HAL open science

Illumination-robust face recognition based on deep convolutional neural networks architectures

Ridha Ilyas Bendjillali, Mohammed Beladgham, Khaled Merit, Abdelmalik Taleb-Ahmed

▶ **To cite this version:**

Ridha Ilyas Bendjillali, Mohammed Beladgham, Khaled Merit, Abdelmalik Taleb-Ahmed. Illumination-robust face recognition based on deep convolutional neural networks architectures. Indonesian Journal of Electrical Engineering and Computer Science, 2020, 18 (2), pp.1015-1027. 10.11591/ijeecs.v18.i2.pp1015-1027 . hal-03564174

HAL Id: hal-03564174

https://uphf.hal.science/hal-03564174v1

Submitted on 10 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Indonesian Journal of Electrical Engineering and Computer Science
Vol. 18, No. 2, May 2020, pp. 1015~1027
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v18.i2.pp1015-1027

Illumination-robust face recognition based on deep convolutional neural networks architectures

Ridha Ilyas Bendjillali[1], Mohammed Beladgham[2], Khaled Merit[3], Abdelmalik Taleb-Ahmed[4]
[1,2,3]Laboratory of TIT, Department of Electrical Engineering, Tahri Mohammed University, Algeria
[4]Laboratory of IEMN DOAE. UMR CNRS 8520, University of Valenciennes, France

Article Info

Article history:

Received Jul 10, 2019
Revised Sep 12, 2019
Accepted Sep 26, 2019

Keywords:

Face recognition
Inception-v3
M-CLAHE
ResNet50
VGG16

ABSTRACT

In the last decade, facial recognition techniques are considered the most important fields of research in biometric technology. In this research paper, we present a Face Recognition (FR) system divided into three steps: The Viola-Jones face detection algorithm, facial image enhancement using Modified Contrast Limited Adaptive Histogram Equalization algorithm (M-CLAHE), and feature learning for classification. For learning the features followed by classification we used VGG16, ResNet50 and Inception-v3 Convolutional Neural Networks (CNN) architectures for the proposed system. Our experimental work was performed on the Extended Yale B database and CMU PIE face database. Finally, the comparison with the other methods on both databases shows the robustness and effectiveness of the proposed approach. Where the Inception-v3 architecture has achieved a rate of 99, 44% and 99, 89% respectively.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Ridha Ilyas Bendjillali,
Laboratory of TIT, Department of Electrical Engineering,
Tahri Mohammed University, Algeria.
Email: rayben43@gmail.com

1. INTRODUCTION

In recent years, high security has become a major problem, not only in military environments but also in many civilian applications. Person identification by face recognition gives a number of solutions to this problem. The study of automatic face recognition has begun to be studied since the early 1970s. The work done by Kanade [1] during his thesis is considered the first in this field. Since then, a lot of research has been done. A face recognition system is an identification and verification system of individuals that can check whether a person belongs to the system database, and identify if this is the case.

The first step in the facial recognition system is face detection. Their effectiveness is directly influenced on the face recognition system. Face authentication essentially depends on the method used to locate the face in the image. Even if many signs of progress in Face Recognition, there are several challenges among them: pose variations, partial occultations, illumination variations, facial expressions.

The second important step of the Face Recognition system is the extraction of facial features. There are mainly three classical methods for face analysis: first is the local approach, second is global approach, and the third is hybrid approach, the local approach is based on the extraction of local feature points of the face (such as the nose, mouth and eyes and to model the relationship between these points, a topological graph is constructed for each face. Local Binary Pattern (LBP) [2], Local Gabor Binary Pattern (LGBP) [3], is the most popular method in the local approach. In the global approach, all pixels of the face are considered and then, dimensional reduction techniques are generally used to reduce pixel redundancy. Principal component analysis (PCA) [4], Linear Discriminant Analysis (LDA) is the most popular methods in the global approach. Hybrid approaches are approaches that combine global and local features to improve the performance of face recognition. Indeed, local characteristics and global characteristics have quite different properties.

These techniques make it possible to obtain a reduced representation of the original data while preserving the most important feature for the treatments. But these approaches are sensitive to changes in the illumination and facial expressions.

In recent years, methods that rely on deep learning particularly Convolutional Neural Networks have achieved great success in several areas, including the classification of images, object detection, facial recognition, and facial expressions. In the last few years, the facial recognition methods witnessed a great development through the evolution of Convolutional Neural Networks architectures like VGG16 [5], ResNet50 [6], and Inception-v3 [7]. These methods avoid the large scale datasets problem and achieved higher accuracy than the classical method.The main contribution of the proposed methodology is to obtain a powerful recognition algorithm with high recognition rate, in this paper we proposed a model by applying Viola-Jones algorithm for two important reasons, firstly to detect faces, secondly to separate the faces from the non-faces, the Modified Contrast Limited Adaptive Histogram Equalization (M-CLAHE) to improve the contrast of the facial image Moreover, we employed VGG-16, InceptionV3, and ResNet50 on face images to extract features and for classification.

The paper is organized as follows: In Section 2 we deal briefly with the related work in field of face recognition, in Section 3 we briefly describe the three research areas related to our work: the steps of face recognition: Viola-Jones face detection, facial image enhancement, Feature learning and classification using Deep convolution Neural Network Architectures. Experimental results are presented in Section 4; Section 5 gives the comparison of the results with other methods, and Section 6 gives the conclusion of this paper.

2. RELATED WORK

In the last decade, many kinds of research on facial recognition techniques have been developed in order to obtain a better recognition algorithm, such as (LDA) and (LBP). These methods have produced good results in ideal conditions. But under the challenging conditions like the changes in illumination and facial expressions, these methods did not provide good results. The CNN based methods have achieved a highly successful in solving several problems of visual information processing like object detection, feature extraction, and image classification.Some researchers have presented the Convolution Neural Network for face recognition as an alternative to traditional methods; because this method has made the FR system faster than the other traditional methods. In this section, we will mention recent research related to FR based on convolutional neural networks CNN's which achieved significant results.

