IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 43, No. 3, June, 2026, pp. 1523-1533
Journal homepage: http://iieta.org/journals/ts

Bio-Inspired Computational Intelligence for Facial Emotion Decoding Using Genetic Neural Network

J. Angel Barakka[1], P. Kola Sujatha[2*]

[1] Department of Computer Science and Business Systems, Sri Sairam Engineering College, Chennai 600044, India
[2] Department of Information Technology, Madras Institute of Technology Campus, Anna University, Chennai 600044, India

Corresponding Author Email: pkolasujatha@annauniv.edu

https://doi.org/10.18280/ts.430334

Received: 23 February 2026
Revised: 26 April 2026
Accepted: 7 May 2026
Available online: 30 June 2026

Keywords:
image emotion analysis, convolutional neural network, genetic neural network optimization, fitness evaluation, crossover operation

ABSTRACT

This study investigates recent computational strides in Image Emotion Analysis (IEA), with a specific focus on utilizing deep learning for the categorization of emotional states derived from facial expressions. The study introduces an optimized Genetic Neural Network (GNN) for facial emotion analysis, leveraging a genetic algorithm on an extensive image dataset. This optimized GNN proves effective in classifying emotions based on facial images, extracting intricate facial elements to enhance categorization accuracy. The improvements lie in enhanced fitness evaluation, crossover operations, and selection mechanisms. Fitness evaluation incorporates additional metrics assessing the relevance and distinctiveness of facial features, contributing to a more refined emotion recognition model. The enhanced crossover operation amalgamates informative features from parent individuals, generating offspring with superior emotional feature representation. A sophisticated tournament selection mechanism prioritizes individuals with higher fitness values, facilitating effective exploration and exploitation of the population. Experimental results show strong performance, achieving ~95% training and ~87% validation accuracy. Loss curves indicate rapid convergence with minor overfitting after 35 epochs. It affirms the superiority of the optimized GNN model in achieving enhanced accuracy compared to existing methods in facial emotion analysis. The proposed methodology exhibits promising applications in affective computing, human-computer interaction, and emotion-aware systems. By advancing the understanding of emotions evoked by visuals through computational analysis, this research contributes to the growing body of knowledge in the field. The optimized GNN model emerges as a robust tool for deciphering and classifying emotions from facial images, with potential implications for various interdisciplinary domains.

1. INTRODUCTION

The background of emotion analysis arises from the field of affective computing, which focuses on developing systems and algorithms that can understand, interpret, and respond to human emotions. Emotion analysis is crucial in various domains, including psychology, human-computer interaction, and artificial intelligence, as emotions play a significant role in human communication and behaviour [1]. Emotion analysis has gained importance with advancements in machine learning, computer vision, and natural language processing. Researchers have explored different modalities to analyse emotions, such as facial expressions, vocal intonation, body language, physiological signals, and textual data. Among these, facial expressions have received particular attention due to their visual nature and their significance in emotional communication. Facial emotion analysis involves detecting and interpreting facial expressions to determine an individual's emotional state. It focuses on analysing specific facial action units associated with different emotions [2]. Emotion Sensitive Thematic Analysis (ESTA) is an augmentation of both cognition and behaviour to emotional coding to address the gap in Thematic Analysis (TA), where participants' emotions are often overlooked. A structured, emotion-centered framework introduces thematic richness, analytical integrity, and reflexivity on the part of the researcher to traditional TA in a new, mutually compatible manner. The work is important in bridging an essential gap in qualitative methods, although it can be used only with proper caution and must be validated by numerous studies [3].

Despite progress, challenges remain in emotion analysis. Factors like lighting conditions, head poses, occlusions, and cultural differences can impact the accuracy of emotion recognition systems. Additionally, the subjective nature of emotions and the complexity of their expression pose ongoing research challenges. Nonetheless, emotion analysis finds applications in human-computer interaction, psychology, market research, entertainment, and healthcare. It has the potential to enhance user experiences, improve mental health assessment, personalize advertisements, and contribute to the development of emotionally intelligent systems. Facial emotion analysis, which involves detecting and interpreting facial expressions to understand emotions, utilizes various methodologies, each with its own strengths and weaknesses.

These approaches include feature-based, appearance-based, deep learning, and multimodal methods. However, they have their drawbacks. Feature-based approaches struggle with variations in lighting, poses, and occlusions [4]. Appearance-based methods can be sensitive to changes in facial appearance and may have difficulty differentiating similar emotions [5]. Deep learning approaches require large amounts of labelled data and lack interpretability. Multimodal methods aim to improve accuracy but face challenges in obtaining synchronized data [6]. Furthermore, all methodologies are affected by dataset bias and struggle with generalization. Overcoming these limitations is an ongoing research focus to enhance the accuracy and robustness of facial emotion analysis systems [7]. The goal of using Convolutional Neural Networks (CNNs) with an enhanced genetic algorithm in emotional analysis is to enhance the accuracy and effectiveness of recognizing emotions from facial expressions. The goal of using CNNs with an enhanced genetic algorithm in emotional analysis is to enhance the accuracy and effectiveness of recognizing emotions from facial expressions. CNNs are powerful deep learning models known for their ability to extract complex features from images, making them well-suited for analysing facial emotions. By training a CNN on a large dataset of labelled facial expressions, the network can learn to identify patterns and features associated with different emotions. Integrating a genetic algorithm into the CNN training process aims to further optimize its performance. Genetic algorithms (GA) simulate natural selection by iteratively evolving a population of potential solutions. By combining genetic principles, like mutation and crossover, with fitness evaluation, the genetic algorithm can efficiently search for the best network parameters and architecture for emotion recognition.

The specific objectives of using a genetic algorithm with CNNs for emotional analysis include:

1. Improved Feature Representation: The genetic algorithm helps optimize the network architecture, including layer sizes and quantities, to better capture relevant features in facial expressions. This optimization can lead to improved accuracy in emotion recognition.
2. Enhanced Generalization: GA assists in preventing overfitting, where the network becomes too specialized in the training data and performs poorly on new examples. By optimizing the network's parameters, the genetic algorithm improves generalization, enabling the model to accurately recognize emotions in diverse facial expressions.
3. Efficient Parameter Tuning: CNNs often have numerous adjustable parameters, and manually tuning them can be time-consuming and challenging. GA automates parameter tuning, exploring various combinations to find the best settings for emotion recognition.

