Neural Computing and Applications
https://doi.org/10.1007/s00521-024-10447-0

ORIGINAL ARTICLE

# Automated face recognition using deep learning technique and center symmetric multivariant local binary pattern

J. C. Sekhar[1] · P. Joel Josephson[2] · A. Chinnasamy[3] · M. Maheswari[4] · S. Sankar[5] · Ruth Ramya Kalangi[6]

Received: 21 October 2023 / Accepted: 11 September 2024


## Abstract

Researchers have recently created several deep learning strategies for various tasks, and facial recognition has made remarkable progress in employing these techniques. Face recognition is a noncontact, nonobligatory, acceptable, and harmonious biometric recognition method with a promising national and social security future. The purpose of this paper is to improve the existing face recognition algorithm, investigate extensive data-driven face recognition methods, and propose a unique automated face recognition methodology based on generative adversarial networks (GANs) and the center symmetric multivariable local binary pattern (CS-MLBP). To begin, this paper employs the center symmetric multivariant local binary pattern (CS-MLBP) algorithm to extract the texture features of the face, addressing the issue that C2DPCA (column-based two-dimensional principle component analysis) does an excellent job of removing the global characteristics of the face but struggles to process the local features of the face under large samples. The extracted texture features are combined with the international features retrieved using C2DPCA to generate a multifeatured face. The proposed method, GAN-CS-MLBP, syndicates the power of GAN with the robustness of CS-MLBP, resulting in an accurate and efficient face recognition system. Deep learning algorithms, mainly neural networks, automatically extract discriminative properties from facial images. The learned features capture low-level information and high-level meanings, permitting the model to distinguish among dissimilar persons more successfully. To assess the proposed technique's GAN-CS-MLBP performance, extensive experiments are performed on benchmark face recognition datasets such as LFW, YTF, and CASIA-WebFace. Giving to the findings, our method exceeds state-of-the-art facial recognition systems in terms of recognition accuracy and resilience. The proposed automatic face recognition system GAN-CS-MLBP provides a solid basis for accurate and efficient face recognition, paving the way for biometric breakthroughs and growing the security and ease of many applications.

**Keywords** Deep learning (DL) · Automated face recognition · Generative adversarial networks (GAN) · Center symmetric multivariable local binary pattern (CS-MLBP)

## 1 Introduction

Human facial expressions are mainly recognized by means of facial emotion recognition (FER) methods. Emotions come in a variation of forms, not all of which are obvious to the human eye [1]. Thus, recommendations of any type can help classify the classification with the support of suitable processes. A wide range of common facial emotions is included in the FER sector, including neutrality, happy, surprise, fear, wrath, grief, and disgust [2]. The dynamic area of retrieving emotions from face cues is presently being examined in psychology, psychiatry, and mental health [3]. Numerous applications, such as augmented reality, smart home, medical, and human–robot interaction (HRI), have great potential for automated facial expression recognition. Since FER comes with a lot of conventions, the common of researchers use it [4]. Creating emotion-specific features is a challenging procedure due to a number of reasons, including multidimensional data, nonlinear interactions among distinct pieces of evidence, and the variation of ways that emotions might arise rapidly in dissimilar contexts [5, 6].

The demand for reliable and accurate systems that can repeatedly categorize and organize human emotions from

Extended author information available on the last page of the article

facial expressions is increasing. Human communication is suggestively influenced by emotions, and this influence has a wide range of applications in domains such as mental health evaluation and human–computer interface [7]. Existing FER models frequently suffer from common difficulties such as noise, feature extraction difficulties, and generalization difficulties. Nonetheless, by distinctive the complex nonlinear features exposed in multimodal data, advances in machine learning (ML) and deep networks provide an attainable option to these challenges. Feature extraction and classification are the two most important methods for detecting emotions [8]. Deep learning (DL) systems, ML, and artificial neural networks are the best options for feature classification, as they provide superior accuracy [9]. The basic goal of many feature engineering and machine learning techniques is to remove complex and nonlinear patterns from multivariate time series data [10].

Expression categorization, feature extraction, and facial and face component recognition are the three main processes in traditional FER approaches [11]. First, from an input image, a face image may be recognized, and from face regions, landmarks or facial characteristics (such as the nose and eyes) can be recognized. Subsequently, a number of temporal and spatial features were obtained from facial components [12]. Thirdly, using the retrieved features, the pretrained FE methods (SVM, random forest, AdaBoost) produce the detection results. Traditional techniques have the drawback of independent feature extraction and classification processes. Improving the system's performance is therefore difficult. End-to-end learning procedures are employed by DL networks to resolve issues with previous methods. In deep learning, the larger the dataset, the better the performance is achieved [13–15]. The authors included noise translations, scaled approaches to increase data size, cropped, normalized, and augmented data to boost the DL performance. Convolutional neural networks (CNNs) are the most effective method for classifying and segmenting data. One of its main advantages is that it can extract features automatically. A DL concept called transfer learning (TL) makes it possible to apply knowledge gained for a particular position to another. Time-saving and increased accuracy were the two key advantages of TL. As many FER techniques as there have been created and published in the literature, most of them do not deal with the hyperparameter tuning process [16, 17]. The identification performance of the DL models is greatly impacted by the values of the hyperparameters; therefore, selecting the appropriate ones is important.

### 1.1 Motivation of this research

Biometric authentication revolutionizes with automated face recognition using deep learning, specifically employing GANs and the CS-MLBP procedure. This innovative method ensures high accuracy and efficiency in identifying individuals from facial data by harnessing the capabilities of GAN for feature extraction and CS-MLBP for robust pattern recognition. This opens up possibilities for advanced security systems and accelerated authentication procedures.

### 1.2 The main contribution of this research

- To improve the existing face recognition algorithm, investigate extensive data-driven face recognition methods, and propose a unique automated face recognition methodology based on GAN and CS-MLBP.
- In this study, research initially extracts the facial texture features using the CS-MLBP technique. This discusses the problem that C2DPCA performs a great job of eliminating the global characteristics of the face but has trouble processing the local features of the face under large samples.
- Finally proposed approach, GAN-CS-MLBP, combines the power of GAN with the robustness of CS-MLBP, resulting in an accurate and efficient face recognition system.

### 1.3 Research question

- What are the specific advantages of using GANs in conjunction with CS-MLBP compared to traditional methods in face recognition, such as eigenfaces or convolutional neural networks (CNNs)?
- How does the integration of generative adversarial networks (GANs) with center symmetric multivariable local binary pattern (CS-MLBP) enhance the accuracy and robustness of automated face recognition systems?
- How does the choice of GAN architecture (e.g., DCGAN and StyleGAN) impact the performance of the face recognition system when combined with CS-MLBP?
- To what extent does the incorporation of CS-MLBP develop the resilience of the face recognition system to common challenges such as variations in illumination, pose, and facial expression?

### 1.4 Problem statement

To overcome these problems, there is an urgent need for a novel strategy that uses cutting-edge technologies like generative adversarial networks (GANs) and center symmetric multivariable local binary pattern (CS-MLBP) to develop the robustness and accuracy of face recognition systems. This project seeks to create an automated face recognition methodology that combines GAN and CS-MLBP methodologies to obtain superior performance in identifying faces under a variety of environmental and contextual settings. The suggested methodology aims to address shortcomings in existing face recognition systems by effectively capturing detailed facial traits, improving discriminative capabilities, and limiting the influence of typical real-world difficulties. This study aims to advance the advanced in automated face recognition by combining the power of GAN for data augmentation and feature enhancement with the discriminative abilities of CS-MLBP, resulting in developed security, surveillance, and biometric authentication applications.

## 2 Literature survey

Abbas et al. [18] distinctive a facial recognition (FR) technique called AFR-Conv that works a convolutional mixer to deal with problems with face occlusion. By assigning weight to each face patch based on importance, combining residual connections with an AdaBoost classifier, a unique architecture for AFR-Conv is created that can routinely classify faces. The benefits of pretrained CNNs are also used by AFR-Conv through the usage of ResNet-50, Inception-v3, and DenseNet-161. Utilizing measurements for precision, recall, detection accuracy, and F1-score, the experimental findings of AFR-Conv are provided. For instance, the recommended method accomplishes a 95.5% accuracy rate on a dataset with 8500 face images.

A whole FR network was developed by Karasugi et al. [19] that can dependably retain accuracy across dissimilar facial demos, even in the presence of mask-obscured images. Their methodology involved creating synthetic datasets by comparing face masks to widely used public datasets that included LFW, CASIA-Web Face, CFP, CPLFW, and CALFW. Then, these artificial datasets were used for testing and training. In addition, they presented a two-component model that contains of an alignment module and a feature extraction module. Deep convolutional neural networks (DCNN) are used to produce a 512-feature vector. FR deployment on heterogeneous IoT platforms was inspected by Elordi et al. [20]. Our foremost challenges are deep neural network (DNN) optimization for many Internet of Things (IoT) devices such as robots, tablets, and smartphones; biometric data management security and user privacy protection; and user-friendly interaction with facial verification systems. They research inspect numerous methods to deal with these problems and recommend a knowledge-based approach for successfully incorporating DNN-powered FR systems into IoT gadgets. To rise user engagement, our method focuses a great priority on the secure management of biometric data and provides real-time feedback. The research presents and assesses state-of-the-art DNN models for FR using hardware platforms for IoT devices from Intel and NVIDIA.

Subsequently the face serves as a fundamental means of conveying our feelings, Monica et al. [21] provide a summary of the procedure of classifying emotions through facial expressions. Thus, to develop accuracy percentage, the first stage of the recommended model uses the histogram of gradients (HOG), a feature that can be identified by deep learning methods such as the linear support vector machine (LSVM), to classify the face. Deep learning methods are then applied to examine the identified individual's emotional state. Images are scaled, noise-reduced using a mean filter, and preprocessed using a basic HAAR classifier software. In terms of emotion detection and facial recognition, the model's accuracy rate is 92%. For the purpose of identifying facial emotions, Huang et al. [22] used a DNN. This work's main goal is to identify the crucial face features of the DNN model for FER. Three neural networks were used in FER: a squeeze-and-excitation network, a convolutional neural network (CNN), and a residual neural network. CNN was trained using face expression datasets from AffectNet and real-world affective faces database (RAF-DB). When evaluated on RAF-DB, the CNN model trained exclusively using AffectNet had an accuracy of 77.37%. On the other hand, the pretrained and transferred model's accuracy was 83.37%.

