Engineered Science
DOI: https://dx.doi.org/10.30919/es1364

Optimized Deep Learning Model for Intensive Care of Neurological Disorders Patients Based on Facial Expression

Dattatray G. Takale,[1,*] Amol V. Dhumane,[2,*] Tushar Jadhav,[3] Amar Buchade,[4] Chitrakant O. Banchhor,[5] Omkaresh Kulkarni[6] and Parikshit N. Mahalle[7,*]

Abstract

Facial expressions play a significant role in nonverbal communication. Reading the facial expressions of those suffering from neurological illnesses is crucial, as they may have significantly reduced verbal communication abilities. Such an assessment requires a thorough examination by medical specialists, which can be expensive and challenging. With the help of low-cost, non-invasive, automated facial expression detection technologies, experts can diagnose neurological disorders. In order to identify the facial expressions of individuals with Parkinson's, stroke, Alzheimer's, and Bell palsy disorders, this research constructs a Fine-Tuned Deep Learning Model (FTDLM). The dataset is initially collected from well-known internet sites. Additionally, using publicly accessible sources, raw photos of the patient's most common facial expressions, such as usual, happy, sad, and angry are gathered. Finding out if it was feasible to identify individual differences when searching for Parkinson's disease symptoms was the aim of the data analysis. Cropping is then taken into consideration in order to alter the image from the input image. Subsequently, the preprocessing method employing a Gaussian filter is examined to eliminate noise. Using FTDLM, the pre-processed image is used to classify the emotions. The New Convolutional Neural Network (NCNN) and the Enhanced Golden Search Algorithm (EGSA) are combined in this suggested model. EGSA is used in the NCNN to pick the hyperparameters. The suggested approach is carried out in Python, and statistical measures of accuracy, sensitivity, specificity, recall, and precision are used to assess performance. Furthermore, it is in contrast to traditional methods.

Keywords: Optimized deep learning model; Convolutional neural network; Enhanced golden search algorithm; Hyperparameter selection.
Received: 15 October 2024; Revised: 20 December 2024; Accepted: 26 December 2024.
Article type: Research article.

1. Introduction

Facial expressions are crucial in both human relationships and behavioral science. Additionally, common symptoms of a wide range of illnesses, including neurological conditions including Parkinson's, stroke, Alzheimer's, and Bell Palsy, can be seen in facial expressions.[1] Medical personnel typically use a combination of expensive, invasive procedures that can be uncomfortable and difficult to administer, along with strict overtime monitoring to identify neurological illness patients. Developing a unique, reasonably priced, and reliable solution is critical.[2] A clinician can use an automatic facial expression recognition system to evaluate patients' overall behavior if they have neurological issues.[3]

Rapid alterations in facial muscle tone provide information about intent, drive, and inner emotional states, frequently in interpersonal interactions.[4] Various neurologic and mental diseases can alter humans' tendency to use facial cues. Alzheimer's,[5] Parkinson's disease, and Bell palsy are a few examples of such disorders. Parkinson's disease is known to cause a flat, expressionless face, also known as a "masked face," which impairs both voluntary and involuntary expressions.[6] Pentland and colleagues reported more than 30 years ago that, despite no differences in objective mood indices,[7] healthcare workers consistently judged patients with Parkinson's disease as more suspicious, nervous, and depressed than cardiac patients.[8] More recent research has

[1] *Department of Computer Engineering, BRACT's Vishwakarma Institute of Information Technology, Pune, 411048, India*
[2] *Symbiosis Institute of Technology, Pune, 412115, India*
[3] *Department of Electronics and Telecommunication Engineering, BRACT's Vishwakarma Institute of Information Technology, Pune, 411048, India*
[4] *Department of Artificial Intelligence and Data Science, BRACT's Vishwakarma Institute of Information Technology, Pune, 411048, India*

demonstrated the adverse effects of hypomimia on interpersonal connections with carers and quality of life.[9,10]

In order to effectively communicate nonverbally and engage in social interactions, one must be able to detect and accurately identify emotions in facial expressions.[11] Thus, a diminished ability to recognize facial emotions may result in issues with forming interpersonal relationships, which is linked to mental health issues and a lower standard of living.[12] Following deep brain stimulation (DBS) surgery, patients may experience challenges in their interpersonal and professional interactions. This could be attributed to abnormalities in facial emotion recognition following surgery. Patients with diseases commonly show deficits in both their ability to mimic and recognize emotional facial expressions,[13] making them an excellent model to study the impact of facial mimicry on the processing of facial emotions. This study examines how manipulating facial mimicry affects the ability to identify emotional expression variations in healthy individuals and patients with illnesses.[14] In line with earlier research, we anticipated a general impairment in identifying emotional expression alterations in patients suffering from different ailments. Additionally, we hypothesized that altering facial mimicry affects patients with different illnesses and healthy controls regarding change detection.[15]

The field of affective computing has come a long way in the last twenty years in creating innovative technology that can simulate emotions and facial expressions. Developing computational models that can identify, understand, and process human emotions is one of the objectives of effective technologies, as it will increase the utility of human-computer interaction. In this work, ideal deep learning models are created to achieve it. An optimization algorithm also empowers the deep learning model and improves its performance.

The research paper's primary contribution and structure are as follows:

- This work creates a fine-tuned deep learning model (FTDLM) to recognize patients' facial expressions (Parkinson's, Stroke, Alzheimer's, and Bell Palsy disorders).
- The data analysis aimed to ascertain whether individual differences might be identified when searching for Parkinson's disease symptoms. Next, cropping is considered to resize the image from the input image. Next, the preprocessing method, which uses a Gaussian filter, is examined to eliminate noise.
- The pre-processed image is used to classify emotions using the FTDLM. This suggested model combines a new convolutional neural network (NCNN) and an enhanced golden search algorithm (EGSA). The EGSA is used in the NCNN to choose the hyperparameters.

Parkinson's disease, stroke, Alzheimer, and Bell palsy affect the facial muscles and thus affect the bodily language. Hence, a client's actual psychological and neurological condition might not be accurately determined. While medical professionals are capable of evaluating patients manually and with a high degree of success, the overall process is time-consuming, imprecise, and often costly, and therefore unavailable to most patients. This points to a need for automated methods that can perform facial expression recognition in a cost-sensitive, non-invasive manner and achieve consistent, valid results. These systems have the capabilities of improving a diagnostic process significantly, as well as early identification of neurological disorders and appropriate management. In addition, they are not only very effective for identification purposes. Still, they are also central to essential service areas such as mental health and patients' evaluation of pain, providing a concrete insight as to whether the patient is emotionally stable and, respectively, how much pain they are enduring. Technological advancement in deep learning, especially in facial expression, still offers an optimized model that incorporates the most recent methods in hyperparameter optimization for facial expressions and accurate techniques to decipher the subtle changes in image formation. This result laid the foundation for a radical change in patient care and treatment plans.

Over the years, numerous studies have explored techniques for identifying patient emotions based on facial expressions, particularly in medical and surveillance contexts. These efforts have leveraged advanced computational approaches such as convolutional neural networks (CNNs), autoencoders (AEs), and preprocessing techniques to address challenges like noise, lighting variations, and complex facial dynamics.

In the study, Huang *et al.*[16] A strategy was proposed for diagnosing auto-Parkinson's Disease (auto-PD) based on mixed emotional facial expressions. The suggested approach was divided into the following four steps: In steps one, using generative adversarial learning to create virtual face images with six fundamental expressions (happiness, disgust, surprise, sadness, and fear) to mimic the premorbid expressions of Parkinson's disease (PD) patients. In second Step, create an efficient screening system to evaluate the quality of the previously created synthetic facial expression photos, then select the best ones for a shortlist; in third Step, utilize a combination of the original PD patient facial expression images, the high-quality synthesized PD patient facial expression images, and in Fourth step, the typical facial expression images from other public face datasets to train a deep feature extractor and facial expression classifier.

Davodabadi *et al.*[17] have outlined two methods for making

[5] *Department of Artificial Intelligence, Vishwakarma Institute of Information Technology, Pune, 411037, India*

[6] *Department of AI & ML, BRACT's Vishwakarma Institute of Information Technology, Pune, 411048, India*

[7] *Research and Development, Vishwakarma Institute of Technology, Pune, 411037, India*

**Email:* dattatraygtakale@gmail.com (D. Takale), amol.dhumane@sitpune.edu.in (A. D. Dhumane), parikshit.mahalle@vit.edu (P. N. Mahalle)

a mental health diagnosis. The initial technique is a support vector machine. This method extracts the properties of an image using a fractal model for classification. Using this method, the Gaussian distribution is used to simulate an image's histogram. The categorization results using various support vector machine kernels were compared. In order to identify mental impairments associated with Alzheimer's disease, step two suggests employing a deep convolutional neural network design.

