

RESEARCH ARTICLE

Oppositional Brain Storm Optimization With Deep Learning Based Facial Emotion Recognition for Autonomous Intelligent Systems

T. PRABHAKARA RAO¹, SATISHKUMAR PATNALA², CH. V. RAGHAVENDRAN³,
E. LAXMI LYDIA⁴, YEONWOO LEE⁵, SRIJANA ACHARYA⁶,
AND JAE-YONG HWANG⁷

¹Department of Computer Science and Engineering, Aditya Engineering College (A), Surampalem 533437, India

²Department of IT, ANITS Engineering College, Visakhapatnam 531163, India

³Department of IT, Aditya College of Engineering and Technology, Surampalem, Andhra Pradesh 533437, India

⁴Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh 532127, India

⁵Department of Information Communication Engineering, Mokpo National University, Muan-gun, Chonnam 58554, South Korea

⁶Department of Convergence Science, Kongju National University, Gongju-si 32588, South Korea

⁷Department of Information and Communication Engineering, Daejeon University, Dong-gu, Daejeon 34520, South Korea


Corresponding authors: Srijana Acharya (srijana@kongju.ac.kr) and Jae-Yong Hwang (platans@dju.kr)

ABSTRACT Autonomous Intelligent Systems (AIS) states to a class of intelligent devices that manage and create decisions independently without human interference. These techniques use numerous models contains artificial intelligence (AI), robotics, machine learning (ML), and sensor fusion, in order to identify its environment, reflect that data, and execute accordingly to achieve specific goals. Facial emotion detection in AIS applies to the capability of AI-driven autonomous machines to detect and interpret human emotions reliant on facial expressions. This technology permits AIS to identify and reply to the emotional conditions of individuals they interconnect with, foremost to more normal and empathetic human-machine communications. So, this study develops an Oppositional Brain Storm Optimizer with Deep Learning based Facial Emotion Recognition (OBSODL-FER) system for AIS. The foremost goal of the OBSODL-FER system is to identify and organize dissimilar classes of facial emotions of the drives in autonomous vehicles. To achieve, the OBSODL-FER approach mainly employs an Xception-based deep convolutional neural networks (CNNs) for feature extractor. Also, the developed OBSODL-FER approach exploits the OBSO system for the hyperparameter selection of the Xception method. Besides, an improved LSTM model (ILSTM) is applied to the classification procedure. Furthermore, a jellyfish search (JFS) optimizer is employed for the optimum hyperparameter selection of the ILSTM technique. The simulation results of the OBSODL-FER approach are verified on a benchmark facial emotion dataset. The experimental results inferred the enhancement of the OBSODL-FER system over other DL algorithms.

INDEX TERMS Autonomous vehicles, intelligent systems, facial emotion recognition, metaheuristics, artificial intelligence.

I. INTRODUCTION

The important robotics research domain is to make intelligent robots proficient in interacting with humans as a companion instead of a machine. With fast development in robotic

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung .

technology, artificial intelligence (AI), and hardware performance, increasingly more collaborative and social robots are designed [1]. Emotions are the level of sentiments such as expressions, thoughts, and psychological changes, which are directed by environments. Emotions positively impact intelligent functions like perceiving, cognition, decision-making, empathic understanding, coping, reasoning, and planning [2].

It can be identified utilizing facial expressions, biological signals, text, gestures, and speech. Facial expressions are unbelievably centrality for understanding human emotions or feelings in human-to-human interaction [3]. Research workers are considered AI as the emergence of the computing revolution that is intellectually and physically identical to human beings. A machine is specifically recognized a driver's facial expression as one method for decreasing the count of dangerous vehicle accidents [4]. The author stated that angry driving actions raise the possibility of car accidents and can be life-threatening to other individuals. A facial expression recognition (FER) technique may support avoiding fatal accidents and rescuing somebody's life from angered drivers. Currently, several researcher workers of FER are severely enhancing new FER methods to achieve the concept of FER method performance [5].

The first technique employed for facial emotion intensity evaluation depends on the distance encouraged [6]. This method utilizes higher-dimensional rates of transformation and regional volumetric distinction maps for classifying and measuring facial expressions. In videos, most techniques utilize Principal Component Analysis (PCA) for characterizing features of facial expressions [7]. PCA is employed to identify the action element for establishing and expressing various facial expressions. Alternative facial expressions have been identified and structured by misuse of PCA to offer a facial action unit [8]. Recently, Deep learning (DL) approaches are revolutionized the computer vision (CV) domain with RNN and CNN. These DL-based techniques can be employed for feature extraction, recognition, and classification tasks [9]. The major benefits of a DL method (CNN) are to address the dependence on physics-based algorithms and decrease the work needed in pre-processing and feature extraction stages. Also, the DL approach is directly for allowing end-to-end learning from the input image [10].

This study presents an Oppositional Brain Storm Optimizer with Deep Learning based Facial Emotion Recognition (OBSODL-FER) method for AIS. The main intention of the OBSODL-FER system is to distinguish and categorize dissimilar classes of facial emotions of the drives in autonomous vehicles. To do so, the OBSODL-FER technique primarily uses an Xception-based deep convolutional neural networks (CNNs) for feature extractor. In addition, the presented OBSODL-FER technique exploits the OBSO system for the hyperparameter selection of the Xception method. Furthermore, an improved LSTM model (ILSTM) is applied to the classification process. Furthermore, the jellyfish search (JFS) optimizer is utilized for the finest hyperparameter selection of the ILSTM method. The simulation results of the OBSODL-FER approach are verified on a benchmark facial emotion database.

