International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *10 August 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *20 October 2023*

Robust Facial Emotion Recognition using Marine Predators Algorithm with Deep Convolutional Neural Network

[1]K. Fathi Mary, [2]L. R. Sudha, [3]W. Paulina Vasanthi
[1]Research Scholar, Department of Computer and Information Science,
Annamalai University, Chidambaram, India.
E-Mail: srfathima17@gmail.com
[2]Associate Professor, Department of Computer Science and Engineering,
Annamalai University, Chidambaram, India.
[3]Assistant Professor, Idhaya College of Arts and Science for Women,
Puducherry.

Abstract—Facial emotion recognition (FER) is a technology that includes the automatic identification and categorization of human sentiments depending on facial emotions. It leverages deep learning (DL), computer vision (CV), and machine learning (ML) methods to analyze the features of an individual's face, like the place of the mouth, eyes, eyebrows, and complete facial movements for determining their emotional conditions. Popular emotions that FER can recognize comprise surprise, sadness, happiness, fear, disgust, and anger. FER employing DL has an advanced application of artificial intelligence (AI) and deep neural networks (DNNs) that contain training methods to automatically recognize and categorize human expressions based on facial expressions. This approach normally comprises convolutional neural networks (CNNs) or highly complex models namely recurrent neural networks (RNNs) and convolutional RNNs (CRNNs) for analyzing and interpreting complex facial features and dynamics. This study introduces a new Robust Facial Emotion Recognition employing the Marine Predators Algorithm with Deep Learning (RFER-MPADL) approach. The main aim of the RFER-MPADL technique is to detect and categorize different kinds of emotions expressed in facial images. To accomplish this, the RFER-MPADL technique initially applies a bilateral filtering (BF) approach for the preprocessing step. Additionally, the RFER-MPADL technique uses the EfficientNet-B0 method for feature extraction. Moreover, the tuning process of the EfficientNet-B0 method was implemented using the MPA. Finally, the classification of facial emotions can be performed by the use of a deep autoencoder (DAE), in turn augments the overall performance of the RFER-MPADL method. The experimental analysis of the RFER-MPADL methodology is assessed on a standard facial expression dataset. The extensive outcomes exhibited the effectual performance of the RFER-MPADL methodology over other methods.

Keywords- Facial Emotion Recognition; Deep Learning; Marine Predators Algorithm; EfficientNet

I. INTRODUCTION

Facial emotions are vital causes in the communication of humans that aid in recognizing the intents of others. Generally, individuals conclude emotional conditions of other people like anger, sorrow and joy employing facial expressions and voice tone [1]. According to various researchers, vocal components transport one-third and non-verbal components transport two-thirds of human messages [2]. When compared to many nonverbal components, facial emotions are chief information networks in personal communication. So, the study of facial expression has gained high attention over previous years with many applications in cognitive sciences, perceptual, computing and computer animations [3]. DNNs or CNNs are implemented in detecting human facial emotions. The study also presented a network taking the assistance of ResNet50 for classifying human facial expressions by employing static imageries [4]. FER deals with common features of FER recognition. Currently, it is growing with fast improvement of AI models such as virtual reality (VR), advanced driver assistant systems (ADASs), human-computer interaction (HCI), augmented reality (AR) and entertainment [5].

A classic FER technique consists of 3 chief phases such as feature detection, extraction, and classification [6]. Generally, the face identification phase could be supplemented by a face alignment point that main idea is to guarantee that all pictures are even. Many FER frameworks employ the open-source application of Viola-Jones and proposed DL-based facial sensors employ models provided by the OpenCV [7]. In feature extraction, customary models for FER use feature descriptors namely local neighbourhood difference binary pattern (LNDBP), local phase quantization (LPQ), local binary pattern (LBP), and histogram of oriented gradient (HOG) [8]. All these

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 10 August 2023 **Revised:** 28 September 2023 **Accepted:** 20 October 2023*

kinds of extractors are hand-designed so it is not capable of simplifying well in dissimilar imaging situations like subjects' ethnicity, occlusion, lighting and much more [9]. For feature classification, the FER technique based on ML employs general classifiers like naive bayes (NB), support vector machine (SVM), etc. The currently presented DL algorithms for FER are completely trainable and attain greater performance [10].

This study introduces a new Robust Facial Emotion Recognition employing the Marine Predators Algorithm with Deep Learning (RFER-MPADL) approach. The main objective of the RFER-MPADL approach is to detect and categorize different kinds of emotions expressed in facial images. To accomplish this, the RFER-MPADL technique initially applies a bilateral filtering (BF) approach for the preprocessing step. Additionally, the RFER-MPADL technique uses the EfficientNet-B0 method for feature extraction. Moreover, the tuning process of the EfficientNet-B0 method was implemented using the MPA. Finally, the classification of facial emotions can be performed by the use of a deep autoencoder (DAE), in turn augments the overall performance of the RFER-MPADL method. The experimental analysis of the RFER-MPADL methodology is assessed on a standard facial expression dataset.

