WILEY | Hindawi

Research Article

Capsule Network-Based Deep Transfer Learning Model for Face Recognition

Keshetti Sreekala,[1] C. Pretty Diana Cyril,[2] S. Neelakandan,[3] Saravanan Chandrasekaran,[4] Ranjan Walia [ID],[5] and Eric Ofori Martinson [ID][6]

[1]*Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Hyderabad 500075, India*
[2]*Department of Computing Technologies, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur 603203, India*
[3]*Department of CSE, RMK Engineering College, Chennai, India*
[4]*Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India*
[5]*Department of Electrical Engineering, Model Institute of Engineering and Technology, Jammu, J&K, India*
[6]*Department of Electronics and Communications Engineering, School of Engineering, All Nations University, Ghana*

Correspondence should be addressed to Eric Ofori Martinson; emartinson@anuc.edu.gh

Face recognition (FR) is a technique for recognizing individuals through the use of face photographs. The FR technology is widely applicable in a variety of fields, including security, biometrics, authentication, law enforcement, smart cards, and surveillance. Recent advances in deep learning (DL) models, particularly convolutional neural networks (CNNs), have demonstrated promising results in the field of FR. CNN models that have been pretrained can be utilized to extract characteristics for effective FR. In this regard, this research introduces the GWOECN-FR approach, a unique grey wolf optimization with an enhanced capsule network-based deep transfer learning model for real-time face recognition. The proposed GWOECN-FR approach is primarily concerned with reliably and rapidly recognizing faces in input photos. Additionally, the GWOECN-FR approach is preprocessed in two steps, namely, data augmentation and noise reduction by bilateral filtering (BF). Additionally, for feature vector extraction, an expanded capsule network (ECN) model can be used. Additionally, grey wolf optimization (GWO) combined with a stacked autoencoder (SAE) model is used to identify and classify faces in images. The GWO algorithm is used to optimize the SAE model's weight and bias settings. The GWOECN-FR technique's performance is validated using a benchmark dataset, and the results are analyzed in a variety of aspects. The GWOECN-FR approach achieved a TST of 0.03 s on the FEI dataset, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet models achieved TSTs of 0.125 s, 0.0051 s, and 0.0062 s, respectively. The experimental results established that the GWOECN-FR technology outperformed more contemporary approaches.

1. Introduction

The face recognition (FR) system is commonly presented with essential facial features including the eyes, nose, and mouth, that is, a nonoccluded face [1]. But a wide variety of circumstances and situations impose people to wear masks on faces that are partially occluded or hidden. This common situation includes laboratories, pandemics, immoderate pollution, or medical operations [2]. For exam-

ple, as per the WHO and Centers for Disease Control and Prevention (CDC), the better method for protecting people from the COVID-19 virus and being infected or preventing the spreading of the disease is practicing social distancing and wearing face masks. Therefore, each country in the world requires people to wear a protective face mask in common places that address the need to understand and investigate FR systems performed with face masks. But performing this safety guiding principle earnestly challenges the current

authentication and security schemes that depend on FR that has already been implemented [3]. Techniques that are at the forefront of lighting and pose research are utilized often all around the world to solve illumination and posture issues. On the basis of this exploration, we are presenting a survey paper that comprises as much literature study as possible for the reader to understand what exactly the variations are that can be caused by variation in illumination and pose, as well as what approaches have been taken up until now to make continuous improvements in existing systems. Mostly, current approaches have been presented to determine whether a face is occluded or not, that is, masked face detection. Even though saving an individual's life is effective, there is a crucial requirement for authenticating persons wearing masks without requiring to expose them. For example, premise access control and immigration points are among various locations in which a subject makes a cooperative presentation to a camera that increases the challenge of face detection since the occluded part is essential for FR and detection [4, 5].

The FR technique comprises four phases that include face alignment, representation, classification (facial feature extraction), and detection [6]. In the FR method, the major problem is the feature representation system utilized for extracting features, with the best approach for a given biometric trait, for representation. Feature extraction is the fundamental step for image classification. Feature extraction means preserving the crucial data that is needed for classification. There are several feature extraction processes that were introduced for usage in a biometric scheme, involving independent component analysis (ICA), principal component analysis (PCA), the histogram method, and local binary patterns (LBP) [7, 8]. In recent times, the convolutional neural network (CNN) displays significant benefits.

Currently, with the development of the DL method, the face detection method accomplishes remarkable outcomes. The CNN method, more commonly known as the deep neural network in computer vision applications, proves a significant benefit of automated visual feature extraction [9]. There are two types of approaches for training CNN for face detection systems, namely, metric learning and classification layer. There are distinct methods for utilizing CNN. Initially, learn the model from scratch. In such cases, the framework of the pretrained method is trained and utilized based on the dataset. Next, transfer learning (TL) uses features from pretrained CNN, where the dataset is larger. At last, CNN is utilized by the TL method to keep the convolution base in its original version and utilize output for feeding the classifier. The pretrained method is utilized as a fixed feature extraction method where the dataset is smaller [10].

This study designs an effective grey wolf optimization with an enhanced capsule network-based deep transfer learning model for real-time face recognition, named the GWOECN-FR technique. The proposed GWOECN-FR technique performs data augmentation and bilateral filtering- (BF-) based noise elimination at the preprocessing step. In addition, the enhanced capsule network (ECN) model can be utilized for extracting feature vectors. In the current investigation, enhanced capsule networks (ECN) are utilized in order to construct an appropriate diagnostic system. The fragmented pixel set of the X-ray image is referred to as a set of nerve cells in ECN. The capsule is the focus of this analysis. For the purpose of the X-ray image, the system utilized a pixel vector as an actuation vector. This pixel vector was encircled by an active capsule and may represent a certain category, such as healthy or COVID-19. Within this example, the capsule output and the coupling coefficient have been multiplied by the capsule routing within a layer. The value of the coupling coefficient is determined by the resistance of the parent capsule when it comes to routing. The low-level COVID-19 diagnosis has been determined based on the routing-by-agreement method, and the high-level capsule activation was responsible for making this determination. Furthermore, grey wolf optimization (GWO) with a stacked autoencoder (SAE) model is applied for the identification and classification of faces where the weight and bias values of the SAE model are chosen by the GWO algorithm. Mirjalili et al. developed the grey wolf optimizer in 2014 as a revolutionary heuristic swarm intelligence optimization technique. As the apex predator in the food chain, the wolf has a remarkable ability to capture prey. GWO is a novel metaheuristic optimization tool. Its guiding premise is to mimic the cooperative hunting behavior of grey wolves in nature. In terms of model structure, GWO differs from others. In order to showcase the supremacy of the GWOECN-FR technique, a comprehensive result analysis is carried out on a benchmark dataset.

2. Related Works

Gwyn et al. [1] provided a comprehensive review of many advanced DL-based facial detection techniques to define accuracy and other metrics that are very efficient. In this work, VGG-16 and VGG-19 show the maximum level of image detection performance and F1 score. Zhu and Jiang [11] aimed to enhance the current face detection method, studied the face detection approach driven by big data, and presented a DL multifeature fusion face detection method driven by big data. The study employs the LBP (local binary pattern) approach for extracting the texture feature of the face, and that is incorporated with the global feature extracted through 2DPCA to multifeature fusion; thus, the fused feature could consider local and global features, which have good detection performance.

