BAU Journal - Science and Technology

Volume 1 | Issue 2
ISSN: 2706-784X

Article 11

June 2020

Blind Image Quality Assessment for Face Pose Problem

Cerine Tafran
PhD Student, Faculty of Science, Beirut Arab University, Beirut, Lebanon, cerine.tafran@bau.edu.lb

Mohamad El-Abed
Associate Professor, Computer & Information Systems Department, Rafik Hariri University, Beirut, Lebanon, alabedma@rhu.edu.lb

Ziad Osman
Professor, Faculty of Engineering, Beirut Arab University, Beirut, Lebanon, zosman@bau.edu.lb

Islam Elkabani
Assistant Professor, Faculty of Science, Beirut Arab University, Beirut, Lebanon, islam.kabani@bau.edu.lb

Blind Image Quality Assessment for Face Pose Problem

Abstract

No-Reference image quality assessment for face images is of high interest since it can be required for biometric systems such as biometric passport applications to increase system performance. This can be achieved by controlling the quality of biometric sample images during enrollment. This paper proposes a novel no-reference image quality assessment method that extracts several image features and uses data mining techniques for detecting the pose variation problem in facial images. Using subsets from three public 2D face databases PUT, ENSIB, and AR, the experimental results recorded a promising accuracy of 97.06% when using the RandomForest Classifier, which outperforms other classifiers

Keywords

Biometrics, No-reference Image Quality Assessment, SIFT, BLIINDS, Pose, Features extraction and classification

1. INTRODUCTION

Recently, developing image quality assessments for predicting image quality has been considered an interesting topic for research. These assessments are very useful in different applications and systems such as biometric systems (Jain, et al., 2004).

It is very important to consider the quality of the image during enrollment and the verification steps while using biometric systems. The captured face images would be affected by many alterations such as: pose, illumination, blurring, sharpness, distance from the camera, *etc*. These alterations will produce low-quality images that will drop significantly the performance of biometric systems (Bharadwaj, Vatsa, & Singh, 2014). To improve the performance of the biometric systems, low-quality samples must be eliminated and replaced with new good quality images. This will improve the system performance and therefore robustness against attacks. Image quality assessment (IQA) is broadly divided into two techniques: subjective and objective (Khodabakhsh, Pedersen, & Busch, 2019).

The subjective evaluation technique is the assessment of the image quality by human observation. It has the advantage of being the most reliable since it is the most accurate method to assess the quality of an image but it is expensive, unsuitable for real-time applications and is time-consuming. Therefore, subjective metrics could not always apply.

To overcome such limitations, efforts have been made to develop an automated objective image quality assessment that can be applied for real-time applications. Various objective IQA techniques have been proposed that can be categorized into full-reference (FR), reduced reference (RR) and no-reference (NR) also called blind (Wang, Bovik, Sheikh, & Simoncelli, 2004). However, the FR-IQA and RR-IQA techniques are limited and cannot be used by biometric recognition systems. Therefore, efforts have been made over the last twenty years by researchers to build NR-IQA algorithms that do not require a reference image (El-Abed, Ninassi, Charrier, & Rosenberger, 2013), (Kerouh, Ziou, & Serir, 2017).

Several efforts were done in predicting the quality of facial images; however, very few are the works related to detecting the pose problem in those images.

In this paper, we propose a novel no-reference image quality assessment (NR-IQA) method for detecting the pose alteration in facial images. This problem is considered an inevitable problem during the enrollment step when using biometric systems because in general it is difficult, if not possible, to control the face rotation when acquiring the human face image. The pose problem addressed in this paper is when a face image is not frontal due to horizontal face rotation with any arbitrary angle during the enrollment step. The presented method extracts features based on image processing (Solomon & Breckon, 2011) and uses data mining techniques (Liao, Chu, & Hsiao, 2012) for classification purposes.

The paper is organized as follows: Section 2 addresses related research on image quality assessments of face problems. Section 3 presents the suggested methodology for no-reference quality metric for detecting the pose problem. Experimental results and analysis are illustrated in Section 4 using selected images from three public face databases. We conclude our work in Section 5 with future scope.

2. RELATED WORK

In real-time applications, faces are subjected to pose, illumination, expressions, etc. Different models deal with different face variations, but few are the research that deals with the pose problem in quality assessment. Several NR quality assessment algorithms (Mittal, Soundararajan, & Bovik, 2013) (Hou, Gao, Tao, & Li, 2014) (Liu, Pedersen, Charrier, & Bours, 2018) were presented in the literature to assess the quality of images. However, those algorithms are not designed for detecting specific alterations affecting the overall performance of biometric authentication systems such as the pose alteration.

Other papers presented in (Ratyal, et al., 2019) and (Sang, Li, & Zhao, 2016) deals with presenting robust face recognition algorithms against pose alteration problem rather than quality assessment metrics for detecting the pose problem in images.

Therefore, we will focus in the rest of this section on only the NR quality assessment algorithms that were designed to be used by biometric authentication systems.

Liu et al. (Liu, Pedersen, Charrier, & Bours, 2018) analyzed and evaluated the performance of 13 selected blind IQA of face images that dealt with different distorted face image. Their contribution can be used for the development of robust quality metrics for face image quality and can also work on other biometric modalities. However, the pose problem was not addressed in any of the selected IQAs.

El-Abed et al. (El-Abed, Charrier, & Rosenberger, 2015) proposed a quality assessment method for face, fingerprint and hand veins images using two types of information: image and pattern-based quality using Scale Invariant Feature Transform SIFT descriptor (Lowe, 2004). This method used six public biometric databases and was intended to detect blurring, Gaussian noise, and scale alterations. However, they did not address the pose problem in facial images.