II. LITERATURE SURVEY

Di Luzio et al. [11] presented new DNNs utilizing parameter randomization in an intricate classification method to identify emotions. Indeed, randomized DNNs signify a remarkable selection for determining the efficiency-to-accuracy balance in real-world utilizations. Punuri et al. [12] introduced a new approach called EfficientNet-XGBoost, which depends on the TL method. EfficientNet-XGBoost was essentially a cascading of the XGBoost and EfficientNet algorithms with specific improvements by analysis that considers an innovation of the effort. To confirm that quickly learned network and to address the vanishing gradient complexity, this architecture integrates FC layers of global average pooling, dense and dropout. EfficientNet could be fine-tuned by changing the upper dense layers as well as cascading the XGBoost method appropriately for FER.

Jain et al. [13] developed an innovative squirrel search optimizer with the DL enabled-FER (SSO-DLFER) method for automatic driving. This technique comprises 2 important methods such as face expression and recognition. The RetinaNet architecture was exploited in the first phase of the face recognition method. To recognize expressions, the presented method implemented the neural architectural search (NASNet) huge feature extraction along with a GRU architecture as an algorithm. To enhance the effectiveness of FER, the SSO-assisted hyperparameter tuning process was executed. Deivendran et al. [14] projected a notion for face and illumination invariant facial emotion credit through the images. The people's faces could be evaluated. Facial expression was employed in the CNNs model for categorizing the obtained images within various emotion classifications. This can be a deep, feed-forward ANN. Efficiently exceeds humanoid involvement as well as exhibits poses another effectiveness.

In [15], an innovative feature descriptor was developed for FER employing an adapted LBP and HOG feature descriptor. Primarily, viola-Jones face recognition is exploited for identifying the face area. Secondarily, this developed adapted HOG and LBP feature descriptor has been employed for removing the features of the recognized mouth, nose and eye areas. The feature extraction of these 3 parts can be combined and decreased its dimensionality through deep-stacked AE (DSA). Lastly, multi-class SVM has been employed for classification and identification. Chirra et al. [16] considered and designed a multi-block DCNN algorithm for FER in human, stylized and virtual features. At the multi-block DCNN, the authors determined 4 blocks with different computational modules for extracting the higher-level features from facial images. In [17], the authors presented an innovative technique to recognize simple facial expressions from textural images employing DL and CNN methods. With this introduced technique textural images of human faces have been produced utilizing the LBP approach and trained the CNN method with the LBP images.

III. THE PROPOSED MODEL

In this article, a new RFER-MPADL methodology is presented. The main aim of the RFER-MPADL technique is to recognize and categorize different kinds of emotions expressed in facial images. The RFER-MPADL methodology has involved four major processes namely, BF-assisted preprocessing, EfficientNetB0-based feature extraction, MPA-related hyperparameter, and DAE-assisted classification. Fig. 1 portrays the complete working-flow of the RFER-MPADL methodology.

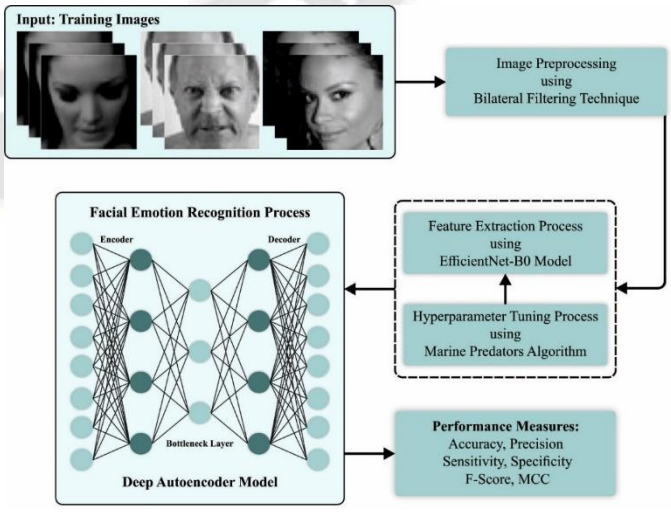

Figure 1. The architecture of the RFER-MPADL method

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *10 August 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *20 October 2023*

A. Preprocessing

Initially, the RFER-MPADL system applies the BF approach for the preprocessing step. BF is an image preprocessing method which smoothens or enhances images while retaining edges and fine details [18]. It works by considering the intensity similarity and spatial proximity between pixels, which applies a weighted average based on the combination of spatial and range domain Gaussian filters. Particularly, this technique is beneficial to reducing stylized rendering, image noise, and tone mapping, where maintaining edges and image features are indispensable. BFs parameter is adjusted for controlling the amount of smoothing and edge preservation, which makes it a valuable tool in CV and image processing applications.