In 2014 K. Simonyan et Al., [5] used VGG-16 architecture for large-scale image classification, where they proved that depth is useful for classification accuracy. In 2015 O. M. Parkhi et Al., [8] they also used the same architecture proposed in [5] performed on Labeled Faces in the Wild dataset (LFW) and YouTube Faces (YTF), this approach has been achieved 98.95% and 97.3% respectively. In 2016 J. C. Chen et Al., [9] proposed an approach based on the DCNN network motivated by [10] achieved 97.45% on (LFW) dataset.

In 2017 J. S. Chan et Al., [11] proposed the Facial Attribute Assistant Network (FAAN) built on the Residual Network ResNet-101 performed on IJB-A, this network has achieved 98.2%. In 2018 A. Zhanfu et Al., [12] proposed an approach based on AlexNet model and Inception-ResNet-V1 model evaluated on the LFW and SLLFW databases; the Inception-ResNet-V1 network achieved a 99.20% and 95.80% recognition rate on LFW and SLLFW respectively.

3. METHODOLOGY

The face recognition system is performed in three main steps: (1) Facial image enhancement, (2) Face detection, (3) Extraction of facial features and classification. In what follows we will explain each step in our system. The main steps are demonstrated as shown in Figure 1.

3.1. Face Detection using the Technique of Viola-Jones

The effectiveness of biometric systems based on face authentication essentially depends on the method used to locate the face in the image. In this regard, we use the Viola-Jones algorithm to efficiently detect various parts of the human faces such as mouth, eyes, nose, eyebrows, lips, ears, etc [13]. Paul Viola and Michael Jones in 2001 who came up with the most effective algorithm to detect the human faces and its parts. This algorithm has been implemented in Matlab using the method vision.CascadeObjectDetector. There are 3 important techniques used by Viola - Jones for the detection of facial parts:

1. Haar-like features are of a rectangular type which is used for the feature extraction thus getting an Integral image [14].
2. Ada boost is artificial intelligence and machine learning method for face detection. The term 'boosted' determines a principle that brings together many algorithms that rely on sets of binary classifiers [2].

3. The third and last step is Cascade classifier that can efficiently combine many features. The term 'cascade' in a classifier determines the several filters on a resultant classifier.

An example of the Viola-Jones method is demonstrated in Figure 2.

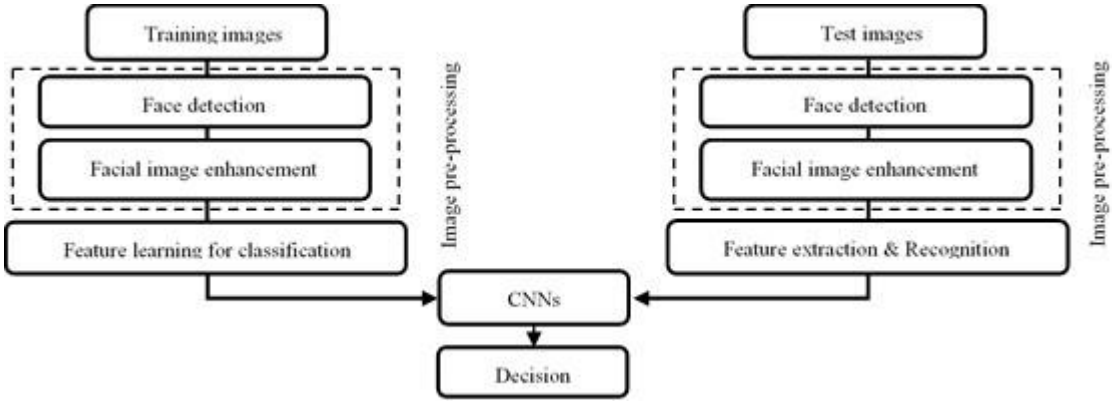


Figure 1. The main steps of the proposed system

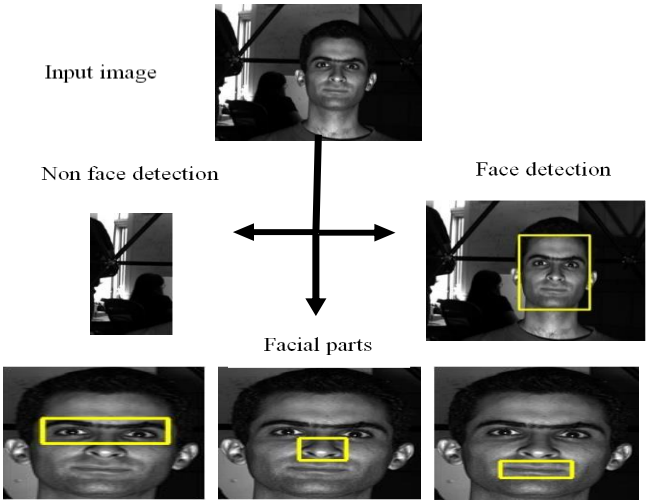


Figure 2. Detection of face and facial parts based on Viola-Jones method

3.2. Enhancement Techniques

Histogram Equalization (HE) is a traditional method for contrast enhancement in a digital image, which modifies the gray level histogram of an image to a uniform distribution [15]. The histogram 'h' of the digital image is defined as the discrete function and is given by (1).

$$h(X_k) = n_k \quad (1)$$

Where Xk is the kth intensity level in the interval [0, L−1] and nk is the number of pixels in the input image. But it has several shortcomings such as the loss of some details in the image, some local areas become brighter than before, and it also fails to conserve the brightness of the image [14]. Another approach has been proposed which is a modification of HE called the adaptive histogram equalization (AHE). This method is based on dividing the input image into small blocks called "tiles"; a local histogram will perform all regions occupying different grayscale ranges using the CDF of a given (PDF). The probability density function (PDF) is defined by (2).

$$P(X_k) = \frac{h(X_k)}{M \times N}, for\ k = 1,2, \dots \dots, L-1 \quad (2)$$

Where M × N is the image size. The cumulative distribution function (CDF) is obtained by (3).

$$C(X_k) = \sum_{j=0}^{k} P(X_j)\,, for\ k = 1,2, \ldots., L-1 \quad (3)$$

But it has several drawbacks such as its high computational cost and noise amplification in the homogeneous regions. In this section, we have presented various techniques for contrast image enhancement such as the following.

