Expert Systems With Applications 233 (2023) 120947

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Facial Micro Emotion Detection and Classification Using Swarm Intelligence based Modified Convolutional Network

Arun A.N.[a,*], Maheswaravenkatesh P.[b], Jayasankar T.[b]

[a] Assistant Professor, Department of Computer Science and Engineering, Sri Venkateswara Institute of Science and Technology, Chennai, India
[b] Assistant Professor (Sr.Gr), Department of Electronics and Communication Engineering University College of Engineering, Bharathidasan Institute of Technology Campus, Anna University, Tiruchirappalli, Tamilnadu, India

A R T I C L E I N F O

Keywords:
Facial Micro Expression Recognition
Deep Learning
Optical Flow
Modified Artificial Bee Colony
Swarm Optimized Convolutional Neural Network

A B S T R A C T

Emotions are what makes us humans. Recognizing human emotions from facial micro expression features help us learn the true emotional state of a person. This technique of classifying the micro expressions can be used in varied application domains like criminology, marketing, job analysis, online learning etc. The field of recognizing micro emotions deals with tracking, recognizing, estimating & sequencing and classifying the recognized expressions. Artificial intelligence plays a crucial function in modern era of technology; micro expression analysis forms an ideal candidate of Deep Learning to correctly recognize these micro expressions when on display. The aim is to build a system that takes in a video data from any source and to recognize the micro expressions exhibited at various points in time. The challenge to overcome is to capture the fast changing expressions and to extract and align these facial features in order to extract suitable frames that provide the information from which the micro expressions can be ascertained by introducing it to a swarm optimization approach called the Artificial Bee Colony Approach. Implemented a novel approach that captures the essence of the micro expressions by an optical flow vector technique, that supplies its input to the modified deep learning Convolutional Neural Network, that in turn, is trained to categorize micro expressions on display. The Convolutional Neural Network combined with the Swarm Optimizer was able to achieve an accuracy of around 99.45% in identifying & classifying the facial micro expressions.

1. Introduction

Recently, there has been a lot of growth in interest in the relatively new study topic of dissecting face micro expression. The main justification for why micro-expression may be naturally distinguished is a crucial emotional indication for various real lives uses (Sergeeva and Sablina, 2020). A type of non-verbal communication is micro expression that unintentionally expresses a person's genuine feelings. Micro-expression differs from macro-expression in three ways: short duration, slight movement, and difficulty hiding (Nguyen and Yan, 2021; Yadav, 2021; Sivaraman and Manickachezian, 2019). A micro expression is a subtle facial expression that lasts for just a fraction on second or low. Most individuals are unaware of them since they occur in several sections of the face. Micro-expressions may be classified into fundamental expressions, much as macro-expressions: happy, sad, furious, afraid, amazed, disgusted, or neutral. It might be difficult to tell when someone is truly feeling something on their face (Xie et al., 2022; Khattak et al., 2022; Takalkar et al., 2021). Therefore, understanding micro expressions helps us in everyday life because it allows us to tell when someone is trying to hide their emotions or deceive us (Adyapady and Annappa, 2023). Ekman created the Micro Expression Training Tool (METT) to assist individuals identify micro expressions in order to address this difficult issue (Li et al., 2021). According to certain studies, it is still challenging for a trained person to recognize facial micro expressions (Ayyalasomayajula et al., 2021). Average percentage of micro-expression identification in a psychological study utilizing the METT micro expression training data set was 50%. The use of more powerful computer vision technologies can address these issues by improving detection and classification performance (Banerjee et al., 2022; Du et al., 2020; Sivaraman and Manickachezian, 2018). A collection of machine learning techniques called "deep learning" are scientifically based on how the brain is organized. From numerous labelled images, modified convolutional neural networks (CNN) can learn hierarchical features (Esmaeili et al., 2022; Ghosh and Sinha, 2022). It was utilized for many

* Corresponding author.
E-mail addresses: arun.svist.cse@gmail.com (A.N. Arun), pmvenkatesh76@gmail.com (P. Maheswaravenkatesh), jayasankar27681@gmail.com (T. Jayasankar).

https://doi.org/10.1016/j.eswa.2023.120947
Received 19 January 2023; Received in revised form 2 May 2023; Accepted 3 July 2023
Available online 4 July 2023
0957-4174/

different things, such object, activity, and location identification, face recognition, and facial expression recognition (Alphonse et al., 2021). A number of image classification tasks have lately shown good results using the modified convolutional neural network (Annadurai et al., 2023). It is a useful technique for recognizing facial micro expressions because to the meticulous employing convolution, pooling, and layered architecture to translate local to global feature learning, which results in excellent visual representation ability.

Facial micro-expressions, which are very brief facial expressions that occur within 1/25th to 1/5th of a second, can reveal a person's true emotions and intentions, even if they are trying to conceal them. Detecting and classifying micro-expressions automatically is a challenging task owing to subtle and fast nature. Traditional computer vision approaches for facial expression recognition are based on handcrafted feature extraction methods, which are often limited in their ability to capture the complex and subtle facial features associated with micro-expressions. The ability to automatically detect and classify facial micro-expressions has significant potential applications in areas such as lie detection, emotion recognition, and human–computer interaction. Moreover, existing techniques for micro-expression recognition are used shallow learning techniques and have limited performance, especially in terms of accuracy and efficiency. Therefore, there is a need to develop more advanced and efficient approaches to increase accuracy as well as efficiency of facial micro expression recognition. Proposed approach, which utilizes swarm intelligence-based modified convolutional networks, is expected to attain higher performance than existing methods by leveraging collective intelligence of a swarm and the powerful feature learning capabilities of deep convolutional networks.

Important research contributions are summarized below:

- Design multistep image pre-processing and feature extraction process that increases the categorization performance.
- In particular Contrast Limited Adaptive Histogram Equalization (CLAHE) (Jebadass and Balasubramaniam, 2022) to remove the noise and contrast levels of the images of identical persons and expressions.
- Then use empirical wavelet transforms (Liu et al., 2022) to extract representative features of pre-processed picture.
- The features are fed into modified convolutional neural network (Xu et al., 2022) for categorization purpose.
- CK+ dataset is used for testing and training. Recognition tests are performed on seven basic expressions, like happy, sad, angry, terrified, astonished, disgusted, or neutral.
- Examples are provided for benchmark comparison and additional testing of the classification system's robustness.