B. Feature extraction

The efficientNet model uses a compound coefficient to uniformly scale the three dimensions of depth (network depth), resolution (input image resolution), and width (network width), thus; the optimal classification impact can be obtained through balancing each dimension [19]. In contrast with classical techniques, this model has a smaller number of parameters and can learn the deep semantic information of images, increasing the accuracy and efficiency of the model. As well, EfficientNet has enhanced transferability and comprises a multi-model mobile inversion bottleneck (MBConv) with the residual model. The architecture includes kxk DepthwiseConv (Swish and BN, the value k is 3or5), squeeze, and excitation (SE) technique, typical 1x1 Conv layer (BN), 1x1 Conv layer (Swish and BN), and dropout layer. This structure takes the network parameter amount while enhancing the ability of feature extraction. Fig. 2 shows the EfficientNetB0 architecture. EfficientNetB0 was a basic architecture for the lightweight network in image classification. EfficientNetB0 includes 9 phases. Stage1 encompasses 3x3 Conv with a stride of 2. Stages 2-8 include repetitive model stacking, and the column layer characterizes the number of times the MB-Conv is repeated. Stage9 includes an FC, 1x1 Conv kernel, and average pooling layer. The number 1 or 6 follows each MBConv in the table. Particularly, the initial Conv layer in the MBConv prolongs the channel of input feature mapping to n times the original. The EfficientNetB1-B7 series of DNN chooses the suitable one in width (the number of channels of the mapping feature), depth (the number of Conv layers), and resolution (the mapping feature's size) depending on width, depth, and resolution of EfficientNetB0.

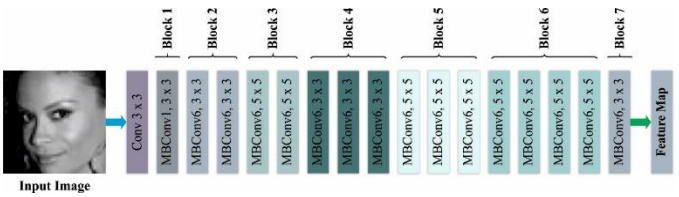

Figure 2. Construction of EfficientNet-B0 approach

C. Parameter tuning using the MPA model

At this phase, the tuning process of the EfficientNet-B0 occurs by employing the MPA. Faramarzi et al. developed the MPA technique that was motivated by hunting schemes of marine predators [20]. Brownian and Lévy signs are measured in MPA. The optimum encounter rate of predators, as well as prey, is occurred.

MPA has a populace-assisted meta-heuristic model. The MPA accumulation procedure begins with a random solution and Eq. (1) employed for activation. In Eq. (1), $X_{\min}$ represents the lesser bound, $X_{\max}$ denotes the greater bound, and $rand$ is a random number.

$$X_0 = X_{\min} + rand(X_{\max} - X_{\min}) \quad (1)$$

There are dual matrices such as $Prey$ and $Elite$ in MPA. These two matrices have the same dimension. The optimal solution is selected as the optimal predator when creating the $Elite$ matrix.

The discovery and hunt of prey are tested through this kind of matrix. $\overrightarrow{X^I}$ symbolizes the upper predator vector, d denotes dimensions and n signifies search agents. Both prey and predator are known as search agents.

$$Elite = \begin{bmatrix} X^I_{1,1} & X^I_{1,2} & \dots & X^I_{1,\mathrm{d}} \\ X^I_{2,2} & X^I_{2,2} & \dots & X^I_{2,\mathrm{d}} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X^I_{n,1} & X^I_{n,2} & \dots & X^I_{n,d} \end{bmatrix}_{nxd} \quad (2)$$

In the Prey matrix, $the\ j$th size of ith prey is shown with $X_{i,j}$. In MPA, the entire accumulation procedure is significant to these dual matrices directly. Predators are utilized for upgrading their locations.

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,\mathrm{d}} \\ X_{2,2} & X_{2,2} & \dots & X_{2,\mathrm{d}} \\ X_{3,1} & X_{3,2} & \dots & X_{3,\mathrm{d}} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,2} & \dots & X_{n,d} \end{bmatrix}_{nxd} \quad (3)$$

These stages are accessible in MPA. The brief details of these stages are described below.