Xu *et al.*[18] have demonstrated using human-computer interaction to display a facial video-based dataset to identify Alzheimer's disease. Alzheimer's disease was considered when designing this interactive test, which measures visual space capacity, attentiveness, facial apraxia, and changes in facial expressions when a task is successful or unsuccessful. One hundred two faces video footage, particular task scores, and emotional self-evaluation comprise the final dataset, which was gathered using this task as the tool. The 5-fold cross-validation method was utilized for feature extraction and classification using random forest (RF) and the enhanced local binary pattern on three orthogonal planes to assess the baseline. Sarang *et al.*[19] have demonstrated the efficiency of a recently created emotion detection and grading system (EDGS) program in recognizing and rating the expressions of people living with Parkinson's disease and healthy individuals. EDGS was a deep learning-based software for detecting and grading emotions. It was designed to identify seven fundamental emotional expressions and grade the degree of happiness and sadness in each expression. It was discovered that both regular people and Parkinson's disease patients could accurately identify feelings of surprise and happiness using the EDGS program.

Ferrer-Cairols *et al.*[20] analyzed the correlation between participants' cortisol levels and the emotion recognition test in individuals with early-stage Alzheimer's disease (AD). The reading the mind in the eyes test (RMET) helps assess emotion recognition, which is frequently compromised in the early stages of AD. Cortisol levels may also function as a biomarker for AD and have an impact on cognition. Participants were categorized into five groups: mild cognitive impairment (MCI) due to AD (n = 25), mild dementia (MD) due to AD (n = 20), MCI non-AD (n = 34), MD non-AD (n = 13), and healthy controls (HC) (n = 16). Complex emotion recognition was assessed using RMET, and plasma cortisol levels were measured by mass spectrometry.

Current challenges in facial expression recognition (FER) for patients with neurological disorders include handling small and diverse datasets through data augmentation, synthetic data generation, and transfer learning to improve model performance. Temporal information, often overlooked, can be integrated using RNNs or LSTMs to enhance pattern recognition in dynamic expressions. Generalization to real-world conditions remains a concern, as variations in lighting, pose, and occlusions impact system accuracy, requiring robust solutions. Detecting subtle facial cues indicative of neurological disorders demands advanced feature extraction and model tuning. Multimodal integration, combining FER with vocal, gestural, or physiological data, offers a comprehensive diagnostic approach. Finally, interpretability and explainability are critical in healthcare applications, necessitating transparent models and visualization tools to build trust and facilitate clinical integration.

2. Mathematical model and method

Numerous classes in the database were examined for this study. It serves to verify the suggested approach. The Karolinska Emotional Dataset for Facial (KEDF) expressions dataset was generated by the clinical neuroscience department of the Karolinska Institute in Sweden and is available to the public. The images, which measure 562 by 762 and feature 70 persons (35 women and 35 men), each have a common human expression (Normal, happy, sad, astonished, terrified, angry, and disgusted). The images were captured using five distinct camera angles.[21] The KEDF dataset, which contains distinct, varied, and high-resolution images, was chosen to train the suggested model. Furthermore, happiness, normalcy, sadness, and fury are the primary emotions people with neurological illnesses display. As a result, we chose only these data classes from the KEDF and arranged them into four groups. Fig. 1 displays the sample photos from the KDEF. The number of classifications, images, and image attributes (color, quality, format, and depth) vary among the databases. This extensive database helps with model learning, but the different shooting angles make it difficult to classify the data.

2.1 Preprocessing-cropping and Gaussian filter

Preprocessing is necessary to improve the model's learning properties during learning. Descriptive measures (mean, standard deviation) provided an overview of variations, while inferential techniques such as analysis of variance (ANOVA) and Pearson’s correlation assessed group differences and relationships with disease severity. In order to manage just the region of interest (ROI) for processing,[22] preprocessing attempts to eliminate unnecessary pixels from the input image. Finding the face and cropping it is the first step. The cropping process accounted for challenges such as illumination and angle fluctuations by employing a series of robust techniques. Histogram equalization was applied for illumination normalization, and facial landmarks were detected using the Dlib library to define the region of interest. Angle fluctuations were corrected through affine transformations, and a dynamic bounding box was used to optimize the cropped region. The Gaussian filter was chosen for noise elimination due to its effectiveness in smoothing high-frequency noise while preserving critical facial features. Its computational efficiency and robustness to illumination variations make it particularly suitable for preprocessing facial images in video surveillance. Compared to other filters, such as the median, mean, or bilateral filters, the Gaussian filter strikes an optimal balance between noise reduction and edge preservation, ensuring

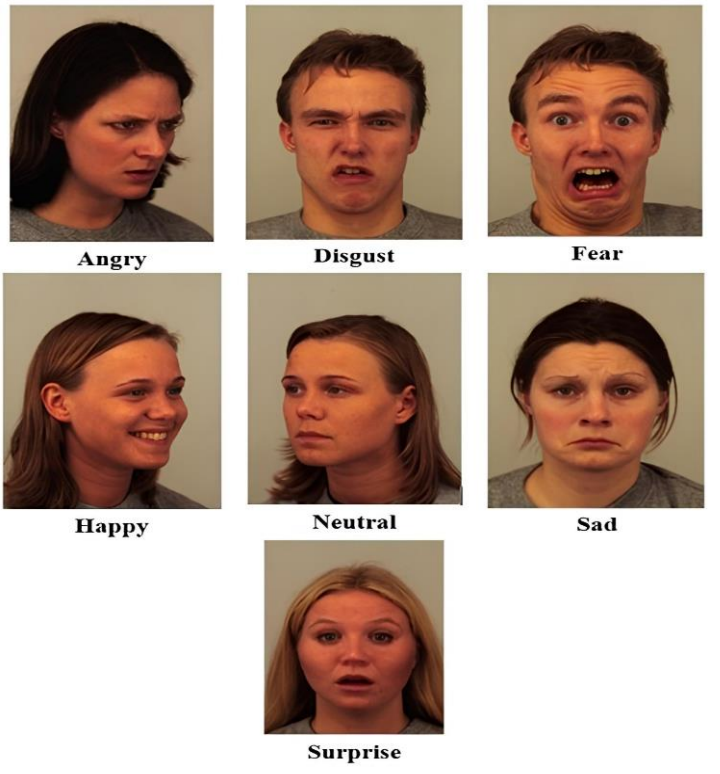

Fig. 1 Input dataset.

reliable preprocessing for subsequent skin detection and encryption processes.

2.2 Convolutional neural network (CNN)

A novel convolutional neural network architecture for patient facial expression recognition is created in this article. CNN, autoencoders, and this architecture are integrated. To accomplish the necessary features in this case, the autoencoders' encoding section.[23] This section considers an input image to generate a high-dimensional feature vector. The features are then combined at different points in time. The background research on CNN is first provided. Living things' typical visual observation architecture serves as the inspiration for CNN, a deep learning architecture. Every layer of a CNN transfers a volume of initiates to the next layer with a variational limit, making a key CNN a progressive of layers. In addition to additional normalization layers, this CNN comprises three primary layers: the convolutional layer, the pooling layer, and the fully connected layer.[24]

Furthermore, the autoencoder is a unique artificial neural

Fig. 2 Pre-processing dataset**.**

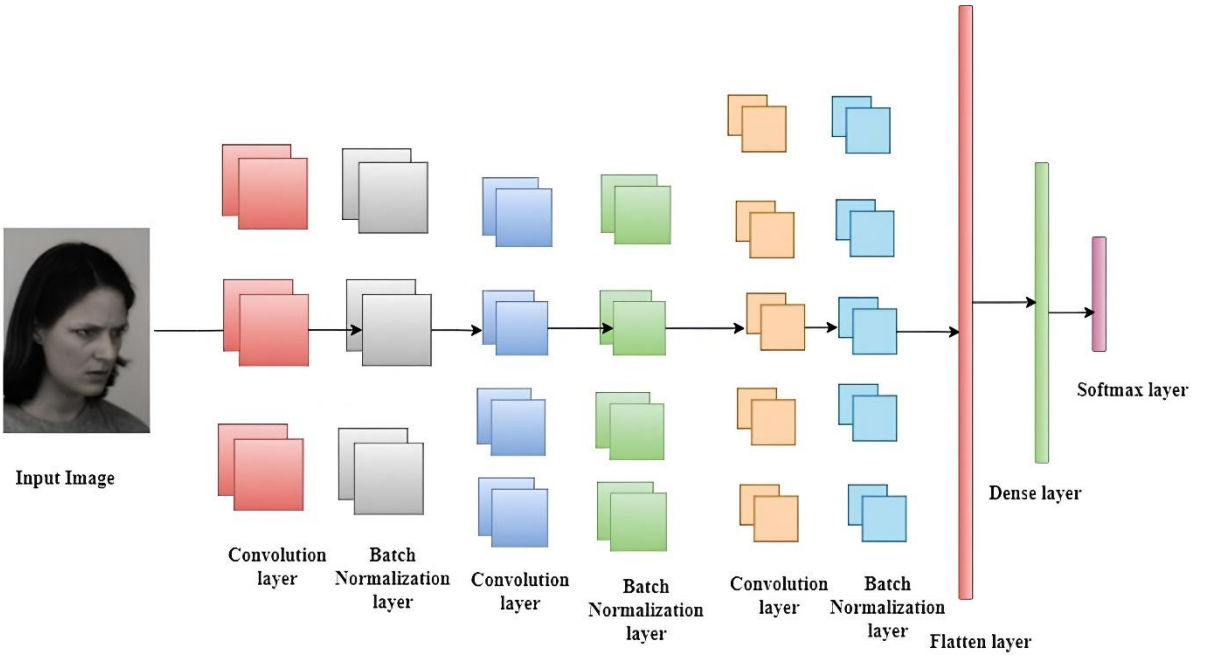

Fig. 3 Encoder-based convolutional neural network.

network (ANN) type trained to produce optimal encodings. There are two stages involved: encoding and decoding. This proposed architecture takes into account the encoding part. Several encoding methods are typically demonstrated, including convolutional, denoising, contractive, zero-biased, and sparse autoencoders. The convolutional encoder is used in this architecture. Max-pooling, fully connected layer, and encoder-based convolution layers are all included in the proposed architecture, as displayed in Fig. 3.