Zhang et al. (Zhang & Wang, 2009) adopted a methodology of three quality features to detect the illumination problem in facial images using asymmetry based quality assessment method which was based on local SIFT features. However, they did not address the pose problem in facial images.

Nikitin et al. (Nikitin, Konushin, & Konushin, 2014) proposed a face image quality assessment method in video-based face verification system that tackles four alteration problems including the pose problem. The quality score was produced by calculating the weight fusion with automatic weight tuning. This method provides a trade-off between efficiency and accuracy, reaching a verification accuracy of 74.46%.

Wong et al. (Wong, Chen, Mau, & S, 2011) introduced a novel patch-based face image quality assessment algorithm that is able of handling different face variation problems including a special type of pose variations, which is the in-plane rotations. This method was evaluated in a video-based face verification setting. It reached a verification accuracy of 86.7%.

Other researches such as in (Abayomi-Alli, Omidiora, Olabiyisi, & Ojo, 2015) and (Nasrollahi & Moeslund, 2008) employed full-reference quality assessments to deal with the pose problem in facial images. The main drawback of this approach is that the original images are required, which makes it not suitable for real-life biometrics applications.

Though, work related to pose variation in the state-of-the-art is considered limited when compared to other research work studying other types of alterations such as blurring, noise, and illumination.

Therefore, in order to contribute of having more work related for pose alteration detection in facial images, we present a method that extracts several image features and use data mining techniques to classify the quality of a face image (good or bad). The method contains features presented in previous work (El-Abed, Charrier, & Rosenberger, 2015), some features presented in the literature by other researchers, and new features (presented in section 3.3.2, 3.3.3, and 3.3.4) that have shown good performance in detecting the pose alteration problem. Furthermore, we show in this paper that some of the features used from the literature that aimed for detecting other types of alterations (such as blurring alteration) can detect as well the pose alteration problem.

3. Proposed NR Quality Metric for Pose Problem

The methodology adopted in this paper is illustrated in Fig. 1 where a set of features are extracted from an input image after face detection and then the quality of the image is predicted and classified as good or bad using different supervised learning algorithms as described in Section 4.2. BLind Image Integrity Notator using discrete cosine transform DCT Statistics BLIINDS-I (Saad, Bovik , & Charrier, 2010), BLIINDS-II (Saad, Bovik, & Charrier, 2011), face, eyes, nose location, and image histogram as presented in Section C.

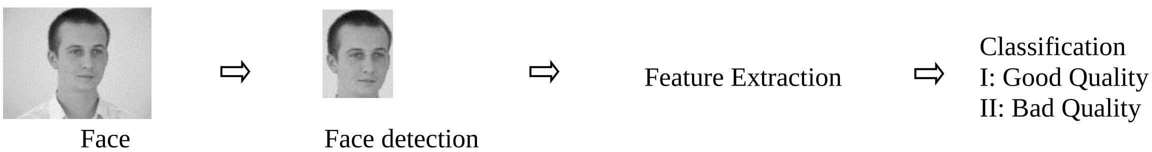

Fig.1: The general scheme of the proposed quality assessment

3.1. Approach to Pose Variations

The pose problem that is addressed in this paper is when the face is rotated to the right or the left as shown in Fig. 2. A posed face appears differently due to changes in viewing conditions, *i.e.* there exists a head rotation. Post-invariance assessment is crucial to a face biometric system because in general, it is difficult to control the face position when acquiring images of human faces. To improve the performance of biometric systems, pose variation is one of the important problems that must be taken into consideration when dealing with quality assessment. Fig. 2 presents sample images from the ENSIB database (Hemery, Rosenberger, & Laurent, 2007).

Fig.2: Example of samples of ENSIB database

3.2. Face, Eye and Nose Detection

Face detection is vital in various applications for computer vision. Face and eye detection were used to minimize the number of keypoints in an image that is not related to the face or eye. While nose detection is used to localize the nose in the detected face.

In this section, the Vision Cascade Object Detector (Viola, Paul, & Jones, 2001) is applied to identify the face, eye and nose location in an image. The Cascade Object Detector employs the Viola-Jones detection algorithm and a trained classification model for detection. The detector is expected to detect the whole face or the whole eye.

3.3. Quality Features

A set of features is extracted to detect the pose variation problem in facial images. These features are listed as follows:

- Features selected from BLIINDS-I (Saad, Bovik , & Charrier, 2010) and BLIINDS-II (Saad, Bovik, & Charrier, 2011).
- Features extracted from the face using the SIFT descriptor.
- Features extracted from the eyes using the SIFT descriptor.
- Features extracted from the nose location in the detected face
- Features derived from the Histogram

3.3.1. Features of BLIINDS-I & BLIINDS-II

Features from BLIINDS-I were tested separately and those that can detect the pose variation problem in facial images were selected. Similarly, for features from BLIINDS-II. As a conclusion, two features from BLIINDS-I, and four features from BLIINDS-II were selected. These features are as follows:

- Features of BLIINDS-I (Saad, Bovik , & Charrier, 2010)
  - DCT – based contrast feature ($\upsilon 1$)
  - DCT – based structure feature ($\upsilon 2$)
- Features of BLIINDS-II (Saad, Bovik, & Charrier, 2011)
  - The Generalized Gaussian Model Shape Parameter
  - The Coefficient of Frequency Variation
  - Orientation Model-Based Feature
  - Energy Sub-band Ratio Measure

3.3.2. Face features extraction

Face features are extracted by using Scale-invariant feature transform (SIFT) (Lowe, 2004), which is a feature detecting algorithm in computer vision that can detect and describe local features in images, and then calculate the association between the two images.

