

New Approaches for Automatic Face Recognition Based on Deep Learning Models and Local Handcrafted ALTP

Abdessalam Hattab^{1,*} and Ali Behloul¹

¹LaSTIC laboratory, computer science department, University of Batna 2, 05000, Algeria

Abstract

Face recognition is one of the most important topics in biometrics, where it achieved great success under controlled scenarios. Still, its accuracy degraded significantly in unconstrained conditions. To meet this challenge, we proposed a handcraft method based on extracting important regions from the face image. We have been using Scale-Invariant Feature Transform (SIFT) besides the Adaptive Local Ternary Patterns (ALTP). We have achieved an accuracy of 99.75% on the ORL database and 94.70% on the FERET database. Then, we proposed a second method based on deep learning to achieve more accurate face recognition. The deep learning models failed to achieve a high accuracy rate because they require a large amount of training data. We used firstly Data Augmentation to solve this failure. However, this solution does not show high performance. Secondly, our proposed ImageNet pre-trained AlexNet-v2 and VGG16 models with LinearSVC increased the accuracy rate to 100% for the both databases.

Keywords: Face recognition, Adaptive Local Ternary Pattern (ALTP), Scale-Invariant Feature Transform (SIFT), Deep Learning, Transfer Learning, LinearSVC.

Received on 29 June 2021, accepted on 15 October 2021, published on 20 October 2021

Copyright © 2021 Abdessalam Hattab *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.20-10-2021.171547

1. Introduction

With the increase of users' personal information in mobility devices, we are obligated to use systems for personal identification. Face recognition is one of the best ways for person identification because that has high accuracy and low intrusiveness [1], and it doesn't need the cooperation of the user [2]. Classical face recognition approaches generally contain two categories: local approach and holistic approach. The holistic approach is based on using the whole face region for recognizing the face [3]. The best two examples in this category are Principal Component Analysis (PCA) [4] and Fisher Linear Discriminant Analysis (LDA) [5].

On the contrary, local approach methods are based on exploiting the local facial features for recognizing the face, such as eyes, ears, nose, mouth, outline of the face, etc. These methods are divided into local appearance-based techniques and key points-based techniques. Local appearance-based technique includes LBP [6], ULBP [7], LTP [8], CS-LTP[9], ALTP [10], and the key points based techniques such as Scale

Invariant Feature Transform (SIFT) [11], as well as Speeded Up Robust Features (SURF) [12]. The local-based approaches are robust to most face recognition challenges (variations in illumination, facial expression, etc.) compared to global-based approaches [13]. The methods using the classical approaches have achieved a very respectable performance for face identification. Nevertheless, in uncontrolled environments (illumination, pose, and expression), the performance degrades little. To solve this problem, we turned to the Deep Learning (DL) approaches. In the last several years, Deep Learning has been achieving a remarkable performance in several domains. DL achieved a high accuracy rate in various fields such as: Healthcare [14][15], biometric recognition [16][17], Text analysis [18][19].

Convolutional Neural Network (CNN) is a Deep Learning approach that is most commonly used to solve computer vision problems, where CNN achieved a high-performance face recognition accuracy in uncontrolled environments [20].

In the last decade, face recognition accuracy achieved great development through the evolution of CNN architectures,

*Corresponding author. Email: a.hattab@univ-batna2.dz

where the Deep CNN architectures appeared such as AlexNet [21][22], VGG16 [23], Inception-v3 [24], ResNet50 [25], and Xception [26]. To achieve a good performance with deep CNN, we need a large database. To overcome this limitation, we used two techniques: transfer learning and data augmentation. Transfer learning is a machine learning technique that uses knowledge learned from one task to help another related task [27]. Using transfer learning helps to train the deep networks if there is insufficient data to train these networks from scratch. Also, we protect computational resources with transfer learning because training a deep network from scratch may take a long time (days to weeks) [28]. As for data augmentation, it is a technique used to increase the volume of the training set by applying random transformations to the original images, such as scaling, rotation, zooming, horizontal or vertical mirroring, etc.

In the last two decades, face recognition has considerably developed where it played an important role in many applications in different fields such as Information Security, Automatic Access Control, Law Enforcement, Smart Cards, Video Surveillance and Smart Home Security Application. However, facial recognition is still one of the main challenging problems to be solved by researchers in the computer vision field. The two important face recognition challenges are the recognition of human faces in unconstrained environments and the training of models using Small-Samples Datasets.

The aim of this study is to propose two methods that can resist these challenges and achieve a high-accuracy face recognition.

In this work, we propose two face recognition methods, the first one is based on both categories of classical approaches. We chose to use the Adaptive Local Ternary Pattern (ALTP) descriptor for the first category because ALTP is more robust to the noise and automatically sets the threshold using Weber's law [29]. The threshold t has a direct effect on noise suppression. For the second category, we use SIFT to detect key points in face images. In the second proposed method based on the deep learning category, we used a pre-trained CNN for extract features and then we used the linearSVC for the classification.

The rest of this paper proceeds as follows: Section 2 describes the related works briefly. We present a background on the face recognition approaches in Section 3. The proposed methods are described in Section 4. In Section 5, experimental results are discussed. Finally, we conclude the paper in Section 6.

2. Related works

Several papers have been published in the literature about face recognition, where many methods had shown very satisfactory results in this field.

Local binary patterns (LBP) [6] is a famous local texture feature descriptor, the first publication mentioned the LBP descriptor emerged by Harwood et al. in the mid-1990s, but in this publication, the LBP descriptor used only to measure the local contrast of an image [30]. In 1996 Ojala et al.

proposed the LBP descriptor as a texture descriptor, and it is used after in several applications like (facial recognition and Facial expression recognition) [31], [32].

In 2015 Shyam et al. [33] proposed a novel method for recognizing faces in uncontrolled environments, based on Local Binary Pattern (LBP) for facial images.

Directional Local Binary Patterns (dLBP) [34] used the central pixel parameter to determine the neighbours in the same orientation, dLBP proposed in 2015. Arisandi et al. [35] developed a real-time phone application for face recognition. The application used Fisherface to recognize students; the accuracy of this application is 90%.

Local ternary patterns (LTP) are extensions of local binary patterns (LBP) to 3 values codes, proposed by Tan et Triggs[8] as an enhancement to the LBP who is sensitive to noise. A hybrid approach based on Improved Kernel Linear Discriminant Analysis (IKLDA) and Probabilistic Neural Networks (PNNs) is proposed by Ouyang et al. [36] For face recognition. This approach can obtain an average recognition accuracy of 83.8% on the YALE database and 97.22% on the ORL database. To enhance the face recognition rate on the ORL database, Kak et al. [37] used Discrete Wavelet Transform with Eigenface; the best face recognition rate result obtained by this method is 99.25%.