Utilizing the activation function, the convolution layer is created. This section explains the layers of the architecture. The integration of autoencoders with CNN enhances the model's ability to extract meaningful features while reducing computational complexity. The encoding section of the autoencoder compresses the input data into a lower-dimensional latent representation, capturing spatial and textural features essential for skin detection and encryption. This process removes noise and redundant information, ensuring that the CNN receives high-quality input. By leveraging the encoder's ability to generalize across variations in lighting and angle, the combined architecture improves accuracy and robustness in privacy protection for surveillance videos.

2.2.1 Convolutional layer

A convolutional network's central design unit is capable of handling challenging jobs. The forward pass function convolves each channel in the convolutional layer throughout the image's height and breadth. The dot product between the image and kernel pixels should be calculated during the backward pass.[25] The gradient of loss, weights, input, and bias should all be calculated in this layer. Here is how the input image's convolution filter is created.

$$F(x,y) * G(x,y) = \sum_{N_1=-\infty}^{\infty} \sum_{N_2=-\infty}^{\infty} f(N_1,N_2).G(x-N_1, y-N_2) \quad (1)$$

A filter mask of a specific size $N_1 \times N_2$ is defined as $G(x,y)$ in the equation above. A weight matrix is incorporated into the design of the filter or channel to extract the necessary data from the standard image matrix. Therefore, various weight combinations are used, each of which can be used to extract a particular set of properties. Furthermore, the suggested classifier made use of several convolutional layers. As We get more profound and intricate, the layers are thought to address a particular problem while extracting the most general elements from the image. Practical techniques such as cushioning and stride were then applied. The process of padding an image matrix in order to preserve the necessary features is known as adding zeros. An image matrix can be moved using a variable called a stride.

2.2.2 Batch normalization

The network's performance and stability are enhanced by its use. It enables layer lowering in a given range for the inputs. This layer has a mean output standard deviation of one and an output activation of zero. It is used to normalize the input of all other layers. The formula for this normalization function is as follows:

$$\hat{x} = \frac{x - e[x]}{\sqrt{var[x] + \epsilon}} \quad (2)$$

Here, ϵ avoids zero division since it is a continuous parameter. The variables $var[x]$ and $e[x]$ represent the current batch's variance and mean. This layer decreases regularization and training time requirements.

2.2.3 Activation function

To control network nonlinearity, activation functions are

incorporated into the suggested classifier.[26] Compared to the sigmoidal function, the Rectified Linear Unit (ReLU) computes more quickly and easily, making it an activation function. The following is the formula for the ReLU function:

$$f(x) = \begin{Bmatrix} x & if\ x > 0 \\ 0 & otherwise \end{Bmatrix} \quad (3)$$

2.2.4 Max pooling layer

This pooling layer shrinks the spatial size of the image to minimize parameter calculation and manage overfitting. If the input image is enormous, the number of trainable features must be lowered. Using a pooling layer causes this issue. Reducing an input image's spatial size is a prerequisite for adding pooling layers. The two common pooling layers are max pooling and average pooling. This paper uses max pooling to help capture important local features in the image and increase the model's generalization ability.

2.2.5 Fully connected layer

This layer is essential for producing an output that matches the number of classes that must be created. During this step, neurons finish flattening and form a vector of fully formed neurons. Neural networks in this layer have entire connections to neurons in the final layer, similar to the problem of coherent neural networks.

2.2.6 SoftMax layer

This layer is the final product. Various events' probability distributions are computed in this layer. The following is the formulation of this SoftMax function of the event,

$$f(x_i) = \frac{Exp(x_i)}{\sum_{j=0}^{k} Exp(x_i)} \quad here, i = 0, 1, 2, \ldots, k \quad (4)$$

The formula calculates the exponential of an input or event and the total of all the input values' exponentials. The SoftMax function's output is then the ratio of the two. One of the main benefits of utilizing the SoftMax activation function is that the computed probabilities will fall between 0 and 1, and the total of all the probabilities will equal 1.

This study ideally chooses CNN's hyperparameter with the application of EGSA. Numerous design techniques make model design challenging. The author is still determining the ideal design architecture for a given model.[27] The configuration's external, non-design factors that are impossible to determine from the data are known as hyperparameters. These two different types of hyperparameters are hyperparameters to the network structure and the trained network. For network structure, activation functions include stride rate, kernel type parameters, hidden layer between input and output layers, padding that adds layers, model management to train nonlinear identification boundaries, and kernel size (size of filter). In the training stage, batch size, momentum with weight parameter, learning rate, and epoch with iterations are hyperparameters of the trained network. Table 1 contains a list of the CNN's chosen hyperparameter.

2.3 Enhanced golden search algorithm (EGSA)

In order to acquire local and global searches, exploration and exploitation should be balanced as much as possible. Local searches in the current place are essential for exploitation. Furthermore, they differ in that one may sacrifice the other when improvising. Therefore, finding the ideal balance between exploitation and exploration is a challenging and important problem for any optimization algorithm.[28] Thus, the following are a few of the global optimization algorithm (GOA's) limitations:

- This approach is simple and keeps the population size constant at each generation. However, it reduces the algorithm's versatility.
- It becomes stuck in local optima and does not respond robustly when trying to achieve global optimization for various functions.
- It has both effective local exploitation capabilities and weak exploitation capabilities.

The oppositional function produces solutions in reverse. This function gives the best CNN hyperparameter solutions while improving the original population. The search process is started with an initial random generation of candidate solutions by the glowworm swarm optimization (GSO),[29] a population-based metaheuristic optimization technique. This algorithm considers the step size variable and upgrades the

Table 1. The hyperparameters of the CNN model.

S. No	Variable Name	Variable
1	Kernel size	Ksize 1 [3,4,5]
2	Hidden Layer	Full hidden 1 [60,100,125]
3	Activation	**"relu"** (ReLU) for efficient training and gradient flow, **"lrelu"** (Leaky ReLU) to prevent dying neurons and enhance feature extraction, and **"elu"** (Exponential Linear Unit) for smoother gradients and faster
4	Kernel size	Ksize 2 [3,4,5]
5	Hidden Layer	Full hidden 2 [60,100,125]
6	Number of filters	Filters 1 [16,32,64,96]
7	Number of filters	Filter 2 [48,64,96,128]
8	Number of filters	Filter 3 [64,96,128]
9	Kernel size	Ksize 3 [3,4,5]

object positions in each iteration till the compensated termination condition. The optimization algorithm comprises mathematical stages, such as exploitation and exploration. It also maintains equilibrium between two contradicting functions. The two primary components of this optimization technique are updating position, creating a population, and evaluating fitness. EGSA is employed to optimize the hyperparameters. The stages of the process are shown as follows:

2.3.1 Phase 1: Initialization with oppositional functions

This method uses the quasi-opposition function to provide the best local and global search results.[30] This algorithm begins the search process with two arbitrarily generated objects in the search space that are connected to the following:

$$O_i = LB_i + RAND \times (UB_i - LB_i),\ i = 1, 2, .., n \quad (5)$$

$$x_i^{Q0} = RAND\left(\frac{LB_i - UB_i}{2}, LB_i - x_i\right),\ i = 1, 2, 3, .., pop \quad (6)$$

where LB_i and UB_i are defined as lower and upper bounds, O_i denotes the position of the objects within the search space UB_i, and x_i^{Q0} denotes the solution based on quasi-oppositional functions.

2.3.2. Phase 2: Fitness computation

This step involves computing the starting population with the objective function and selecting the object with the best fitness value. The fitness function is used to train and validate the suggested model. The low parameters of the utility function show how well the model's predictions for facial remarks matched reality. The fitness function, therefore, calculates the forecast accuracy. The fitness function is regarded as a mean square error.

$$FF = \frac{1}{N}\sum_{I=1}^{N}(t_i - p_i)^2 \quad (7)$$

Here, N is the total number of features, P_i defines the expected parameters, and t_i represents the valid parameters.

