sustainability

MDPI

Article

Manta Ray Foraging Optimization with Transfer Learning Driven Facial Emotion Recognition

Anwer Mustafa Hilal [1,*], Dalia H. Elkamchouchi [2], Saud S. Alotaibi [3], Mohammed Maray [4], Mahmoud Othman [5], Amgad Atta Abdelmageed [1], Abu Sarwar Zamani [1] and Mohamed I. Eldesouki [6]

[1] Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj 16278, Saudi Arabia
[2] Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[3] Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Mecca 24382, Saudi Arabia
[4] Department of Information Systems, College of Computer Science, King Khalid University, Abha 62529, Saudi Arabia
[5] Department of Computer Science, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo 11835, Egypt
[6] Department of Information System, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, AlKharj 16278, Saudi Arabia
* Correspondence: a.hilal@psau.edu.sa

Abstract: Recently, facial expression-based emotion recognition techniques obtained excellent outcomes in several real-time applications such as healthcare, surveillance, etc. Machine-learning (ML) and deep-learning (DL) approaches can be widely employed for facial image analysis and emotion recognition problems. Therefore, this study develops a Transfer Learning Driven Facial Emotion Recognition for Advanced Driver Assistance System (TLDFER-ADAS) technique. The TLDFER-ADAS technique helps proper driving and determines the different types of drivers' emotions. The TLDFER-ADAS technique initially performs contrast enhancement procedures to enhance image quality. In the TLDFER-ADAS technique, the Xception model was applied to derive feature vectors. For driver emotion classification, manta ray foraging optimization (MRFO) with the quantum dot neural network (QDNN) model was exploited in this work. The experimental result analysis of the TLDFER-ADAS technique was performed on FER-2013 and CK+ datasets. The comparison study demonstrated the promising performance of the proposed model, with maximum accuracy of 99.31% and 99.29% on FER-2013 and CK+ datasets, respectively.

Keywords: transfer learning; metaheuristics; facial emotion recognition; driver assistance system; smart cities

check for updates

Citation: Mustafa Hilal, A.; Elkamchouchi, D.H.; Alotaibi, S.S.; Maray, M.; Othman, M.; Abdelmageed, A.A.; Zamani, A.S.; Eldesouki, M.I. Manta Ray Foraging Optimization with Transfer Learning Driven Facial Emotion Recognition. *Sustainability* **2022**, *14*, 14308. https://doi.org/10.3390/su142114308

Academic Editors: Amjad Ali, Farman Ali, Jin-Ghoo Choi and Muhammad Shafiq

Received: 4 October 2022
Accepted: 27 October 2022
Published: 2 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the drastic increase in urban population, smart cities become popular in several sectors such as environmental sustainability, healthcare transportation, etc. [1–3]. The transformation of traditional cities into smart cities will depend greatly on modern technology in computing paradigms, particularly machine learning (ML), Internet-of-Things (IoT), data mining (DM), and artificial intelligence (AI) [4]. In the future, each conglomerate network of traditional cities (for example, electricity, information, transportation, and so on) will be assisted by a large number of IoT devices, generating a wide variety of heterogeneous and unstructured data in return [5,6]. Drivers' emotional state could affect their driving capability while driving vehicles. Because of the growing complexity of vehicles, identifying the emotions of drivers has become increasingly prominent [7]. To guarantee a more pleasant and secure ride, good infotainment could accurately identify the driver's emotional state beforehand, making adjustments to the vehicle dynamics [8–10]. In smart

cars, it is crucial to identify the emotion of the driver since the vehicle could make decisions regarding what to do in specific situations based on the driver's psychological states (for instance autonomous driving, driving modes, and mood-altering songs) [11]. Emotions like disgust, sadness, fear, and anger influence the driver's ability to cause road accidents. This system might assist to avoid road accidents and control the functions of the vehicles based on the driver's emotions [12].

Driver emotion detection in advanced driver assistance systems (ADAS) can be achieved by the facial expression recognition (FER) technique [13]. FER can be accomplished by using a manual feature extraction method with ML-based classifier implementation or using deep neural network (DNN) or a hybrid model including SVM as classifiers to convolution neural network (CNN). The computer vision (CV) based DL method is widely employed for emotion monitoring and FER [14,15]. This study analysis of deep FER summarized facial expression dataset and collection environments, namely the Internet or laboratories. But the present work is implemented mainly on lab-captured datasets because of the deficiency of real-time datasets [16,17]. There exists no on-road driver facial data accessible for driver FER tasks, and the driving task might suppress facial expression. Because of these problems, there is a lack of on-road driver FER studies essential for automotive human-machine systems [18]. However, various aspects affect this type of model, from predicting the appropriate class of expression to accomplishing higher performance.

Though several FER models are available in the literature, there is still a need to improve recognition performance. At the same time, the trial-and-error selection of model parameters is a tedious process. Therefore, this study develops a Transfer Learning Driven Facial Emotion Recognition for Advanced Driver Assistance System (TLDFER-ADAS) technique. The TLDFER-ADAS technique helps proper driving and determines the emotion of drivers, which assures road safety. The TLDFER-ADAS technique initially performs contrast enhancement procedures to enhance image quality. In the TLDFER-ADAS technique, the Xception model was applied to derive feature vectors. For driver emotion classification, manta ray foraging optimization (MRFO) with a quantum dot neural network (QDNN) model was exploited in this work. The proposed system can be used to classify the different emotions of drivers such as anger, disgust, fear, happiness, sadness, surprise, and neutral. The experimental result analysis of the TLDFER-ADAS technique was performed on a benchmark dataset. In short, the paper's contributions can be summarized as follows.

- An intelligent TLDFER-ADAS technique encompassing preprocessing, Xception feature extraction, QDNN classification, and MRFO parameter tuning is presented for facial emotion classification;
- To the best of our knowledge, the presented TLDFER-ADAS technique never existed in the literature;
- Parameter tuning of the QDNN model using the MRFO algorithm helps in accomplishing significant classification performance;
- The emotion recognition performance was validated on two facial datasets: FER-2013 and CK+ datasets.

The rest of the paper is organized as follows: Section 2 offers a brief survey of existing works; Section 3 introduces the proposed model; Section 4 provides a detailed performance validation; and Section 5 draws the conclusions.

2. Related Works

Naqvi et al. [19] presented a multimodal technique for aggressive behavior detection of drivers from remote places. This model depends upon the variations in gaze and facial emotions of drivers while driving with the use of near-infrared (NIR) camera sensors and an illuminator installed in vehicle. Driver's aggressive and normal time-series data were collected while playing car racing and truck driving computer games, respectively, while using a driving game simulator. Paikrao et al. [20] proposed a frontend processing model for stress emotion identification in various noisy environments. The proposed model analyzed noisy speech emotions with the presence of background noise, extracted

features using Mel-frequency cepstral coefficients (MFCC) features, and evaluated the overall system performance.

Jeong and Ko [21] developed a faster FER technique to monitor a driver's emotion that was able to operate with a lower-specification device mounted on vehicles. For these purposes, a hierarchical weighted random forest (WRF) classification was trained by using the similarity of the sample dataset to increase its performance. Initially, facial landmarks discovered from geometric features and input images were extracted, which considered the spatial location among landmarks. In [22], a novel hybrid network structure was developed based on DNN and SVM to forecast among six as well as seven drivers' emotions in dissimilar illumination conditions, poses, and occlusions to accomplish this purpose. To define the emotion, a combination of LBP and Gabor features was exploited for determining the feature and categorized by means of an SVM classification integrated with the CNN. Xiao et al. [23] developed a facial expression-based-on-road-driver-emotion detection network named FERDERnet. This technique split the on-road driver FER tasks into three mechanisms: a face recognition model which identifies the driver's face, an augmentation-based resampling model which implements resampling and data augmentation, and an emotion recognition model that adopts a DCNN pre-trained on CK+ and FER data and later finetuned as a backbone for recognizing driver emotions.

Mehendale [24] developed a FER technique based on CNNs (FERC). The method can be classified into different parts: Initially, it eliminates the background from the picture, and then concentrate on extracting facial feature vectors. In FERC method, expressional vector (EV) is utilized for finding the five distinct kinds of regular facial expressions. The two-level CNN functions sequentially, and the final layer of perceptron alters the exponent values and weights with all the iterations. Generally, FERC varies from following strategy with single-level CNN, thus enhancing the performance.

Naqvi et al. [19] designed a multi-modal based technique to remotely identify driver aggressiveness for managing this problem. The presented technique depends on variations in the facial and gaze emotions of driver while driving with the use of a near-infrared (NIR) camera sensor and an illuminator mounted on vehicles. Normal time-series and driver aggression datasets are gathered while using a driving game simulator, playing truck driving computer, and car racing games, correspondingly. Oh et al. [25] presented a DL-based driver's real emotion recognizer (DRER), to identify the drivers' real emotions that could not be wholly recognized according to facial expressions. Li et al. [26] proposed a cognitive-feature-augmented driver emotion detection model that depends on deep networks and emotional cognitive methods. The convolution technique was selected to create the approach for driver emotion recognition, while concurrently considering cognitive process characteristics and the driver's facial expression.