Ameen et al. [38] proposed a hybrid method for face recognition using the Local Ternary Pattern (LTP) and Singular Value Decomposition (SVD), the recognition rate on the ORL database was 98%. Aldhahab et al. [39] proposed a system for face recognition based on facial parts detection (nose, both eyes, and mouth) integrated with Vector Quantization (VQ), the recognition rate on the ORL database was surpassed 98%.

Yang et al.[10] proposed a new feature descriptor called ALTP. This descriptor had good face recognition performance where it achieved a maximum accuracy of 97.5% on the ORL database and 94% on FERET databases. The Scale Invariant Feature Transform (SIFT) algorithm is based on key points extracted from the image. SIFT is presented by David Lowe in 1999 [40] to detect local image features; five years later, David Lowe improved the SIFT algorithm for feature matching in natural images [11]. SIFT features are invariant against image zooming, translation, and rotation. Bicego et al. [41] used the SIFT operator in face authentication and obtained a good accuracy on BANCA databases. Geng et al. [42] proposed two approaches: The Keypoints-Preserving-SIFT (KPSIFT) and Partial-Descriptor-SIFT (PDSIFT); the approaches achieved good results in face recognition. Rotation-invariant features based on directional coding proposed by Ousliman et al. [43] for texture classification; this method can obtain an average recognition accuracy of 90.63% on the YALE B database.

Tabejamaat et al. [44] (2020) suggested a novel face recognition approach named Local Comparative Decimal Pattern. This approach is based on extract features texture from face images, and it gives a good accuracy under illumination and poses variations.

In 2020, a novel local descriptor named Dense Local Graph Structure (D-LGS) was proposed by Kumar et al. [45]. They used this descriptor for face recognition. The descriptor

generates the descriptive image from the input image by applying a bilinear interpolation to increase the pixel density. D-LGS gave acceptable results in both constrained and unconstrained environments. Alternative representations of facial images used by Qin et al. [46] for face recognition. This method achieved maximum accuracy of 95 % on the ORL database.

Mixed Transform and Multilayer Sigmoid Neural Network Classifier was proposed by Sapijaszko et al. [47] for face recognition. This method used the grayscaleing algorithm to enhance the face image. The method extracts features using the two-dimensional DWT (Discrete Wavelet Transform) and the two-dimensional DCT (discrete cosine transform). For the classification, they used a multilayer sigmoid neural network. The Convolutional Neural Networks (CNN) based methods have obtained accuracy better than many well-known classical methods. In [48], authors proposed a method based on CNN's where it has been achieved 98.3 % on the ORL database.

Several Deep Convolutional Neural Network (DCNN) models have won great success in the facial identification field, such as VGG16, ResNet50, and Inception-v3 [49]. In 2017 Wang et al.[50] proposed a hybrid method based on LBP and DCNN. The LBP features of the faces are used as inputs of the DCNN, and then we train this network with the LBP features. Finally, they use the trained network for face identifications; this method achieved 95.6% on the Feret database.

Zeng et al. [51] proposed combining traditional (handcrafted) and Deep Learning features for face recognition. The authors used transfer learning to extract deep features. In [52] authors proposed a face identification algorithm based on transfer learning and proposed a sample expansion method. A deep convolutional neural network is pre-trained on a common multi-sample face dataset and then applied the model to the target data set. Then, a sample expansion method called K Class Feature Transfer (KCFT) was used to enrich intra-class; this algorithm achieved better than 98% on the ORL database.

3. Background

Several methods were proposed for face recognition. We can divide then into two categories :

3.1. Classical methods (handcraft methods)

Many methods have been developed to achieve a better face recognition accuracy based on handcrafted feature extractions. The most famous of these methods are mentioned as follows :

a) Local binary patterns (LBP) :

The LBP operator calculates the relationship between the pixel and its immediate neighbours, the special local structure of the image is defined as (1).

$$LBP_{R,N} = \sum_{k=0}^{N-1} s(p_k - p_c) 2^k, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

P_c represents the grey intensity value of the centre pixel, p_k represents the grey intensity value of the neighbour pixels of p_c , N represents the number of pixels equally spaced on a centre of a circle of radius R (see Figure 1).

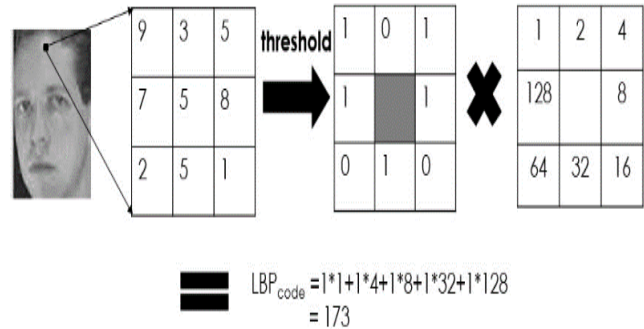


Figure 1. Illustration of the basic local binary patterns operator.

b) Local ternary patterns (LTP) :

Local ternary patterns (LTP) are extensions of local binary patterns (LBP) to 3 values codes, proposed by Tan et Triggs [8] as an enhancement to the LBP who is sensitive to noise. LTP assigns the value 0 to the pixels whose value is in a zone of width $\pm t$ around the central pixel (p_c), assigns the value 1 for the pixels above and -1 for the pixels below.

The LTP code of a pixel (u) is calculated as follows:

$$S(u, p_c, t) = \begin{cases} 1, & u \geq p_c + t \\ 0, & |u - p_c| < t \\ -1, & u \leq p_c - t \end{cases} \quad (2)$$

Figure 2. illustrates an example of the LTP operator with $t=5$.

Each ternary pattern is divided into positive and negative halves Figure 3. We calculate both histograms, a histogram of the positive parts and a histogram of the negative parts, concatenate both histograms to get the final LTP descriptor of an image.

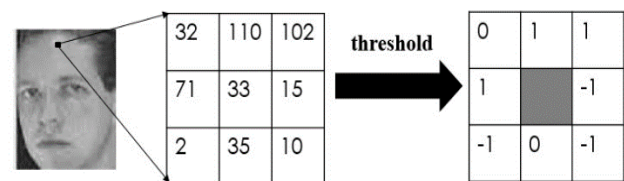


Figure 2. Illustration of the basic Local ternary patterns operator with $t=5$.

