

A Case Study on Various Preprocessing Methods and their Impact on Face Recognition using Random Forest

Hanan M. S. Algharib
Computer Engineering
Ankara Yıldırım Beyazıt university
Ankara/ Turkey

Shafqat Ur Rehman
Computer Engineering
Ankara Yıldırım Beyazıt university
Ankara/ Turkey

ABSTRACT

The Random forest is a well-known powerful classifier, that used to classify a wide range of patterns in our daily life for different purposes, it enters into many fields such as images and objects classification. In this paper, we studied the impact of a five-common preprocessing method in face recognition on the random forest performance, the study included applying five different pre-processing methods (Single Scale Retinex, Discrete Cosine Transform, wavelet Denoising, Gradient faces, and the method proposed by tan and et Known as pp chain or TT), each one has applied separately with a general random forest as a classifier, we computed the error rate for each method. The study was conducted on a face recognition system under occlusion and illumination variation. All experiments were done using MATLAB and Extended Yale B database.

General Terms

Face Recognition, Occlusion, Illumination Variation.

Keywords

Random Forest, Gradient Faces, Wavelet Denoising, Discrete Cosine Transform, Single Scale Retinex, TT (tan and et).

1. INTRODUCTION

Classification of Patterns is a very important term and an essential matter in the machine learning, the urgent need for strong Classifiers leads us to think about the possibility of searching for irreplaceable classifiers, Testing the strength of classifiers under different challenges is the focus of our goal. In this regard, we have conducted an actual study on a face recognition system includes human faces with frontal pose under different illumination and occlusion. The preprocessing phase includes applying five preprocessing separately each method has its impact on the random forest. The choice of the best preprocessing method that increase the efficiency of a random forest is the focus of our work. Unstructured forests have a very high velocity, especially in the classification task that require speed such as face recognition at airports and this observation should be taken into consideration. Random trees have been applied in many fields and have been adopted as a strong classifier. Based on the rationale the combination of learning models increases the classification accuracy which is called bagging, the main goal of bagging is to average noisy unbiased models in order to create a model with low invariance.

2. PREVIOUS WORKS

A Survey of some previous researches, [1] Ghoshal (2009) have used random forests to classify faces with occlusion, it also used LBP and Gabor filter as preprocessing. In [2] presents an algorithm for face recognition using modified

local binary pattern (MLBP) and random forest (RF). In [3] suggested to detect and recognize human faces using random forests for mobile robots. In [4] proposed a Gabor wavelet transform to extract features from a small dataset of images, then applied the Random Forest algorithm for classification based on the regression technique. In [5] compute the performance of Random Forest (RF) and Support Vector Machine (SVM) to classify faces. In [6] proposed Principal Component Analysis (PCA) and Eigenface to extract features, then applied classifiers K-Nearest Neighbors (K-NN), Random Forest and K-Star for recognition. In [7] suggested combination of Principal Component Analysis (PCA), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) as a feature extractors and random forest as classifier. In [8] proposed algorithm to image fusion between the visible and thermal IR face images based on a-torus wavelet transform and random forest (RF). In [9] face recognition system was introduced using random forests as a classifier with Histograms of oriented gradients as feature representation. In [11] it focuses to enhance illumination variation using adaptive multiscale retinex as a preprocessing to overcome the lighting condition effects. In [11] proposed to combine two methods to image enhancement Based on variants of Retinex algorithms and image epitome. In [12] applied the retinex model on color constancy and lightness and test in an image to study the Performance of a Center/Surround Retinex and its Properties. [13] In this paper 14 normalization algorithm were applied on a face verification system as preprocessing and compared under the illumination condition through 3 verification methods. In [14] a book by Gonzales and earth, explained the Retinex algorithm and discuss its work and advantages from different wide aspects. While in [15] presented a new approach based on discrete cosine transform (DCT) used for image normalization illumination under different lighting conditions. In [16] Robust approach for illumination preprocessing based on Discrete cosine transform (DCT) applied on sample of images with proper scaling of low frequency DCT coefficients. [17] This paper proposed algorithm to extract local information depending on block-based discrete cosine transform. [18] it used DCT as a feature extractor from the images that suffer from poor illumination. In [19] (DCT) is employed to address the human face images in difficult lighting conditions. In Reference [21] a paper, explain a method of a wavelet-based illumination normalization for human face recognition under different directions and strength of light. In [21] The human face images are preprocessed and normalized by a group of combination of preprocessing normalization method (DCT-WD-Gradient face-LCE-Weber faces) to eliminate the negative impacts of illumination. In [22] book discusses in

detail Photometric Normalization Techniques for inconstant Illumination Examines the impacts of preprocessing (WD-DCT-SSR) and other methods on the images that Distorted images by illumination. In [23] this paper addressed the problem of illumination based on method called "Gradientfaces" The proposed method has the ability to extract illumination insensitive. [24] comparative study, it discussed a group of illumination variation preprocessing (HE, LT, GIC, DGD, LoG, SSR, GHP, SQI, LDCT, LTV, LN and TT) with different three classes: gradient and edge extraction, gray-level transformation and face reflection field estimation. In [25] face image with different illumination solved using DCT coefficients in logarithm domain. In [26] Discussed a lot of normalization preprocessing methods for practical FRS, in local face representations, one of presented methods is a preprocessing chain which is consist of a series of sequential processing (Gamma Correction, Difference of Gaussian (DoG), Filtering Masking, Contrast Equalization), that eliminates uncontrolled lighting conditions.