3.2.1 Contrast Limited Adaptive Histogram Equalization (CLAHE)

The contrast limited adaptive histogram equalization (CLAHE) is an improvement of adaptive histogram equalization (AHE). This method is mainly adopted on clipping the histogram at a predefined value in order to limit contrast amplification before computing the CDF value this method is called the clip limit [16]. The method is fast, easy to implement, and fully automatic [17].

3.2.2 Modified Contrast Limited Adaptive Histogram Equalization (M-CLAHE)

In order to avoid the noise problem of CLAHE after the enhancement process, a filtering operation is performed. But the low-pass filtering can create the loss of image details and thus the loss of information, therefore a discriminative low pass filtering process is applied only on noisy regions while the other regions remain unchanged [18]. The algorithm of M-CLAHE is shown in Figure 3.

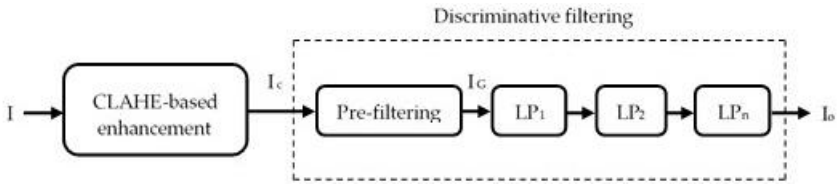


Figure 3. Modified CLAHE of the proposed system

The first stage, called the pre-filtering stage, performs a slight smoothing operation of the enhanced image using the Gaussian blur. The LP_1 can perform moderate filtering and the strength of the low-pass filtering operation can progress in the next stages [19].

Like the AHE, the original source low contrast input image 'I' is fed to the HE-based enhancement block. The noise generated through the enhancement block is removed by the discriminative filtering process.

3.2.3 Brightness Preserving Bi-Histogram Equalization (BBHE)

Brightness Preserving Bi-Histogram Equalization (BBHE) is another extension of the histogram equalization (HE)-based contrast enhancement to avoid the shortcomings of HE, the BBHE algorithm divides the input image histogram into two sub-images, based on the mean of the input image, the basic ideas of the BBHE method of decomposing the original image into two sub-images and then equalizes the histograms of the sub-images separately, thus preserving the mean brightness of the image [20]. Let I_m be the mean of the input image I, such that $I_m \in \{I_0, I_1, \ldots, I_{L-1}\}$. Based on the mean, the original image is decomposed as shown in (4).

$$I = I_L \cup I_U \quad (4)$$

Where,

$$I_L = \{I(i,j) |\, I(i,j) \leq I_m, \forall\, I(i,j) \in I\} \quad (5)$$

And

$$I_U = \{I(i,j) |\, I(i,j) > I_m, \forall\, I(i,j) \in I\} \quad (6)$$

3.2.4 Gaussian Based Image Enhancement (GBIE)

This method is a widely used effect in image processing typically to reduce image noise and reduce detail. This method has achieved great success in image enhancement. The window used in this algorithm uses is of 5x5 pixel window, where a block of 25 pixels of the original image is convolved with Gaussian kernel of size 5x5. The formula of a Gaussian function in 1D is obtained by (7).

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \tag{7}$$

The 2D is a product of two such Gaussian functions, the formula of 2D Gaussian function is defined by (8).

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{8}$$

Where x and y are the spatial coordinates and σ is the standard deviation of the Gaussian distribution. The Gaussian convolution matrix is defined by (9).

$$G(x,y) = I(x,y) \otimes \mathrm{g(x,y)} \tag{9}$$

Where ⊗ indicate the convolution. g(x, y) and I(x, y) are the Gaussian kernels and the original image respectively and G(x, y) is the convolved output.The dynamic range of gray level input values is compressed to controlled levels after applying the logarithmic transform on the image which is done by (10).

$$G_L(x,y) = K \times \log 2\,[1 + G(x,y)] \tag{10}$$

Where K is a constant and $G_L\,(x,y)$ is the logarithmic transformed image. In order to get the pixels value in the grayscale range [0 255], the logarithmic corrupted image is scaled by a scaling factor of 32. Finally, the output image is given by (11).

$$I\,'(x,y) = \frac{F_{max}}{G_{L\,max} - G_{L\,min}} [G_L\,(x,y) - G_{L\,min}] \tag{11}$$

Where F_{max} is the maximum intensity. $G_{L\,min}$ and $G_{L\,max}$ are the minima. $G_{L\,max}$ is the values of a logarithmic transformed image and $I\,'(x,y)$ is the enhanced image [21].

3.3. Classification using Deep Convolution Neural Network

The Convolutional Neural Networks (CNN) is a type of Artificial Neural Networks (ANN), which have wide uses in several areas such as image classification, decision-making, etc. It adopts the state of the art machine learning tasks. CNN's are algorithms that can identify faces, character, human pose, tumors, street signs, and so on [22]. Its design is based on some basic layers which are as follows:

3.3.1 Convolution Layer (ConvL)

The most important operation on CNN is the ConvL, the ConvL performs the core building block of a Convolutional Network that does most of the computational heavy lifting [23]. Like the traditional neural network, the input of each ConvL is the output of the upper layer [24], on the one hand, in the ConvL, each of the feature graphs corresponds to a convolution kernel of the same size, on the other hand a convolution operation is done between each of the feature maps of the ConvoL is and the feature map of the previous layer [25], after that, we add a bias and then add the corresponding element obtained by function activation.

The mathematical expression of the layer [26] is:

$$x_i^l = f\left(\sum_{i \in M_j^{l-1}} x_i^{l-1} k_{ij}^l + b_j^l\right) \tag{12}$$

Where l represents the layer. *f* represents the activation function. *k* is the convolution kernel, *b* is the bias. And *Mj* represents the feature map.