2.3.3 Phase 3: Gloden variation

The third stage involves sorting the items according to their fitness function and changing the object with the lowest fitness using a random solution.

2.3.4 Phase 4: Step size computation

The step size operator is considered in each iteration of the optimization process to modify the objects to the ideal solution. There are three components to the step size operator. In the first part, the transform operator that reduces iteratively to balance the algorithm's local and global searches defined the previous variable of the step size, which is different. The distance between the object's current location and its best position to date was determined in this section by calculating the cosine of a random parameter with a range of 0 to 1. In the final part, the sine of a random parameter between 0 and 1 is multiplied to determine the distance between the current position of the ith object and the ideal position attained among all objects. The step size operator is generated randomly in the first optimization iteration and updated using the following equations as needed. Fig. 4 shows the steps of proposed Architecture.

$$S_{Ti}(T+1) = t.S_{Ti}(T) + C_1.cos(R1).\left(Obest_i - x_i(t) + C_2.cos(R2).(Obest_i - x_i(t))\right) \quad (8)$$

Here, t is a transfer operator that changes the focus of search from exploitation to exploration; $Obest_i$ is described as the object's ideal final location. The random numbers in the range of (0,1) are designated as $R2$ and $R1$, and the random numbers between 0 and 1 are designated as C_1 and C_2. This transfer operator improves search performance and manages the ratio of local search in subsequent iterations to global search in initial iterations. Typically, this transfer function is decreasing and can be calculated with the following formula:

$$T = 100 \times \left(-20 \times \frac{T}{T_{MAX}}\right) \quad (9)$$

Here, the maximum number of iterations is denoted by T_{MAX}.

2.3.5 Phase 5: Step size limitation

Every iteration of the method works by controlling the distance each object travels in each dimension problem. Thanks to this stochastic variable step size, the objects can handle more comprehensive cycles in the issue space. A necessary gap is designed for the object clamp movement associated with it to prevent these oscillations and lessen divergence and explosion.

$$-S_{Timax} \le S_{Ti} \le S_{Timax} \quad (10)$$

Here, S_{Timax} is a defined maximum movement produced that characterizes the maximum variation of an item throughout an iteration while considering positional coordinates. The formulation of this process is as follows:

$$S_{Timax} = 0.1 \times (UB_i - LB_i) \quad (11)$$

2.3.6 Phase 6: Position updating

During this stage, the item travels to the global optimum in the search space associated with the equation below.

$$O_i(T+1) = O_i(T) + S_{Ti}(T+1) \quad (12)$$

2.3.7 Phase 7: Termination condition

This stage involves verifying the termination condition. Convergence occurs when the maximum number of iterations is reached. Ultimately, the best options are stored and considered to recognize facial expressions.

The EGSA is employed to optimize the CNN's hyperparameters, including both continuous (*e.g.*, learning rate) and discrete (*e.g.*, number of filters, activation functions) types. EGSA iteratively evaluates candidate values by applying a Golden Section Search strategy, narrowing the search space based on performance metrics such as validation accuracy. By systematically refining the hyperparameter configurations, EGSA ensures that the CNN achieves optimal performance. The integration of EGSA with the CNN results in a robust and efficient model fine-tuned for the specific task of facial expression recognition.

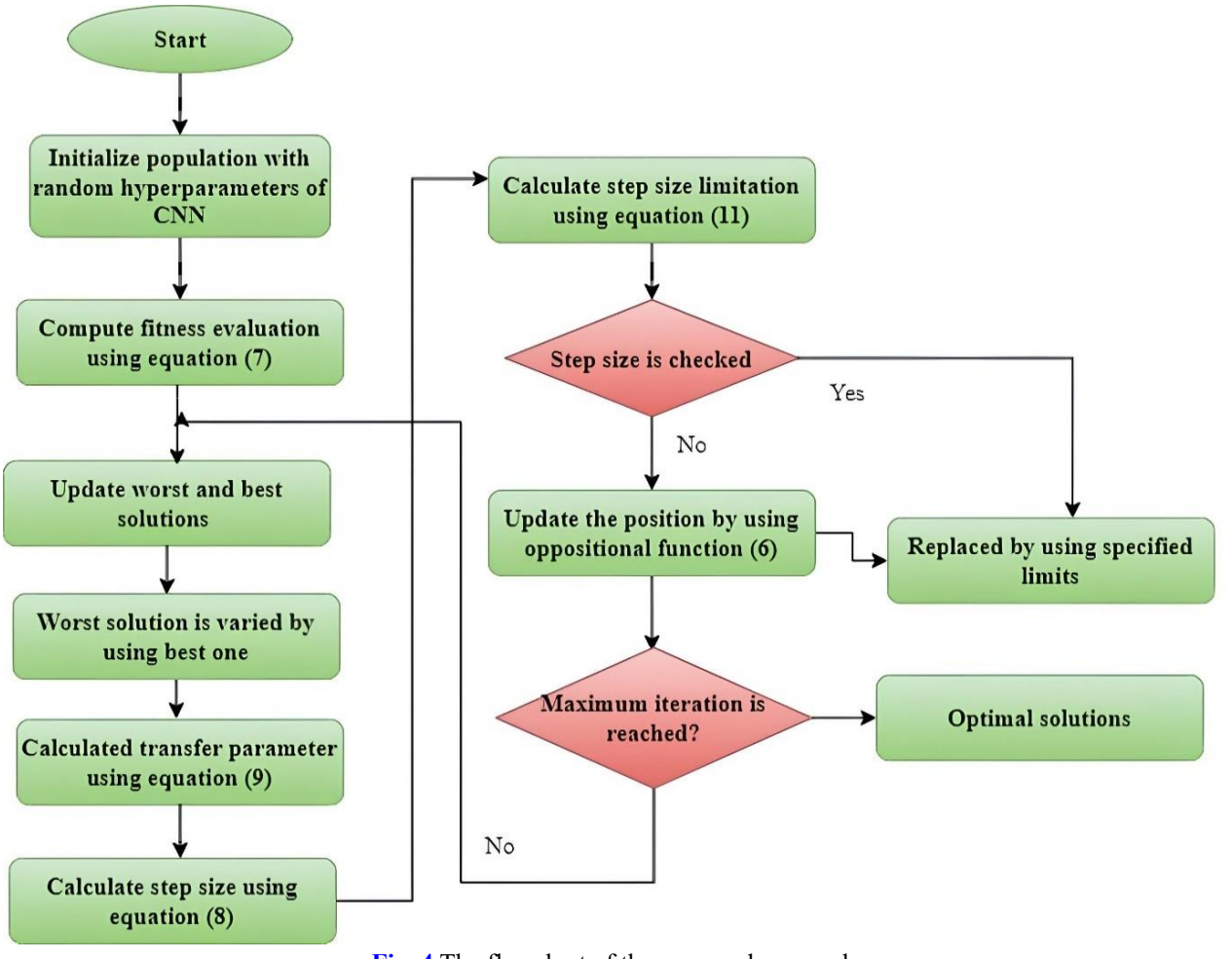

Fig. 4 The flowchart of the proposed approach.

2.4 Proposed architecture

Patients with Parkinson's, stroke, Alzheimer's, and Bell palsy disorders would have their facial expressions identified using the suggested method. First, familiar web sources are used to compile the dataset. Furthermore, publicly available databases gather unprocessed photos of facial expressions that often represent the patient, such as usual, happy, sad, and angry. The data analysis aimed to determine the extent to which individuals could be distinguished from one another when searching for Parkinson's disease symptoms. The entire design of the suggested method is shown in Fig. 5.

The image is then cropped from the input image by considering cropping. Next, a Gaussian filter is used as part of the preprocessing approach to eliminate noise. FTDLM is used to classify the emotions in the preprocessed image. Combining an EGSA and an NCNN, this model is suggested. With the NCNN, EGSA is used to choose the hyperparameters. The training and testing parts of the suggested classifier are now completed. The model is trained on 80% of the photos throughout the training phase. During testing, 20% of the photos are used to evaluate the model according to the facial emotions of the patients.

3. Experimental evaluation

The assessment and rationale for the suggested approach are examined in this section. The suggested approach is carried out in a virtual environment running Python 3.7, which is set up with an Intel (r) Xeon (r) X5560 CPU, 8.00GB of install RAM, 2.80GHz clock speed, and GRX GeForce 107 graphics card. The suggested classifier's evaluation, training, and design also use several libraries and frameworks, such as the TensorFlow-GPU version 2.0.0 and Kera's GPU frontend. Mean square error and exponential gradient optimization algorithm (EGOA) with an initial learning rate of 10^{-4} are used to calculate the loss of the projected classifier and upgrade its weights during training. Performance and comparative analysis are used to validate the suggested classifier. Several perspectives are taken into consideration when evaluating the input database. Fig. 6a shows the sample input image (anger) with different faces.