Face Features were extracted by following the steps below and as summarized in Fig.3:

- Detecting the face from an image by utilizing the detection method.
- Cropping the detected face.
- Dividing the cropped image into two blocks (Left & Right).
- Mirroring one of the blocks horizontally (in our experiment we mirrored the right block).
- Applying SIFT and calculating the number of match keypoints in each block, the number of associations of these match keypoints, and another set of features.
- Selecting features using the ranker.
- Using classifiers for feature evaluation.

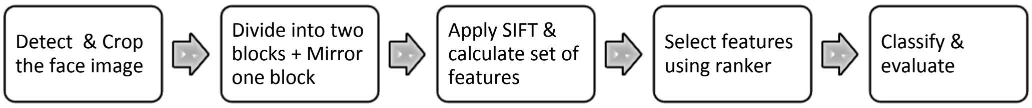

Fig.3: Face association features approach

In Fig.4, the number of association indicate common features between two different images and horizontal mirroring to one of the sides is done for better detection of common features. The resulted number of associations between the two blocks is considered one of the features to be used in classification.

Number of association is 11   Number of association is 3

Fig.4 : Examples of associations between the two blocks of a face image (the image on the left is a frontal image with association 11, while the image on the right is an image with pose and association 3)

The following features are considered to contribute to the quality assessment:

– The number of match key points and association of match key points between the two blocks
– The average of the key points detected in the two blocks.

3.3.3. Eyes features extraction

These features are extracted by using (SIFT) (Lowe, 2004) and then calculating the association between the two images.

Eye Features were extracted by following the steps below and as summarized in Fig.5:

– Detecting the face from an image by utilizing the detection method.
– Cropping the detected face.
– Dividing the cropped image into four blocks (Left & Right, Up & Down).
– Detecting the eyes from the upper blocks by utilizing the detection method.
– Cropping the detected eyes blocks.

– Mirroring one of the blocks horizontally (in our experiment we mirrored the right block).
– Applying SIFT and calculating the number of match keypoints in each block, the number of associations of these match keypoints, and another set of features.
– Selecting features using the ranker.
– Using classifiers for feature evaluation.

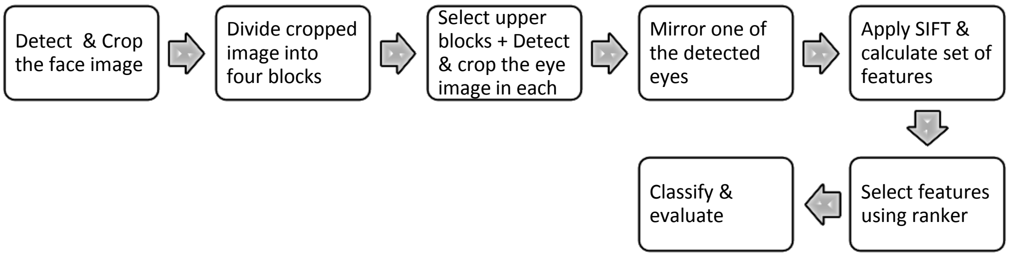

Fig.5: Eye association features approach

The following features are considered to contribute to the quality assessment:
– The number of match keypoints and association of match keypoints between the two blocks.
– The average, median and standard deviation of scales related to the keypoints detected in the four blocks.
– The average of the scales related to the keypoints in the eyes blocks.
– The average of the scales related to the block sizes of the eyes.
– The absolute difference between the scales related to the block sizes of the eyes.
– The skewness of scales related to the keypoints detected in the four blocks.
– The kurtosis of scales related to the keypoints detected in the four blocks.

Fig.6 illustrates the association of match keypoints between the two eye blocks. The resulted number of associations between the two blocks is considered as a new feature to be used later in classification.

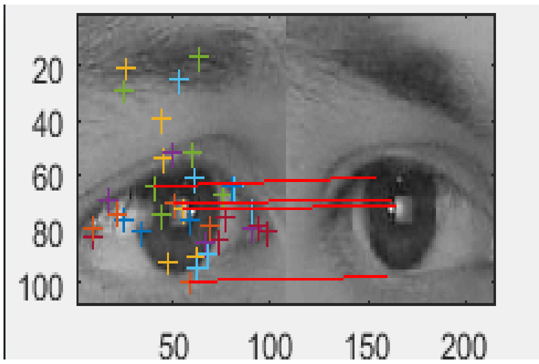

Fig.6 : Examples of associations between the two eyes blocks

Horizontal mirroring was made to one of the divided blocks, as mentioned above. Mirroring was needed to have a twin block that will result in an increase in the number of associations between the two-twin block as shown in Fig.6. If the face is frontal the number of associations is high since the blocks will be the mirror to each other. If pose exists, the two blocks will not look alike and therefore, the number of associations between the two blocks is low.

3.3.4. Nose location feature

The nose location feature is extracted by using the Vision Cascade Object Detector (Viola, Paul, & Jones, 2001), which employs the Viola-Jones detection algorithm to find the location of the nose in the copped detected face image.

The following attributed are used to obtain the nose feature: The size of the x-axis of the nose location nose(x), the horizontal size of the face block face(x), and finally, the feature was obtained by calculating the following (Face(x)/2-nose(x)).

This provides us with the distance between the detected center point of the nose and the midline of the image. In the ideal case, the distance should be equal to 0, any value other than 0 means that the face is rotated either to the right or to the left. As this distance increase, this means that the face is rotated at higher angles.

