International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V71I3P243

Volume 71 Issue 3, 407-422, March 2023

Original Article

MF-FaceNet: An Intuitive Age-Invariant Face Recognition through Multi-Feature and Multi-Fusion CNNs

M. Rajababu[1], K. Srinivas[2], H. Ravisankar[3]

[1]*Department of CSE, JNTUH, Hyderabad, Telangana, India.*
[2]*Department of CSE, V.R. Siddhartha Engineering College, Andhra Pradesh, India.*
[3]*CTRI, Andhra Pradesh, India.*

[1]*Corresponding Author : raajababu.makineedi@gmail.com*

Received: 01 January 2023 Revised: 15 March 2023 Accepted: 22 March 2023 Published: 25 March 2023

Abstract *- In contrast to general face recognition, Age-Invariant Face Recognition (AIFR) seeks to match faces with a significant age difference. Due to the various shapes and textures of the human body, one of its most important steps is the capacity to recognize the various facial features. Poor recognition performance is the result of all these variances. Previous discriminative techniques have typically focused on dividing facial characteristics into age-related and age-invariant components, which results in the loss of face identification information. In terms of facial recognition, deep learning is thought to be a promising method. Several obstacles still exist, despite the advances made in this area. Deep learning has been seen as a promising technique in terms of face recognition. Unfortunately, despite the advances made in this area, there are still several difficulties. This paper presents a Multi-Feature and Multi-Fusion CNN method for accurate AIFR. This proposed system is named as MF-FaceNet. In this proposed system, features of the face images are extracted from the augmented, grey level and Local Binary Pattern (LBP)image datasets. Experiments are conducted for each extracted feature with three individual CNN models such as VGG-16, DenseNet-201 and ResNet-50. In the next phase of experimental investigation, extracted features are fused, and the investigation is carried out with the fusion CNN models like VDe-23, VRe-23, DeRe-23 and VDeRe-23.For this experimentation with individual and Fusion CNN models, extensive results are obtained with the aid of input standard face datasets such as FGNET, CACD and a Live dataset AS-23. Among all three datasets, the best-performed dataset is the AS-23 live dataset, and compared with the three CNN models, the highest testing accuracy of face recognition is obtained from DenseNet-201(94.27%,93.97%,90.07%). In the AS-23 live dataset, in comparison with three fusion models, the VDeRe-23 CNN feature fusion model achieved the best face recognition testing accuracies, such as 94.02%, 91.27% and 90.7%, respectively. Among all the comparisons on three datasets, the DenseNet-201 achieved the best face recognition accuracy, and for the fusion model, VDeRe-23 gave the best recognition accuracy. Our method enhances the AIFR performance on the College dataset (AS-23) when compared to other well-known datasets. Our approach performs better than the current approaches and offers a high level of recognition accuracy.*

Keywords *- Age-invariant Face Recognition, Convolutional Neural Network Feature Extraction, Feature Fusion, Local Binary Pattern.*

1. Introduction

Face Recognition (FR) is a field of biometrics that deals with extracting facial features to authenticate and identify a person. Recent research on FR [1,49] achieved remarkable results in developing deep learning methods. The FR performance, however, is still below the level of acceptance because of the impact of ageing and variations in facial features. This necessitates the development of an AIFR for the anti-aging process. It is an extremely practical technology that can be used in crucial real-world scenarios where age correction is necessary, such as criminal and missing people identification and biometric security systems. An AIFR system's critical phase is feature extraction. It deals with extracting data that best represents the image and is age-invariant across age progression. The two main categories of AIFR approaches are generative and discriminative. The generative techniques [11,40] first model the face's ageing process to create a face image of the desired age.

The major goal of discriminative methods [14–20], in contrast to generative methods, is to identify an important feature for identity recognition that is robust to ageing. In recent years, deep learning techniques have made it possible to advance discriminative methods [3–9]. There are two separate types of age- and face-related pieces of information contained in the deep features that have been extracted at

various ages. In order to distinguish the identity-dependent components from the extracted facial features, the cross-age discriminative models focus on this. Although these techniques have achieved satisfactory performance on the AIFR by decomposing face features to obtain age-invariant identity features, learning age-invariant patterns is still difficult when the age range is broad, and the intra-class variations between face images from a single person are also large.

The proposed system is the MF-FaceNet system, which reduces the intra-class variations while learning more robust and discriminative features. More discriminative features of the face images are extracted from augmented, grey level, and LBP face image datasets to improve the accuracy.

The following list summarizes the principal objectives of our study:

1. Design an AIFR system with a seven-stage process.
2. Conversion of features
3. Experiment with the individual models.
4. Fusion of CNN models.
5. Feature fusion
6. Fusion of CNN models for feature fusion.

The rest of this paper is organized as follows: Section 2 presents current works of art in AIFR along with their limitations. A brief overview of the new suggested method is presented in Section 3. The experimental findings and analysis are reported in Section 4. Finally, Section 5 concludes the study.

2. Related Work

The most recent research on AIFR is covered in this section, along with its limitations. The two main kinds of conventional methods for age-invariant face recognition are generative methods and discriminative methods. Generative methods synthesize the input image to create the face image with the desired age by simulating the ageing process. Lanitis et al. [10] suggested an automatic age simulation framework utilizing a statistical age model to simulate the effects of ageing on input face images.

Park et al. [11] suggested a more reliable 3D face model improve recognition performance by considering both adult and developmental ageing effects on face photos into account. They tested their technique using the FG-NET, MORPH-II, and BROWNS datasets and discovered that it. Due to the drawbacks of generation methods, researchers frequently employ discriminative techniques to address the AIFR issue effectively. Discriminative approaches use more sophisticated feature representation and learning techniques to overcome AIFR. In order to recognize faces, discriminative approaches must be used to extract age-invariant facial features. In recent years, the community proposed several discriminative methods.

Ling et al. [12] studied the performance against age gaps in face recognition using more powerful face descriptors. In a novel approach called Hidden Factor Analysis (HFA), developed by Gong et al. [41], face images are highlighted as identity and ageing components to discriminate between ageing variations and individual features. The identity factor is constant throughout life; however, the aging factor changes with time. Chen et al. [14] employed a novel coding scheme to transform low-level information into an age-independent representation. The Cross-Age Celebrity Dataset (CACD), a massive cross-age face dataset, was additionally introduced. Performance improvements were made in both the studies they ran on the open FG-NET dataset and their own proprietary dataset. In a recent study, Verma et al. [15] fused the CSLBP features with a weighted KNN classifier to build a reliable technique for face verification. The verification accuracy was improved to 80.7% by testing it on the MORPH-II database.

Due to the great potential of CNN in classification problems, deep learning approaches have gained popularity in computer vision applications. By directly taking images, both recognition and classification can be accomplished simultaneously using Deep CNN. Unlike in CNNs, where no feature extraction is necessary, the traditional method relies on carefully built features, which must be extracted by hand. Automatically, CNNs pick up on these details. CNN's retrieved features are more relevant to categorization and more resistant to age differences.

Consequently, researchers presented their AIFR approaches based on deep learning. A new deep CNN model was proposed by Li et al. [16], which simultaneously learns features, optimizes distance metrics, and adjusts thresholds. Their method outperformed all others on the MORPH-II dataset with a Rank-1 recognition rate of 93.65%. A more reliable and effective approach, called Latent Factor Guided-CNN(LF-CNN), was proposed by Wen et al. [17] using a carefully designed CNN model. They reported accuracy for the FGNET, CACD-VS, and MORPH Album2 datasets of 97.51%, 88.11%, and 99.45%, respectively.

Zheng et al. [18] developed a new network named AE-CNN to separate the age factor from the image. They performed an experiment using the MORPH-II dataset, and the results showed that their method had a comparative identification rate of 98.13%. An innovative method for face verification with significant age discrepancies was reported by Bianco [19] using a deep CNN trained in an s Siamese network that ended with a contrastive loss. By adding a layer to inject new features, they were able to boost face verification performance even further. Li et al. [20] proposed a novel approach for joint learning and measuring similarity by combining feature learning with distance metric learning within a deep CNN.