3. The Proposed Model

In this study, we have introduced a new TLDFER-ADAS technique for smart city environments. The TLDFER-ADAS technique aids in proper driving and determines the emotions of drivers, which assures road safety. The TLDFER-ADAS technique encompasses a sequence of operations like contrast enhancement, Xception-based extracting features, QDNN-based emotion classification, and MRFO-hyperparameter tuning. Figure 1 defines the overall block diagram of TLDFER-ADAS system.

3.1. Contrast Enhancement

The TLDFER-ADAS technique originally performed contrast enhancement procedure for enhancing image quality. The contrast limited adaptive histogram equalization (CLAHE) technique enhances the lower contrast problem for digital images. CLAHE performed better than adaptive histogram equalization (AHE) and normal histogram equalization (HE) [27]. Generally, CLAHE works with the constraint of contrast enhancement viz., generally accomplished by ordinary HE that results in noise enhancement. Therefore, with the constraint of contrast enhancement in HE, the desired outcome is accomplished in the case

where noise performs a pivotal role by enhancing contrast, especially in healthcare images. In general, contrast enhancement is defined as the slope of function viz., input image intensity value to a desirable resulting image intensity. The contrast is limited by the constraint of the slope of the related function. Additionally, contrast enhancement is strongly connected to the histogram height at that intensity value. Therefore, limiting the slope and clipping the height of histogram is a similar function that controls contrast enhancement.

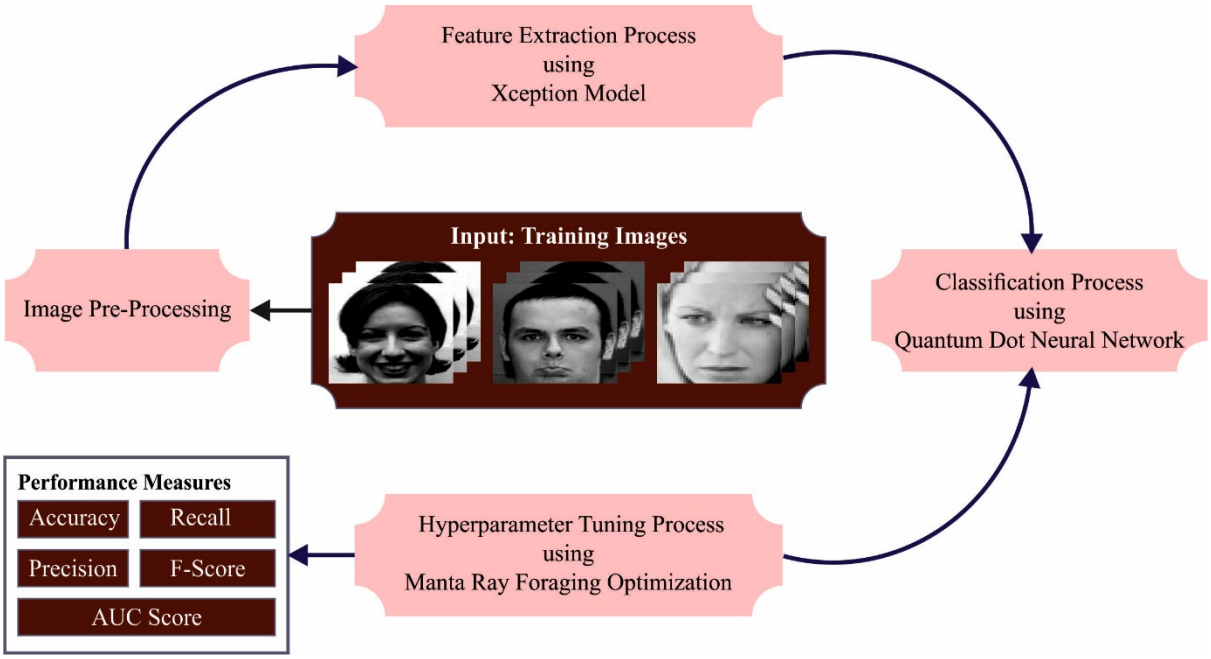


Figure 1. Overall block diagram of TLDFER-ADAS system.

3.2. Xception Based Feature Extraction

In the TLDFER-ADAS technique, the Xception model is applied to produce feature vectors. In this study, the Xception network with a depthwise separable convolutional model was adopted as the CNN for the classification of driver emotions [28]. Xception is intended for separating spatial correlations and cross-channel correlations completely. Thus, it applies depth-wise separable convolutions comprising pointwise and depth-wise layers. The depth-wise part separates the channel completely and performs a size 3×3 convolutional process. This technique generates one feature map for all channels. The computation amount is decreased in a convolution of 3×3 size that needs significant computation, thereby avoiding the bottleneck phenomenon [29]. The pointwise layer implements a 1×1 convolution process on all channels of the depth-wise outputs. By using the depth-wise separable convolution, extracting feature was effectively implemented, and the amount of computation is decreased. Furthermore, a deep network was created by using approaches like batch normalization to Xception and skip connection of ResNet.

To adjust the hyperparameters, the Adam optimizer is used. Adam is a first order gradient based stochastic objective function optimization method [30]. Adam integrates the benefits of the RMSProp and AdaGrad methods; the former is utilized for sparse gradient problems, and the latter is utilized for unfixed and non-linear optimization problems. Adam has the advantage of low memory requirements, easier implementation, and high computing efficiency. Its gradient diagonal scaling is invariant, hence it is appropriate for resolving problems with parameters or large-scale data. For distinct variables, Adam can upgrade the weight of the NN iteratively and adjust the learning rate adaptively based on the training dataset.

3.3. Driver Emotion Recognition

For driver emotion classification, the QDNN model was exploited in this work. The QDNN comprises of single electron stuck inside a cage of atoms [31]. In each ANN architecture, the neuron obtains input from other processors via weighted connection, and calculates the output that is passed through other neurons. The calculated output y_i of *i-th* neurons is implemented on the signals $\{y_j\}$ from other neurons as:

$$y_i = \sum_j w_{ij} f_j(y_j) \tag{1}$$

In Equation (1), w_{ij} represents the weight between *j-th* to *i-th* neurons and f_j shows the activation function for *j-th* neurons. Consequently, the mathematical formula can be represented in the following equation:

$$|\Omega\left(\mathcal{Y}_f,\ S\right)\rangle = \lim_{N\to\infty} \int_{(y_0,0)}^{y_{N+1}\mathcal{Y}_f,S} dy_1 \cdots dy_N(s)\left(\frac{m}{2\,\pi i n\,\Delta t}\right)^{\frac{N+1}{2}} \times \exp\left(\frac{i\Delta t}{n}\sum_{j=0}^{N}\left[\frac{m}{2}\left(\frac{y_{j+1}-y_j}{\Delta t}\right)^2 - V\left(y_j\right)\right]\right|\Omega(y_0,0)\rangle \tag{2}$$

From the abovementioned, $|\Omega(y_0,0)\}$ indicates the input state of quantum model, and $\left|\Omega\left(y_f,\ S\right)\right\rangle$ denotes the output state, at time $t = S$. G shows the Green's function that propagates the system forward in time, from an early location y_0 at time $t = 0$ to last location y_f at time $t = S$. Equation (2) formulates G in Feynman path integral formulation of quantum mechanics. The dots are strongly associated with one another such that tunneling can be possible between any two adjacent dot molecules.

$$|\Omega(\sigma_z(N,\ \Delta t),\ S)\} = \sum_{\sigma_z^{(j\Delta t)}} \exp\left(\frac{i}{h}\sum_j [K\sigma_X(i\Delta t) + \in (j\Delta t)\sigma_z(j\Delta t)]\right) I[\sigma_z(t)]|\Omega(\sigma_z(0)),0\rangle \tag{3}$$

where (t) represents a finite set of sums over state of the polarization, σ_z, at every time slice $j\Delta t$ at every time slice, the polarization is $+1$ or -1. Figure 2 showcases the framework of QDNN technique [32].

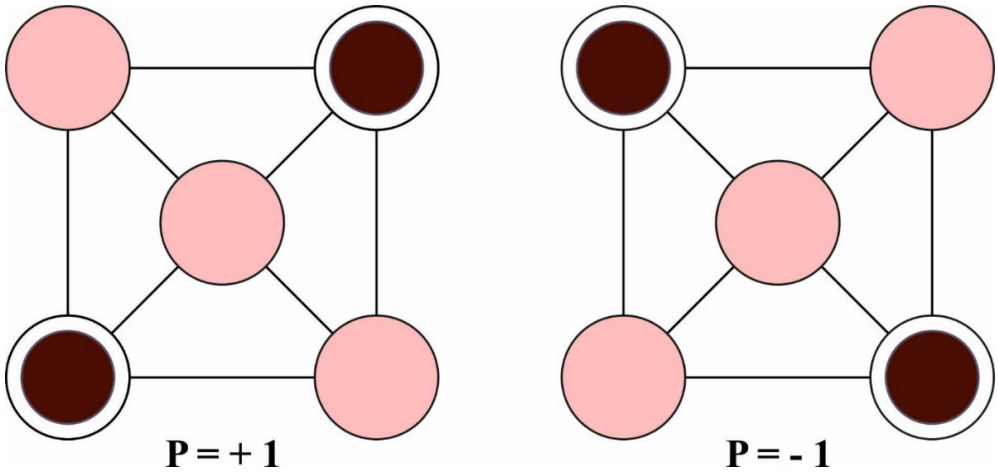


Figure 2. Structure of QDNN.