Al-Waisy et al. [12] proposed an architecture based on merging the advantage of the local handcrafted feature descriptor with the DBN for addressing the face detection problem in unrestrained conditions. Initially, a novel multimodal local feature extraction method based on merging the advantage of the curvelet transform with the fractal dimension has been presented and called the curvelet-fractal method. Mao et al. [13] used the differentially private method to enable the privacy-preserving edge-based training of the DNN face detection method. In the training process, DNN has split among the edge server and the user device so that model parameters and private data are secured, with a smaller cost of local computation.

Sharma and Kumar [14] developed a three-dimensional face reconstruction and a sequential DL-based architecture for face detection. It employs the reflection principle to generate the reconstructed points in 3D with the midface plane. From the reconstructed face, a sequential DL architecture is designed for recognizing the person, gender, emotion, and occlusion. The presented method uses the concept of triplet loss training, VAE, and BiLSMT. Anand et al. [15] focused on using one of the advanced ML methods in face detection to achieve maximum performance. Then, we generated our own dataset and trained it on the GoogLeNet (inception) DL method with the Caffe and Nvidia DIGITS architecture.

3. The Proposed Model

In this study, a new GWOECN-FR technique has been developed for the rapid and prompt identification of facial images. The presented GWOECN-FR technique initially performs data augmentation and BF-based noise elimination to preprocess the facial input images. Afterward, the ECN model is applied to produce a useful set of feature vectors. Besides, the SAE model is employed to recognize the faces and the performance of the SAE model can be improved by adjusting its parameters using the GWO algorithm. Figure 1 illustrates the overall process of the GWOECN-FR technique.

3.1. Data Augmentation and Noise Removal. Since DL models require a large number of training instances, data augmentation becomes essential. In this work, geometric transformation, filtering, and brightness operations are performed to augment the images. Image rotation, zooming, translation, brightness, and filtering operations are carried out to increase the number of variations of the images. Then, the BF technique can be employed to eradicate the existence of noise involved in it.

Assume that F refers to the multichannel images, and consider that W exists as slide windows of a fixed size $n \times n$. Let the pixel from W signify the Cartesian coordinates, denoted by $u = (u_1, u_2) \in Y^2$, the place of pixels F_u from W, where $Y = \{0, 1, \cdots, n-1\}$ is endowed with the standard order. The BF exchanges the central pixels of all the filter windows by the weighted average of their neighbor color pixels. The weighted function was planned for smoothing from the area of related colors but keeping edges intact by the heavily weighted individual's pixel which is both spatially close and photometrically related to the central pixels [16]. Represent by $\|\cdot\|_2$ the Euclidean norms and F_u the central pixels of concern. Afterward, the weighted $\mathscr{W}(F_u, F_v)$ equivalent to some pixel F_v in accordance with F_u refers to the product of 2 elements, 1 spatial and 1 photometrical,

$$\mathscr{W}(F_u, F_v) = \mathscr{W}_s(F_u, F_v)\mathscr{W}_p(F_u, F_v), \tag{1}$$

but the spatial element $\mathscr{W}_s(F_u, F_v)$ is offered as

$$\mathscr{W}_s(F_u, F_v) = e^{-\|u-v\|_2^2/2\sigma_s^2}, \tag{2}$$

and the photometrical element $\mathscr{W}_p(F_u, F_v)$ is provided by

$$\mathscr{W}_p(F_u, F_v) = e^{-\Delta E_{Lab}(F_u, F_v)^2/2\sigma_p^2}, \tag{3}$$

where $\Delta E_{Lab} = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{1/2}$ demonstrates the perceptual color error from the $L^* a^* b^*$ color space, and $\sigma_s, \sigma_p > 0$. The color vector outcome $\widetilde{F_u}$ of filtering is calculated utilizing the normalization weight, so it can be provided as

$$\bar{F}_u = \frac{\sum_{F_V \in w} \mathscr{W}(F_u, F_v) F_v}{\sum_{F_V \in w} \mathscr{W}(F_u, F_v)}. \tag{4}$$

The $\mathscr{W}_s$ weighted purpose reduces as the spatial distance from the image among u and v improves, and the $\mathscr{W}_p$ weighted function reduces as the perceptual color variance among the color vector enhances. The spatial element reduces the control of the furthest pixel decreasing blurring, but the photometric element decreases the control of the individual's pixels that are perceptually distinct in accordance with one under process. During this approach, only perceptually related regions of pixels were averaged together and the sharpness of edges is maintained. The parameters σ_s and σ_p were utilized for adjusting the control of the spatial and photometric elements correspondingly. A rough threshold is assumed to identify pixels appropriately closer or related to the central one. Notice that if $\sigma_p \longrightarrow \infty$, the BF methods include Gaussian filtering, and if $\sigma_s \longrightarrow \infty$, the filter methods range filters without spatial notion. During the case, if $\sigma_p \longrightarrow \infty$ and $\sigma_s \longrightarrow \infty$ are combined, the BF performs as AMF.

3.2. Feature Extraction Using the ECN Model. During the feature extraction process, the preprocessed facial image is passed into the ECN model to generate feature vectors. The enhanced capsule network (ECN) system has been employed. In our method, the split pixel set of the images is labeled as a group of nerve cells respective to the capsule [17]. The pixel vectors are utilized as an activation vector enclosed through an active capsule; also, there might be a certain class, namely, tumor or healthy, for an image-related pixel vector segmentation representing the overall length. The capsule routing in a layer is implemented through the multiplication of the coupling coefficient (CC) and capsule output. The values of CC can be defined as the resistance of the parent capsule for routing. The lower-level tumor diagnoses are determined as higher-level capsule activation via a top-down feedback method that is named "routing-by-agreement" [17]. Assume $Y_i \in [healthy, tumor]$ as output capsule i, and we_{ij} denotes the weight matrix as follows:

$$\hat{y}_{(ji)} = we_{ij}y_i, \tag{5}$$

where $\hat{y}_{(ji)}$ determines the detection vector that diagnoses the output parent capsule j with capsule i, and the pixel
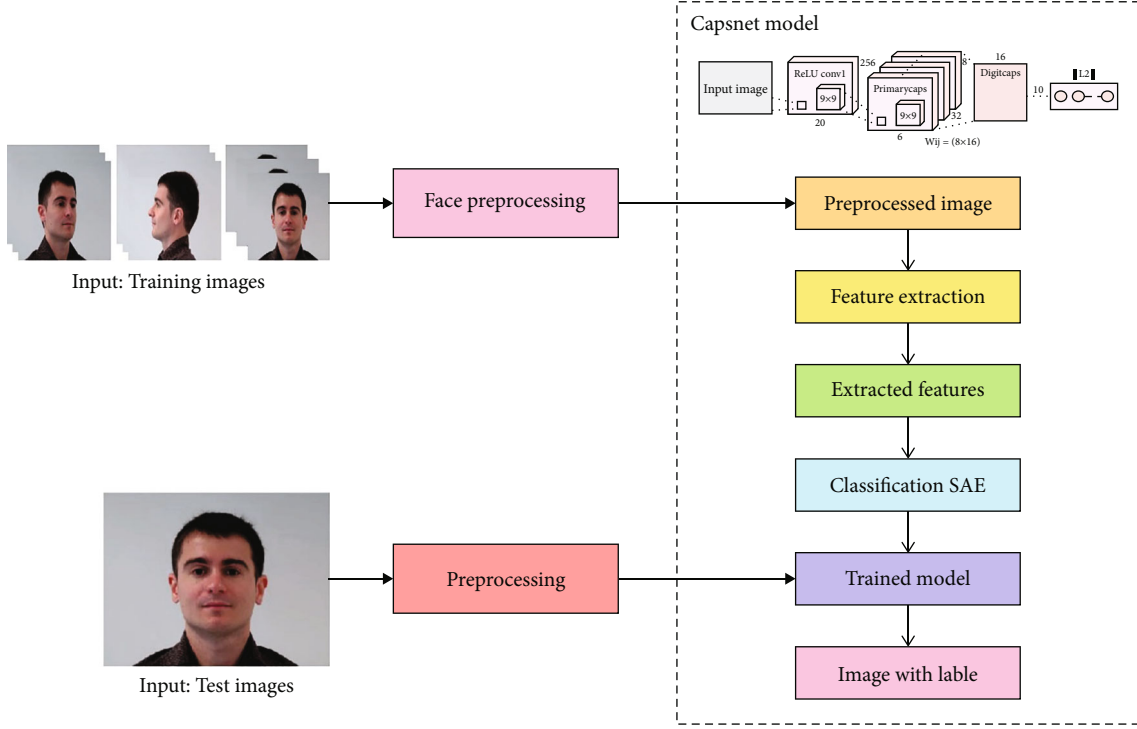
FIGURE 1: Overall process of the GWOECN-FR technique.

range has been applied for evaluating the weight quantity. The quantity for the weight was enhanced as long as the value has been reduced or the pixel is possibly included in the positive group. The softmax process is utilized by the previous layer capsule, and the possible parent capsule as a coefficient is encoded as c_{ij} where main logits b_{ij} display the log preceding possibility of routing capsule i in the prior layer to capsule j in the subsequent layer. In general, the "routing-by-agreement" method was implemented by logits of the capsule in each layer:

$$c_{ij} = \frac{e^{b_{ij}}}{\sum_i e^{b_{ij}}}. \tag{6}$$

The preceding layer illustrates an essential component in the computation of the input of parent capsule j, i.e., accomplished by

$$s_j = \sum_i c_{ij} \widehat{y}_{(j|i)}. \tag{7}$$

The compression value of the pixel vector was determined within $[0, 1]$ using a nonlinear process named squashing. The computational operation is expressed by

$$va_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \times \frac{s_j}{\varepsilon + \|s_j\|^2}, \tag{8}$$

where $\varepsilon = 10^{-7}$. And the following layer capsule was accomplished by

$$a_{ij} = va_j \times \widehat{y}_{(ji)}. \tag{9}$$

The whole capsule classification is considered the margin loss (Loss_k) in the class capsule k for the capsule network according to the loss:

$$\begin{aligned}\text{Loss}_k = T_k \ \max \ (0, m^+ - \|va_k\|)^2 \\ + \lambda(1 - T_k) \max \ (0, \|va_k\| - m^-)^2,\end{aligned} \tag{10}$$

where T_k denotes the instant existence in class capsule k, and λ, m^-, and m^+ determine hyperparameter assistance. The training of ECN can be determined according to six hundred iterations according to the Adam optimizer for the optimum of the hyperparameter using the amount of learning rates. Adam is a stochastic gradient descent replacement optimization technique for training deep learning models. Adam combines the finest features of the AdaGrad and RMSProp methods to provide an optimization technique that can handle sparse gradients on noisy issues. The Adam optimizer combines two gradient descent methodologies: Momentum. This algorithm is used to accelerate the gradient descent technique by taking the "exponentially weighted average" of the gradients into account. Using averages causes the algorithm to converge faster to the minima. The Adam optimizer produces better results than other optimization algorithms, takes less time to compute, and requires fewer

parameters to tune. Because of this, Adam is suggested as the default optimizer for the majority of applications.

3.3. Image Classification. At the final stage, the SAE model receives the feature vectors as input to allot proper class labels to the facial images. AE is a type of unsupervised learning infrastructure which retains 3 states such as input, hidden, and output states [17]. An autoencoder is an unsupervised learning strategy for neural networks that learns efficient data representations (encoding) by training the network to disregard signal "noise." Autoencoders can be used for image denoising, image compression, and, in some situations, picture data synthesis. An autoencoder is made up of three major components: an encoder, a code, and a decoder. The initial data is converted into a coded result, and the network's successive layers extend it into a finished output. A "denoising" autoencoder can help you understand autoencoders. The denoising autoencoder refines the output by combining original and noisy input. Autoencoders are useful in image processing, classification, and other elements of machine learning. The procedure of trained AE has 2 parts, encoded and decoded. The encoded part was utilized to map the input data to hidden representations, and the decoded part was signified as recreating input information in the hidden representations. To provide the unlabeled input dataset $\{x_n\}_{n=1}^{N}$, where $x_n \in R^{m\times1}$, h_n implies the hidden encoded vector computed in x_n, and $\widehat{x}_n$ stands for the decoded vector of the resultant state. Therefore, the encoding method is as follows:

$$h_n = f(W_1 x_n + b_1), \tag{11}$$

where f signifies the encoded functions, W_1 refers to the weighted matrix of encoding, and b_1 implies the bias vectors.

The decoded procedure can be determined as

$$\widehat{x}_n = g(W_2 h_n + b_2), \tag{12}$$

where g denotes the decoded functions, W_2 demonstrated the weighted matrix of decoding, and b_2 implies the bias vectors. The parameter set of AE was the optimization to minimize the reconstruction error:

$$\varnothing(\Theta) = \arg\ \min_{\theta,\theta'}\frac{1}{n}\sum_{i=1}^{n}L(x^i, \widehat{x}^i), \tag{13}$$

where L stands for a loss function $L(x, \widehat{x}) = \|x - \widehat{x}\|^2$.