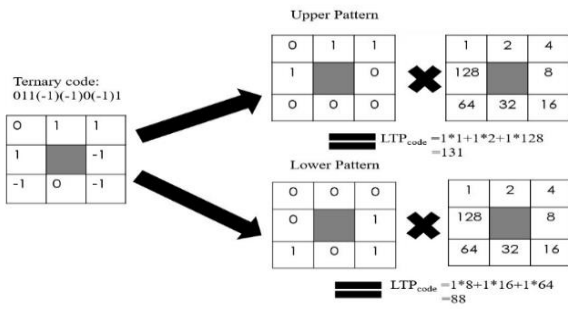


Figure 3. Illustration of Splitting a Local ternary patterns code into positive and negative code

3.2 Deep learning methods

Recently, deep learning has won great success in various fields such as computer visions, natural language processing, etc.

Deep Convolutional Neural Networks (DCNN) achieved high accuracy in image classification and face recognition fields.

A DCNN architecture is based on some basic layers:

- Convolution Layer (ConvL): is the most important layer on DCNN, used to generate Feature maps based on the features extracted from an input image.
- Maxpooling Layer: The feature maps generated by the Convolution layer are generally high dimensions. To reduce their dimensionality and obtain the most important features, we use the Maxpooling Layer.
- Fully Connected Layer: It is the same as a multi-layer perceptron (MLP), which connects each Neuron in a fully connected layer to all activations in the previous layer.

Recently several Deep CNN models have been proposed for image recognition. In this paper, we mainly focus on AlexNet-v2 [22] and VGG16[23].

VGG16 is a deeper CNN model proposed by Simonyan et al. [23] in 2014, which contains 13 convolutional layers, followed by ReLU, five max polling layers, and three fully connected layers. VGG16 are divided into five groups, and a max-pooling layer ends each group. The input images in the VGG16 model are 224 x 224 dimensions, and the final features are 512 x 7 x 7 before they enter into the first fully connected layers. The architecture of VGG-16 is shown in figure 4.

AlexNet-v2 is a deeper CNN model proposed by Simonyan et al. [23] in 2014, which contains five convolutional layers, followed by ReLU, three Maxpolling layers, and three fully-connected layers. The entered images in the AlexNet-v2 model are 224 x 224 dimensions, and the final features are 256 x 6 x 6 before they enter into the first fully connected layers. The architecture of AlexNet-v2 is shown in figure 5.

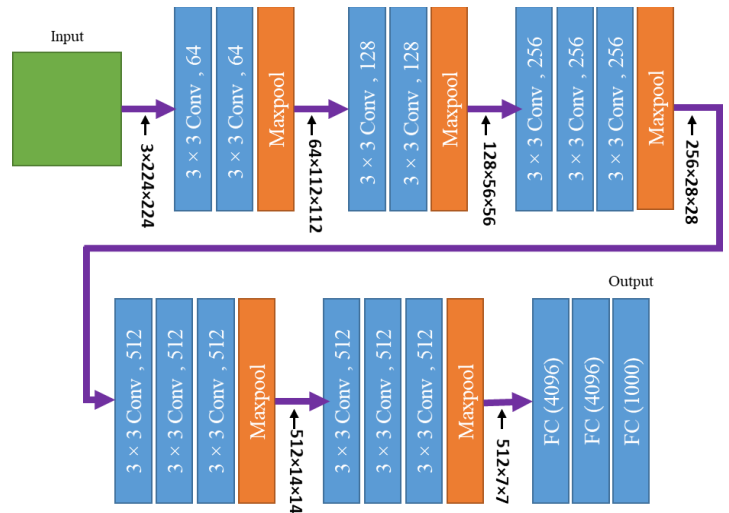


Figure 4. The architecture of VGG16[23]

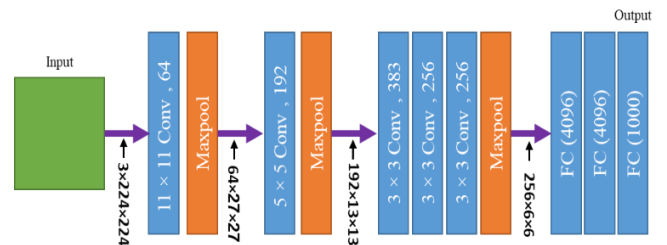


Figure 5. The architecture of AlexNet-v2[22]

Despite that, the lesser amount of training data and the limitation of the computing resources becomes a handicap to train a deep CNN. To solve this problem, researchers use data augmentation or transfer learning. As DCNN needs many samples to be trained, transfer learning can be applied to train DCNN with lesser amounts of data and without High-performance computing resources.

4. The proposed methods

In this section, we propos a robust face recognition method based on handcrafted feature extraction. To develop an accurate method that achieves more accuracy than the handcraft method proposed, we have applied Deep Learning.

4.1 Handcraft method

We present a hybrid method to recognize faces by combining two techniques: key-points extraction and features extraction. Our method is based on extracting the important regions from the face image by using the key-points extraction technique and then applying the feature extraction technique for the regions extracted in the first step.

We apply Scale-Invariant Feature Transform (SIFT) to extract key points, and then we extract a block of pixels around each key-points. Also, we use Adaptive Local Ternary Patterns (ALTP) to extract features from the extracted blocks.

After that, we count the ALTP histogram of each block. We used the X2 distance between the ALTP histograms of the testing image and each person's ALTP histograms that have been stored in the training phase to recognize faces. Figure 6. shows a detailed diagram of our method. So our proposed method is based on three main steps that can be thoroughly described as follows:

a) Extract Key-points and get blocks

The extraction of the SIFT key-points is described in the following steps:

- Constructing the Difference of Gaussian Pyramid:

First, we represent the image in scale-space where the scale space of an image is defined as $L(x, y, \sigma)$, which is calculated by the convolution of a variable-scale Gaussian $G(x, y, \sigma)$ with an entered image $I(x, y)$ as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

Whereas $G(x, y, \sigma)$ is calculated as follows:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{4}$$

Second, we calculate the Difference of Gaussian between two nearby scales :

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \tag{5}$$

Extrema detection :

To detect the local maxima and minima in the DOG pyramid, each pixel is compared to its (3×3) neighbours in the scale below, eight neighbours in the current image, and (3×3) neighbours in the scale above. The point is an extrema if its value is the maximum or minimum.

- Key points elimination:

Discard key points with Low-Contrast and used Hessian matrix to eliminating edge responses.

After the key points are extracted, we select 32×32 pixels around each key point, to compute features.