2.1 The Aim of study

The main goal is to measure the impact of five suggested preprocessing methods on the random forest performance.

The preprocessing method that improves the performance of random forests will be considered as the best preprocessing method suitable to apply with the random Forest classifier in case of occlusion and illumination variation. Random forest will also be applied without any of these preprocessing methods to demonstrate their performance when it standalone. The study will be conducted on a face recognition system under difficult conditions of illumination variation and occlusion. The main motive of this paper is to introduce a new and unique case study of random forest efficiency as a strong classifier with different preprocessing methods and examine the impacts. The processing and classification stages are important stages in any classification system where they are complementary processes in any recognition system. This study could provide a Foundation stone in face recognition system that uses the random forest classifier, it focused on two Common problems which are occlusion and illumination variation, also Constitute a basis for information that can be used as a reference to future studies.

3. THE PROPOSED METHODS

Five different powerful preprocessing methods that are used to solve illumination variation problem were applied in this study.

- Single Scale Retinex (SSR)
- Discreet Cosine Transform (DCT)
- Wavelet Denoising (WD)
- Gradient Faces (GF)
- PP Chain (TT)

The following figure illustrates two important stages in the recognition system and the impact between them. The pre-processing stage where applied five different pre-processing, which they are a Collection of powerful pre-processing methods used in illumination varying for recognition purpose, and the classification stage where the random forest as classification algorithm used to make classification decision based on a number of random decision trees.



Faces under Random occlusion and uncontrolled lighting.

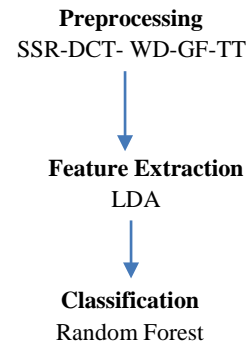


Fig 1: Diagram to explain the proposed Techniques

The figure above explains the five suggested preprocessing methods in the preprocessing stage, and the random forest in the classification stage. In our face recognition system, we use the Linear Discriminant Analysis (LDA) as a necessary step for feature extraction and dimension reduction and to reduce the huge image dimensions.

3.1 The imaging model and retinex theory

Illumination model or lighting model is a well-known model in image processing, many of illumination preprocessing methods were derived from this model, it based on two factors L (luminance) and R (Reflectance).

$$I(x,y) = R(x,y) * L(x,y) \quad (1)$$

$L(x,y)$ is the luminance (amount of lighting) is a dynamic (changes slowly with the spatial position of the image I) Contains a blurred image , represents Low frequency phenomenon. $R(x,y)$ is the reflectance at each point (Structure of object surface) illumination invariant or insensitive measure (Key facial features), the edges in the image $I(x,y)$ also represent the edges in the reflectance. $R(x,y)$ represents high frequency phenomenon. Retinex stands for Retina and cortex, Proposed by Edwin land in 1986 ,it is a model of lightness and human vision perception. Retinex improves visual rendering of an image when lighting conditions are not good. To bridge the gap between images and human observation of scenes. A lot of illumination enhancement preprocessing methods developed based on this theory.

3.1 Single Scale Retinex

Single scale retinex also called as center/surround retinex it is straightforward method, it is photometric based techniques derived from imaging model. offered by Edwin Land in 1986 based on his experimentation involving color constancy. Retinex algorithm, is used to enhance images based on Retinex Technique, which mainly suggested to be used in image improving with regard to the brightness, contrast and sharpness of an image mostly through compression in the

dynamic range [11]. The impact of SSR on the image can be illustrates in the figure (2). SSR were applied on a set of images belongs to the same person under different illumination.



Fig 2: The Impact of SSR

Based on the imaging model in eq (1) that mentioned above, Where $I(x, y)$ represents the 2D input image to be enhanced. we convert the image into logarithmic mode as in eq 2, to separate between R and L.

$$\log I(x, y) = \log R(x, y) + \log L(x, y) \quad (2)$$

$$R'(x, y) = \log I(x, y) - \log [I(x, y) * k(x, y)] \quad (3)$$

where “*” denotes the convolution operator, $K(x,y)$ is a smoothing kernel, which can be computed by eq (4). $R'(x,y)$ refer to the illumination invariant reflectance output of the SSR algorithm.

$$K(x, y) = k \exp\left(-\frac{r^2}{c^2}\right) \quad (4)$$

C can be compute by eq(3.5) It is calculated in a manner that it is equal to 1.

$$\iint K(x, y) dx dy = 1 \quad (5)$$

While, $r = (x^2 + y^2)^{1/2}$, $I(x,y)$ is orginal image, $R(x,y)$ is reflectance and $L(x,y)$ represent illumination .

3.2 Discreet Cosine Transform (DCT)

The discrete-cosine-transform-(DCT) is a photometric normalization technique recently introduced by Chen et al in 2006. The technique also relies on the retinex theory and imaging model [22] it supposes a similar assumption about illumination characteristics in varying appearance. The technique assume that the variations of illumination are closely related, to the low frequency coefficients of the DCT transform. these low frequency coefficients are discarded before converting the image back to the spatial domain by using the inverse DCT to obtain face representation under constant illumination, the following figure shows the impact of DCT on the human face Under different possibilities of illumination belongs to the same person, where the upper images represent the images under illumination and the images on the lower rows represent the images that has been processed by DCT.