II. RELATED WORKS

Jain et al. [11] developed a new Squirrel Search Optimizer with DL Enabled FER (SSO-DL-FER) method for Autonomous Vehicle Driver. This method can be divided into

2 algorithms. Primarily, the RetinaNet method is used at the primary step of the face identification method. Next, the study used the NASNet massive feature extractor with the GRU algorithm. In [12], proposed a method of developing a "proof of concept" technique was defined for recognizing expression in horses. This method is designed with two components one is the model and another one is the detector. The model has been a CNN that forecasts the feelings of those horses. The detector is a rapid region-based CNN to identify horses in an image. When two elements are validated, they have been integrated into an analytical system, which can recognize horse feelings. Chauhan et al. [13] employed a CNN to make an image classifier that identifies objects and categorizes waste products. Eventually, an effective technique and waste disposal method is delivered that might be used to enhance cost-efficiency and performance.

Ge et al. [14] introduced an occluded expression detection approach that depends on the created countermeasure network. This technique can be classified into face image recognition and restoration. Initially, this study reviews the investigation status of the deep-FER approach. Next, the recent FER technique depends on DL and can be categorized into two namely Static FER and dynamic FER. Shahid and Yan [15] presented a SqueezeExpNet framework that can acquire the benefit of global and local facial data for an extremely accurate FER technique. Primarily, a weighted mask of 3D face landmarks was made and element-wise multiplication was employed. Next, the face spatial image was provided as input and its augmentations. Lastly, an RNN is developed for collaborating the identified data from dual phases instead of easily utilizing the SoftMax function. The authors in [16], presented an Assistive Robotics named Lightweight emotion recognition (LEMON) that applies CV, DL method, and image processing to identify facial expressions. Besides, Dilated Convolution increases receptive fields rapidly with maintaining resolution, minimum memory cost and computation. At last, Exponential Linear Unit (ELU) was employed for the primary layers.

Alenazy and Alqahtani [17] concentrated on a semi-supervised DBN algorithm for predicting the FER. The GSA was used for improving any parameters in the DBN model. The 2D-Discrete Wavelet Transform (2D-DWT) and HOG both can be utilized for extracting features from the furrow patches, cheek, brow, lip, and eye. Ni et al. [18] introduced a cross-modality attention-based CNN (CM-CNN) for FER. The authors developed a new cross-modality attention-integrated network for improving spatial relationships. Lastly, the CM-CNN was enhanced with an increased focal loss that become more interested in facial expressions by poor emotion intensities.

III. THE PROPOSED MODEL

The study mainly concentrates on the growth of the OBSODL-FER method for AIS. The purpose of the OBSODL-FER method is to identify and classify dissimilar types of facial emotions of the drives in autonomous vehicles.

The proposed model incorporates 4 phases of processes as Xception model, OBSO-based parameter tuning, ILSTM classification, and JFS-based hyperparameter tuning. Fig. 1 shows the working flow of the OBSODL-FER technique.

A. FEATURE EXTRACTOR

Initially, the Xception architecture is utilized for feature extraction purposes. Google researchers developed this XceptionNet architecture [19]. It has three phases namely the entry, middle, and exit flow of XceptionNet architecture, which is an addition to the Inception framework. As in the Inception framework, feature space is compressed by multiplying 1×1 depthwise convolution across dissimilar channels. Other parallel convolutions are concatenated, to provide a deep and wide Inception framework. Fig. 2 displays the infrastructure of Xception. In the XceptionNet model, depthwise separable convolution is followed by the pointwise convolution layer. A point-wise convolution is given in green and a depthwise (channelwise) separable convolution is given in blue. In depthwise or channel-wise convolutions, if 64 channels are available, and a 3×3 filter size is utilized, $64 \times 3 \times 3$ spatial convolutions will be implemented. On the other hand, pointwise convolution exploits a 1×1 size filter exploited across the depth. In this work, images first enter the entry layer, followed by middle flow, and exit from the exit layer.

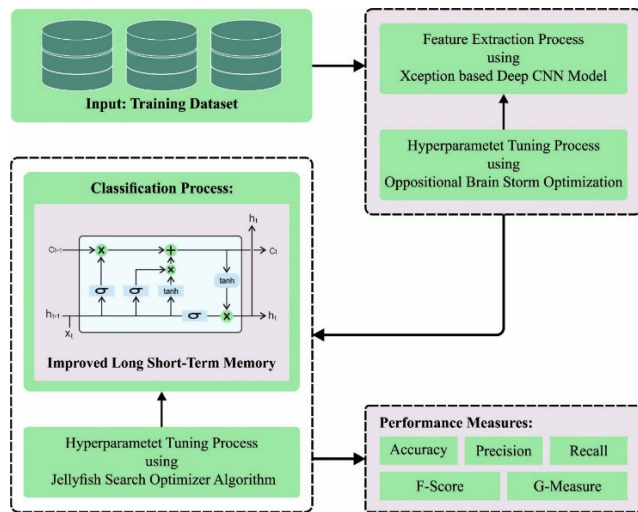


FIGURE 1. Workflow of OBSODL-FER approach.