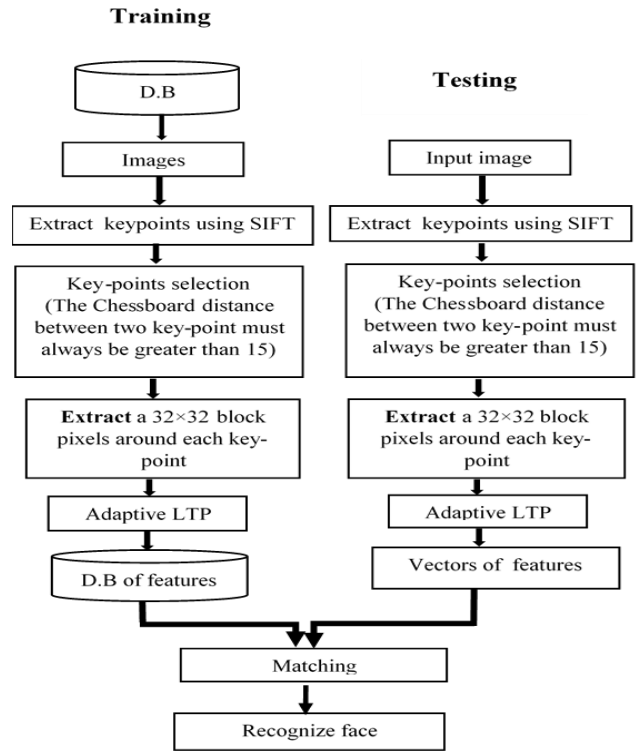


Figure 6. diagram of the proposed method

b) calculate ALTP of the blocks extracted and count the histograms

Local Ternary Patterns (LTP) uses thresholds to prevent the influence of noise, but the manually set of threshold in real application uses different data sets containing different faces is not realistic. ALTP solves the problem of manually threshold selection in LTP. The strategy of automatic threshold selection in ALTP inspired by Weber's Law is as follows:

$$t = p_c \times k \tag{7}$$

Where t is the threshold, k is Weber's law parameter, and p_c represents the grey intensity value of the centre pixel. After ALTP codes are calculated, count the histogram of each code.

c) Matching

We calculate the sum of the minimum X2 histogram distance between the histograms of the testing image and the histograms stored of each person, and then we choose the person with the smallest distance.

4.2 deep learning method

We have analysed CNN architectures to obtain a high accuracy for face recognition. Firstly, we trained deep CNN models without data augmentation; Then, we added data augmentation techniques. Finally, we used transfer learning.

DEEP CNN :

After we trained (VGG16, AlexNet-v2) with data, we extracted the features before they entered into the first fully connected layers, then we used the linearSVC¹ algorithm for classification or used PCA to reduce the dimensionality and linearSCV for the classification.

DEEP CNN with DATA AUGMENTATION:

We use data augmentation to achieve a high face recognition accuracy rate because it has been successfully applied in the literature to increase the accuracy rate of the different deep learning models[53].

We augmented our database using Flip, Zoom, and Random Rotation Augmentation, and we used the augmented data to train VGG16 and AlexNet-v2 model. We extracted the features after they were leaving from the last Max Pooling layer, then we used the linearSVC classifier or used PCA to reduce the dimensionality and then the linearSCV classifier.

PRE-TRAINED DEEP CNN :

We used the pre-trained model (Alex-Net-v2 or VGG16). The models have been trained on ImageNet Dataset[54]. ImageNet Dataset contains 1000 categories and 1.4 million images, 17% of all images contain at least one face [55]. We obtained the features from the last max pooling layer then used the linearSVC or used PCA to reduce the dimensionality and linearSCV for classification.

5. Experiments

This work proposes two methods for face recognition; the recognition rate performance of the proposed methods was tested on ORL and FERET databases, where we take a subset for each database for testing, whereas the four rest subsets are used for training.

We illustrate the efficiency of our methods by comparing their performance with some recent works.

The handcraft methods were carried out using a PC with Intel Core i7-8550U CPU and 16 GB of RAM. The programming language was python 3.8.

We have run all the deep learning experiments on Google Colab² with the help of the open-source libraries PyTorch[56] and Keras[57].

5.1 Data sets:

We conducted evaluation experiments on two different face datasets:

The ORL database

The ORL database³ was published by AT&T Laboratories of Cambridge university. ORL database was collected between

April 1992 and April 1994. The ORL database comprises 400 images with a size of (112×92) with 256 grey levels per pixel. There are 40 individuals, ten images per person. Figure 7. shows some images from the ORL database.

The images were taken at different facial expressions with a variation of time, and some persons have worn glasses. All images were collected on a dark background. In our experiments, this database is divided into five sets (Sub1, Sub2, Sub3, Sub4, and Sub5). Each subset is composed of 80 images, two images per person.

The FERET database:

The FERET database[58] was developed by the Facial Recognition Technology (FERET) program. The scientists used the FERET database to develop face recognition algorithms, test and evaluate them.

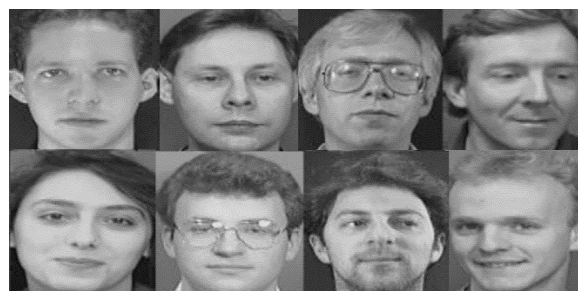


Figure 7. Faces from the ORL database



Figure 8. Faces from the FERET database

In our experiments, we gathered a subset that contains 1400 images of 200 individuals. Each individual has seven face images of different poses whose names are marked with two characters: “ba”, “bd”, “be”, “bf”, “bg”, “bj” and “bk” to indicate different face poses. Figure 8 depicts images of a person in this subset. There are variations of illumination, facial expression, and pose in this subset. The images of this subset are divided into five sets (Sub1, Sub2, Sub3, Sub4, and Sub5).

5.2 Experimental results :

¹“sklearn.svm.LinearSVC — scikit-learn 0.24.2 documentation”, <https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html>, Accessed on 06/09/2021

²“Google Colaboratory “, <https://colab.research.google.com>,

³“AT&T laboratories of cambridge university ,the database of faces,” <http://cam-orl.co.uk/facedatabase.html>, Accessed on 06/09/2021

To prove the effectiveness of our methods, we applied them to the ORL database and FERET database. We use five experiments in each case, where we take a subset for testing, whereas the four rest subsets are used for training.

a) EXPERIMENT ON ORL DATASET

The ORL database is selected to test the recognition accuracy of the proposed methods because it has differences in facial expressions (smiling/not smiling, eyes open/eyes closed) and face specifics such as wearing glasses or no glasses on.

i. Face recognition based on Handcraft method:

In the ORL database, the face recognition rate of the SIFT is better than that of the LBP and ALTP, but it is approximately the same as the LTP. Still, its performance is worse than the handcraft method proposed in this paper, which obtained an accuracy rate of 99.75%. Table 1 shows that.

Table 1. Recognition accuracy ratio of handcraft methods on the ORL database.