Fig 3: The Impact of DCT

The following steps explain how DCT works:

1- First step, separates between the reflectance (R) and luminance (L) by apply the logarithm (log) to the input image $I(x, y)$, next step, the all image is converted to frequency domain by apply the DCT transform, where the coefficients of DCT manipulated until, reaching the goal of obtaining a constant illumination, the initial DCT coefficient $C(0,0)$ is set to as the eq (6) :

$$C(0,0) = \log \mu \cdot \sqrt{MN} \quad (6)$$

MN refers to the dimensions of the input image $I(x, y)$, μ is the average value of $I(x, y)$. The number of DCT coefficients is defined previously from the DCT coefficient matrix in a snaky way. encoding the low-frequency information of the image is then set to zero. In the final step, the modified matrix of DCT coefficients is converted again to the local domain by taking the inverse of DCT to generate the representation of the facial image under illumination varying.

3.3 Wavelet Denoising (WD)

This technique presented by (Zhang et al., 2009), it is one of the photometric normalization technique, the wavelet-denoising approach is exploited to get a face representation under illumination invariant [22] . To clarify the WD impact on illumination variation we applied it on a set of images belong to the same person of class 8 under different illumination as shown in the following figure.

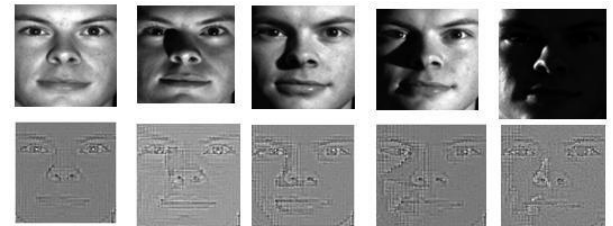


Fig 4: The Impact of WD

The wavelet denoising depends on four major steps here is an important operation applied during denoising model they are: (2D DWT), thresholding, 2DIDWT. The following figure in the next page summarizes the general steps of wavelet denoising, where it starts with input image in logarithm domain then decompose the image and apply thresholding then reconstruct the image to obtain the output image in local domain. (see the figure 5)

The wavelet denoising steps during applied on the image

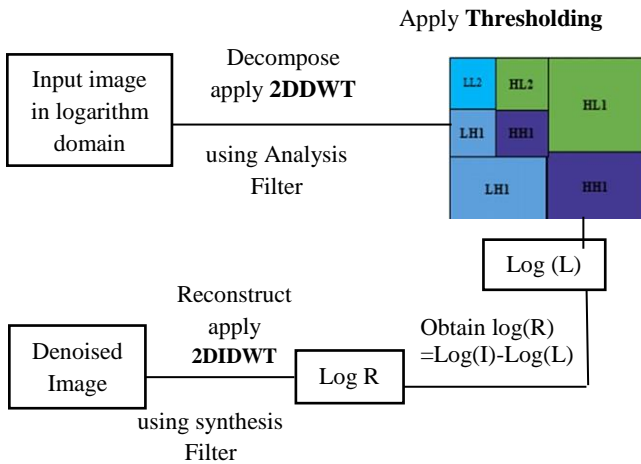


Fig 5: The Steps of WD

3.4 Gradient Faces (GF)

Gradientfaces is also known as an illumination insensitive measure this technique proposes the principle that based on extract illumination insensitive facial feature under different lighting situations in gradient domain. There are many ways to compute the gradient of image. However, the numerical calculation of derivative directly without smoothing (gradient) is typically not the ideal computation. To compute the gradient in more stable way, we firstly smooth the image by using Gaussian kernel function with a convolution type smoothing, the numerical calculation of gradient will be more stable in calculation [22]. moreover, the main feature for using Gaussian kernel is multiplied, also robustness of Gradientfaces is very high to deal with the image noise in satisfied, it can eliminate the effects of shadows. The steps of applying the Gradientfaces illustrated below, the following figure illustrates the Gradientfaces impact.

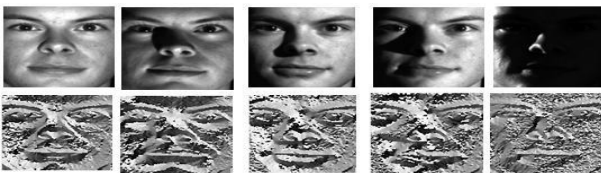


Fig 6: The Impact of Gradient face

1- The first step starts with Smoothing operation which applied on the input image with convolving Gaussian kernel function

$$I' = I * G(x, y, \sigma) \quad (7)$$

Where * is a convolution operator, while $G(x, y, \sigma)$ Contains the following equation

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp(-x^2 - y^2) \quad (8)$$

Where it represents the Gaussian kernel function with standard deviation σ .