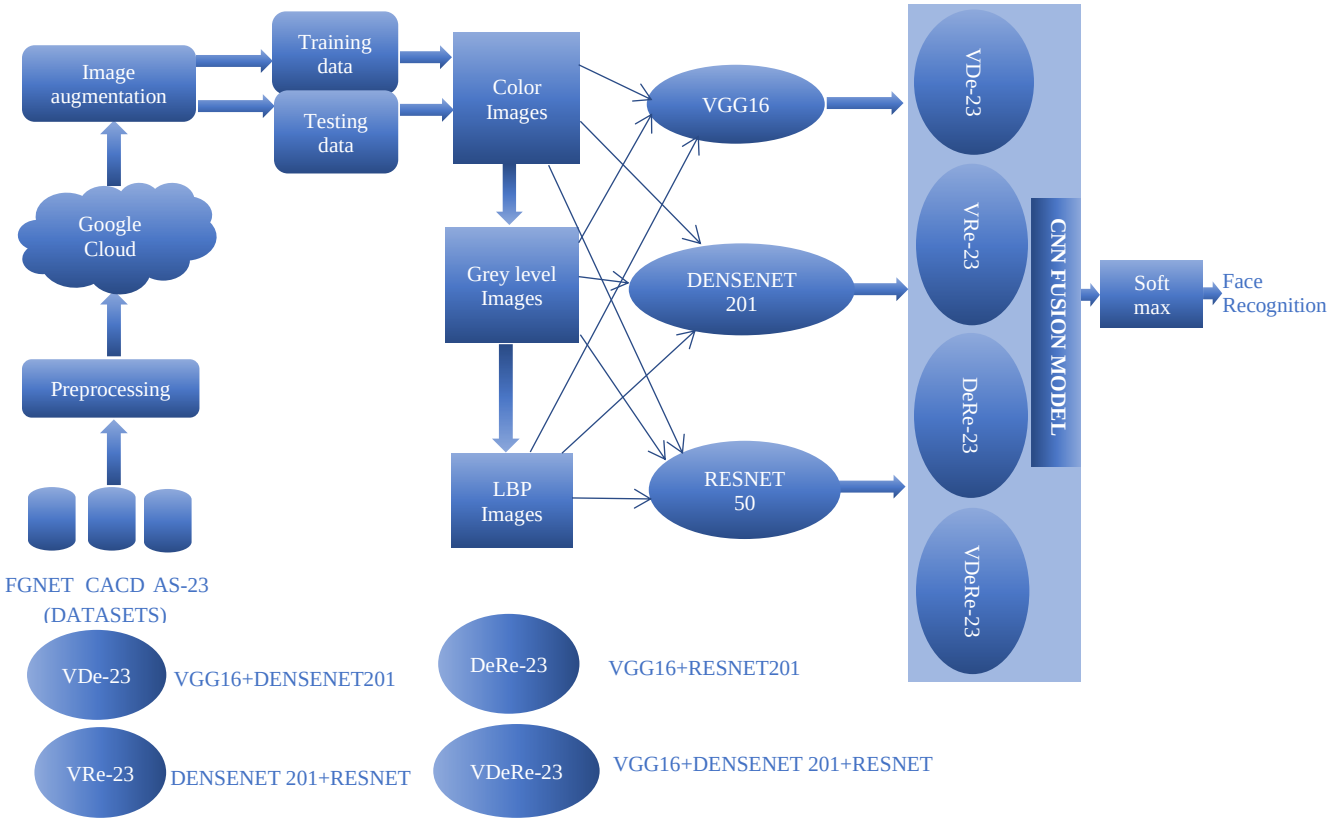


Fig. 1 Block diagram for fusion models of MF-FaceNet

Coupled Auto-Encoder Network (CAN) is a novel neural network model that Xu et al. [21] implemented for AIFR. It consists of encoders connected by two shallow neural networks. Orthogonal Embedding CNN (OE-CNN) is a novel network that Wang et al. [42] developed to improve the efficiency of AIFR. They analyzed the efficiency of their method by testing on MORPH-II, FG-NET, CACD-VS and Labelled Faces in the Wild Home datasets, where they attained recognition accuracies of 98.5%,58.1%,99.5% and 99.4%, respectively. Using a 7-layer CNN architecture, Nimbarte and Bhoyar[23] developed a new approach to AIFR. The Rank-1 recognition accuracy of their method was 76.6% on the FGNET and 92.50% on the MORPH. There are three stages to the new AIFR procedure developed by Afroze et al. [24]. Frangi2D normalization is used to standardize the age ranges of all faces, then ageing-invariant features are retrieved, and finally, a Sparse Representation Classifier is used for face recognition (SRC). An improved recognition rate of 84.79% was achieved for the MORPH-II dataset. With the goal of selecting the most important features for face recognition in mind, a new approach for AIFR developed by Moustafa et al. [25] is based on deep feature optimization with the help of a Genetic algorithm (GA). The Recognition accuracy was increased to 86.2% and 96%, respectively, when their approach was tested on the FGNET and MORPH-II datasets.

A sophisticated feature encoding technique for feature extraction using a pre-trained CNN model was introduced by Shakeel and Lam [43]. They tested their approach on datasets, and the best Rank-1 recognition percentage they found was 98.67%. Decorrelated Adversarial Learning (DAL) is a unique approach to the decorrelated adversarial regularization framework that Wang et al. [27] brought to AIFR with the goal of regularizing the learning of components.

Additionally, the method uses an adversarial process to reduce the correlation between age and identity features. They also introduced Batch Canonical Correlation Analysis (BCCA), an extension of CCA primarily useful for correlation regularization. Extensive testing showed that on the three widely used datasets of FG-NET, MORPH-II, and CACD-VS, the recognition rates for Rank-1 were 94.5, 98.97%, and 99.6%, respectively.

A four-stage AIFR system was proposed by Moustafa et al. [28]. After the image preprocessing step for face detection, the VGG-Face CNN model was used for feature extraction, yielding highly accurate discriminative identifiers. After that, the feature vector dimensions are reduced with an effective multi-Discriminant correlation analysis (MDCA) technique before fusion.

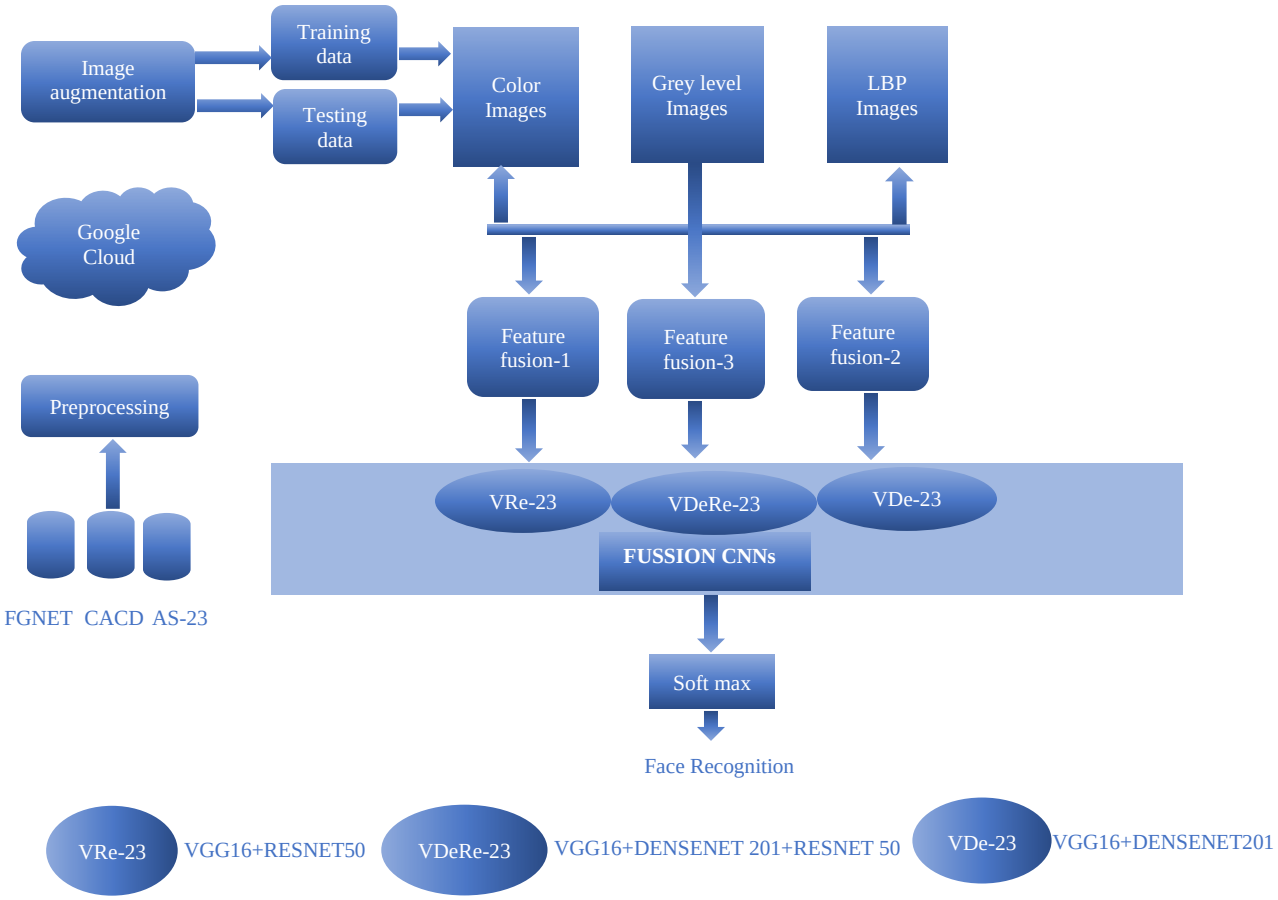


Fig. 2 Schematic diagram for feature fusion models of MF-FaceNet

Ultimately, the provided method classified data using the KNN and SVM algorithms. They evaluated this method and found that it was accurate to the tune of 81.5% on FGNET and 96.50% on MORPH-II. For the purpose of multi-age face recognition, Du et al.[44] introduced a unique framework called Age Factor Removal Network (AFRN), which combines transfer learning and adversarial learning to take on AIFR tasks. Specifically, this model uses adversarial learning to bring together two separate networks—a feature generator network (G) and an age discriminator network (D). Huang and Hu [30] introduced a unique end-to-end CNN-based model, the Age Adversarial Convolutional Neural Network (AA-CNN), trained using age-labelled and identity-labelled datasets. Yan, Chenggang, et al.[31]proposed a novel Multi-feature Fusion and Decomposition (MFD) framework for age-invariant face recognition, which learns more discriminative and robust features and reduces the intra-class variants. An age-invariant face identification system using high-dimensional binary pattern features and a parallel deep convolutional neural network was proposed by Dharavath et al. [46]. With FGNET, they reached a high identification accuracy level of 81.25%, and with the CACD dataset, the MAP was improved from 0.54 to 0.81.

3. Proposed Method

This section provides a concise summary of our proposed method and discusses the underlying algorithms. The proposed MF-FaceNet system, intelligently recognizing the faces with the three datasets, is shown in Figures 1&2. The proposed system is dependent on the seven stages for recognizing faces.