If $Iter < ((\mathrm{Max}\_Iter)/3)$, stage 1 is understood. $Iter$ signifies the present round number and Max_Iter denotes the upper round number. The optimal approach is single where in predator must stay. In calculated method Eq. (4) of stage 1, R_B describes a vector describing Brownian motion as well as an arbitrary number depending on a normal dispersion. P symbolizes an endless number at a range value of 0.5. R denotes an arbitrary number amid 0, 1.

$$\overrightarrow{stepsize_i} = \vec{R}_{\mathrm{B}} \otimes (\overrightarrow{Elite_i} - \vec{R}_{\mathrm{B}} \otimes \overrightarrow{Prey_i}) i = 1, \dots n \quad (4)$$
$$\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P.\vec{R} \otimes \overrightarrow{stepsize_i}$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *10 August 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *20 October 2023*

If $((\text{Max\_}Iter)/3) < Iter < ((2\text{Max\_}Iter)/3$, stage 2 happens. When the prey's dislocation is Lévy, the predator's dislocation needs to be Brownian. The concern is exploitation that is expressed in Eq. (5), and the duty of the predator can also exploration that is represented in Eq. (6). $\vec{R}_L$ is an arbitrary vector number signifying Lévy displacement in Eq. (5). Prey and $\vec{R}_L$ multiplication represent prey move and it is signified by totalling step size to prey place. CF represents an adaptive parameter in Eq. (6). The step size is measured through CF. The $Elite$ and $\vec{R}_{\mathrm{B}}$ multiplication denotes predator move.

$$\overrightarrow{stepsize}_{\mathrm{i}} = \vec{R}_{\mathrm{L}} \otimes \left(\overrightarrow{Elite}_{\mathrm{i}} - \vec{R}_{\mathrm{L}} \otimes \overrightarrow{Prey}_i\right) i = 1, \ldots \frac{n}{2} \quad (5)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Prey}_i + P.\vec{R} \otimes \overrightarrow{stepsize}_i$$

$$\overrightarrow{stepsize}_{\mathrm{i}} = \vec{R}_{\mathrm{B}} \otimes \left(\vec{R}_{\mathrm{B}} \otimes \overrightarrow{Elite}_{\mathrm{i}} - \overrightarrow{Prey}_i\right) i = n/2, \ldots n$$

$$\overrightarrow{Prey}_i = \overrightarrow{Elite}_i + P.\mathrm{CF} \otimes \overrightarrow{stepsize}_i \quad (6)$$

$$CF = \left(1 - \frac{Iter}{\text{Max\_}Iier}\right)^{\left(2\frac{Iter}{\text{Max\_}Iter}\right)}$$

If $> ((2\text{Max\_}Iter)/3)$, stage 3 is understood. As optimal plan, the predator's move is Lévy in Eq. (7), $Elite$ and $\vec{R}_L$ multiplication indicates predator move.

$$\overrightarrow{stepsize}_{\mathrm{i}} = \vec{R}_{\mathrm{L}} \otimes \left(\vec{R}_{\mathrm{L}} \otimes \overrightarrow{Elite}_{\mathrm{i}} - \overrightarrow{Prey}_i\right) i = 1, \ldots n \quad (7)$$

$$\overrightarrow{Prey}_i = \overrightarrow{Elite}_i + P.\mathrm{CF} \otimes \overrightarrow{stepsize}_i$$

The Fish Aggregating Devices (FADs) may upset marine predator's development of activity. It is well known as FAD's impact on MPA. Then it is calculated by employing Eq. (8) with values of 0.2. $\vec{U}$ represents a binary vector with array values 0 and 1. r means arbitrary number range among 0 and 1. $\vec{X}_{\min}$ is a vector that symbolizes lesser bounds of sizes. $\vec{X}_{\max}$ is a vector symbolizing greater bounds of sizes. $r1$ and $r2$ show the prey matrix's arbitrary directories.

$$\overrightarrow{Prey}_{\mathrm{i}} = \begin{cases} \overrightarrow{Prey}_{\mathrm{i}} + CF\left[\vec{X}_{min} + \vec{R} \otimes \left(\vec{X}_{max} - \vec{X}_{min}\right)\right] \otimes \vec{U}, r \le FADs \\ \overrightarrow{Prey}_{\mathrm{i}} + \left[FADs\ (1-r) + r\right](\overrightarrow{Prey}_{\mathrm{r1}} - \overrightarrow{Prey}_{\mathrm{r2}}), r > FADs \end{cases} \quad (8)$$

The fitness selection is the major factor manipulating the accuracy of the MPA technique. The selection process of hyperparameter includes the solution encoding to estimate the candidate solution effectiveness. The MPA model considers accuracy as a crucial criterion for designing the fitness function.

$$Fitness = \max(P) \quad (9)$$

$$P = \frac{TP}{TP + FP} \quad (10)$$

Here, TP and FP portrays the true and false positive values.

