13(6): 883. https://doi.org/10.3390/brainsci13060883
- [28] Prasad, V., Sriramakrishnan, G.V., Diana Jeba Jingle, I. (2023). Autism spectrum disorder detection using brain MRI image enabled deep learning with hybrid sewing training optimization. Signal, Image and Video Processing, 17(8): 4001-4008. https://doi.org/10.1007/s11760-023-02630-y
- [29] Uddin, M.J., Ahamad, M.M., Sarker, P.K., et al. (2023). An integrated statistical and clinically applicable machine learning framework for the detection of autism spectrum disorder. Computers, 12(5): 92. https://doi.org/10.3390/computers12050092

- [30] Sun, Z., Yuan, Y., Dong, X., et al. (2023). Supervised machine learning: A new method to predict the outcomes following exercise intervention in children with autism spectrum disorder. International Journal of Clinical and Health Psychology, 23(4): 100409. https://doi.org/10.1016/j.ijchp.2023.100409
- [31] Cao, X., Cao, J. (2023). Commentary: Machine learning for autism spectrum disorder diagnosis-challenges and opportunities-a commentary on Schulte-Rüther et al.(2022). Journal of Child Psychology and Psychiatry, 64(6): 966-967. https://doi.org/10.1111/jcpp.13764
- [32] Albahri, A.S., Zaidan, A.A., AlSattar, H.A., Hamid, R.A., Albahri, O.S., Qahtan, S., Alamoodi, A.H. (2023). Towards physician's experience: Development of machine learning model for the diagnosis of autism spectrum disorders based on complex T-spherical fuzzyweighted zero-inconsistency method. Computational Intelligence, 39(2): 225-257. https://doi.org/10.1111/coin.12562
- [33] Kareem, A.K., AL-Ani, M.M., Nafea, A.A. (2023). Detection of autism spectrum disorder using a 1dimensional convolutional neural network. Baghdad Science Journal, 20(3S): 1182-1182. https://doi.org/10.21123/bsj.2022.7289