1) Preprocess the image datasets prepared for the experimentation.
2) Convert the features from original data as Augmented, Grey level and LBP images and these images are generated as three individual datasets.
3) Classify the face image datasets using three individual CNNs such as VGG-16, DenseNet-201 and ResNet-50.
4) Combine the three CNN models with various combinations like VDe-23, VRe-23, DeRe-23 and VDeRe-23 fusion models for the further classification of images.
5) The extracted features like augmented, grey level and LBP are combined as feature fusion of all three types of features. The feature fusion images are further applied to the fusion CNN models.

6) All the individual, fusion and feature fusion CNN models are compared with the performance measures and the best CNN model is taken for the next steps.
7) Finally, all the CNN Models are applied to the SoftMax classification to recognize the faces of three datasets. The best dataset has been taken as the best recognition accuracies among the three datasets.

The system MF-FaceNet is illustrated with several steps to find face recognition as follows:

3.1. Datasets

This study uses publicly accessible datasets FGNET and CACD and a live dataset AS-23. The FGNET dataset included a total of 1,002 images from 82 unique people.

Each person, on average, has 12 images, with a wide variation between 6 and 18 images. A major limitation is the limited number of subjects whose images make up the FG-NET collection (82). While the ages of FG-NET participants range from 0 to 69, over half are younger than 13 years old. In this dataset, among the 82 persons, 1002 images, for experimentation, the best possible 10 persons,150 images, are considered.

In CACD [14], there are a total of 2000 subjects, which contain 163,446 images of those subjects as celebrities. Between 2004 and 2013, the entire collection was captured. People in the sample range in age from 12 to 62. In this dataset, among the 2000 persons, 1,63,446 images, for experimentation, the best possible 10 persons,160 images are considered.

The dataset AS-23 is the college student dataset; the students' images are collected from the ages of 2 to 21 years. There is a total of 100 subjects, which contains 10000 images of those subjects as students. Between 2003 and 2023, the entire collection was captured. In this dataset, among the 100 students, 10000 images, for experimentation, the best possible 10 students,100 images are considered. The sample images of three datasets are shown in Figure 3.

3.2. Preprocessing

Preprocessing is done to lower the level of uncertainty in subsequent face recognition procedures. Here illumination normalization and noise removal are used for this. Noise removal is a crucial step in face recognition that helps to increase recognition accuracy. For better experiments, the raw images are prepared for adjusted for brightness and contrast and enhance the image quality.

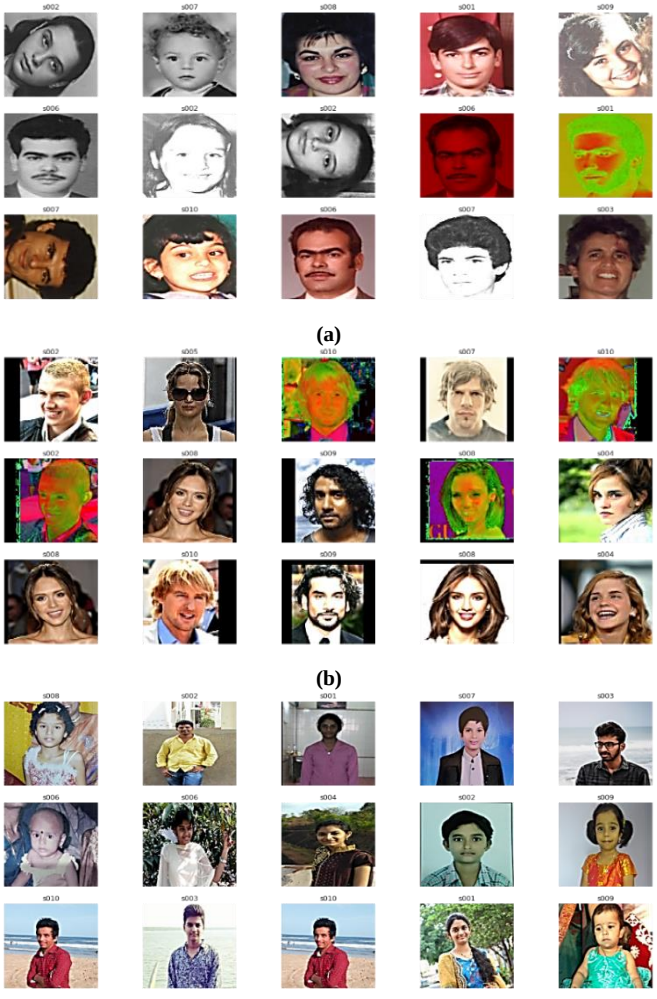
The inputs for CNN experiments are used to resize the recovered images into 224x224 as they come in various image sizes. Since the recovered images come in a variety of sizes, they are converted into images with a size of 224X 224 pixels using the inputs for CNN. Tests and every dataset used in the system are considered into 10 different folders, and each folder is represented as one class. So, 10 classes are represented in all three datasets for the experimentation.

3.3. Google cloud

Face image datasets were created by dividing face images into different categories of subjects. The high-end local servers and Google Cloud [33] for each user's unique account and Google Drive are then used to store these images. The Google Co-Lab was used to conduct in-depth research in order to analyze the results.

3.4. Data Augmentation

The dataset's considered images are the best possible, so the number of images is insufficient to train the CNN model. Data augmentation is considered a technique to enhance the images for better performance. Here the augmented technique is applied to the three datasets and improves the level of images. The techniques applied are resizing(224X224), rotation (90^0,180^0,270^0), shift (left and right), shearing, flip (horizontal and vertical), image noising, image blurring, scaling and zooming (20%,40% and 60%)



(a) **(b)** **(c)**

Fig. 3 sample images of a person from (a) FG-NET (b) CACD (c) AS-23

3.5. Training and Test Split

When the image augmentation is completed, the FGNET dataset gets a total image of 12800. The CACD dataset gets a total image of 11550, and the AS-23 dataset gets a total image of 7830. These final augmented images' training and testing datasets are divided into a 70:30 ratio. This splitting ratio yields the best outcomes in deep learning. The CNNs were then subjected to the next round of training and testing using these training and test images. The train and test split of three datasets is shown in Table 1.

Table 1. Train and test split of three datasets

S. No.	Dataset	Training set	Testing set	Total images
1	FGNET	9100	3700	12800
2	CACD	7850	3700	11550
3	AS-23	5490	2340	7830

3.6. Original Images

The produced enhanced images from all three datasets are considered the original images. These images are regarded as having unique RGB (Red-Green-Blue) colour images. These colour images are very key images in the entire system for considering further experimentations.

3.7. Feature Extraction from Grey-Level Images

Red-Green-Blue (RGB) images are more informative than grayscale ones (RGB). Nonetheless, they are frequently used in image processing because doing so is quicker and uses less space when dealing with complex computations. After data augmentation, grey-level data sets are created for all three datasets. The next level of experimentation is done with all individual, fusion and feature fusion CNN models.

3.8. Feature Extraction from LBP Images

LBP is a frequently employed descriptor to record texture information from the provided target image. The LBP coding of that pixel is determined by comparing a specific pixel's value to those of nearby pixels. As seen in Fig. 4, the left portion of the diagram depicts all the pixel values of a local area, while the right part offers the binary LBP coding of the central pixel.

24	15	40
31	31	57
25	14	19

Example Image

0	0	1
1		1
0	0	0

LBP

	19	

LBP Value

Fig. 4 Computation of LBP code

The LBP operator is used to characterise an image's local texture components. It can extract a lot of texture detail information. The LBP value, which characterises the local texture feature of the image, is obtained by comparing the values of the image's core pixel to those of its eight surrounding pixels. Its value is set to 1 if the surrounding pixels' grey values are greater than or equal to those of the centre pixels. The equation to find the LBP image is

$$LBP(x,y) = \sum_{i=0}^{i} k * (p - c) * 2^i \quad (1)$$

Where (x, y) stands for the centre pixel, p for the neighbouring pixel, and c for the middle pixel. Here, K(x), whose binary representation is based on the deviation from the central pixel, is defined as follows:

$$K(x) = \begin{cases} 1, & if\ x \geq 0 \\ 0, & else \end{cases} \quad (2)$$

Concatenating all of these binary codes starting at the top-left in a clockwise direction yields a binary number, and the labelling is done using the associated decimal value. Local Binary Patterns, or LBP codes, are the terms used to describe the resulting binary numbers.

For the next level of experimentation with all individual, fusion, and feature fusion CNN models, the LBP image datasets for the three datasets are constructed.

3.9. Face Image Recognition using CNNs

CNNs are a crucial component of the MF-FaceNet system. Here VGG-16[47], DenseNet-201[37,38] and ResNet-50[48] CNNs are the individual CNNs. Important experiments for recognizing faces from face datasets are being conducted using both these individual CNNs and the fusion of these CNNs. All four fusion models[39] and their suggested names are described in Table 2.

Each CNN fusion model integrates the individual CNN models, and the resulting model is then applied to training and testing experiments. The CNN used for recognizing the face images from the datasets is illustrated in Figure 3. It can take the input from the three image face datasets FGNET, CACD and AS-23. The convolution layer is applied to the dataset images to recover the picture features. In order to filter the picture values, the maximum pooling layer receives the image characteristics obtained from the preceding layer. The SoftMax layer, the last layer, uses multi-class classification to identify the many person face classes to categorize the input photos.