	<i>LBP</i>	<i>LTP</i>	<i>ALTP</i>	<i>SIFT</i>	<i>Handcraft Proposed method</i>
Subset 1	87.50%	96.25%	91.25%	98.75%	100%
Subset 2	95.00%	96.25%	92.50%	96.25%	98.75%
Subset 3	92.50%	96.25%	98.75%	96.25%	100%
Subset 4	91.25%	97.50%	92.50%	96.25%	100%
Subset 5	90.00%	97.50%	97.50%	96.25%	100%
Average	91.25%	96.75%	94.50%	96.75%	99.75%

ii. Face recognition method based on DEEP CNN:

Firstly, we have applied several deep CNN models for face identification, which we trained without data augmentation; secondly, we analysed the accuracy of the CNN models with data augmentation techniques. Finally, we used ImageNet pre-trained models.

• Face recognition using deep CNN models without Data Augmentation

We trained (AlexNet-v2, VGG16, ResNet50, Xception, Inception-v3) with the ORL database, and then we extracted the features before entering into the first fully connected layers. We used the linearSVC algorithm for classification without/with PCA to reduce the dimensionality. The ResNet50 and Inception-v3 achieved an acceptable face recognition accuracy rate compared with the others models, in which ResNet50 achieved 92.00% and Inception-v3 87.75%. Still, its accuracy is lower than our proposed handcraft method. Table 2 shows that.

The deep CNN models failed to achieve a high accuracy rate because they require a large amount of data for training, but the ORL is a very small database.

• Face recognition using deep CNN with Data Augmentation

To improve the accuracy rate of our deep learning method, we used the data augmentation technique.

We augmented the training data using Flip, Zoom, and Random Rotation, and then we used the same steps presented in the precedent experiment. This experiment's face recognition accuracy rate is better than the precedent experiment, in which ResNet50 achieved 97.50% and Inception-v3 98.50%. However, it is significantly lower than our handcraft method. Table 3 shows that.

Although we used data augmentation, the face recognition accuracy rates are still less than our proposed handcraft method, which can be explained by the close-up face images used in this experiment.

The face images in the ORL database are so close-up that the flipping and the rotation do not add significant variations in the images. On the other hand, the zoom applied in these images deletes important regions from the face images.

• Face recognition using PRE-TRAINED DEEP CNN

In this experiment, we used the pre-trained models (AlexNet-v2 or VGG16) for face recognition, and then we compared the accuracy rate of both models with other pre-trained models (ResNet50, Xception, Inception-v3) and with the handcraft method proposed in this paper.

We obtained the features before entering the first fully connected layer. Then we used two experiments to classify the data. In the first experiment, we used the linearSVC classifier, but in the second experiment, we applied PCA to reduce the dimensionality, and then we used the linearSVC classifier.

The face recognition accuracy in this experiment is very high than the last two previous experiments (we obtained an accuracy of 100 %), and it is better than our proposed handcraft method. Table 4 shows that.

b) EXPERIMENT ON FERET DATASET

A subset of the FERET database is selected to test and evaluate the proposed method's recognition accuracy because there are variations of pose, facial expression, and illumination.

i. Face recognition based on Handcraft method:

Our handcraft method obtained an accuracy rate of 94.63% in the FERET database, which surpasses those obtained by LBP, LTP, ALTP, and SIFT, Table 5 shows that.

ii. Face recognition method based on DEEP CNN:

We repeat the same experiments with the FERET database. We have trained several deep CNN models without data augmentation; then, we analysed the accuracy of the CNN models with data augmentation techniques. Finally, we used the ImageNet pre-trained model.

Table 2. Recognition accuracy ratio of SIFT + ALTP method and different deep learning models on the ORL database.

	SIFT + ALTP	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3	
		Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	100%	65.00%	11.25%	33.75%	33.75%	82.50%	81.25%	98.75%	100.00%	96.25%	93.75%
Sub 2	98.75%	76.25%	16.25%	23.75%	25.00%	87.50%	88.75%	6.25%	6.25%	88.75%	93.75%
Sub 3	100%	71.25%	16.25%	45.00%	47.50%	100.00%	98.75%	8.75%	8.75%	78.75%	83.75%
Sub 4	100%	86.25%	11.25%	23.75%	28.75%	91.25%	93.75%	61.25%	61.25%	92.50%	85.00%
Sub 5	100%	57.50%	15.00%	5.00%	3.75%	97.50%	97.50%	38.75%	38.75%	82.50%	81.25%
Average ± SD	99.75% ±0.56%	71.25% ±10.93%	14.00% ± 2.56%	26.25% ±14.76%	27.75% ±15.90%	91.75% ±7.16%	92.00% ±7.16%	42.75% ±38.67%	43.00% ±39.13%	87.75% ±7.15%	87.50% ±5.86%

Table 3. Recognition accuracy ratio of SIFT +ALTP method and different deep learning models with data augmentation on the ORL database, and the difference of accuracies with various deep learning models trained without data augmentation.

	SIFT + ALTP	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3	
		Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	100%	82.50%	18.75%	76.25%	76.25%	98.75%	98.75%	96.25%	63.75%	98.75%	97.50%
Sub 2	98.75%	78.75%	18.75%	76.25%	77.50%	77.50%	100.00%	10.00%	10.00%	96.25%	95.00%
Sub 3	100%	72.50%	16.25%	90.00%	93.75%	100.00%	97.50%	22.50%	22.50%	100.00%	98.75%
Sub 4	100%	77.50%	15.00%	41.25%	42.50%	93.75%	92.50%	100.00%	100.00%	98.75%	96.25%
Sub 5	100%	75.00%	13.75%	5.00%	5.00%	98.75%	98.75%	96.25%	95.00%	98.75%	100.00%
Average ±SD	99.75% ±0.56%	77.25% ±3.79%	16.50% ±2.24%	57.75% ±34.57%	59.00% ±35.48%	93.75% ±9.40%	97.50% ±2.93%	65.00% ±44.75%	58.25% ±41.02%	98.50% ±1.37%	97.50% ±1.98%
Difference		6.00%	2.50%	31.50%	31.25%	2.00%	5.50%	22.25%	15.25%	10.75%	10.00%

Table 4. recognition accuracy ratio of SIFT + ALTP method and different deep learning models with transfer learning on the ORL database, and the difference of accuracies with various deep learning models trained with data augmentation.

