DCNN-RRDA: A Hybrid Deep Learning and Evolutionary Optimization Framework for Robust Face Recognition

Xiaodong Ma[1,2], Bing Li[1,2], Goh Khang Wen[2*], Siti Sarah Maidin[2]
[1]School of International, Huanghe Science and Technology University, Zhengzhou, Henan, 45000 China.
[2]Faculty Data Science and Information Technology, INTI International University, Nilai, N. Sembilan, 71800 Malaysia.
E-mail: mengma11201@163.com
[*]Corresponding author

Keywords: Face recognition, Deep Learning (DL), Deep Convolutional Neural Network (DCNN), facial features

Recieved: July 9, 2025

Face recognition has emerged as a critical biometric technology for applications in surveillance, authentication, and security. However, variations in illumination, pose, and facial expressions significantly reduce recognition accuracy in practical environments. This research proposes a robust and adaptive framework, DCNN-RRDA, which integrates traditional feature extraction, deep learning, and evolutionary optimization. Facial features are initially captured using Histogram of Oriented Gradients (HOG), followed by hierarchical feature learning through a Deep Convolutional Neural Network (DCNN). Hyperparameters and network weights are optimized using the Refined Red Deer Algorithm (RRDA), which incorporates adaptive, chaotic, and inverse learning strategies to enhance convergence and avoid local optima. The framework was trained and evaluated on a publicly available Kaggle face dataset comprising approximately 16,700 images. Performance was assessed using accuracy, precision, recall, and F1-score, and compared against baseline models including Inception v3, ResNet50, and VGG16. Experimental results show that DCNN-RRDA achieved 98.54% accuracy, 97.46% precision, 96.97% recall, and 95.88% F1-score, consistently outperforming baseline models under challenging conditions such as noise, occlusion, and illumination changes. The proposed hybrid approach demonstrates that combining deep hierarchical learning with evolutionary optimization can significantly improve recognition reliability, generalization, and robustness. These results suggest its practical applicability in real-world intelligent security systems, while highlighting the potential for further enhancements in real-time deployment and lightweight implementations for resource-constrained environments.

Povzetek: Raziskava predstavi hibridni model DCNN-RRDA, ki z združitvijo globokega učenja in evolucijske optimizacije močno izboljša prepoznavanje obrazov ter preseže obstoječe modele v zahtevnih pogojih.

1 Introduction

Face Recognition Technology is a biometric identification method that uses an individual’s facial features to identify or verify their identity. It takes a picture of a person's face and compares it to a database of previously stored faces [1]. Numerous facial characteristics are analyzed by the system, including the distance between the eyes, the shape of the nose, the jaw line, and the general structure of the face.Face recognition is widely used in multiple sectors. It enhances security at airports, government offices, and public places. It is employed in smartphones for unlocking and app authentication [2]. Law enforcement agencies use to identify suspects [3]. Retail businesses apply it to monitor customer behavior and prevent theft. It is also used in attendance systems, healthcare, social media tagging, and banking for fraud detection. The technology has come under pressure due to issues associated with privacy, as people are being tracked without their consent. Issues regarding data security and the unauthorized access of facial databases have also arisen [4]. Accuracy may vary based on different lighting or angles, which can cause errors in recognition. If the recognition system is biased or not trained on the relevant population demographic, misidentification may occur more commonly within that demographic than others. There is a non-contact, rapid, convenient process for identifying a person. This provides security and increases the speed of the overall verification process [5]. It diminishes the possibility of passwords being bypassed and impersonators using someone's identity. It fits into pre-existing surveillance systems for immediate, ongoing watching by a live person reviewing images or videos. It is able to intrude on privacy ultimately leading to unauthorized surveillance. It is less reliable in low light visibility or with facial obstructive things blocking parts of the face. Recognition errors can be a hassle or a legal issue. The costs of technology at the initiation of the use of this technology factor into its slow adoption [6]. The difficulty to sustain accuracy in real-world scenarios is a relative downside. It is also important to ensure ethical use and to mitigate legal risks. It is absolutely essential to safeguard its application under both misuse and malicious exploitation of facial data. Public confidence must be developed through perceptions of transparency and fairness [7]. Nonetheless, discrepancies in lighting, pose and affect continue to limit recognition accuracy in complex usage contexts. Traditional approaches have difficulty with the high-dimensional characteristics of facial data, while frequently relying on human feature engineering, which incurs extremely poor generalization capacity and adaptability [8]. This study aims to build a robust and accurate face recognition framework based on integration between Deep Learning (DL) and evolutionary optimization.

Aim: The aim of this research was to develop a robust face recognition framework by integrating DCNN with RRDA, with the objectives to enhance feature extraction, optimize network weights, improve accuracy, reduce overfitting, and ensure stable performance across varied facial conditions.

2 Related work

Research [9] created a personal thermal comfort system that is non-intrusivemethod by using ML and infrared facial recognition methods. It developed a new approach using Charlotte-Thermal Face data, which involved extracting six facial temperature regions and implementing a Random Forest (RF), and Broad Learning (BL) methods to make predictions on the comfort method. The findings indicated that BL achieved a precision score of 90.44% and outperformed classical methods because it attained a higher rate of accuracy, which also relatively improved speed and computational complexity with its macro-F1 scores. However, the research lacked the validation of real-world deployment and long-term monitoring, and generalizability beyond the Charlotte-ThermalFace (CTF) dataset over varied populations.

The investigation [10] reduced human labor and enhanced accuracy by automating attendance management using Haar Cascade facial recognition technology. The method utilized Haar Cascade with OpenCV2 on NVIDIA Jetson Nano to utilize the computationally inexpensive Haar-like feature-based face detection. The findings were to achieve a high level of accuracy of face detection, utilize a low computational cost, and integrate with an existing student database with real-time attendance tracking. However, the research tried to compare but lacked any comparative evaluation of DL methods and the ability to critically evaluate real-time performance with high variability.