Reddi et al. [23] presented a simple approach to facial expression recognition using few-parameter CNNs and transfer learning. The proposed CNN architecture was simultaneously trained for real-time detection on the FER-2013, JAFFE, and CK + datasets, broadening the range of emotional expressions that could be identified. The model successfully identified emotions such as happiness, sadness, surprise, fear, anger, disgust, or neutrality. A method for Henry gas solubility optimization with deep learning-based FER (HGSO-DLFER) was designed and described by AlEisa et al. [24] for human–computer interaction (HCI). Classifying and organizing dissimilar kinds of facial emotions is the impartial of the HGSO-DLFER method. The adaptive fuzzy filtering (AFF) used in the HGSO-DLFER technique for noise removal enables this. Moreover, feature vector improvement is carried out using the MobileNet model, and the hyperparameter scan is about choice by the HGSO algorithm. Using an auto encoder (AE) classifier with a Nadam optimizer, the HGSO-DLFER method distinguishes face emotions. To assistance with a better understanding of the FER consequences by using the HGSO-DLFER method, a comprehensive experimental analysis is accomplished.

The impartial of Kundu et al. [25] is to use CNN to generate a FER system. They emphasize the importance it has been to analyze human emotions as a nonverbal communication tool throughout human existence. Their aim is to improve a real-time detection system by using CNNs and computer vision (CV) libraries. This technique uses speech and facial expression data to track human emotions more efficiently. Wang et al. [26] presented a new unsupervised domain-specific data augmentation method aimed at developing the recognition accuracy of multimodal face recognition systems. This method leverages information from multiple sources to accomplish its objective. The proposed technique firstly creates realistic synthetic data by preserving the data distribution of each modality in a domain-specific manner. It then creates novel examples using generative models. The efficacy of this approach is established by experimental findings on multiple benchmark datasets.

Zhang et al. [26] present a meta-learning architecture designed precisely for face recognition to solve the domain gap problematic caused by modifications in illumination, expression, and position. The recommended technique uses domain-specific calibration to dynamically modify the feature representation particular to each domain once meta-learning acquires a domain-invariant feature representation. The experimental findings display that the recommended method outperforms many benchmark datasets.

### 2.1 Limitations of existing system

- These systems may misidentify due to differences in accuracy caused by things like occlusions, illumination, angles, and facial expressions.
- Facial recognition technology founds significant privacy concerns as it can be used without individuals' acceptance, possibly leading to misuse or prohibited surveillance.
- The legal and regulatory framework surrounding facial recognition technology is continually enhancing, generating uncertainty about its ethical application, data protection, and accountability.
- Spoofing attacks, in which attackers provide modified or synthesized facial images, are able to deceive automated face recognition systems.

## 3 Proposed system

This section describes a revolutionary automated facial recognition system based on the GAN and CS-MLBP. The features of the facial texture are extracted in the major step of this assignment utilizing the CS-MLBP method. The global features found by C2DPCA are combined with the recovered texture features to harvest a comprehensive set of multiple components. The recommended method, GAN-CS-MLBP, combines the advantages of GAN and CS-MLBP to create an accurate and efficient face recognition system. Automatic discriminative feature extraction from facial images is achieved using deep learning techniques, mostly neural networks. The learned features enable the model to distinguish between many people more effectively by capturing low-level information and high-level meanings. The block diagram for the GAN-CS-MLBP approach is displayed in Fig. 1.

### 3.1 Dataset

This section provides an overview of face datasets, encompassing images, videos, and diverse faces, essential to deep learning. Additionally, we showcase the results of deep learning algorithms on various databases, such as CASIA-WebFace, YTF, and LFW.

#### 3.1.1 LFW dataset

A database of images of faces called LFW (Labeled Faces in the Wild) was established to look at the issue of unrestricted face recognition [27]. For the purpose of gathering data, almost 13,000 face images were collected from the internet. There are 1680 subjects in the collection, and there are two or more distinct images for each subject. Only participants with a minimum of 70 face samples were taken for consideration in this experiment (Fig. 2).

#### 3.1.2 YouTube faces dataset

There are 3425 YouTube videos in total in the collection. A portion of the celebrities at the LFW is featured in the videos [29]. There are 15,955 people in the videos. We used still images from the video in our investigation (Fig. 3).

#### 3.1.3 CASIA-WebFace dataset

The training dataset for facial recognition tasks is the CASIA-WebFace dataset [31], whereas the test set is CASIA-WebFace. The CASIA-WebFace database contains 494,414 unaligned face images of 10,575 individuals, each with an image size of 256 X 256 and a variety of ages, expressions, and lighting conditions. CelebA features 202,599 portraits of 10,177 celebrities. However, as shown in Fig. 4, there are times when the identity label is inaccurate, and image quality in CASIA-WebFace varies substantially.

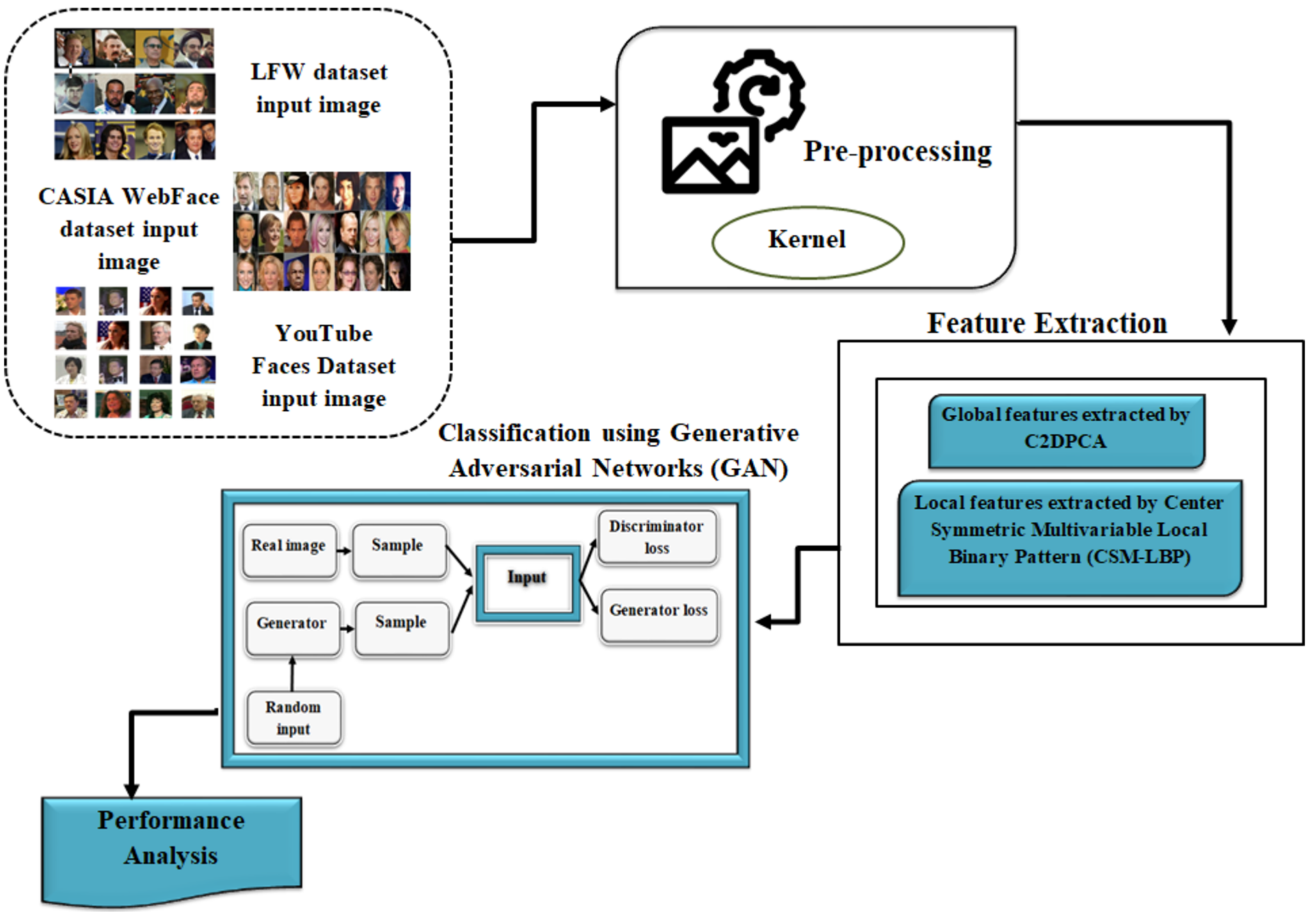


**Fig. 1** Proposed method of GAN-CS-MLBP

### 3.2 Preprocessing

In image processing, the goal of applying MF to input images is to remove noise, which improves data quality [33]. This important preprocessing method is used to decrease noise and improve the overall quality of the data. MF approaches, in contrast to other smoothing techniques, are effective at preserving edges and small features while effectively reducing impulsive noise, such as the salt-and-pepper noise that is typically produced in images. It is especially useful in scenarios where noise reduction is essential because of its technique, which maintains the integrity of important image features while replacing all pixel values with the median value from a neighboring neighborhood. In order to develop the quality and resilience of the data before further display or analysis, MF is widely used in domains including medical imaging, remote sensing, and CV.