At the last level, the MRFO algorithm is used as a parameter tuning technique. The stimulation of MRFO is based on the smart foraging performances of MR. It takes 3 exclusive foraging rules of manta ray (MR) to detect a better food source [33]. MRFO can be functioned by 3 foraging performances like Cyclone foraging, Somersault foraging, and Chain foraging. A few numerical approaches are provided as follows. Once the plankton concentration has been increased, afterward the position is optimum where all

the positions are upgraded by remarkable solution predictable. This numerical approach to chain foraging is written as:

$$C_x^{dim}(n+1) = C_x^{dim}(n) + rand \cdot C_{best}^{dim}(n) + \varphi C_{best}^{dim}(n) - C_x^{dim}(n), \quad X = 1 \tag{4}$$

$$C_x^{dim}(n+1) = C_x^{dim}(n) + rand \cdot C_{x-1}^{dim}(n) - C_x^{dim}(n) + \varphi C_{best}^{dim}(n) - C_x^{dim}(n) X = 2 \ldots\ldots N \tag{5}$$

whereas, signifies the location of xth individual at time n in dim implies the dimensional, $rand$ denotes the arbitrary vector in zero, and one, φ signifies the weighted coefficient, represents the plankton with maximal concentration. The numerical notion of spiral-shaped events of MRs is defined in the subsequent equation:

$$C_x(n+1) = C_{best} + rand \cdot \left(C_{x-1}(n) - C_x(n) + r^{at} \cdot \cos(2\pi t) \cdot (C_{best} - C_x(n))\right) \tag{6}$$

$$D_x(n+1) = D_{best} + rand \cdot \left(D_{x-1}(n) - D_x(n) + r^{at} \cdot \cos(2\pi t) \cdot (D_{best} - D_x(n))\right) \tag{7}$$

This performance is upgraded to d space. The arithmetical method of cyclone foraging was defined as:

$$C_x^{dim}(n+1) = C_{best}^{dim} + rand \cdot \left(C_{best}^{d}(n) - C_x^{d}(n) + \alpha C_{best}^{dim}(n) - C_x^{dim}(n)\right), \; X = 1 \tag{8}$$

$$C_x^{dim}(n+1) = C_{best}^{dim} + rand \cdot \left(C_{best}^{d}(n) - C_x^{d}(n) + \alpha C_{best}^{dim}(n) - C_x^{dim}(n)\right), \; X = 2 \ldots\ldots N \tag{9}$$

$$A = 2E^{rand1 \frac{T-t+1}{T}} \cdot Sin(2\pi rand1) \tag{10}$$

whereas α represents the weighted coefficient, T stands for the superior number of iterations, and $rand1$ represents the rand value in zero and one [34]. All the individuals search for a new location apart from the state-of-the-art improved individual, with transfer of a new arbitrary place from the searching space location:

$$C_r^{dim} = LB^{dim} + rand\left(UB^{dim} - LB^{dim}\right) \tag{11}$$

whereas $rand$ denotes the random position, LB and UB signify the lower as well as upper limits of dimensional, correspondingly. All the MR proposes to move and somersault to a new location as follows:

$$C_x^{dim}(n+1) = C_x^{dim}(n) + som \cdot rand2 \cdot \left(C_{best}^{dim}(n) - rand3 C_x^{dim}(n)\right), \; X = 1 \ldots N \tag{12}$$

whereas som depicts the somersault factor that chooses a somersault threshold of MRs and $Som = rand2$ and $rand3$ determine 2 random values in zero and one.

The MRFO methodology develops a fitness function (FF) for realizing higher classification results. It defines a positive integer for exemplifying a good efficiency of candidate results. During this case, the minimized classifier error rate was supposed that FF is written in Equation (13).

$$fitness(x_i) = \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \tag{13}$$

4. Results and Discussion

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU. In this section, a brief emotion recognition performance of the TLDFER-ADAS approach is tested on two databases as FER-2013 and CK+ databases. The first FER-2013

dataset has 35,527 images and the CK+ dataset holds 636 images. The details related to the database are given in Table 1. Figure 3 illustrates some sample images.

Table 1. Dataset details.

Class	FER-2013	CK+ (Last Frame)
Angry	4593	45
Disgust	547	59
Fear	5121	25
Happy	8989	69
Sad	6077	28
Surprise	4002	83
Neutral	6198	327
Total Number of Samples	**35,527**	**636**



Figure 3. Sample images.

The confusion matrices provided by the TLDFER-ADAS model on the FER-2013 dataset are reported in Figure 4. The results implied the improved emotion recognition outcomes of the TLDFER-ADAS model. It is noticeable that the TLDFER-ADAS model has proficiently identified seven distinct classes.

Table 2 and Figure 5 demonstrate a detailed outcome of the TLDFER-ADAS approach on the test FER-2013 dataset. The experimentation values inferred the enhanced emotion classification outcomes of the TLDFER-ADAS model. For instance, in the entire dataset, the TLDFER-ADAS model has reached average $accu_y$ of 99.29%, $prec_n$ of 95.80%, $reca_l$ of 96.69%, F_{score} of 96.22%, and AUC_{score} of 98.14%. Likewise, on 70% of TR dataset, the TLDFER-ADAS algorithm has attained average $accu_y$ of 99.29%, $prec_n$ of 95.69%, $reca_l$ of 96.68%, F_{score} of 96.15%, and AUC_{score} of 98.13%. At last, on 30% of TS dataset, the TLDFER-ADAS system has resulted in average $accu_y$ of 99.31%, $prec_n$ of 96.06%, $reca_l$ of 96.71%, F_{score} of 96.36%, and AUC_{score} of 98.15%.

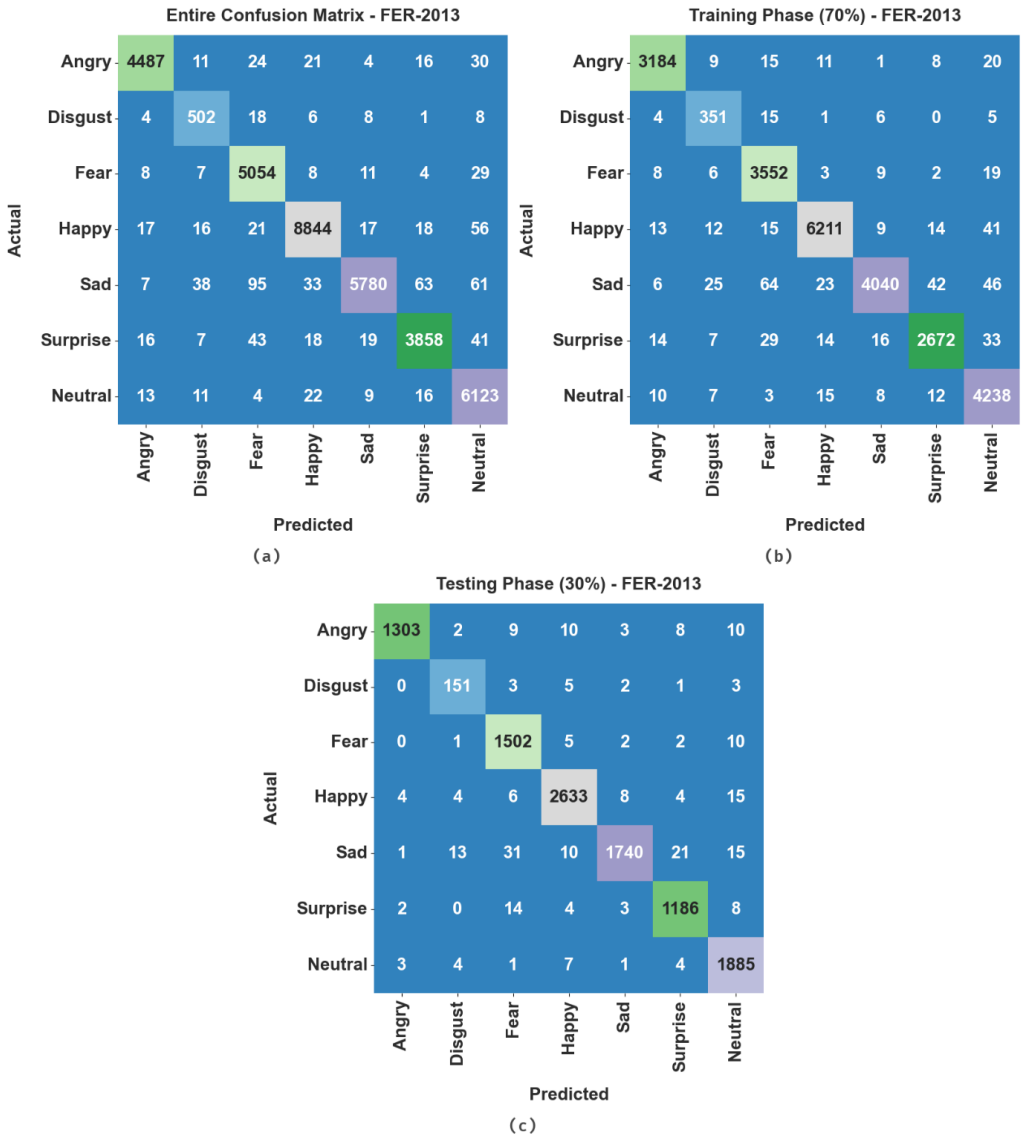


Figure 4. Confusion matrices of TLDFER-ADAS system under FER-2013 dataset: (**a**) entire database; (**b**) 70% of TR database; and (**c**) 30% of TS database.