B. HYPERPARAMETER OPTIMIZATION

In this work, the OBSO algorithm is applied for the parameter range of the Xception model. BSO is a new optimization algorithm which incorporates the device of beetle foraging along with the Particle Swarm optimizer (PSO) method based on the biological nature, a metaheuristic optimization technique depends on the forage character of beetles is derived [20]. Two antennae of beetles were used for exploring the neighbouring region. Once the antenna detects strong food concentration, then beetles move to the antennae. The

beetle's location from S - dimension space at t as x^t , along with the beetle location at $t + 1$ is given as follows:

$$x^{t+1} = x^t + \delta^t * b * \text{sign}(f(x_{rt}) - f(x_{lt})) \quad (1)$$

$$x_{rt} = x^t + d^t * \vec{b}; x_{lt} = x^t - d^t * \vec{b} \quad (2)$$

$$\vec{b} = \frac{\text{rands}(s, 1)}{\|\text{rands}(s, 1)\|}; \delta^t = 0.95\delta^{t-1};$$

$$d^t = 0.95d^{t-1} + 0.01 \quad (3)$$

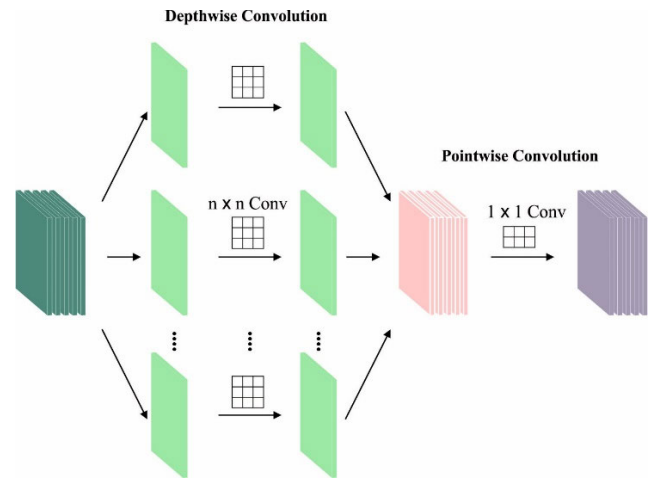


FIGURE 2. Structure of Xception.

Here, \vec{b} refers to a random exploring direction of beetles at S -dimension space, and rand shows a random integer. δ^t denotes the search length of the beetle, and d^t signifies the noticeable distance of the antennae. The δ^t and d^t values are set to maximum and gradually diminished. The primary goal is to develop historical search limits from the optimization model to take over huge areas and moved to the local extreme. x_{lt} and x_{rt} are the predicted location of the right and left antennae of beetles and higher concentrations of food flavor at 2 positions are represented as $f(x_{rt})$ and $f(x_{lt})$.

Like the PSO model, each beetle presented in BSO is a possible solution to optimization problems, and different beetles distribute information amongst themselves. The procedure of enhancing the beetle speed was evaluated by the tendency of BSO to search for global extrema and look for single extrema. Different from this, the position updating of beetles could not be calculated by extending the velocity and data attained by respective antennae. Therefore, it exploits the loosening method for smart computation mechanisms. $X = (X_1, X_2, \dots, X_n)$ is used to characterize the beetle swarms together with the population size of n from the S -dimension searching range, and $X_i = (x_{i1}, x_{i2}, \dots, x_{iS})^T$ denotes the S -dimension vector and position of beetles i in S -dimension search space and a possible solution to the optimization problem. $V_i = (v_{i1}, v_{i2}, \dots, v_{iS})^T$ specifies the speed parameter of beetle i . Single extrema was signified as $P_i = (p_{i1}, p_{i2}, \dots, p_{iS})^T$, along with global extrema is represented as $P_g = (p_{g1}, p_{g2}, \dots, p_{gS})^T$. Thus, the updates of speed

and location of BSO are shown below:

$$x_{is}^{k+1} = x_{is}^k + \lambda v_{is}^k + (1 - \lambda) \xi_{is}^k, \quad (4)$$

$$v_{is}^{k+1} = \omega v_{is}^k + c_1 r_1 (p_{is}^k - x_{is}^k) + c_2 r_2 (p_{gs}^k - x_{gs}^k), \quad (5)$$

$$\xi_{is}^{k+1} = \delta^k * v_{is}^k * \text{sign} \left(f(x_{rs}^k) - f(x_{ls}^k) \right), \quad (6)$$

$$x_{rs}^{k+1} = x_{rs}^k + v_{is}^k * \frac{\bar{d}}{2}; x_{ls}^{k+1} = x_{ls}^k - v_{is}^k * \frac{\bar{d}}{2}, \quad (7)$$

where $s = 1, 2, \dots, S; i = 1, 2, \dots, n; k$ represent the processing duration. ξ_{is} denotes the single portion of displacement. Accordingly, inertia weight (ω) and loosening factor (λ) were adjusted. r_1 and r_2 denote the random function within [1, 0]. Attributes c_1 and c_2 evaluate the effects of single and global extrema of beetles. The semantics of δ^k, d, x and $f(x)$ are comparable to the fundamentals of BAS.

C. OPTIMAL ILSTM-BASED RECOGNITION

For the recognition process, the ILSTM model is applied in this study. In standard LSTM, the resultant node of the standard method is equivalent to the input feature counts, eventually, only the final nodes are obtained and the remaining nodes are lost [21]. In the first case, the removal of nodes may lose data, ensuing in worse training outcomes; conversely, discarded nodes participated in the computing method that uses computational resources and reduces the training speed. It is realized that the quantity of data kept is inversely proportional to the input feature counts as exposed in Eq. (8). It is recognized that the additional input feature counts, the worst the training outcome of the standard method, and the slower the training speed.

$$\text{information_reserve} = \frac{1}{\text{se_length}} * \text{information} \quad (8)$$

$$X = [x_1 x_2 \dots x_{\text{seq_length}}] \quad (9)$$

Afterwards, the outcome of the enhanced LSTM network method is expressed in Eq. (10).

$$\text{lstm_out} = [h] \quad (10)$$

The fundamental phases to enhance the model are written as Algorithm 1. In the first case, an enhanced method could not discard nodes and whole data is achieved which makes sure that the enhanced method is an optimum outcome than the standard LSTM method. An enhanced method has a single LSTM cell, hence it filters the interference data and stores the essential data for the classification outcome of the model and is continuing to improve.