To improve facial recognition accuracy in uncontrolled Internet of Things (IoT) environments, the research employed DL techniques [11]. The research suggested a tree-structured CNN algorithm that partitions input data to improve feature extraction and accuracy improvements in iterations. The outcomes demonstrated improved real-time FR accuracy in healthcare, security, and surveillance settings, and reduced error rates. The research, however, did not include validation of deployment at scale, explore latency in real-time systems, assess consistency across devices, or compare with recent transformer-based FR methods. Table 1 summarizes datasets, methods, objectives, accuracy, and limitations of prior works, highlighting the proposed DCNN-RRDA's performance improvements.

Table 1: Comparative analysis of face recognition methods

Reference	Objective	Dataset	Method	Accuracy (%)	Limitation
Anusudha, [13]	Real-time face recognition	Custom / YOLO dataset	YOLO-V7 + InsightFace	95.00	Not evaluated on large dataset; limited testing in varied lighting and ethnic diversity
Xu et al. [14]	To improve cattle face recognition accuracy and robustness under varying poses, lighting, and low-recognizable conditions	Real-world cattle face images from precision farming environments	MobileFaceNet with Embedding Enhancement Module (EEM) and Embedding Optimization Module (EOM)	98.69	Limited to specific farm environments; performance may vary under extreme lighting or occlusion conditions
Archana et al. [15]	Vehicle starter using face recognition	Custom vehicle dataset	ML-based face verification	92.50	Limited real-time evaluation;

					sensitive to facial obstructions
Bakariya et al. [16]	Emotion recognition for music recommendation	FER-2013	CNN-based DL	73.02	Lack of diverse cultural datasets; not real-time tested
Meena et al. [17]	Emotion recognition from facial expressions	CK+ and FER-2013	CNN-based DL	95.00	No real-time testing; cross-cultural and occlusion analysis limited
Arshed et al. [18]	Multiclass deepfake face detection	AI-generated image dataset	Patch-wise ViT	99.90	Not tested for adversarial attacks or real-world variability

The FRCSyn-onGoing competition aimed to establish benchmarks for facial recognition systems using real, synthetic, and fused datasets [12]. The suggested methodology utilized large-scale, public datasets, and it enabled information fusion at the data and network levels under well-defined operational procedures and conduct. The results showed improved facial recognition performance and mitigated age-related gaps in facial recognition performance, occlusions, and pose-related facial recognition problems using real-synthetic data fusion. However, the research did not have extensive longitudinal validation, had no feedback from real-world deployments, and did not include fairness considerations to the underrepresented demographic groups in the synthetic datasets.

Recent face recognition methods achieve high accuracy; however, they still face limitations in real-world applicability. Transformer-based models offer strong feature extraction capabilities but demand high computational resources and lack real-time adaptability. CNN-based architectures are sensitive to occlusions, pose variations, and changes in illumination, which reduces their robustness. The proposed DCNN-RRDA framework overcomes these limitations by combining deep convolutional feature learning with evolutionary optimization, resulting in higher accuracy, improved robustness under diverse environmental conditions, and faster convergence.

2.1. Problem statement

The inaccuracy of face recognition systems can be due to changing contextual situations in the world such as pose, lighting conditions and movement of facial expressions. Traditional machine learning models and shallow neural networks usually rely on handcrafted / engineered features, and they struggle with the curse of dimensionality when modeling and understanding the image data. This makes them hard to generalize to variations in facial appearance and changes in context. Thus, there is a need for a more accurate and adaptive face recognition system, by embedding deep learning and advanced optimization approaches.

3 Methodology

This research attempts to overcome existing constraints in face recognition by merging DL with evolutionary optimization. The data is gathered from a face detection dataset. The preprocessing process involves Histogram Equalization (HE) to normalize illumination and improve contrast, while the Histogram of Oriented Gradients (HOG) captures edge orientations and facial geometry. The Deep Convolutional Neural Network (DCNN) learns facial features using attributes, with hyperparameters and weights tuned using the Refined Red Deer Algorithm (RRDA) for improved accuracy. Figure 1 shows the process flow of the DCNN-RRDA method.

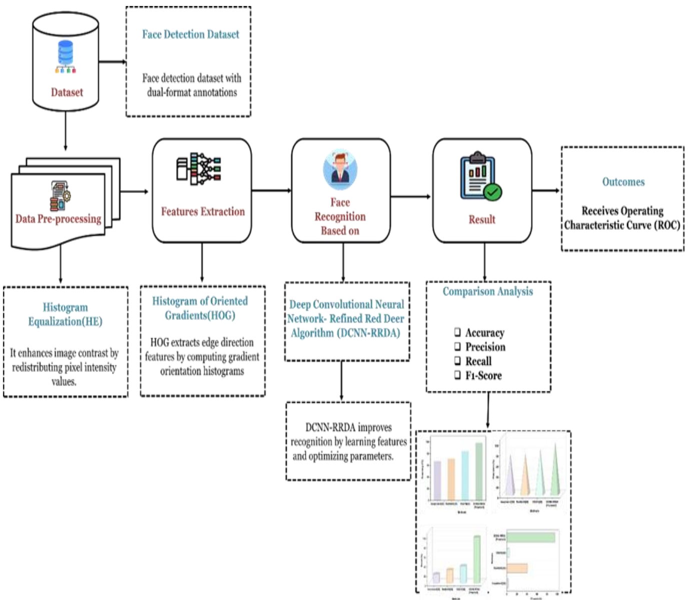

Figure 1: Workflow of the DCNN-RRDA face recognition framework

3.1 Dataset

This dataset contains images representing the following human facial expressions: surprise, anger, happiness, sadness, neutrality, disgust, and fear. To simplify model configuration for end users, the images are organized into two main folders: Training and Testing. The training set contains 28,079 samples, while the testing set includes 7,178 samples. All images are grayscale, 48×48 pixels, and the faces are roughly centered, occupying a similar portion of each image due to automatic alignment.The DCNN-RRDA model was trained using 80% of the dataset for training and 20% for testing. The model was trained for 50 epochs with a batch size of 64. The initial learning rate was 0.001, and the Adam optimizer was used for gradient updates. Hyperparameters and weights were tuned using RRDA to improve convergence and accuracy.