Table 2. Fusion models and their suggested names

S.No.	Fusion model	Suggested name
1	VGG16+ DenseNet-201	VDe-23
2	VGG16+ ResNet-50	VRe-23
3	DenseNet-201+ ResNet-50	DeRe-23
4	VGG16+DenseNet-201+ResNet-50	VDeRe-23

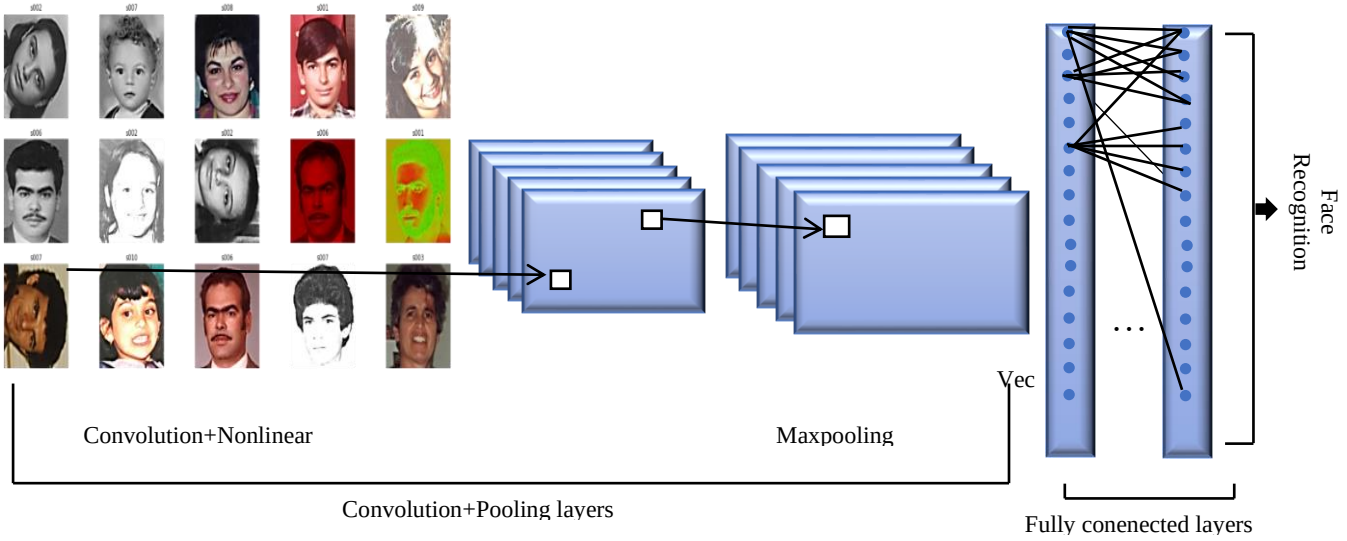


Fig. 5 Convolutional Neural Network architecture

In the classification of input photos, both individual and fusion models clearly have an advantage. Face recognition benefits greatly from DL technology. Recent studies have shown that when there are smaller connections between the layers near the input and the layers near the output, CNNs can learn considerably more thoroughly, precisely, and effectively. Table 3 provides the number of model parameters of all CNN models used in the experimentation.

3.10. Feature Fusion

The collected features like augmented, grey level and LBP images are converted from one feature to another feature. The three types of features are specifically applied to individual CNNs such as VGG-16, DenseNet-201 and ResNet-50. The same three features are further applied to fusion CNN models such as VDe-23, RDe-23, DeRe-23 and VDeRe-23.

The major concept is feature fusion in this system, which has taken as the combinations of all features. So, these features are combined as feature fusion-1 with the combination of augmented and grey levels. Next, feature fusion-2 is combined with grey level and LBP images and finally, feature fusion -3, which has the combination of all three features, combined with augmented, grey level and LBP images.

The feature fusion-1 dataset is applied to the fusion CNN Model VDe-23, the dataset feature fusion-2 is now applied to another fusion CNN model VRe-23, and finally, dataset feature fusion-3 is now applied to the final fusion model VDeRe-23. In this feature fusion, a number of CNN models have been experimented with, and all the experiments have produced several performance results. In these performance results, all the results are compared with each other. Finally, with this experimentation, the best feature fusion and fusion CNN model is considered for face recognition.

3.11. Face Recognition

The final layer is then applied to SoftMax for multi-class classification for each and every CNN, including individual and fusion models. With the use of performance measurements, faces are recognised in this multiclass classification, and CNN offers the highest level of accuracy.

4. Results

This experimentation was carried out using a three-stage MF-FaceNet system on two public faces datasets such as FGNET, CACD and a live dataset AS-23.In the first stage, for each extracted feature, experiments are conducted with three individual CNN models such as VGG16, DenseNet-201 and ResNet-50.In the second stage of experimental investigation, extracted features are fused, and the investigation is carried out with the fusion CNN models like VDe-23, VRe-23, DeRe-23 and VDeRe-23.In the third stage, features are combined as feature fusion-1 with the combination of grey level and LBP, feature fusion-2 combined with augmented and grey level images, and feature fusion-3 combined with augmented, grey level and LBP images. The feature fusion-1 dataset is applied to the fusion CNN Model VDe-23, the dataset feature fusion-2 is now applied to another fusion CNN model VRe-23, and finally, dataset feature fusion-3 is now applied to the final fusion model VDeRe-23.AUC stands for the area under the (ROC) curve. Generally, the higher the AUC score, the better a classifier performs for the given task. The recommended dataset generated the maximum overall accuracy of 97.99% for DenseNet-201, with individual CNN accuracy of 99.31% and integrated CNN accuracy of 91.72%. The comparison results on three datasets are presented in Tables 4, 5, 6, and 7. The analysis of comparative metrics with individual CNN models is described in Figures 7, 8, and 9, and with fusion models are described in Figures 10,11 and 12. Finally, the analysis of comparative metrics with feature fusion CNN models is described in Figure 13.

Table 3. Number of model parameters for the proposed CNN models

CNN Model	Introduced year	Total parameters	Trainable parameters	Nontrainable parameters	Layers
VGG-16 [47]	2014	1,53,14,391	5,99,703	1,47,14,688	16
DenseNet-201 [37,38]	2018	1,94,29,463	11,07,479	1,83,21,984	201
ResNet-50 [48]	2016	25,933,130	23,69,175	2,35,64,800	50
Proposed Fusion Models					
VDe-23	2023	34,742,374	1,705,702	33,036,672	217
VRe-23		41,246,886	2,967,398	38,279,488	201
DeRe-23		45,361,958	3,475,174	41,886,784	251
VDeRe-23		15,313,546	598,858	14,714,688	267

Table 4. Accuracy and AUC comparison for various CNN architectures on FGNET Dataset

CNN Model	Type of image features	Training accuracy in %	Testing accuracy in %	AUC in %
VGG-16	Augmented Images	97	88	99.38
	Grey level images	95	89	99.23
	LBP Images	75	70	95.44
DenseNet-201	Augmented Images	99	90	99.32
	Grey level images	98	89	99.02
	LBP Images	76.5	79.2	95.15
ResNet-50	Augmented Images	98	86.54	99.12
	Grey level images	95.77	88.05	99.17
	LBP Images	74.21	78.27	96.71
Fusion CNN Model	**Type of image features**	**Training accuracy in %**	**Testing accuracy in %**	**AUC in %**
VDe-23	Augmented Images	99.2	88.64	97.78
	Grey level images	99.6	89	97.98
	LBP Images	72.11	62	91.72
VRe-23	Augmented Images	97.99	86.05	97.71
	Grey level images	99.6	86	96.63
	LBP Images	74	78	89.14
DeRE-23	Augmented Images	96	89	98.12
	Grey level images	97.87	86	97.71
	LBP Images	74	73.2	92.67
VDeRe-23	Augmented Images	99.68	89.54	99.08
	Grey level images	99.78	90.78	98.69
	LBP Images	91.79	80.41	95.27

Table 5. Accuracy and AUC comparison for various CNN architectures on CACD Dataset

CNN Model	Type of image features	Training accuracy in %	Testing accuracy in %	AUC in %
VGG-16	Augmented Images	98.6	90	98.92
	Grey level images	97.85	90.43	96.66
	LBP Images	80.76	76.35	99.19
DenseNet-201	Augmented Images	99.39	89.65	99.09
	Grey level images	98.64	91.24	97.8
	LBP Images	96.15	83.05	99.1
ResNet-50	Augmented Images	98.09	89.21	98.99
	Grey level images	97.25	88.56	88.75
	LBP Images	52.18	55.62	99.42
Fusion CNN	**Type of image features**	**Training accuracy in %**	**Testing accuracy in %**	**AUC in %**
VDe-23	Augmented Images	92.19	83.21	99.42
	Grey level images	99.63	89.35	98.52
	LBP Images	74.11	63.05	99.09
VRe-23	Augmented Images	99.39	88.08	97.71
	Grey level images	97.26	71.35	98.47
	LBP Images	78	68	82.65
DeRe-23	Augmented Images	78.32	65.29	94.64
	Grey level images	94	79	97.42
	LBP Images	82	78	83.2
VDeRe-23	Augmented Images	99.76	90.02	98.21
	Grey level images	99.75	86.83	97.42
	LBP Images	94.23	89.23	95.1

Table 6. Accuracy and AUC comparison for various CNN architectures on AS-23 Dataset

CNN Model	Type of image features	Training accuracy in %	Testing accuracy in %	AUC in %
VGG-16	Augmented Images	98.67	93.8	99.69
	Grey level images	98.83	93.29	99.65
	LBP Images	89.04	88.54	96.66
DenseNet-201	Augmented Images	99.47	94.27	99.81
	Grey level images	98.85	93.97	99.75
	LBP Images	93.74	90.07	99.31
ResNet-50	Augmented Images	98.47	94.27	98.86
	Grey level images	96.2	90.12	98.65
	LBP Images	91.74	89.07	98.6
Fusion CNN	**Type of image features**	**Training accuracy in %**	**Testing accuracy in %**	**AUC in %**
VDe-23	Augmented Images	97.19	93.21	99.89
	Grey Level images	99.63	92.35	99.75
	LBP Images	93.11	90.05	99.63
VRe-23	Augmented Images	99.39	95.07	97.71
	Grey level images	97.26	91.56	97.98
	LBP Images	93.48	89.04	98.07
DeRe-23	Augmented Images	98	94.87	94.64
	Grey level images	94	92.17	97.42
	LBP Images	92.43	89.05	97.37
VDeRe-23	Augmented Images	99.76	94.02	99.61
	Grey level images	99.75	91.27	99.45
	LBP Images	94.23	90.7	99.71