### 3.3 Multi feature fusion

Face recognition technology has multiple elements that represent different human face characteristics. By efficiently employing these aspects, FR will progress. The multifeatured fusion method uses the fusion feature, which involves extracting facial features at various scales and resolutions. To improve overall performance, it fuses local and global elements. In this study, recognition performance in the vast data environment has been enhanced by the fusion features created by integrating various features from global parts obtained using the C2DPCA technique and local features retrieved using the LBP algorithm. Face recognition techniques employ international and local feature extraction methods, two different feature extraction methods. PCA, LDA, and C2DPCA are the main methods for extracting global features. Global features can recognize an images texture, shape, and other specifics. However, resolving the problem of excessive processing brought on by the bulk of the image data can take time and effort. Reduced processing and simple extraction of local features make them useful for recognition. Combining

**Fig. 2** Several LFW dataset examples of faces [28]

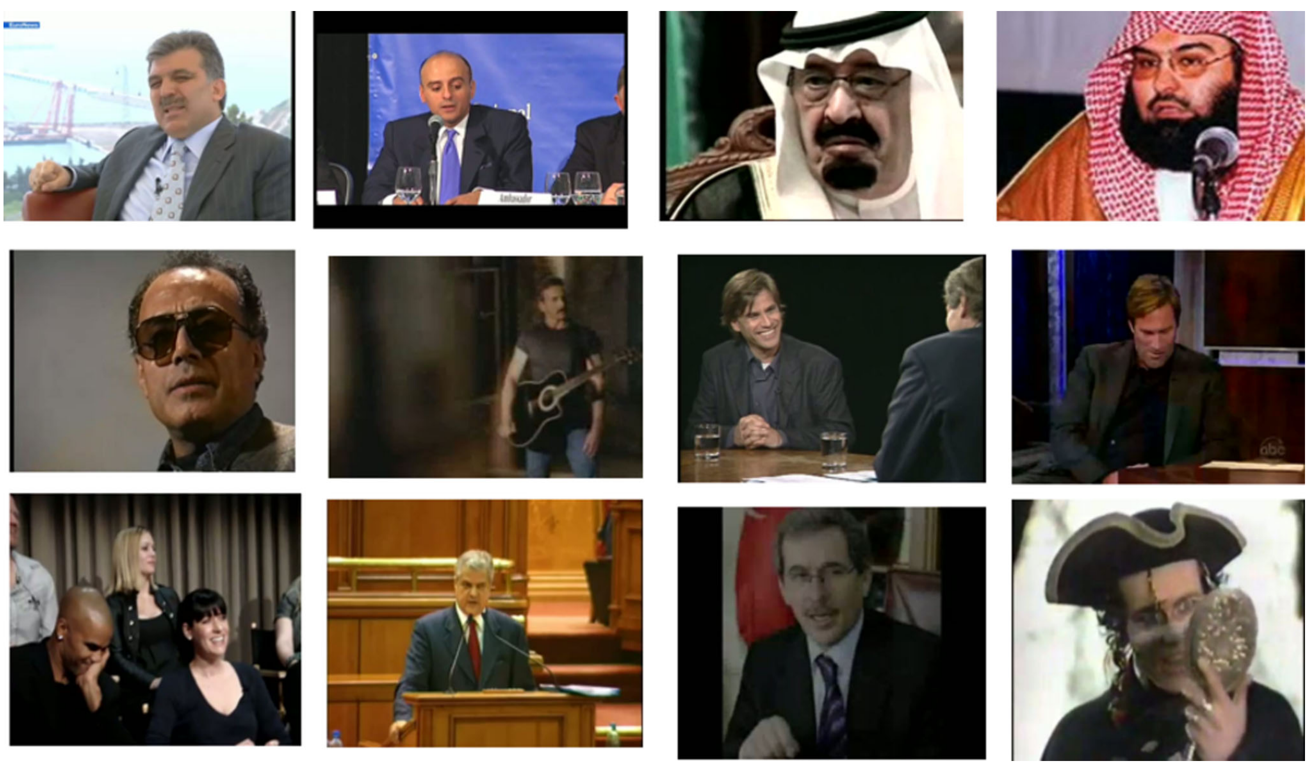

**Fig. 3** Several samples of face images from the YTF dataset [30]

different local elements can also improve facial recognition accuracy. The major drawback of local features is that they are overly susceptible to them, making them weak against local lighting, expression, posture, and other

**Fig. 4** Several representative face images from the CASIA-WebFace dataset [32]

circumstances. It is crucial to consider combining a part that offers both benefits to integrate the help of local and global features. Next, this study describes extracting international and local features before presenting the feature fusion approach.

#### 3.3.1 C2DPCA to extract global facial features

Particularly, the covariance matrix can only be approximated using the singular value decomposition (SVD) methodology. This results from the PCA method's initial step, which drastically raises the computing complexity by reducing the two-dimensional matrix to a single dimension. The image covariance array is computed by the C2DPCA method directly from the initial image array [34]. The resulting covariance array is more accurate than PCA, although being less in size. Additionally, and perhaps more critically, building the eigenvector of the image covariance matrix takes less time to compute. Large training samples are more suited for C2DPCA.

The typical projection vector YofH, where is the unitized column vector, can be produced using C2DPCA by projecting the image matrix $I = M \times S$ onto Y using the linear transformation $Z = IS$. Using C2DPCA, the optimal projection space should be found. The definition criteria serves the following purposes:

$$K(Y) = \mathrm{us}(R) \tag{1}$$

$$R = E[(I - EI)Y][(I - EI)Y]^U \tag{2}$$

where $R_i$ is the eigenvector's covariance matrix produced following the projection of the training example. The trace of Q is $\mathrm{us}(R)$. The following describes the image's total walk matrix:

$$H = E\left[(I - EI)^U (I - EI)\right] \tag{3}$$

Consequently, the overall walking matrix of the training sample is

$$H = \frac{1}{L}\sum_{i=1}^{Q}\sum_{j=1}^{R}\left(T_j^i - T\right)^U\left(T_j^i - T\right),\ i = 1, 2, \ldots,\quad Q;\ j = 1, 2, \ldots, R \tag{4}$$

$$T = \frac{1}{I}\sum_{i=1}^{M}\sum_{j=1}^{L} T_j^i \tag{5}$$

where Q is the entire amount of training examples, Q is the total amount of examples of each type, $T$ is the overall mean of the training examples, $T_j^i$ is sample j of class i, and $L = Q \times R$, is the entire number of training examples. H is undoubtedly a nonnegative definite matrix of the kind $S \times S$. The matrix of generic walking criteria is then (6) represented as:

$$K(I) = I^U HI \tag{6}$$

$I_i$ s is the vector of normalized columns. After projecting the training image matrix to I, the walking matrix of the best projection axis reaches its maximum, despite the fact that the criterion function maximizes the normalized eigenvector $I_i$ s. In other words, the greatest eigenvalue is used to calculate the eigenvector of H. Standard orthogonal projection axes $I_1, I_2, \ldots, I_Q$ are typically discovered to be the optimal projection axes in actual applications. Create a linear combination as the ideal space, or the principal component, by computing the eigenvalues of G, choosing the first $q(q << N)$ bigger eigenvalues based on the contribution rate, and calculating the related eigenvector $Q_1, Q_2, \ldots, Q_q$. The ideal area maximizes the function criteria.

### 3.4 Multivariate local binary pattern

Arco Lucifer [35] created the multivariate local binary pattern operator, which represents the intricate local interactions between pixels in three distinct bands. It takes into account both spatial interactions between rounds as well as interactions within a single round. As a result, the local neighbors visible in all three rounds make up the neighborhood set for a given pixel.

$$MLBP = \sum_{i=1}^{Q=1} \begin{array}{r} \mathrm{sign}(h_i^{c1} - h_d^{c1}) + \mathrm{sign}(h_i^{c2} - h_d^{c1}) \\ +\mathrm{sign}(h_i^{c3} - h_d^{c1}) + \mathrm{sign}(h_i^{c1} - h_d^{c2}) \\ +\mathrm{sign}(h_i^{c2} - h_d^{c2}) + \mathrm{sign}(h_i^{c3} - h_d^{c2}) \\ +\mathrm{sign}(h_i^{c1} - h_d^{c3}) + \mathrm{sign}(h_i^{c2} - h_d^{c3}) \\ +\mathrm{sign}(h_i^{c3} - h_d^{c3}) \end{array} \tag{7}$$

Nine potential combinations of these bands, chosen from Eq. 7, are used to calculate the local threshold. The operator describing a local color texture is the outcome of this. Computed across an entire or partial image, the color texture measure is the histogram of $MLBP_d$ occurrence. This single distribution has $Q \times 3^2$ bins (P = 8 produces 72 bins, for example).

#### 3.4.1 Center symmetric multivariable local binary pattern

The CS-MLBP is an additional modified variant of MLBP. The model for the PASCAL database's object recognition was built by Marko Heikkila. The original MLBP, a highly wordy feature, could be better for flat graphics. Rather than comparing each individual pixel's gray level value to the center pixel, this approach evaluates pairs of symmetric center pixels. The gradient operator is interlaced with the basic ideas of CS-MLBP. It clarifies the differences in grayscale between neighboring pairs of opposite pixels. Because of this, CS-MLBP makes use of both MLBP and gradient-based features. The boundaries and different textures are also noticed.

$$CS - LBP_{q,s,u} = \sum_{i=0}^{M/2-1} t(|h_i - h_i(M/2)|)2^i, \; t(y) = \begin{cases} 1 & \text{if } y \geq u \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Here, $h_i$ and $h_{i+m/2}$ represent the gray levels of the center-symmetric pairs of pixels (M whole), uniformly distributed over a radius s circle. When compared to basic MLBP, it also makes computations less complex.

### 3.5 Generative adversarial networks

Two neural networks make up GAN [36]: a discriminator neural network (D) and a generator neural network (G). Figure 5 depicts the overall block diagram of the GAN used to generate fake images. By using random noise (r) as input, the generator seeks to create fake images, represented as G(r), that imitate the real images in the dataset (x). On the other hand, the discriminator, D, aims to distinguish among the images created by G and the actual images. The discriminator determines the likelihood that a sample is taken from the actual dataset instead of the artificial images created by G by analyzing both real and fake images.

When the discriminator determines that an image is real, it will provide a probability value of 1, and when it finds an imitation image, it will produce a value of 0. The generator attempts to lower the discriminator's accuracy, but the discriminator's objective is to categorize an image correctly as many times as possible when it receives it as input. Both networks are thus in competition with one another to determine who can more successfully complete their specific goal. The discriminator receives training aimed at optimizing its ability to discriminate among real and fake images, while the generator is trained to reduce 1-D(G(r)), or the probability that the false images it produces would be identified as such by the discriminator. The two networks thus engage in a minimax game with one another, which is represented mathematically by the value function displayed in (9).