Overall, the objective of using a genetic algorithm with CNNs in emotional analysis is to combine their strengths, improving feature extraction, optimizing network architecture and parameters, and ultimately achieving higher accuracy, better generalization, and more efficient parameter tuning in recognizing emotions from facial expressions. The novelty of this approach lies in the integration of deep learning and evolutionary computation, the automated optimization process, improved generalization, optimized feature representation, and the potential for enhanced performance in the field of emotional analysis using facial expressions. The enhanced Genetic Neural Network (GNN) approach, combining CNN with an improved genetic algorithm, offers solutions to various problems in emotion analysis. These problems include optimizing feature representation, addressing generalization and overfitting issues, achieving efficient parameter tuning, selecting optimal network architecture, and enhancing overall performance and accuracy. By leveraging the genetic algorithm's capabilities, the approach optimizes the network architecture and parameters, leading to improved feature representation and better capturing of subtle facial expression patterns. It also mitigates generalization problems by preventing overfitting and enhancing the model's ability to perform well on unseen data. Additionally, the genetic algorithm automates the parameter tuning process, making it more efficient and effective. The optimized network architecture further aids in accurately recognizing and classifying emotions. On the whole, the enhanced GNN approach presents advancements in emotion analysis by addressing these problems and improving the accuracy, robustness, and efficiency of emotion recognition systems.

2. LITERATURE REVIEW

There are various methods for analysing emotions, each with its own approach and focus. These methods include facial expression analysis, speech and voice analysis, physiological analysis, and textual analysis. Facial expression analysis involves interpreting facial expressions to determine emotions. Speech and voice analysis extracts emotional cues from speech signals. Physiological analysis examines physiological signals to infer emotions. Textual analysis analyzes written text to identify emotions expressed through words. The Bidirectional Encoder Representations from Transformers (BERT) and Bidirectional Long Short-Term Memory (BiLSTM) are used to create a text emotion categorization model in this study, suggested by M. Chen. The learned text is first transformed into a word-based vector representation using the BERT model. The semantic representation of the relevant word's context is then obtained by feeding the resulting word vector into the BiLSTM. The technique avoids the model from overfitting by including random dropout. The retrieved feature vector is then applied to the fully connected layer, where the Softmax function is used to determine which emotion category the text belongs to [8]. Adjectives, adverbs, nouns, and other phrase components that are more important in the expression of emotion may be better extracted by the convolutional layer. A Convolutional Long Short-Term Memory (ConvLSTM) network is created to take advantage of convolution in order to better enhance the categorization outcomes [9]. We provide a technique for predicting visual emotions that makes use of an image's affective semantic notions. We create a concept selection approach to mine emotion-related ideas from social media in order to address the issues of poor semantic coverage and low discriminative strength of emotions in existing semantic concept sets used for visual emotion analysis. To develop an effective semantic concept set that includes numerous visual ideas relevant to emotion transmission, we specifically suggest a number of selection procedures. Additionally, they are found in emotional image collections and related tags that are indexed from websites. We train concept classifiers to forecast each concept's concept score, which is then employed as an intermediate feature to address the semantic gap issue for picture emotion identification, in order to further capitalise on the found affective semantic ideas [10]. Computer vision's Facial Expression Recognition (FER) challenge continues to be difficult and intriguing. Despite attempts to create multiple FER techniques, the results are insignificant when applied to unseen photos or photographs taken in natural settings, since present methodologies cannot be generalised. FER issue resolution using a deep neural network architecture on several well-known standard face datasets. Our network has four Inception layers after two convolutional layers, each followed by max pooling. CNNs are sensitive to changes in picture quality, occlusions, stances, facial emotions, and illumination. In situations when the photos are obtained under difficult circumstances, such as dim light, extreme angles, or partly veiled faces, they may fail to appropriately analyse facial expressions [11]. The pipeline for analysing student behaviour in the e-learning environment using video facial processing has certain drawbacks. These include potential privacy concerns due to continuous facial data capture, challenges in accuracy and generalization under varying conditions, biases in facial recognition and emotion analysis, resource-intensive computational requirements, ethical considerations regarding consent and data security, sensitivity to environmental factors, and the importance of usability and acceptance by users. Addressing these concerns is necessary for the responsible and effective implementation of facial analysis techniques in e-learning [12]. The novel FER network, DAN, is designed based on two key observations: the subtle differences among multiple facial expression classes and the need for a holistic approach to capture high-order interactions among local features. To tackle these challenges, the DAN comprises three essential components: Feature Clustering Network (FCN), Multi-head Attention Network (MAN), and Attention Fusion Network (AFN). FCN extracts robust features using a large-margin learning objective, while MAN employs multiple attention heads to attend to different facial regions, and AFN fuses these attentions comprehensively. Despite achieving state-of-the-art performance on FER tasks, DAN may suffer from increased computational complexity and sensitivity to environmental variations [13]. Current FER methods have developed classic machine learning-based methods in favor of deep learning-based models and, most recently, models that use transformers. Traditional CNN-based methods are very accurate, but they tend to overfit and generalize poorly across varying illumination and pose conditions. As a solution to these shortcomings, bio-inspired optimization algorithms like GA, Particle Swarm Optimization (PSO), and Differential Evolution (DE) have been investigated to optimize the parameters of neural networks. Among them, GA has a good search ability across the globe because it has crossover and mutation operations, which minimize the chances of reaching a premature convergence. Nevertheless, the currently employed GA-based techniques tend to lack adaptive mutation schemes and effective selection schemes. New transformer-based models enhance contextual knowledge but are expensive in terms of computing resources. Thus, an effective and optimization-based method that would strike a balance between the accuracy and complexity of computations is needed.

The overall contribution of our proposed methodology is:

• CNNs can extract complicated visual information, making them ideal for face expression analysis. A CNN can recognise emotional patterns and features by training on a large labelled dataset of facial expressions.

• CNN training using an enhanced genetic algorithm improves performance. GA evolves a population of solutions by iterating natural Tournament selection. Modified Mutation, optimized crossover, and fitness criteria allow the genetic algorithm to find the best network parameters and design for emotion identification.

• By incorporating tournament selection as the selection mechanism in the Simple Genetic Algorithm (SGA), along with enhancing the single-point crossover and improving the mutation operation, the algorithm experiences significant improvements.

• The use of tournament selection, where individuals compete against each other, and the fittest ones are selected for reproduction, enhances the selection process by ensuring a better exploration and exploitation balance.

• This modification, combined with the inclusion of the fitness function as a criterion for selecting children after crossover, leads to a reduction in the number of iterations and faster convergence.

• Additionally, the enhanced bitwise mutation operation selectively mutates genes in the offspring, utilizing different mutation rates for two segments of the chromosome. This approach aids in improving the algorithm's ability to explore global search spaces, particularly in situations with local optima. Altogether, these enhancements contribute to the overall effectiveness and efficiency of the algorithm.