Table 7. Accuracy and AUC comparison for feature fusion CNNs

Dataset	Feature Fusion CNN Model	Feature Fusion	Training accuracy in %	Testing accuracy in %	AUC in %
FGNET	VDe-23	Grey level +LBP Images	98.67	87.23	95.25
	VRe-23	Augmented + Grey level Images	94.13	89.12	94.35
	VDeRe-23	Augmented+ Grey level +LBP Images	98.76	87.25	98.01
CACD	VDe-23	Grey level + LBP Images	98.62	83.21	87.43
	VRe-23	Augmented+ Grey level Images	91.1	81.32	97.42
	VDeRe-23	Augmented+ Grey level +LBP Images	98.57	90.02	97.99
AS-23	VDe-23	Grey level +LBP Images	98.67	98.8	99.69
	VRe-23	Augmented+ Grey level Images	98.83	99.29	99.65
	VDeRe-23	Augmented+ Grey level +LBP Images	99.76	98.54	99.65

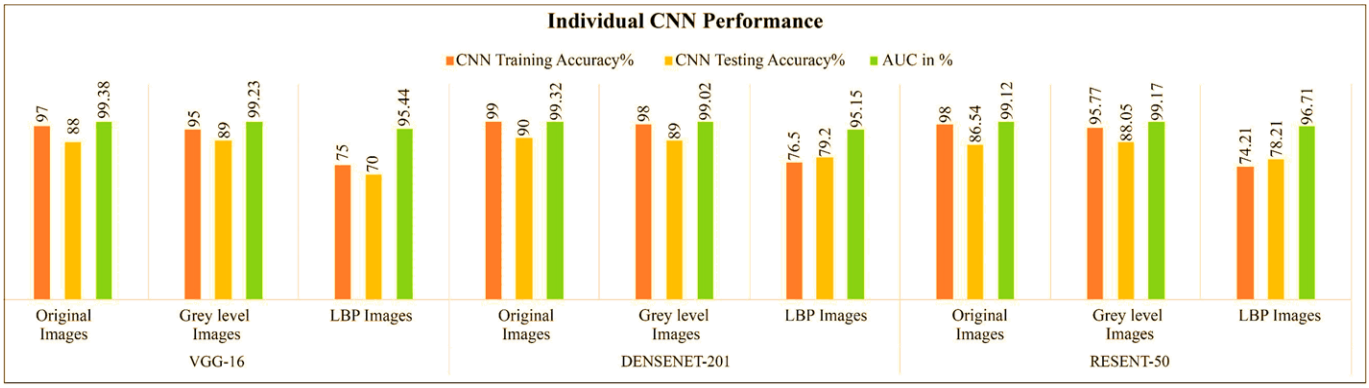


Fig. 6 Individual CNN performance on FGNET Dataset

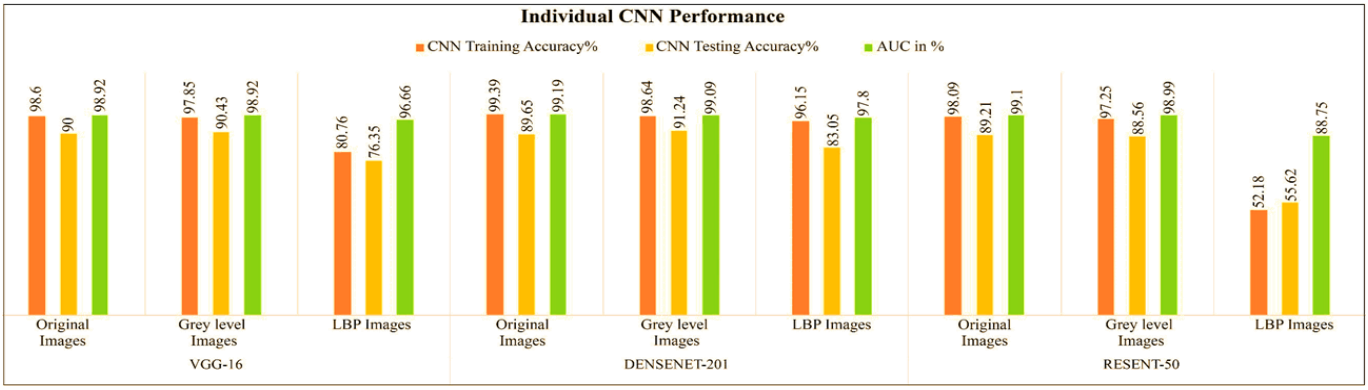


Fig. 7 Individual CNN performance on CACD Dataset

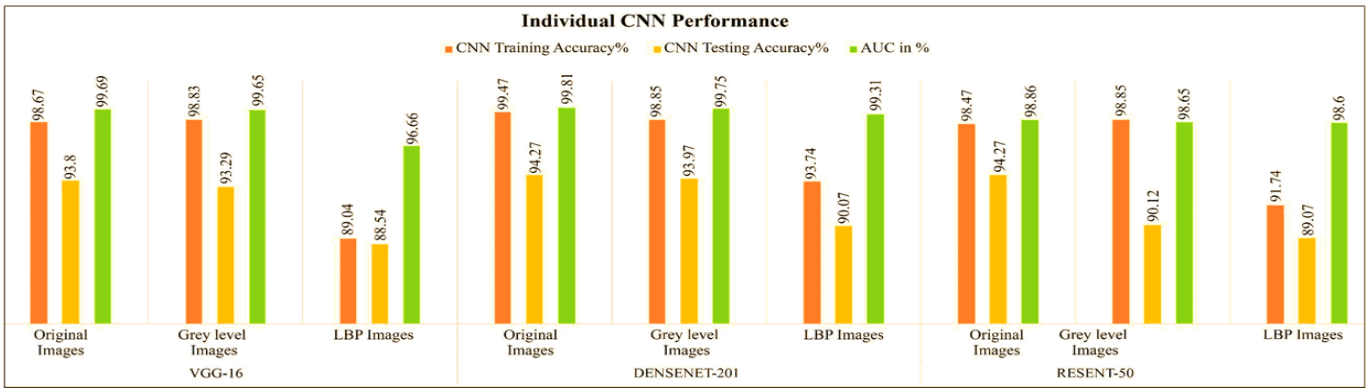


Fig. 8 Individual CNN performance on AS-23 Dataset.

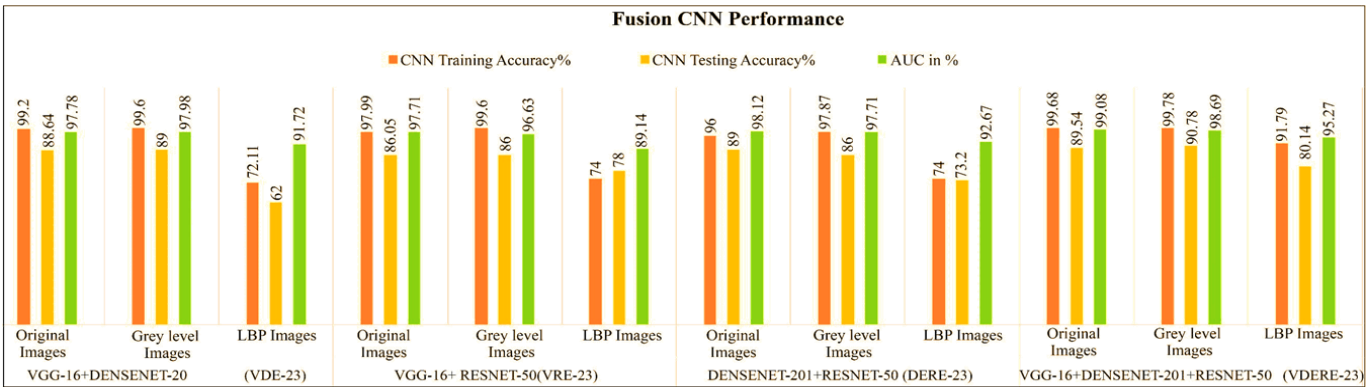


Fig. 9 Fusion CNN performance on FGNET Dataset.

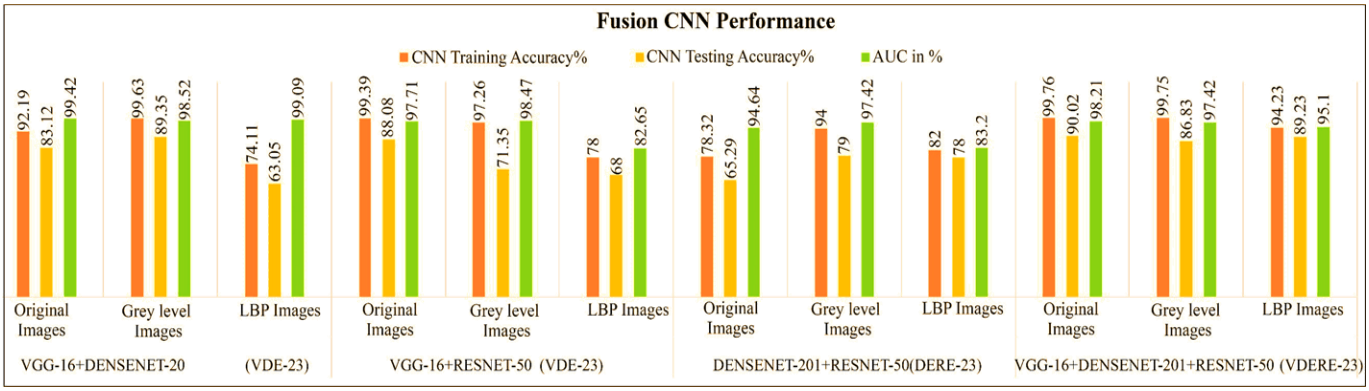


Fig. 10 Fusion CNN performance on CACD Dataset.