3.3.2 Maxpooling Layer

The output feature maps obtained after the calculation of the ConvL are generally not much reduced in dimension. If the dimension does not change, there will be a great amount of computation need to do, and the network learning process will become very difficult, more likely to get a reasonable result [26]. The pooling layer is another important concept of CNN's simplifies the output by performing non-linear down-sampling, reducing the number of parameters that the network needs to learn, and don't change the number of feature graphs; the pooling layer is sampled with the maximum value, the sampling size is 2x2.

3.3.3 Fully Connected Layer

For the network, after several convolutions and max-pooling layers, the high-level reasoning in the neural network is done via fully connected layers. Neurons in a fully connected layer have connections to all activations in the previous layer, and these full-connected layers form a multi-layer perceptron (MLP), which plays the role of a classifier. There are several models of the CNN that have been successful in the field of facial recognition; among these models we mention the VGG16, ResNet50, and Inception-v3. In what follows: we detailed each of these architectures.

3.4. Very Deep Convolution Network (VGG)

The VGG16 architecture is proposed by Simonyan and Zisser-man [5], the input size of this model is fixed as 224 × 224. The images are passed through a stack of convolutional (convs) layers, where a small receptive filters of size 3 × 3. Further 1 × 1 convolution filters are also used where a linear transformation of input channels (followed by non-linearity). And in order to preserve the spatial resolution after convolution, the padding of 1 pixel for 3 × 3 conv layers is employed; the spatial pooling is executed by 5 max-pooling layers. The Max-pooling is performed over a 2×2 pixel window, with stride 2, a stack of convs layers followed by three fully-connected (FC) layers have been utilized such as the first two FC layers have 4096 channels, and the third FC layer has 1000 channels (one for each class). The final layer of this architecture is the softmax layer. The VGG16 architecture that is used in this paper is shown in Figure 4.

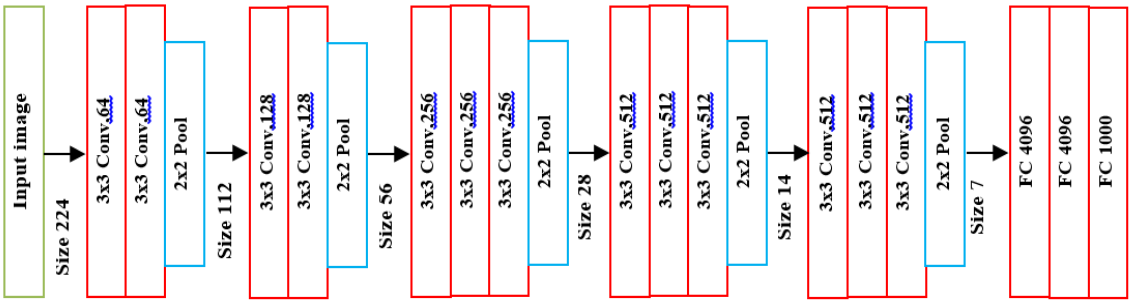


Figure 4. The proposed VGG16 architecture

3.5. Deep Residual Network Architectures (ResNet)

The ResNet 50 is a 50 layer of Residual Network proposed by He et al. [6]. Like the VGG nets the size of convolutional layers is 3×3 filters, the input size of this model is fixed as 224 × 224. and they follow some simple designs such as: For the layers having the same number of filters has the same output. The number of filters is doubled if the convolved output size is halved such that the time complexity per layer is preserved. The model ends with an average pooling layer and a 1000-way fully-connected layer with softmax. This model has fewer filters and lower complexity than VGG nets and there are other variants like ResNet101 and ResNet152. Figure 5 show the configuration layers of ResNet50 network.

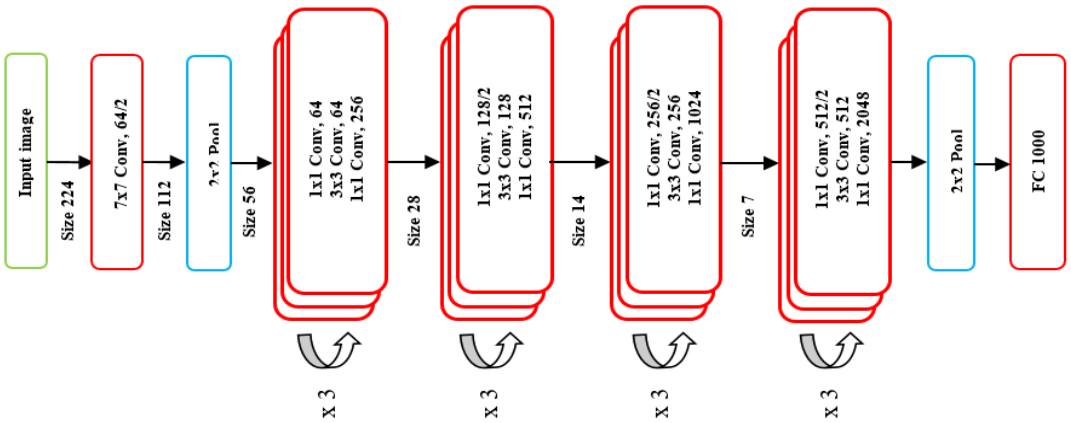


Figure 5. The proposed Resnet50 CNN architecture

3.6. Inception-v3 (GoogLeNet)

The Inception-v3 is a type of convolutional neural network was introduced as GoogLeNet by Szegedy et al. [7, 27]. This network contains 48 layers deep and can classify images into 1000 classes the input size of this model is fixed as 299 × 299. This model based on a multi-scale approach wich multiple classifier structures while combined with multiple sources for backpropagation. The Inception-v3 model increases both the width and depth of the network without causing penalties. In this model, multiple Inception layers are applied in convolution on the input feature map in different scales for allowing making more complex decisions. Figure 6 shows the architecture of Inception-v3.