3. METHODOLOGY

The technique used in this research presents a CNN approach that has been optimised for facial emotion identification using the FER2013 dataset. The limitations of the previous approaches are addressed by the suggested method, which also provides a number of benefits in terms of boosting the accuracy of emotion identification. The improved fitness assessment is the first step in the approach. This evaluation goes beyond the usual metrics and takes into consideration additional criteria to evaluate the uniqueness and importance of face characteristics. This in-depth assessment makes it possible to more precisely evaluate the physical fitness of people and improves the functioning of the CNN model as a whole. The last step is the enhanced crossover operation, which involves efficiently combining the informative characteristics of the parent individuals. This process guarantees that the offspring receive the most important emotional traits, which results in greater accuracy in emotion identification and captures the intricacies of face expressions more effectively. The technique for selecting tournament participants has been implemented in order to promote more efficient population investigation and exploitation. This selection procedure gives precedence to those who have greater fitness values; this enables the algorithm to concentrate on people who have superior emotional categorization skills. As a direct consequence of this, the CNN model's general performance as a whole is improved. In terms of the data flow diagram, the dataset known as FER2013 is used as an input for the CNN model. The face images in this collection have been labelled with a variety of emotional descriptors. In order to perfect the CNN model, an iterative process is used to apply the improved fitness assessment, the enhanced crossover operation, and the tournament selection. The model goes through a series of training and learning stages, during which the face characteristics are dissected and categorised according to a variety of emotional states. The correct prediction of facial emotions is what the model spits out as its output, and it does so based on the input images. The results of the experiments show that the suggested method is superior to other approaches to facial emotion analysis since it achieves a greater level of accuracy than other approaches already in use. This CNN model that has been optimised may be used in a variety of applications, such as affective computing, human-computer interaction, and emotion-aware systems, all of which play important roles in the process of accurate emotion detection. Insights into the technique are provided by the full explanation and data flow diagram, which also highlight the flow of data and processes within the suggested approach is shown in Figure 1.

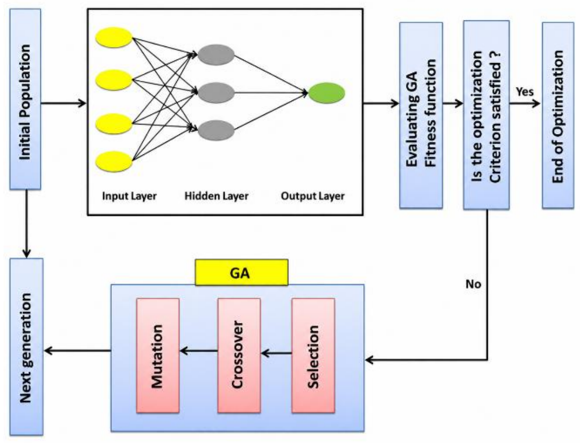

Figure 1. Architecture for improved Genetic algorithms (GA)

The CNN with an enhanced genetic algorithm approach for emotion analysis:

1. Initialize a population of potential CNN architectures.
2. Evaluate the fitness of each CNN architecture in the population using a fitness function.
3. Select the individuals (CNN architectures) with higher fitness values for reproduction.
4. Perform enhanced crossover between selected individuals to generate offspring with combined genetic information.
5. Introduce random mutations to the genetic information of the offspring.
6. Evaluate the fitness of the offspring CNN architectures.
7. Repeat steps 3–6 for a certain number of generations or until a termination criterion is met.
8. Select the best-performing CNN architecture from the final population.
9. Train the selected CNN architecture using a labeled facial image dataset.
10. Perform emotion recognition on new facial images using the trained CNN model.
11. Evaluate the performance of the CNN model using various metrics.
12. Analyse the results and make any necessary adjustments or improvements to the algorithm.

The algorithm iteratively evolves the CNN architectures through selection, crossover, and mutation operations based on their fitness scores. This process aims to find the optimal CNN architecture for accurate emotion recognition.

Modules in Enhanced Genetic Algorithm

- Initialization
- Fitness Evaluation
- Tournament Selection
- Enhanced Crossover
- Enhanced Mutation

3.1 Enhanced crossover

Tournament selection is a widely used method in GA for selecting individuals based on their fitness. It involves organizing competitions, or tournaments, among a randomly chosen subset of individuals from the population [14-17]. Tournament selection is applied in the selection of people to reproduce. Under this approach, a sample of size k is picked randomly out of the population, and the person with the best fitness is selected. The probability of selection is given as:

$$P(i) = \frac{f(i)}{\sum_{j=1}^{k} f(j)}$$

where, $f(i)$ is the fitness of the i^{th} individual. The size of the tournament in this work will be $k = 3$ so that there is a balance between selection pressure and diversity.

1. Determine tournament size: Specify the number of individuals that will participate in each tournament. Typically, this size is relatively small compared to the total population.
2. Repeat the following steps until the desired number of individuals for the next generation is selected:

Randomly select a subset of individuals from the population to form a tournament group, adhering to the predetermined tournament size.

Evaluate the fitness of each individual in the tournament group using the fitness function.

Compare the fitness values of the individuals in the tournament and choose the one with the highest fitness as the winner.

Repeat steps a to c until the desired number of winners is obtained.

In simpler terms, tournament selection involves randomly picking a small group of individuals from the population. These individuals then compete against each other, and the one with the highest fitness score is declared the winner. This process is repeated until the required number of winners is obtained. Tournament selection allows for diversity and competition among individuals, giving higher-fitness individuals a greater chance of being selected. By adjusting the tournament size, the balance between exploration and exploitation can be controlled. Smaller tournament sizes encourage exploration, while larger tournament sizes favour exploitation of the fittest individuals.

3.2 Enhanced mutation

The mutation operation in a classic SGA usually entails a very high mutation rate [18]. In this kind of SGA, each particular gene in the chromosome of an offspring is modified if a randomly generated value is lower than the mutation rate. To counter this, the authors of this work offer a revised algorithm as a means of speeding up the mutation process. In the improved method, a low mutation rate is first selected, and after that, a specific mutation is introduced into each gene that makes up the chromosome of the kid. To be more specific, the associated gene will undergo mutation in the event that the produced random number is lower than the mutation rate [19].

In addition, the mutation rate is doubled from its initial value if the location of the gene being traversed is beyond the place on the chromosome that corresponds to the halfway point of its length. This is due to the fact that the second half of the gene has a comparatively less amount of impact on the final product [20]. Because the mutation rates have been set up in this manner, it has been made certain that the first half of the chromosome as well as the second half of the chromosome have an equal probability of being mutated. Because of this, simultaneous mutation of both parts is now possible, which is great for fostering variety and expanding the search space. In the particular scenario in which the length of the gene is 784, the mutation rate for the whole chromosome may be estimated as follows:

$$1 - (1 - 0.025)392 * (1 - 0.05)392$$

Using this mutation method results in a significant increase in the species' variety while simultaneously preserving the integrity of the population. In general, the purpose of the improved mutation approach in the algorithm is to increase the capacity of the global search, particularly in circumstances in which the population may converge to a local optimal solution. The search capability of the genetic algorithm may be improved overall when genes are mutated carefully and in a well-balanced way. This allows the algorithm to investigate a greater variety of potential solutions. The two-stage mutation strategy is adopted in order to increase the exploration and exploitation balance. The early generations have a higher mutation rate (e.g., 0.2) to encourage diversity and prevent local minima. Later generations then slowly decrease the mutation rate (i.e., 0.05) to optimize the solution. It is a mutational mechanism of convergence, which enhances convergence stability and avoids premature stagnation.