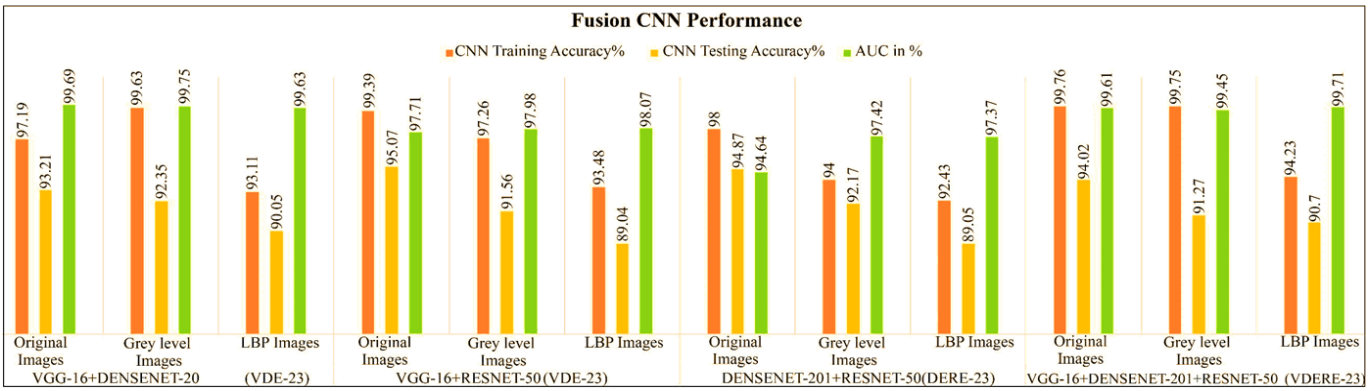


Fig. 11 Fusion CNN performance on AS-23 Dataset

The best training accuracy was obtained using the FGNET dataset with CNN DenseNet-201, with values of 99.39%, 98.64% and 96.15% and the best testing accuracy with values of 89.65%, 91.24% and 83.05%, respectively, with augmented, grey level and LBP images and with the fusion CNN Models, VDeRe-23 achieved the best training accuracy of 99.68%,99.78% and 91.79% and best testing accuracy of 89.54%,90.78% and 80.41% respectively.

The best training accuracy was obtained using the CACD dataset with CNN DenseNet-201, with values of 99.39%, 98.64% and 96.15% and the best testing accuracy with values of 89.65%, 91.24% and 83.05%, respectively, with augmented, grey level and LBP images and with the fusion CNN Models, VDeRe-23 achieved the best training accuracy of 99.76%,99.75% and 94.23% and best testing accuracy of 90.02%,86.83% and 89.23% respectively.

The best training accuracy was obtained using the CACD dataset with CNN DenseNet-201, with values of 99.47%,98.85% and 93.74% and best testing accuracy of 94.27%, 93.97% and 90.07%, respectively, with augmented, grey level and LBP images and with the fusion CNN Models, VDeRe-23 achieved the best training accuracy of 99.76%,99.75% and 94.23% and best testing accuracy of 94.02%,91.27% and 90.7% respectively. The AS-23 dataset with the feature fusion CNN, VDeRe-23, generated the highest training accuracy of 99.76% and the best testing accuracy of 98.54% in all analysis results and comparisons.

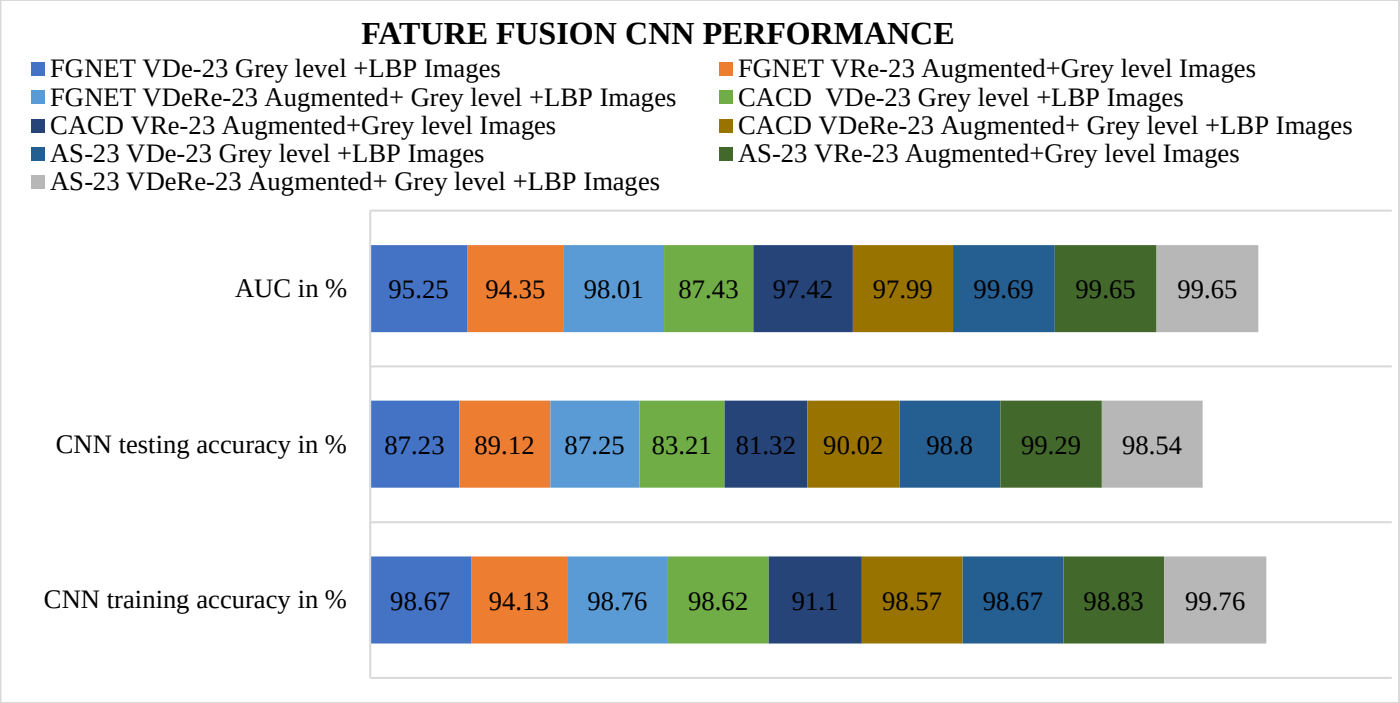


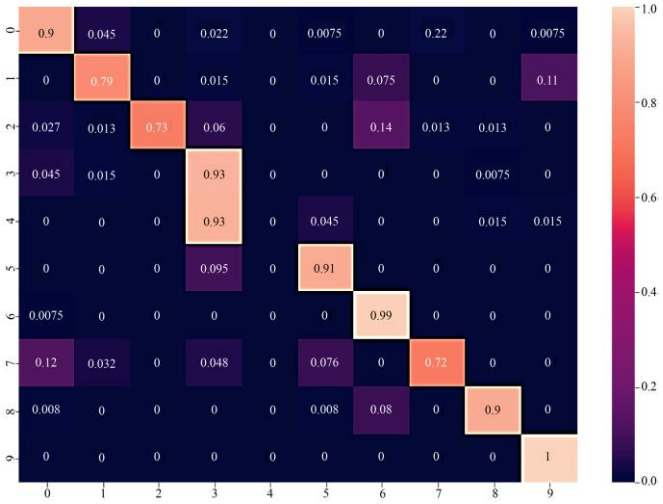
Fig. 12 Feature Fusion CNN performance comparision

4.1. Result Analysis

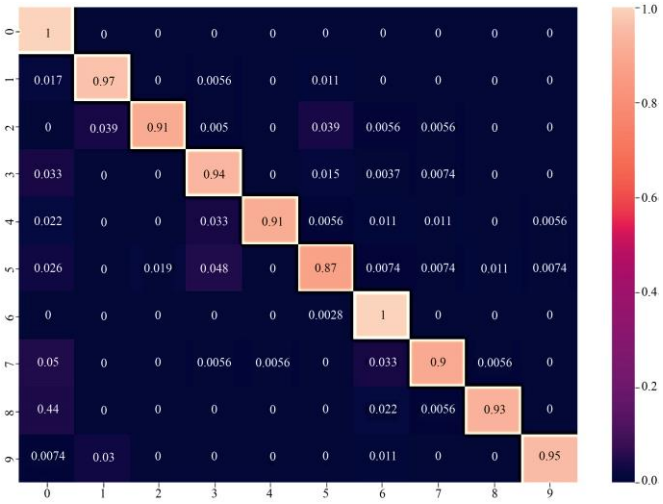
Among all the experimentation results on three datasets, the CNN DenseNet-201 achieved the maximum accuracy of 99.47%,98.85% and 93.74% and best testing accuracy of 94.27%, 93.97% and 90.07%, respectively, with augmented, grey level and LBP images. With the fusion CNN Models, VDeRe-23 achieved the best training accuracy of 99.76%,99.75% and 94.23% and the best testing accuracy of 94.02%,91.27% and 90.7%, respectively, on AS-23. The best accuracies (training:99.76%, testing:98.54%) were achieved using the AS-23 dataset with the feature fusion CNN, VDeRe-23 with augmented, grey level and LBP feature fusion. The performance measures True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are used to calculate accuracy. For the classes listed for each dataset, confusion matrices were created. The effectiveness of the assessment model can be evaluated using a confusion matrix as matrix. Every column is a predicted/estimated class, and every row is a real/true class. The projected values were contrasted with the actual ones. Thus, several diagonal elements are needed for the proper models. According to our model, all classes have diagonal values that are close to unity. The confusion matrices of the best models are shown in Figure 14.