The remaining section of this manuscript is organized as: the Literature survey is portrayed in section 2, the proposed methodology is illustrated in section 3, result with discussion is explained in section 4, and lastly, this manuscript is concluded in section 5.

2. Related work

Several research were suggested in the literature related to deep learning based Facial Micro Emotion Detection; a few recent works are reviewed here,

Arul Vinayakam Rajasimman et al., (2022) have presented Robust Facial Expression Recognition Utilizing an Evolutionary Approach along Deep Learning Method. This paper presents a unique Robust Facial Expression Recognition RFER-EADL method. RFER-EADL uses DL models and computer vision to identify different kinds of emotions. To standardize intensity and contrast identical people photos, emotions, RFER-EADL first performs histogram equalization. The densely connected network (DenseNet-169) method based on deep convolution neural networks was then used in conjunction along COA method for hyper parameter tuning. Finally, expression recognition and classification are performed using TLBO using long short-term memory (LSTM) method. The proposed method offers better accuracy and lesser recall.

Sikkandar, and Thiyagarajan, (2021) have presented Deep learning depended facial expression recognition utilizing enhanced Cat Swarm Optimization. The ICSO algorithm, a modified version of CSO approach, was proposed in this study as a revolutionary method for recognizing human face expressions. The suggested system uses an input image to find comparable pictures in data set and to identify person's emotional state based on their expressions. The DCNN technique was utilized to extract deep characteristics present in the facial image. It was suggested that ICSO pick out the best features from a face picture were distinguish a person's appearance on their face in a particular way. The proposed method provides lower F1 score and higher precision.

Durga and Rajesh, (2022) have presented A Res-Net deep learning depended facial recognition design for future multimedia applications. CNN technique has been recognized to use deep micro-facial emotion. The 2D-ResNet CNN was the first to extract multiple class features. A 2D-ResNet CNN multiple class classifiers were presented to identify the mask able images of facial emotions. By reducing over fitting and contentious issues, the 2D-ResNet deep learning method was trained using open-source public dataset JAFFE. The proposed method provide lower accuracy and higher precision.

Lakshmi and Ponnusamy, (2021) have presented Facial emotion recognition utilizing modified HOG and LBP features along deep stacked auto encoders. To locate eye, nose, and mouth regions utilizing the Viola-Jones algorithm, the face region was first detected using viola-Jones face detection. The identified area was then enhanced using a Butterworth high pass filter. The proposed modified HOG and features of identified eye, nose, and mouth areas were extracted utilizing LBP feature descriptors. Using Deep Stacked Auto Encoders, the retrieved features from these three regions were combined and their dimensionality was decreased. Finally, classification and recognition are accomplished utilizing multiple classes Support Vector Machine. The proposed method provides higher accuracy and lower precision.

Jarraya et al., (2020) have presented Compound emotion recognition of autistic children during meltdown crisis utilizing deep spatio-temporal analysis of facial geometric features. In this work, deep spatiotemporal geometric elements of autistic children's micro expressions during a meltdown crisis are experimentally evaluated. To achieve this, you should compare the CER performance to other collections of micro expression features in order to find the characteristics that autistic children's meltdowns from normal condition are most likely to have, and most accurate classifier. The proposed method provides higher recall and lower f1 score.

Li et al., (2020) have presented Joint local and global information learning along single apex frame identification for micro-expression recognition. To identify apex frame, this research first suggests measuring pixel-level change amounts at frequency domain. It outperforms previous apex frame spotting algorithms depending on spatio-temporal change information in terms of apex frame spotting when frequency information is included. Because not all regions contribute equally to ME recognition and few areas don't even possess emotional information, this research also presents a joint feature learning architecture combining local and global information to recognize MEs. More specifically, proposed model combines the global information learnt from entire face with local information learned from facial areas giving most significant emotion info. The proposed method provides higher precision and lower accuracy.

Thuseethan et al., (2020) have presented Complex emotion profiling: An incremental active learning depended algorithm along sparse annotations. In this research, suggest a comprehensive architecture to gradually and actively profile complicated emotions in wild using scant data. Three key elements make up strategy: a pre-processing unit, optimization unit, and active learning unit. In complicated emotion images retrieved from an uncontrolled dataset, the preprocessing unit eliminates variations. With the aid of optimization unit and our innovative incremental active learning algorithm, were successfully forecast complex emotions that exist in wild. Estimation utilizing various complex emotions benchmark data sets demonstrates suggested algorithm effectively profiles complicated emotions at levels that were comparable to human perception capacity. The proposed method provides higher F1 score and lower recall. Table1 shows that the Literature comparison table.

3. Proposed methodology

During this investigation, a novel MCNN based Artificial Bee Colony (ABC) (Kaya et al., 2022; Hossain et al., 2023; Nisamudeen and Zhang, 2023) method was recognized for emotion recognition in facial pictures. The presented MCNN-ABC method utilizes an HE procedure. The EWT (Zahara et al., 2020) is then employed for extract features. Finally, the MCNN is utilized to recognize and classify emotional facial expressions. At last the hyper parameters are tuned using ABC algorithm. Fig. 1 shows that the Block diagram of the RFER-EADL algorithm.

3.1. Preprocessing using CLAHE

The technique that overcomes the less contrast problem for face expression photos is called Contrast Limited Adaptive Histogram Equalization (CLAHE). CLAHE is superior than Adaptive Histogram Equalization (AHE) and conventional Histogram Equalization (HE), particularly in medical imaging, where it outperforms both of those techniques. In essence, CLAHE works by restricting the contrast enhancement and eliminating noise, which are typically carried out by standard HE and result in further noise enhancement. Therefore, in situations where noise became too prominent through limiting contrast enhancement at HE, desired outcomes were attained by enhancing contrast in facial emotion images. The functions of slope connecting the input picture intensity rate to desired resulting picture intensities are used to roughly define contrast enhancement. By lowering slope of related purpose, contrast can be reduced. Additionally, histogram height for intensity rate directly correlates with contrast enhancement. Therefore, the same functions that control the contrast amplification and noise removal processes are limiting slope and clipping the histogram height. In light of the necessity for contrast by choosing clip parameter, user may restrict the contrast. In this CLAHE technique to remove noise and enhance facial emotion image. Initially, CLAHE is utilized for generating Contrast image C_{image} and it is given in the following equation (1)

$$C_{image} = Input\ liver\ image{*}CLAHE \tag{1}$$

Then the output pre-processing image are obtained based on the following equation (2)

$$Output\ pre-processed\ image = U{*}C_{image} + V \tag{2}$$

where U and V are linear coefficients and it is calculated based on the following equation (3)

$$F(U,V) = [Output\ pre-processed\ image - C_{image}]^2 + \xi{*}U^2 \tag{3}$$

Finally, the pre-processed facial emotion image is presented towards feature extraction stage.