	SIFT + ALTP	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3	
		Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	100%	100.00%	100.00%	100.00%	100.00%	93.75%	98.75%	98.75%	100.00%	100.00%	100.00%
Sub 2	98.75%	100.00%	100.00%	97.50%	100.00%	96.25%	97.50%	100.00%	98.75%	97.50%	98.75%
Sub 3	100%	100.00%	100.00%	100.00%	100.00%	95.00%	100.00%	98.75%	98.75%	98.75%	98.75%
Sub 4	100%	100.00%	100.00%	100.00%	100.00%	96.25%	86.25%	100.00%	100.00%	98.75%	98.75%
Sub 5	100%	100.00%	100.00%	100.00%	100.00%	95.00%	97.50%	98.75%	100.00%	98.75%	98.75%
Average ±SD	99.75% ±0.56%	100.00% ±0.00%	100.00% ±0.00%	99.50% ±1.12%	100.00% ±0.00%	95.25% ±1.50%	96.00% ±5.55%	99.25% ±0.68%	99.50% ±0.68%	98.75% ±0.88%	99.00% ±0.56%
Difference		22.75%	83.50%	41.75%	41.00%	1.50%	-1.50%	34.25%	41.25%	0.25%	1.50%

Table 5. Recognition accuracy ratio of different methods on the FERET database.

	<i>LBP</i>	<i>LTP</i>	<i>ALTP</i>	<i>SIFT</i>	<i>Proposed method</i>
Subset 1	86.00%	99.50%	95.25%	88.75%	99.75%
Subset 2	79.50%	98.75%	87.25%	87.75%	98.00%
Subset 3	42.50%	52.25%	65.75%	67.00%	87.50%
Subset 4	45.50%	51.75%	69.75%	70.50%	90.00%
Subset 5	81.75%	98.75%	90.25%	88.75%	98.25%
Average	67.05%	80.20%	81.65%	80.55%	94.70%

• **Face recognition using deep CNN models without Data Augmentation:**

We trained the same previous models with the FERET database without data augmentation, and then we classified data with the same classifier.

The face recognition rates of the Xception and ResNet50 surpass those obtained by the other deep CNN models; when the former achieved 95.50% and the latter achieved 81.90%. Nevertheless, their performance is lower than our proposed handcraft method Table 6 shows that.

We know that deep learning models are very data-hungry, but the FERET dataset used in this experiment is small. We used the data augmentation technique to solve this problem to obtain a high face recognition accuracy rate.

• **Face recognition using deep CNN with Data Augmentation:**

We used the data augmentation technique to improve the face recognition accuracy rate of the deep CNN models used in this experiment.

We trained the models with the FERET database with data augmentation, and then we classified data with the same classifier. The accuracy rates of the Xception, ResNet50, and Inception-v3 with data augmentation surpass those obtained without data augmentation (95.70%, 87.70%, and 81.50%, respectively). Still, their performance is significantly worse than the handcraft method proposed in this paper, Table 7 shows that.

Although the data augmentation, the improvement of the accuracy rate is quite small because the training set contains many persons (200 persons) with only five samples per person, with significant differences in expression, illumination, and pose.

Table 6. Recognition accuracy ratio of SIFT + ALTP method and different deep learning models on the FERET database.

	SIFT + ALTP	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3	
		Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	99.75%	99.75%	99.50%	4.00%	4.00%	99.75%	99.50%	100.00%	100.00%	98.75%	97.75%
Sub 2	98.00%	0.50%	0.50%	1.00%	1.00%	99.25%	98.25%	99.50%	99.50%	82.50%	79.50%
Sub 3	87.50%	49.25%	49.00%	0.50%	0.50%	61.50%	59.75%	89.25%	92.75%	55.25%	55.25%
Sub 4	90.00%	44.25%	44.50%	1.00%	1.00%	53.50%	48.75%	86.75%	87.50%	49.75%	46.50%
Sub 5	98.25%	96.75%	95.75%	1.00%	1.00%	95.50%	94.00%	97.00%	97.75%	89.50%	89.25%
Average	94.70%	58.10%	57.85%	1.50%	1.50%	81.90%	80.05%	94.50%	95.50%	75.15%	73.65%
SD	±5.54%	±41.28%	±40.98%	±1.41%	±1.41%	±22.51%	±23.96%	±6.11%	±5.31%	±21.55%	±21.99%

Table 7. Recognition accuracy ratio of SIFT +ALTP method and different deep learning models with data augmentation on the FERET database, and the difference of accuracies with various deep learning models trained without data augmentation.

	SIFT + ALTP	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3	
		Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	99.75%	99.25%	99.75%	55.25%	55.50%	100.00%	99.75%	100.00%	100.00%	99.75%	99.50%
Sub 2	98.00%	0.50%	0.50%	0.50%	0.50%	99.50%	98.00%	99.50%	99.25%	97.50%	96.00%
Sub 3	87.50%	48.75%	47.75%	1.00%	1.00%	70.50%	56.00%	88.75%	92.75%	63.50%	58.50%
Sub 4	90.00%	48.75%	48.25%	1.00%	1.00%	71.00%	60.75%	85.50%	87.75%	50.50%	50.25%
Sub 5	98.25%	96.50%	96.00%	1.00%	1.00%	97.50%	95.00%	99.25%	98.75%	96.25%	97.75%
Average	94.70%	58.75%	58.45%	11.75%	11.80%	87.70%	81.90%	94.60%	95.70%	81.50%	80.40%
SD	±5.54%	±40.80%	±40.90%	±24.32%	±24.43%	±15.50%	±21.61%	±6.93%	±5.30%	±22.87%	±23.97%
Difference		0.65%	0.60%	10.25%	10.30%	5.80%	1.85%	0.10%	0.20%	6.35%	6.75%

Table 8. Recognition accuracy ratio of SIFT +ALTP method and different deep learning models with data augmentation on the FERET database, and the difference of accuracies with various deep learning models with data augmentation

	ALEXNET-v2		VGG16		ResNet50		Xception		Inception-v3		
	SIFT + ALTP	Linear SVC	Linea SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA	Linear SVC	Linear SVC+PCA
Sub 1	99.75%	100.00%	100.00%	100.00%	100.00%	83.75%	99.75%	100.00%	100.00%	100.00%	100.00%
Sub 2	98.00%	100.00%	100.00%	100.00%	100.00%	62.75%	96.75%	99.75%	100.00%	99.25%	99.50%
Sub 3	87.50%	100.00%	99.50%	99.50%	99.50%	62.75%	78.75%	98.50%	97.75%	95.50%	96.25%
Sub 4	90.00%	100.00%	99.50%	99.50%	99.75%	66.25%	79.25%	97.75%	98.25%	97.75%	97.50%
Sub 5	98.25%	100.00%	100.00%	99.75%	99.75%	88.00%	97.00%	99.75%	99.50%	100.00%	100.00%
Average	94.70%	100.00%	99.80%	99.75%	99.80%	72.70%	90.30%	99.15%	99.10%	98.50%	98.65%
SD	±5.54%	± 0.00%	± 0.24%	± 0.22%	± 0.19%	±10.92%	±09.29%	±0.87%	±0.93%	±1.71%	±1.519%
Difference		41.25%	41.35%	88.00%	88.00%	-15.00%	8.40%	4.55%	3.40%	17.00%	18.25%

• Face recognition using PRE-TRAINED DEEP CNN:

In this experiment, we used the ImageNet pre-trained models, and then we compared their accuracy rate with the accuracy rate of the models trained on the augmented FERET database. Also, we compared the performance of the pre-trained models with the handcraft method proposed in this paper.