Table 2. Result analysis of TLDFER-ADAS system with distinct classes under FER-2013 dataset.

FER-2013 Dataset					
Class	**Accuracy**	**Precision**	**Recall**	**F-Score**	**AUC Score**
Entire Dataset					
Angry	99.52	98.57	97.69	98.13	98.74
Disgust	99.62	84.80	91.77	88.15	95.76
Fear	99.23	96.10	98.69	97.38	99.01
Happy	99.29	98.79	98.39	98.59	98.99
Sad	98.97	98.84	95.11	96.94	97.44
Surprise	99.26	97.03	96.40	96.72	98.01
Neutral	99.16	96.46	98.79	97.61	99.01
Average	**99.29**	**95.80**	**96.69**	**96.22**	**98.14**

Table 2. *Cont.*

FER-2013 Dataset					
Class	**Accuracy**	**Precision**	**Recall**	**F-Score**	**AUC Score**
Training Phase (70%)					
Angry	99.52	98.30	98.03	98.17	98.89
Disgust	99.61	84.17	91.88	87.86	95.81
Fear	99.24	96.18	98.69	97.42	99.02
Happy	99.31	98.93	98.35	98.64	99.00
Sad	98.97	98.80	95.15	96.94	97.46
Surprise	99.23	97.16	95.94	96.55	97.79
Neutral	99.12	96.27	98.72	97.48	98.96
Average	**99.29**	**95.69**	**96.68**	**96.15**	**98.13**
Testing Phase (30%)					
Angry	99.51	99.24	96.88	98.04	98.38
Disgust	99.64	86.29	91.52	88.82	95.64
Fear	99.21	95.91	98.69	97.28	98.99
Happy	99.23	98.47	98.47	98.47	98.98
Sad	98.97	98.92	95.03	96.94	97.41
Surprise	99.33	96.74	97.45	97.09	98.51
Neutral	99.24	96.87	98.95	97.90	99.13
Average	**99.31**	**96.06**	**96.71**	**96.36**	**98.15**

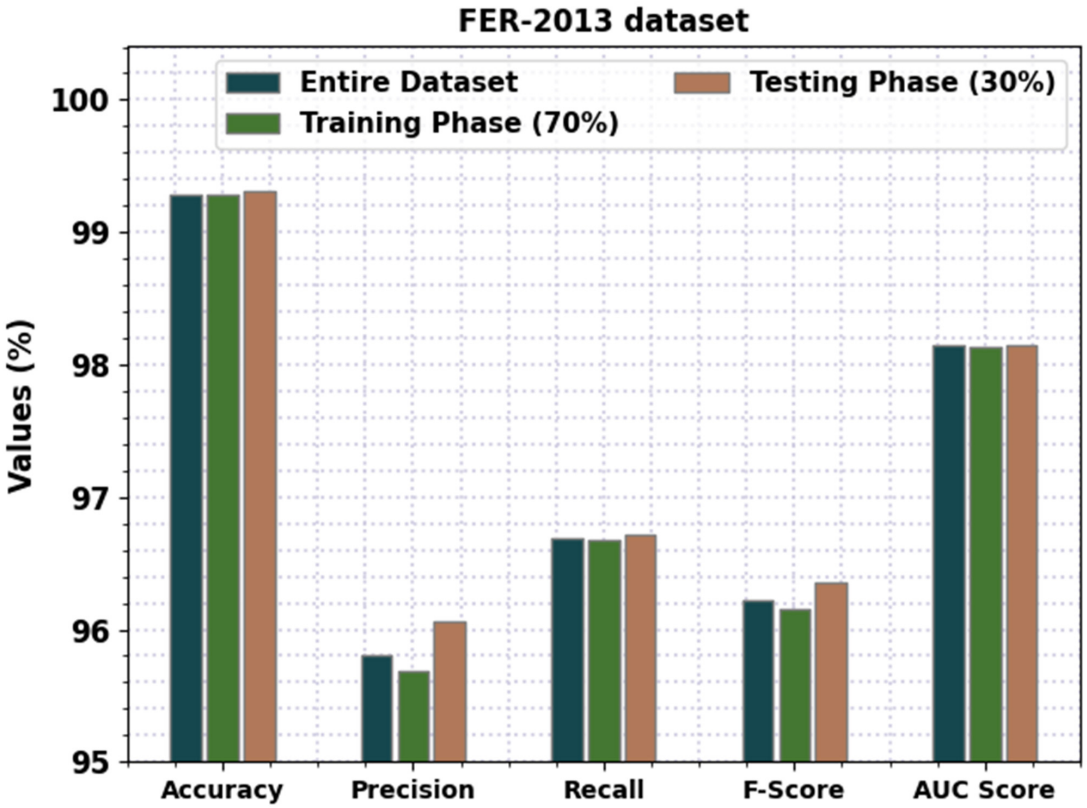


Figure 5. Result analysis of TLDFER-ADAS system under FER-2013 dataset.

The training accuracy (TR_{acc}) and validation accuracy (VL_{acc}) acquired by the TLDFER-ADAS system in FER-2013 database is exhibited in Figure 6. The simulation result pointed out the TLDFER-ADAS methodology has gained improved values of TR_{acc} and VL_{acc}. Notably, the VL_{acc} appears to exist superior to TR_{acc}.

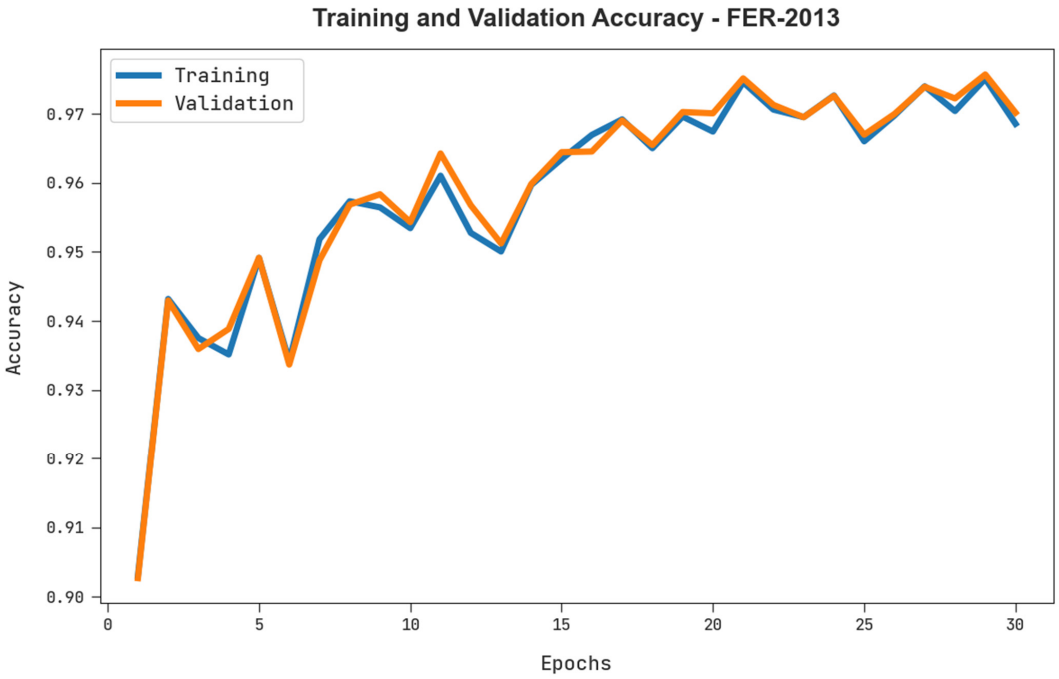


Figure 6. TR_{acc} and VL_{loss} analysis of TLDFER-ADAS system under FER-2013 dataset.

The training loss (TR_{loss}) and validation loss (VL_{loss}) accomplished by the TLDFER-ADAS system under FER-2013 database are revealed in Figure 7. The simulation result referred that the TLDFER-ADAS system has gained reduced values of TR_{loss} and VL_{loss}. In certain, the VL_{loss} is lesser than TR_{loss}.