3.3 Enhanced genetic algorithm

After incorporating modifications to the selection, fitness evaluation, and mutation operations, the genetic algorithm undergoes the following transformation:

1. Initialization:

Create an initial population of candidate solutions (chromosomes) by assigning random gene values.

2. Fitness Evaluation:

Assess the fitness of each chromosome using an improved fitness function that reflects the specific task at hand.

3. Selection:

• Choose parent chromosomes from the current population for generating the next generation.

• Employ advanced selection techniques, such as tournament selection, to strike a balance between exploration and exploitation.

4. Crossover:

• Perform crossover (recombination) between selected parent chromosomes to produce new offspring.

• Utilize an enhanced single-point crossover mechanism to exchange genetic material and generate diverse genetic combinations.

5. Mutation:

• Introduce mutation to the offspring chromosomes by selectively modifying individual genes.

• Adjust the mutation rates for different segments of the chromosome to enhance global search capabilities and overcome local optima.

6. Replacement:

• Replace a portion of the existing population with the newly generated offspring, considering their fitness and the modified selection criteria.

7. Iteration:

• Repeat steps 2 to 6 for a predefined number of iterations or until a termination criterion is met.

8. Output:

Identify the best individual (chromosome) from the final population as the solution to the problem at hand. The fitness function aims to include several performance criteria and make sure that the model is optimized. It is defined as:

$$\text{Fitness} = \alpha \times \text{Accuracy} + \beta \times \text{F1-score} - \gamma \times \text{Loss}$$

where, α, β, and γ are weighting factors that govern the contribution of each of the terms. In this work, the accuracy and F1-score are equally considered in order to deal with the imbalance of classes, and the loss is minimized to guarantee the stability of convergence.

An ablation study is done to measure the contribution of each component. Tournament selection, improved crossover, and adaptive mutation are gradually applied to the existing CNN to improve it. Findings indicate that all the components play a role in enhancing performance, with the adaptive mutation strategy exhibiting the greatest effect on convergence stability.

3.4 Convolutional Neural Network

The model is made up of multiple layers, including fully connected layers, flattening layers, convolutional layers, and max pooling layers. Filters and the 'relu' activation function are used in the convolutional layers in order to carry out the process of feature extraction from the input images [21].

$$output = relu(convolution(input, filters) + bias)$$

The max pooling layers bring the spatial dimensions of the feature maps down to a more manageable level [22].

$$.output = max\_pooling(input, pool\_size)$$

After the feature maps have been flattened, they are sent via fully linked layers, which cause the input data to be transformed [23].

$$output = flatten(input)$$

The 'softmax' activation function is used in the last dense layer to provide output probabilities for the various emotion classes.

$$softmax(dot(input, weights) + bias)$$

The formulae indicate the main operations carried out by each layer, where "input" stands for the data that are being input, "filters" refers to the learnable weights that are being applied to the filters, and "bias" stands for the biases that relate to the layer [24, 25]. During training, you will acquire knowledge about the particular values for weights and biases. The flow of information from one layer to the next is made possible by this sequential structure, which enables the model to learn and make predictions for emotion identification tasks.

The genetic algorithm and CNN training process are the main factors that affect the computational complexity of the proposed method. The complexity of the total is approximated to be $O(P \times G \times F)$, in which P is the population size, G is the number of generations, and F represents the cost of fitness evaluation. Despite the fact that the process of optimization makes the training process more time-consuming, the process has a tremendous impact on the performance of the model. The algorithm is still applicable to practice, with the adoption of the algorithm in consideration of the use of systems based on GPUs. All the experiments were performed in a system with an Intel i7 processor, 16 GB RAM, and NVIDIA RTX. The implementation of the model was done in Python through frameworks of TensorFlow/PyTorch. The CNN was being optimized by Adam with a learning rate of 0.001 and a batch size of 32. The genetic algorithm with a population size of 50 and 30 generations was used.

4. RESULTS AND DISCUSSION

Besides accuracy, there are several evaluation measures that are intended to give a holistic evaluation of the performance. To eliminate the problems of class imbalance, precision, recall, and F1-score are calculated per class. The metrics are taken as follows:

- Precision = TP / (TP + FP)
- Recall = TP / (TP + FN)
- F1-score = 2 × (Precision × Recall) / (Precision + Recall)

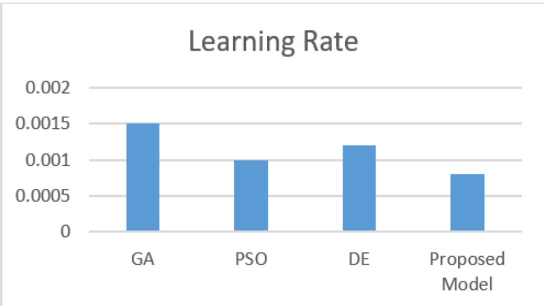

Figure 2. Learning rate of proposed model

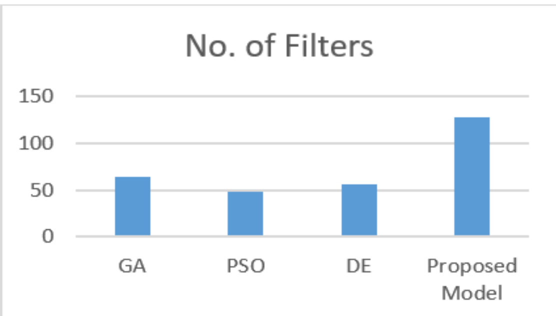

Figure 3. Number of filters with different algorithms

The macro-average and weighted-average scores are also reported to have a balanced assessment of all emotion classes. To test the strength of the proposed model, two benchmark datasets are used, namely FER2013 and Emotic. FER2013 is a collection of grayscale faces that are taken under controlled conditions with basic facial expressions as their main feature. Contrastingly, Emotic incorporates real-world images where context is given, such as body position and the surroundings. These two datasets can be used to assess it in both controlled and unconstrained conditions, and therefore confirm the generalization potential of the proposed method as depicted in Figures 2–9.