D. DAE-based classification model

Finally, the classification of facial emotions can be performed by the use of the DAE model. The AE has a common form of ANN that is normally utilized in unsupervised ML [21]. It contains an input layer of d neurons that signify the count of the feature and is encrypted by sending it to a hidden layer (HL) l_1 in the dimension of $d/2$ neurons. An additional layer with $l_1/2$ size neurons is supplemented that also indicates the encoded data symbol of unique input. Then, AE is learned to rebuild inputs from bottleneck by consuming hidden decoding layers. Every layer is monitored by dropout HLs (not presented in the image) to preclude information's overfitting. The encoder plans input vector x_i to a hidden illustration unit that indicates the bottleneck layer's latent space which can be given in Eq. (11).

$$l_i = f(x_i) = g(Wx_i + b_i) \quad (11)$$

Whereas W represents the weight matrix, g is the activation function and b is the bias vector. The recognized activation function is hyperbolic tangent ($tanh$) and rectified linear unit (ReLU) that is utilized in the technique. The research has found that merging these functions provides improved outcomes. ReLU in Eq. (12) is a non-sequential activation role and is appropriate in MPL. It comes with numerous HLs due to its fastness. In addition, it aids in decreasing error gradient problems as well as state disappearing. The tanh in Eq. (13) is similar to the sigmoid function and it varies in the output limit $(-1,1)$. The value of sigmoid is $(0,1)$.

$$g(x) = ReLU(x) = \max(0, x) \quad (12)$$

$$g(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (13)$$

where χ is input to the function.

The decoder record l_n rebuilds x' of similar input size x:

$$lx' = g'(W'l_i + b_i') \quad (14)$$

By employing error backpropagation in the learning procedure, the autoencoder attempts to reduce normal rebuilding error amid input and recreated output in Eq. (15)

$$RE(x, x') = \frac{1}{m}\left\|\sum_{1}^{m}\left(x_i - x_{\mathrm{i}}^i\right)\right\|^2 \quad (15)$$

Whereas m is the sample number in above mentioned Eq. (15).

IV. Results and Discussion

In this section, the FER assessment of the BTS-AERODL method is examined with the Kaggle dataset [22]. It comprises 32298 instances with 7 classes as defined in Table 1. Fig. 3 illustrates the sample images.

TABLE I. Dataset Description

Class	Instance Numbers
Anger	4486
Contempt	491
Disgust	4625
Fear	8094
Happy	5591
Sadness	5424
Surprise	3587

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *10 August 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *20 October 2023*

Total Instances	**32298**

Figure 3. Sample images

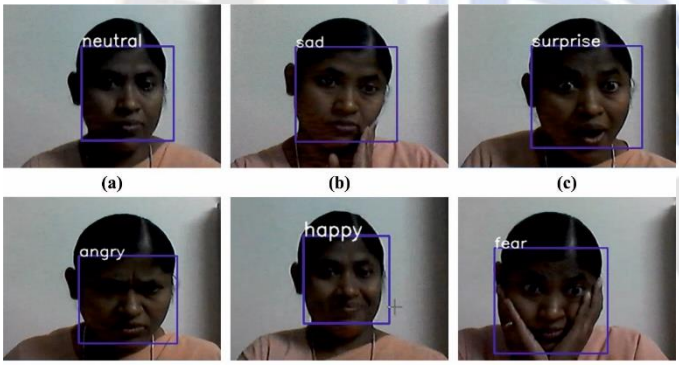

Figure 4. Test on Real-Time Images a) Emotion-Neutral b) Emotion-Sad c) Emotion-Surprise d) Emotion-Angry e) Emotion-Happy f) Emotion-Fear

Fig. 4 represents a visualization result of the BTS-AERODL method on the real-time images. The figure illustrated that the BTS-AERODL model effectually recognizes facial emotions.

Fig. 5 exhibits the classifier analysis of the BTS-AERODL approach in the test database. Figs. 5a-5b shows the confusion matrices given for the BTS-AERODL approach at 70:30 of the TR/TS set. This figure demonstrated that the BTS-AERODL approach can be appropriately categorized and identified with seven classes. Likewise, Fig. 5c pointed out the PR-curve of the BTS-AERODL technique. The figure denoted that the BTS-AERODL methodology attains considerable PR performance with every class. Lastly, Fig. 5d reveals the ROC analysis of the BTS-AERODL methodology. This figure displays that the BTS-AERODL methodology accelerates effective outcomes with better ROC values in diverse classes.

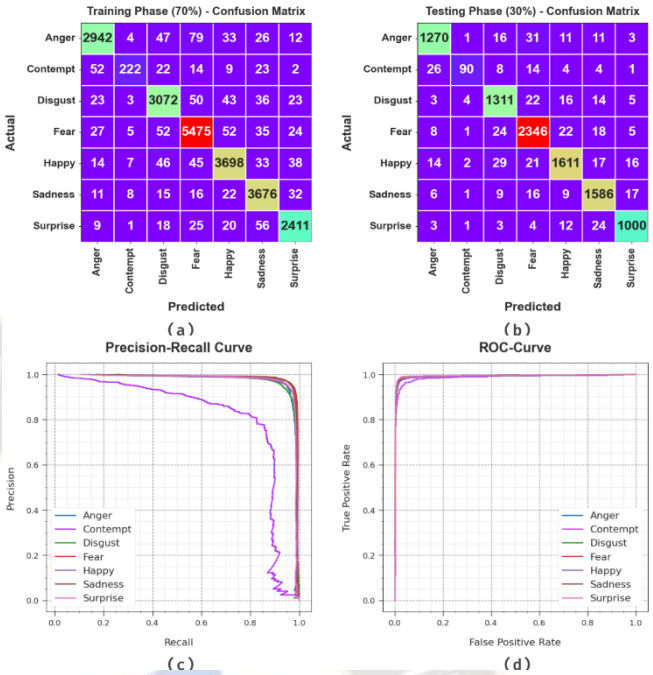

Figure 5. Classifier analysis of (a-b) Confusion matrices with 70:30 of TR/TS set, (c-d) Curves of PR and ROC