Subsequently, the input photographs are routed to the preprocessing phase for image enhancement and cropping. The results of the picture enhancement are displayed in Fig. 2. Fig. 6 (b) shows the cropping of the face from the input image. Several performance metrics, including F1 Score, mean square error, mean absolute error, recall, accuracy, specificity, precision, and mean square error, are considered to validate the suggested technique. Table 2 shows validation of accuracy, precision, and specificity.

Figure 7 calculates and shows the accuracy metric. We

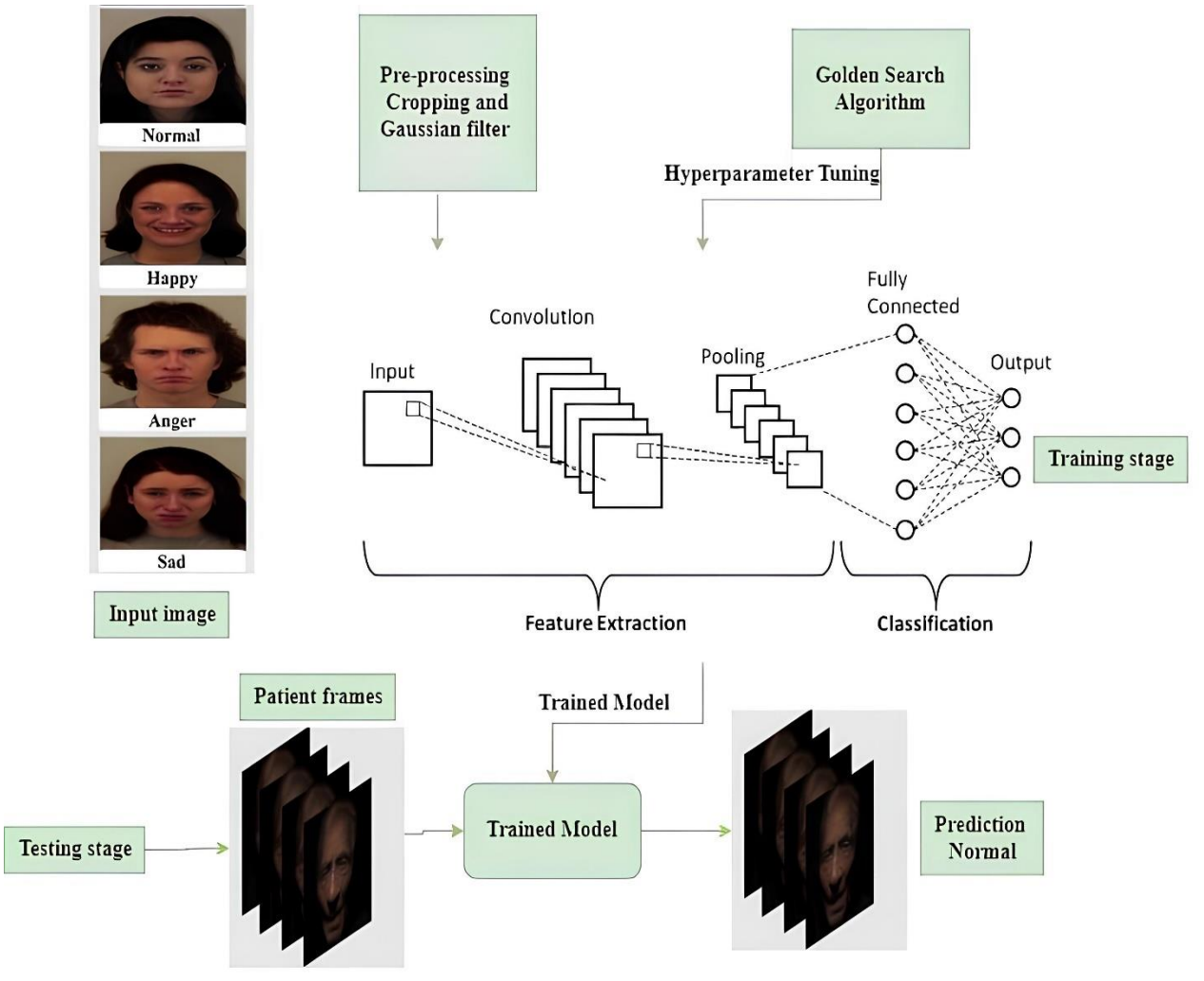

Fig. 5 Proposed architecture.

Table 2. Validation of accuracy, precision, and specificity.

Measure	Methods	Training percentage				
		0.5	0.6	0.7	0.8	0.9
Accuracy	CNN_GSA	0.44	0.58	0.85	0.85	0.95
	CNN_HBO	0.32	0.39	0.52	0.52	0.52
	CNN_PSO	0.34	0.35	0.64	0.64	0.64
	CNN_GA	0.17	0.22	0.42	0.42	0.43
Precision	CNN_GSA	0.12	0.30	0.30	0.78	0.95
	CNN_HBO	0.18	0.06	0.20	0.20	0.48
	CNN_PSO	0.19	0.08	0.20	0.20	0.43
	CNN_GA	0.18	0.02	0.20	0.20	0.64
Specificity	CNN_GSA	0.41	0.47	0.63	0.64	0.90
	CNN_HBO	0.11	0.20	0.31	0.31	0.37
	CNN_PSO	0.17	0.25	0.47	0.48	0.54
	CNN_GA	0.11	0.11	0.32	0.32	0.40

compare the proposed approach with the existing approaches used by CNN-HBO,[28] CNN-PSO,[29] and CNN-GA.[30] With a training percentage of 0.9, the recommended classifier yielded a 0.95. Furthermore, at 0.9 training percentage, the accuracy of the existing methods was 0.5299, 0.6484, and 0.435714. This validation showed that the recommended classifier has good accuracy at 0.9 training %. Fig. 8 calculates and shows the precision metric. We compare the proposed approach with the existing approaches used by CNN-HBO, CNN-PSO, and CNN-GA. With a training percentage of 0.9, the recommended

(a)

(b)

Fig. 6 Input image (a) and cropping image (b) for different classes.

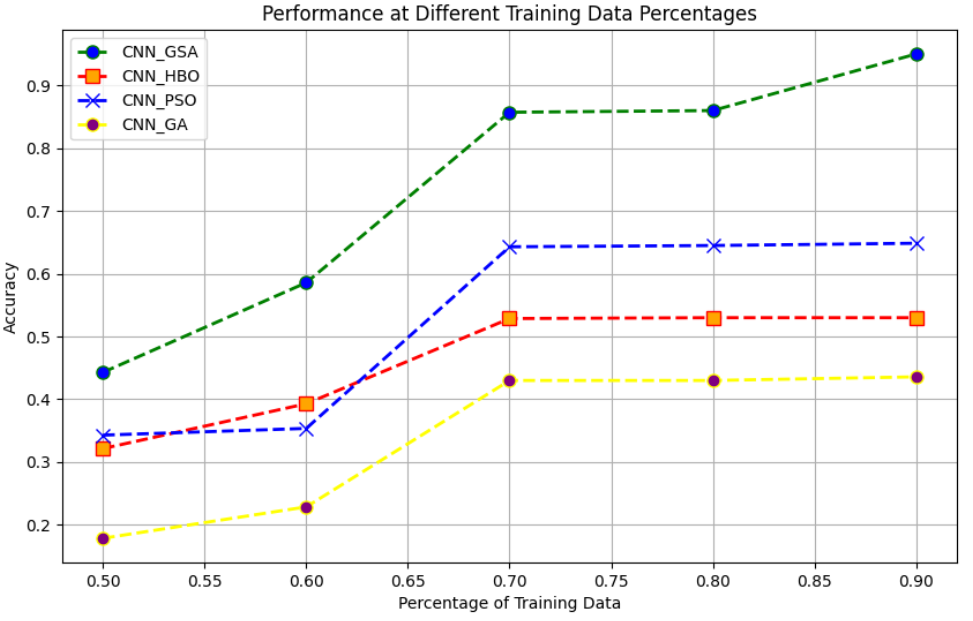

Fig. 7 The accuracy at different training data percentages.

classifier yielded a 0.95. Furthermore, at 0.9 training percentage, the accuracy of the existing methods was 0.480, 0.4325, and 0.640. This validation showed that the recommended classifier has good precision at 0.9 training %. The specificity metric is calculated and shown in Fig. 9. We compare the proposed approach with the existing approaches used by CNN-HBO, CNN-PSO, and CNN-GA. With a training percentage of 0.9, the recommended classifier yielded a 0.90. Furthermore, the present techniques' accuracy at 0.9 training % was 0.379, 0.548, and 0.4041. This validation showed that the recommended classifier has reasonable specificity at 0.9 training percentage.