The face recognition accuracy rate of the pre-trained AlexNet-v2, VGG16, Xception, and inception-v3 models are very higher than the models trained with the FERET database. They have also achieved an accuracy better than our proposed handcraft method. In which AlexNet-v2 achieved an accuracy of 100%. Table 8 illustrates that.

c) Comparison to State-of-the-art methods :

In this part, we compared our proposed methods with several state-of-the-art techniques used in the face recognition field. As it is clear from Table 9, the average face recognition rates of our methods in the ORL database surpassed the average face recognition rates of recent works. AlexNet-v2 and VGG16 with transfer learning and LinearSVC classifier can reach a face recognition accuracy of 100 % in the ORL database.

In addition, our proposed methods obtain higher face recognition accuracy rates in the FERET database than some state-of-the-art techniques. Table 9 shows that.

Our handcraft proposed method achieves an accuracy rate of 94.70% in the FERET database, while both methods AlexNet-v2 and VGG16 with transfer learning and LinearSVC achieve 100% top accuracy.

VGG16 with transfer learning, PCA, and LinearSVC classifier with one image in the training set and six images in the testing set achieved an extremely high accuracy rate (accuracy of 99.33%).

6. Conclusion and Future Work

This research proposes two novel face recognition methods: The first one is based on handcraft features extraction, and the second is based on deep learning features extraction.

The handcraft-based proposed method does not calculate the ALTP for the entire image but only for the important parts of the image, which are extracted based on SIFT Key points. The deep learning-based method used the ImageNet pre-trained AlexNet-v2 or VGG16 models to extract features and LinearSVC for the classification.

The experimental results demonstrate that our methods have a strong resistance to the difficult lighting conditions, variations of pose, and facial expression; our methods are shown to outperform the state-of-the-art face recognition methods in recognition accuracy rate.

In the future work, we intend to investigate the adaptation of the proposed methods for iris recognition. Then we plan to extend our methods for combining two modalities, the face and both irises.

Table 9. The average accuracy of state-of-the-art face recognition methods compared with our proposed methods

Database	years	Method	Recognition rate	Evaluation protocol training: testing
ORL	2016	DCT/VQ[39]	98.25%	Cross-validation 8:2
	2017	CNN[48]	98.30%	8:2
	2017	Deep CNN [50]	96.6%	4:8
	2018	DWT-2 levels –Bior5.5 filter-PCA[37]	98.25%	8:2
	2019	DC[43]	98.61%	8:2
	2019	Transfer Learning + KCFT [52]	97.77 %	1:9
	2020	IKLDA+PNN[36]	97.22%	8:2
	2020	Mixed Transform and Multilayer Sigmoid NN Classifier [47]	98.8%	8:2
	Proposed	ALTP+SIFT	99.75%	Cross-validation 8:2
	Proposed	pre-trained VGG16 + linearSVC	99.50%	Cross-validation 8:2
pre-trained VGG16 +PCA+ linearSVC		100%	Cross-validation 8:2	
pre-trained AlexNet-v2 + linearSVC		100%	Cross-validation 8:2	
pre-trained AlexNet-v2 +PCA+ linearSVC		100%	Cross-validation 8:2	
FERET	2015	dLBP α [34]	94.50%	5:2
	2017	Deep CNN [50]	95.6%	2/3:1/3
	2018	TDL [51]	93.90%	1:6
	2019	DC[43]	93.46%	5:2
	2019	Transfer Learning + KCFT [52]	93.04%	1:6
	2020	AR based on CR[46]	76.50%	5:2
	2020	AR based on LILS[46]	78.00%	5:2
	Proposed	ALTP+SIFT	94.70%	Cross-validation 5:2
	Proposed	pre-trained VGG16 + linearSVC	99.75%	Cross-validation 5:2
		pre-trained VGG16 +PCA+ linearSVC	99.80%	Cross-validation 5:2
		pre-trained AlexNet-v2 + linearSVC	100.00%	Cross-validation 5:2
		pre-trained AlexNet-v2 +PCA+ linearSVC	99.80%	Cross-validation 5:2
		pre-trained VGG16 + linearSVC	97.83%	1:6
pre-trained VGG16 +PCA+ linearSVC		99.33%	1:6	
pre-trained AlexNet-v2 + linearSVC		98.67%	1:6	
pre-trained AlexNet-v2 +PCA+ linearSVC		98.50%	1:6	

References

- [1] S. Akbar, A. Ahmad, M. Hayat, and F. Ali, "Face recognition using hybrid feature space in conjunction with support vector machine," *J. Appl. Environ. Biol. Sci.*, vol. 5, no. 7, pp. 28–36, 2015.
- [2] G. L. Marcialis and F. Roli, "Fusion of face recognition algorithms for video-based surveillance systems," in *Multisensor Surveillance Systems*, Springer, 2003, pp. 235–249.
- [3] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 10, pp. 1042–1052, 1993.
- [4] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, 1997.
- [6] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
- [7] T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.
- [8] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. image Process.*, vol. 19, no. 6, pp. 1635–1650, 2010.
- [9] R. Gupta, H. Patil, and A. Mittal, "Robust order-based methods for feature description," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 334–341.
- [10] W. Yang, Z. Wang, and B. Zhang, "Face recognition using adaptive local ternary patterns method," *Neurocomputing*, vol. 213, pp. 183–190, 2016.