TABLE II. FER ANALYSIS OF BTS-AERODL SYSTEM ON 70:30 OF TR/TS SET

Class	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	F_{score}	MCC
TR Phase (70%)						
Anger	98.51	95.58	93.60	99.30	94.58	93.73
Contempt	99.34	88.80	64.53	99.87	74.75	75.39
Disgust	98.33	93.89	94.52	98.97	94.20	93.23
Fear	98.12	95.99	96.56	98.65	96.27	95.02
Happy	98.40	95.38	95.28	99.04	95.33	94.37
Sadness	98.62	94.62	97.25	98.89	95.92	95.10
Surprise	98.85	94.85	94.92	99.35	94.88	94.24
Average	98.59	94.16	90.95	99.15	92.28	91.58
TS Phase (30%)						
Anger	98.63	95.49	94.56	99.28	95.02	94.23
Contempt	99.31	90.00	61.22	99.90	72.87	73.92
Disgust	98.42	93.64	95.35	98.93	94.49	93.57
Fear	98.08	95.60	96.78	98.51	96.19	94.91
Happy	98.21	95.61	94.21	99.07	94.90	93.83
Sadness	98.49	94.74	96.47	98.91	95.60	94.70
Surprise	99.03	95.51	95.51	99.46	95.51	94.97
Average	98.60	94.37	90.59	99.15	92.08	91.45

The FER analysis of the BTS-AERODL approach with 70:30 of TR/TS set is examined in Table 2 and Fig. 6. The obtained output shows that the BTS-AERODL approach obtains effectual performance with seven classes. On 70% of the TR Phase, the BTS-AERODL technique attains an average $accu_y$ of 98.59%, $prec_n$ of 94.16%, $sens_y$ of 90.95%, $spec_y$ of 99.15%, F_{score} of 92.28%, and *MCC* of 91.58%. On the other hand, with 30% of TS Phase, the BTS-AERODL technique achieves an average $accu_y$ of 98.60%, $prec_n$ of 94.37%, $sens_y$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *10 August 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *20 October 2023*

of 90.59%, $spec_y$ of 99.15%, F_{score} of 92.08%, and *MCC* of 91.45% respectively.

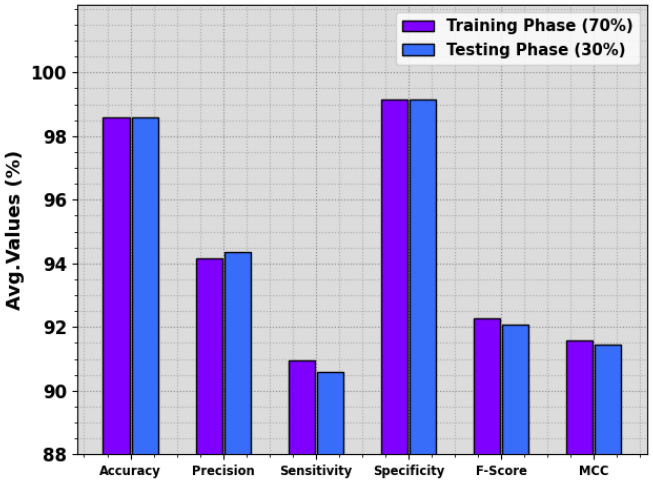

Figure 6. Average of BTS-AERODL model at 70:30 of TR/TS set

To evaluate the effectiveness of the BTS-AERODL system, $accu_y$ curves are generated for the testing (TS) and training (TR) phases, as illustrated in Fig. 7. Two curves provide valued insights into the technique's capability and learning evolution in generalization. While increasing the number of epochs, an observable escalation in this TR and TS $accu_y$ curves has been obvious. This enrichment exhibits the model's potential to greatly identify patterns in terms of the databases of TR and TS.

Figure 7. $Accu_y$ curve of BTS-AERODL approach

Fig. 8 exhibits an analysis of the BTS-AERODL approach, and loss values through the training manner. The minimizing trend at TR loss over epochs pointed out that the model often enhances their weights for reducing prediction errors under both TR and TS databases. The loss analysis deliberates how the model could be fitted to the TR database. Mainly, the TS and TR loss are constantly lower, presenting the efficacious learning patterns of the model being in two databases. Likewise, this can display the variation of the model in reducing variances among predicted and original training labels.