3.3.5. Histogram features

In this section, we present the extracted features from the first and second-order histograms. From first-order histogram, we calculated the Kurtosis, mean, median and the standard deviation.

Fig. 7 represents from left to right a detected frontal face and its left and right sides consecutively when divided into 2 blocks.

Frontal Face Left side Right side

Fig.7: Frontal face and its two sides

Fig. 8, Fig. 9, and Fig. 10 represent the histogram of the images presented in Fig. 7 consecutively. A histogram is a graphical representation (Ioannidis, 2003) of exposed pixels in the image, where black areas or shadows are represented on the left side, the highlights or the bright areas are represented on the right side, and mid-tones which are neither dark nor light are represented in the middle portion of the histogram. The high of the peaks reach denotes the number of pixels in that specific tone.

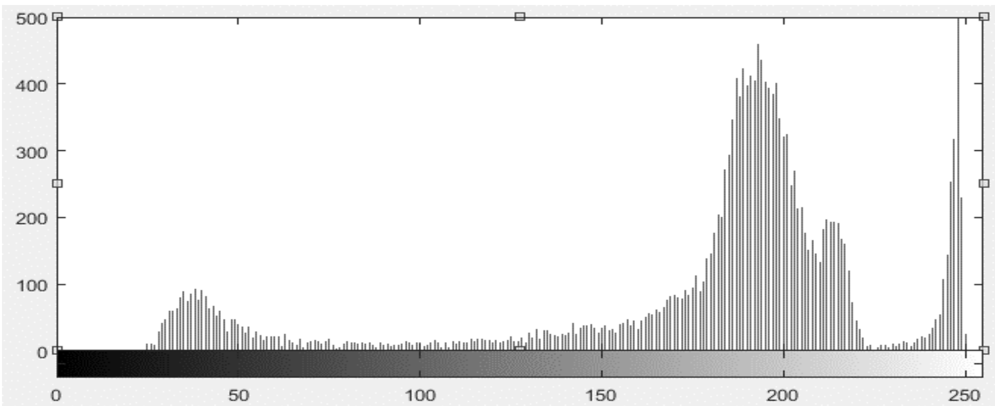

Fig.8: Histogram of Frontal face

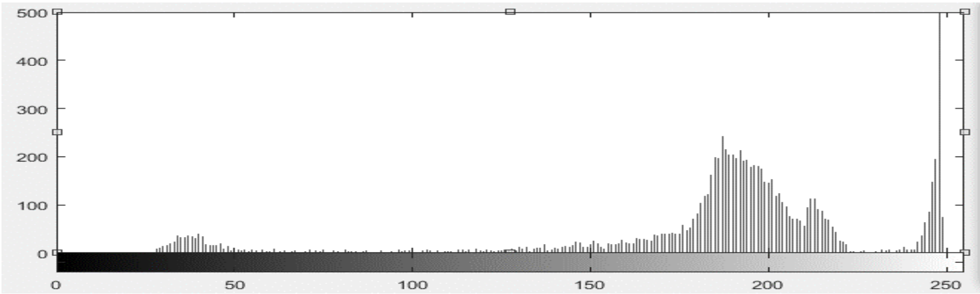

Fig.9: Histogram of the left side of frontal face

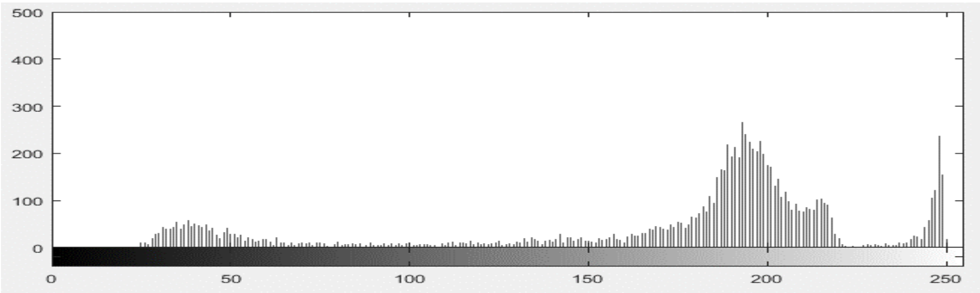

Fig.10: Histogram of the right side of frontal face

The two blocks of the right and the left side of the face are expected to have similar number of pixels in all tones, resulting in two similar histograms (equal number of pixels from 0-255) as in Fig. 9 and Fig. 10, and the difference between these two histograms show difference in the two images, indicating the existence of pose in the image as in Fig.13 and Fig. 14.

Fig. 11 represents from left to right a detected face with pose problem and its left and right sides consecutively when divided into 2 blocks.