2- Compute the Gradient of image I by feeding the soothed image through convolution operation with the derivation of Gaussian kernel function in the x, y direction

$$I_x = I' * G_x(x, y, \sigma) \text{ and } I_y = I' * G_y(x, y, \sigma) \quad (9)$$

Where : $G_x(x, y, \sigma)$, $G_y(x, y, \sigma)$ are the derivation of Gaussian kernel function in the x, y directions Respectively 3- compute the illumination in insensitive measure by following

$$G = \arctan(I_x/I_y) \in [0, 2\pi] \quad (10)$$

3.6 PP chain (TT)

also known as (TT) Proportion of its inventors (Tan and Triggs) [22], as shown by the name it is a chain of preprocessing steps, one of illumination normalization method, it consists of a series Sequential stages (Gamma Correction, Difference of Gaussian (DoG), Filtering Masking, Contrast Equalization). This chain is suggested to be applied like any preprocessing before feature extraction, the main goal is to eliminate the effects of illumination, local shadowing and highlights with keeping the essential parts of facial appearance.

1- **Gamma Correction** γ : the first step starts with applying the Gamma correction, which is a nonlinear transformation method in gray-level it replaces the gray-level of image I by I^γ .

Where $\gamma \in [0, 1]$ the parameter can be defined by the user to enhances the darkness or shadowed regions and press it in the bright regions and at highlights in the local dynamic range of the image, the default value of γ usually equal to 0.2.

2- Difference of Gaussian (DoG) Filtering:

Gamma correction couldn't eliminate the impact of all intensity gradients such as shading influences, Shading comes from the surface structure that effected by illumination gradients. this make the spatial information in frequency domain hard to separate from those effects. To solve this problem, we using the Filtering process but, there are some factors must be considered when selecting filters, for example using High pass filtering may removes both the useful and unwanted information, this make the recognition task more complex. In contrast suppress the high frequencies reduces the noise without effecting on the important recognition aspects. DoG filtering is a proper filter to keep the fine details. Dog filter has two filters to deals with high frequency and low frequencies at the same time so this make it as a strong filter that gives the important details that needed in recognition [22,26].

3-Masking

Masking is an optional step, we need to apply the mask in case of getting facial regions such as hair, beard that may considered as irrelevant or too variable, so this area needed to be mask out and the mask should be applied on this point [26].

4- Contrast Equalization

It is the final step of preprocessing chain, it basically proposed to deal with any extreme values may be generated by the filtering process, it takes extreme noise in the dominant noise source and makes it approximately uniform [26]. The figure (7) shows the impact of TT on the person belongs to class 8.



Fig 7: The Impact of TT

4. FEATURE EXTRACTION

Feature extraction is the task of reducing the high dimension of training data to a set of useful features. we applied LDA as a necessary step before the classification step using random forest. In our face recognition system we used the Linear Discriminant Analysis (LDA) to reduce the huge image dimensions, thus restricting the classification of selected features in a limited number of samples.

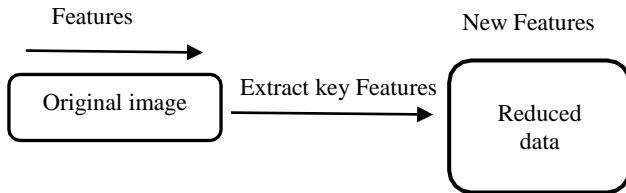


Fig 8: Feature extraction

5. RANDOM FOREST ALGORITHM

It is an ensemble learning method based on generate many classifiers and aggregate their results. Random forest (RF) is a supervised classification algorithm and a strong modern approach to data classification and modeling. As illustrated from its name “forest” it’s a group of trees, each tree consists of splitting node, internal node or decision node and leaf node. The image is classified by sending it down to the terminal nodes of trees and aggregating the reached leaf distributions based on voting. Randomness can be at two points during training: in subsampling the training data so that each tree is grown using a different subset. each subset represents sub node that may selected to be split more based on GiNi index, this process repeated in recursive manner. The classification results from each tree are collected for an input image, then the majority voting is gathered to give the resulting class label. random forest is grown using randomization strategy. The leaf nodes are the final node that used to make the final decision. Each internal node contains a test to select the best splits of data to be classified.

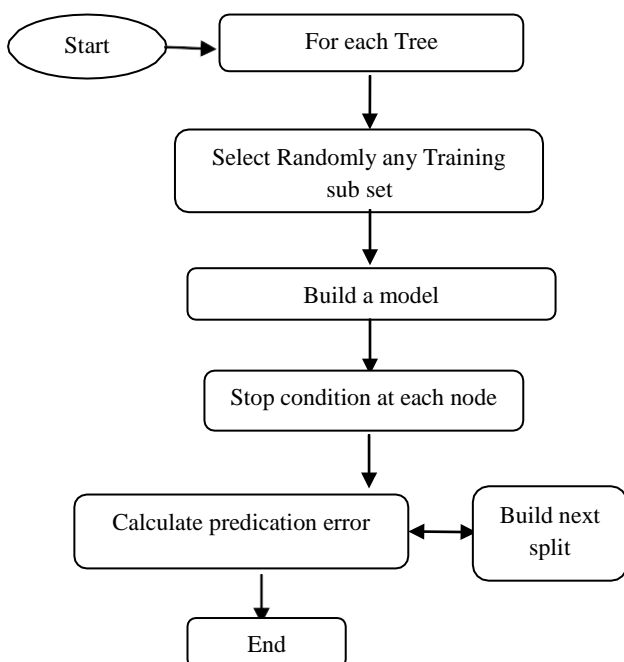


Fig 9: The Flow chart of Random forest

6. IMPLEMENTATION

The focus of our study on the random forests efficiency and the impact of different preprocessing on random forest performance, we divided the experiments into 3 as following:

- Experiment 1:** Compute the error rate for each preprocessing with Random forest (Applied on subset 1 (faces with illumination variation).
 - Single Scale Retinex (SSR).
 - Discreet Cosine Transform (DCT).
 - wavelet Denoising (WD).
 - Gradient Faces (GF).
 - PP Chain (TT).
- Experiment 2:** Compute the error rate for each preprocessing with RF under different occlusion (Applied on subset 1+2).
- Experiment 3:** compute the Preprocessing time and classification time.