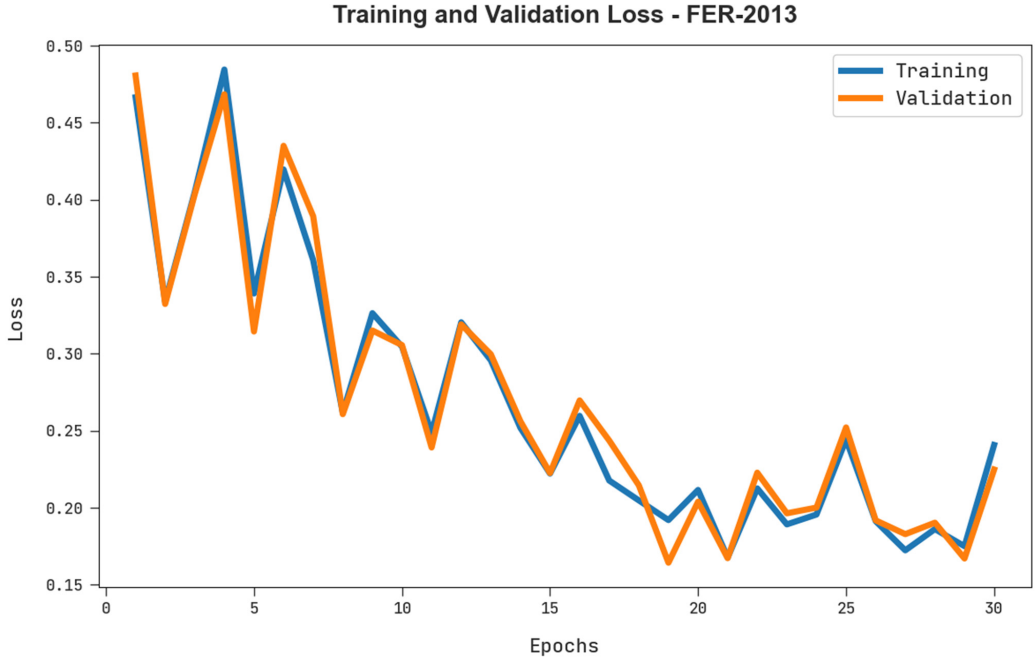


Figure 7. TR_{loss} and VL_{loss} analysis of TLDFER-ADAS system under FER-2013 dataset.

The confusion matrices gained by the TLDFER-ADAS approach on the CK+ dataset is conveyed in Figure 8. The outcomes inferred the better emotion recognition outcome of the TLDFER-ADAS method. It is obvious that the TLDFER-ADAS technique has capably recognized seven distinct classes.

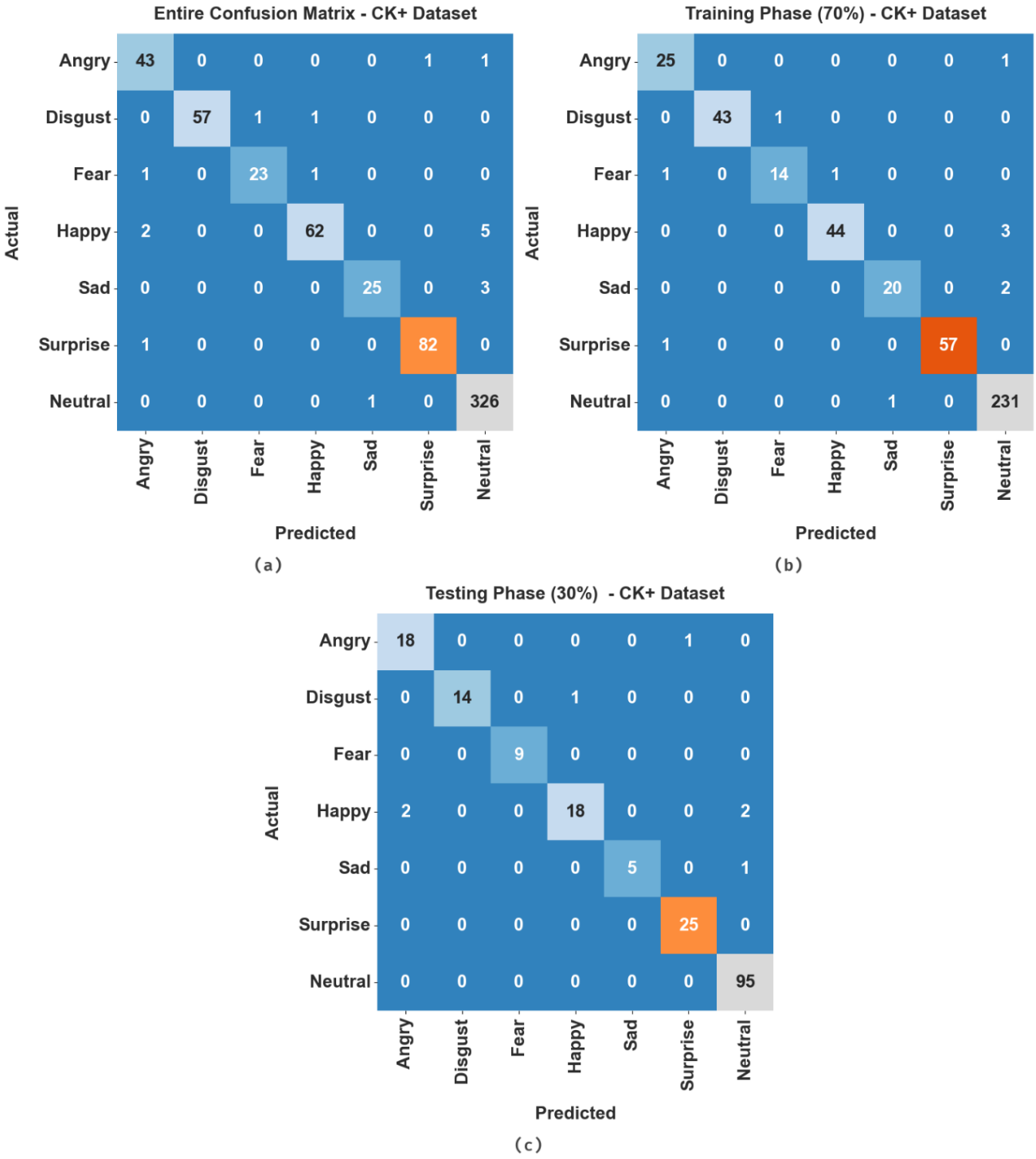


Figure 8. Confusion matrices of TLDFER-ADAS system under CK+ dataset: (**a**) entire database; (**b**) 70% of TR database; and (**c**) 30% of TS database.

Table 3 and Figure 9 reveal thorough results of the TLDFER-ADAS approach on the test CK+ dataset. These results represented the improved emotion classification outcomes of the TLDFER-ADAS model. For instance, on entire dataset, the TLDFER-ADAS technique has accomplished average $accu_y$ of 99.19%, $prec_n$ of 96.64%, $reca_l$ of 94.54%, F_{score} of 95.53%, and AUC_{score} of 96.95%. Likewise, on 70% of TR dataset, the TLDFER-ADAS system has depicted average $accu_y$ of 99.29%, $prec_n$ of 96.63%, $reca_l$ of 94.82%, F_{score} of 95.69%, and AUC_{score} of 97.12%. Lastly, on 30% of TS dataset, the TLDFER-ADAS approach has exhibited average $accu_y$ of 98.95%, $prec_n$ of 96.83%, $reca_l$ of 93.32%, F_{score} of 94.87%, and AUC_{score} of 96.27%.

Table 3. Result analysis of TLDFER-ADAS system with distinct classes under CK+ dataset.

CK+ Dataset					
Class	**Accuracy**	**Precision**	**Recall**	**F-Score**	**AUC Score**
Entire Dataset					
Angry	99.06	91.49	95.56	93.48	97.44
Disgust	99.69	100.00	96.61	98.28	98.31
Fear	99.53	95.83	92.00	93.88	95.92
Happy	98.58	96.88	89.86	93.23	94.75
Sad	99.37	96.15	89.29	92.59	94.56
Surprise	99.69	98.80	98.80	98.80	99.31
Neutral	98.43	97.31	99.69	98.49	98.39
Average	**99.19**	**96.64**	**94.54**	**95.53**	**96.95**
Training Phase (70%)					
Angry	99.33	92.59	96.15	94.34	97.84
Disgust	99.78	100.00	97.73	98.85	98.86
Fear	99.33	93.33	87.50	90.32	93.63
Happy	99.10	97.78	93.62	95.65	96.68
Sad	99.33	95.24	90.91	93.02	95.34
Surprise	99.78	100.00	98.28	99.13	99.14
Neutral	98.43	97.47	99.57	98.51	98.38
Average	**99.29**	**96.63**	**94.82**	**95.69**	**97.12**
Testing Phase (30%)					
Angry	98.43	90.00	94.74	92.31	96.79
Disgust	99.48	100.00	93.33	96.55	96.67
Fear	100.00	100.00	100.00	100.00	100.00
Happy	97.38	94.74	81.82	87.80	90.61
Sad	99.48	100.00	83.33	90.91	91.67
Surprise	99.48	96.15	100.00	98.04	99.70
Neutral	98.43	96.94	100.00	98.45	98.44
Average	**98.95**	**96.83**	**93.32**	**94.87**	**96.27**

The TR_{acc} and VL_{acc} acquired by the TLDFER-ADAS methodology under CK+ database is displayed in Figure 10. The simulation result stated that the TLDFER-ADAS algorithm has gained higher values of TR_{acc} and VL_{acc}. In certain instances, the VL_{acc} appeared better than that of TR_{acc}.