$p_r(r)$ indicates the distribution of noise (r), whereas $p_{data}(x)$ illustrates the data distribution of true images (x).

$$\min_G \max_D V(D,G) = \mathbb{R}_{x \sim p_{data}^{(x)}}[\log D(x)] + \mathbb{R}_{r \sim p_r^{(r)}}[\log(1 - D(G(z)))] \tag{9}$$

After completing the required training, the generator will be able to use noise signals 'n' to create artificial images that look natural and realistic, and the ability of 'D' to distinguish between deep fake and actual images will also increase as shown in Table 1.

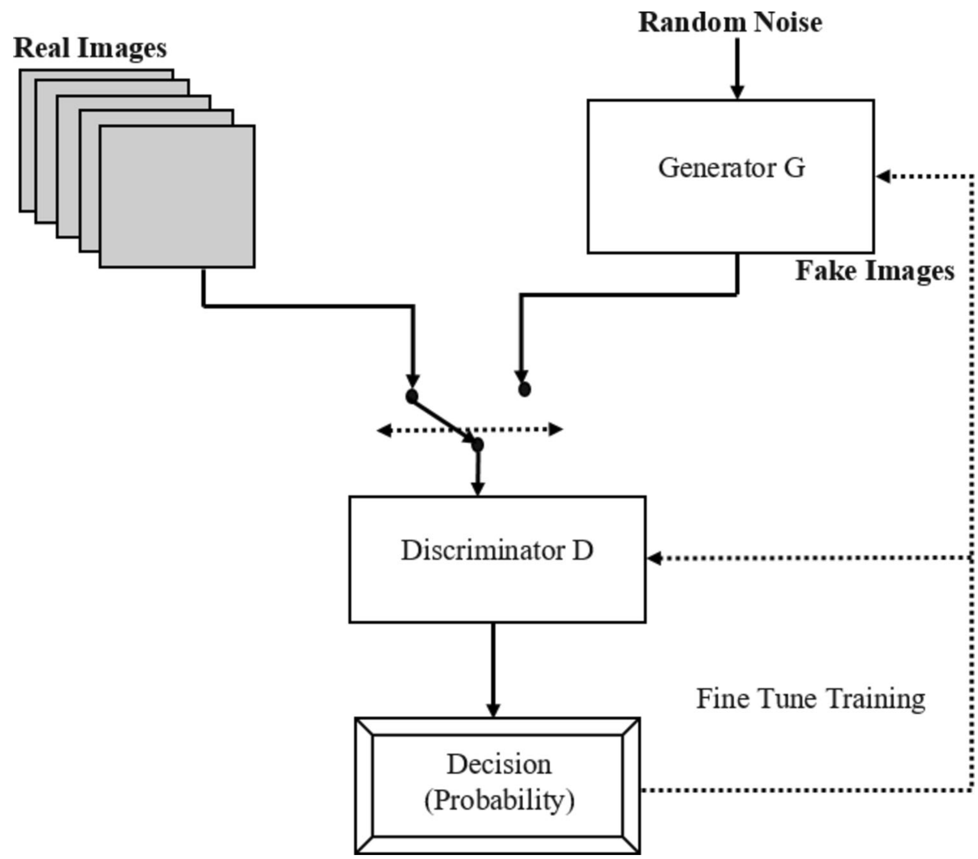


**Fig. 5** Architecture of generative adversarial network (GAN)

**Table 1** Description of mathematical symbol

| Symbol | Description |
|---|---|
| $Z = IS$ | Linear transformation |
| $I = M \times S$ | Image matrix |
| $R_i$ | Eigenvector's |
| Q | Entire amount |
| $L = Q \times R,$ | Total number of training samples |
| $I_i$ s | Vector of normalized columns |
| s | Circle of radius |
| G | Generator neural network |
| D | Discriminator neural network |
| r | Random noise |
| G(r) | Fake images |
| x | Real image dataset |
| $p_{data}(x)$ | Data distribution |

#### 3.5.1 Proposed model environmental setup

**Generator Network**: This network generates fake data from random noise

Input: Random noise vector (e.g., 100-dimensional)

Layers: Dense layer followed by a reshape layer to form a low-resolution image

Transposed convolutional (deconvolutional) layers, often with batch normalization and ReLU activation

Output layer with a tanh activation function to produce an image

**Discriminator Network:** This network tries to distinguish among real and fake data

Input: Image (e.g., $64 \times 64$ or $128 \times 128$ pixels)

Layers: Convolutional layers, often with Leaky ReLU activation and dropout for regularization. Dense layer followed by a sigmoid activation function to output a probability

**Loss Functions**

Generator Loss: Binary cross-entropy loss between the discriminator's output for generated images and the label of real (1)

Discriminator Loss: Binary cross-entropy loss for real images labeled as real (1) and generated images labeled as fake (0)

**Optimization Algorithms**

Typically, Adam optimizer is used with parameters:

Learning rate: 0.0002, Beta1: 0.5, Beta2: 0.999

| Hyperparameters |
|---|
| Noise vector size: 100, Batch size: 64, Number of epochs: 200, Learning rate: 0.0002 |

While the research focuses on the combination of generative adversarial networks (GANs) and center symmetric multivariable local binary pattern (CS-MLBP) for automated face recognition, it does not discuss potential adversarial attacks or responses. Adversarial assaults are a serious threat to face recognition systems because they can modify input data and fool the system's decision-making process. Adding noise, occlusion, disguise, or developing synthetic faces to deceive the identification system may jeopardize the suggested method's reliability and efficacy. There are various tactics that can be used to strengthen the facial recognition system's resilience against malicious attacks. These contain of:

- *Adversarial Training* By including adversarial instances in the training process, the model can be trained to categorize and counter probable attacks.
- *Robust Feature Extraction* The robustness of the system can be enhanced by employing more resilient feature extraction methods that are less susceptible to adversarial perturbations.
- *Defense Mechanisms* Adversarial attacks can be recognized and lessened by putting defense mechanisms like input decontamination, feature randomization, and model concentration into practice.
- *Data augmentation* By adding varied adversarial samples to the training set, the model's capability for simplification and face recognition in a range of scenarios can be enhanced.
- *Ensemble Methods* By lessening the effect of adversarial cases, combining several recognition models or classifiers helps increase the robustness of the system.

### 3.6 Strengths of the proposed method

- GANs can create synthetic data, removing the necessity on large-scale real-world datasets for training purposes. This is mostly helpful when labeled datasets are scarce or unapproachable, as the GAN can accompaniment the training data.
- GAN-generated images are more resilient to adversarial attacks than standard methods, hence increasing the face recognition system's security against management efforts.
- GANs may generate synthetic data that closely matches actual photos to train the recognition system with different facial traits. CS-MLBP retrieves detailed facial patterns with feature extraction. Even under adverse lighting, facial expressions, and occlusions, this combination improves face recognition accuracy.

### 3.7 Weakness of the proposed method

- Face recognition technology, particularly when used in automated systems, raises serious privacy concerns. There is a possibility of sensitive personal data being misused or unauthorized access, thus violating privacy rights and raising surveillance issues.
- The use of facial recognition technology in many contexts raises ethical concerns about permission, autonomy, and the possibility of discrimination or bias, especially against underprivileged groups. GAN-based algorithms may mistakenly learn and perpetuate biases in training data.
- GAN-based models are inherently complex and difficult to interpret. This lack of explainability can be problematic, particularly in applications requiring openness and interpretability, such as legal or forensic situations.

### 3.8 Implications of the proposed model

The use of facial recognition technology poses serious privacy concerns because it collects and processes people's biometric data without their explicit agreement. Individuals are at risk of being monitored and tracked without their knowledge or consent. Ethical issues require individuals to be fully informed about the collection, storage, and use of their facial data. Obtaining informed consent from people whose faces are being identified is critical to ensure ethical use of technology. Storing face data in databases raises security concerns, such as data leaks and identity theft. Robust security measures must be in place to protect biometric data from unauthorized access or malicious usage. The growing use of face recognition technology in public places raises concerns about mass surveillance and infringement on civil freedoms. To preserve individuals' privacy and freedom of movement, precise norms and regulations governing the use of face recognition in public places are required.

## 4 Result and discussion

This section discusses a Initially, the experimental setting was explained; subsequently, the comparative methodologies were explained; and lastly, the performance metrics were explained.

### 4.1 Comparative methods

The proposed GAN-CS-MLBP is compared with a number of well-established methods, some of which are used in this study's analysis: the deep belief network (DBN) [37], the spectral regression discriminant analysis network (SRDA-Net) [38], the discriminative covariance oriented representation learning (DCRL) [39], and the vector of locally aggregated descriptor encoded DCNN feature (VLAD-DCNN) [40].

### 4.2 Performance metrics

Recognition rate analysis, equal error rate analysis, information rate, time elapsed, and accuracy are usually calculated performance metrics. These metrics are typically created from four key measurements connected to a binary classification outcome (positive/negative). These assessments include false positives (FP) and false negatives (FN), which show incorrect identifications of emotional states, in addition to true positives (TP) and true negatives (TN), which represent correctly detected emotional states.

### 4.3 Face recognition rate analysis

The face recognition rate for automated face recognition denotes to a system or algorithm's correctness and effectiveness in precisely detecting and validating persons' faces from an input image or video footage. It counts the number of times a system correctly matches a face to a known distinctiveness within a specified dataset or situation. A great face recognition rate describes a dependable and efficient face recognition system, whereas a low face recognition rate shows possible flaws or limitations in the algorithm's efficiency.

$$\text{Recognition rate} = \frac{TP + TN}{P + N} \tag{10}$$

Figure 6 and Table 2 found a contrast of the proposed GAN-CS-MLBP technique's face recognition rate to other known approaches. The graph shows that the deep learning method increases efficiency while maintaining the face recognition rate. The graph shows the face recognition rate of the proposed GAN-CS-MLBP utilizing LFW, YTF, and CASIA-WebFace datasets. For instance, the proposed GAN-CS-MLBP, with LFW dataset, has a face recognition rate of 89.325%, with YTF and CASIA-WebFace it is 89.526% and 90.526%, respectively. The proposed method has the maximum face recognition rate related to the existing approaches.