Finally, the JFS algorithm selects the parameters related to the ILSTM model. Jellyfish motion was categorized into three types [22]. First, the movement of jellyfish towards the direction of the ocean current. Next, jellyfish might be inside its swarm. Then, the switching between both possibilities can be managed by the time control (TC) mechanism. At last, the movement of jellyfish towards the location of abundant food sources where the FF characterizes the quantity of

Algorithm 1 Pseudocode of ILSTM technique

```

Input: x
Output: out
(batchsize, 1, seqlength * embedding) ← x.shape;
LSTM (seq_length * embedding, hidden_size, 2)
← LSTM
((2, batchSize, hidden_size),
(2, batchSize, hidden_size)) ← hidden
lstmout, bidden ← LSTM(x, hidden);
Linear (← hidden_size, outputsize) ← FC;
fc_out ← lstmout * FC;
out ← fcout;
End;
Return outcome;

```

food as follows:

$$JF_i(It + 1) = 4Y_0(1 - JF_i) s, 0 \leq Y_0 \leq 1, j = 1, 2, \dots, N_{JF} \quad (11)$$

In Eq. (11), $JF_i(t + 1)$ represents the chaotic counterpart of (i) jellyfishes; It refers to the iteration value; and Y_0 refers a random integer within [0 and 1]; N_{JF} shows the amount of jellyfishes. The TC value at every iteration (It) is differed from 0 to 1 over time which can be calculated by Eq. (12):

$$TC(It) = \left| \left(1 - \frac{It}{It^{\max}} \right) \times (-1 + 2 \times z_1) \right| \quad (12)$$

where It^{\max} denotes the maximal iteration count; z_1 shows the random integer that is generated by uniformly distributed random integer [0, 1]. If the $TC(It)$ value is greater than 0.5, each position of jellyfish was expressed in the following:

$$JF_i(It + 1) = Z \times (JF_{Best} - 3 \times Z \times AV) + JF_{is}(It) \quad (13)$$

In Eq. (13), Z denotes the random number that is generated by the uniform distribution within [1, 0]; JF_{Best} shows the total best location of jellyfish with the maximum value of food source; AV indicates the average location of the jellyfish.

If the $TC(It)$ values are lesser than 0.5, the new position of the jellyfish was established according to the kind of motion inside the swarm. The position updating with the passive type of movement can be expressed as follows:

$$JF_i(It + 1) = 0.1 \times Z \times (U_b - L_b) + JF_i(It) \quad (14)$$

In Eq. (14), U_b and L_b denote the high and small boundaries of the control variables:

$$JF_i(It + 1) = \begin{cases} JF_i(It) + R(JF_j(It) - JF_i(It)) & \text{iff } (JF_i) \geq f(JF_j) \\ JF_i(It) + R(JF_i(It) - JF_j(It)) & \text{iff } (JF_i) < f(JF_j) \end{cases} \quad (15)$$

In Eq. (15), f denotes the amount of food source that shows the FF of optimization problems.

As control variables in the vector of JF_i position in every iteration (It) is checked and set at the adjacent limits:

$$JF(It)_{i,d} = U_{b,d} \text{ or } L_{b,d} \text{ if } JF(It)_{i,d} > U_{b,d} \text{ or } JF(It)_{i,d} < L_{b,d}, \text{ respectively} \tag{16}$$

In Eq. (16), d denotes the control variables inside the vector of jellyfish position.

The fitness function (FF) is a significant feature in the JFS approach. An encode result has been used to evaluate the goodness of candidate performances. Currently, the accuracy standards are the main state utilized in order to plan a FF.

$$Fitness = \max(P) \tag{17}$$

$$P = \frac{TP}{TP + FP} \tag{18}$$

Here, FP and TP represents the false and true positive value.

IV. EXPERIMENTAL VALIDATION

The experimental results of the OBSODL_FER model are verified on dual databases like KDEF and KMU_FED datasets. The KDEF dataset holds 4900 samples with 7 classes. Next, the KMU_FED dataset includes 1106 samples with 6 classes.

Fig. 3 shows the classifier results of the OBSODL_FER approach on the database of KDEF. Figs. 3a-3b portrays the confusion matrix obtainable by the OBSODL_FER model at 70:30 of the TRA set/ TES set. The result indirect that the OBSODL_FER system has been recognised and considered all 7 class labels. Simultaneously, Fig. 3c determines the PR study of the OBSODL_FER model. The outcome portrayed that the OBSODL_FER procedure has gained higher PR accuracy under 7 class labels. Finally, Fig. 3d portrays the ROC review of the OBSODL_FER technique. The outcome specified that the OBSODL_FER procedure has been a promise results with the uppermost ROC values under 7 classes.

The FER results of the OBSODL_FER model are portrayed under the KDEF dataset in Table 1 and Fig. 4. The experimental results inferred that the OBSODL_FER system reached enhanced performance under all kinds of facial emotions. On 70% of the TRA set, the OBSODL_FER tactic delivers average $accu_y$, $prec_n$, $reca_l$, F_{score} , and $G_{measure}$ of 99.58%, 98.55%, 98.55%, 98.55%, and 98.55% correspondingly. Furthermore, on 30% of TES set, the OBSODL_FER methodology delivers average $accu_y$, $prec_n$, $reca_l$, F_{score} , and $G_{measure}$ of 99.75%, 99.12%, 99.11%, 99.11%, and 99.11% correspondingly.