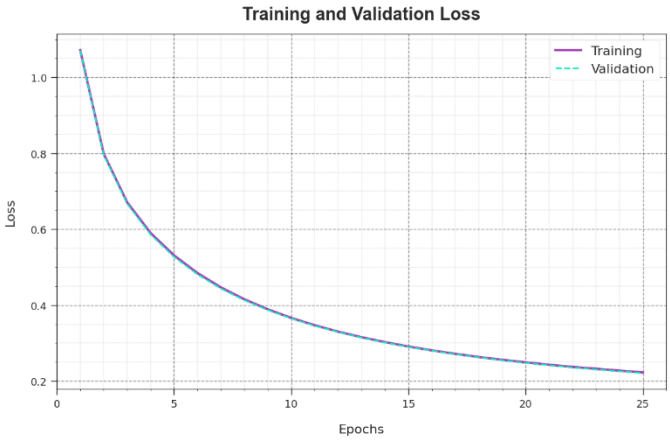

Figure 8. Loss curve of BTS-AERODL approach

TABLE III. $Accu_y$ OUTCOME OF BTS-AERODL MODEL COMPARED WITH OTHER EXISTING SYSTEMS

Methods	$Accu_y$ (%)
Unsupervised DA	66.82
BoW Model	69.15
VGG+SVM	67.96
GN Model	66.76
FER-SoC	67.60
FER-ACN	71.69
Aff-Wild2	76.71
IFER-ODTFL	94.23
RFER-MPADL	98.60

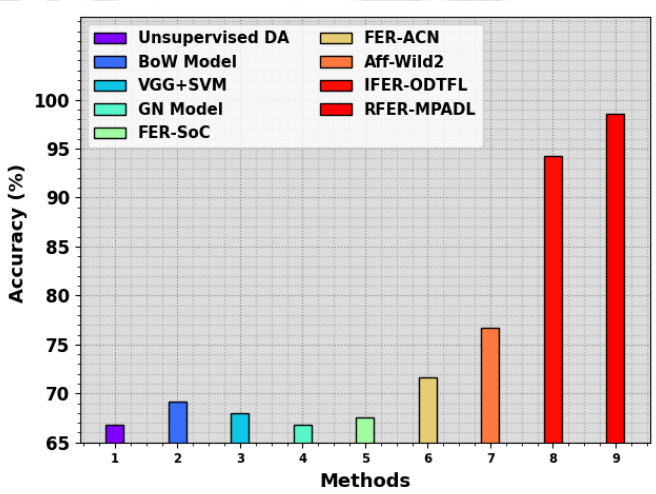

Figure 9. $Accu_y$ analysis of the BTS-AERODL approach with other present methods

Table 3 and Fig. 9 highlight a comprehensive FER assessment of the BTS-AERODL approach with existing models [23]. The outputs indicate that the Unsupervised DA, BoW, VGG+SVM, GN, and FER-SoC approaches obtain reduced

performance with lower $accu_y$ values of 66.82%, 69.15%, 67.96%, 66.76%, and 67.60% respectively. Concurrently, the FER-CAN and Aff-Wild2 methodologies accomplish moderately improved outcomes with $accu_y$ of 71.69% and 76.71% respectively. Although the IFER-ODTFL approach achieves a near-optimal $accu_y$ of 94.23%, the BTS-AERODL approach obtains superior performance over other existing systems with a greater $accu_y$ of 98.60%. Thus, the BTS-AERODL system can be employed for an accurate FER detection process.

V. Conclusion

In this study, a new RFER-MPADL approach is presented. The main objective of the RFER-MPADL algorithm is to recognize and categorize different kinds of emotions expressed in facial images. To accomplish this, the RFER-MPADL technique initially applies the BF approach for the preprocessing step. In addition, the RFER-MPADL technique uses the EfficientNetB0 model for feature extraction. Moreover, the hyperparameter tuning of the EfficientNetB0 takes place using the MPA. Finally, the classification of facial emotions can be performed by the use of DAE, in turn augments the complete effectiveness of the RFER-MPADL technique. The investigational assessment of the RFER-MPADL technique is examined with a standard facial expression dataset. The investigational outputs exhibited the effectual performance of the RFER-MPADL approach with other models.

References

[1] Gupta, S., Kumar, P. and Tekchandani, R.K., 2023. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications*, *82*(8), pp.11365-11394.

[2] Khattak, A., Asghar, M.Z., Ali, M. and Batool, U., 2022. An efficient deep learning technique for facial emotion recognition. *Multimedia Tools and Applications*, pp.1-35.

[3] Pan, J., Fang, W., Zhang, Z., Chen, B., Zhang, Z. and Wang, S., 2023. Multimodal emotion recognition based on facial expressions, speech, and EEG. *IEEE Open Journal of Engineering in Medicine and Biology*.