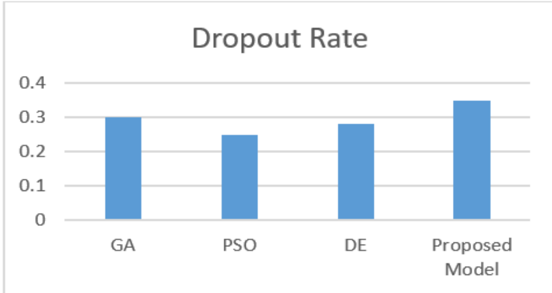

Figure 4. Dropout w.r.t to different algorithms

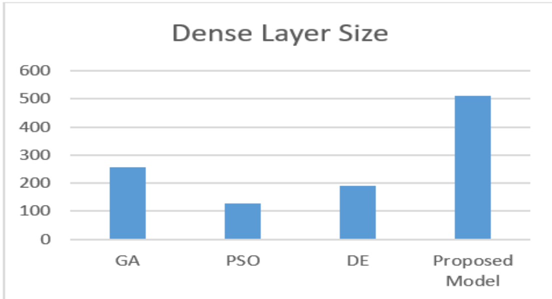

Figure 5. Dense layer size w.r.t proposed and other algorithms

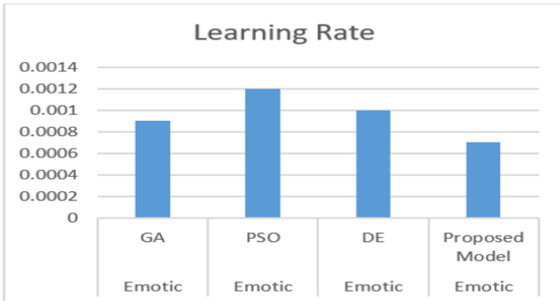

Figure 6. Learning rate of the Emotic dataset w.r.t other algorithms

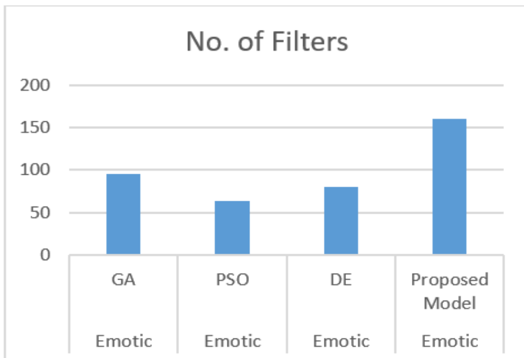

Figure 7. Number of filters of Emotic w.r.t proposed algorithms

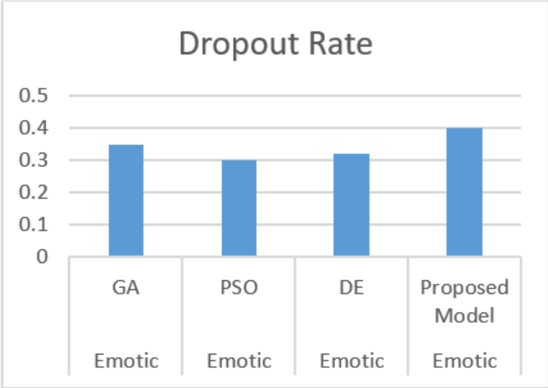

Figure 8. Dropout of Emotic w.r.t to different algorithms

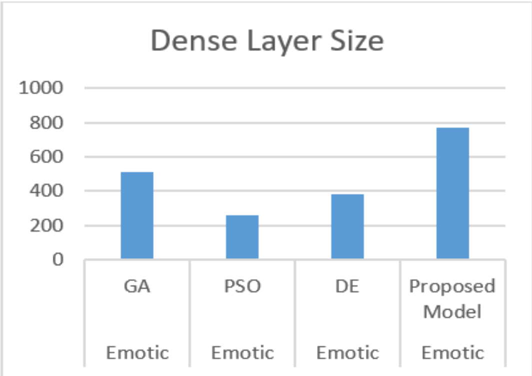

Figure 9. Dense layer size Emotic size w.r.t proposed and other algorithms

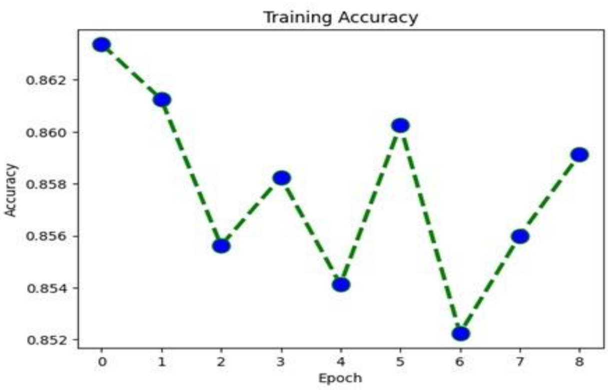

Figure 10. Accuracy using Emotic dataset

The enhanced GNN is then applied to the Emotic dataset [26], where we obtained 87 percent accuracy with 8 epochs. Achieving high accuracy with a small number of epochs might indicate that your model has been effectively regularized. Regularization techniques, such as dropout or weight decay, can help prevent overfitting, allowing the model to generalize well from a limited amount of training data. The loss and accuracy is shown in Figures 10 and 11. The Emotic dataset comprises images saved as NumPy arrays, and each image is associated with an extensive set of annotations. These annotations encompass 26 different emotion categories, as well as the three typical continuous dimensions of Valence, Arousal, and Dominance.

Training the model to identify the patterns in the images that match the various emotions, as presented in Figure 10. After training, the model is evaluated using a separate portion of the dataset that it has never seen before. This evaluation dataset is used to assess the model's performance. For each image in the evaluation dataset, the model predicts an emotion label based on the patterns it has learned during training. The accuracy is then calculated by dividing the number of correct predictions by the total number of predictions made. It is expressed as a percentage.

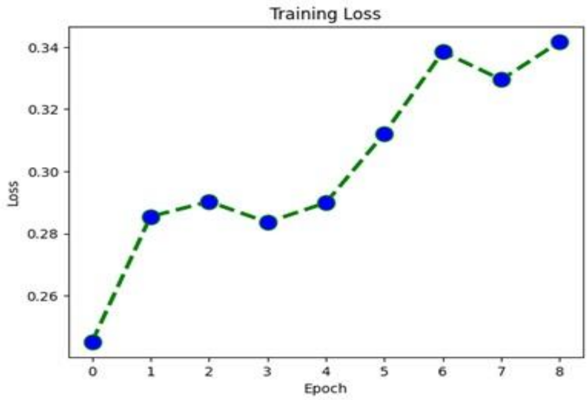

Figure 11. Loss using Emotic dataset

The improved GNN is then used on the FER2013 data, as presented in Figure 11. This emotional analysis dataset has 35,685 samples of grayscale images of faces with a resolution of 48 by 48 pixels, and it is split into a train dataset and a test dataset. Then, the FER2013 dataset [27], where images are sorted into several categories according to the feeling conveyed by the subjects' facial expressions (joy, neutral, sorrow, rage, surprise, disgust, and fear).