Fig. 5 displays the accuracy of training TR_{accu_y} and VL_{accu_y} of the OBSODL_FER model on the dataset of KDEF. The TL_{accu_y} is definite by the assessment of the OBSODL_FER procedure on the TRA database while the VL_{accu_y} is intended by estimating the solution on a distinct dataset of TES. The results show that TR_{accu_y} and VL_{accu_y} growth with a gain in epochs. As an outcome, the

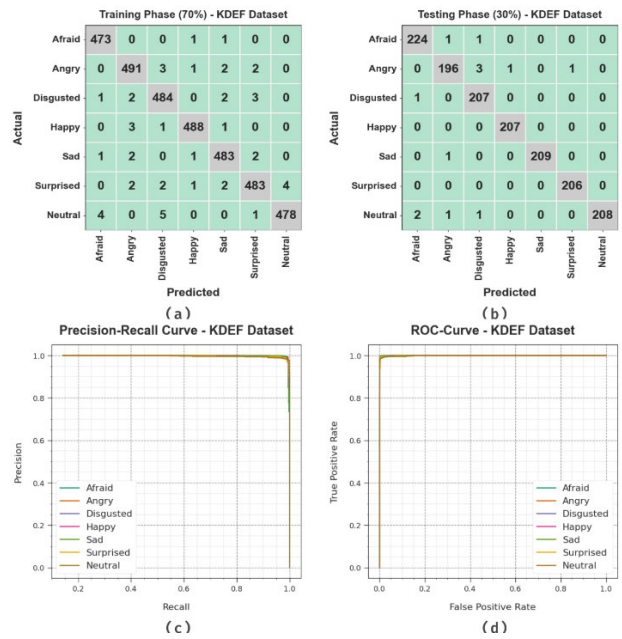


FIGURE 3. Performances on KDEF dataset (a-b) Confusion matrices, (c) PR curve, and (d) ROC.

TABLE 1. FER outcome of OBSODL_FER system on the KDEF dataset.

KDEF Dataset					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	$G_{measure}$
TRA Phase (70%)					
Afraid	99.77	98.75	99.58	99.16	99.16
Angry	99.50	98.20	98.40	98.30	98.30
Disgusted	99.45	97.78	98.37	98.07	98.08
Happy	99.74	99.19	98.99	99.09	99.09
Sad	99.59	98.37	98.77	98.57	98.57
Surprised	99.45	98.37	97.77	98.07	98.07
Neutral	99.59	99.17	97.95	98.56	98.56
Average	99.58	98.55	98.55	98.55	98.55
TES Phase (30%)					
Afraid	99.66	98.68	99.12	98.90	98.90
Angry	99.46	98.49	97.51	98.00	98.00
Disgusted	99.59	97.64	99.52	98.57	98.58
Happy	99.93	99.52	100.00	99.76	99.76
Sad	99.93	100.00	99.52	99.76	99.76
Surprised	99.93	99.52	100.00	99.76	99.76
Neutral	99.73	100.00	98.11	99.05	99.05
Average	99.75	99.12	99.11	99.11	99.11

solution of the OBSODL_FER method gets enlarged on the TRA and TES datasets with a growth in the amount of epochs.

In Fig. 6, the TR_{loss} and VR_{loss} results of the OBSODL_FER system on the KDEF dataset are exposed.

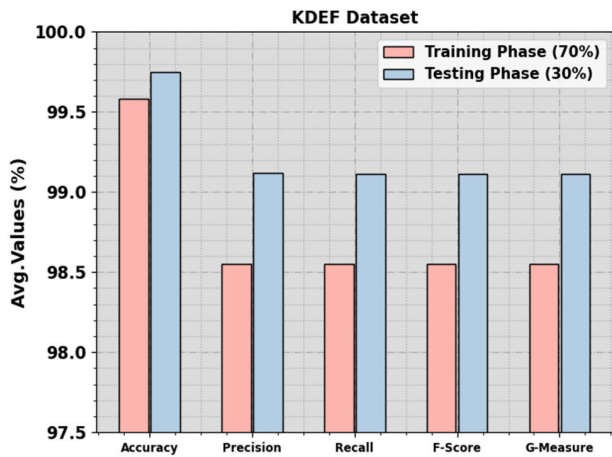


FIGURE 4. Average outcome of the OBSODL_FER technique on the KDEF dataset.

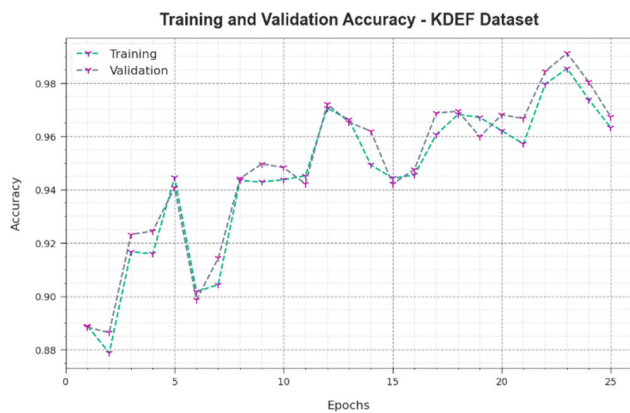


FIGURE 5. Accy curve of OBSODL_FER technique on KDEF dataset.



FIGURE 6. Loss curve of OBSODL_FER technique on the KDEF dataset.