The infrastructure of SAE is stacked n AEs to n hidden states with an unsupervised statewise learning technique and next is fine-tuned by a supervised approach. Therefore, the SAE-based technique is separated into 3 phases:

(i) Train the primary AE by input information and attain the learn feature vectors

(ii) The feature vector of the previous state was utilized as input to the next stage, and this process was repeated till the training ends
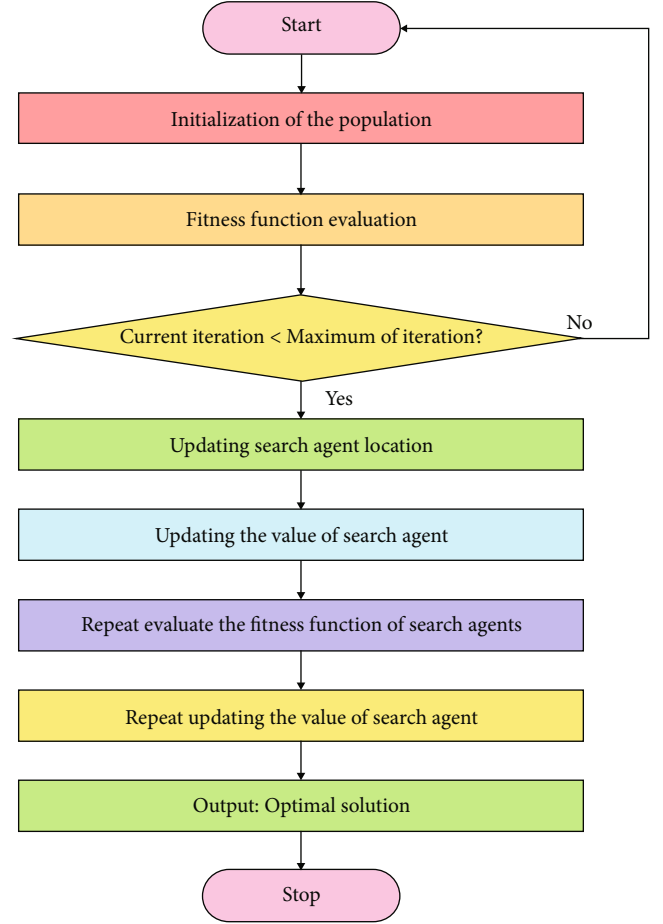

FIGURE 2: GWO flowchart.

Population initialization: grey wolves $X_i(i = 1, 2, \cdots, n)$
Parameter initialization: a, A, and C
Determine fitness values of all searching agents
X_α = optimal searching agent
X_β = second optimal searching agent
X_δ = third optimal searching agent
While (t < Max_number_iterations)
   For every searching agent
     Upgrade the location of the present searching agent
   End for
   Upgrade a, A, and C
   Determine fitness values of all searching agents
   Upgrade X_α, X_β, and X_δ
   Increment t
End while
Return X_α

ALGORITHM 1: Pseudocode of the GWO algorithm.

(iii) Afterward, each hidden state is trained, and the BP technique was utilized for minimizing the cost function and upgrading the weight with a labeled trained set for achieving fine-tuning

Figure 3: Sample images.

3.4. Parameter Tuning Using the GWO Algorithm. For optimally adjusting the parameters (such as weight and bias) involved in the SAE model, the GWO algorithm is applied to it. Recently, a novel SI-optimized technique is called GWO established by Mirjalili et al. [18]. In fact, it can be an original method which accelerates the social and hunting hierarchy of GW by default. For developing the social performance of GW, it can be categorized into 4 states such as α, β, δ, and ω. αconsidered the optimum solutions executed by β and δ, correspondingly, and the residual solution derived inω.The 3 primary fittest wolves namedα, β, and δ neighboring the prey supportω for identifying the food from the challenging region. During the surrounding step, wolves increase the place of β or δ as illustrated in

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_{\mathrm{p}(t)} - \vec{X}(t) \right|, \tag{14}$$

$$\vec{X}(t+1) = \vec{X}_{\mathrm{p}(t)} - \vec{A} \cdot \vec{D}, \tag{15}$$

where t refers to the present iteration, $\vec{X}_{\mathrm{p}(t)}$ denotes the present place of prey, and $\vec{X}(t)$ signifies the current place of wolves. $\vec{D}$ implies the distance between wolves and prey, and coefficient vectors $\vec{A}$ and $\vec{C}$ are resultant from the mathematical process as illustrated as follows [19]. Figure 2 showcases the flowchart of the GWO technique and pseudocode of GWO discussed in Algorithm 1.

$$\begin{aligned} \vec{A} &= 2\vec{a}\,\vec{r_1} - \vec{a}, \\ \vec{C} &= 2\,\vec{r_2}, \end{aligned} \tag{16}$$

where $\vec{r_1}$ and $\vec{r_2}$ imply the 2 vectors created in zero and one from an arbitrary fashion, and the element of $\vec{a}$ has decreased linearly from two to zero to all the iterations. At this point, α, β, and δ define the location nearer to the place

of prey. During the case of hunting, the top 3 were optimum solutions and residual wolves ω are appropriate for replacing the fundamental of the 3 initial optimum wolves. The place of wolves was upgraded on the fundamental of

$$\vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X} - \vec{X} \right|, \tag{17}$$

$$\vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X} - \vec{X} \right|, \tag{18}$$

$$\vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X} - \vec{X} \right|, \tag{19}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_2 \cdot \left(\vec{D}_\alpha \right), \tag{20}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \left(\vec{D}_\beta \right), \tag{21}$$

$$\vec{X}_3 = \vec{X} - \vec{A} \cdot \left(\vec{D}_\delta \right), \tag{22}$$

$$\vec{X}(t+1) = \frac{\vec{X_1} + \vec{X} + \vec{X_3}}{3}, \tag{23}$$

where $\vec{X}_\alpha$ implies the place of α; $\vec{X}_\beta$ defines the place of β; $\vec{X_\delta}$ stands for the place of δ; $\vec{X}$ indicates the place of existing solutions; $\vec{C_1}$, $\vec{C_2}$, and $\vec{C_3}$ signify the vectors created from an arbitrary fashion. Here, $\vec{A}_1$, $\vec{A}_2$, and $\vec{A}_3$ are demonstrated as arbitrary vectors, and t denotes the amount of rounds. The step size of ωwolves was executed; then,α, β, and δ are illustrated as in Equations (17)–(19), correspondingly. Afterward, the resultant places of ωwolves are evaluated on the fundamental of Equations (20)–(23).