### 4.4 Equal error rate analysis (EER)

The equal error rate is a performance statistic that is regularly used in binary classification jobs and biometric security systems to evaluate the efficacy of identity or verification systems. It indicates that the false acceptance rate (FAR) and false rejection rate (FRR) are in balance. When assessing the way the system functions, this parameter is important. The most effective compromise between these two error rates is the EER, which indicates the point at which the system has an equal chance of rejecting a matching face and admitting a nonmatching face. In other words, the EER is the level of accuracy at which the system strikes a reasonable balance among

**Fig. 6** Face recognition rate analysis for GAN-CS-MLBP method with existing systems

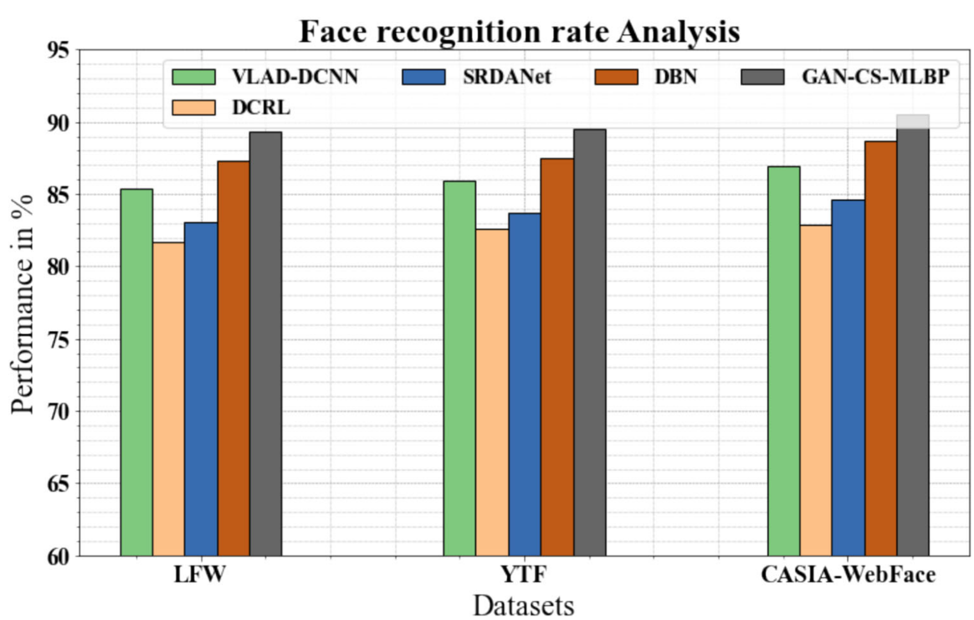

**Table 2** Face recognition rate analysis for GAN-CS-MLBP method with existing systems

| Datasets | VLAD-DCNN | DCRL | SRDANet | DBN | GAN-CS-MLBP |
|---|---|---|---|---|---|
| LFW | 85.324 | 81.726 | 83.029 | 87.314 | 89.325 |
| YTF | 85.938 | 82.633 | 83.728 | 87.526 | 89.526 |
| CASIA-WebFace | 86.938 | 82.917 | 84.627 | 88.672 | 90.526 |

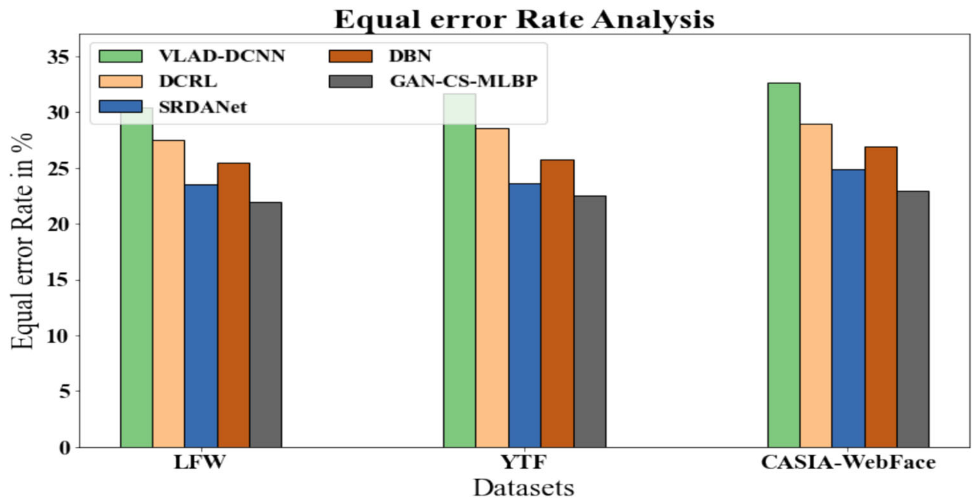


**Fig. 7** Equal error rate analysis for GAN-CS-MLBP technique with existing systems

**Table 3** Equal error rate analysis for GAN-CS-MLBP technique with existing systems

| Datasets | VLAD-DCNN | DCRL | SRDANet | DBN | GAN-CS-MLBP |
|---|---|---|---|---|---|
| LFW | 30.456 | 27.459 | 23.526 | 25.425 | 21.928 |
| YTF | 31.655 | 28.564 | 23.637 | 25.726 | 22.526 |
| CASIA-WebFace | 32.675 | 28.976 | 24.827 | 26.928 | 22.917 |

accepting impostors and rejecting legitimate users in face recognition applications.

$$\text{Equal Error Rate } (EER) = \frac{FAR + FRR}{2} \tag{11}$$

Figure 7 and Table 3 present an association between the suggested GAN-CS-MLBP method equal error rate and other established techniques. As the graph illustrates, the deep learning approach performs better while lowering the equal error rate. The graph displays the GAN-CS-MLBP's suggested equal error rate for LFW, YTF, and CASIA-WebFace datasets. For instance, the equal error rate for the GAN-CS-MLBP with LFW dataset is 21.928%, while the equal error rate for the YTF is 22.526%. Similarly, the GAN-CS-MLBP, with CASIA-WebFace dataset, has an equal error rate value of 22.917%. The proposed method has the minimal error rate compared to the existing techniques.

### 4.5 Information ratio analysis

The automated face recognition information ratio is a statistic used to assess the effectiveness of a face recognition system. It compares the amount of useful information gained from the system (such as correctly recognized faces) to the amount of irrelevant or misleading information (such as false positives or false negatives) it generates to determine the system's effectiveness in accurately identifying and verifying individuals. A more excellent information ratio suggests that the face recognition system is more trustworthy and efficient.

Figure 8 and Table 4 associate the information ratio of the proposed GAN-CS-MLBP method with other recognized techniques. The graph displays that even though the deep learning technique is more effective, the information ratio is preserved. The information ratio of the proposed GAN-CS-MLBP is displayed on the graph using the LFW, YTF, and CASIA-WebFace datasets. For instance, the YTF has an information ratio of 93.627%, while the GAN-CS-MLBP, using the LFW dataset, has an information ratio of 92.563%. Similarly, the GAN-CS-MLBP, with CASIA-WebFace dataset, has an Information ratio of 94.627%. The proposed technique has the maximum information ratio rate associated to the existing methods.

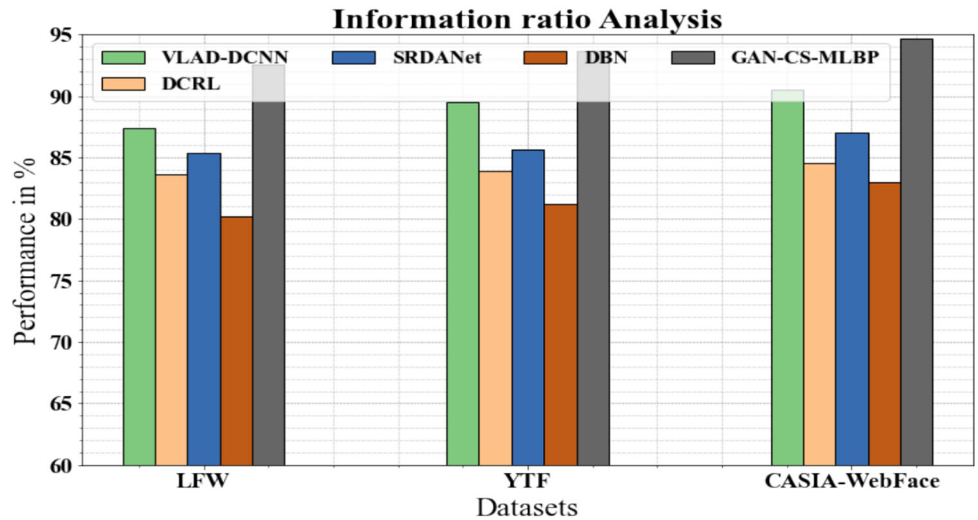


**Fig. 8** Information ratio analysis for GAN-CS-MLBP technique with existing systems

**Table 4** Information ratio analysis for GAN-CS-MLBP method with existing systems

| Datasets | VLAD-DCNN | DCRL | SRDANet | DBN | GAN-CS-MLBP |
|---|---|---|---|---|---|
| LFW | 87.425 | 83.625 | 85.324 | 80.234 | 92.563 |
| YTF | 89.526 | 83.927 | 85.672 | 81.245 | 93.627 |
| CASIA-WebFace | 90.533 | 84.526 | 86.983 | 82.938 | 94.627 |

### 4.6 Time-elapsed analysis

Measuring and analyzing the performance and efficiency of a face recognition system over a certain length of time is denoted to as time-elapsed analysis for automated face recognition. It inspects how long the approach takes to correctly classify and verify individuals' faces from a given dataset. This analysis aids in determining the system's speed, dependability, and scalability, allowing adjustments to be made to optimize its performance in real-world conditions.

The time-elapsed analysis of the suggested GAN-CS-MLBP technique is compared to the current methods in Table 5 and Fig. 9. The data unequivocally demonstrate that the GAN-CS-MLBP strategy has done better than any other approach. Using the LFW, YTF, and CASIA-WebFace datasets, the time-elapsed analysis of the recommended GAN-CS-MLBP is displayed in the graph. For example, the GAN-CS-MLBP, with LFW dataset, has an elapsed time of 2.737 s, whereas the dataset YTF has an elapsed time of 2.425 s. Likewise, the GAN-CS-MLBP, with CASIA-WebFace dataset, has an elapsed time of 3.029 s. The proposed method has the minimum elapsed time connected to the existing methods.