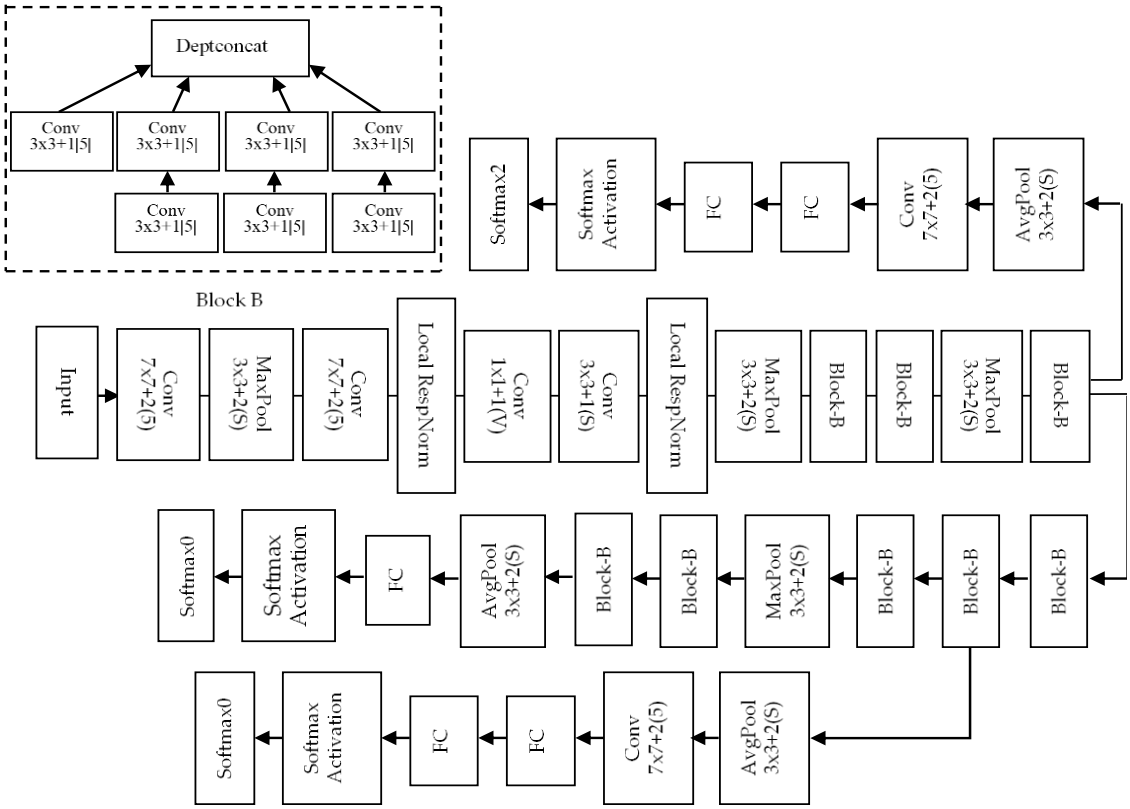


Figure 6. The proposed Inception-v3 CNN architecture

4. RESULTS AND DISCUSSION

In this paper, the tests were performed on a personal computer PC 64 bits system with I7 3.0 GHz processor and 16 GB of RAM using MATLAB R2019a.

4.1. Performance Comparison of Enhancement Techniques

In our experiment different type of image enhancement algorithms have been used, the comparison and the judgment of each of the others are given by the calculation of the parameters Absolute Mean Brightness Error (AMBE), and Peak Signal to Noise Ratio (PSNR) [28].

4.1.1 PSNR

It is often easy to define the (PSNR) via the Mean squared error (MSE), the MSE is defined as follows:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(I(i,j) - \hat{I}(i,j) \right)^2 \tag{13}$$

Which requires two M x N grayscale images I and Î. The PSNR is defined as:

$$PSNR = 10 \log_{10} \left(\frac{(255)^2}{MSE} \right) \tag{14}$$

The greatest value of PSNR indicates the better contrast enhancement algorithm.

4.1.2 AMBE (Absolute Mean Brightness Error)

The absolute mean brightness error (AMBE) is a parameter proposed to rate the performance in preserving image brightness. The (AMBE) is defined by the following:

$$AMBE\ (x,y)=|\ X_m - Y_m\ | \tag{15}$$

Where Xm mean intensity of the input image. Ym mean intensities of the output image [29].

On the contrary to PSNR, the least value of AMBE indicates better brightness preservation; the Table 1 and Table 2 shows the AMBE and PSNR results.

Table 1. The Absolute mean Brightness Error (AMBE) Values for Different Algorithms. CLAHE; M-CLAHE; GBIE; BBHE

Algorithms	Average AMBE	
	Extended Yale B Database	CMU PIE Database
CLAHE	4.19	7.28
M-CLAHE	4.07	7.07
GBIE	34.42	45.92
BBHE	18.79	23.68

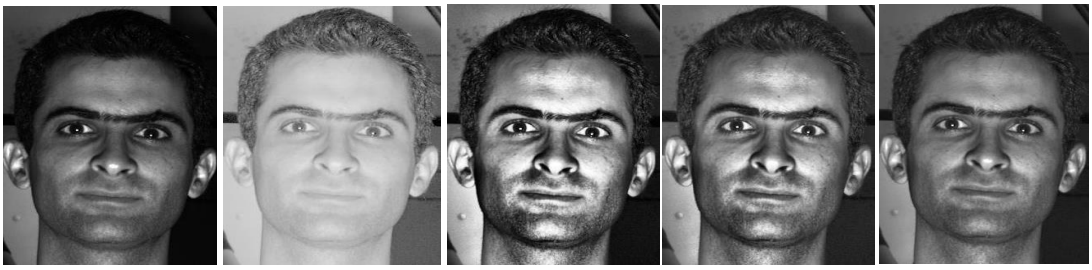
Table 2. The Peak Signal to Noise Ratio (PSNR) values for Different Algorithms. CLAHE; M-CLAHE; GBIE; BBHE

Algorithms	Average PSNR	
	Extended Yale B Database	CMU PIE Database
CLAHE	41.92	31.77
M-CLAHE	45.26	34.32
BBHE	27.11	20.56
GBIE	24.30	18.43

4.2. The Visual Comparison

In this section, the facial image after enhancement is shown (Figure 7 and 8); the main goal of the visual comparison is to judge if the enhanced facial image has a more natural appearance and is visually acceptable to the human eye.