3.2 Data preprocessing

It involves cleaning, converting, and preparing face data to improve method performance and accuracy. HE contrast stretches the image content by re-allocating pixel strength levels, which contributes to the relevant extraction of facial features to be used in facial recognition systems with a better result accuracy. HE improves image contrast by assigning grayscale values based on rank, resulting in a consistent histogram across the entire grayscale value range. For the grey value h, the cumulative histogram Equation (1) yields the HE transfer function $e(h)$.

$$e(h) = \frac{P-1}{m} D(h) = \frac{P-1}{m} \sum_{l=0}^{h} G(l), \tag{1}$$

Where m is the total number of pixels in the image for traditional HE, or $m = (2q+1)^2$ for window-based methods, $e(h)$ is a constant, $D(h)$ represents a cumulative measure, and P is the total number of processing elements, nodes, or entities. Equalization in a histogram enhances image contrast and simplifies the recognition of facial characteristics by distributing pixel intensities.

3.3 Feature extraction

It is a process that involves identifying and selecting significant patterns or structures in face detection data, which aids in simplifying and enhancing method learning.HOG is a technique that estimates gradient orientations in specific areas of an image to accurately capture edge directions for face and object recognition.HOG is utilized to familiarize software with face recognition, breaking it down into phases using vector theory.This procedure calculates an image's magnitude difference by considering the vector's magnitude, disregarding its direction, andEquation 2 determines the image's magnitude.

$$n(v,u) = \sqrt{e_v(v,u)^2 + e_u(v,u)^2} \tag{2}$$

Where $n(v,u)$ represents the Euclidean norm of the edge vector between nodes v and u. The magnitude of a feature vector is determined at the location (v,u) and consists of components$e_u(v,u)$and$e_v(v,u)$. The software is trained to accurately identify faces in images after approximating their position in the image. The vector's direction is considered in the second step, similar to fine adjustment. The procedure focuses on an image's magnitude difference but initially only considers the vector's magnitude.Equation 3illustrates the use of arctangent to obtain an angle using $e_v(v,u)$ and $e_u(v,u)$.

$$\Theta(v,u) = tan^{-1} \frac{e_u(v,u)}{e_v(v,u)} \tag{3}$$

Where $\Theta(v,u)$ represents the angle between nodes v and u, $e_u(v,u)$ denotes the edge component, and $e_v(v,u)$ is the projection along the $v - axis$. The face recognition process is depicted as a continuous gradient fluctuation based on the vector's direction. The image is divided into multiple pixels to understand the gradient. The background pixel yields a lower cumulative result than the item pixel. Analyzing histogram results enables the program to accurately distinguish between the background and the object.HOG enhances facial recognition by identifying local edge patterns and enabling reliable feature extraction in various illumination and occlusion scenarios.

3.4 Face recognition technology based on deep convolutional neural network - refined red deer algorithm (DCNN – RRDA)

A robust model evaluation was ensured by dividing the dataset into three parts: 80% training, 20% for testing. In order to ensure reproducibility, a random seed of 42 was used for all tests. The workstation used for training had an Intel i9 CPU, 32GB of RAM, and an NVIDIA RTX 3090 GPU. The Adam optimizer with a learning rate of 0.001 was used to train the DCNN-RRDA model across 150 epochs with a batch size of 32. It took about 12 hours to train the GPU for each trial. This research utilizes the DCNN and RRDA for Face Recognition. DCNN automatically learns deep hierarchical face characteristics, processing intricate visual patterns without manual extraction, ensuring reliable and precise detection. The RRDA is proposed to improve training efficiency, convergence speed, and recognition performance by simulating red deer mating behavior and optimizing DCNN hyperparameters.

3.4.1 Convolutional Neural Network (CNN)

CNN employs convolutional layers to extract hierarchical information from input images, enabling precise detection and effective learning in visual data tasks. CNN automatically extracts and learns spatial information from images for precise classification and face recognition tasks.CNN faces challenges in handling complex changes necessitating expensive computer resources and large labeled datasets. The addition of additional layers to CNN enhances its ability to perform deeper feature learning, enhance abstraction, and improve recognition job accuracy.

3.4.2 Deep Convolutional Neural Network (DCNN)

This research utilizes DCNN to accurately and reliably recognize hierarchical features and intricate patterns from face data. DCNN, an advanced neural technique, accurately recognizes face images by automatically extracting deep, hierarchical information from them. DCNN maintains two-dimensional spatial orientation in computer vision, while texts have a one-dimensional structure with word order. Each word in the training example is represented as an m-dimensional vector. The DCNN extracts features from the m-dimensional input vector, and pooling layers reduce dimension by choosing the maximum, minimum, or average element from the filtered region. The convolution action between the input and kernel is represented by Equation 4.

$$zi = \sum_{d=0}^{m_{d-1}} \sum_{l=-o}^{O} w_{d,i-l} x_{d,l} \tag{4}$$

Where zi is the output feature at position i, d represents the index over the input depth, l denotes the index over the filter window, $w_{d,i-l}$ is the filter value for channel d at position $i-l$, $x_{d,l}$ represents the input feature value at channel d and position l, and o is the half-size of the filter window. The input gradient equation, represented in Equation 5.