### 4.7 Accuracy analysis

Accuracy analysis for automated face recognition refers to determining and assessing the effectiveness and reliability of a facial recognition system in correctly categorizing individuals based on their facial traits. It involves evaluating the system's capability to match and categorize faces against a known database or set of reference images. Metrics including recall, precision, false-positive rate, true-positive rate, and entire accuracy are included in accuracy analysis and provide information about the effectiveness of the system as well as possible practical uses. This study assistances gauges the system's accuracy levels and can guide developments in algorithm progress, training data, and complete performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

The accuracy of the GAN-CS-MLBP method is associated to numerous other approaches in Fig. 10 and

**Table 5** Time-elapsed analysis for GAN-CS-MLBP technique with existing systems

| Datasets | VLAD-DCNN | DCRL | SRDANet | DBN | GAN-CS-MLBP |
|---|---|---|---|---|---|
| LFW | 6.673 | 5.526 | 3.245 | 4.526 | 2.737 |
| YTF | 7.415 | 5.927 | 3.627 | 4.922 | 2.425 |
| CASIA-WebFace | 7.817 | 6.021 | 4.029 | 5.123 | 3.029 |

Fig. 9 Time-elapsed analysis for GAN-CS-MLBP technique with existing systems

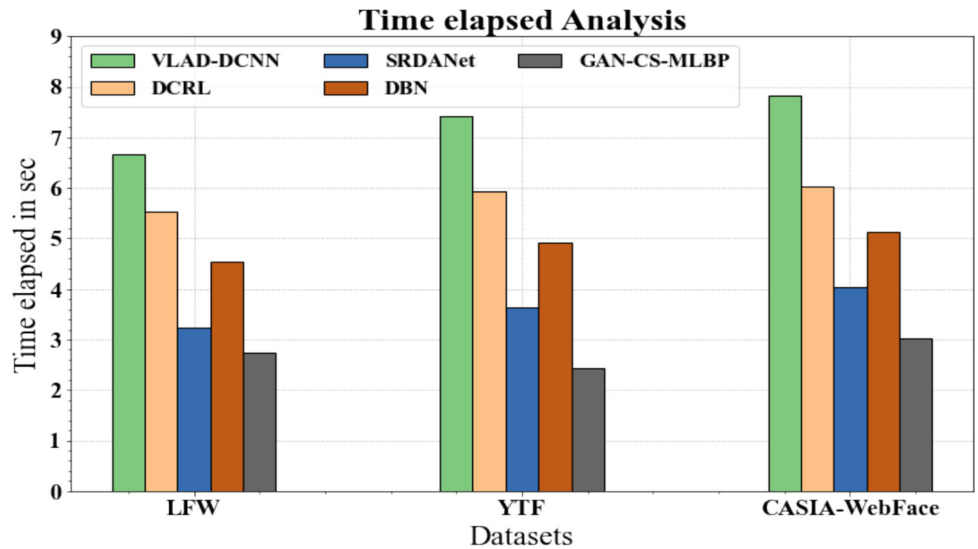


Fig. 10 Accuracy analysis for GAN-CS-MLBP technique for CASIA-WebFace dataset

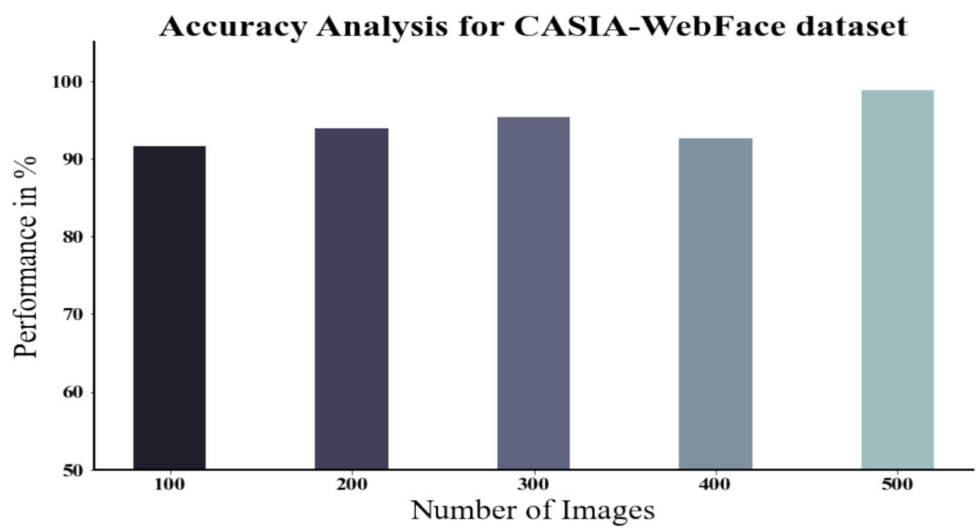


Table 6 Accuracy analysis for GAN-CS-MLBP technique for CASIA-WebFace dataset

| Number of images | Accuracy |
|---|---|
| 100 | 91.627 |
| 200 | 93.928 |
| 300 | 95.324 |
| 400 | 92.635 |
| 500 | 98.873 |

Table 6. The graph displays how accuracy is maintained while efficiency is improved with the deep learning approach. The GAN-CS-MLBP model has an accuracy of 91.627%, 93.928%, 95.324%, and 92.635% for 100, 200, 300, and 400 data. With a wide range of data sizes, the GAN-CS-MLBP model outperformed the others. With 500 data, the accuracy of the GAN-CS-MLBP model is 98.873%.

### 4.8 Root-mean-square error

A quantitative assessment technique for calculating the accuracy of facial recognition systems is root-mean-square error (RMSE) analysis for automated face recognition. It takes some work to find the square root of the average squared dissimilarities between the expected and real facial features or identities. A quantitative indicator of the entire discrepancy or variation among expected and actual results in facial recognition is provided by the measurement of RMSE. This creates it possible for researchers and programmers to assess and contrast the efficiency of numerous facial recognition algorithms or models.

The only difference between the mean absolute error (MAE) and the mean square error (MSE) is the square root representation. Equation (13) can be used to design the mean absolute error.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad (13)$$

RMSE is a measurement that ranges from 0 to, and lower RMSE values are preferred.

An association of the proposed GAN-CS-MLBP method's RMSE with other known techniques is displayed in Fig. 11 and Table 7. The graph shows that the deep learning method performs better with the least amount of RMSE. The graph displays the RMSE of the proposed GAN-CS-MLBP using LFW, YTF, and CASIA-WebFace datasets. For example, the GAN-CS-MLBP, with LFW dataset, has an RMSE value of 34.827%, whereas with YTF dataset it has an RMSE of 35.425%. Similarly, the GAN-CS-MLBP, with CASIA-WebFace dataset, has an RMSE value of 36.524%.

### 4.9 Computational efficiency

In Fig. 12, the computational efficiency of the GAN-CS-MLBP methodology is compared to many other methods that are currently in use. The graph illustrates the improved efficiency of the deep learning approach while maintaining computational efficiency. The GAN-CS-MLBP model has a computational efficiency of 84.82%, 86.42%, 85.66%, and 87.75% for 100, 200, 300, and 400 data. With a range of data sizes, the GAN-CS-MLBP model outperformed the others. Under 500 data, the computational efficiency of the GAN-CS-MLBP model is 89.62%.

### 4.10 Scalability

Figure 13 present a scalability comparison of the GAN-CS-MLBP technique with other current approaches. The graph shows how the deep learning method maintains scalability while offering increased efficiency. The GAN-CS-MLBP model has a scalability of 87.53%, 89.53%, 88.29%, and 90.17% for 100, 200, 300, and 400 data. With a range of data sizes, the GAN-CS-MLBP model outperformed the others. Under 500 data, the scalability of the GAN-CS-MLBP model is 91.43%.

Figure 14 shows the training and validation accuracy of the GAN-CS-MLBP system on an 80:20 training set to validation set ratio. The training accuracy is ascertained by assessing the GAN-CS-MLBP approach on the training dataset, whereas the validation accuracy is ascertained by assessing the performance on an alternative testing dataset. The findings show that as epochs increase, so do training and validation accuracy. Consequently, the GAN-CS-MLBP approach performs better on the training and validation dataset as the number of epochs increases. The GAN-CS-MLBP system's training and validation losses are shown in Fig. 14b using an 80:20 split of the training and validation sets. We estimate the difference between original values and anticipated performance inside the training data using a statistic called training loss, and we utilize a metric called validation loss to evaluate the performance of the GAN-CS-MLBP approach on individual validation data. The outcomes demonstrate the improved performance and classification accuracy of the GAN-CS-MLBP technique by showing that both training loss and validation loss reduce as the number of epochs increases. The technique's superiority in identifying patterns and correlations is highlighted by the declining values of the validation loss and training loss.