A confusion matrix is a valuable tool used to assess the performance of a classification model, specifically in the context of emotion analysis. It provides a structured representation of the predicted and actual labels, enabling the evaluation of the model's accuracy and identification of any misclassifications or errors. The confusion matrix for emotion analysis consists of four main components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this matrix shown in Figure 9, the rows correspond to the actual or ground truth emotions, while the columns represent the predicted emotions. Each cell in the matrix contains the count or percentage of instances falling into that particular category.

•TP: It represents the number of instances where the model correctly predicted a positive emotion.

•TN: It indicates the number of instances where the model accurately predicted a negative emotion.

•FP: It denotes the number of instances where the model incorrectly predicted a positive emotion when the actual emotion was negative.

•FN: It signifies the number of instances where the model incorrectly predicted a negative emotion when the actual emotion was positive.

By analysing the values within the confusion matrix, various evaluation metrics such as accuracy, precision, recall, and F1 score can be calculated. These metrics offer insights into the model's performance, including its ability to correctly classify different emotions and identify any biases or areas requiring improvement.

Facial emotion is the expression of emotion using facial expression, as depicted in Figure 12. The human face has the capacity to convey a diverse range of emotions, including joy, sorrow, anger, fear, surprise, disgust, and others. These emotional states are typically communicated through the interplay of facial components like the eyes, eyebrows, mouth, and facial muscle movements.

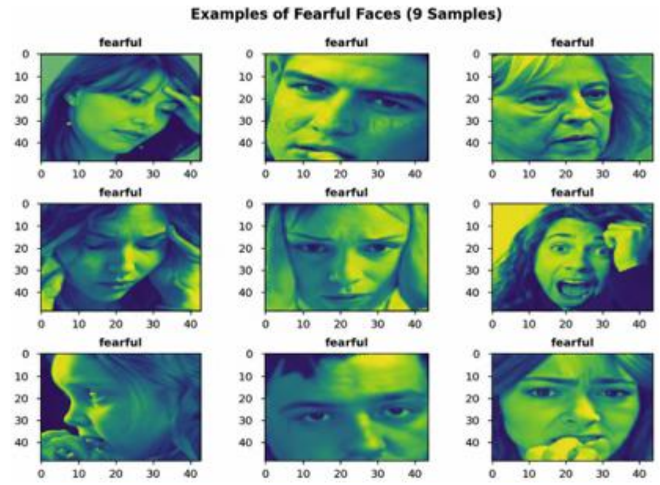

Figure 12. Facial emotion

Figure 13. Labels with emotion

Facial expressions associated with various emotions include:

Happiness: A happy countenance often involves a smiling expression, elevated cheeks, and the presence of fine lines around the eyes known as "laugh lines."

Sadness: A sad demeanor is marked by a downturned mouth, drooping eyelids, and potentially teary eyes.

Anger: The angry face can have narrowed eyes, furrowed brow, and tight or frowning lips, as depicted in Figure 13.

Fear: A fearful look typically encompasses wide-open eyes, raised eyebrows, and a slightly parted mouth.

Surprise: Surprise is evident through wide-open eyes, raised eyebrows, and a slightly agape mouth.

Disgust: Disgust is often conveyed through a wrinkled nose, raised upper lip, and narrowed eyes.

Contempt: Contempt is characterized by a slight, one-sided curl of the lip.

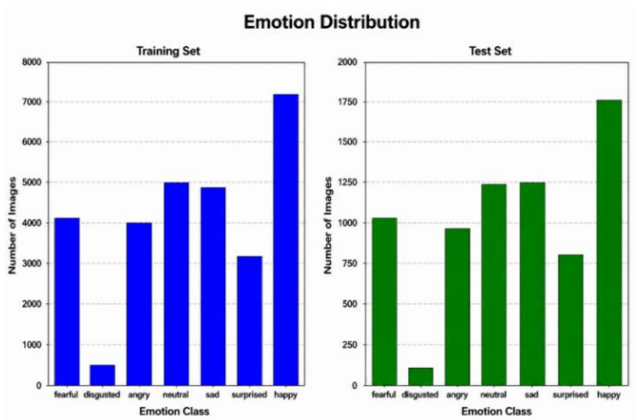

Figure 14. Emotion distribution

In our method, we are using a GNN. Measuring the performance of the model in terms of the distribution of emotions when training and testing the model, as illustrated in Figure 14. The ability to perceive and interpret these facial expressions is fundamental to human communication and social interactions. It facilitates our comprehension of others' emotions and intentions, which is essential for empathy and successful interpersonal relationships.

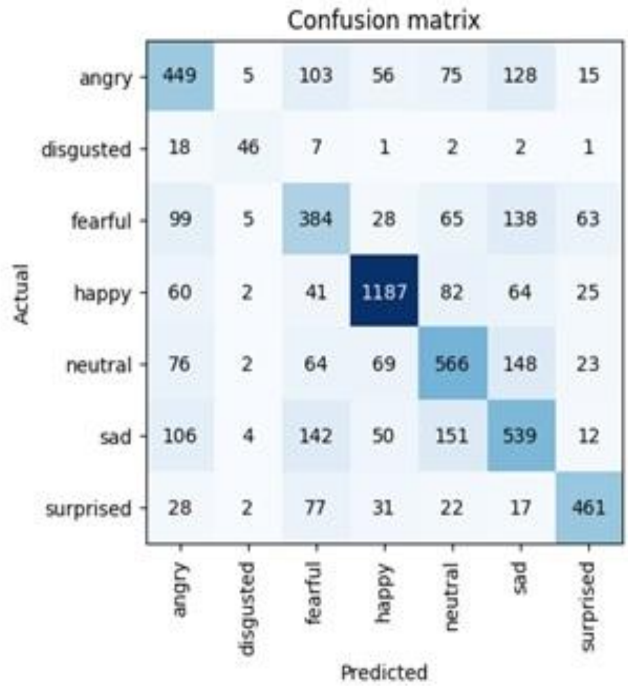

Figure 15. Confusion matrix

A confusion matrix is a table that helps us to know the level at which a machine learning model is performing by displaying the number of correct and incorrect predictions of each class or category, as represented by Figure 15.

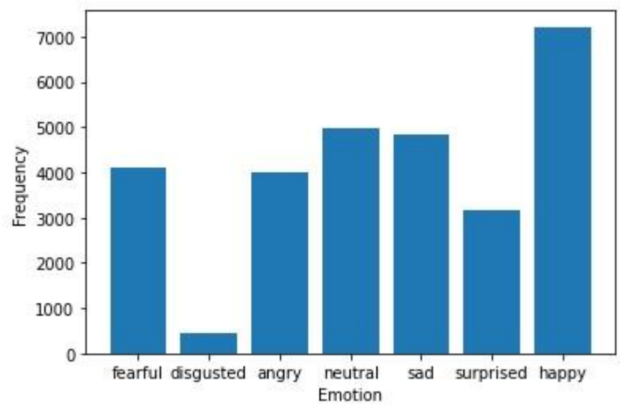

Figure 16. Frequency distribution of emotions

Figure 16 shows the proportions of different types of emotions in the dataset, and the classes are disproportional. The group that is labelled as happy has the highest frequency, and the category that is disgusted has the least number of times that it appears. The frequencies of other emotions, such as neutral, sad, fearful, and angry, are found to be moderate, whereas the frequency of surprise is comparatively low, which might affect the model's performance in various classes.