Frontal Face Left side Right side

Fig.11: Posed face and its two sides

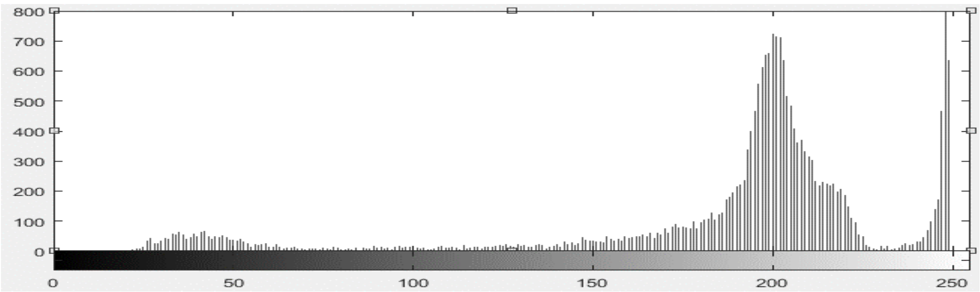

Fig.12: Histogram of Posed face

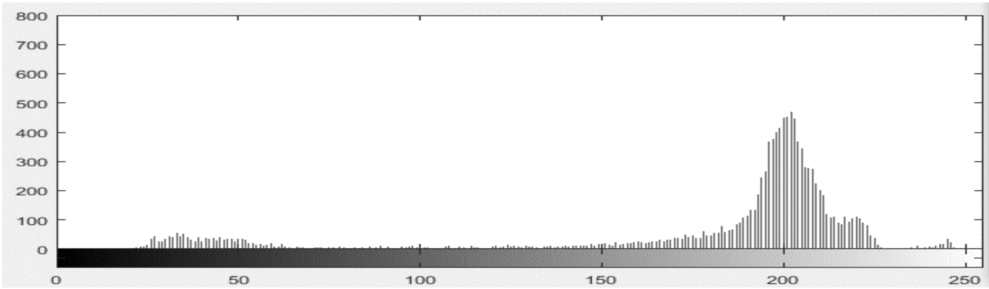

Fig.13: Histogram of the left side of Posed face

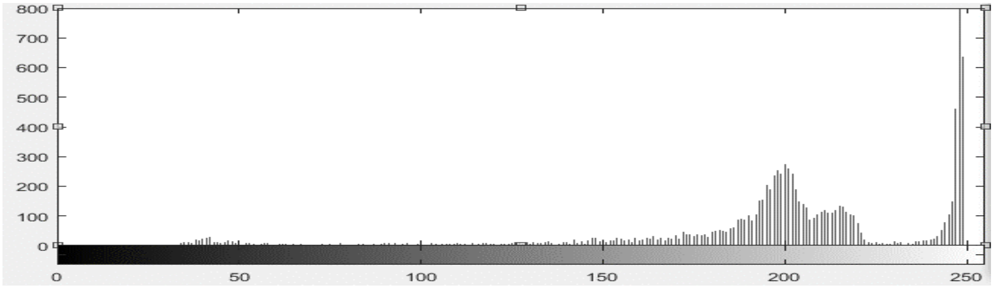

Fig.14: Histogram of the right side of Posed face

From the second-order histogram, the Gray-Level Co-Occurrence Matrix (GLCM) is used. It is formed from a gray-scale image and has various features that can be extracted from the probabilistic value p (i, j). It calculates the frequent occurrence of a pixel with the gray-level value i in adjacent to pixels with value j either horizontally, vertically, or diagonally. It proved to be a well known statistical method of extracting textural features from images (mathworks, n.d.), (Haralick, Shanmugan, & Dinstein, 1973). In this study, only four features are implemented using the graycoprops function in Matlab as presented in Fig. 15.

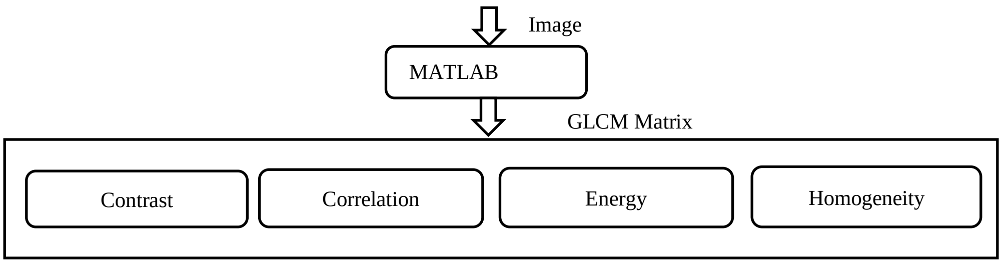

Fig.15: The four utilized features from GLCM

These statistics are presented below as follows:

➢ Contrast

This is also called "sum of squares variance". Contrast is a measure of local-level variations in the (GLCM) where images of high contrast take high values.

The returned measure is the intensity contrast between two adjacent pixels over the whole image and it ranges between 0 and (size(GLCM,1)-1)^2.

$\sum_{i,j=0}^{N-1} |i-j|^2 p(i,j)$ (1)

where p(i,j) is the probability value between 0 and 1. i and j are the gray level values in the image, such that i is the row number and j is the column number.

N is the number of rows or columns. The summation is from 0 to (N-1) since the first cell in the upper left of the GLCM is numbered (0,0), not (1,1).

- Correlation

Correlation is a feature that calculates the correlation between pixels in two different directions. It returns an amount indicating how correlated a pixel is to its neighbor over the whole image. This amount ranges between -1 and 1.

$\sum_{i,j=0}^{N-1} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$ (2)

where μ is the mean based on the reference pixels, and σ is the variance. p(i,j) is the probability value from the GLCM.

- Energy

Energy is a feature that calculates how smooth the image is. The less smooth the region is, the more uniformly distributed P(i, j) and the lower will be the value of this feature. It ranges between 0 and 1.

$\sum_{i,j=0}^{N-1} p(i,j)^2$ (3)

where p(i,j) is the probability value, i and j are the gray level values in the image, such that i is the row number and j is the column number. N is the number of rows or columns.

- Homogeneity

Homogeneity is a feature that provides information about the contrast of the image. It calculates the nearness of the distribution of elements in the GLCM to the GLCM diagonal. The returned measure ranges between 0 and 1, where 1 is for a diagonal GLCM.

$\sum_{i,j=0}^{N-1} \frac{p(i,j)}{1+|i-j|}$ (4)

where p(i,j) is the probability value, i and j are the gray level values in the image, such that i is the row number and j is the column number.