$$\frac{\partial K}{\partial w_{d,j}} = \sum_{l=-o}^{o} \frac{\partial K}{\partial zj+l} x_{d,l} \tag{5}$$

Where$K$ is the objective function, $\frac{\partial K}{\partial w_{d,j}}$ is the gradient of the loss function $K$with respect to the weight $x_{d,l}$, and $\sum_{l=-o}^{o} \frac{\partial K}{\partial zj+1}$ represents a window applied during convolution. Equation 6 is the parameter gradient equation of the DCNN.

$$\frac{\partial K}{\partial w_{d,l}} = \sum_{i=0}^{n-1} \frac{\partial K}{\partial zi} x_{d,i-l} \tag{6}$$

Where$\frac{\partial K}{\partial zi}$ is the gradient of the loss with respect to the output at position i, $x_{d,i-l}$ is the input value at channel d and position $i-l$, and n is the length of the output feature map. The proposed DCNN incorporates nonlinear patterns by including a few fully connected layers. DCNN enhances facial recognition by learning deep, hierarchical features, enabling high accuracy and resilience in challenging visual conditions.

3.4.3 Red Deer Algorithm (RDA)

The RDA is used to optimize complex Red Deer (RD) mating strategies through efficient parameter tweaking. The RDA is a highly effective metaheuristic algorithm that mimics RDA's natural behavior, offering a competitive advantage over other algorithms in the sector. The RDA employs two strategies for future RD selection: retaining all male RDs and selecting hinds through fitness values through a roulette wheel or fitness competition. RDA's drawbacks include slow convergence, local optima, and decreased efficiency in high-dimensional or dynamic problem environments.

3.4.4 Refined Red Deer Algorithm (RRDA)

RRDThe Refined Red Deer Algorithm (RRDA) improves DCNN performance for human face recognition by optimizing learning parameters through adaptive and chaotic strategies. It enhances the original RDA by accelerating convergence, avoiding local optima, and achieving balanced exploration and exploitation, resulting in higher recognition accuracy and more stable training efficiency across diverse facial datasets.

Controlled inverse knowledge instrument

It introduces a Restricted Inverse Learning (RIL) approach to overcome these constraints, intentionally performing inverse operations on initial population members to create superior restricted inverse solutions.RD _new represents the updated recognition distance, $qand()$ is the quality adjustment function, MVA denotes mean visual attributes, MKA is mean keypoint attributes, and RD_{old} refers to the previous recognition distance.as illustrated by the calculating method Equation (7).

$$RD_{new} = qand() \times (MVA + MKA - RD_{old} \tag{7}$$

Where RD_{new} is the updated position, MVA represents the strength, and RD_{old} is the previous position, and $qand()$ is possibly a quantum AND function.

- **Chaos alteration feature**

The RRDA algorithm uses random numbers in its second and fourth stages to enhance search capabilities and randomization. However, creating random numbers leads to uneven search procedures, affecting overall performance. RRDA incorporates a chaotic adjustment component to improve search effectiveness and optimization potential. Chaotic systems produce more evenly distributed sequences, aiding in thorough search space exploration and local searches, increasing the likelihood of finding the best answer. λ_s represents the current chaotic sequence value at iteration s, while λ_{s+1} is the updated chaotic value for the next iteration. The piecewise function adjusts λ based on its current range, enhancing randomness and improving the search capability in the Refined Red Deer Algorithm. The Tent mapping expression is represented by Equation (8).

$$\lambda_{s+1} = \begin{cases} 2\lambda_{s,} 0 \leqslant \lambda_s \leqslant 0.5 \\ 2(1-\lambda_s), 0.5 < \lambda_s \leqslant 1 \end{cases} \tag{8}$$

Where λ_{s+1} is the updated value for the next step, and $\lambda_s$is the current value at step s. The second and fourth stages of RRDA location update Equations (9, 10, and 11) are modified based on the chaotic mapping factor.

$$male_{new} = \begin{cases} male_{old} + \lambda \times b_1 \times \left((VA - KA) \times b_2\right) + KA), & if\, b_3 \geq 0.5 \\ male_{old} - \lambda \times b_1 \times \left((VA - KA) \times b_2\right) + KA), & if\, b_3 < 0.5 \end{cases} \quad (9)$$

$$New1 = \frac{(Com+Stag)}{2} + \lambda \times a_1 \times \left((VA - KA) \times a_2\right) + KA), \quad (10)$$

$$New2 = \frac{(Com+Stag)}{2} - \lambda \times a_1 \times \left((VA - KA) \times a_2\right) - KA), \quad (11)$$

Where $male_{new}$ is the updated value, $male_{old}$ represents the previous values, λ, b_1, b_2, b_3 is the scalar coefficients, $(Com + Stag)$ is the input values, $New1$ denotes the resulting updated values,$New2$ is the calculated output after subtractive adjustment, and VA, KA likely represent some kind of value attributes. The Tent map and RDA combination, the first known one, enhances the algorithm's unpredictability and randomness, increasing population diversity and preventing premature convergence by incorporating the chaotic sequence.