3.2. Feature extractions utilizing empirical wavelet transform (EWT)

For optimum derivation of features associated to facial pictures, empirical wavelet transform (EWT) model are utilized. EWT is employed to extract Haralick texture from pre-processing image. At the time of extraction, the EWT emulated certain features, like it lessen the difficulty and over fitting problem while classification. The mean, contrast, homogeneity is the Haralick texture features that are extracted through empirical wavelet transforms. To extract image feature, the given steps are applied;

First, set the count of empirical wavelet transforms at original image. The preprocessed image denotes a. Then, scale 1 dimensional Fourier transforms for each row R of a indicates $A(R,\omega)$, columns C of a specifies $A(\omega,C)$, the empirical wavelet transforms transform including matrix operations exhibited in equation (4), (5) as follows,

$$A_{mean\ \ row} = \frac{1}{num_{R\omega}} \sum_{R=0}^{num_{R\omega}} A(R,\omega) \tag{4}$$

$$A_{\mathrm{cotrast\ column}} = \frac{1}{num_{C\omega}} \sum_{C=0}^{num_{C\omega}} A(\omega,C) \tag{5}$$

where $num_{R\omega}$ and $num_{C\omega}$ denotes number of rows, columns. The image boundaries detect the Haralick texture features along homogeneity with respect to rows and columns of empirical wavelet transforms associated with filter bank $\left\{\zeta_1^R, \{\xi_{\mathrm{homo}}^R\}_{\mathrm{homo}}^{Num_R}\right\}$ and $\left\{\zeta_1^C, \{\xi_{\mathrm{homo}}^C\}_{\mathrm{homo}}^{Num_C}\right\}$ respectively. Where homo implies homogeneity feature, Haralick texture feature as well as homogeneity features are extracted via EWT is expressed in eqn (6),

$$A_{\mathrm{homo}} = \sum_{x=0}^{j=1} \sum_{y=0}^{i=1} \frac{1}{1+(x-y)^2} g(x-y) \tag{6}$$

Let i,j specifies input picture pixels coordinates along middle of local area picture, x,y specifies preprocessed imageries. The feature extracted imageries including rows $\left\{\zeta_1^R, \{\xi_{\mathrm{homo}}^R\}_{output}^{Num_R}\right\}$ gives num_R+1 output feature extracted images using mean, contrast, homogeneity. This num_R+1 output feature extracting imageries and columns $\left\{\zeta_1^C, \{\xi_{\mathrm{homo}}^C\}_{output}^{Num_C}\right\}$ gives $(num_R+1)(num_C+1)$ along sub band imageries.

3.3. Facial expression classification using modified convolutional neural network

The updated CNN classifier is employed to advance the micro facial emotion categorization utilizing characteristics extracted from face picture. The ABC aims to find optimum weights for MCNN tuning classifier to categorizing facial emotions, like happy, sad, angry, terrified, astonished, disgusted, or neutral facial emotions, extracted features are given to the modified Convolutional Neural Network.

Table 1
Literature comparison table.

Author	Method	disadvantage	Advantages
Arul Vinayakam Rajasimman et al., (2022)	Evolutionary Algorithm with Deep Learning	Lower Recall	Higher Accuracy
Sikkandar and Thiyagarajan, (2021)	Deep convolution neural network	Lower F1 Score	Higher Precision
Durga and Rajesh, (2022)	2D-ResNet convolutional neural network	Lower Accuracy	Higher Recall
Lakshmi and Ponnusamy, (2021)	Histogram of Oriented Gradients. Local Binary pattern. Deep stacked auto encoder.	Lower Precision	Higher Accuracy
Jarraya et al., (2020)	Deep Spatio-Temporal	Lower F1 Score	Higher Recall
Li et al., (2020)	Joint Local as well as Global Information Learning	Lower Accuracy	Higher Precision
Thuseethan et al., (2020)	ACTIVE LEARNING UNIT	Lower Recall	Higher F1 Score