The above features are summarized in Table 1 and are used as input to our binary classifier. The time complexity of extracting these features is O(n2) where n is the width and the height of the image (Appiah & Acquah, 2018) (Lowe, 2004).

Table 1: Features Summery

Type of Feature	*Feature*
BLIIND	· DCTCF: DCT – based contrast feature (υ1) · DCTKURT: DCT – based structure feature (υ2) · DCTGGM: The Generalized Gaussian Model Shape Parameter: · DCTCFV: Coefficient of Frequency Variation · DCTOMF: Orientation Model-Based Feature · DCTESUB: Energy Sub-band Ratio Measure
Facial-based Features	· ASSOCF: Association of match keypoints between the two blocks · NUMF: Number of match keypoints between the two blocks · MEANTB: Average of scales related to the keypoints detected in the two blocks
Eye-based Features	· ASSOCE: Association of match keypoints between the two blocks · NUME: Number of match keypoints between the two blocks · MEANFB: Average of keypoints detected in the four blocks · MEDIANFB: Median of keypoints detected in the four blocks · STDFB: Standard deviation of keypoints detected in the four blocks · MEANEB: Average of keypoints in the eyes blocks. · MEANBS: Average of the scales related to the block sizes of the eyes. · ABSDIF: Absolute difference between the scales related to the block sizes of the eyes. · SKEWB: Skewness of scales related to the keypoints detected in the four blocks. · KURTB: Kurtosis of scales related to the keypoints detected in the four blocks.
Nose Location Feature	· NOSELOC = Difference between x-coordinates of the face divided by 2 and nose x-coordinates.
Histogram Feature	· KURTHIST: Kurtosis of the histogram of the image · MEANHIST: Average of the histogram of the image · MEDIANHIST: Median of the histogram of the image · STDHIST: Standard deviation of the histogram of the image · CONTHIST: Contrast features from the image · CORRHIST: Correlation features from the image · ENGYHIST: Energy features from the image · HOMOHIST: Homogeneity features from the image

3.4. Classifiers

In this study, different supervised learning algorithms are adopted from WEKA (Frank, Hall, & Witten, 2016) such as Logistic Regression, Naïve Bayes, Multilayer Perceptron, Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), etc. to build a prediction model for classifying images as of good or bad quality.

To improve the predictive performance, ensemble learning Random Forest (Patel, 2017), one of the ensembles learning algorithms in WEKA is used to compare results.

Random Forests is one of the most useful models. It creates random forests by bagging ensembles of random trees and uses averaging to improve the predictive accuracy and control over-fitting, what makes them so great is that it corrects the over cross-validation fitting of a single decision tree model by using Bagging. The collected dataset was tested for over-fitting by using 10-fold cross-validation and the results were as follows: Maximum accuracy was 100% and Minimum accuracy about 94%, with an average of 97%.

The Random Forest classifier showed the highest accuracy among the other classifiers using 10-fold cross-validation classification as shown in Section IV-B.

The binary classification is calculated by finding the average of the accuracy calculated in each of the ten folds as in equation (6):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

where,

- TP = True Positives (good quality image classified as good quality),
- TN = True Negatives (bad quality image classified as bad quality),
- FP = False Positives (bad quality image classified as good quality), and
- FN = False Negatives (good quality image classified as bad quality).

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the selected databases along with the adopted classifiers are presented. The experiment is implemented using a laptop with an Intel(R) Core™ i7-3630QM CPU operating @ 2.10GHz.

4.1. Database

Many publicly available facial databases exist but very few are the databases that have sufficient image quality variability to meet the objectives of this research work. Due to the lack of public available benchmark having the pose problem, we have used subsets from three databases: PUT (Kasiński, Florek, & Schmidt, 2008), AR (Martinez & Benavente, 1998), and ENSIB (Hemery, Rosenberger, & Laurent, 2007) databases were used for experiments. Selection of images from these databases was combined into a dataset containing a set of original images with frontal face images, and another set of images with posed faces.

The set of original images are considered as good quality images, and the set of images with the posed faces are considered as bad quality images.

The AR database constitutes of 120 individuals and 26 samples per individual. These images are captured in changed situations of illumination, expression, and occlusion. Each image is 768 x 576 pixels in size.

The PUT database is of color, high-resolution face images. It contains images of 100 individuals and 22 images per person. Images were taken in partially controlled conditions while people were moving their heads without any constraint to the pose or expression. Each image is 2048 × 1536 pixels in size.

ENSIB database is composed of 100 individuals with 40 different views. Images were acquired for each individual using a webcam to record a video of 401 x 401 pixels. Individuals were asked to turn their heads from the left profile to the right profile in two seconds.

We used a subset of PUT, AR, and ENSIB databases. The selection was made subjectively by selecting images for each subject from each database where one is considered as the original image and a range from 1 to 5 images for each subject considered as posed image. Fig. 16 presents a sample of selected images from each database.

Fig.16: Three sample images from PUT, AR, and ENSIB databases from left to right consecutively.

The resolution of the images in ENSIB Database is 401x401 pixels, 2048x1536 the resolution of the images in the PUT database, 768x576 the resolution of the AR database. Images from different databases were resized to 400x400 pixels before running the face detection algorithm to ensure the same resolution for all the images. Then, the Viola-Jones detection algorithm was run on the selection and all images, the images were no detection of a face, eye or nose were removed. These removed images were considered as bad quality images and are not suitable and therefore, they were discarded from our dataset. Finally, our new dataset of original and posed images is constructed as follows: The portion of the selected good quality images (original images) consists of 87 images from the PUT database, 207 from the AR database, and 51 images from the ENSIB database, reaching a total of 345 original images. As for the pose images, a total of 438 images is collected by selecting a portion of 330 images from the PUT database, and 108 from the ENSIB database. The AR database does not contain images with pose problems but it is used to increase the number of original images to have a sufficient size of images with good quality. Samples from the PUT database are presented in Fig. 17.