In Table 3, Validation of F1 score, Recall, mean absolute error (MAE) and mean square error (MSE). Fig. 10 shows the F1 measure computed and displayed. CNN-HBO, CNN-PSO, and CNN-GA's present practices are compared with the proposed strategy. The recommended classifier achieved a 0.90 at a training percentage of 0.9. Furthermore, at 0.9 training percentage, the current methods' F1-scores were 0.360, 0.4402, and 0.360. According to this validation, the proposed classifier had an excellent F1 measure at 0.9 training %. The recall measure is calculated and shown in Fig. 11. The present practices of CNN-HBO, CNN-PSO, and CNN-GA are compared with the proposed strategy. The recommended classifier achieved a 0.95 at a training percentage of 0.9. Furthermore, the recall of the existing techniques was 0.428, 0.6484, and 0.435714, with a 0.9 training percentage. This validation revealed that the proposed classifier had an excellent recall at 0.9 training %. The MAE metric is calculated and presented in Fig. 12. CNN-HBO,

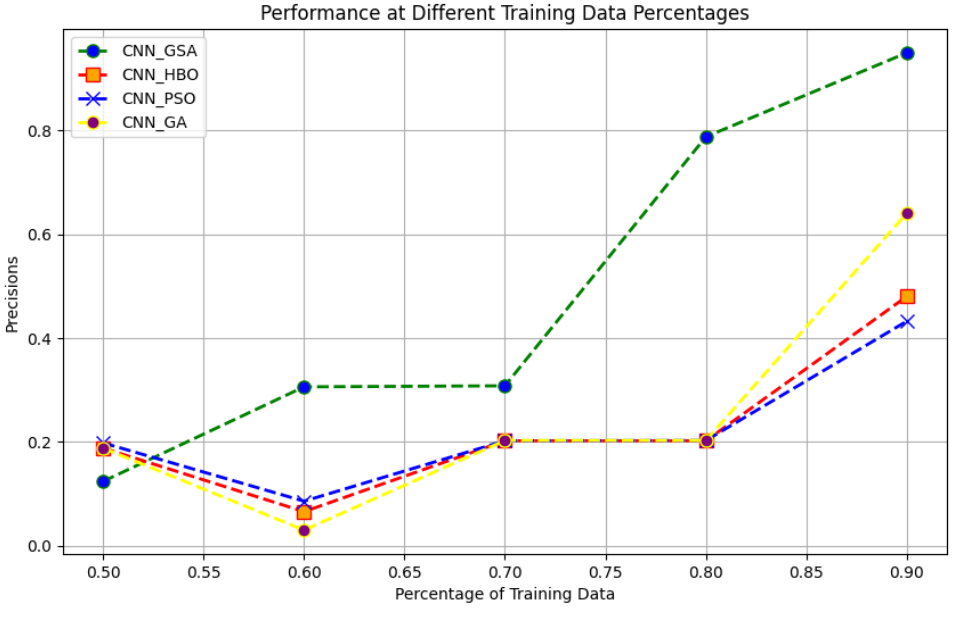

Fig. 8 The precision at different training data percentages.

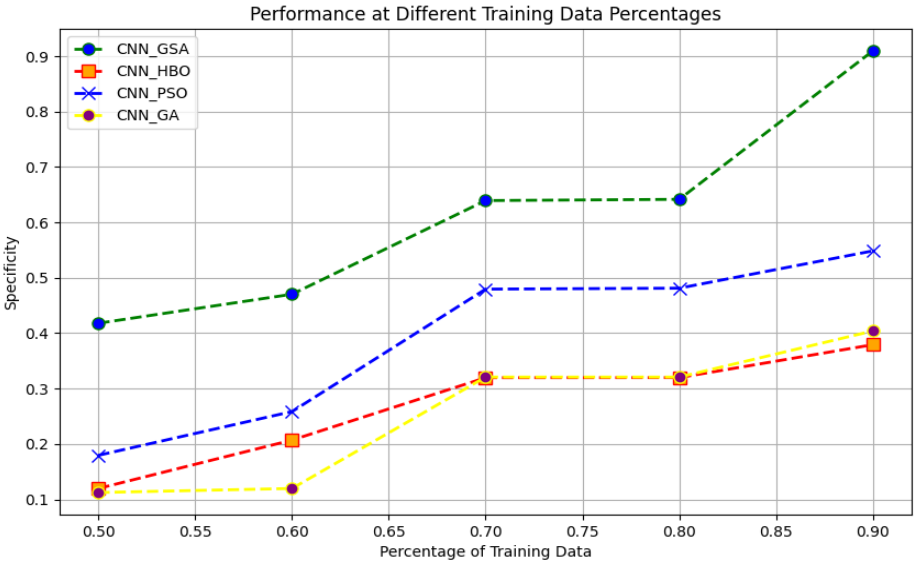

Fig. 9 The specificity at different training data percentages.

Table 3. Validation of F1 score, Recall, mean absolute error (MAE) and mean square error (MSE).

Measure	Methods	Training percentage				
		0.5	0.6	0.7	0.8	0.9
F1_score	CNN_GSA	0.35	0.35	0.58	0.66	0.90
	CNN_HBO	0.17	0.17	0.17	0.19	0.36
	CNN_PSO	0.35	0.35	0.35	0.39	0.44
	CNN_GA	0.17	0.17	0.17	0.19	0.36
Recall	CNN_GSA	0.44	0.58	0.85	0.85	0.95
	CNN_HBO	0.22	0.29	0.42	0.42	0.42
	CNN_PSO	0.34	0.35	0.64	0.64	0.64
	CNN_GA	0.17	0.22	0.42	0.42	0.43
MAE	CNN_GSA	0.92	1.04	1.14	1.14	1.14
	CNN_HBO	1.85	2.085	2.28	2.28	2.28
	CNN_PSO	1.52	1.65	1.71	1.71	1.85
	CNN_GA	1.85	1.94	2.20	2.41	2.99
MSE	CNN_GSA	2.5	3.9	4.00	4.01	4.72
	CNN_HBO	5	7.8	8.01	8.02	9.45
	CNN_PSO	3.89	4.00	4.02	4.8	5
	CNN_GA	5	5.63	8.47	10.36	12.99

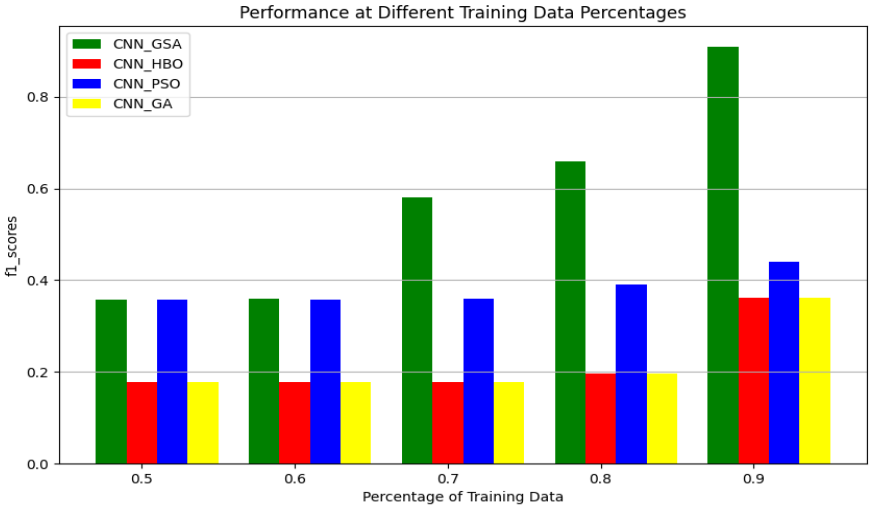

Fig. 10 F1_scores at different training data percentages.

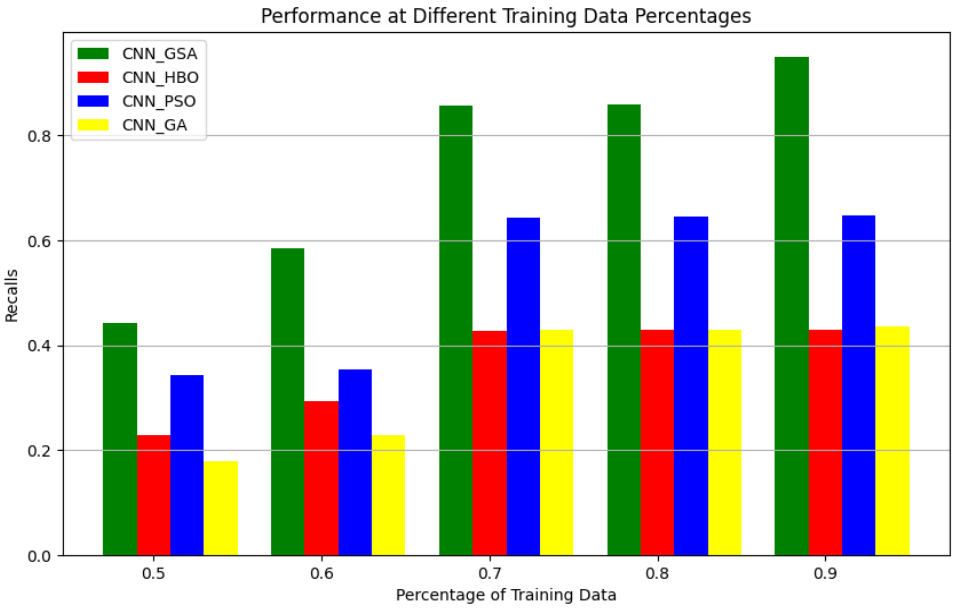

Fig. 11 Recalls at different training data percentages.