Based on visual results comparison shown in (Figure 7) and (Figure 8) of both databases. We conclude that the M-CLAHE technique provides better visual quality and a more natural appearance compared with other techniques.



(a) **(b)** **(c)** **(d)** **(e)**

Figure 7. The results obtained from different algorithms on the Extended Yale B database: (a) original image, (b) GBIE; (c) BBHE, (d) CLAHE, and (e) M-CLAHE

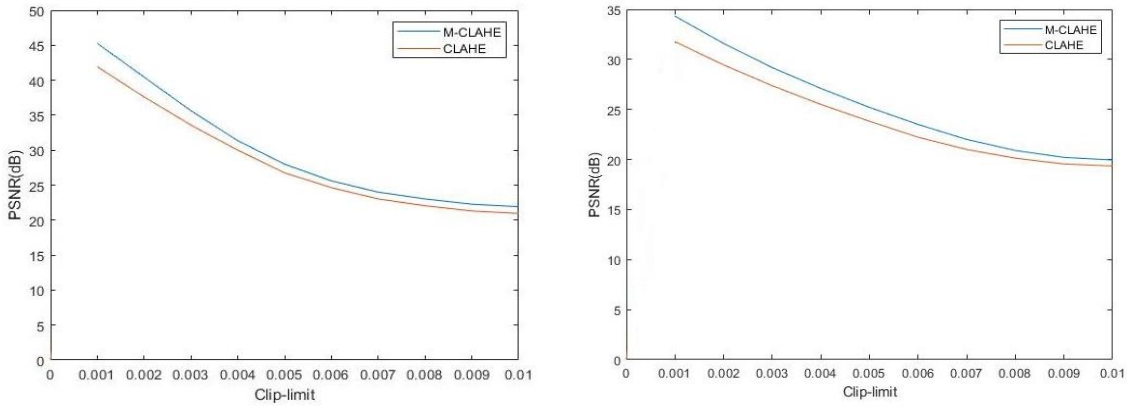
(a) **(b)** **(c)** **(d)** **(e)**

Figure 8. The results obtained from different algorithms on the CMU PIE database: (a) original image, (b) GBIE; (c) BBHE, (d) CLAHE, and (e) M-CLAHE

4.3. The Choice of Clip-Limit (CL) Value

After the visual comparison, we have focused our attention on two other parameters; the impact of the clip-limit (CL) value and block size (Bs) of the CLAHE and M-CLAHE algorithms. As a first step, we fixed the value of BS to [8 8] and varied the CL from 0.001 to 0.010 and for each variation. The PSNR results of CLAHE and M-CLAHE for the Extended Yale B and CMU PIE databases are shown in Figure 9.

From Figure 9 above it can be concluded that the CLAHE and M-CLAHE algorithms achieved the highest PSNR value at CL= 0.001 on the Extended Yale B database and CMU PIE respectively.



(a) (b)

Figure 9. Peak signal to noise ratio (PSNR) values of each variation of CLAHE and M-CLAHE for the (a) Extended Yale B database, (b) CMU PIE database

4.4. The Choice of Block Size

The choice of block size also affects the performance of the contrast enhancement algorithm, and in order to find the appropriate block size, we fixed the clip limit value at 0.01 and varied the block size from [2 2] to [128 128], and for each size variation, we calculate the PSNR values (see Table 3).

From the results shown in the table, it is clear that the PSNR has the highest value when the block size of [8 8] for the Extended Yale B and the CMU PIE databases was used.

Table 3. PSNR Values of Different Block Sizes

Average PSNR	Extended Yale B		CMU PIE	
Block Size	M-CLAHE	CLAHE	M-CLAHE	CLAHE
2 × 2	18.71	16.57	17.80	17.11
4 × 4	20.19	18.91	18.99	18.05
8 × 8	21.92	19.95	20.95	19.35
16 × 16	18.53	17.24	17.93	17.22
32 × 32	16.73	16.36	16.28	16.00
64 × 64	15.66	15.19	15.15	15.06
128 × 128	12.39	11.28	11.55	11.13

4.5. The Extended Yale B Face Database

The extended Yale Face Database B [30] contains 16128 images taken from 28 human subjects under 9 different poses, 64 different illumination conditions, this last divided into five subgroups according to the angle between the camera axis and the light source direction. An illustration of the database is shown in Figure 10; Table 4 shows the number of images and the lighting angles in each Subgroup.

Table 4. Lighting Angles of Each Subgroup Images

Subgroup	Number of Images	Lighting Angles
1	7	< 12°
2	12	20°- 45°
3	12	35°- 50°
4	14	60°- 77°
5	19	> 77

Figure 10. A preview image of the Extended Yale B face database

4.6. The CMU PIE Database

The Carnegie Mellon University (CMU) Pose, Illumination and Expression (PIE) face database [31] contain 41,368 facial images of size 640×480 pixels captured from 68 subjects under 13 poses and 43 illumination conditions and 4 different facial expressions. An illustration of the database is shown in Figure 11.

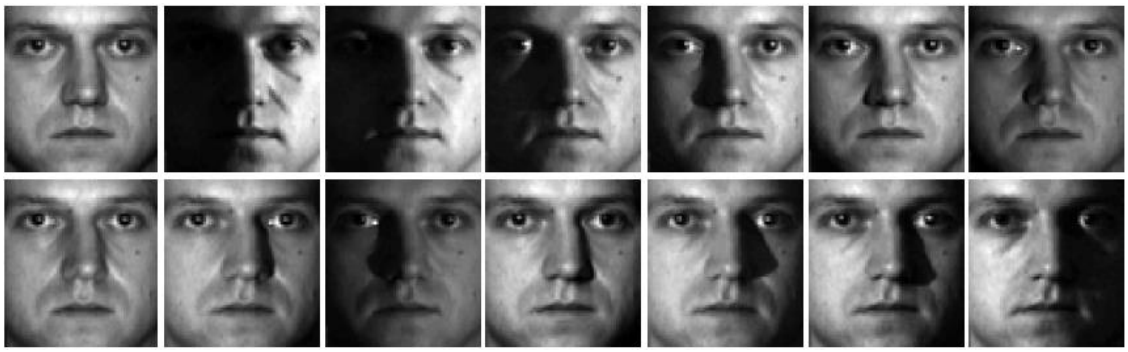


Figure 11. A preview image from the CMU PIE database

4.7. Experimental Steps of Both Databases

Firstly, process the pictures from both databases as follows: the size of all the images was reduced to 224x224 pixel size for VGG16 and ResNet50 and 299x299 for Inception-v3. After that, normalize the input image data from [0-255]. Finally, 20% of the human faces in the database were selected as the training set, and the remaining 80% was used as the test set. During network training, we have employed the mini-batch gradient descend technique [32] with 500 batch size, of epochs being 100 with a weight decay of 0.95 and 0.001 of learning rates.