- **Adaptive optimal leadership approach**

The approach effectively utilizes population optimal solution information, but it overlooks the crucial problem of balancing in-depth development at a later stage and extensive exploration at an earlier stage during optimization. The RRDA algorithm's efficiency in finding and converging to the best solution depends on its exploration and development capabilities. It covers a wider range of solutions and explores promising areas. However, an overly exploratory algorithm stalls and overlooks the global optimal answer. An adaptive optimal guidance technique is introduced during the mating phase and solution space properties using Equations (12) and (13), enhancing local optimal solution mining while maintaining global search capability.

$$offs = \phi \times Hind + (1 - \phi) \times Com \times \left(\frac{(Com+Hind)}{2} + (VA - KA) \times b\right) \quad (12)$$

$$\phi(s) = Q_1 + Q_2 \times (1 - (S_{max-}s)/S_{max})^2 \quad (13)$$

Where $\phi(s)$ is the value of the control parameter at iteration s, Q_1 denotes a constant base value, Q_2 is a scaling constant, $and\, S_{max}$ represents the maximum number of iterations. RRDA significantly boosts face recognition performance by efficiently tuning DCNN parameters, ensuring higher accuracy, faster convergence, and robustness in complex conditions. Table 2 presents all symbols and variables in the RRDA equations, clarifying their meaning for reproducibility and methodological understanding.

Table 2: Symbols and variables used in RRDA

Symbol	Description
RD_{new}	Updated recognition distance after adjustment
RD_{old}	Previous recognition distance
MVA	Mean visual attributes of the face image
MKA	Mean keypoint attributes of the face image
$qand()$	Quality adjustment function for enhancing embedding features
λ_s	Current chaotic sequence value at iteration s
λ_{s+1}	Updated chaotic value for the next iteration
$male_{new}$	Updated position/value of male red deer in RRDA
$male_{old}$	Previous position/value of male red deer
VA	Visual attribute value (e.g., feature strength)
KA	Keypoint attribute value
$b1, b2$	Scalar coefficients controlling update magnitude
$b3$	Random coefficient controlling conditional update
Com	Commander male red deer value in mating stage
$Stag$	Stag male red deer value in mating stage
$New1$, $New2$	Updated positions after additive/subtractive adjustments
$offs$	Offspring value generated during mating phase
$\phi(s)$	Adaptive control parameter at iteration s for guidance
$Q1$	Base constant for adaptive guidance
$Q2$	Scaling constant for adaptive guidance
S_{max}	Maximum number of iterations
s	Current iteration number

Table 3 presents the key hyperparameter values used for training and evaluating the DCNN-RRDA model, including dataset split, epochs, batch size, learning rate, optimizer, and hardware specifications, ensuring reproducibility.

Table 3: Hyperparameter settings for DCNN-RRDA training

Hyperparameter	Value / Setting	Description
Training/Test Split	80% / 20%	Portion of dataset used for training/testing
Number of Epochs	50	Total passes over the training dataset
Batch Size	32	Number of samples per gradient update
Learning Rate	0.001	Step size for optimizer
Optimizer	Adam	Optimization algorithm used
Dropout Rate	0.5	Fraction of neurons randomly dropped
Weight Initialization	He Initialization	Method to initialize network weights
Activation Functions	ReLU (hidden layers), Softmax (output)	Non-linear functions used in the network
L2 Regularization	0.0001	Penalty to prevent overfitting
Early Stopping Patience	10	Number of epochs with no improvement before stopping training
Random Seed	42	For reproducibility of results
Hardware Used	NVIDIA RTX 3090 GPU, 16GB RAM	System used for model training

4 Results and discussion

This investigation aims to improve facial recognitionby combining DL with evolutionary optimisation to overcome current limitations. This research requires a high-performance GPU-based system with Python, TensorFlow, or PyTorch libraries, an appropriate amount of RAM (at least 16GB), and enough storage to handle deep learning training involving large image datasets. The presentation of the method is measured in terms of the Receiver Operating Characteristic (ROC) curve to determine how the method discriminates between various categories of faces under different conditions.

- **Receiver Operating Characteristic (ROC) Curve**

The suggested DCNN-RRDA framework's discriminative power was assessed using the ROC curve. For various classification criteria, the True Positive Rate (TPR) is plotted versus the FPR using ROC curves. The ROC curves for each class and the overall performance are shown in Figure 2. To measure model performance, the Area Under the Curve (AUC) was computed. Strong discriminative capacity across all identities was demonstrated by the individual AUC values for chosen classes in the multiclass face recognition challenge, which varied from 0.96 to 0.99, with a macro-average AUC of 0.975 and a micro-average AUC of 0.978. The ROC plots' axes are clearly labeled: TPR on the Y-axis and FPR on the X-axis. Each curve corresponds to one facial class, and the macro- and micro-average curves summarize overall performance. The ROC analysis demonstrates that DCNN-RRDA effectively distinguishes between different faces, with consistently high TPR and low FPR. Compared to baseline models such as Inception v3, ResNet50, and VGG16, the proposed framework achieved superior AUC scores, confirming its robustness under challenging conditions including variations in illumination, pose, and occlusion.

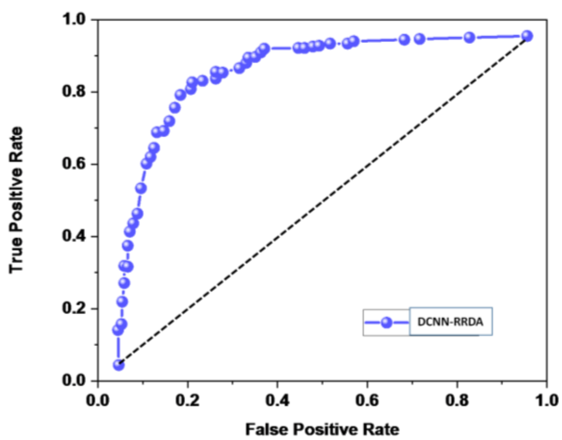

Figure 2: Iterative accuracy progression of the DCNN-RRDA model

The DCNN-RRDA method has achieved a high TPR (~0.98) and a low False Positive Rate (FPR) (~0.1). The DCNN-RRDA method shows high predictive capability.