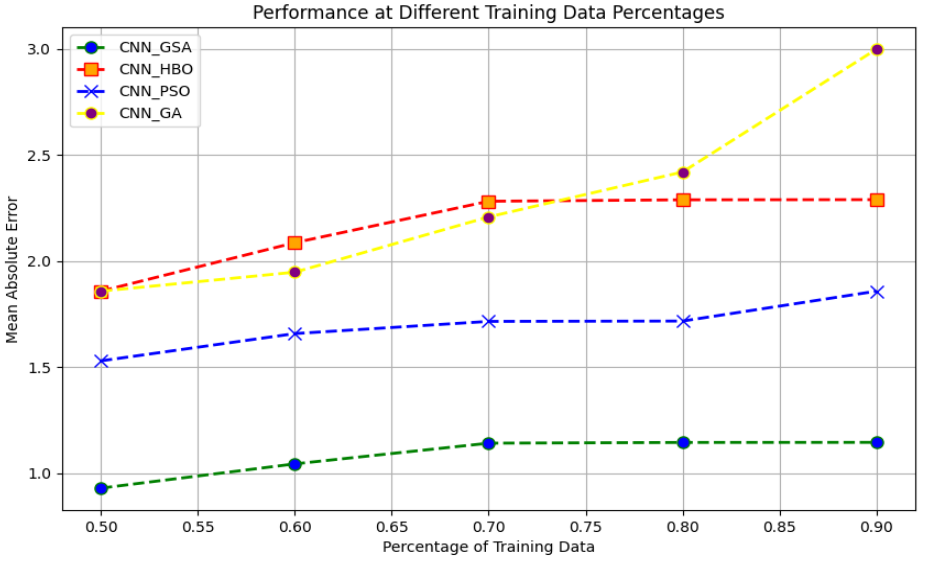

Fig. 12 Mean absolute error at different training data percentages.

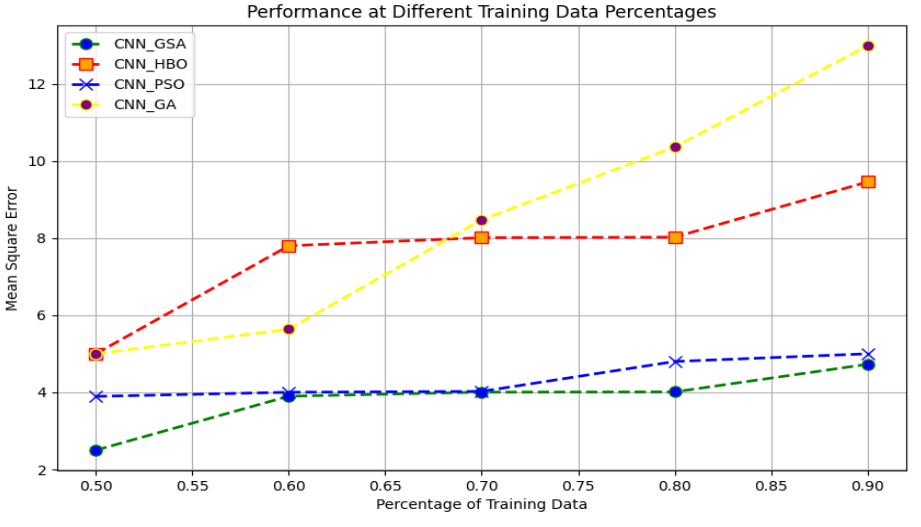

Fig. 13 Mean square error at different training data percentages.

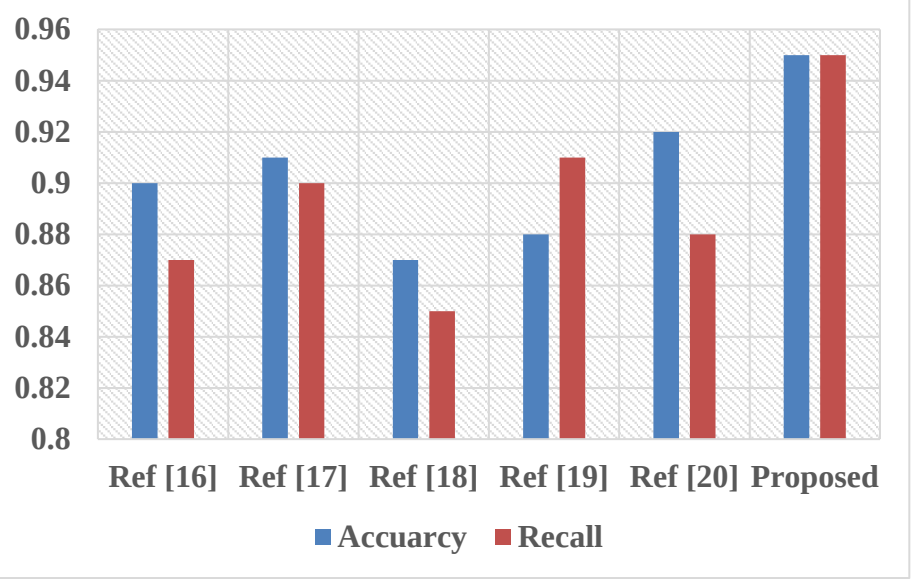

Fig. 14 Comparison analysis of different methods.

CNN-PSO, and CNN-GA's present practices are compared with the proposed strategy. The recommended classifier achieved a 1.14 at a training percentage of 0.9. Furthermore, at 0.9 training percentage, the current approaches' MAEs were 2.288, 1.85, and 2.99. This validation showed a low MAE of 0.9 for the recommended classifier's training %. The MSE measure is calculated and presented in Fig. 13. The proposed strategy is compared with the existing strategies used by CNN-HBO, CNN-PSO, and CNN-GA. With the recommended classifier, a 4.72 was obtained with a training percentage of 0.9. Furthermore, the MSE of the existing methods was 5, 12.990, and 9.45 at 0.9 training percentage. Based on this validation, the recommended classifier has a low MSE at 0.9 training percentage.

Table 4. Comparative evaluation of various methods.

Reference	Accuracy	Recall
[16]	0.9	0.87
[17]	0.91	0.9
[18]	0.87	0.85
[19]	0.88	0.91
[20]	0.92	0.88
Proposed	0.95	0.95

The proposed classifier is displayed and compared with the published papers in Fig. 14 and Table 4. Four academic articles are compared to it. The suggested classifier obtained a score of 0.95, whereas the published publications received scores of 0.87, 0.9, 0.85, 0.91, and 0.88. This validation showed that the recommended approach yielded the best outcomes.

4. Conclusion

In this paper, an FTDLM for diagnosing patients' facial expressions (including those with Parkinson's, stroke, Alzheimer's, and Bell palsy) has been developed. The dataset was first collected from familiar web sources. Furthermore, raw photos of the patient's most common facial expressions, such as usual, happy, sad, and angry, are gathered from publicly available sources. The data analysis aimed to determine if it was possible to identify individual differences when searching for Parkinson's disease symptoms. Following that, cropping was taken into consideration in order to resize the image from the input image. Next, the preprocessing method, which uses a Gaussian filter, is examined for eliminating noise. Using FTDLM, the preprocessed image has been used to classify emotions. This suggested model combines an EGSA with an NCNN. EGSA is used in the NCNN to choose the hyperparameters. With a high accuracy of 0.95, the proposed classifier was successful. Real-time data will be taken into consideration for this validation in the future.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] B. Torres Mendonça De Melo Fádel, R. L. Santos De Carvalho, T. T. Belfort Almeida Dos Santos, M. C. N. Dourado, Facial expression recognition in Alzheimer's disease: a systematic review, *Journal of Clinical and Experimental Neuropsychology*, 2019, **41**, 192-203, doi: 10.1080/13803395.2018.1501001.

[2] B. Güntekin, L. Hanoğlu, T. Aktürk, E. Fide, D. D. Emek-Savaş, E. Ruşen, E. Yıldırım, G. G. Yener, Impairment in recognition of emotional facial expressions in Alzheimer's disease is represented by EEG theta and alpha responses, *Psychophysiology*, 2019, **56**, e13434, doi: 10.1111/psyp.13434.