4.8. Results with and without Contrast Enhancement

In order to demonstrate the effect of M-CLAHE based contrast enhancement method on the recognition rate, we suggested two methods; the first was used without the application of M-CLAHE. The recognition rate results with and without the application of M-CLAHE enhancement algorithm for the Extended Yale B database are shown in Table 5.

Table 5. Results without (a) and with (b) the Application of M-CLAHE on Extended Yale B. RR-Recognition Rate

Methods	S1 a	S1 b	S2 a	S2 b	S3 a	S3 b	S4 a	S4 b	S5 a	S5 b	AVG a	AVG b
VGG16	100	100	100	100	97.80	98.30	84.55	95.6	75.7	92.5	91.61	97.28
ResNet50	100	100	100	100	98.37	99.60	86.61	96.33	79.6	95.8	92.92	98.35
Inception-v3	100	100	100	100	98.81	100	88.73	99.20	83.7	98.00	94.25	99.44

The performance comparison between the two methods (with M-CLAHE and without M-CLAHE) in the Extended Yale B database, show an improvement in the results of each of the three architectures used, for VGG16 the recognition rate of the three subgroups S3, S4, and S5 which has challenging lighting conditions is improved by 0.5 %, 11.05%, 16.8%, respectively, for ResNet50 the RR is improved by 1.23%, 9.72%, 16.2%, and for Inception-v3 1.19%, 10.47%, 14.3%. The recognition rate results with and without the application of M-CLAHE enhancement algorithm for the CMU PIE database are shown in Table 6.

Table 6. Results without (a) and with (b) the Application of M-CLAHE on CMU PIE. RR—Recognition Rate

Methods	1 a	1 b	5 a	5 b	10 a	10 b	20 a	20 b	30 a	30 b
VGG16	88.10	90.20	89.45	93.61	91.57	95.72	93.31	96.39	93.87	97.41
ResNet50	90.10	91.60	91.44	96.43	92.80	97.28	94.48	98.47	95.66	99.53
Inception-v3	91.80	93.80	92.80	98.51	93.70	99.42	95.50	99.75	96.24	99.89

Through the results shown in the tables above, first, we conclude that by increasing the number of samples used in training, the recognition rate increases, second the performance comparison between the two methods (with M-CLAHE and without M-CLAHE) in the CMU PIE database, show an improvement in the results of each of the three architectures used, the recognition rate is improved by 3.54%, 3.87%, 3.65% for VGG16, ResNet50, Inception-v3 respectively.

5. COMPARISON WITH OTHER METHODS

In order to prove the effectiveness of the proposed system, the recognition rate is compared with the existing methods for face recognition. Tables 7 and 8 show the performance comparison of the recognition rate between our method and with other methods for the Extended Yale B and the CMU PIE databases.

Table 7. Shows the Comparison between Different Methods and Our Method for the Extended Yale B Face Database

Methods	Recognition Rate %
LSP [33]	85.6
POEM [34]	90.5
BLCP [35]	98.3
VGG16	97.28
ResNet50	98.35
Inception-v3	99.44

Table 8. Shows the Comparison between Different Methods and Our Method for the CMU PIE Face Database

Methods	Recognition Rate %
Low rank matrix approximation [36]	90.55
VGG-Face [37]	93.16
Volterrafaces [38]	97.1
VGG16	97.41
ResNet50	99.53
Inception-v3	99.89

To confirm the effectiveness of our proposed system, we compared our method with differents methods were employed the Extended Yale B and CMU PIE databases. After the comparison process, it is clear that the recognition rate of our method outperforms the methods mentioned above. Thus we can conclude that our method is robust against pose variation, Illumination changes, and facial expression.

6. CONCLUSIONS

A face recognition (FR) method is presented in this paper. The proposed system based on the Viola-Jones face detection algorithm, and facial contrast enhancement, and three deep convolution neural network architectures. We compared performance between all image enhancement methods the evaluation and judgment of each method were done by the calculation of the (AMBE) and (PSNR) parameters through these results we concluded that Modified CLAHE outperforms all other techniques in term of contrast and brightness. For feature learning and classification Comprehensive experimentation has been performed using VGG16, ResNet50, and Inception-v3 CNN architecture. 20% of the samples from each subject for Extended Yale B and CMU PIE databases were selected as the training set, and the remaining 80% was used as the test set. Moreover, the performance of the proposed method in this paper is compared with the state-of-the-art methods on these databases it is clear that our method outperforms all the other method mentioned above.

ACKNOWLEDGEMENTS

This work is supported by a research project about design and implementation of a surveillance system based on biometric systems for the detection and recognition of individuals and abnormal behaviors (N° A25N01UN080120180002).