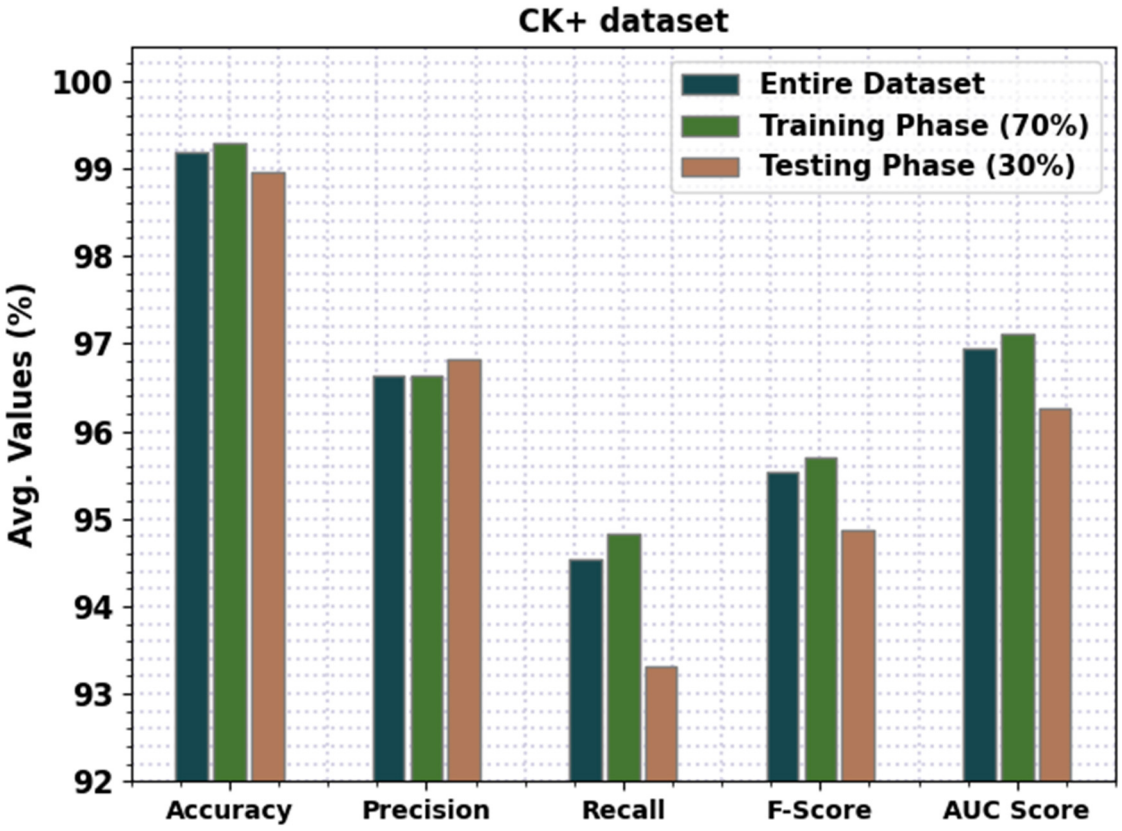


Figure 9. Average analysis of TLDFER-ADAS system under CK+ dataset.

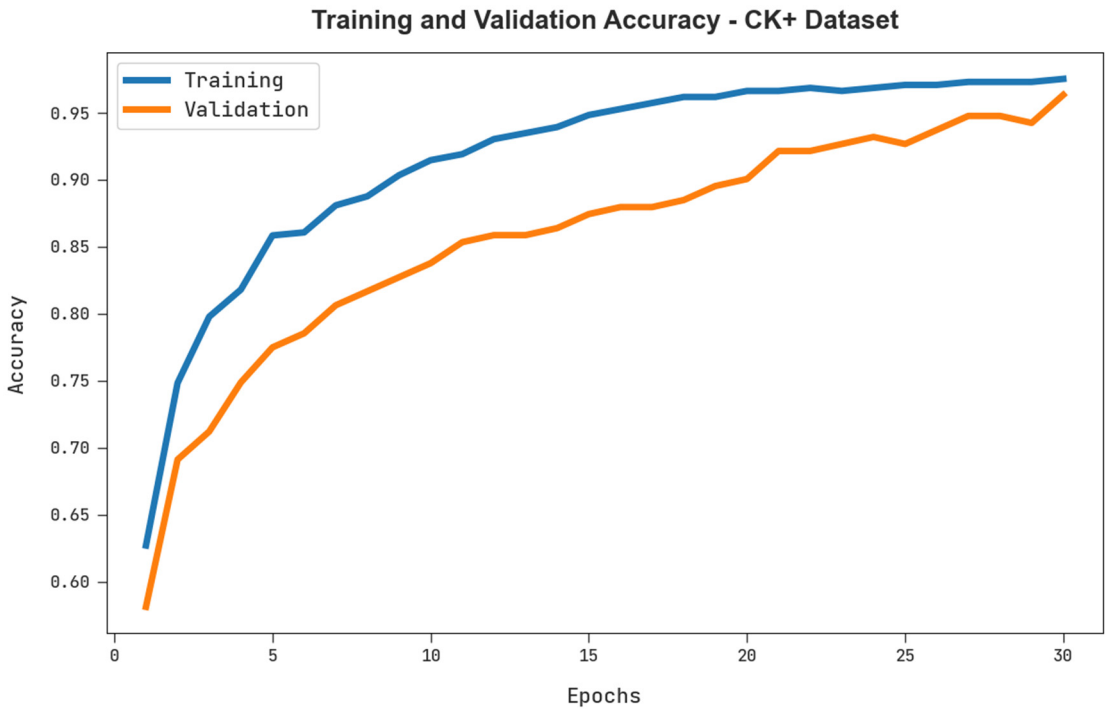


Figure 10. TR_{acc} and VL_{acc} analysis of TLDFER-ADAS system under CK+ dataset.

The TR_{loss} and VL_{loss} attained by the TLDFER-ADAS method under CK+ database are portrayed in Figure 11. The simulation results outperformed that of the TLDFER-ADAS methodology and achieved decreased values of TR_{loss} and VL_{loss}. In particular, the VL_{loss} is less than TR_{loss}.

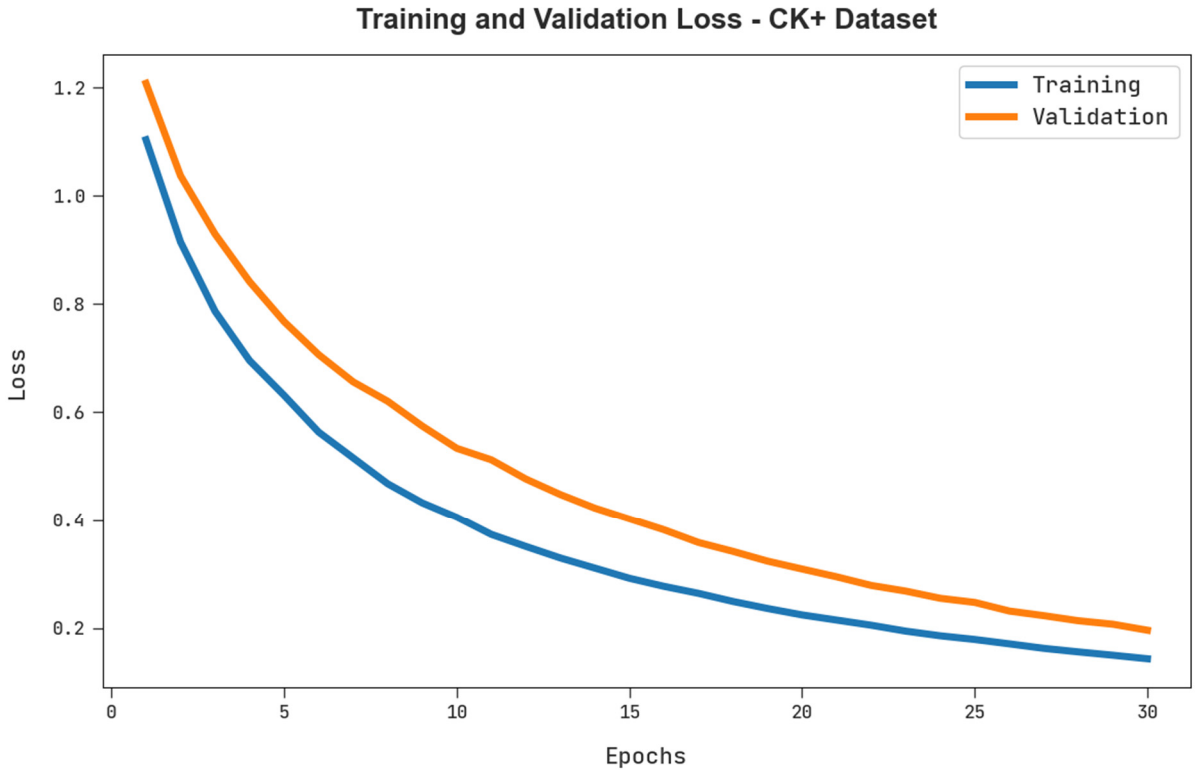


Figure 11. TR_{loss} and VL_{loss} analysis of TLDFER-ADAS system under CK+ dataset.