[3] J. Cárdenas, M. J. Blanca, F. Carvajal, S. Rubio, C. Pedraza, Emotional processing in healthy ageing, mild cognitive

impairment, and Alzheimer's disease, *International Journal of Environmental Research and Public Health*, 2021, **18**, 2770, doi: 10.3390/ijerph18052770.
[4] E. Fide, D. D. Emek-Savaş, T. Aktürk, B. Güntekin, L. Hanoğlu, G. G. Yener, Electrophysiological evidence of altered facial expressions recognition in Alzheimer's disease: a comprehensive ERP study, *Clinical Neurophysiology*, 2019, **130**, 1813-1824, doi: 10.1016/j.clinph.2019.06.229.
[5] L. C. Jiskoot, J. M. Poos, M. E. Vollebergh, S. Franzen, J. van Hemmen, J. M. Papma, J. C. van Swieten, R. P. C. Kessels, E. van den Berg, Emotion recognition of morphed facial expressions in presymptomatic and symptomatic frontotemporal dementia, and Alzheimer's dementia, *Journal of Neurology*, 2021, **268**, 102-113, doi: 10.1007/s00415-020-10096-y.
[6] A. de la Torre-Luque, A. Viera-Campos, A. C. Bilderbeck, M. T. Carreras, J. Vivancos, C. M. Diaz-Caneja, M. Aghajani, I. M. J. Saris, A. Raslescu, A. Malik, J. Clark, B. W. J. H. Penninx, N. van der Wee, I. W. Rossum, B. Sommer, H. Marston, G. R. Dawson, M. J. Kas, J. L. Ayuso-Mateos, C. Arango, Relationships between social withdrawal and facial emotion recognition in neuropsychiatric disorders, *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 2022, **113**, 110463, doi: 10.1016/j.pnpbp.2021.110463.
[7] M. Takahashi, S. Kitamura, K. Matsuoka, H. Yoshikawa, F. Yasuno, M. Makinodan, S. Kimoto, T. Miyasaka, K. Kichikawa, T. Kishimoto, Uncinate fasciculus disruption relates to poor recognition of negative facial emotions in Alzheimer's disease: a cross-sectional diffusion tensor imaging study, *Psychogeriatrics*, 2020, **20**, 296-303, doi: 10.1111/psyg.12498.
[8] H. Chainay, F. Gaubert, Affective and cognitive theory of mind in Alzheimer's disease: the role of executive functions, *Journal of Clinical and Experimental Neuropsychology*, 2020, **42**, 371-386, doi: 10.1080/13803395.2020.1726293.
[9] U. Nam, K. Lee, H. Ko, J. Y. Lee, E. C. Lee, Analyzing facial and eye movements to screen for Alzheimer's disease, *Sensors*, 2020, **20**, 5349, doi: 10.3390/s20185349.
[10] K. M. Heilman, S. E. Nadeau, Emotional and neuropsychiatric disorders associated with Alzheimer's disease, *Neurotherapeutics*, 2022, **19**, 99-116, doi: 10.1007/s13311-021-01172-w.
[11] J. Bek, E. Poliakoff, K. Lander, Measuring emotion recognition by people with Parkinson's disease using eye-tracking with dynamic facial expressions, *Journal of Neuroscience Methods*, 2020, **331**, 108524, doi: 10.1016/j.jneumeth.2019.108524.
[12] G. Mattavelli, E. Barvas, C. Longo, F. Zappini, D. Ottaviani, M. C. Malaguti, M. Pellegrini, C. Papagno, Facial expressions recognition and discrimination in Parkinson's disease, *Journal of Neuropsychology*, 2021, **15**, 46-68, doi: 10.1111/jnp.12209.
[13] R. Dan, F. Růžička, O. Bezdicek, J. Roth, E. Růžička, J. Vymazal, G. Goelman, R. Jech, Impact of dopamine and cognitive impairment on neural reactivity to facial emotion in Parkinson's disease, *European Neuropsychopharmacology*, 2019, **29**, 1258-1272, doi: 10.1016/j.euroneuro.2019.09.003.
[14] S. P. Coundouris, J. D. Henry, A. C. Lehn, Moving beyond basic emotions in Parkinson's disease, *British Journal of Clinical Psychology*, 2022, **61**, 647-665, doi: 10.1111/bjc.12354.
[15] J. Jakubowski, A. Potulska-Chromik, K. Białek, M. Nojszewska, A. Kostera-Pruszczyk, A study on the possible diagnosis of Parkinson's disease on the basis of facial image analysis, *Electronics*, 2021, **10**, 2832, doi: 10.3390/electronics10222832.
[16] W. Huang, W. Xu, R. Wan, P. Zhang, Y. Zha, M. Pang, Auto diagnosis of Parkinson's disease *via* a deep learning model based on mixed emotional facial expressions, *IEEE Journal of Biomedical and Health Informatics*, 2024, **28**, 2547-2557, doi: 10.1109/JBHI.2023.3239780.
[17] A. Davodabadi, B. Daneshian, S. Saati, S. Razavyan, Mathematical model and artificial intelligence for diagnosis of Alzheimer's disease, *European Physical Journal Plus*, 2023, **138**, 474, doi: 10.1140/epjp/s13360-023-04128-5.
[18] T. Xu, X. Wang, X. Lun, H. Pan, Z. Wang, ADReFV: Face video dataset based on human-computer interaction for Alzheimer's disease recognition, *Computer Animation and Virtual Worlds*, 2023, **34**, e2127, doi: 10.1002/cav.2127.
[19] S. Sarang, B. Sonawane, P. Sharma, R. Yeradkar, To study the effect of a newly developed emotion detection and grading system software for identifying and grading expressions of patients with Parkinson's disease, *Multimedia Tools and Applications*, 2024, **83**, 22855-22874, doi: 10.1007/s11042-023-16156-5.
[20] I. Ferrer-Cairols, L. Ferré-González, G. García-Lluch, C. Peña-Bautista, L. Álvarez-Sánchez, M. Baquero, C. Cháfer-Pericás, Emotion recognition and baseline cortisol levels relationship in early Alzheimer disease, *Biological Psychology*, 2023, **177**, 108511, doi: 10.1016/j.biopsycho.2023.108511.
[21] S. Barra, S. Hossain, C. Pero, S. Umer, A facial expression recognition approach for social IoT frameworks, *Big Data Research*, 2022, **30**, 100353, doi: 10.1016/j.bdr.2022.100353.
[22] Y. Li, K. Guo, Y. Lu, L. Liu, Cropping and attention based approach for masked face recognition, *Applied Intelligence*, 2021, **51**, 3012-3025, doi: 10.1007/s10489-020-02100-9.
[23] Y. Wang, X. Qiao, G.-G. Wang, Architecture evolution of convolutional neural network using monarch butterfly optimization, *Journal of Ambient Intelligence and Humanized Computing*, 2023, **14**, 12257-12271, doi: 10.1007/s12652-022-03766-4.
[24] E. Başaran, A new brain tumor diagnostic model: selection of textural feature extraction algorithms and convolution neural network features with optimization algorithms, *Computers in Biology and Medicine*, 2022, **148**, 105857, doi: 10.1016/j.compbiomed.2022.105857.
[25] F. Chen, C. Yang, M. Khishe, Diagnose Parkinson's disease and cleft lip and palate using deep convolutional neural networks evolved by IP-based chimp optimization algorithm, *Biomedical Signal Processing and Control*, 2022, **77**, 103688, doi: 10.1016/j.bspc.2022.103688.
[26] D. Wang, C. Xiang, Y. Pan, A. Chen, X. Zhou, Y. Zhang, A deep convolutional neural network for topology optimization with perceptible generalization ability, *Engineering Optimization*,

2022, **54**, 973-988, doi: 10.1080/0305215x.2021.1902998.
[27] R. Ali, J. H. Chuah, M. S. Abu Talip, N. Mokhtar, M. Ali Shoaib, Structural crack detection using deep convolutional neural networks, *Automation in Construction*, 2022, **133**, 103989, doi: 10.1016/j.autcon.2021.103989.
[28] M. Noroozi, H. Mohammadi, E. Efatinasab, A. Lashgari, M. Eslami, B. Khan, Golden search optimization algorithm, *IEEE Access*, 2022, **10**, 37515-37532, doi: 10.1109/ACCESS.2022.3162853.
[29] D. Swaminathan, A. Rajagopalan, O. D. Montoya, S. Arul, L. F. Grisales-Noreña, Distribution network reconfiguration based on hybrid golden flower algorithm for smart cities evolution, *Energies*, 2023, **16**, 2454, doi: 10.3390/en16052454.
[30] M. F. Ahmad, N. A. M. Isa, W. H. Lim, K. M. Ang, Differential evolution with modified initialization scheme using chaotic oppositional based learning strategy, *Alexandria Engineering Journal*, 2022, **61**, 11835-11858, doi: 10.1016/j.aej.2022.05.028.