The GWO approach derives a Feedback Framework (FF) for obtaining higher classification performance. It defines a positive integer for representing the optimum efficiency of the candidate solution. During this study, the minimized
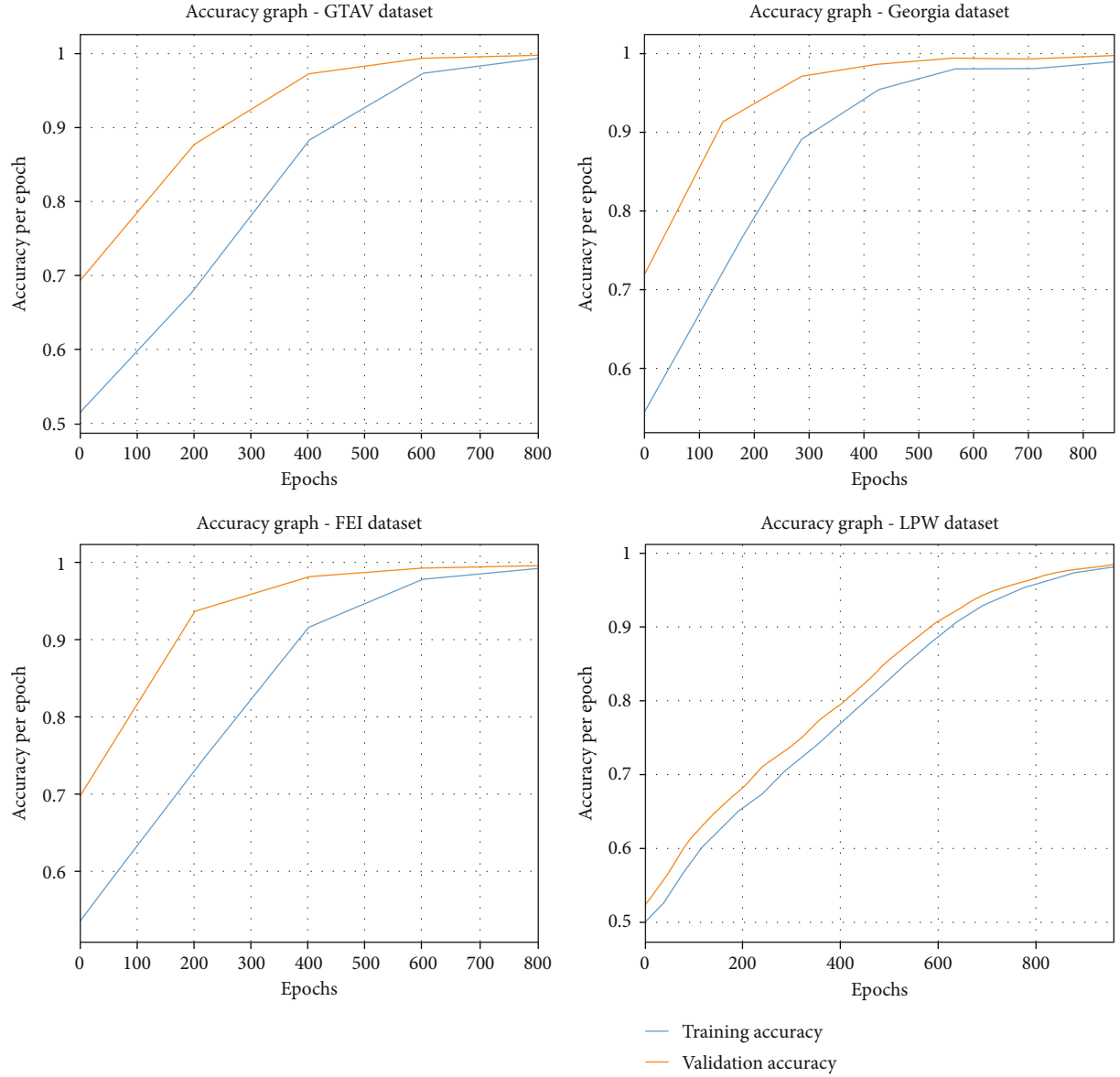
FIGURE 4: Accuracy analysis of the GWOECN-FR technique under 4 datasets.

classification error rate assumed as FF is provided in Equation (24). The optimum solution is a lesser error rate, and the worst solution gains a higher error rate.

$$\begin{aligned} \text{fitness}(x_i) &= \text{Classifier Error Rate}(x_i) \\ &= \frac{\text{number of misclassified instances}}{\text{Total number of instances}} * 100. \end{aligned} \tag{24}$$

4. Performance Validation

In this section, the result analysis of the GWOECN-FR model is carried out using four benchmark datasets. The first GTAV face dataset [20] includes images of 44 persons. The next Georgia Tech face database [21] includes a collection of images for 50 persons. The third FEI face database [22]

comprises 14 sets of images from 200 people. Finally, the Labeled Faces in the Wild (LFW) database [23] includes 13K facial images gathered from the Internet. A few sample images are demonstrated in Figure 3.

The overall accuracy outcome analysis of the GWOECN-FR technique under four datasets is portrayed in Figure 4. The results demonstrated that the GWOECN-FR algorithm has accomplished improved validation accuracy compared to training accuracy [24–30]. It can be also observable that the accuracy values get saturated with the count of epochs.

The overall loss outcome analysis of the GWOECN-FR system under four datasets is depicted in Figure 5. The figure revealed that the GWOECN-FR approach has denoted the reduced validation loss over the training loss. It is additionally noticed that the loss values get saturated with the count of epochs.
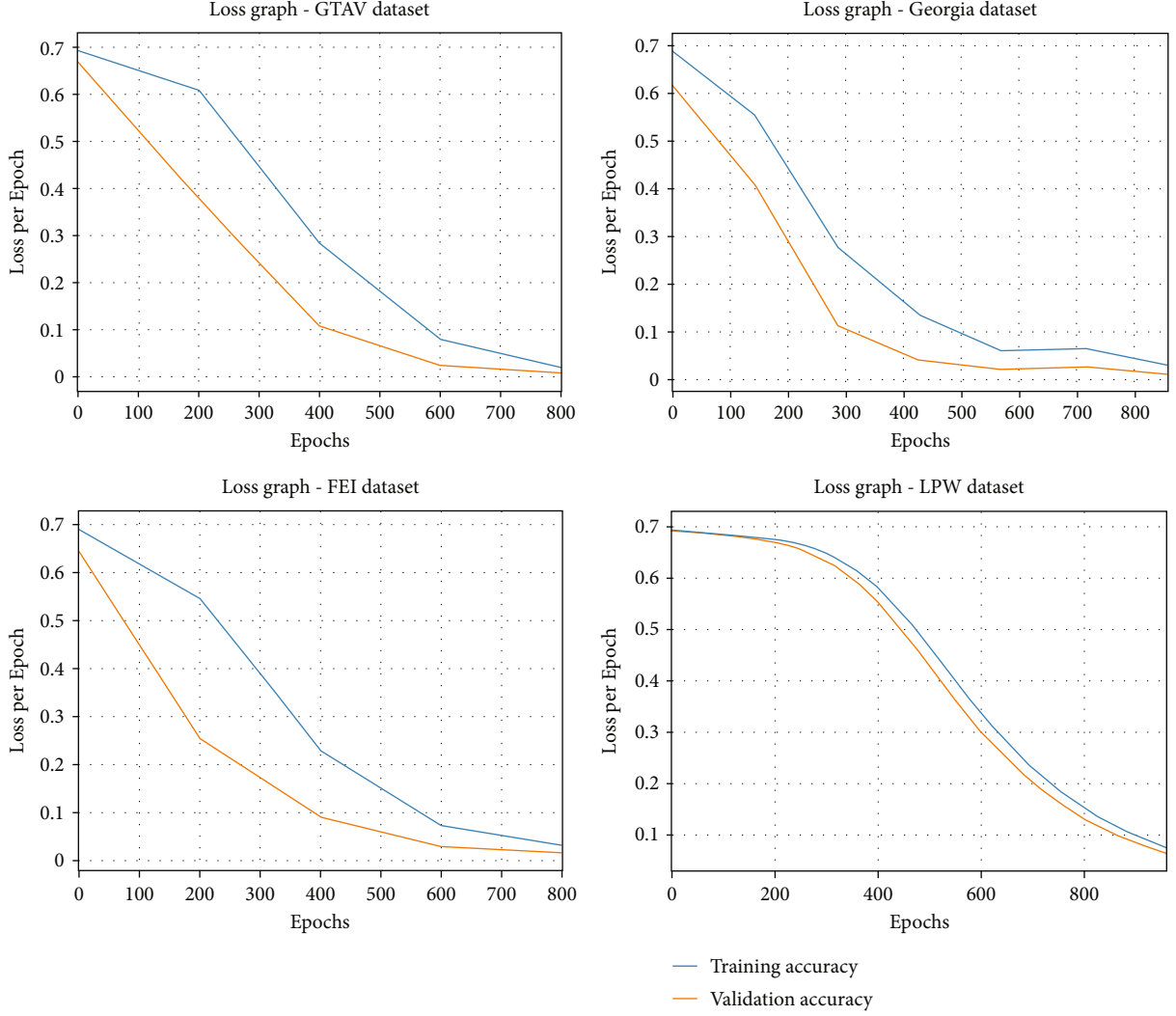
Figure 5: Loss analysis of the GWOECN-FR technique under 4 datasets.