Table 8. The comparison of the obtained results of the proposed method with the existing methods

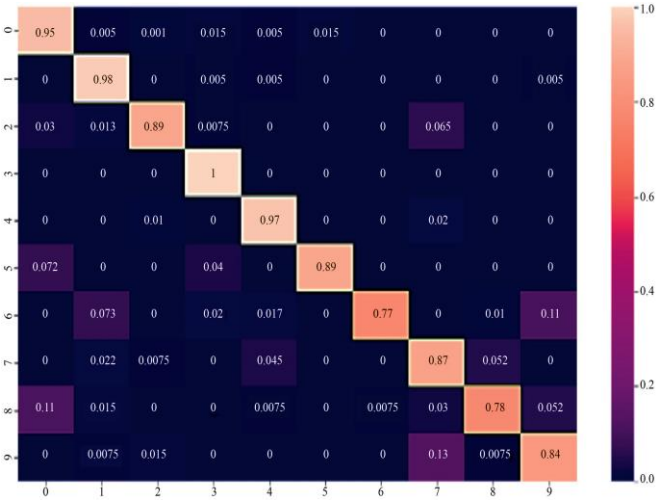
Dataset	Author	Method	Accuracy in %
FGNET	Xu et al. [21]	CAN (2017)	86.5
CACD			77
FGNET	Wen et al. [17]	LF-CNN (2016)	88.10
CACD			93.8
FGNET	Huang and Hu [30]	AA-CNN (2020)	89.34
CACD			95.1
CACD	Yan, Chenggang, et al. [31]	MFD (2022)	97.51
FGNET	Dharavath et al. [46]	Parallel Deep CNN (2022)	81.25
CACD			81
FGNET	Ours	Proposed Method (MF-FaceNet)	90.78
CACD	Ours	Proposed Method (MF-FaceNet)	91.24
AS-23	Ours	Proposed Method (MF-FaceNet)	98.54



(a)

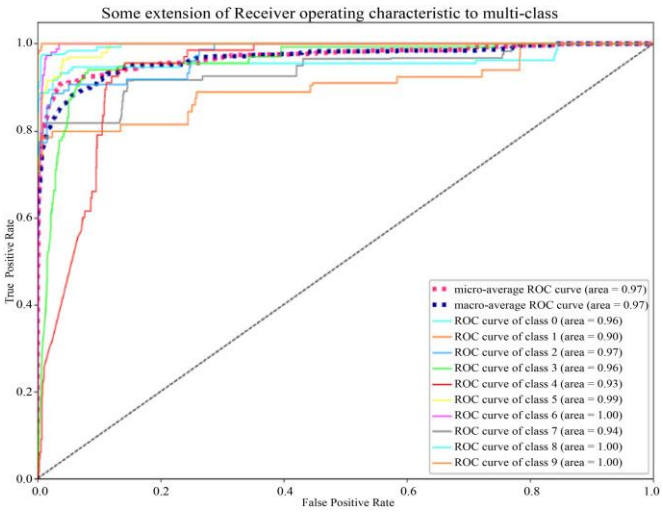


(b)

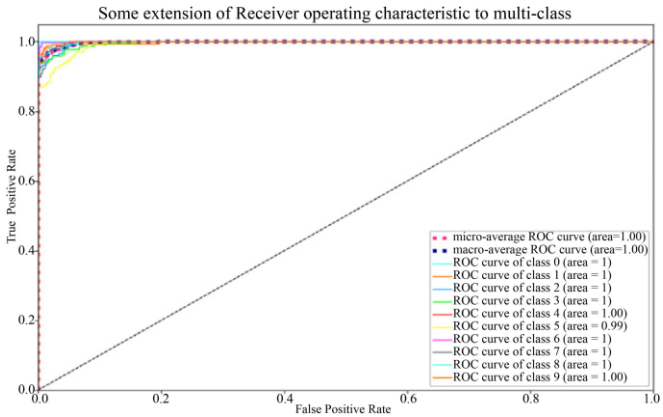


(c)

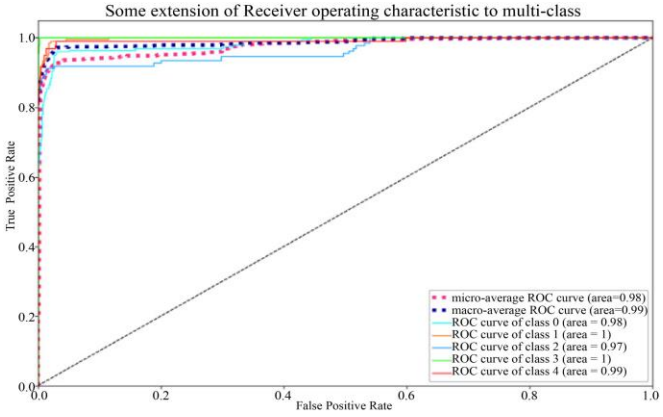
Fig. 14 The confusion matrices of AS-23 dataset with (a) Individual CNN(DenseNet-201)(b) Fusion CNN(VDeRe-23) (c) feature fusion CNN(VDeRe-23)



(a)



(b)



(c)

Fig. 15 The ROC Curves of AS-23 dataset with (a) Individual CNN(DenseNet-201) (b) Fusion CNN(VDeRe-23) (c) feature fusion CNN(VDeRe-23)

4.2. Discussions

The three-individual, four-fusion, and three-feature fusion CNN models are the main components of the MF-FaceNet system. The individual CNN, DenseNet-201, and the fusion CNN VDeRe-23 successfully recognize faces on the

AS-23 dataset with a total of ten experiments. Using the FGNET and CACD datasets, the proposed model's performance is compared with those of several AIFR models, including the most recent AA-CNN and MFD. The results are shown in Table 8. The prosed method is trained on augmented, grey and LBP cross-age face datasets. This shows that our method has good face feature learning capability to achieve good results on cross-age datasets. With all these comparisons, it is demonstrated that our approach outperforms other methods such as LF-CNN, AACNN, and MFD by achieving higher recognition accuracy of 90.78% on FGNET,90.43% on CACD AND 98.54% on AS-23 datasets. These outcomes further demonstrate the efficacy of the suggested approach.

5. Conclusion

A new framework for age-invariant face recognition called the Multi-Feature, and Multi-Fusion CNN is proposed. In this proposed system, features of the face images are extracted from augmented, Grey level, and LBP image face datasets. For each extracted feature, experiments are conducted with three individual CNN models such as VGG-16, DenseNet-201 and ResNet-50, and four fusion models such as VDe-23, VRe-23, DeRe-23 and VDeRe-23.For this experimentation with individual and Fusion CNN models, extensive results are obtained with the standard face datasets such as FGNET, CACD and a Live dataset AS-23. Among all three datasets, the best-performed dataset is the AS-23 live dataset and compared with three CNN models, the highest testing accuracy of face recognition is obtained from DenseNet-201(94.27%,93.97%,90.07%). In the AS-23 live dataset, in comparison with three fusion models, the VDeRe-23 CNN feature fusion model achieved the best face recognition testing accuracies, such as 94.02%, 91.27% and 90.7%, respectively. Among all the comparisons in each and every dataset, the DenseNet-201 achieved the best face recognition accuracy; for the fusion model, VDeRe-23 gives the best recognition accuracy. Compared to well-known datasets, our method improves the AIFR performance on the college dataset (AS-23). Our approach performs better than the existing methods and offers a high level of recognition accuracy. In future works, the proposed CNN models can be applied in Multi-Task Learning (MTL) that can simultaneously learn identity-sensitive features and age-sensitive features to obtain robust AIFR along with age classification.

References

[1] Shui-Guang Tong, Yuan-Yuan Huang, and Zhe-Ming Tong, "A Robust Face Recognition Method Combining LBP with Multi-Mirror Symmetry for Images with Various Face Interferences," *International Journal of Automation and Computing*, vol. 16, pp. 671-682, 2019. [CrossRef] [Google Scholar] [Publisher link]

[2] P.Bala Saiteja et al., "Enhanced Security for ATM Transactions using Facial Verification," *SSRG International Journal of Electronics and Communication Engineering*, vol. 3, no. 3, pp. 5-7, 2016. [CrossRef] [Publisher link]

[3] Yi Sun et al., "Deep Learning Face Representation by Joint Identification-Verification," *Advances in Neural Information Processing Systems*, vol. 27, 2014. [Google Scholar] [Publisher link]

[4] Anil K. Jain, Karthik Nandakumar, and Arun Ross, "50 Years of Biometric Research: Accomplishments, Challenges, and Opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80-105, 2016. [CrossRef] [Google Scholar] [Publisher link]

[5] Yi Sun et al., "Deepid3: Face Recognition with Very Deep Neural Networks," *arXiv:1502.00873*, 2015. [CrossRef] [Google Scholar] [Publisher link]

[6] Omkar M. Parkhi, Andrea Vedaldi, and Andrew Zisserman, "Deep Face Recognition," *Proceedings of the British Machine Vision Conference (BMVC)*, Swansea, pp. 41.1–41.12, 2015. [CrossRef] [Google Scholar] [Publisher link]

[7] Florian Schroff, Dmitry Kalenichenko, and James Philbin "FaceNet: A Unified Embedding for Face Recognition and Clustering," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 815–823, 2015. [CrossRef] [Google Scholar] [Publisher link]

[8] Qiong Cao et al., "VGGFace2: A Dataset for Recognising Faces Across Pose and Age," *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition, IEEE*, pp. 67-74, 2018. [CrossRef] [Google Scholar] [Publisher link]