6.1 Database

Experiments will be conducted on Extended Yale B database which contains a lot of frontal human faces under strong effects for lighting. The number of stored images is about 2432 image of 38 persons with frontal pose, each person has 64 images were captured under different illumination conditions. We have added a block of random occlusion on frontal faces from 10% to 50%, to measure the efficiency of the random forest as classifier even under Random occlusion. The database is divided into five subsets based on the degree of the light source direction and the central camera axis ($0^{\circ}, 12^{\circ}, 25^{\circ}, 50^{\circ}, 77^{\circ}, 90^{\circ}$). Fig.10, the image size is resized to 128 x 128. Subset 1, subset 2 are images with simple illumination effects, the effects are gradually increased until subset 5 which is contain the strong illumination effects, the subset 3 are images with different occlusion.

e.g. (Lighting $50 < \theta > 77$)
(Occlusion=50%)



Fig 10: Samples of challenging faces

6.2 Computation the Error rate

We compute the error rate to refer the number of images that classified incorrectly and usually calculate as shown in equation below.

$$ER = \frac{\text{Number of Classified images incorrectly}}{\text{Total number of testing images}} * 100$$

We dependent the equation above to calculate the accuracy of random forest performance as a classifier with each preprocessing. In contrast, the recognition rate (RR) can be calculated instead of the error rate as found in some research

but in this case the equation will vary, it will have computed as the recognized images correctly divided by total images. However, in our study, we calculated the error Rate not the Recognition Rate, since the focus of our study is to measure performance accurately to determine the rate of error and to know which of the preprocessing give a large error rate.

7. EXPERIMENTS

7.1 Impact of Preprocessing Methods

We computed the Error Rate (ER) or accuracy for Five preprocessing methods, also we used Subset 1 (samples with illumination variation) as training set and the whole database as testing set. we used five readings for each preprocessing for each different number of trees and computed the average for each trial to get more accurate results. The table below shows the average results of each preprocessing, from the table we can note that the ER increases in case of no preprocessing and in case of using minimum number of trees, in our experiment the minimum number of trees was equal to 10 The WD an TT shows a good result with minimum number of trees, while the other methods shown a good result from 20 number of trees to above. The results of this experiment is explained in table 1 in the next page.

There is no one static reading every time, but there is a certain reading range, due to the randomized strategy of the random forests in every classification process.

7.1.1 Evaluation of Experiment 1 In Table (1)

The Error rate of random forest is increased without any preprocessing, also increased when the number of trees reaches to the minimum, while the error rate Significantly decreased when the number of trees ≥ 500 . All the preprocessing methods (TT, WD, GF, SSR,DCT) Sequentially, give a good result when the number of trees is equal to 50 or more. The PP chain (TT) and WD gives a good impact on random forest even in little number of trees = 10. In case the number of trees = 10, and without any preprocessing the AVG of ER almost equal to 36%. The following bar chart explains the error rate for each preprocessing.

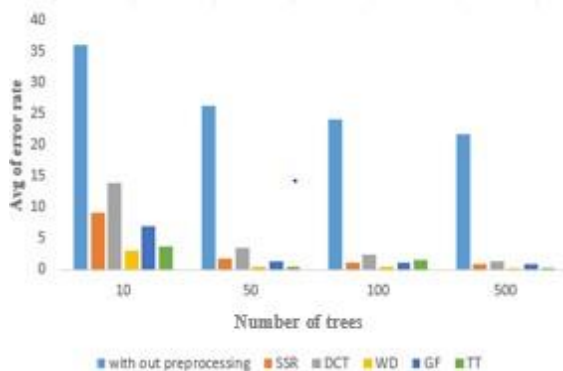


Fig 11: Bar chart of experiment 1 results

As shown in the figure above, no preprocessing gives the high error rate, while the (TT,WD) shown an excellent result under minimum number of trees, while (GF,SSR,DCT) did not give a good results with minimum trees =10.

7.2 Preprocessing methods under occlusion

we Computed the error rate for each preprocessing under different occlusion and static number of trees = 500, we used Subset 1 and subset 2 (samples with illumination variation) as training set while, subset 3 (samples with different occlusion)

as testing. The goal to show the performance of RF under occlusion and how preprocessing may be affected by occlusion, as we know the suggested pre-processing methods is mainly applied for the illumination variation problem not for occlusion, but some methods may deal with occlusion as a noise and can enhances the result. The table (2) in the page 8 shows the results of each preprocessing under different occlusion.

The Random Forest is able to classify the occluded faces with different ratios of occlusion even without pre-processing, but the error rate increases when there is no preprocessing, also the number of trees that used in classification is another factor, since the minimum number of trees gives the higher error rate even if there a preprocessing.