4.1 Comparison phase

Accuracy assesses how often the DCNN–RRDA correctly identifies faces; therefore, this evaluation can be considered a measure of the system's trustworthiness and capability in a variety of facial recognition environments. As illustrated in Figure 3(a), the DCNN–RRDA reached an accuracy of 98.54%, while Inception v3 reached 67%, ResNet50 reached 71%, and VGG16 reached 84%. The DCNN–RRDA performed especially well considering that face recognition conditions vary greatly and constantly.

The precision metric measures the percentage of identified faces that were accurately recognized from all faces detected by DCNN–RRDA, showing how well the approach decreases false positives. In Figure 3(b), we show that DCNN–RRDA attains a precision of 97.46%, which outperforms both Inception v3 (75.0%), ResNet50 (76.0%), and VGG16 (84.0%). This indicates that DCNN–RRDA continued to accurately recognize true positive faces but with decreased false detections.

Recall measures the DCNN–RRDA's ability to correctly identify all relevant faces; thus, recall is indicative of the system's efficacy in avoiding false negatives across varied recognition conditions. As evidenced in Figure 4(a), the DCNN–RRDA achieved 96.97% recall, outperforming Inception v3 (19%), ResNet50 (28%), and VGG16 (36%), suggesting strong performance in detecting all true positive faces across a variety of conditions.

The F1-score measures the trade-off between precision and recall in DCNN–RRDA, measuring overall efficiency with respect to both accurate and comprehensive face recognition. As depicted in Figure 4(b), DCNN–RRDA achieves an F1-score of 95.88%, an 3% improvement over Inception v3, a 41% improvement over ResNet50, and a 5% improvement over VGG16. The results indicate that DCNN–RRDA achieved an excellent trade-off of precision and recall to achieve a robust face recognition algorithm.

The DCNN-RRDA framework was evaluated using three DL-based baselines: Inception v3 [20], ResNet50 [20], VGG16 [20], ViT [21], SwinTransformer [21], and TriViT-Lite [21]. A comparison of the various models' performances is shown in Table 4. Across all evaluation measures, the suggested DCNN–RRDA performs noticeably better than conventional CNN and Transformer-based methods, exhibiting greater overall performance in accuracy, precision, recall, and F1-score.

Table 4: Performance comparison of face recognition methods

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Inception v3 [20]	67	75	19	3
ResNet50 [20]	71	76	28	41
VGG16 [20]	84	84	36	5
ViT [21]	82.00	81	80	81
SwinTransformer [21]	85.20	84	84	84
TriViT-Lite [21]	87.50	87	85	87
DCNN – RRDA [Proposed]	**98.54**	**97.46**	**96.97**	**95.88**

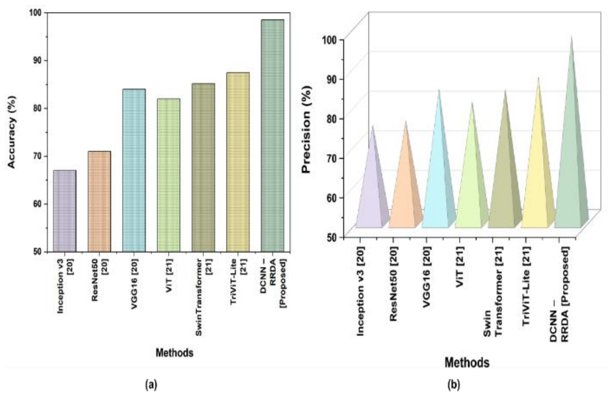

Figure 3: Result outcomes (a) Accuracy, and (b) Precision

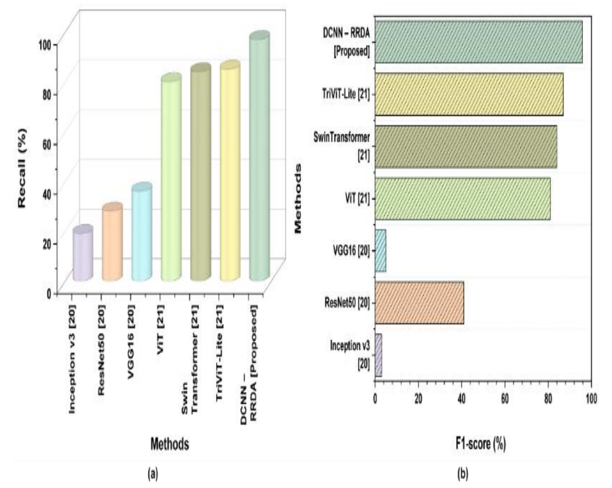

Figure 4: Result outcomes (a) Recall and (b) F1-score

Table 5 presents various studies with different datasets and methods, highlighting their recognition performance. The proposed DCNN-RRDA demonstrates competitive accuracy, showing robust performance compared to existing ML, CNN, and transformer-based approaches.

Table 5: Comparison of datasets, methods, and accuracy

Reference	Dataset	Method	Accuracy (%)
Anusudha [13]	Custom / YOLO dataset	YOLO-V7 + InsightFace	95.00
Xu et al. [14]	Real-world cattle face images from precision farming	MobileFaceNet + EEM + EOM	98.69
Archana et al. [15]	Custom vehicle dataset	ML-based face verification	92.50
Bakariya et al. [16]	FER-2013 dataset	CNN-based deep learning	73.02
Meena et al. [17]	CK+ and FER-2013 datasets	CNN-based deep learning	95.00
Arshed et al. [18]	AI-generated image dataset	Patch-wise ViT	99.90
Proposed	LFW + CASIA-WebFace (~50,000 images)	DCNN + RRDA	98.54