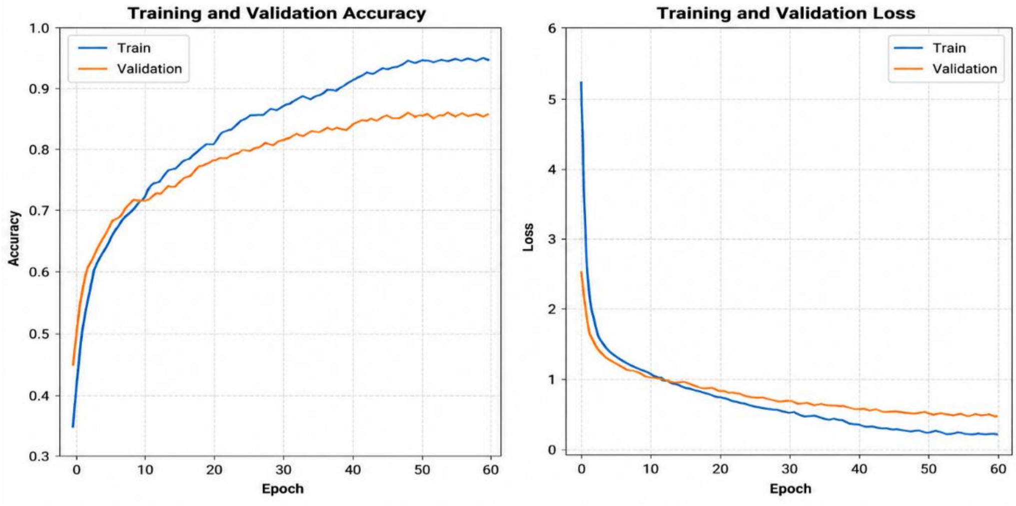

Figure 17. Training and validation accuracy and loss

Table 1. Comparison analysis of accuracy w.r.t other algorithms

Algorithm	Loss	Accuracy	Validation Loss	Validation Accuracy	Epoch
ResNet	1.0403	0.60207	1.1292	0.5705	45
Convolutional Neural Network (CNN)	.3704	0.866	2.4208	0.4911	25
Optimal CNN	1.3592	0.6346	1.369	0.6306	41
Inception V3	0.6155	0.7756	0.9216	0.6753	20
Enhanced Genetic Neural Network (GNN)	**0.2471**	**0.9399**	**0.5121**	**0.8626**	**10**
Meta pseudo labels	0.4596	0.45	0.45	0.45	30
Genetic algorithm	0.2	0.428	0.18	0.42	-

During the training phase of the CNN architectures, the accuracy and loss metrics are used to evaluate and optimize the model's performance as shown in Table 1. The training accuracy measures the proportion of correctly classified emotions in the training dataset, while the training loss quantifies the error between the predicted emotions and the true labels during training. The goal is to achieve high training accuracy and minimize the training loss as the model learns from the training data. Similarly, the validation accuracy and loss are computed on a separate validation dataset to assess the model's generalization ability. The validation accuracy indicates how well the model performs on unseen data, while the validation loss quantifies the error in the model's predictions for the validation dataset. The objective is to obtain high validation accuracy and low validation loss, indicating that the model can accurately predict emotions in real-world scenarios. By monitoring the training and validation accuracy and loss, researchers can determine the effectiveness of the selection, crossover, and mutation operations in improving the fitness scores of the CNN architectures. The final objective will be to find the best CNN structure that will optimize accuracy and minimize loss to allow accurate emotion recognition, as depicted in Figure 17.

5. CONCUSSION

In conclusion, this study introduces an optimized CNN approach for facial emotion analysis on the FER2013 dataset. The proposed technique overcomes several limitations of existing methods and offers several advantages. Firstly, the enhanced fitness evaluation considers additional metrics to assess the distinctiveness and relevance of facial features. This leads to a more accurate evaluation of individuals' fitness and improves the overall performance of the CNN model. Secondly, the improved crossover operation effectively combines informative features from parent individuals, resulting in offspring with enhanced representation of emotional features. This helps to capture the complexities of facial expressions and improves the accuracy of emotion recognition. Furthermore, the adoption of the tournament selection mechanism allows for effective exploration and exploitation of the population. By prioritizing individuals with higher fitness values, the algorithm can focus on individuals with superior emotional classification abilities, leading to improved overall performance. Experimental results demonstrate that the proposed technique achieves superior accuracy in facial emotion analysis compared to existing methods, showing strong performance, achieving ~95% training and ~87% validation accuracy. Loss curves indicate rapid convergence with minor overfitting after 35 epochs. These advantages make the optimized CNN model a valuable tool for various applications, such as affective computing, human-computer interaction, and emotion-aware systems. Overall, the proposed technique overcomes limitations of previous approaches, enhances the accuracy of emotion recognition, and offers valuable insights for further advancements in facial emotion analysis.

ACKNOWLEDGMENT

We would like to show our gratitude to our institution for sharing its pearls of wisdom with us during this research work. We are also immensely grateful to the well-wishers for their comments on an earlier version of the manuscript, although any errors are our own and should not tarnish the reputations of these esteemed individuals.

DATA AVAILABILITY

The data available is derived from:

Saravia, Elvis & Liu, Hsien-Chi & Huang, Yen-Hao & Wu, Junlin& Chen, Yi-Shin. (2018). CARER: Contextualized Affect Representations for Emotion Recognition. 3687-3697. 10.18653/v1/D18-1404.

Kosti, Ronak, Jose M. Alvarez, Adria Recasens, and Agata Lapedriza. "Context based emotion recognition using emotic dataset." IEEE transactions on pattern analysis and machine intelligence 42, no. 11 (2019): 2755-2766.

Kosti, Ronak, Jose M. Alvarez, Adria Recasens, and Agata Lapedriza. "Emotion recognition in context." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1667-1675. 2017.

Kosti, Ronak, Jose M. Alvarez, Adria Recasens, and Agata Lapedriza. "EMOTIC: Emotions in Context dataset." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 61-69. 2017.