7.2.1 Evaluation of Experiment 2 In Table (2)

As we seen in the table (2) the Random forest does not affect by occlusion, no matter how much the occlusion ratio is, RF shows a good classification even under 50% occlusion, but the problem arises in the case of the absence of preprocessing, as seen from the results table (2) the higher error rate in case the absence of preprocessing. Another important problem can arise in case of choice a sensitive preprocessing for occlusion such as SSR it gives a negative impact on classification process. SSR gives a negative impact that are greater than the impact in case of absence of preprocessing itself. DCT and GF had shown the best results under occlusion and classification with random forest, this is due to its powerful nature to resist the occlusion. The TT and WD had shown also a positive impact.

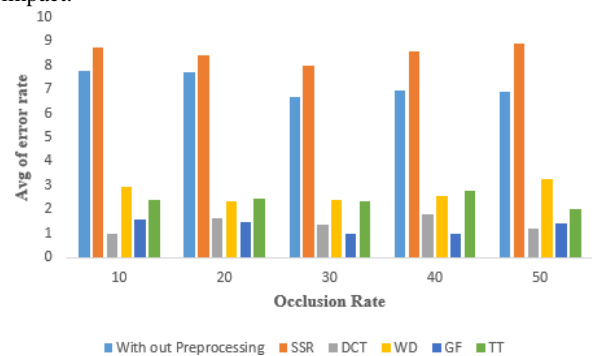


Fig 12: Bar chart experiment 2 results

As shown in the figure above, the SSR is a case sensitive to occlusion, the DCT and GF had given the best result with RF, also the TT and WD had given a positive impact with RF, while in case of absence the preprocessing the RF gives a high error rate of classification.

In this Experiment, we applied the techniques below as a preprocessing and the random forest as a classifier with number of trees from **10 to 500**, applied on the **subset 1** as training and the all database as testing.

Table 1. Impact of preprocessing methods on ER (Experiment 1)

Preprocessing	Trees Number	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	AVG
No pre-processing	10	37.39	37.58	35.64	35.22	34.62	36.09
	50	27.14	24.05	26.77	27.42	25.57	26.19
	100	23.68	24.33	24.14	24	24.06	24.04
	500	21.51	22.16	21.09	21.23	23.08	21.81
SSR	10	9.56	9.23	9.37	8.68	8.67	9.10
	50	1.94	1.66	2.17	1.34	1.29	1.68
	100	1.06	0.92	1.34	0.97	1.15	1.08
	500	1.25	1.20	0.78	0.74	0.88	0.97
DCT	10	14.04	15.05	12.60	14.13	13.67	13.89
	50	3.23	3.97	3.23	3.42	3.51	3.47
	100	2.49	2.68	2.49	2.86	2.17	2.53
	500	1.71	1.52	1.39	1.38	1.20	1.44
WD	10	3.46	3.23	3.09	2.35	3.19	3.06
	50	0.42	0.32	0.60	0.74	0.60	0.53
	100	0.37	0.42	0.46	0.32	0.37	0.38
	500	0.32	0.37	0.37	0.32	0.23	0.32
GF	10	8.86	6.37	6.42	6.14	6.83	6.92
	50	1.20	1.66	1.11	2.21	0.92	1.42
	100	0.97	1.29	1.06	0.78	1.20	1.06
	500	1.02	0.69	0.83	0.60	1.11	0.85
PP-Chain (TT)	10	4.34	3.83	3.28	3.60	3.19	3.64
	50	0.55	0.69	0.60	0.41	0.46	0.54
	100	0.32	0.32	3.37	3.32	0.18	1.50
	500	0.23	0.32	0.27	0.41	0.32	0.31

In this Experiment, we applied the techniques below as a preprocessing and the random forest as a classifier with number of trees = 500, applied on the subset 1 and subset 2 as training and the subset 3 as testing under different occlusion from 10% to 50% .

Table 2. preprocessing methods Under Different Occlusion (Experiment 2)

Preprocessing	Occlusion Ratio (%)	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	AVG
No preprocessing	10	8.57	7.47	8.13	9.01	5.71	7.77
	20	9.45	7.69	7.03	7.91	6.59	7.73
	30	8.13	7.69	4.17	6.15	7.25	6.67
	40	7.47	8.13	5.71	6.81	6.59	6.94
	50	7.91	6.59	8.13	5.71	6.15	6.89
SSR	10	7	9	7	11.8	9	8.76
	20	7.47	7.91	8.35	7.47	10.76	8.39
	30	7.25	7.91	8.35	8.79	7.69	7.99
	40	8.35	7.25	9.23	9.23	8.79	8.57
	50	9.01	8.35	9.89	8.79	8.57	8.92
DCT	10	1.31	1.31	0.87	0.65	0.87	1.00
	20	1.97	2.19	1.09	1.97	1.09	1.66
	30	1.09	0.65	2.85	1.09	1.09	1.35
	40	1.31	1.97	2.19	1.75	1.75	1.79
	50	1.53	1.31	1.97	1.09	0.21	1.22
WD	10	3.07	3.29	3.51	2.19	2.63	2.93
	20	2.41	1.31	2.19	3.29	2.41	2.32
	30	2.41	1.75	1.97	3.07	2.85	2.41
	40	1.53	2.85	3.07	2.41	2.85	2.54
	50	2.85	4.61	2.41	2.63	3.73	3.24
GF	10	1.31	1.97	1.75	1.09	1.75	1.57
	20	1.97	1.75	1.09	0.87	1.75	1.48
	30	2.41	0.43	0.31	0.43	1.31	0.97
	40	1.09	1.31	1.09	0.44	1.09	1.00
	50	1.97	1.53	1.31	1.09	1.31	1.44
TT	10	1.97	2.63	2.41	2.63	2.41	2.41
	20	2.41	3.29	1.31	2.19	3.07	2.45
	30	2.63	2.41	2.85	2.63	1.09	2.32
	40	3.07	2.85	2.19	3.07	2.85	2.80
	50	1.53	3.29	1.97	2.63	0.65	2.01