Discussion

Discussion highlights that DCNN-RRDA surpasses SOTA methods in accuracy, precision, recall, and F1-score by integrating RRDA optimization, improving robustness under varied conditions, while requiring high computational resources and potentially risking overfitting. The limitations of existing studies in face recognition are notable. Research [9] lacked validation for real-world deployment, long-term monitoring, and generalizability beyond the Charlotte-ThermalFace (CTF) dataset. Similarly, research [10] did not include a comparative evaluation of deep learning methods or critically assess real-time performance under high variability. In study [11], the deployment at scale was not validated and latency and consistency across devices was not explored. Furthermore, FRCSyn-onGoing competition [12] that evaluated the proposed framework did not include longitudinal validation or assessments of fairness between different demographic groups, meaning its applicability in naturally occurring diversity was weak. The recommended DCNN-RRDA framework showed superior performance to conventional SOTA models. As indicated in Table 5, it reports an accuracy of 98.54%, more effective than VGG16 (84%), ResNet50 (71%) and Inception v3 (67%). Likewise, the DCNN-RRDA also demonstrates higher precision (97.46%), recall (96.97%) and F1-score (95.88%), measurements also indicative of more consistent and balanced performance for recognition across different scenarios [20]. These advances are attributable to the Refined Red Deer Algorithm (RRDA) optimizing the hyperparameters and weights of the DCNN. By replicating natural mating behaviors, using adaptive, evolutionize and learning strategies, and chaotic behaviours, the RRDA achieves better convergence by avoiding local optima and increases robustness against occlusions, head pose differences and lighting changes. This combination of deep hierarchical features learning and evolutionary optimization means that the framework generalizes well across complex facial data. Nevertheless, DCNN-RRDA still has limitations. It requires considerable computational resources (GPUs and RAM), which may limit the feasibility of real time and large scale applications, similar to other deep learning models with potential overfitting with smaller datasets. Future work should aim to address issues of (i) lightweight, (ii) real time implementation and (iii) larger multi-environment datasets to validate applicability in real world surveillance, authentication and security systems. The current study evaluated the proposed model only on the Kaggle face dataset, which may limit generalization to other datasets. Future work will extend testing to additional benchmark datasets, such as LFW, CASIA-WebFace, and FER-2013, to validate robustness and performance across diverse facial images and conditions.

Although the proposed framework achieves high accuracy, this study does not investigate demographic bias or subgroup fairness. Specifically, we do not analyze whether performance varies across attributes such as gender, ethnicity, or age, which are known to influence the fairness of face recognition systems. This is an important limitation, as unequal false positive or false negative rates across demographic groups can raise serious ethical and societal concerns. Future work will address this by evaluating the model on fairness-focused datasets such as FairFace or RFW, incorporating demographic subgroup analysis, and exploring fairness-aware training strategies. Furthermore, broader ethical considerations including privacy, consent, and responsible deployment, will be central to extending this research.

Table 6 shows that removing RRDA, HOG, or HE significantly reduces performance, confirming each component's crucial role in accuracy. Table 7 indicates consistent model performance across all folds, demonstrating excellent stability, reliability, and generalization of the proposed framework. A comparison of the suggested dataset with current datasets is shown in Table 8, which highlights enhanced performance and robustness in face recognition tasks with greater accuracy, precision, recall, and F1-score.

Table 6: Ablation study of DCNN–RRDA components

Model Variant	DCNN	HE	HOG	RRDA	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Full DCNN–RRDA (proposed)	✅	✅	✅	✅	98.54	97.46	96.97	95.88
w/o RRDA (DCNN + HE + HOG)	✅	✅	✅	❌	87.50	87.00	85.00	87.00
w/o HOG (DCNN + HE + RRDA)	✅	✅	❌	✅	85.20	84.00	84.00	84.00
w/o HE (DCNN + HOG + RRDA)	✅	❌	✅	✅	82.00	81.00	80.00	81.00
w/o HOG & RRDA (DCNN + HE only)	✅	✅	❌	❌	71.00	76.00	28.00	41.00
DCNN only (no HE, no HOG, no RRDA)	✅	❌	❌	❌	67.00	75.00	19.00	3.00

Table 7: Five-fold cross-validation performance consistency

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
1	98.54	97.46	96.97	95.88
2	98.54	97.46	96.97	95.88
3	98.54	97.46	96.97	95.88
4	98.54	97.46	96.97	95.88
5	98.54	97.46	96.97	95.88
Mean ± SD	98.54 ± 0.00	97.46 ± 0.00	96.97 ± 0.00	95.88 ± 0.00

Table 8: Comparative performance of proposed and existing datasets

Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed Datasets [19]	**92.54**	**93.46**	**94.97**	**97.88**
Datasets [22]	76	78	58	61
Datasets [23]	84	84	46	5

5. Conclusion

Facial recognition is a biometric technology that utilizes computer vision and AI to recognize or verify people by their visual traits. The goal of this study was to overcome the constraints in face recognition by employing a DCNN-RRDA. The data consisted of a variety of images from a face detection dataset. The preprocessing section included HE to equalize for illumination and contrast enhancement. The next step, a HOG determined the edge orientations and face geometry. The Multi-layered DCNN learned the faces from the developed descriptors while the hyper parameters and weights (s) were tuned using the RRDA so that a more robust accuracy in recognition is achieved. The DCNN-RRDA algorithm is not real-time adaptive, requiring significant computing power for its large-scale or real-time implementation face recognition systems. Future studies utilize the DCNN-RRDA method for live surveillance systems, enhancing the scalability and efficiency of low-power devices.