An overall $accu_y$ comparison of the TLDFER-ADAS model with other models on two datasets is given in Table 4. In Figure 12, brief emotion recognition results of the TLDFER-ADAS model with existing approaches are provided on FER-2013 dataset. The results implied that the improved FRCNN model has demonstrated the least improved results whereas the DNN and PGC models have exhibited certainly enhanced performance. On the other hand, the Asm-VM and FPD-NN models have demonstrated reasonable results with $accu_y$ of 97.91% and 97.03%, respectively. Finally, the TLDFER-ADAS model outperformed the other models with maximum $accu_y$ of 99.31%.

Table 4. Accuracy analysis of TLDFER-ADAS approach with other algorithms under two datasets.

Accuracy (%)		
Methods	**FER-2013**	**CK+ (Last Frame)**
TLDFER-ADAS	99.31	99.29
DNN	95.65	96.03
Asm-SVM	97.91	94.29
PGC	96.39	95.81
FPD-NN	97.03	96.22
Improved FRCNN	94.78	94.29

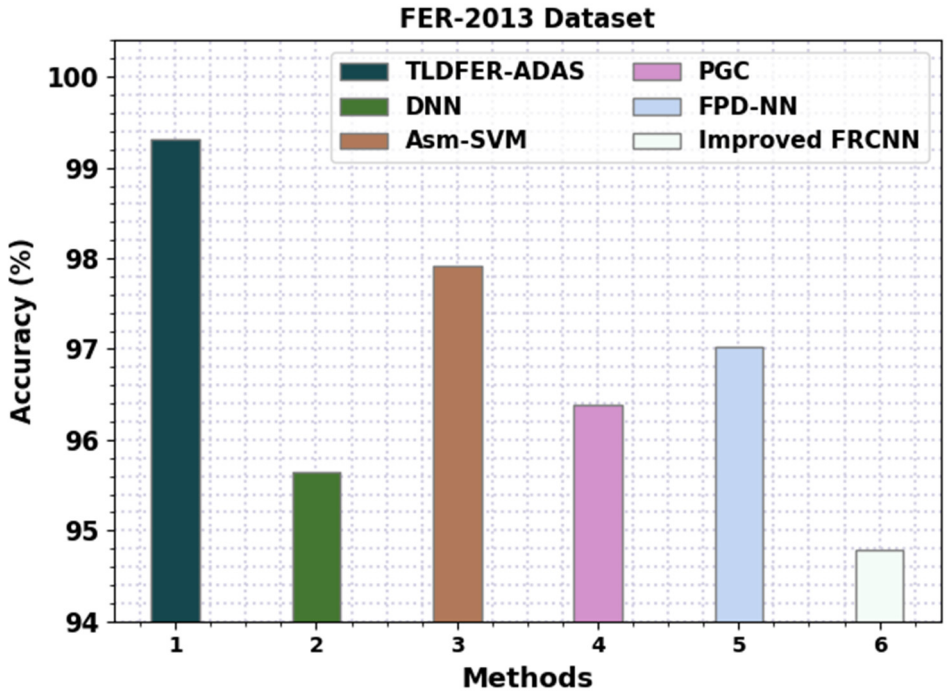


Figure 12. $Accu_y$ analysis of TLDFER-ADAS approach under FER-2013 dataset.

In Figure 13, a detailed comparative study of the TLDFER-ADAS model with existing approaches is given on CK+ dataset. The obtained values indicated that the improved FR-CNN model has exhibited minimum outcomes whereas the Asm-VM and PGC models have displayed surely improved performance. At the same time, the DNN and FPD-NN models have confirmed equitable results with $accu_y$ of 96.03% and 96.22% respectively. Finally, the TLDFER-ADAS model has outpaced the other models with maximum $accu_y$ of 99.29%.

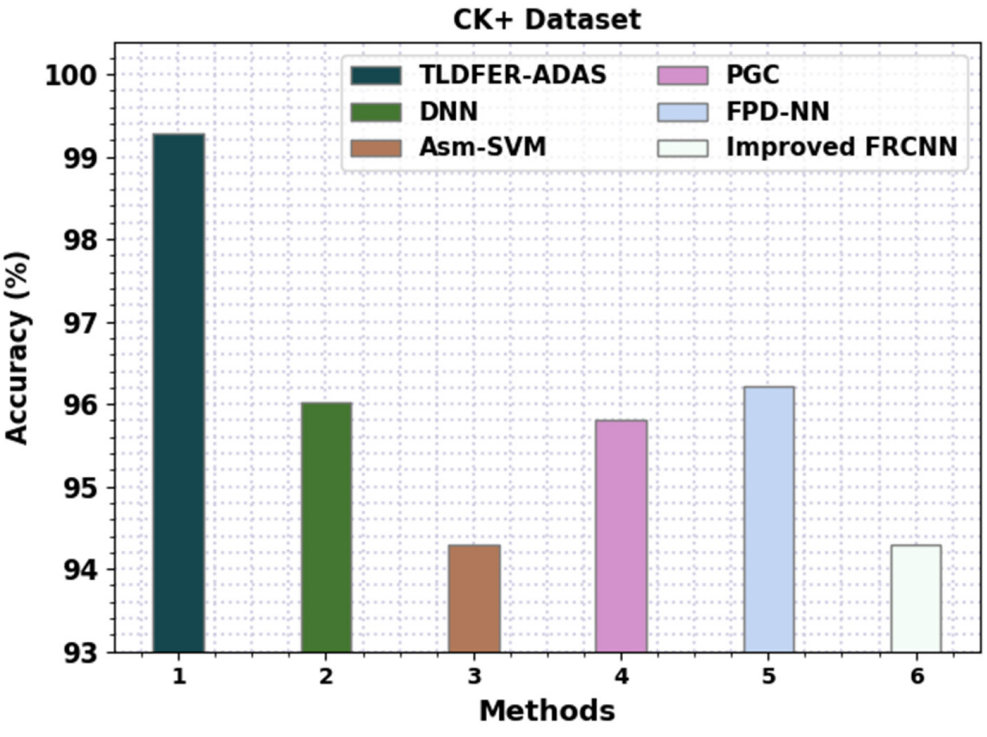


Figure 13. $Accu_y$ analysis of TLDFER-ADAS approach under CK+ dataset.

As shown in figures, the proposed model has shown enhanced performance on both datasets. For instance, on FER-2013 dataset, the proposed model has obtained maximum accuracy of 99.31% whereas the existing Asm-SVM model has attained accuracy of 97.91%, indicating an improvement of 1.4%. Similarly, on CK+ dataset, the proposed model has obtained maximum accuracy of 99.29% whereas the existing FPD-NN model has attained accuracy of 96.22%, indicating an improvement of 3.07%. These results confirmed the enhanced performance of the TLDFER-ADAS model over other models. The enhanced performance of the proposed model is due to the parameter tuning process.

5. Conclusions

In this study, we have introduced a new TLDFER-ADAS technique for smart city environment. The TLDFER-ADAS technique aids in proper driving and determines the emotion of drivers, which assures road safety. The TLDFER-ADAS technique originally performed contrast enhancement procedures for enhancing the image quality. In the TLDFER-ADAS technique, the Xception model was applied to produce feature vectors. For driver emotion classification, the MRFO with QDNN model was exploited in this work. The experimental result analysis of the TLDFER-ADAS technique was performed on a benchmark dataset. A widespread experimental analysis demonstrated the enhancements of the TLDFER-ADAS technique over other techniques.

In future, the emotion detection results can be improvised using the ensemble deep learning classifiers with optimal hyperparameter tuning strategies. In addition, the proposed model can be tested on large-scale real-time datasets in future. Moreover, the proposed model can be extended to the use of hybrid metaheuristic optimizers.

Author Contributions: Conceptualization, A.M.H.; Data curation, M.M.; Formal analysis, A.A.A. and M.I.E.; Funding acquisition, D.H.E.; Methodology, A.M.H. and M.I.E.; Project administration, A.M.H.; Resources, A.A.A. and A.S.Z.; Software, M.O., A.A.A. and A.S.Z.; Supervision, S.S.A.; Validation, A.M.H., S.S.A., M.M. and A.S.Z.; Visualization, M.O.; Writing–original draft, D.H.E., S.S.A. and M.M.; Writing–review & editing, M.O. and M.I.E. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Small Groups Project under grant number (168/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R238), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4210118DSR52).