The TR_{loss} defines the fault amongst the predictive performance as well as unique values on the TRA dataset. The VR_{loss} signify the amount of the solution of the OBSODL_FER approach on distinct validation data. The outcomes designate that the TR_{loss} and VR_{loss} tend to

decrease with rising epochs. It defined the amended performance of the OBSODL_FER technique and its capability to make precise identification. The declined value of TR_{loss} and VR_{loss} demonstrates the enhanced solution of the OBSODL_FER methodology in seizing designs and relations.

The higher performance of the OBSODL_FER model on the KDEF dataset is certified utilizing a comparison study, as shown in Fig. 7 [11s]. The outcomes specified that the MULTICNN, RTCNN, ALEXNET_LDA, and MSLBP_SVM techniques are reported worse performance with $accu_y$ of 89.34%, 87.74%, 88.84%, and 89.23% correspondingly. Meanwhile, the HDNN and DL_FER models have shown moderate performance. Although the SSO_DLFER technique accomplishes near optimal $accu_y$ of 99.69%, the OBSODL_FER methodology gains a higher $accu_y$ value of 99.75%.

Fig. 8 exhibits the classifier results of the OBSODL_FER system on KMU_FED datasets. Figs. 8a-8b depicts the confusion matrix offered by the OBSODL_FER model at 70:30 of the TRA set/TES set. The result indicated that the OBSODL_FER technique has perceived and categorized all 6 classes. Simultaneously, Fig. 8c displays the PR inspection of the OBSODL_FER method. The result showed that the OBSODL_FER technique has achieved high PR effects below 6 classes. Finally, Fig. 8d depicts the ROC inspection of the OBSODL_FER system. The result showed that the OBSODL_FER model has ensued in promising effects with high ROC values below 6 classes.

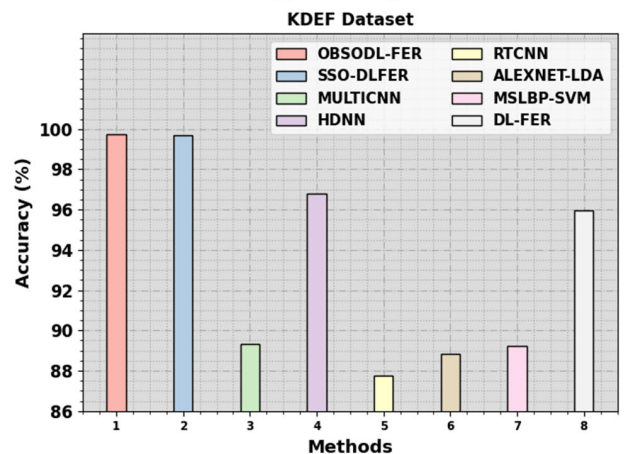


FIGURE 7. Accy outcome of OBSODL_FER technique on the KDEF dataset.

The FER results of the OBSODL_FER method are portrayed under the KMU_FED dataset in Table 2 and Fig. 9. The result showed that the OBSODL_FER model achieved enhanced performance under all kinds of facial emotions. On 70% of the TRA set, the OBSODL_FER process offers average $accu_y$, $prec_n$, $reca_l$, F_{score} , and $G_{measure}$ of 99.61%, 98.83%, 98.86%, 98.84%, and 98.84% correspondingly. Furthermore, on 30% of the TES set, the OBSODL_FER system

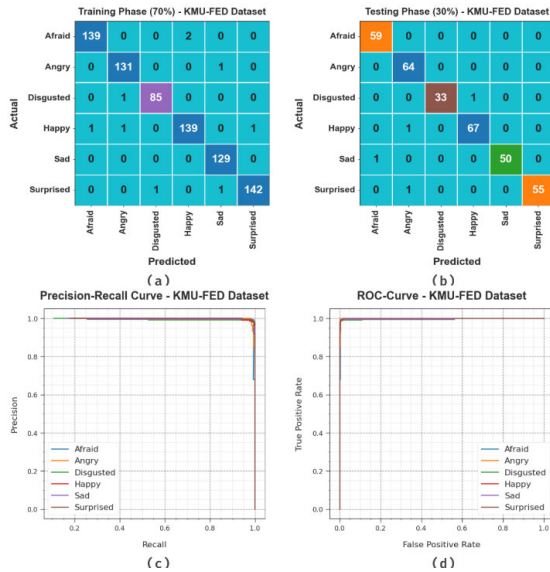


FIGURE 8. Performances on KMU_FED dataset (a_b) Confusion matrices, (c) PR_curve, and (d) ROC.

offers average $accu_y$, $prec_n$, $reca_l$, F_{score} , and $G_{measure}$ of 99.60%, 98.97%, 98.64%, 98.79%, and 98.80% respectively.

TABLE 2. FER outcome of OBSODL_FER technique on the KMU_FED dataset.

KMU_FED Dataset					
Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	$G_{measure}$
TRA Phase (70%)					
Afraid	99.61	99.29	98.58	98.93	98.93
Angry	99.61	98.50	99.24	98.87	98.87
Disgusted	99.74	98.84	98.84	98.84	98.84
Happy	99.35	98.58	97.89	98.23	98.23
Sad	99.74	98.47	100.00	99.23	99.23
Surprised	99.61	99.30	98.61	98.95	98.96
Average	99.61	98.83	98.86	98.84	98.84
TES Phase (30%)					
Afraid	99.70	98.33	100.00	99.16	99.16
Angry	99.40	96.97	100.00	98.46	98.47
Disgusted	99.70	100.00	97.06	98.51	98.52
Happy	99.40	98.53	98.53	98.53	98.53
Sad	99.70	100.00	98.04	99.01	99.01
Surprised	99.70	100.00	98.21	99.10	99.10
Average	99.60	98.97	98.64	98.79	98.80

Fig. 10 displays the accuracy of training of TR_{accu_y} and VL_{accu_y} of the OBSODL_FER approach on the KMU_FED dataset. The TL_{accu_y} is defined by the estimation of the OBSODL_FER procedure on the TRA dataset whereas the

VL_{accu_y} is multiplied by assessing the solution on a distinct testing database. The results display that TR_{accu_y} and VL_{accu_y} rise with a gain in epochs. As an outcome, the solution of the OBSODL_FER technique acquires amended on the datasets of TRA and TES with a growth in the number of epochs.