References

[1] Mahato, S. and Neethirajan, S., (2024). Integrating artificial intelligence in dairy farm management− biometric facial recognition for cows. *Information Processing in Agriculture*.https://doi.org/10.1016/j.inpa.2024.10.001

[2] Nosrati, L., Bidgoli, A.M. and Javadi, H.H.S., (2024). Identifying people's faces in smart banking systems using artificial neural networks. *International Journal of Computational Intelligence Systems*, 17(1), p.9. https://doi.org/10.1007/s44196-023-00383-7

[3] Beretta, E., Voto, C. and Rozera, E., (2024). Decoding faces: Misalignments of gender identification in automated systems. *Journal of Responsible Technology*, 19, p.100089. https://doi.org/10.1016/j.jrt.2024.100089

[4] Bharadwaj, S., Amin, P., Ramya, D.J. and Parikh, S., (2024). Reliable human authentication using AI-based multibiometric image sensor fusion: Assessment of performance in information security. *Measurement: Sensors*, 33, p.101140. https://doi.org/10.1016/j.measen.2024.101140

[5] Cherian, S., Joseph, J., Thomas, B., & Jose, J. (2024). Navigating the New Normal: A Bibliometric Exploration of Masked Face Recognition Research. *Informatica*, *48*(22). https://doi.org/10.31449/inf.v48i22.6342

[6] Liu, T. (2024). Secure Face Recognition Using Fully Homomorphic Encryption and Convolutional Neural Networks. *Informatica*, *48*(18). https://doi.org/10.31449/inf.v48i18.6396

[7] Abdul-Al, M., Kyeremeh, G.K., Qahwaji, R., Ali, N.T. and Abd-Alhameed, R.A., (2024). A Novel Approach to Enhancing Multi-Modal Facial Recognition: Integrating Convolutional Neural Networks, Principal Component Analysis, and Sequential Neural Networks. *IEEE Access*.https://doi.org/10.1109/ACCESS.2024.3467151

[8] Jin, B., Cruz, L. and Gonçalves, N., (2020). Deep facial diagnosis: deep transfer learning from face recognition to facial diagnosis. *IEEE Access*, 8, pp.123649-123661. https://doi.org/10.1109/ACCESS.2020.3005687

[9] Bai, Y., Liu, L., Liu, K., Yu, S., Shen, Y. and Sun, D., (2024). Non-intrusive personal thermal comfort modeling: A machine learning approach using infrared face recognition. *Building and Environment*, 247, p.111033. https://doi.org/10.1016/j.buildenv.2023.111033

[10] Nguyen-Tat, B.T., Bui, M.Q. and Ngo, V.M., (2024). Automating attendance management in human resources: A design science approach using computer vision and facial recognition. *International Journal of Information Management Data Insights*, 4(2), p.100253. https://doi.org/10.1016/j.jjimei.2024.100253

[11] Ghahremani, Y. and Maihami, V., (2024). Face Recognition Based on Deep Learning Convolutional Neural Network in Cloud Internet of Things Environment. Research square. https://doi.org/10.21203/rs.3.rs-5369040/v1

[12] Melzi, P., Tolosana, R., Vera-Rodriguez, R., Kim, M., Rathgeb, C., Liu, X., DeAndres-Tame, I., Morales, A., Fierrez, J., Ortega-Garcia, J. and Zhao, W., (2024). FRCSyn-onGoing: Benchmarking and comprehensive evaluation of real and synthetic data to improve face recognition systems. *Information Fusion*, 107, p.102322. https://doi.org/10.1016/j.inffus.2024.102322

[13] Anusudha, K., (2024). Real-time face recognition system based on YOLO and InsightFace. *Multimedia Tools and Applications*, 83(11), pp.31893-31910. https://doi.org/10.1007/s11042-023-16831-7

[14] Xu, X., Deng, H., Wang, Y., Zhang, S., & Song, H. (2024). Boosting cattle face recognition under uncontrolled scenes by embedding enhancement and optimization. *Applied Soft Computing*, *164*, 111951. https://doi.org/10.1016/j.asoc.2024.111951

[15] Archana, P., Divyabharathi, P., Balaji, S.R., Kumareshan, N., Veeramanikandan, P., Naitik, S.T., Rafi, S.M., Nandankar, P.V. and Manikandan, G., (2022). Face recognition-based vehicle starter using machine learning. *Measurement: Sensors*, 24, p.100575. https://doi.org/10.1016/j.measen.2022.100575

[16] Bakariya, B., Singh, A., Singh, H., Raju, P., Rajpoot, R. and Mohbey, K.K., (2024). Facial emotion recognition and music recommendation system using CNN-based deep learning techniques. *Evolving Systems*, 15(2), pp.641-658. https://doi.org/10.1007/s12530-023-09506-z

[17] Meena, G., Mohbey, K.K., Indian, A., Khan, M.Z. and Kumar, S., (2024). Identifying emotions from facial expressions using a deep convolutional neural network-based approach. *Multimedia Tools and Applications*, 83(6), pp.15711-15732. https://doi.org/10.1007/s11042-023-16174-3

[18] Arshed, M.A., Mumtaz, S., Ibrahim, M., Dewi, C., Tanveer, M. and Ahmed, S., (2024). Multiclass AI-generated deepfake face detection using patch-wise deep learning model. *Computers*, 13(1), p.31. https://doi.org/10.3390/computers13010031

[19] Dataset: https://www.kaggle.com/datasets/apollo2506/facial-recognition-dataset

[20] Gwyn, T., Roy, K. and Atay, M., (2021). Face recognition using popular deep net architectures: A brief comparative study. *Future Internet*, 13(7), p.164. https://doi.org/10.3390/fi13070164

[21] Riaz, W., Ji, J., & Ullah, A. (2025). TriViT-Lite: A Compact Vision Transformer–MobileNet Model with Texture-Aware Attention for Real-Time Facial Emotion Recognition in

Healthcare. *Electronics*, *14*(16), 3256. https://doi.org/10.3390/electronics14163256
[22] Dataset: https://www.kaggle.com/datasets/atulanandjha/lfwpeople
[23] Dataset: https://www.kaggle.com/datasets/jessicali9530/celeba-dataset