7.3 TIME COMPUTATION

In the experiment, we computed the preprocessing time and classification time. in order to compute the total recognition time.

- Software & Hardware Specification

We applied on computer with Core i7, windows 7, RAM 6 Giga. MATLAB 2012

7.3.1 Preprocessing Time

The time taken to apply the preprocessing on the whole Database images, as we see from the table below the system is very fast without a Preprocessing it takes about 8 seconds, while the DCT has consumed the longest preprocessing time, while the Gradient faces (GF) has consumed the shortest preprocessing time compared with the other methods, this is maybe depending on the algorithm complexity that are adopted by each preprocessing.

Table 3. Preprocessing Time for each preprocessing

PREPROCESSING NAME	PREPROCESSING TIME
Without Preprocessing	8 Sec
SSR	22 Sec
DCT	70 Sec
WD	37.5 Sec
GF	16.9 Sec
TT	25 Sec

From the table above, we note that DCT and WD we note that DCT and WD have taken a long time in preprocessing, this depends on the algorithm complexity which contains a lot of steps includes decompose and reconstruct the image again (see the DCT and WD steps on page 3). The following Bar chart shows the preprocessing time for each preprocessing applied, we can note that DCT is the highest Preprocessing Time while without preprocessing is the fast preprocessing, GF is the lowest preprocessing Time.

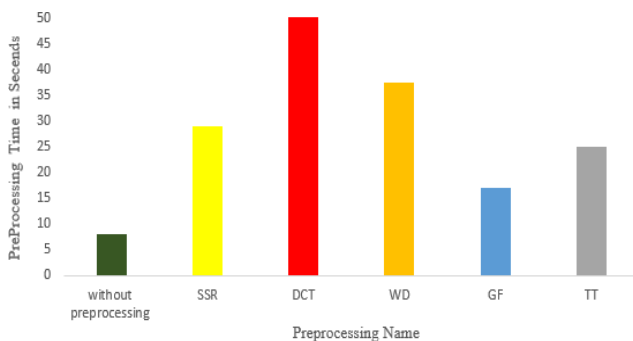


Fig13: Bar chart for preprocessing time

7.3.2 Classification Time

Is the time taken by the random forest classifier to train and classify the images in the database, the random forest is known as a fast and strong classifier, it classifies faces even under occlusion. we noted this from the results, it works efficiently even under 50% occlusion. The Classification Time and the error rate are affected by the number of trees that used in the classification process, when the number of trees is low the classification time is decreased while, the error rate is increased.

Table 4. Classification Time for each number of trees

NUMBER OF TREES	CLASSIFICATION TIME
10	0.06 Sec
50	0.25 Sec
100	0.47 Sec
500	2.28 Sec
1000	4.47 Sec

As shown in the table above when the number of trees =10 the classification time becomes very fast, while the number of trees = 1000 the time increases until 4.47. As we have seen from the previous experiments It is important to note that the minimum number of trees gives a fast classification time with a high error rate and Vice versa.

The following Bar chart shows the Classification time for each Number of Trees, we can note that when Number of trees =1000 we get a slow Classification Time, while when number of trees =10 we get a fast Classification Time, but with a higher error rate as we mentioned before.

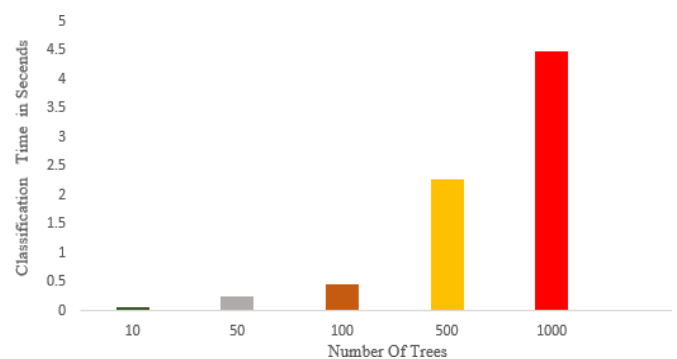


Fig 14: Bar chart for classification time

- Total Time

The Total Time is computed by Adding the training time to classification time.

8. CONCLUSION

In this paper, we studied the impact of five common preprocessing methods in face recognition on the random forest performance. The study was conducted on Extended Yale B database which contains faces under occlusion and different illumination. We applied three different experiments. Based on the practical experiments we got the following results: In illumination variation experiment, we noticed that TT and WD have shown an excellent result with random forest even under using a little number of trees =10.