Fig.17: Sample images from the PUT database
(Left: good quality and Right: bad quality)

4.2. Classification Results

Different learning classifiers were adopted on the new datasets using WEKA with 10-fold-cross-validation accuracy. Table 2 presents the comparison of accuracy between different classifiers.

Table 2: Comparing Accuracy Using Different Classifiers

Classifier	10-fold cross-validation accuracy (%)
Logistic	**96.2963**
Simple Logistic	95.2746
Multilayer Perception	95.1469
RandomCommittee	**96.424**
RandomTree	91.4432
NaiveBayes	89.9106
Decision Stump	84.802
SGD	95.0192
LMT	95.4023
SMO	94.636
LWL	90.2937
RandomForest	**96.6794**

To improve accuracy, the Ensemble model was used and compared with the previous classifiers. RandomForest is one of the Ensemble models that is presented in WEKA, it performs a 96.6794% accuracy which is higher accuracy than the above-mentioned classifiers.

All the runs were completed using 28 features. To find the positive or negative impact of these features on the model, we used a filter attribute selection method (InfoGainAttributeEval available in WEKA) to rank these features according to their impact. InfoGainAttributeEval is a popular feature selection technique to calculate the information gain. The entry values of the information gain will range from 0 (no information) to 1 (maximum information). Higher information gain means that the attribute contribute more information and can be selected, whereas those with low information can be removed (Brownlee, 2016).

The ranking of these features is shown in Fig. 18 where the top features should have the higher impact on our proposed method. It confirms that the nose feature has the highest impact on the classification, as there is an ideal case when the distance between the detected center point and the midline of the image is 0. This feature certainly improved the accuracy of the models used.

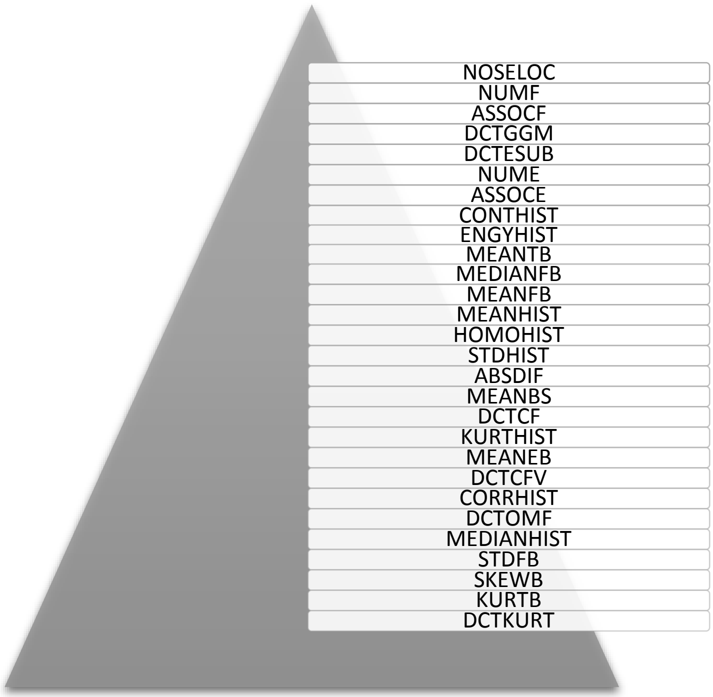

Fig.18: Ranking of Features

The num and association features of the face and the eyes have also a high impact on the method after the nose location feature. These features are the result of horizontally mirroring the image, providing a twin block that resulted in increasing the number of associations between the two-twin block since if pose exists, the two blocks will not look alike and therefore, the number of associations between the two blocks is low.

We excluded four irrelevant features which are ranked as the last features. This improved the accuracy of the RandomForest ensemble learning reaching an accuracy of 97.06% with 10-fold cross-validation.

In order to justify that our dataset size is sufficient, the concept of the learning curve was applied to monitor its performance. A learning curve is a concept that graphically depicts the relationship between the database size and the accuracy of the classifiers over a defined period.

Fig.19 shows the learning curve of several classifier when the dataset is divided into a proportional number of original and pose images. As shown, the curve is improving in performance reaching almost saturation indicating the point at which incrementing the data size will not be beneficial.

We calculated the accuracy by increasing the dataset size proportionally. Starting with the size of 50/63, where 50 indicates the size of the original images in the dataset and 63 is the size of the posed images in the dataset and ending with the full dataset 354/438. 63 is proportion to 50 with respect to the data size of the two databases.