[4] Gaddam, D.K.R., Ansari, M.D., Vuppala, S., Gunjan, V.K. and Sati, M.M., 2022. Human facial emotion detection using deep learning. In *ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications* (pp. 1417-1427). Springer Singapore.

[5] Yildirim, E., Akbulut, F.P. and Catal, C., 2023. Analysis of facial emotion expression in eating occasions using deep learning. *Multimedia Tools and Applications*, pp.1-13.

[6] Arora, T.K., Chaubey, P.K., Raman, M.S., Kumar, B., Nagesh, Y., Anjani, P.K., Ahmed, H.M., Hashmi, A., Balamuralitharan, S. and Debtera, B., 2022. Optimal facial feature based emotional recognition using deep learning algorithm. *Computational Intelligence and Neuroscience*, *2022*.

[7] Bashir, M.F., Javed, A.R., Arshad, M.U., Gadekallu, T.R., Shahzad, W. and Beg, M.O., 2023. Context-aware Emotion Detection from Low-resource Urdu Language Using Deep Neural Network. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(5), pp.1-30.

[8] Minaee, S., Minaei, M. and Abdolrashidi, A., 2021. Deep-emotion: Facial expression recognition using attentional convolutional network. *Sensors*, *21*(9), p.3046.

[9] Jayanthi, K. and Mohan, S., 2022. An integrated framework for emotion recognition using speech and static images with deep classifier fusion approach. *International Journal of Information Technology*, *14*(7), pp.3401-3411.

[10] Poulose, A., Reddy, C.S., Kim, J.H. and Han, D.S., 2021, August. Foreground Extraction Based Facial Emotion Recognition Using Deep Learning Xception Model. In *2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 356-360). IEEE.

[11] Di Luzio, F., Rosato, A. and Panella, M., 2023. A randomized deep neural network for emotion recognition with landmarks detection. *Biomedical Signal Processing and Control*, *81*, p.104418.

[12] Punuri, S.B., Kuanar, S.K., Kolhar, M., Mishra, T.K., Alameen, A., Mohapatra, H. and Mishra, S.R., 2023. Efficient net-XGBoost: an implementation for facial emotion recognition using transfer learning. *Mathematics*, *11*(3), p.776.

[13] Jain, D.K., Dutta, A.K., Verdú, E., Alsubai, S. and Sait, A.R.W., 2023. An automated hyperparameter tuned deep learning model enabled facial emotion recognition for autonomous vehicle drivers. *Image and Vision Computing*, *133*, p.104659.

[14] Deivendran, P., Babu, P.S., Malathi, G., Anbazhagan, K. and Kumar, R.S., 2023. Emotion Recognition for Challenged People Facial Appearance in Social using Neural Network. *arXiv preprint arXiv:2305.06842*.

[15] Lakshmi, D. and Ponnusamy, R., 2021. Facial emotion recognition using modified HOG and LBP features with deep stacked autoencoders. *Microprocessors and Microsystems*, *82*, p.103834.

[16] Chirra, V.R.R., Uyyala, S.R. and Kolli, V.K.K., 2021. Virtual facial expression recognition using deep CNN with ensemble learning. *Journal of Ambient Intelligence and Humanized Computing*, pp.1-19.

[17] Mukhopadhyay, M., Dey, A., Shaw, R.N. and Ghosh, A., 2021, September. Facial emotion recognition based on textural pattern and convolutional neural network. In *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 1-6). IEEE.

[18] Asokan, A. and Anitha, J., 2020. Adaptive Cuckoo Search based optimal bilateral filtering for denoising of satellite images. *ISA transactions*, *100*, pp.308-321.

[19] Routray, N., Rout, S.K., Sahu, B., Panda, S.K. and Godavarthi, D., 2023. Ensemble Learning with Symbiotic Organism Search Optimization Algorithm for Breast Cancer

Classification & Risk Identification of Other Organs on Histopathological Images. *IEEE Access*.

[20] Baştemur Kaya, C., 2023. A Novel Hybrid Method Based on the Marine Predators Algorithm and Adaptive Neuro-Fuzzy Inference System for the Identification of Nonlinear Systems. *Symmetry*, *15*(9), p.1765.

[21] Al Sawafi, Y., Touzene, A. and Hedjam, R., 2023. Hybrid Deep Learning-Based Intrusion Detection System for RPL IoT Networks. *Journal of Sensor and Actuator Networks*, *12*(2), p.21.

[22] https://www.kaggle.com/datasets/msambare/fer2013

[23] Albraikan, A.A., Alzahrani, J.S., Alshahrani, R., Yafoz, A., Alsini, R., Hilal, A.M., Alkhayyat, A. and Gupta, D., 2022. Intelligent facial expression recognition and classification using optimal deep transfer learning model. *Image and Vision Computing*, *128*, p.104583.