Table 1: Accuracy analysis of the GWOECN-FR technique with existing approaches under four datasets.

Dataset	Accuracy (%)			
	AlexNet-SVM	ResNet-SVM	AlexNet	GWOECN-FR
GTAV	99.42	99.54	99.78	99.82
Georgia	96.09	96.18	98.71	99.43
FEI	97.33	98.37	98.75	99.56
LPW	93.80	93.99	95.42	98.38

Table 1 and Figure 6 report the enhanced accuracy examination of the GWOECN-FR model with recent methods on four datasets [31–34]. The results indicated the betterment of the GWOECN-FR model on all the test datasets compared to existing methods. For instance, on the GTAV dataset, the GWOECN-FR model has obtained higher accuracy of 99.82%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet models have attained lower accuracy values of 99.42%, 99.54%, and 99.78%, respectively.

Meanwhile, on the FEI dataset, the GWOECN-FR model has gained increased accuracy of 99.56%, whereas the Alex-Net-SVM, ResNet-SVM, and AlexNet models have accomplished reduced accuracy values of 97.33%, 98.37%, and 98.75%, respectively. Eventually, on the LPW dataset, the GWOECN-FR model has resulted in better accuracy of 98.38%, whereas the AlexNet-SVM, ResNet-SVM, and Alex-Net models have offered reduced accuracy values of 93.80%, 93.99%, and 95.42%, respectively.

Table 2 and Figure 7 examine the enhanced precision examination of the GWOECN-FR technique with recent methods on four datasets. The results revealed the betterment of the GWOECN-FR method on all the test datasets compared to existing techniques [35, 36]. For instance, on the GTAV dataset, the GWOECN-FR method has obtained higher precision of 99.32%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet techniques have reached minimal precision values of 96.81%, 98.84%, and 98.75%, respectively. In the meantime, on the FEI dataset, the GWOECN-FR model has gained increased precision of 98.97%, whereas
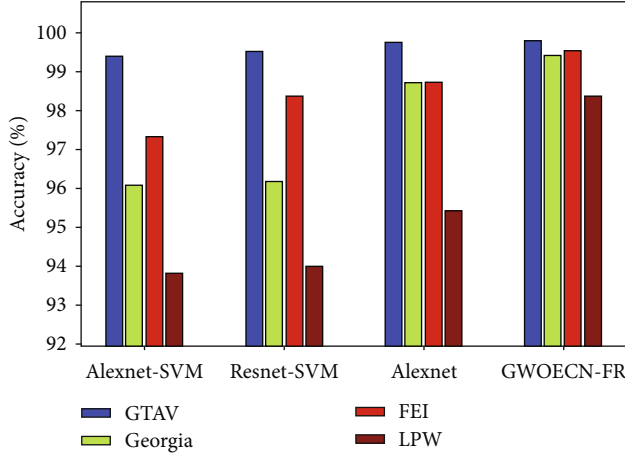
FIGURE 6: Accuracy analysis of the GWOECN-FR technique with recent approaches.

TABLE 2: Precision analysis of the GWOECN-FR technique with existing approaches under four datasets.

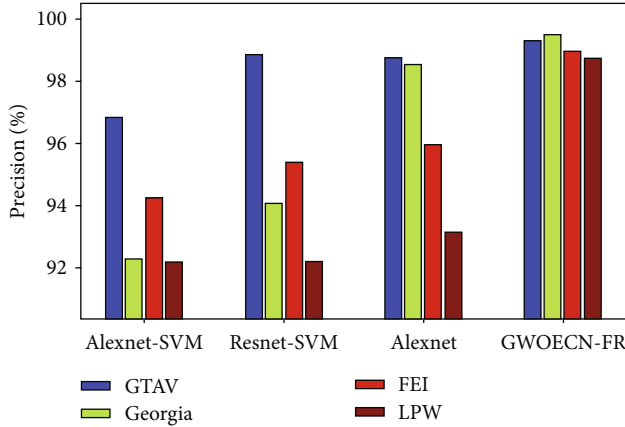| Dataset	Precision (%)			
	AlexNet-SVM	ResNet-SVM	AlexNet	GWOECN-FR
GTAV	96.81	98.84	98.75	99.32
Georgia	92.28	94.07	98.52	99.51
FEI	94.26	95.40	95.96	98.97
LPW	92.18	92.18	93.13	98.74

FIGURE 7: Precision analysis of the GWOECN-FR technique with recent approaches.

the AlexNet-SVM, ResNet-SVM, and AlexNet techniques have accomplished lower precision values of 94.26%, 95.40%, and 95.96%, respectively. Finally, on the LPW dataset, the GWOECN-FR approach has resulted in better precision of 98.74%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet algorithms have obtainable reduced precision values of 92.18%, 92.18%, and 93.13% correspondingly.

Table 3 and Figure 8 demonstrate the improved recall examination of the GWOECN-FR model with recent tech-

TABLE 3: Recall analysis of the GWOECN-FR technique with existing approaches under four datasets.

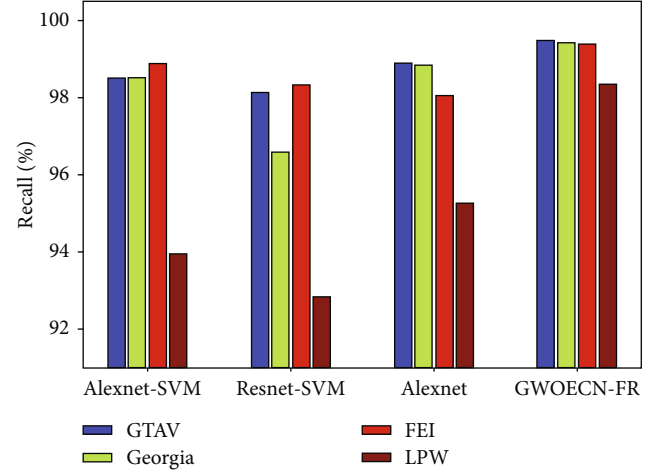| Dataset	Recall (%)			
	AlexNet-SVM	ResNet-SVM	AlexNet	GWOECN-FR
GTAV	98.52	98.14	98.89	99.51
Georgia	98.55	96.60	98.86	99.43
FEI	98.89	98.33	98.05	99.42
LPW	93.97	92.84	95.28	98.36

FIGURE 8: Recall analysis of the GWOECN-FR technique with recent approaches.

niques on four datasets. The results showed the betterment of the GWOECN-FR technique on all the test datasets compared to existing techniques. For instance, on the GTAV dataset, the GWOECN-FR approach has attained superior recall of 99.51%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet techniques have attained lesser recall values of 98.52%, 98.14%, and 98.89% correspondingly. Afterward, on the FEI dataset, the GWOECN-FR model has gained increased recall of 99.42%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet methodologies have accomplished decreased recall values of 98.89%, 98.33%, and 98.05% correspondingly. At last, on the LPW dataset, the GWOECN-FR approach has resulted in better recall of 98.36%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet techniques have offered reduced recall values of 93.97%, 92.84%, and 95.28%, respectively.