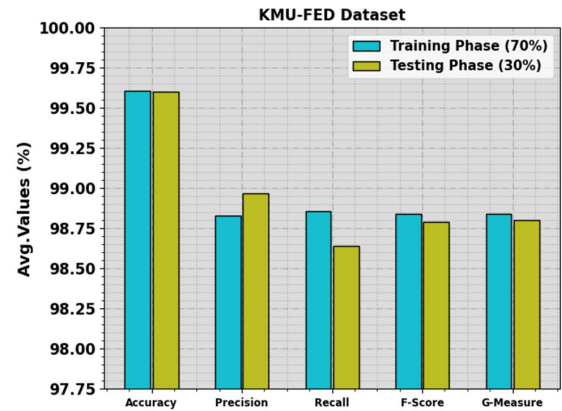


FIGURE 9. Average outcome of the OBSODL_FER method on the KMU_FED dataset.

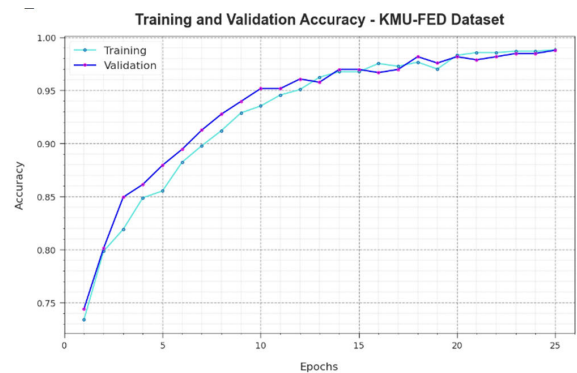


FIGURE 10. Accu_y curve of OBSODL_FER technique on KMU_FED dataset.



FIGURE 11. Loss curve of OBSODL_FER technique on KMU_FED dataset.

In Fig. 11, the TR_{loss} and VR_{loss} outcomes of the OBSODL_FER technique on the KMU_FED dataset are shown. The TR_{loss} describes the fault amongst the predictive performance as well as values of unique on the TRA

dataset. The VR_{loss} represent the size of the solution of the OBSODL_FER method on individual validation data. The results indicate that the TR_{loss} and VR_{loss} incline to decrease with rising epochs. It represented the enhanced solution of the OBSODL_FER technique and its capability to produce precise identification. The condensed value of TR_{loss} and VR_{loss} demonstrates the enhanced solution of the OBSODL_FER technique in seizing patterns and relations.

The superior outcome of the OBSODL_FER approach on the KMU_FED database is ensured utilizing a comparison study, as shown in Fig. 12. The outcomes definite that the SqueezeNet, d-RFs, FTDRF, and WRF models have reported worse performance with $accu_y$ of 89.54%, 91.65%, 93.17%, and 93.46% respectively. Meanwhile, the MobileNetV3 and LMRF models have shown moderate performance. Although the SSO_DLFEF technique accomplishes near optimal $accu_y$ of 99.50%, the OBSODL_FER technique gains a higher $accu_y$ value of 99.61%.

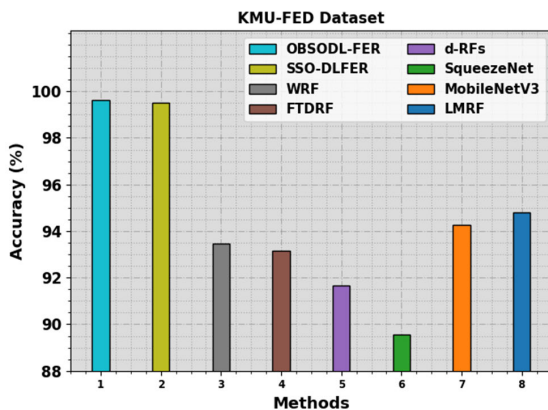


FIGURE 12. $Accu_y$ outcome of OBSODL_FER technique on KMU_FED dataset.

These outcomes proved the greater performance of the OBSODL_FER procedure over other techniques.

V. CONCLUSION

This study focuses on the development of the OBSODL-FER method for AIS. The purpose of the OBSODL-FER method is to identify and classify dissimilar types of facial emotions of the drives in autonomous vehicles. The proposed model incorporates 4 phases of processes as Xception model, OBSO-based parameter tuning, ILSTM identification, and JFS-based hyperparameter tuning. In this research, the presented OBSODL-FER technique exploits the OBSO technique for the hyperparameter range of the Xception algorithm. Finally, the JFS model has been employed for the finest hyperparameter selection of the ILSTM methodology. The simulation outcomes of the OBSODL-FER method is verified on a benchmark facial emotion dataset. The experimentation outcomes concluded the enhancement of the OBSODL-FER approach than other DL procedures.