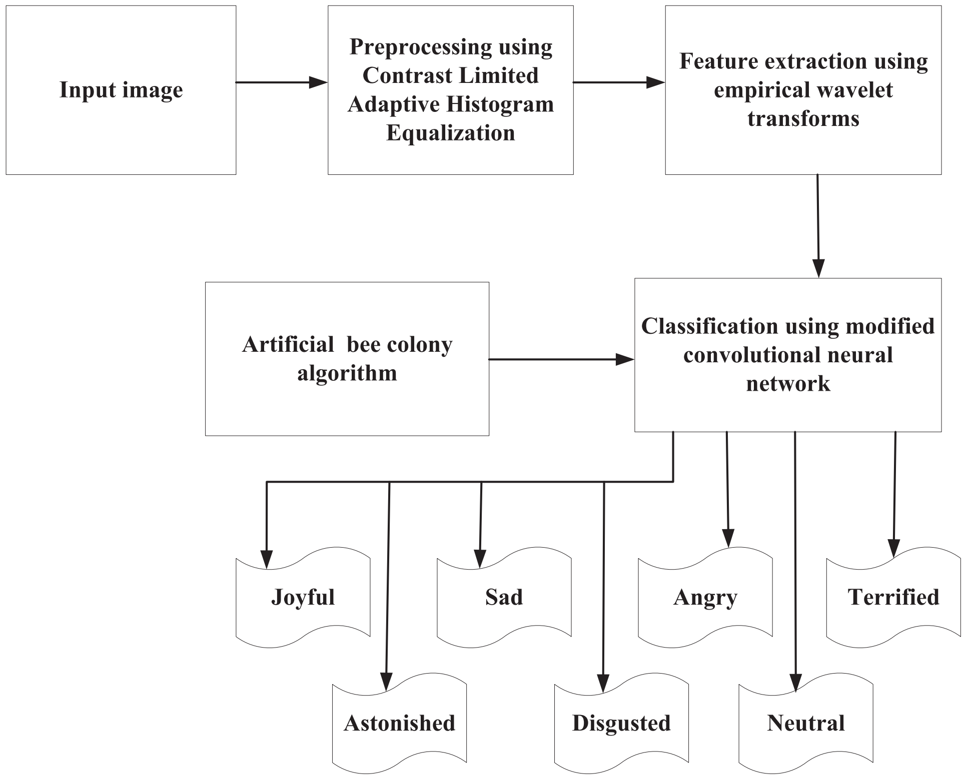

Fig. 1. Block diagram of the RFER-EADL algorithm.

3.4. *Structure of deep CNN*

The Facial Micro Emotion categorization is carried out by MCNN utilizing morphological information as input. The three layers of MCNN are convolutional layers, pooling layers, and fully connected (FC) layers. The neurons in the MCNN's succeeding layers are coupled to a patch of neurons, as opposed to the regular NN, in which one neuron is linked to other. Each layer in the MCNN performs its own duty, including feature acquisition, subsampling, and categorization.

3.5. *Conv layers*

Convolution layer's goal is to extract limited features that are hidden inside the input feature vector before to applying convolution filters. Receptive fields transmit input characteristics to the following layer, and trainable weights enables connection to the succeeding layers. The convolutional operator used in the conv layer's function convolves input data along kernel filter and weights that is computed optimally utilizing ABC. As a result, a collection of restricted features is generated at the end of each set of layers, which is then provided at input to following convolutional units. The MCNN has a certain number of convolutional layers. It is important to remember that the accuracy of classification is correlated with convolution count. The MCNN classifier's accuracy is inversely related to the total convolution count. The output from b^{th} convolutional layer is given at equation (7),

$$\left(\overline{C}_h^b\right)_{R,S} = \left(A_h^b\right)_{R,S} + \sum_{f_1=1}^{u_1^{f_1-1}} \sum_{f_2=-u_1^b}^{u_1^b} \sum_{f_3=-u_2^b}^{J_2^b} \left(\beta_{h,u_1}^b\right)_{\gamma,\nu} * \left(\overline{C}_{u_1}^{b-1}\right)_{R+u_2,Q+u_3} \tag{7}$$

here $*$ denotes convolutional operator. From the input of previous convolutional layer, the convolutional layer b extracts local patterns, from the input of earlier convolutional layer. $\left(\overline{C}_h^b\right)_{R,S}$, represents fixed feature map; output from b^{th} convolution is placed in (R,S). The output from earlier $(b-1)^{th}$ convolution input to b^{th} convolutional layer. The weights of convolutional a layer that is denoted at β_{h,u_1}^b, indicates weights of b^{th} convolution and signifies the bias of b^{th} convolutional layer at A_h^b. h count of convolution and notations, like $f_1 f_{2,}\ and f_3$ represent feature maps. The feature maps are results of applying a convolutional filter on the input feature vector to each individual convolutional layer. In order to extract characteristics at all dimensions, neurons present in convolutional layers is further systematized in three dimensions, like width, height, and depth. Convolutional layer output is passed to ReLU layer. It is given at equation (8),

$$\overline{C}_h^b = hm\left(\overline{C}_h^{b-1}\right) \tag{8}$$

The capacity of ReLU layer can handle a large number of layers in the classifier enhances deep CNN for extracting relevant characteristics on categorization.

3.6. *POOL layers*

The POOL layer, which is thought to function consistently, is a non-parametric layer free of bias and weight. 100% interconnected layers: The categorization is done in FC layer, and the convolution layers are employed to extract confined features from the features. The FC layer categorization permits the feature categorization to proclaim the patients with happy, sad, angry, afraid, amazed, disgusted, or neutral emotions. FC layer output is given at equation (9),

$$\overline{C}_h^b = \alpha \overline{C}_h^b\ with \sum_{f_1=1}^{u_1^{f_1-1}} \sum_{f_2=-u_1^b}^{u_1^b} \sum_{f_3=-u_2^b}^{J_2^b} \left(\beta_{h,u_1}^b\right)_{\gamma,\nu} * \left(\overline{C}_{u_1}^{b-1}\right)_{R+u_2,Q+u_3} \tag{9}$$

The accurate categorization of facial emotion in terms of weights as well as biases used in classifier, it is determined by utilizing ABC algorithm. The achieved outcomes specify the MCNN method has effectually predicted facial emotion. To get more accurate evaluation of the facial emotion detection weight parameter of MCNN $\vec{C}_h^b$ method are enhanced using Artificial Bee Colony approach. It is described below,

3.6.1. Step by step procedure of Artificial bee Colony (ABC) approach for optimizing MCNN

The ABC is used to optimize the parameters of MCNN method to getting optimal limits. These limits are optimized to calculating optimum limits on promising accurate prediction for facial emotions. ABC approach is meta-heuristic algorithm which enthused through intelligent foraging behavior of a honeybee swarm. Three types of bees are included in ABC's artificial bee colony: Observer bees observing employed bees dance within the hive to choose a food source, scout bees searching randomly for food sources, and employed bees tied to specific food sources. Unemployed bees are a term used to describe both scouts and bystanders. All food source positions are discovered through scout bees. In ABC, food source position represents a possible solution for the prediction of facial emotions and nectar content of a food source correlates with effectiveness (fitness) of corresponding solution. Since every employed bee is linked to one and only source of food, employed bees count equals food sources count. The procedure of Artificial Bee Colony are given below,

Step 1: Initialization

Based on upper and lower boundaries of scout bees' production power and control settings, food sources' population was initialized at random. The initial populace of bees is given at eqn (10),

$$I = \frac{Q_s}{4\pi d^2} \tag{10}$$

here Q_s, d represents lower and upper bounds.