Table 4 and Figure 9 define the increased F1 score examination of the GWOECN-FR system with recent algorithms on four datasets. The outcomes referred to the betterment of the GWOECN-FR method on all the test datasets compared to existing methods. For instance, on the GTAV dataset, the GWOECN-FR method has achieved a higher F1 score of 99.31%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet models have reached lower F1 score values of 96.69%, 98.05%, and 98.99% correspondingly. Likewise, on the FEI dataset, the GWOECN-FR system has gained a higher F1 score of 99.52%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet models have accomplished minimal F1 score

TABLE 4: F1 score analysis of the GWOECN-FR technique with existing approaches under four datasets.

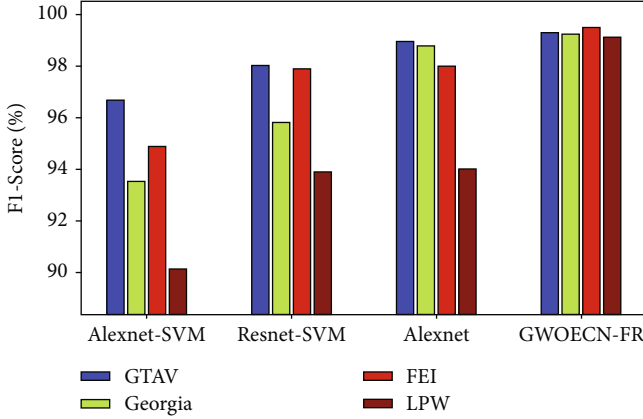| Dataset	F1 score (%)			
	AlexNet-SVM	ResNet-SVM	AlexNet	GWOECN-FR
GTAV	96.69	98.05	98.99	99.31
Georgia	93.58	95.84	98.80	99.26
FEI	94.90	97.91	98.01	99.52
LPW	90.18	93.95	94.05	99.16

FIGURE 9: F1 score analysis of the GWOECN-FR technique with recent approaches.

TABLE 5: Testing time analysis of the GWOECN-FR technique with recent methods.

Dataset	Testing time (sec)			
	AlexNet-SVM	ResNet-SVM	AlexNet	GWOECN-FR
GTAV	0.100	0.061	0.080	0.04
Georgia	0.113	0.049	0.096	0.02
FEI	0.125	0.051	0.062	0.03
LPW	0.100	0.054	0.075	0.03

values of 94.90%, 97.91%, and 98.01%, respectively. Lastly, on the LPW dataset, the GWOECN-FR methodology has resulted in a better F1 score of 99.16%, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet techniques have accessible reduced F1 score values of 90.18%, 93.95%, and 94.05%, respectively.

Finally, a detailed testing time (TST) examination of the GWOECN-FR technique with recent methods [24] is performed in Table 5 and Figure 10. The results indicated that the GWOECN-FR technique has obtained effectual outcomes with the least TST under all datasets. For instance, on the GTAV dataset, the GWOECN-FR technique has resulted in a lower TST of 0.04 s, whereas the AlexNet-SVM, ResNet-SVM, and AlexNet models have attained increased TSTs of 0.100 s, 0.0361 s, and 0.080 s, respectively.

Similarly, on the FEI dataset, the GWOECN-FR technique has achieved a reduced TST of 0.03 s, whereas the
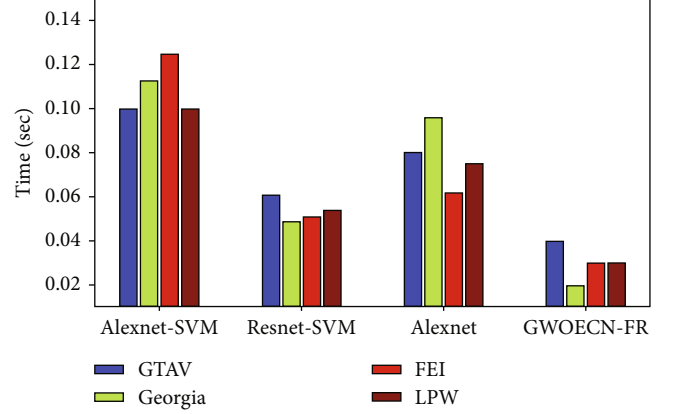

FIGURE 10: Testing time analysis of the GWOECN-FR technique with recent methods.

AlexNet-SVM, ResNet-SVM, and AlexNet models have obtained higher TSTs of 0.125 s, 0.0051 s, and 0.0062 s, respectively. From the abovementioned result analysis, it is ensured that the GWO-ECN-FR technique has accomplished the maximum FR outcome over the other techniques.

5. Conclusion

In this study, a new GWOECN-FR technique has been developed for the rapid and prompt identification of facial images. The presented GWOECN-FR technique comprises several stages of operations such as data augmentation, BF-based noise elimination, ECN-based feature extraction, SAE-based classification, and GWO-based parameter tuning. Moreover, the GWO algorithm is utilized to optimally modify the weight and bias values of the SAE model. In order to showcase the supremacy of the GWOECN-FR technique, a comprehensive result analysis is carried out on a benchmark dataset. On the FEI dataset, the GWOECN-FR method got a lower TST of 0.03 s, while the AlexNet-SVM, ResNet-SVM, and AlexNet models all got higher TSTs of 0.125 s, 0.0051 s, and 0.0062 s, respectively. The experimental results demonstrated the betterment of the GWOECN-FR technique over the recent approaches. Therefore, the GWOECN-FR technique can be applied as an effective tool for FR. In the future, hybrid DL models with hyperparameter optimizers can be involved to improve the recognition performance.

Data Availability

The manuscript contains all of the data.

Conflicts of Interest

The authors state that they do not have any conflicts of interest.

Authors' Contributions

The contributions of the authors involved in this study are as follows: conceptualization, KS and CP; methodology, SNK and SC; validation, RW; data curation, EOM; writing—original draft preparation, KS and SNK; writing—review and editing, CP and SC; and supervision, RW and EOM.

References

[1] T. Gwyn, K. Roy, and M. Atay, "Face recognition using popular deep net architectures: a brief comparative study," *Future Internet*, vol. 13, no. 7, p. 164, 2021.