### 4.11 Discussion

Utilizing a combination of deep learning methods, precisely GAN and CS-MLBP, the study recommends an original technique for automated face identification. In suggestion to conventional methods, the results display significant improvements in facial recognition accuracy. The model professionally learns and creates realistic facial

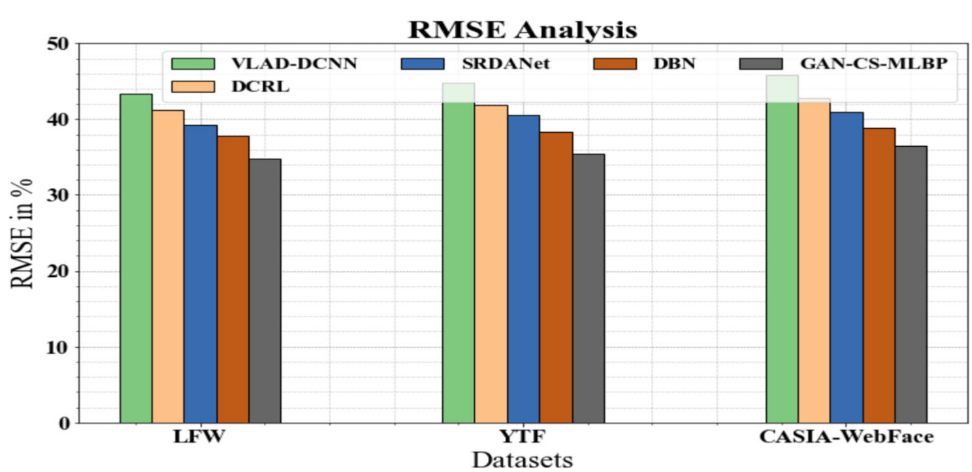


**Fig. 11** RMSE analysis for GAN-CS-MLBP method with existing systems

**Table 7** RMSE analysis for GAN-CS-MLBP technique with existing systems

| Datasets | VLAD-DCNN | DCRL | SRDANet | DBN | GAN-CS-MLBP |
|---|---|---|---|---|---|
| LFW | 43.324 | 41.234 | 39.222 | 37.827 | 34.827 |
| YTF | 44.725 | 41.827 | 40.526 | 38.324 | 35.425 |
| CASIA-WebFace | 45.827 | 42.827 | 40.927 | 38.928 | 36.524 |

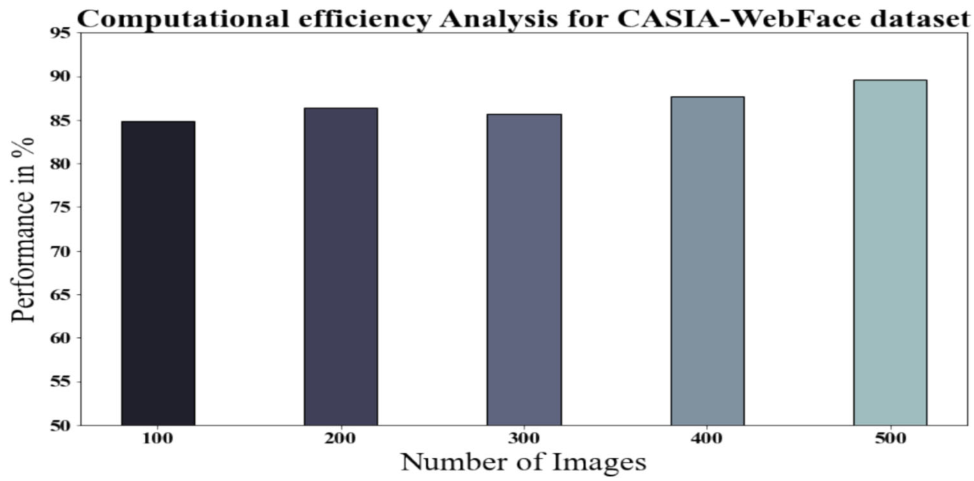


**Fig. 12** Computational efficiency for GAN-CS-MLBP technique for CASIA-WebFace dataset

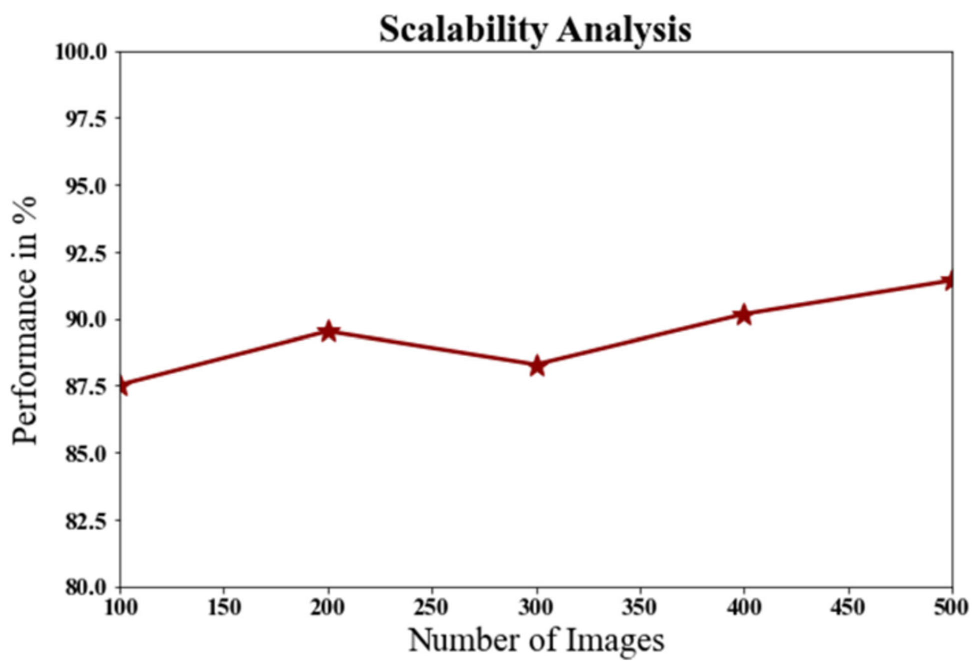


**Fig. 13** Scalability for GAN-CS-MLBP technique for CASIA-WebFace dataset

features by using the power of GANs, growing the resilience of the recognition system. Moreover, the combination of CS-MLBP rises the ability to capture intricate face features, resulting in enhanced performance across a range of scenarios involving changes in lighting and facial expressions. The findings designate that this hybrid technique may achieve new performance in face recognition tasks, with promising applications in biometric identification systems, security, and surveillance. With scores of 96.324% for LFW, 97.627% for YTF, and 98.873% for CASIA-WebFace, the presented model had the highest accuracy. Our proposed model performed well, with even more extraordinary results in practical implementations; further investigation and optimization may follow.

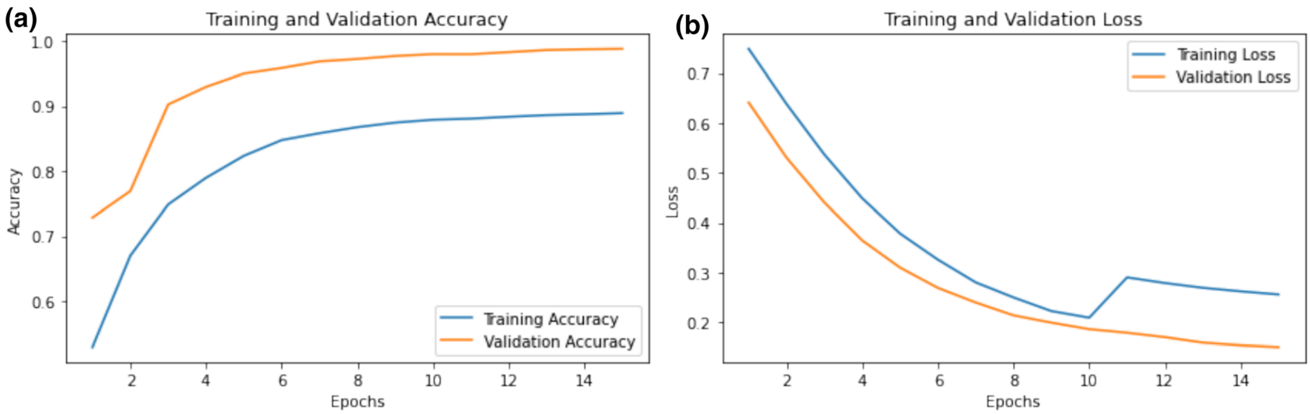


**Fig. 14** **a** Training and testing accuracy analysis **b** Training and testing loss analysis

## 5 Conclusion

Automated facial recognition technology has progressive rapidly, with numerous possible uses. Law enforcement, security, and digital identity verification use it extensively. Ethical problems, including privacy violations and prejudices, must be addressed and regulated to ensure fair and responsible usage of this technology. Technology developers, lawmakers, and privacy activists must continue research, improvement, and collaboration to balance innovation and privacy in automated facial recognition. This study recommends two novel face recognition approaches: the CS-MLBP and GAN. Column-based two-dimensional principle component analysis stands out as an effective technique for efficiently extracting the basic features of a face from enormous datasets. Nevertheless, it requirements help managing detailed local information. In this study, the global components extracted by C2DPCA are combined with the regional characteristics of the face's texture using the CS-MLBP method. A robust and accurate face recognition system called GAN-CS-MLBP combines CS-MLBP and GAN. Deep learning systems, mainly neural networks, automatically extract discriminative features from facial images. The learned features let the model differentiate people by capturing low-level information and high-level meanings. In-depth tests are done on benchmark face recognition datasets like LFW, YTF, and CASIA-WebFace to evaluate how well the recommended GAN-CS-MLBP technique performs. Future research will associate the recommended system to numerous current state-of-the-art systems and evaluate it against other databases.

### 5.1 Advantages of proposed method

- Center symmetric multivariant local binary pattern and deep learning algorithms work together to improve accuracy and provide reliable face recognition results.
- Robust recognition across a range of ambient situations, lighting variations, and facial expressions is made feasible by the combination of deep learning and local binary pattern.
- Compared to manual identification methods, automated processing saves time and resources by simplifying face recognition processes.

### 5.2 Limitations

- GANs, like other deep learning models, are vulnerable to adversarial assaults. Small variations in input photos can result in misclassification or erroneous recognition.
- The use of face recognition technology raises ethical concerns about permission, surveillance, and possible biases. There is a possibility of discrimination or unfair treatment, particularly for underrepresented populations.
- GANs are known for their lack of interpretability. Understanding why a GAN-based facial recognition system makes a specific judgment might be difficult, raising questions about responsibility and reliability.

**Data availability** The dataset used in this paper are available from: https://www.kaggle.com/datasets/debarghamitraroy/casia-webface.