Step 2: Random Generation

After initialization, input parameters have been generated randomly. The maximum fitness values are selected using accurate hyper parameter. Here, randomly produce the populace for accurate prediction of facial emotions.

Step 3: Fitness Function

The iterative search procedures are initiated, the initial population is established along with fitness, and subpopulation colour is assigned. The ranges of iterative search then begin with 1 and go up to highest repetition number. The fitness function is estimated to achieve objective function, like an accurate prediction of facial emotion to achieve the optimum value. Utilizing ABC, the MCNN $\vec{C}_h^b$ weight parameters are optimised. It is given at eqn (11),

$$Fitness = optimization[\vec{C}_h^b] \tag{11}$$

Step 4: Unemployed foragers

Jobless foragers contains binary types of bees includes scouts and onlookers. Exploring and exploiting source of food are the major task. Foragers looking for work have binary options. It first transforms into a scout that haphazardly looks for fresh sources of food near nest. After monitoring the waggle dances of hired bee, an observer now estimates the nectar amount of food source and chooses the food source based on profitability. The jobless foragers appearances are computed through using eqn (12),

$$y_i^{g+1} = y_i^g + d * \frac{\gamma_i^g * \left(\beta_{gpop} * \left(y_{gpop}^g - y_i^g\right) l_i + \beta_{bj} * \left(y_{bj}^g - y_i^g\right)\right)}{2} \tag{12}$$

where y_i^g specifies earlier position i^{th} scouts, d represents uniformly distribution random count of selection at $[0, 1]$, γ_i^g denotes arbitrary coefficient at i^{th} specifies bees at repetition g, β_{gpop} and β_{bj} denotes point in random among present location and target facial emotions,y_{gpop}^g signifies best optimum location, y_{bj}^g denotes position of j^{th} jobless foragers.

Step 5: Update the position of Unemployed foragers

Every jobless forager at ABC may revise their status through improved by predetermined trials number, specified through user of ABC approach. Hence position updation of unemployed foragers is expressed at eqn (13),

$$\gamma_i^g = d * \mathrm{Tan}sig\left(1 - \frac{g}{g_{\max}}\right)\eta_i \tag{13}$$

here d represents random count of normal distribution at $[0, 3]$, Tan*sig* signifies tangent sigmoid functions, g denotes present repetition count and $g_{\max}$ indicates maximum number of iterations.

Step 6: Employed foragers

The honeybees found food source, which also known as employed bees, are equal to the food sources count. From this facial emotions were assessed. The working bees keep track of the food source information and distribute it to others with a given probability. When source of food is gone, the employed bee will become scout. Exploration on employed foragers for better solutions are obtained by equation (14),

$$\beta = \frac{1}{1 + NI_j^i} \tag{14}$$

here β represents random selecting a source of food j, i signifies employed bees and NI_j^i denotes new source of food. Fig. 2 displays flowchart for ABC for optimizing MCNN.

Step 7: Update the position of employed foragers for optimizing $\tau_{\ln}^*$

In ABC, the probability values derived from the fitness values given by employed bees determine employed foragers' current position. A fitness depended selection method is used. It is acquired utilizing eqn (15),

$$N(M) = \left\{R\left(\frac{S}{2} * \left(1 - \frac{g}{g_{\max}}\right)\right) + 1\right\} \tag{15}$$

here future location count predicted at every repetition is specified at $N(M)$, g denotes present repetition count, $g_{\max}$ represents best optimum locations, S signifies maximum repetitions count and R indicates round.

Step 8: Termination Condition

The optimal hyper-parameter $\vec{C}_h^b$ are selected MCNN based on ABC, will repeat step 3 till halting criteria. Finally MCNN predicts facial macro emotions through utilizing ABC.

4. Results and discussion

The experiment was performed on a computer having an Intel i7 core processor with 32 GB of RAM and an Nvidia model GTX 1080 GPU which has 2560 CUDA Cores and a memory bandwidth of 320 GB/s. The experiment was also ported to run on TPU architecture on the Google cloud platform. A total of 28,709 image files were taken to train the network in 7 of the micro expressional classes. The training takes around an average of 42 h on the GPU setup and close to 7 h on the Google cloud v3-8 TPU architecture employing 8, v2 cores.

The training set contains a mix of all homologically aligned images from which the micro expressions are identified.

The hyper parameter values that were able to produce enhanced and optimal outcomes are displayed at Table 2.

4.1. Dataset description

The experimental rationale of RFER-EADL method utilized the CK + dataset, which consists of 837 images with seven class labels, as shown in Table 3. Fig. 3 displays a few illustrations. 28,000 labelled photos make up the training set for the FER-2013 data set. 3500 photographs are included in test set, whereas the development set contains 3500 labelled pictures. FER-2013 assigns one of seven emotions to every image: happy, sad, furious, afraid, amazed, disgusted, or neutral. By this 837 images are used for testing and 27,163 images are used for training.