The other methods like GF and SSR have shown a good result when using number of trees equal to 50 or above, while DCT has shown the lowest performance with random forest compared with the other methods, also in case of absence preprocessing method with random forest we get the highest error rate. In occlusion experiment, GF and DCT have shown the best results with random forest even under 50% occlusion, TT and WD have shown a good result with random forest, while SSR and absence preprocessing method

, we get the highest error rate. In time experiment, we noticed in case of there is no preprocessing we get the fastest preprocessing time, the GF has shown the fastest preprocessing time compared with the other methods, while the DCT has shown the slowest preprocessing time.

9. FUTURE WORK

The pattern recognition is a very wide area, where the human faces are considered as a special case of pattern recognition. In the future work, we plan to conduct a methodology based on several classifiers to explain how to choose the appropriate preprocessing for each classifier, in other areas of patterns classification.

10. REFERENCES

- [1] V. Ghosal, "Efficient Face Recognition System Using Random Forests," Work, 2009.
- [2] B. O. Connor and K. Roy, "Facial Recognition using Modified Local Binary Pattern and Random Forest," vol. 4, no. 6, pp. 25–33, 2013.
- [3] H. O. A. Achen and P. R. O. F. G. E. L. Akemeyer, "Detection and Recognition of Human Faces using Random Forests for a Mobile Robot."
- [4] V. Vishwakarma and A. K. Srivastava, "Performance Improvements in Face Classification using Random Forest," vol. 2, no. 3, pp. 2384–2388, 2012.
- [5] E. Kremic and A. Subasi, "Performance of Random Forest and SVM in Face Recognition," vol. 13, no. 2, pp. 287–293, 2016.
- [6] F. Classifiers et al., "A Facial Recognition System Based on Techniques of Principal Component Analysis and Auto faces with K-NN , K-Star and Random," vol. 12, no. April, pp. 7–14, 2013.
- [7] V. W. Parate, M. T. Scholar, and V. Branch, "PCA , DCT and DWT based Face Recognition System using Random Forest Classifier," vol. 3, no. 6, 2015.
- [8] A. Seal, D. Bhattacharjee, and M. Nasipuri, "Human Face Recognition using Random Forest based Fusion of A-trous Wavelet Transform Coefficients from Thermal and Visible ," AEUE - Int. J. Electron. Common., no. May, 2016.
- [9] A. I. Salhi, M. Kardouchi, and N. Belacel, "Fast and efficient face recognition system using Random Forest and Histograms of Oriented Gradients," pp. 293–303, 2012.
- [10] K. A. Ishak, S. A. Samad, M. A. Hannan, and M. M. Sani, "Image Enhancement Technique using Adaptive Multiscale Retinex for Face Recognition Systems 2 The Original Multiscale Retinex.
- [11] B. Academy, O. F. Sciences, and I. Technologies, "Image Enhancement Using Retinex Algorithms and Epitomic Representation Violeta Bogdanova," vol. 10, no. 3, pp. 10–19, 2010.
- [12] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and Performance of a Center / Surround Retinex," vol. 6, no. 3, pp. 451–462, 1997.
- [13] M. Leszczyński, "Image Preprocessing for Illumination Invariant Face Verification," pp. 19–25
- [14] C. Constancy, "Fast RETINEX for Color Image Enhancement: Methods and Algorithms," pp. 1–23.
- [15] V. S. Manjula, "Image Normalization Robust using Histogram Equalization and Logarithm Transform Frequency DCT Coefficients for Illumination in Facial Images," vol. 3, no. 11, pp. 1891–1895, 2014.
- [16] V. P. Vishwakarma, S. Pandey, and M. N. Gupta, "Adaptive Histogram Equalization and Logarithm Transform with Rescaled Low Frequency DCT Coefficients for Illumination Normalization," vol. 1, no. 1, pp. 1–5, 2009.
- [17] H. K. Ekenel and R. Stiefelham, "local appearance based face recognition using discrete cosine transform."
- [18] Z. Sufyanu, F. S. Mohamad, A. A. Yusuf, and B. Mustafa, "discreet Cosine Transform," no. February, 2016.
- [19] W. Chen, M. J. Er, and S. Wu, "Transform in Logarithm Domain," vol. 36, no. 2, pp. 458–466, 2006.
- [20] O. Of, "A Wavelet-Based Image Preprocessing Method for Illumination Insensitive Face Recognition," vol. 1731, pp. 1711–1731, 2015.
- [21] A. P. C. S. Devi and V. Mahesh, "performance evaluation of illumination normalization techniques for face recognition," vol. 1, no. 2, pp. 11–20, 2013.
- [22] "Photometric Normalization Techniques for Illumination Invariance AVTOR : Vitomir Štruc INTERNAL REPORT : LUKS Photometric Normalization Techniques for Illumination."
- [23] C. N. Raghu, "Illumination Insensitive Face Recognition Using Gradientfaces," no. 1, pp. 38–44, 2012.
- [24] H. Han, S. Shan, X. Chen, and W. Gao, "A comparative study on illumination preprocessing in face recognition,"
- [25] Z. Yang, X. He, W. Xiong, and X. Nie, "Face Recognition under Varying Illumination Using
- [26] X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions," pp. 1–16, 2009.