[2] A. Alzu'bi, F. Albalas, T. Al-Hadhrami, L. B. Younis, and A. Bashayreh, "Masked face recognition using deep learning: a review," *Electronics*, vol. 10, no. 21, p. 2666, 2021.

[3] Y. Liu, M. Lin, W. Huang, and J. Liang, "A physiognomy based method for facial feature extraction and recognition," *Journal of Visual Languages and Computing*, vol. 43, pp. 103–109, 2017.

[4] C. Boufenar, A. Kerboua, and M. Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition," *Cognitive Systems Research*, vol. 50, pp. 180–195, 2018.

[5] G. Guo and N. Zhang, "A survey on deep learning based face recognition," *Computer Vision and Image Understanding*, vol. 189, article 102805, 2019.

[6] D. S. Trigueros, L. Meng, and M. Hartnett, "Face recognition: from traditional to deep learning methods," 2018, https://arxiv.org/abs/1811.00116.

[7] A. J. O'Toole and C. D. Castillo, "Face recognition by humans and machines: three fundamental advances from deep learning," *Annual Review of Vision Science*, vol. 7, no. 1, pp. 543–570, 2021.

[8] M. Wang and W. Deng, "Deep face recognition: a survey," *Neurocomputing*, vol. 429, pp. 215–244, 2021.

[9] Q. Wang and G. Guo, "Benchmarking deep learning techniques for face recognition," *Journal of Visual Communication and Image Representation*, vol. 65, article 102663, 2019.

[10] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: a review," *Electronics*, vol. 9, no. 8, p. 1188, 2020.

[11] Y. Zhu and Y. Jiang, "Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data," *Image and Vision Computing*, vol. 104, article 104023, 2020.

[12] A. S. Al-Waisy, R. Qahwaji, S. Ipson, and S. Al-Fahdawi, "A multimodal deep learning framework using local feature representations for face recognition," *Machine Vision and Applications*, vol. 29, no. 1, pp. 35–54, 2018.

[13] Y. Mao, S. Yi, Q. Li, J. Feng, F. Xu, and S. Zhong, "A privacy-preserving deep learning approach for face recognition with edge computing," in *Proc. USENIX Workshop Hot Topics Edge Comput. (HotEdge)*, pp. 1–6, BOSTON, MA, USA, 2018.

[14] S. Sharma and V. Kumar, "Voxel-based 3D face reconstruction and its application to face recognition using sequential deep learning," *Multimedia Tools and Applications*, vol. 79, no. 25-26, pp. 17303–17330, 2020.

[15] R. Anand, T. Shanthi, M. S. Nithish, and S. Lakshman, "Face recognition and classification using GoogLeNET architecture," in *Soft Computing for Problem Solving*, Advances in Intelligent Systems and Computing, K. Das, J. Bansal, K. Deep, A. Nagar, P. Pathipooranam, and R. Naidu, Eds., pp. 261–269, Springer, Singapore, 2020.

[16] P. Naveen and P. Sivakumar, "Adaptive morphological and bilateral filtering with ensemble convolutional neural network for pose-invariant face recognition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 11, pp. 10023–10033, 2021.

[17] C. A. O. Jian-fang, Y. A. N. Min-min, T. I. A. N. Xiao-dong, J. I. A. Yi-ming, and J. I. N. Meng-yan, "A dynasty classification algorithm of ancient murals based on adaptively enhanced capsule network," *Journal of Graphics*, vol. 42, no. 5, p. 744, 2021.

[18] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[19] J. C. Bansal and S. Singh, "A better exploration strategy in grey wolf optimizer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1099–1118, 2021.

[20] G. Sunitha, A. K. S. Pundir, S. Hemalatha, and V. Kumar, "Intelligent deep learning based ethnicity recognition and classification using facial images," *Image and Vision Computing*, vol. 121, article 104404, 2022.

[21] S. Bhargava, K. Mohan, N. R. Robert, and S. Upadhye, "Optimal stacked sparse autoencoder based traffic flow prediction in intelligent transportation systems," in *Virtual and Augmented Reality for Automobile Industry: Innovation Vision and Applications*, A. E. Hassanien, D. Gupta, A. Khanna, and A. Slowik, Eds., vol. 412 of Studies in Systems, Decision and Control, Springer, Cham, 2022.

[22] A. A. A.-R. Akeji, A. K. S. Pundir, and R. Walia, "LSGDM with biogeography-based optimization (BBO) model for healthcare applications," *Journal of Healthcare Engineering*, vol. 2022, Article ID 2170839, 11 pages, 2022.

[23] S. Neelakandan and D. Paulraj, "An automated exploring and learning model for data prediction using balanced CA-SVM," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 5, pp. 4979–4990, 2021.

[24] M. Ranjith Kumar, V. C. S. Rao, R. Anand, and H. Singh, "Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm," *Measurement*, vol. 183, article 109804, 2021.

[25] T. Kavitha, P. P. Mathai, C. Karthikeyan et al., "Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 1, pp. 113–129, 2022.

[26] C. P. D. Cyril, J. R. Beulah, P. Mohan, A. Harshavardhan, and D. Sivabalaselvamani, "An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM," *Concurrent Engineering Research and Applications*, vol. 29, no. 4, pp. 386–395, 2021.

[27] S. Neelakandan, A. Arun, R. R. Bhukya, B. M. Hardas, T. C. A. Kumar, and M. Ashok, "An automated word embedding with parameter tuned model for web crawling," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1617–1632, 2022.

[28] D. K. Jain, S. K. K. Sah Tyagi, M. Prakash, and L. Natrayan, "Metaheuristic optimization-based resource allocation technique for cybertwin-driven 6G on IoE environment," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4884–4892, 2022.

[29] D. Venu, A. V. R. Mayuri, G. L. N. Murthy, and N. Arulkumar, "An efficient low complexity compression based optimal

homomorphic encryption for secure fiber optic communication," *Optik*, vol. 252, article 168545, 2022.

[30] G. Reshma, C. Al-Atroshi, V. K. Nassa, B. Geetha, and S. Neelakandan, "Deep learning-based skin lesion diagnosis model using dermoscopic images," *Intelligent Automation & Soft Computing*, vol. 31, no. 1, pp. 621–634, 2022.

[31] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," 2007, https://hal.inria.fr/inria-00321923.

[32] O. I. Khalaf, Y. Alotaibi, S. Alghamdi, and M. Rajagopal, "Chaotic search-and-rescue-optimization-based multi-hop data transmission protocol for underwater wireless sensor networks," *Sensors*, vol. 22, no. 8, p. 2867, 2022.

[33] S. Almabdy and L. Elrefaei, "Deep convolutional neural network-based approaches for face recognition," *Applied Sciences*, vol. 9, no. 20, p. 4397, 2019.

[34] N. Soni, M. Kumar, and G. Mathur, "Face recognition using SOM neural network with different facial feature extraction techniques," *International Journal of Computer Applications*, vol. 76, no. 3, pp. 7–11, 2013.

[35] Q. Wang, "An analysis of the Viola-Jones face detection algorithm," *Image Processing On Line*, vol. 4, pp. 128–148, 2014.

[36] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: deep hypersphere embedding for face recognition," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 212–220, Honolulu, HI, USA, 2017.