REFERENCES

- [1] J. Bao, X. Tao, and Y. Zhou, "An emotion recognition method based on eye movement and audiovisual features in MOOC learning environment," *IEEE Trans. Computat. Social Syst.*, vol. 11, no. 1, pp. 171–183, Feb. 2024.
- [2] C. Pabba and P. Kumar, "An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition," *Expert Syst.*, vol. 39, no. 1, Jan. 2022, Art. no. e12839.
- [3] F. M. Talaat, "Real-time facial emotion recognition system among children with autism based on deep learning and IoT," *Neural Comput. Appl.*, vol. 35, no. 17, pp. 12717–12728, Jun. 2023.
- [4] Y. Kumar, S. K. Verma, and S. Sharma, "Multi-pose facial expression recognition using hybrid deep learning model with improved variant of gravitational search algorithm," *Int. Arab J. Inf. Technol.*, vol. 19, no. 2, pp. 281–287, 2022.
- [5] H. A. Meshram, M. G. Sonkusare, P. Acharya, and S. Prakash, "Facial emotional expression regulation to control the semi-autonomous vehicle driving," in *Proc. IEEE Int. Conf. for Innov. Technol. (INOCON)*, Nov. 2020, pp. 1–5.
- [6] K. Teoh, R. Ismail, S. Naziri, R. Hussin, M. Isa, and M. Basir, "Face recognition and identification using deep learning approach," *J. Phys., Conf. Ser.*, vol. 1755, no. 1, Feb. 2021, Art. no. 012006.
- [7] M. Rakhra, D. Singh, A. Singh, K. D. Garg, and D. Gupta, "Face recognition with smart security system," in *Proc. 10th Int. Conf. Rel., INFOCOM Technol. Optim. Trends Future Directions (ICRITO)*, Oct. 2022, pp. 1–6.
- [8] X. Wang, X. Chen, and C. Cao, "Human emotion recognition by optimally fusing facial expression and speech feature," *Signal Process., Image Commun.*, vol. 84, May 2020, Art. no. 115831.
- [9] N. K. Benamara, M. Val-Calvo, J. R. Álvarez-Sánchez, A. Díaz-Morcillo, J. M. Fernández-Vicente, E. Fernández-Jover, and T. B. Stambouli, "Real-time facial expression recognition using smoothed deep neural network ensemble," *Integr. Comput.-Aided Eng.*, vol. 28, no. 1, pp. 97–111, Dec. 2020.
- [10] B. Kocacinar, B. Tas, F. P. Akbulut, C. Catal, and D. Mishra, "A real-time CNN-based lightweight mobile masked face recognition system," *IEEE Access*, vol. 10, pp. 63496–63507, 2022.
- [11] D. K. Jain, A. K. Dutta, E. Verdú, S. Alsbai, and A. R. W. Sait, "An automated hyperparameter tuned deep learning model enabled facial emotion recognition for autonomous vehicle drivers," *Image Vis. Comput.*, vol. 133, May 2023, Art. no. 104659.
- [12] L. A. Corujo, E. Kieson, T. Schloesser, and P. A. Gloor, "Emotion recognition in horses with convolutional neural networks," *Future Internet*, vol. 13, no. 10, p. 250, Sep. 2021.
- [13] R. Chauhan, S. Shighra, H. Madkhali, L. Nguyen, and M. Prasad, "Efficient future waste management: A learning-based approach with deep neural networks for smart system (LADS)," *Appl. Sci.*, vol. 13, no. 7, p. 4140, Mar. 2023.
- [14] H. Ge, Z. Zhu, Y. Dai, B. Wang, and X. Wu, "Facial expression recognition based on deep learning," *Comput. Methods Programs Biomed.*, vol. 215, May 2022, Art. no. 106621.
- [15] A. R. Shahid and H. Yan, "SqueezeExpNet: Dual-stage convolutional neural network for accurate facial expression recognition with attention mechanism," *Knowl.-Based Syst.*, vol. 269, Jun. 2023, Art. no. 110451.
- [16] R. R. Devaram, G. Beraldo, R. De Benedictis, M. Mongiovì, and A. Cesta, "LEMON: A lightweight facial emotion recognition system for assistive robotics based on dilated residual convolutional neural networks," *Sensors*, vol. 22, no. 9, p. 3366, Apr. 2022.
- [17] W. M. Alenazy and A. S. Alqahtani, "Gravitational search algorithm based optimized deep learning model with diverse set of features for facial expression recognition," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 2, pp. 1631–1646, Feb. 2021.
- [18] R. Ni, B. Yang, X. Zhou, A. Cangelosi, and X. Liu, "Facial expression recognition through cross-modality attention fusion," *IEEE Trans. Cogn. Develop. Syst.*, vol. 15, no. 1, pp. 175–185, Mar. 2023.
- [19] K. K. Singh, M. Siddhartha, and A. Singh, "Diagnosis of coronavirus disease (COVID-19) from chest X-ray images using modified XceptionNet," *Romanian J. Inf. Sci. Technol.*, vol. 23, no. 657, pp. 91–115, 2020.
- [20] K. Rajesh Kumar and M. Vijayakumar, "Optimization of cognitive femto-cell network via oppositional beetle swarm optimization algorithm," *Intell. Autom. Soft Comput.*, vol. 36, no. 1, pp. 819–832, 2023.
- [21] H. Xiao, M. A. Sotelo, Y. Ma, B. Cao, Y. Zhou, Y. Xu, R. Wang, and Z. Li, "An improved LSTM model for behavior recognition of intelligent vehicles," *IEEE Access*, vol. 8, pp. 101514–101527, 2020.
- [22] A. Shaheen, R. El-Seheimy, S. Kamel, and A. Selim, "Reliability enhancement and power loss reduction in medium voltage distribution feeders using modified jellyfish optimization," *Alexandria Eng. J.*, vol. 75, pp. 363–381, Jul. 2023.