- [11] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [12] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*, 2006, pp. 404–417.
- [13] M. El Aroussi, M. El Hassouni, S. Ghouzali, M. Rziza, and D. Aboutajdine, "Local appearance based face recognition method using block based steerable pyramid transform," *Signal Processing*, vol. 91, no. 1, pp. 38–50, 2011.
- [14] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automated detection of mild and multi-class diabetic eye diseases using deep learning," *Heal. Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–9, 2020.
- [15] J. He, J. Rong, L. Sun, H. Wang, Y. Zhang, and J. Ma, "A framework for cardiac arrhythmia detection from IoT-based ECGs," *World Wide Web*, vol. 23, no. 5, pp. 2835–2850, 2020.
- [16] R. A. Priyadarshini, S. Arivazhagan, and M. Arun, "A deep learning approach for person identification using ear biometrics," *Appl. Intell.*, vol. 51, no. 4, pp. 2161–2172, 2021.
- [17] J. Jayanthi, E. L. Lydia, N. Krishnaraj, T. Jayasankar, R. L. Babu, and R. A. Suji, "An effective deep learning features based integrated framework for iris detection and recognition," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, pp. 3271–3281, 2021.
- [18] H. Jiang, R. Zhou, L. Zhang, H. Wang, and Y. Zhang, "Sentence level topic models for associated topics extraction," *World Wide Web*, vol. 22, no. 6, pp. 2545–2560, 2019.
- [19] R. Singh, Y. Zhang, H. Wang, Y. Miao, and K. Ahmed, "Deep learning for Antisocial Behaviour Analysis on Social Media," in *2020 24th International Conference Information Visualisation (IV)*, 2020, pp. 428–434.
- [20] S. Almabdy and L. Elrefaie, "Deep convolutional neural network-based approaches for face recognition," *Appl. Sci.*, vol. 9, no. 20, p. 4397, 2019.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.
- [22] A. Krizhevsky, "One weird trick for parallelizing convolutional neural networks," *arXiv Prepr. arXiv1404.5997*, 2014.
- [23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Prepr. arXiv1409.1556*, 2014.
- [24] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [26] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [27] O. Lucena, A. Junior, V. Moia, R. Souza, E. Valle, and R. Lotufo, "Transfer learning using convolutional neural networks for face anti-spoofing," in *International conference image analysis and recognition*, 2017, pp. 27–34.
- [28] S. Masood, T. Luthra, H. Sundriyal, and M. Ahmed, "Identification of diabetic retinopathy in eye images using transfer learning," in *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 2017, pp. 1183–1187.
- [29] A. K. Jain, *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.
- [30] D. Harwood, T. Ojala, M. Pietikäinen, S. Kelman, and L. Davis, "Texture classification by center-symmetric autocorrelation, using Kullback discrimination of distributions," *Pattern Recognit. Lett.*, vol. 16, no. 1, pp. 1–10, 1995.
- [31] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *European conference on computer vision*, 2004, pp. 469–481.
- [32] X. Feng, M. Pietikainen, and A. Hadid, "Facial expression recognition with local binary patterns and linear programming," *Pattern Recognit. Image Anal. C/C Raspoznaniye Obraz. I Anal. Izobr.*, vol. 15, no. 2, p. 546, 2005.
- [33] R. Shyam and Y. N. Singh, "Face recognition using augmented local binary pattern and Bray Curtis dissimilarity metric," in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 779–784.
- [34] Y. Kaya, Ö. F. Ertuğrul, and R. Tekin, "Two novel local binary pattern descriptors for texture analysis," *Appl. Soft Comput.*, vol. 34, pp. 728–735, 2015.
- [35] D. Arisandi, M. F. Syahputra, I. L. Putri, S. Purnamawati, R. F. Rahmat, and P. P. Sari, "A real time mobile-based face recognition with fisherface methods," in *Journal of Physics: Conference Series*, 2018, vol. 978, no. 1, p. 12038.
- [36] A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang, "A hybrid improved kernel LDA and PNN algorithm for efficient face recognition," *Neurocomputing*, vol. 393, pp. 214–222, 2020.
- [37] S. F. Kak, F. M. Mustafa, and P. R. Valente, "Discrete wavelet transform with eigenface to enhance face recognition rate," *Acad. J. Nawroz Univ.*, vol. 7, no. 4, pp. 9–17, 2018.
- [38] H. A. Ameen and A. N. Hashim, "Face Recognition using Hybrid Methods of Transformation and Statistical Measures," *MS&E*, vol. 765, no. 1, p. 12068, 2020.
- [39] A. Aldhahab, T. Al Obaidi, and W. B. Mikhael, "Employing vector quantization on detected facial parts for face recognition," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2016, pp. 1233–1237.
- [40] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the seventh IEEE international conference on computer vision*, 1999, vol. 2, pp. 1150–1157.
- [41] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the use of SIFT features for face authentication," in *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, 2006, p. 35.
- [42] C. Geng and X. Jiang, "SIFT features for face recognition," in *2009 2nd IEEE international conference on computer science and information technology*, 2009, pp. 598–602.
- [43] F. Ouslimani, A. Ouslimani, and Z. Ameer, "Rotation-invariant features based on directional coding for texture classification," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6393–6400, 2019.
- [44] M. Tabejamaat, A. Mousavi, and M. L. Gavrilova, "Local comparative decimal pattern for face recognition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 34, no. 12, p. 2056006, 2020.
- [45] D. Kumar, J. Garain, D. R. Kisku, J. K. Sing, and P. Gupta, "Unconstrained and constrained face recognition using dense local descriptor with ensemble framework," *Neurocomputing*, vol. 408, pp. 273–284, 2020.
- [46] Y. Qin, L. Sun, and Y. Xu, "Exploring of alternative

- representations of facial images for face recognition,” *Int. J. Mach. Learn. Cybern.*, pp. 1–7, 2020.
- [47] G. M. Sapijaszko and W. B. Mikhael, “Facial Recognition System Using Mixed Transform and Multilayer Sigmoid Neural Network Classifier,” *Circuits, Syst. Signal Process.*, vol. 39, pp. 6142–6161, 2020.
- [48] P. Kamencay, M. Benčo, T. Miždoš, and R. Radil, “A new method for face recognition using convolutional neural network,” 2017.
- [49] R. I. Bendjillali, M. Beladgham, K. Merit, and A. Taleb-Ahmed, “Illumination-robust face recognition based on deep convolutional neural networks architectures,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 2, pp. 1015–1027, 2020.
- [50] M. Wang, Z. Wang, and J. Li, “Deep convolutional neural network applies to face recognition in small and medium databases,” in *2017 4th International Conference on Systems and Informatics (ICSAI)*, 2017, pp. 1368–1372.
- [51] J. Zeng, X. Zhao, J. Gan, C. Mai, Y. Zhai, and F. Wang, “Deep convolutional neural network used in single sample per person face recognition,” *Comput. Intell. Neurosci.*, vol. 2018, 2018.
- [52] R. Min, S. Xu, and Z. Cui, “Single-sample face recognition based on feature expansion,” *IEEE Access*, vol. 7, pp. 45219–45229, 2019.
- [53] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” *arXiv Prepr. arXiv1712.04621*, 2017.
- [54] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248–255.
- [55] K. Yang, J. Yau, L. Fei-Fei, J. Deng, and O. Russakovsky, “A Study of Face Obfuscation in ImageNet,” *arXiv Prepr. arXiv2103.06191*, 2021.
- [56] A. Paszke *et al.*, “Automatic differentiation in pytorch,” 2017.
- [57] F. Chollet and others, “keras.” 2015.
- [58] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, “The FERET database and evaluation procedure for face-recognition algorithms,” *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, 1998.