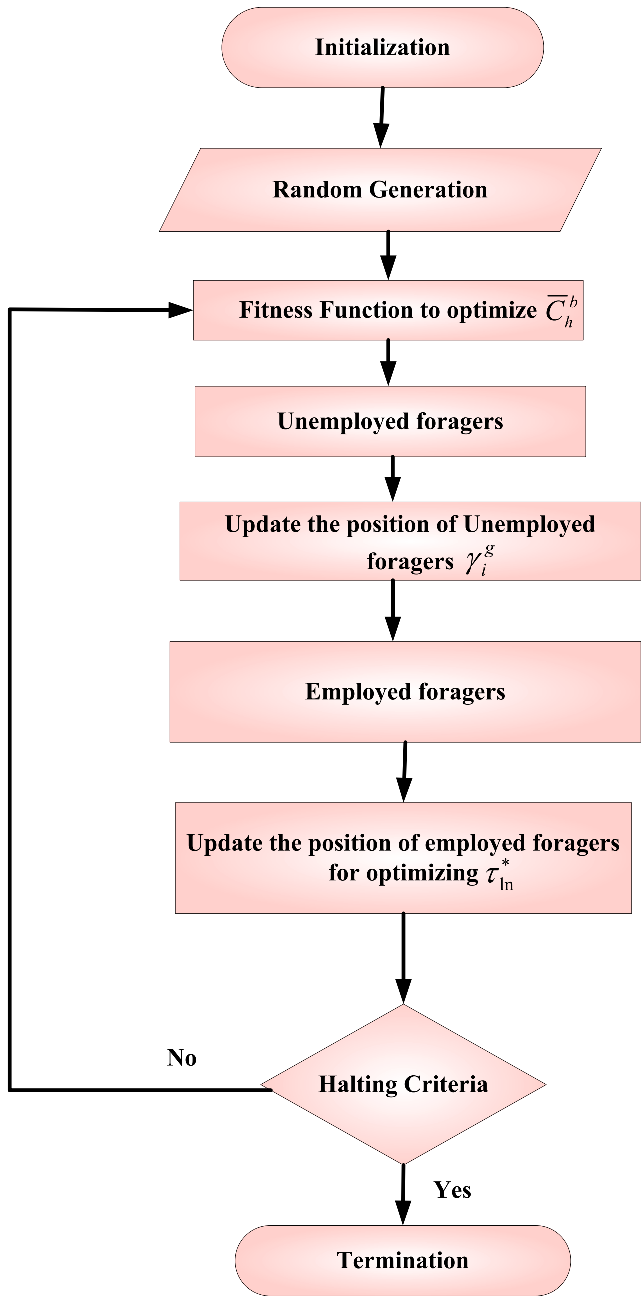

Fig. 2. Artificial bee colony approach for optimizing MCNN.

Table 2
Hyper parameters with optimal values obtained using ABC algorithm.

Parameters	Optimal Values
Total Number Of Layers	159
Fully Connected Layers	3
Optimizer	Artificial bee colony
Filter Dimension Size	3×3
Batch Size	64
Stride	2
Learning Rate	$1.46e^{-3}$

4.2. Performance metrics

Several parameters have been utilized to conduct experiment, estimate scheme performance. To measure performance of proposed method with accuracy, f-measure, precision, sensitivity, specificity, and error rate are evaluated. Mathematical values as well as the definitions are given below,

Table 3
Dataset details.

Description	Number of images
Joyful	69
Sad	28
Angry	45
Terrified	25
Astonished	18
Disgusted	59
Neutral	593
Total	837

4.2.1. Accuracy

A percentage of exact predictions to the total count of records in data set equals accuracy. It is given at eqn (16),

$$Accuracy = \frac{(T(P) + T(N))}{(T(P) + T(N) + F(P) + F(N))} \tag{16}$$

$(T(P))$ True Positive, $(T(N))$ True Negative, $(F(P))$ False Positive, $(F(N))$ False Negative respectively.

4.2.2. Precision

Precision is described as the classifier’s capacity to calculate the typical assault under normal circumstances. It is given in eqn (17),

$$Precision = T(P)/(T(P) + F(P)) \tag{17}$$

4.2.3. Sensitivity

Sensitivity refers to the amount of actual positives that are accurately foreseeable. It is given in eqn (18),

$$Sensitivity = \frac{T(P)}{F(N) + F(P)} \tag{18}$$

4.2.4. Specificity

Specificity is also called TN ratio. Thus it is given in eqn (19),

$$Specificity = \frac{F(N)}{F(P) + F(N)} \tag{19}$$

4.2.5. F-Measure

As the harmonic mean of recall and precision, F-Measure is definite.

$$F - Measure = \frac{2\,T(P)}{2\,T(P) + F(P) + F(N)} \tag{20}$$

Table 4 shows the accuracy of the Facial Micro Emotion Detection. Here, the proposed FMEC-MCNN-ABC method attains higher accuracy 19.33%, 23.46%, 24.55% and 26.46% for joyful; higher accuracy 25.25%, 26.28%, 27.42% and 28.52% for sad; higher accuracy 28.33%, 29.46%, 30.55 and 32.35% for terrified; higher accuracy 22.33%, 23.46%, 25.55% and 27.35% for astonished; higher accuracy 25.33%, 26.46%, 27.05% and 28.25% for disgusted; higher accuracy 28.33%, 29.46%, 30.55% and 32.25% for neutral is compared with existing techniques like, FMEC-EADL-TLBO-LSTM, FMEC-DCNN-ICSO, FMEC-2D-ResNet-CNN and FMEC-MCNN (without optimization) respectively.

Table 5 shows the precision of the Facial Micro Emotion Detection. Here, the proposed FMEC-MCNN-ABC method attains higher precision 29.13%, 33.06%, 34.05% and 36.93% for joyful; higher precision 35.25%, 36.08%, 37.44% and 38.14% for sad; higher precision 38.13%, 39.66%, 40.15 and 42.35% for terrified; higher precision 32.88%, 33.46%, 35.35% and 37.29% for astonished; higher precision 35.03%, 36.06%, 37.48% and 38.15% for disgusted; higher precision 38.73%, 39.42%, 40.25% and 42.47% for neutral is compared with existing techniques like, FMEC-EADL-TLBO-LSTM, FMEC-DCNN-ICSO, FMEC-

Fig. 3. Sample images of the dataset.

Table 4
Performance analysis of accuracy.

Accuracy (%) Methods	Joyful	Sad	Terrified	Astonished	Disgusted	Neutral
FMEC-EADL-TLBO-LSTM	98.23%	97.36%	98.34%	96.35%	97.25%	98.26%
FMEC-DCNN-ICSO	95.24%	97.24%	94.25%	97.35%	98.34%	93.83%
FMEC-2D-ResNet-CNN	98.33%	92.45%	96.39%	97.35%	94.77%	96.36%
FMEC-MCNN (without optimization)	97.34%	96.35%	98.34%	97.87%	98.35%	98.34%
FMEC-MCNN-ABC (with optimization) (Proposed)	99.34%	99.27%	99.83%	99.11%	99.56%	99.38%

Table 5
Performance analysis of precision.

Precision (%) Methods	Joyful	Sad	Terrified	Astonished	Disgusted	Neutral
FMEC-EADL-TLBO-LSTM	94.73%	93.39%	96.94%	95.35%	95.70%	93.27%
FMEC-DCNN-ICSO	94.35%	96.47%	93.27%	95.85%	96.24%	94.03%
FMEC-2D-ResNet-CNN	91.63%	93.75%	95.09%	94.67%	93.70%	92.06%
FMEC-MCNN (without optimization)	94.84%	95.95%	94.39%	95.07%	96.75%	93.73%
FMEC-MCNN-ABC (with optimization) (Proposed)	99.49%	99.07%	99.74%	99.37%	99.90%	99.25%

2D-ResNet-CNN and FMEC-MCNN (without optimization) respectively.