REFERENCES

[1] Lihong Zhao, Yu Cai, Xinhe Xu. "Face Recognition Based on Correlation of Wavelet Transform Images", 2006, 6th World Congress on Intelligent Control and Automation, 2006.
[2] Rabul Saikia, Aditya Bihar Kandali. "*DWT-ELBP based model for face recognition*", 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017.
[3] H. Nguyen, L. Bai, and L. Shen, "Local gabor binary pattern whitened pca: A novel approach for face recognition from single image per person," Advances in Biometrics, pp. 269–278, 2009.
[4] Meihua Wang, Hong Jiang, Ying Li. "*Face recognition based on DWT/DCT and SVM*", 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), 2010.
[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv Prepr. arXiv1409.1556, 2014.
[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "*Deep residual learning for image recognition*." In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
[7] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "*Going deeper with convolutions*." In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015.
[8] O. M. Parkhi, et Al, "*Deep Face Recognition*," in British Machine Vision Conference, vol.1, no. 3, pp. 1-12, 2015.
[9] J. C. Chen, et Al, "Unconstrained face verification using Deep CNN features," in WACV, March 2016, pp. 1–9.
[10] D. Yi, et Al. "Learning face representation from scratch." arXiv preprint arXiv:1411.7923, 2014.
[11] J. S. Chan, G. S. Hsu, H. C. Shie, Y. X. Chen, "*Face Recognition by Facial Attribute Assisted Network*, " 2017 IEEE International Conference on Image Processing (ICIP),pages 3825-3829, Sept. 2017.
[12] Zhanfu A, et Al,13th IEEE International Conference on Automatic Face & Gesture Recognition, pp 416-422, 2018.
[13] Hajar Filali, Jamal Riffi, Adnane Mohamed Mahraz, Hamid Tairi. "*Multiple face detection based on machine learning*", 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), 2018.
[14] Bendjillali, R.I.; Beladgham, M.; Merit, K.; Taleb-Ahmed, "*A Wavelet-Based Facial recognition*." in 6th International Conference on Control Engineering & Information Technology (CEIT), IEEE 2018.
[15] Lee, P.-H.; Wu, S.-W.; Hung, Y.-P. "Illumination compensation using oriented local histogram equalization and its application to face recognition." *IEEE Trans. Image Process*. 2012, 21, pp. 4280–4289.
[16] Huang, L.; Zhao, W.; Wang, J.; Sun, Z. "Combination of contrast limited adaptive histogram equalization and discrete wavelet transform for image enhancement." IET Image Process. 2015, 9, pp. 908–915.
[17] Bendjillali, R.I.; Beladgham, M.; Merit, K.; Taleb-Ahmed, A. "Improved Facial Expression Recognition Based on DWT Feature for Deep CNN." Electronics 2019, 8, 324.
[18] B. Nahar, "Contrast enhancement with the noise removal by a discriminative filtering process," Concordia University, Montreal, 2012.
[19] Neha Sharma, et Al. "*A Comparative Analysis of Various Image Enhancement Techniques for Facial Images*" International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp 2279-2284, India, Aug. 2015.
[20] M. Yeong-Taekgi, "Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1-8, 1997.
[21] M. C. Hanumantharaju, et Al, "Design of Novel Algorithm and Architecture for Gaussian Based Color Image Enhancement System for Real-Time Applications," Advances in Computing, Communication, and Control, Springer Berlin Heidelberg, pp. 595-608, 2013.
[22] Arel, D; et Al, T.P. "Deep machine learning-A new frontier in artificial intelligence research [research frontier]." *IEEE Comput. Intell. Mag. 2010*, 5,pp. 13–18.
[23] Shijie Qiao, Jie Ma. "A Face Recognition System Based on Convolution Neural Network", Chinese Automation Congress (CAC), 2018.
[24] Minjun Wang, Zhihui Wang, Jinlin Li. "*Deep convolutional neural network applies to face recognition in small and medium databases*", 2017 4th International Conference on Systems and Informatics (ICSAI), 2017.
[25] Bendjillali, R.I.; Beladgham, M.; Merit, K.; et al. "*Facial Expression Recognition Based on DWT Feature for Deep CNN*." in 6th International Conference on Control, Decision and Information Technologies (CoDIT), pp.344-348, IEEE 2019.
[26] K. Yan, S. Huang, Y. Song, W. Liu, N. Fan " *Face Recognition Based on Convolution Neural Network*" Proceedings of the 36th Chinese Control Conference, on, pp, 4077-4081,IEEE, 2017.
[27] Christian Szegedy, et Al. "*Rethinking the inception architecture for computer vision."* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 2818–2826, 2016.
[28] Sahu, S. "Comparative Analysis of Image Enhancement Techniques for Ultrasound Liver Image." *Int. J. Electr.Comput. Eng. 2012*, 2,pp. 792–797.
[29] Kim, M.; Chung, M.G. "Recursively Separated and Weighted Histogram Equalization for Brightness Preservation and Contrast Enhancement." *IEEE Trans. Consum. Electron. 2008*, 54,pp. 1389–1397.

[30] A. Georghiades, et Al. "From few to many: illumination cone models for face recognition under variable lighting and pose." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):pp.643–660, 2001.
[31] T. Sim, S. Baker, and M. Bsat. "*The CMU pose, illumination, and expression (PIE) database*.' In Fifth IEEE International Conference on Automatic Face Gesture Recognition, 2002.
[32] Sergey Ioffe and Christian Szegedy. 'Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167, 2015.
[33] Yinyan Jiang, et Al, "*Log-domain polynomial filters for illumination-robust face recognition*," in Acoustics, Speech, and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014, pp. 504–508.
[34] Ngoc-Son Vu and Alice Caplier, "Enhanced patterns of oriented edge magnitudes for face recognition and image matching," Image Processing, IEEE Transactions on, vol. 21, no. 3, pp. 1352–1365, 2012.
[35] Yichuan Wang, et Al. "*Illumination-robust face recognition with Block-based Local Contrast Patterns*". IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Pp. 1418-1422, USA, 2017.
[36] Chia-Po Wei, Chih-Fan Chen, and Yu-Chiang Frank Wang, "Robust Face Recognition with Structurally Incoherent Low-Rank Matrix Decomposition" *IEEE Trans on Image Processing*, vol. 23, no. 8, August 2014.
[37] M. M. Ghazi, and H. K. Ekenel, "A Comprehensive Analysis of Deep Learning Based Representation for Face Recognition", arXiv: 1606.02894v1, 2016.
[38] L. Chen, "A fair comparison should be based on the same protocol-comments on" trainable convolution filters and their application to face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 622-623, 2014.