## Declarations

**Conflict of interest** The authors declare no conflict of interests.

## References

1. Mukhiddinov M, Djuraev O, Akhmedov F, Mukhamadiyev A, Cho J (2023) Masked face emotion recognition based on facial landmarks and deep learning approaches for visually impaired people. Sensors 23(3):1080. https://doi.org/10.3390/s23031080

2. Gupta S, Kumar P, Tekchandani RK (2023) Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimed Tools Appl 82(8):11365–11394. https://doi.org/10.1007/s11042-022-13558-9
3. Poulose A, Reddy CS, Kim JH, Han DS (2021, August) Foreground extraction based facial emotion recognition using deep learning xception model. In: 2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), IEEE, 356–360. https://doi.org/10.1109/ICUFN49451.2021.9528706
4. Gaddam DKR, Ansari MD, Vuppala S, Gunjan VK, Sati MM (2022) Human facial emotion detection using deep learning. In: ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications, Springer Singapore, 1417–1427. https://doi.org/10.1007/978-981-16-3690-5_136
5. Hossain S, Umer S, Rout RK, Tanveer M (2023) Fine-grained image analysis for facial expression recognition using deep convolutional neural networks with bilinear pooling. Appl Soft Comput 134:109997. https://doi.org/10.1016/j.asoc.2023.109997
6. Chaudhari A, Bhatt C, Krishna A, Travieso-González CM (2023) Facial emotion recognition with inter-modality-attention-transformer-based self-supervised learning. Electronics 12(2):288. https://doi.org/10.3390/electronics12020288
7. Chen H, Wang T, Chen T, Deng W (2023) Hyperspectral image classification based on fusing S3-PCA, 2D-SSA and random patch network. Remote Sensing 15(13):3402. https://doi.org/10.3390/rs15133402
8. Duan Z, Song P, Yang C, Deng L, Jiang Y, Deng F, Deng W (2022) The impact of hyperglycaemic crisis episodes on long-term outcomes for inpatients presenting with acute organ injury: a prospective, multicentre follow-up study. Front Endocrinol 13:1057089. https://doi.org/10.3389/fendo.2022.1057089
9. Bharti SK, Varadhaganapathy S, Gupta RK, Shukla PK, Bouye M, Hingaa SK, Mahmoud A (2022) Text-based emotion recognition using deep learning approach. Comput Intell Neurosci 2022(1):2645381. https://doi.org/10.1155/2022/2645381
10. Lasri I, Riadsolh A, Elbelkacemi M (2023) Facial emotion recognition of deaf and hard-of-hearing students for engagement detection using deep learning. Educ Inf Technol 28(4):4069–4092. https://doi.org/10.1007/s10639-022-11370-4
11. Farzin S, Anaraki MV, Naeimi M, Zandifar S (2022) Prediction of groundwater table and drought analysis; a new hybridization strategy based on bi-directional long short-term model and the Harris hawk optimization algorithm. J Water Clim Change 13(5):2233–2254. https://doi.org/10.2166/wcc.2022.066
12. Stoean C, Zivkovic M, Bozovic A, Bacanin N, Strulak-Wójcikiewicz R, Antonijevic M, Stoean R (2023) Metaheuristic-based hyperparameter tuning for recurrent deep learning: application to the prediction of solar energy generation. Axioms 12(3):266. https://doi.org/10.3390/axioms12030266
13. He Z, Li Z, Yang F, Wang L, Li J, Zhou C, Pan J (2020) Advances in multimodal emotion recognition based on brain–computer interfaces. Brain Sci 10(10):687. https://doi.org/10.3390/brainsci10100687
14. Veni S, Anand R, Mohan D, PAUL E (2021, March) Feature fusion in multimodal emotion recognition system for enhancement of human-machine interaction. In: IOP conference series: materials science and engineering, IOP Publishing, 1084(1):012004. https://doi.org/10.1088/1757-899X/1084/1/012004
15. Nan Y, Ju J, Hua Q, Zhang H, Wang B (2022) A-MobileNet: an approach of facial expression recognition. Alex Eng J 61(6):4435–4444. https://doi.org/10.1016/j.aej.2021.09.066
16. Mohan K, Seal A, Krejcar O, Yazidi A (2021) FER-net: facial expression recognition using deep neural net. Neural Comput Appl 33(15):9125–9136. https://doi.org/10.1007/s00521-020-05676-y
17. Fard AP, Mahoor MH (2022) Ad-corre: adaptive correlation-based loss for facial expression recognition in the wild. IEEE Access 10:26756–26768. https://doi.org/10.1109/ACCESS.2022.3156598
18. Abbas Q, Albalawi TS, Perumal G, Celebi ME (2023) Automatic face recognition system using deep convolutional mixer architecture and adaboost classifier. Appl Sci 13(17):9880. https://doi.org/10.3390/app13179880
19. Karasugi IPA, Williem (2020, August) Face mask invariant end-to-end face recognition. In: European Conference on Computer Vision, Cham: Springer International Publishing, 261–276. https://doi.org/10.1007/978-3-030-68238-5_19
20. Elordi U, Lunerti C, Unzueta L, Goenetxea J, Aranjuelo N, Bertelsen A, Arganda-Carreras I (2021) Designing automated deployment strategies of face recognition solutions in heterogeneous IoT platforms. Information 12(12):532. https://doi.org/10.3390/info12120532
21. Monica S, Roseline Mary R (2022, July) Face and emotion recognition from real-time facial expressions using deep learning algorithms. In: Congress on Intelligent Systems: Proceedings of CIS 2021, Volume 1, Singapore: Springer Nature Singapore, 451–460. https://doi.org/10.1007/978-981-16-9416-5_32
22. Huang ZY, Chiang CC, Chen JH, Chen YC, Chung HL, Cai YP, Hsu HC (2023) A study on computer vision for facial emotion recognition. Sci Rep 13(1):8425. https://doi.org/10.1038/s41598-023-35446-4
23. Reddi PS, Krishna AS (2023) Cnn implementing transfer learning for facial emotion recognition. Int J Intell Syst Appl Eng 11(4s):35–45
24. AlEisa HN, Alrowais F, Negm N, Almalki N, Khalid M, Marzouk R, Alneil AA (2023) Henry gas solubility optimization with deep learning based facial emotion recognition for human computer interface. IEEE Access 11:62233–62241. https://doi.org/10.1109/ACCESS.2023.3284457
25. Kundu P, Kundu P, Mallik S, Bhowmick S, Mandal P, Banerjee H, Pal SB (2022) Facial expression recognition using convoluted neural network (CNN). In: Cyber Intelligence and Information Retrieval: Proceedings of CIIR 2021, Springer Singapore, 81–88. https://doi.org/10.1007/978-981-16-4284-5_8
26. Lijuan Y, Yanhu Z (2023) A face recognition algorithm based on the combine of image feature compensation and improved PSO. Sci Rep 13(1):12372. https://doi.org/10.1038/s41598-023-39607-3
27. Huang GB, Mattar M, Berg T, Learned-Miller E (2008) Labeled faces in the wild: a database forstudying face recognition in unconstrained environments. In: Workshop on faces in'Real-Life'Images: detection, alignment, and recognition.
28 https://www.kaggle.com/datasets/quadeer15sh/lfw-facial-recognition
29. Wolf L, Hassner T, Maoz I (2011, June) Face recognition in unconstrained videos with matched background similarity. In: CVPR 2011, IEEE, 529–534. https://doi.org/10.1109/CVPR.2011.5995566
30 https://www.kaggle.com/datasets/nouhabannour/youtube-faces
31. Yi D, Lei Z, Liao S, Li SZ (2014) Learning face representation from scratch. arXiv preprint arXiv:1411.7923.
32 https://www.kaggle.com/datasets/nhatdealin/casiawebface-dataset-crop
33. Kansizoglou I, Misirlis E, Tsintotas K, Gasteratos A (2022) Continuous emotion recognition for long-term behavior modeling through recurrent neural networks. Technologies 10(3):59. https://doi.org/10.3390/technologies10030059
34. Yang J, Zhang D, Frangi AF, Yang JY (2004) Two-dimensional PCA: a new approach to appearance-based face representation

and recognition. IEEE Trans Pattern Anal Mach Intell 26(1):131–137. https://doi.org/10.1109/TPAMI.2004.1261097
35. Lucieer A, Stein A, Fisher P (2005) Multivariate texture-based segmentation of remotely sensed imagery for extraction of objects and their uncertainty. Int J Remote Sens 26(14):2917–2936. https://doi.org/10.1080/01431160500057723
36. Remya Revi K, Vidya KR, Wilscy M (2021) Detection of deepfake images created using generative adversarial networks: A review. In: Second International Conference on Networks and Advances in Computational Technologies: NetACT 19, Springer International Publishing, 25–35. https://doi.org/10.1007/978-3-030-49500-8_3
37 Babu K, Kumar C, Kannaiyaraju C (2022) Face recognition system using deep belief network and particle swarm optimization. Intell Autom Soft Comput 33(1):317–328. https://doi.org/10.32604/iasc.2022.023756
38. Tian L, Fan C, Ming Y, Shi J (2015) SRDANet: an efficient deep learning algorithm for face analysis. In: Intelligent Robotics and Applications: 8th International Conference, ICIRA 2015, Portsmouth, UK, August 24–27, 2015, Proceedings, Part I 8, Springer International Publishing, 499–510. https://doi.org/10.1007/978-3-319-22879-2_46
39. Wang W, Wang R, Shan S, Chen X (2017) Discriminative covariance oriented representation learning for face recognition with image sets. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5599–5608. https://doi.org/10.1109/CVPR.2017.609
40. Zheng J, Chen JC, Bodla N, Patel VM, Chellappa R (2016, December) Vlad encoded deep convolutional features for unconstrained face verification. In: 2016 23rd international conference on pattern recognition (ICPR), IEEE, 4101–4106. https://doi.org/10.1109/ICIP.2016.7532906



## Authors and Affiliations

**J. C. Sekhar[1] · P. Joel Josephson[2] · A. Chinnasamy[3] · M. Maheswari[4] · S. Sankar[5] · Ruth Ramya Kalangi[6]**

✉ J. C. Sekhar
jcsekhar9@nriit.ac.in

P. Joel Josephson
joeljosephsonp@gmail.com

A. Chinnasamy
chinnasamyamb@gmail.com

M. Maheswari
m.mahe05@gmail.com

S. Sankar
sankars.sse@saveetha.com

Ruth Ramya Kalangi
ramya_cse@kluniversity.in

1 Department of Computer Science and Engineering, NRI Institute of Technology, Guntur, Andhra Pradesh, India

2 Department of Electronics and Communication Engineering, Malla Reddy Engineering College, Hyderabad, Telangana, India

3 Department of Data Science and Business Systems, School of Computing, SRMIST, Kattankulathur, Chennai, India

4 Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, India

5 Department of Computer Science and Engineering, Saveetha School of Engineering, SIMATS, Chennai, India

6 Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vijayawada, Andhra Pradesh, India