Table 6 shows the f1-score of the Facial Micro Emotion Detection. Here, the proposed FMEC-MCNN-ABC method attains higher f1-score 25.43%, 27.96%, 28.25% and 30.25% for joyful; higher f1-score 21.05%, 24.78%, 26.40% and 28.28% for sad; higher f1-score 29.03%, 33.06%, 35.85 and 37.56% for terrified; higher f1-score 30.82%, 32.06%, 34.75% and 36.28% for astonished; higher f1-score 36.43%, 38.04%, 40.45% and 42.15% for disgusted; higher f1-score 32.43%, 35.47%, 38.85% and 40.14% for neutral is compared with existing techniques like, FMEC-EADL-TLBO-LSTM, FMEC-DCNN-ICSO, FMEC-2D-ResNet-CNN and FMEC-MCNN (without optimization) respectively.

Table 7 shows the sensitivity of the Facial Micro Emotion Detection. Here, the proposed FMEC-MCNN-ABC method attains higher sensitivity 35.03%, 37.26%, 38.85% and 39.45% for joyful; higher sensitivity 31.47%, 34.78%, 36.46% and 38.67% for sad; higher sensitivity 39.83%, 43.83%, 45.05 and 47.36 for terrified; higher sensitivity 27.02%, 32.86%, 34.15% and 36.45% for astonished; higher sensitivity

Table 6
Performance analysis of f1-score.

F1-Score (%) Methods	Joyful	Sad	Terrified	Astonished	Disgusted	Neutral
FMEC-EADL-TLBO-LSTM	95.84%	94.40%	95.05%	96.46%	96.81%	94.38%
FMEC-DCNN-ICSO	95.46%	97.58%	94.38%	96.96%	97.35%	95.14%
FMEC-2D-ResNet-CNN	92.74%	94.86%	96.10%	95.78%	94.81%	93.17%
FMEC-MCNN (without optimization)	98.95%	96.06%	97.40%	95.67%	98.75%	97.73%
FMEC-MCNN-ABC (with optimization) (Proposed)	99.09%	99.87%	99.70%	99.84%	99.45%	99.18%

Table 7
Performance analysis of sensitivity.

Sensitivity (%) Methods	Joyful	Sad	Terrified	Astonished	Disgusted	Neutral
FMEC-EADL-TLBO-LSTM	96.84%	97.24%	96.85%	97.40%	97.62%	95.39%
FMEC-DCNN-ICSO	96.76%	98.50%	95.32%	97.91%	98.38%	96.74%
FMEC-2D-ResNet-CNN	93.77%	95.81%	97.90%	96.72%	95.91%	94.14%
FMEC-MCNN (without optimization)	97.05%	98.56%	97.74%	96.61%	98.05%	98.72%
FMEC-MCNN-ABC (with optimization) (Proposed)	99.43%	99.85%	99.15%	99.86%	99.42%	99.96%

26.03%, 28.45%, 30.25% and 32.45% for disgusted; higher sensitivity 30.63%, 32.57%, 34.92% and 36.87% for neutral is compared with existing techniques like, FMEC-EADL-TLBO-LSTM, FMEC-DCNN-ICSO, FMEC-2D-ResNet-CNN and FMEC-MCNN (without optimization) respectively.

The various hyper parameters like: total number of layers on the deep learning neural network, the number of fully connected layers, dimension of the filters, the batch sizes to consider, the stride of the kernel in scanning the images, the optimization function utilized and the learning rate of the network were considered as in the process for optimization. The proposed technique was able to extract the micro expressions from the images that were assigned to the classifier and to accurately place classification labels from the given set of categories. The class of assignment and the levels of accuracy in placing the labels were represented as a graph as shown in Table 4. Thus results after hyper parameter tuning shows an important improvement in accuracy classification on facial micro expressions that were considered for the study.

Further research is carried out by us to induct systems that can detect more types of complex facial emotions and micro expressions. The training phase can also be modified to encompass features of transfer learning so that when new micro expressional classes are added, it would significantly reduce the time associated for training the convolutional network to stabilize and produce statistically valid classifications.

5. Conclusion

In this study, the ABC Approach was combined with modified convolutional neural network architecture to recognize and classify facial micro expressions. The work mainly involves improving on the efficiency of classification of the expressions by tuning the hyper parameters of the network so that it performs with improved accuracy. This has the added advantage of the network converging within a short amount of time with optimal results rather than manually tweaking the hyper parameters to obtain improvements. This method was compared with other studies employing identical data sets and techniques and was found that the proposed work produces much greater accuracy on detection and classification. The study was also done by comparing the network performance before and after the hyper parameters were tuned and was found that the proposed ABC algorithm along with the modified convolutional network produces better classification outputs on all the classes that were studied and achieves a better rate of success when the hyper parameters of the CNN was optimized.

6. Funding information

There is no particular grant for this research from governmental, private, or nonprofit funding organizations.

7. Data availability statement

As there were no new data generated or examined in this study, data sharing is not applicable to this article.