[9] Xiang Wu et al., "A Light CNN for Deep Face Representation Noisy Labels," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2884–2896, 2018. [CrossRef] [Google Scholar] [Publisher link]

[10] Chandan A D et al., "Survey Paper on Vehicle Security using Facial Recognition & Password," *SSRG International Journal of Electronics and Communication Engineering*, vol. 9, no. 6, pp. 5-9, 2022. [CrossRef] [Publisher link]

[11] Unsang Park, Yiying Tong, and Anil K. Jain, "Age-Invariant Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 947–954, 2010. [CrossRef] [Google Scholar] [Publisher link]

[12] Haibin Ling et al., "Face Verification Across Age Progression Using Discriminative Methods," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 1, pp. 82–91, 2010. [CrossRef] [Google Scholar] [Publisher link]

[13] Joyassree Sen et al., "Face Recognition Using Deep Convolutional Network and One-shot Learning," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 4, pp. 23-29, 2020. [CrossRef] [Google Scholar] [Publisher link]

[14] Bor-Chun Chen, Chu-Song Chen, and Winston H. Hsu, "Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval," *Lecture Notes in Computer Science*, vol. 8694, pp. 768–783, 2014. [CrossRef] [Google Scholar] [Publisher link]

[15] Garima Verma et al., "A Technique for Face Verification Across Age Progression with Large Age Gap," *4th International Conference on Signal Processing, Computing and Control (ISPCC)*, vol. 2017, pp. 603–607, 2017. [CrossRef] [Google Scholar] [Publisher link]

[16] Ya Li et al., "A Deep Joint Learning Approach for Age Invariant Face Verification," *CCF Chinese Conference on Computer Vision*, Springer, Berlin, Heidelberg, pp. 296-305, 2015. [CrossRef] [Google Scholar] [Publisher link]

[17] Yandong Wen, Zhifeng Li, and Yu Qiao, "Latent Factor Guided Convolutional Neural Networks for Age-Invariant Face Recognition," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4893–4901, 2016. [CrossRef] [Google Scholar] [Publisher link]

[18] Nazia Begum, and Syed Noorullah, "Biometric Uncorrelated Face Recognition Using Unsupervised Learning," *International Journal of Computer Engineering in Research Trends*, vol. 7, no. 3, pp. 13–18, 2020. [CrossRef] [Publisher link]

[19] Simone Bianco, "Large Age-Gap Face Verification by Feature Injection in Deep Networks," *Pattern Recognition Letters*, vol. 90, pp. 36–42, 2017. [CrossRef] [Google Scholar] [Publisher link]

[20] Ya Li et al., "Distance Metric Optimization Driven Convolutional Neural Network for Age Invariant Face Recognition," *Pattern Recognition*, vol. 75, pp. 51–62, 2018. [CrossRef] [Google Scholar] [Publisher link]

[21] Chenfei Xu, Qihe Liu, and Mao Ye, "Age Invariant Face Recognition and Retrieval by Coupled Auto-Encoder Networks," *Neurocomputing*, vol. 222, pp. 62–71, 2017. [CrossRef] [Google Scholar] [Publisher link]

[22] Opeyemi Oyelesi, and Akingbade Kayode Francis, "Face Recognition for Access Control using PCA Algorithm," *SSRG International Journal of VLSI & Signal Processing*, vol. 4, no. 2, pp. 22-27, 2017. [CrossRef] [Publisher link]

[23] Sai Vamshi Challamalla et al., "Online Attendance Management System Using Face Recognition Techniques," *International Journal of Computer Engineering in Research Trends*, vol. 9, no. 8, pp. 158–169, 2022. [CrossRef] [Publisher link]

[24] Sabah Afroze et al., "Age Invariant Face Recognition using Frangi2D Binary Pattern," *ACM International Conference Proceeding Series*, pp. 8–13, 2019. [CrossRef] [Google Scholar] [Publisher link]

[25] Amal A. Moustafa, Ahmed Elnakib, and Nihal F. F. Areed, "Optimization of Deep Learning Features for Age-Invariant Face Recognition," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 2, pp. 1833–1841, 2020. [CrossRef] [Google Scholar] [Publisher link]

[26] Ankit Jain, Kirti Bhatia, Rohini Sharma, and Shalini Bhadola, "An Overview on Facial Expression Perception Mechanisms," *SSRG International Journal of Computer Science and Engineering*, vol. 6, no. 4, pp. 19-24, 2019. [CrossRef] [Publisher link]

[27] Hao Wang et al., "Decorrelated Adversarial Learning for Age-Invariant Face Recognition," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, vol. 2019, pp. 3522–3531, 2019. [CrossRef] [Google Scholar] [Publisher link]

[28] Amal A. Moustafa, Ahmed Elnakib, and Nihal F. F. Areed, "Age-Invariant Face Recognition Based on Deep Features Analysis," *Signal, Image Video Processing*, vol. 14, no. 5, pp. 1027–1034, 2020. [CrossRef] [Google Scholar] [Publisher link]

[29] Bhumika Pathya, and Sumita Nainan, "Performance Evaluation of Face Recognition using LBP, PCA and SVM," *SSRG International Journal of Computer Science and Engineering*, vol. 3, no. 4, pp. 58-61, 2016. [CrossRef] [Google Scholar] [Publisher link]

[30] Yangjian Huang, and Haifeng Hu, "A Parallel Architecture of Age Adversarial Convolutional Neural Network for Cross-Age Face Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 1, pp. 148-159, 2021. [CrossRef] [Google Scholar] [Publisher link]

[31] Chenggang Yan et al., "Age-Invariant Face Recognition by Multi-Feature Fusionand Decomposition with Self-Attention," *ACM Transactions on Multimedia Computing, Communications, and Applications*, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher link]

[32] D. J. Samatha Naidu, and R. Lokesh, "Missing Child Identification System using Deep Learning with VGG-FACE Recognition Technique," *SSRG International Journal of Computer Science and Engineering*, vol. 9, no. 9, pp. 1-11, 2022. [CrossRef] [Publisher link]

[33] Mubarak Albarka Umar et al., "Fighting Crime and Insecurity in Nigeria: An Intelligent Approach," *International Journal of Computer Engineering in Research Trends*, vol. 8, no. 1, pp. 6–14, 2021. [CrossRef] [Google Scholar] [Publisher link]

[34] Caifeng Shan, Shaogang Gong, and Peter W. McOwan, "Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009. [CrossRef] [Google Scholar] [Publisher link]

[35] R Sindhoori, "Digital image processing. Multi feature face recognition in PSO -SVM," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 1, no. 3, pp. 1-6, 2014. [CrossRef] [Publisher link]

[36] Prasanna Rajendra, Niket Worlikar, Yash Mahajan, Siddhesh Swami, "Smart Surveillance Using OpenCV, Motion Analysis And Facial Landmark," *SSRG International Journal of VLSI & Signal Processing*, vol. 7, no. 1, pp. 11-14, 2020. [CrossRef] [Publisher link]

[37] Gao Huang et al., "Densely Connected Convolutional Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 2261-2269, 2017. [CrossRef] [Google Scholar] [Publisher link]

[38] Noha Ghatwary, Xujiong Ye, and Massoud Zolgharni, "Esophageal Abnormality Detection Using Densenet Based Faster R-CNN with Gabor Features," *IEEE Access*, vol. 7, pp. 84374–84385, 2019. [CrossRef] [Google Scholar] [Publisher link]

[39] D. Betteena Sheryl Fernando et al., "Face Recognition For Home Security," *SSRG International Journal of Computer Science and Engineering*, vol. 6, no. 10, pp. 7-12, 2019. [CrossRef] [Publisher link]

[40] A. Lanitis, C.J. Taylor, and T.F. Cootes, "Toward Automatic Simulation of Aging Effects on Face Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442–455, 2002. [CrossRef] [Google Scholar] [Publisher link]

[41] Dihong Gong et al., "Hidden Factor Analysis for Age Invariant Face Recognition," *IEEE International Conference on Computer Vision*, pp. 2872–2879, 2013. [CrossRef] [Google Scholar] [Publisher link]

[42] Yitong Wang et al., "Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition," *European Conference on Computer Vision*, vol. 11219, pp. 764–779, 2018. [CrossRef] [Google Scholar] [Publisher link]

[43] M. Saad Shakeel, and Kin-Man Lam, "Deep-Feature Encoding-Based Discriminative Model for Age-Invariant Face Recognition," *Pattern Recognition*, vol. 93, pp. 442–457, 2019. [CrossRef] [Google Scholar] [Publisher link]

[44] Lingshuang Du, Haifeng Hu, and Yongbo Wu, "Age Factor Removal Network Based on Transfer Learning and Adversarial Learning for Cross-Age Face Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 9, pp. 2830-2842, 2020. [CrossRef] [Google Scholar] [Publisher link]

[45] Shiji S K, and Dr. S.H Krishnaveni, "An Analytical Approach for Reconstruction of Cosmetic Surgery Images using EUCLBP and SIFT," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 9, no. 8, pp. 60-71, 2022. [CrossRef] [Publisher link]

[46] Krishna Dharavath, Amarnath Gaini, and Viod Adla, "A Parallel Deep Learning Approach for Age Invariant Face Recognition system," *2022 International Conference on Computer Communication and Informatics, IEEE*, pp. 1-6, 2022. [CrossRef] [Google Scholar] [Publisher link]

[47] Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556*, 2015. [CrossRef] [Google Scholar] [Publisher link]

[48] Kaiming He et al., "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770–778, 2015. [CrossRef] [Google Scholar] [Publisher link]

[49] Hae-Min Moon, Chang Ho Seo, and Sung Bum Pan, "A Face Recognition System based on Convolution Neural Network Using Multiple Distance Face," *Soft Computing*, vol. 21, no. 17, pp. 4995-5002, 2017. [CrossRef] [Google Scholar] [Publisher link]