Institutional Review Board Statement: This article does not contain any studies with human participants performed by any of the authors.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated during the current study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kandeel, A.A.; Abbas, H.M.; Hassanein, H.S. Explainable model selection of a convolutional neural network for driver's facial emotion identification. In Proceedings of the International Conference on Pattern Recognition, Milan, Italy, 10–15 January 2021; Springer: Cham, Germany, 2021; pp. 699–713.
2. Gera, D.; Balasubramanian, S.; Jami, A. CERN: Compact facial expression recognition net. *Pattern Recognit. Lett.* **2022**, *155*, 9–18. [CrossRef]
3. Li, W.; Cui, Y.; Ma, Y.; Chen, X.; Li, G.; Zeng, G.; Guo, G.; Cao, D. A Spontaneous Driver Emotion Facial Expression (Defe) Dataset for Intelligent Vehicles: Emotions Triggered by Video-Audio Clips in Driving Scenarios. In *IEEE Transactions on Affective Computing*; IEEE: Piscataway Township, NJ, USA, 2021.
4. Bodapati, J.D.; Naik, D.S.; Suvarna, B.; Naralasetti, V. A Deep Learning Framework with Cross Pooled Soft Attention for Facial Expression Recognition. *J. Inst. Eng. Ser. B* **2022**, *103*, 1395–1405. [CrossRef]

5. Deng, W.; Wu, R. Real-time driver-drowsiness detection system using facial features. *IEEE Access* **2019**, *7*, 118727–118738. [CrossRef]
6. Dias, W.; Andaló, F.; Padilha, R.; Bertocco, G.; Almeida, W.; Costa, P.; Rocha, A. Cross-dataset emotion recognition from facial expressions through convolutional neural networks. *J. Vis. Commun. Image Represent.* **2022**, *82*, 103395. [CrossRef]
7. Yan, K.; Zheng, W.; Zhang, T.; Zong, Y.; Cui, Z. Cross-database non-frontal facial expression recognition based on transductive deep transfer learning. *arXiv* **2018**, arXiv:1811.12774.
8. Jabbar, R.; Al-Khalifa, K.; Kharbeche, M.; Alhajyaseen, W.; Jafari, M.; Jiang, S. Real-time driver drowsiness detection for android application using deep neural networks techniques. *Procedia Comput. Sci.* **2018**, *130*, 400–407. [CrossRef]
9. Hung, J.C.; Chang, J.W. Multi-level transfer learning for improving the performance of deep neural networks: Theory and practice from the tasks of facial emotion recognition and named entity recognition. *Appl. Soft Comput.* **2021**, *109*, 107491. [CrossRef]
10. Alzubi, J.A.; Jain, R.; Alzubi, O.; Thareja, A.; Upadhyay, Y. Distracted driver detection using compressed energy efficient convolutional neural network. *J. Intell. Fuzzy Syst.* **2022**, *42*, 1253–1265. [CrossRef]
11. Yang, H.; Zhang, Z.; Yin, L. Identity-adaptive facial expression recognition through expression regeneration using conditional generative adversarial networks. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 294–301.
12. Khanzada, A.; Bai, C.; Celepcikay, F.T. Facial expression recognition with deep learning. *arXiv* **2020**, arXiv:2004.11823.
13. Hossain, S.; Umer, S.; Asari, V.; Rout, R.K. A unified framework of deep learning-based facial expression recognition system for diversified applications. *Appl. Sci.* **2021**, *11*, 9174. [CrossRef]
14. Macalisang, J.R.; Alon, A.S.; Jardiniano, M.F.; Evangelista, D.C.P.; Castro, J.C.; Tria, M.L. Drive-Awake: A YOLOv3 Machine Vision Inference Approach of Eyes Closure for Drowsy Driving Detection. In Proceedings of the 2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 13–15 September 2021; IEEE: Piscataway Township, NJ, USA, 2021; pp. 1–5.
15. Rescigno, M.; Spezialetti, M.; Rossi, S. Personalized models for facial emotion recognition through transfer learning. *Multimed. Tools Appl.* **2020**, *79*, 35811–35828. [CrossRef]
16. Hou, M.; Wang, M.; Zhao, W.; Ni, Q.; Cai, Z.; Kong, X. A lightweight framework for abnormal driving behavior detection. *Comput. Commun.* **2022**, *184*, 128–136. [CrossRef]
17. Shao, J.; Qian, Y. Three convolutional neural network models for facial expression recognition in the wild. *Neurocomputing* **2019**, *355*, 82–92. [CrossRef]
18. Sini, J.; Marceddu, A.C.; Violante, M.; Dessì, R. Passengers' emotions recognition to improve social acceptance of autonomous driving vehicles. In *Progresses in Artificial Intelligence and Neural Systems*; Springer: Singapore, 2021; pp. 25–32.
19. Naqvi, R.A.; Arsalan, M.; Rehman, A.; Rehman, A.U.; Loh, W.K.; Paul, A. Deep learning-based drivers emotion classification system in time series data for remote applications. *Remote Sens.* **2020**, *12*, 587. [CrossRef]
20. Paikrao, P.; Mukherjee, A.; Jain, D.K.; Chatterjee, P.; Alnumay, W. Smart emotion recognition framework: A secured IOVT perspective. In *IEEE Consumer Electronics Magazine*; IEEE: Piscataway Township, NJ, USA, 2021.
21. Jeong, M.; Ko, B.C. Driver's facial expression recognition in real-time for safe driving. *Sensors* **2018**, *18*, 4270. [CrossRef] [PubMed]
22. Sukhavasi, S.B.; Sukhavasi, S.B.; Elleithy, K.; El-Sayed, A.; Elleithy, A. A hybrid model for driver emotion detection using feature fusion approach. *Int. J. Environ. Res. Public Health* **2022**, *19*, 3085. [CrossRef]
23. Xiao, H.; Li, W.; Zeng, G.; Wu, Y.; Xue, J.; Zhang, J.; Li, C.; Guo, G. On-Road Driver Emotion Recognition Using Facial Expression. *Appl. Sci.* **2022**, *12*, 807. [CrossRef]
24. Mehendale, N. Facial emotion recognition using convolutional neural networks (FERC). *SN Appl. Sci.* **2020**, *2*, 446. [CrossRef]
25. Oh, G.; Ryu, J.; Jeong, E.; Yang, J.H.; Hwang, S.; Lee, S.; Lim, S. Drer: Deep learning–based driver's real emotion recognizer. *Sensors* **2021**, *21*, 2166. [CrossRef]
26. Li, W.; Zeng, G.; Zhang, J.; Xu, Y.; Xing, Y.; Zhou, R.; Guo, G.; Shen, Y.; Cao, D.; Wang, F.Y. CogEmoNet: A Cognitive-Feature-Augmented Driver Emotion Recognition Model for Smart Cockpit. *IEEE Trans. Comput. Soc. Syst.* **2021**, *9*, 667–678. [CrossRef]
27. Ma, J.; Fan, X.; Yang, S.X.; Zhang, X.; Zhu, X. Contrast limited adaptive histogram equalization-based fusion in YIQ and HSI color spaces for underwater image enhancement. *Int. J. Pattern Recognit. Artif. Intell.* **2018**, *32*, 1854018. [CrossRef]
28. Lo, W.W.; Yang, X.; Wang, Y. An xception convolutional neural network for malware classification with transfer learning. In Proceedings of the 2019 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Canary Island, Spain, 24–26 June 2019; IEEE: Piscataway Township, NJ, USA, 2019; pp. 1–5.
29. Dharmawan, W.; Nambo, H. End-to-End Xception model implementation on Carla Self Driving Car in moderate dense environment. In Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference, Kobe, Japan, 21–23 December 2019; pp. 139–143.
30. Zhang, L.; Wang, M.; Fu, Y.; Ding, Y. A Forest Fire Recognition Method Using UAV Images Based on Transfer Learning. *Forests* **2022**, *13*, 975. [CrossRef]
31. Jeswal, S.K.; Chakraverty, S. Recent developments and applications in quantum neural network: A review. *Arch. Comput. Methods Eng.* **2019**, *26*, 793–807. [CrossRef]
32. Tate, N.; Miyata, Y.; Sakai, S.I.; Nakamura, A.; Shimomura, S.; Nishimura, T.; Kozuka, J.; Ogura, Y.; Tanida, J. Quantitative analysis of nonlinear optical input/output of a quantum-dot network based on the echo state property. *Opt. Express* **2022**, *30*, 14669–14676. [CrossRef]

33. Houssein, E.H.; Zaki, G.N.; Diab, A.A.Z.; Younis, E.M. An efficient Manta Ray Foraging Optimization algorithm for parameter extraction of three-diode photovoltaic model. *Comput. Electr. Eng.* **2021**, *94*, 107304. [CrossRef]
34. Hu, G.; Li, M.; Wang, X.; Wei, G.; Chang, C.T. An enhanced manta ray foraging optimization algorithm for shape optimization of complex CCG-Ball curves. *Knowl. Based Syst.* **2022**, *240*, 108071. [CrossRef]