CRediT authorship contribution statement

Arun A.N.: Conceptualization, Methodology, Writing – original draft. **Maheswaravenkatesh P.:** Supervision. **Jayasankar T.:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Adyapady, R. R., & Annappa, B. (2023). A comprehensive review of facial expression recognition techniques. *Multimedia Systems, 29*(1), 73–103.
Alphonse, A. S., Shankar, K., Jeyasheela Rakkini, M. J., Ananthakrishnan, S., Athisayamani, S., Robert Singh, A., & Gobi, R. (2021). A multi-scale and rotation-invariant phase pattern (MRIPP) and a stack of restricted Boltzmann machine (RBM) with preprocessing for facial expression classification. *Journal of Ambient Intelligence and Humanized Computing, 12*, 3447–3463.
Annadurai, S., Arock, M., & Vadivel, A. (2023). Real and fake emotion detection using enhanced boosted support vector machine algorithm. *Multimedia Tools and Applications, 82*(1), 1333–1353.
Arul Vinayakam Rajasimman, M., Manoharan, R. K., Subramani, N., Aridoss, M., & Galety, M. G. (2022). Robust facial expression recognition using an evolutionary algorithm with a deep learning model. *Applied Sciences, 13*(1), 468.
Ayyalasomayajula, S. C., Ionescu, B., & Ionescu, D. (2021). A cnn approach to micro-expressions detection. In *2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI)* (pp. 345–350). IEEE.
Banerjee, R., De, S., & Dey, S. (2022). Wtaof-ilpb based feature learning and lfssoa-rbfnn based classification for facial micro-expression recognition. *Wireless Personal Communications, 127*(3), 2285–2304.
Du, G., Long, S., & Yuan, H. (2020). Non-contact emotion recognition combining heart rate and facial expression for interactive gaming environments. *IEEE Access, 8*, 11896–11906.
Durga, B. K., & Rajesh, V. (2022). A ResNet deep learning based facial recognition design for future multimedia applications. *Computers and Electrical Engineering, 104*, Article 108384.
Esmaeili, V., Mohassel Feghhi, M., & Shahdi, S. O. (2022). A comprehensive survey on facial micro-expression: Approaches and databases. *Multimedia Tools and Applications, 81*(28), 40089–40134.
Ghosh, R., & Sinha, D. (2022). Human emotion recognition by analyzing facial expressions, heart rate and blogs using deep learning method. *Innovations in Systems and Software Engineering*, 1–9.
Hossain, S., Umer, S., Rout, R. K., & Tanveer, M. (2023). Fine-grained image analysis for facial expression recognition using deep convolutional neural networks with bilinear pooling. *Applied Soft Computing, 109997*.
Jarraya, S. K., Masmoudi, M., & Hammami, M. (2020). Compound emotion recognition of autistic children during meltdown crisis based on deep spatio-temporal analysis of facial geometric features. *IEEE Access, 8*, 69311–69326.
Jebadass, J. R., & Balasubramaniam, P. (2022). Low contrast enhancement technique for color images using interval-valued intuitionistic fuzzy sets with contrast limited adaptive histogram equalization. *Soft Computing, 26*(10), 4949–4960.
Kaya, E., Gorkemli, B., Akay, B., & Karaboga, D. (2022). A review on the studies employing artificial bee colony algorithm to solve combinatorial optimization problems. *Engineering Applications of Artificial Intelligence, 115*, Article 105311.
Khattak, A., Asghar, M. Z., Ali, M., & Batool, U. (2022). An efficient deep learning technique for facial emotion recognition. *Multimedia Tools and Applications*, 1–35.
Lakshmi, D., & Ponnusamy, R. (2021). Facial emotion recognition using modified HOG and LBP features with deep stacked autoencoders. *Microprocessors and Microsystems, 82*, Article 103834.

Li, Y., Huang, X., & Zhao, G. (2020). Joint local and global information learning with single apex frame detection for micro-expression recognition. *IEEE Transactions on Image Processing, 30*, 249–263.

Li, Y., Huang, X., & Zhao, G. (2021). Micro-expression action unit detection with spatial and channel attention. *Neurocomputing, 436*, 221–231.

Liu, Q., Yang, J., & Zhang, K. (2022). An improved empirical wavelet transform and sensitive components selecting method for bearing fault. *Measurement, 187*, Article 110348.

Nguyen, M., & Yan, W. Q. (2022, June). Temporal colour-coded facial-expression recognition using convolutional neural network. In *Science and Technologies for Smart Cities: 7th EAI International Conference, SmartCity360°, Virtual Event, December 2-4, 2021, Proceedings* (pp. 41-54). Cham: Springer International Publishing.

Nisamudeen, H. P., & Zhang, L. (2023). Deep learning-based facial expression recognition. In *Machine Learning, Multi Agent and Cyber Physical Systems: Proceedings of the 15th International FLINS Conference (FLINS 2022)* (pp. 269-276).

Sergeeva, A. D., & Sablina, V. A. (2020). Eye landmarks detection technology for facial micro-expressions analysis. In *2020 9th Mediterranean Conference on Embedded Computing (MECO)* (pp. 1–4). IEEE.

Sikkandar, H., & Thiyagarajan, R. (2021). Deep learning based facial expression recognition using improved Cat Swarm Optimization. *Journal of Ambient Intelligence and Humanized Computing, 12*, 3037–3053.

Sivaraman, E., & Manickachezian, R. (2018). Unevenness measurement using the support vector machine and dynamic multiservice load balancing with modified genetic algorithm in cloud-based multimedia system. *International Journal of Computer Aided Engineering and Technology, 10*(6), 732–747.

Sivaraman, E., & Manickachezian, R. (2019). Intelligent decision making service framework based on analytic hierarchy process in cloud environment. *International Journal of Networking and Virtual Organisations, 21*(2), 221–236.

Takalkar, M. A., Thuseethan, S., Rajasegarar, S., Chaczko, Z., Xu, M., & Yearwood, J. (2021). LGAttNet: Automatic micro-expression detection using dual-stream local and global attentions. *Knowledge-Based Systems, 212*, Article 106566.

Thuseethan, S., Rajasegarar, S., & Yearwood, J. (2020). Complex emotion profiling: An incremental active learning based approach with sparse annotations. *IEEE Access, 8*, 147711–147727.

Xie, H. X., Lo, L., Shuai, H. H., & Cheng, W. H. (2022). An overview of facial micro-expression analysis: Data, methodology and challenge. *IEEE Transactions on Affective Computing*.

Xu, H., Tang, T., Zhang, B., & Liu, Y. (2022). Identification of two-phase flow regime in the energy industry based on modified convolutional neural network. *Progress in Nuclear Energy, 147*, Article 104191.

Yadav, S. P. (2021). Emotion recognition model based on facial expressions. *Multimedia Tools and Applications, 80*(17), 26357–26379.

Zahara, L., Musa, P., Wibowo, E. P., Karim, I., & Musa, S. B. (2020). The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm based Raspberry Pi. In *2020 Fifth international conference on informatics and computing (ICIC)* (pp. 1–9). IEEE.