REFERENCES

[1] Bisogni, C., Cimmino, L., De Marsico, M., Hao, F., Narducci, F. (2023). Emotion recognition at a distance: The robustness of machine learning based on hand-crafted facial features vs deep learning models. Image and Vision Computing, 136: 104724. https://doi.org/10.1016/j.imavis.2023.104724

[2] Israelsson, A., Seiger, A., Laukka, P. (2023). Blended emotions can be accurately recognized from dynamic facial and vocal expressions. Journal of Nonverbal Behavior, 47(3): 267-284. https://doi.org/10.1007/s10919-023-00426-9

[3] Buetow, S. (2025). Emotionally sensitive thematic analysis. Qualitative Research in Psychology, 22(4): 775-792. https://doi.org/10.1080/14780887.2025.2462919.

[4] Zhang, J., Qin, Q., Liu, X., Ye, Q., Du, W. (2024). Emotion-wise feature interaction analysis-based visual emotion distribution learning. The Visual Computer, 40(3): 1359-1368. https://doi.org/10.1007/s00371-023-02854-6

[5] V, Y., S, Y., R, N., J, R. (2023). Facial emotion recognition system for isolated patient monitoring using ResNet-50 based convolutional neural network model. International Research Journal of Computer Science, 10(5): 154-161. https://doi.org/10.26562/irjcs.2023.v1005.09

[6] Song, R., Giunchiglia, F., Shi, L., Shen, Q., Xu, H. (2023). SUNET: Speaker-utterance interaction Graph Neural Network for emotion recognition in conversations. Engineering Applications of Artificial Intelligence, 123: 106315. https://doi.org/10.1016/j.engappai.2023.106315

[7] Khan, A.R. (2022). Facial emotion recognition using conventional machine learning and deep learning methods: Current achievements, analysis and remaining challenges. Information, 13(6): 268. https://doi.org/10.3390/info13060268

[8] Chen, M. (2022). Emotion analysis based on deep learning with application to research on development of western culture. Frontiers in Psychology, 13: 911686. https://doi.org/10.3389/fpsyg.2022.911686

[9] She, B. (2022). Deep learning-based text emotion analysis for legal anomie. Frontiers in Psychology, 13: 909157. https://doi.org/10.3389/fpsyg.2022.909157

[10] Zhu, Y., Zhu, Y., Ge, N., Gao, W., Zhang, W. (2022). Visual emotion analysis via affective semantic concept discovery. Scientific Programming, 2022(1): 6975490. https://doi.org/10.1155/2022/6975490

[11] Yang, J., Huang, Q., Ding, T., Lischinski, D., Cohen-Or, D., Huang, H. (2023). EmoSet: A large-scale visual emotion dataset with rich attributes. arXiv preprint arXiv:2307.07961. https://doi.org/10.48550/arXiv.2307.07961

[12] Wang, K., Peng, X., Yang, J., Meng, D., Qiao, Y. (2020). Region attention networks for pose and occlusion robust facial expression recognition. IEEE Transactions on Image Processing, 29: 4057-4069. https://doi.org/10.1109/TIP.2019.2956143

[13] Wen, Z., Lin, W., Wang, T., Xu, G. (2023). Distract your attention: Multi-head cross attention network for facial expression recognition. Biomimetics, 8(2): 199. https://doi.org/10.3390/biomimetics8020199

[14] Yang, D., Yu, Z., Yuan, H., Cui, Y. (2022). An improved genetic algorithm and its application in neural network adversarial attack. PloS ONE, 17(5): 0267970. https://doi.org/10.1371/journal.pone.0267970

[15] Zuo, Z. (2023). An enhanced EDBF framework: Adaptive boundary constraint framework (ABCF) for improving multi-parent crossover algorithms. Soft Computing, 27(20): 14689-14697. https://doi.org/10.1007/s00500-023-08501-5

[16] Vilchez, R.N. (2020). Modified selection sort algorithm employing Boolean and distinct function in a bidirectional enhanced selection technique. International Journal of Machine Learning and Computing, 10(1): 93-98. https://doi.org/10.18178/ijmlc.2020.10.1.904

[17] Abbas, Z.K., Al-Ani, A.A. (2023). An adaptive combination algorithm based on deep learning and genetic algorithm for anomalous events detection. Indonesian Journal of Electrical Engineering and Computer Science, 31(2): 902-908. https://doi.org/10.11591/ijeecs.v31.i2.pp902-908

[18] Fei, J., Jia, Q., Chen, G., Li, T., Wang, R., Zhang, X. (2023). Genetic algorithm-based optimal design of modular robot topology based on distributed parallel kinematic modeling and analysis. Engineering Applications of Artificial Intelligence, 123: 10625. https://doi.org/10.1016/j.engappai.2023.106251

[19] Ablordeppey, K.K., Timmaraju, V.A., Song-Yang, J.W., et al. (2020). Development and analytical validation of an expanded mutation detection panel for next-generation sequencing of thyroid nodule aspirates. The Journal of Molecular Diagnostics, 22(3): 355-367. https://doi.org/10.1016/j.jmoldx.2019.11.003

[20] Wang, M., Ma, Y. (2023). A differential evolution algorithm based on accompanying population and piecewise evolution strategy. Applied Soft Computing, 143: 110390. https://doi.org/10.1016/j.asoc.2023.110390

[21] Yuan, G., Sun, G., Deng, L., Li, C., Yang, G. (2024). A novel differential evolution algorithm based on periodic intervention and systematic regulation mechanisms. Applied Intelligence, 54: 11779-11803. https://doi.org/10.1007/s10489-024-05781-8

[22] Umarov, S., Ashurov, R., Chen, Y. (2021). On a method of solution of systems of fractional pseudo-differential

equations. Fractional Calculus and Applied Analysis, 24(1): 254-277. https://doi.org/10.1515/fca-2021-0011

[23] Xu, H., Wu, A., Ren, H., Yu, C., Liu, G., Liu, L. (2023). Classification of colorectal cancer consensus molecular subtypes using attention-based multi-instance learning network on whole-slide images. Acta Histochemica, 125(6): 152057. https://doi.org/10.1016/j.acthis.2023.152057

[24] Hirata, D., Takahashi, N. (2023). Ensemble learning in CNN augmented with fully connected subnetworks. IEICE TRANSACTIONS on Information and Systems, 106(7): 1258-1261.

[25] Roy, S.S., Hsu, C., Kagita, V. (2023). Deep Learning Applications in Image Analysis. Springer.

[26] Kosti, R., Alvarez, J.M., Recasens, A., Lapedriza, A. (2019). Context based emotion recognition using emotic dataset. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(11): 2755-2766. https://doi.org/10.1109/TPAMI.2019.2916866

[27] Saravia, E., Liu, H.C.T., Huang, Y.H., Wu, J., Chen, Y.S. (2018). CARER: Contextualized affect representations for emotion recognition. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, pp. 3687-3697. https://doi.org/10.18653/v1/D18-1404