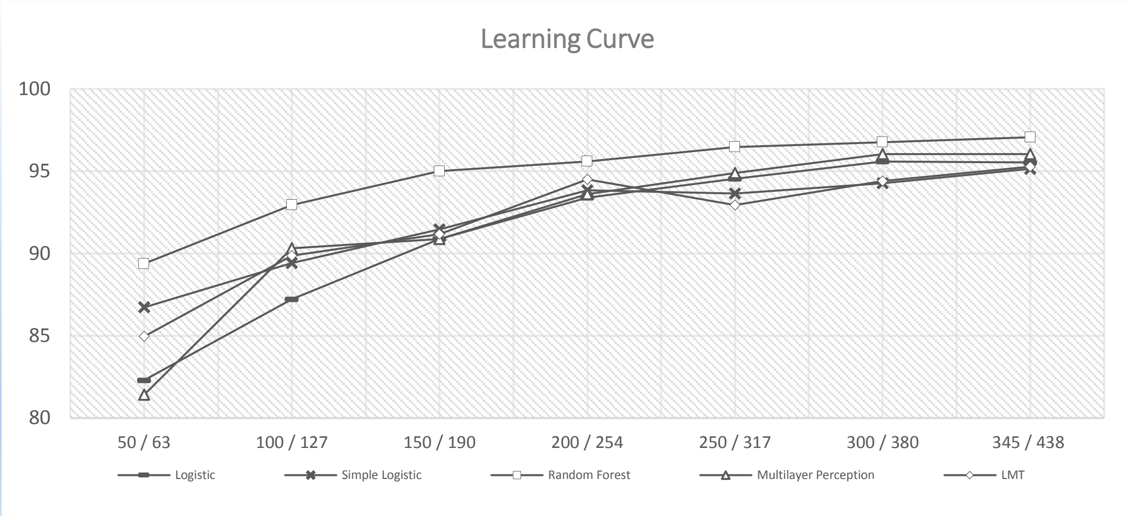

Fig.19: Comparison between different classifiers with the proportion number of datasets

5. Conclusion and Future Scope

Excessive efforts are recommended to develop no-reference quality assessments that can detect the pose problems during enrollment, which will improve the performance of facial recognition systems.

In this paper, a novel no-reference facial quality assessment is proposed that detects the pose problem. The method presents additional features to existing work towards detecting pose variation problems in facial images. Furthermore, this paper showed that some of the features used from the literature that aimed for detecting other types of alterations (such as blurring alteration) can detect as well the pose alteration problem. Several classifiers were tested and compared. The RandomForest classifier outperformed other classifiers and reached an accuracy rate of 97.06%.

Our future work consists of using quality assessment for detecting fake face images, and testing deep learning approaches using Convolution Neural Network to compare the obtained results with the results presented in this paper.

References

- Frank, E., Hall, M. A., & Witten, I. H. (2016). *The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques* (Fourth ed.). Morgan Kaufmann.
- Abayomi-Alli, A., Omidiora, E., Olabiyisi, S., & Ojo, J. (2015). Adaptive Regression Splines Models for Predicting Facial Image Verification and Quality Assessment Scores. *Balkan Journal of Electrical & Computer Engineering, III*.
- Appiah, O., & Hayfron-Acquah, J. B. (2018). Fast Generation of Image's Histogram Using Approximation Technique for Image Processing Algorithms. *International Journal of Image, Graphics and Signal Processing*, *10*(3), 25.
- Bharadwaj, S., Vatsa, M., & Singh, R. (2014). Biometric quality: a review of fingerprint, iris, and face. *EURASIP journal on Image and Video Processing*, *2014*(1), 34.
- Brownlee, J. (2016, July). *How to Perform Feature Selection With Machine Learning Data in Weka*. Retrieved from Machine Learning Mastery: www. machinelearningmastery.com
- Charrier, C., Saadane, A., & Fernandez-Maloigne, C. (Sep 2017). No-Reference Learningbased and Human Visual-based Image Quality Assessment Metric. *19th International Conference on Image Analysis and Processing.* Catania, Italy.
- El-Abed, M., Charrier, C., & Rosenberger, C. (2015). Quality assessment of image-based biometric information. *EURASIP Journal on Image and Video Processing*, *2015*(1), 3.

- El-Abed, M., Ninassi, A., Charrier, C., & Rosenberger, C. (2013). Fingerprint Quality Assessment using a No-Reference Image Quality Metric. *IEEE 21st European Signal Processing Conference (EUSIPCO).* Marrakech, Morocco.
- Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, And Cybernetics*, (6), 610-621.
- Hemery, B., Rosenberger, C., & Laurent, H. (2007). The ENSIB database: a benchmark for face recognition. In *International Symposium on Signal Processing and its Applications (ISSPA), special session "Performance Evaluation and Benchmarking of Image and Video Processing* (pp. 37-40).
- Hou, W., Gao, X., Tao, D., & Li, X. (2014). Blind image quality assessment via deep learning. *IEEE transactions On Neural Networks And Learning Systems*, *26*(6), 1275-1286.
- Ioannidis, Y. (2003, January). The history of histograms (abridged). In *Proceedings 2003 VLDB Conference* (pp. 19-30). Morgan Kaufmann.
- Jain, A., Pankanti, S., Prabhakar, S., Hong, L., Ross, A., & Wayman, J. (2004). Biometrics: A Grand Challenge. *International Conference on Pattern Recognition,.* Cambridge, UK.
- Kasiński, A., Florek, A., & Schmidt, A. (2008). The PUT Face Database. Retrieved from www.degruyter.com
- Kerouh, F., Ziou, D., & Serir, A. (2018). Histogram modelling-based no reference blur quality measure. *Signal Processing: Image Communication*, *60*, 22-28.
- Khodabakhsh, A., Pedersen, M., & Busch, C. (2019, May). Subjective Versus Objective Face Image Quality Evaluation For Face Recognition. In *Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications* (pp. 36-42).
- Liao, S. H., Chu, P. H., & Hsiao, P. Y. (2012). Data mining techniques and applications–A decade review from 2000 to 2011. *Expert systems with applications*, *39*(12), 11303-11311.
- Liu, X., Pedersen, M., Charrier, C. M., & Bours, P. (2018). Performance evaluation of no-reference image quality metrics for face biometric images. *Journal of Electronic Imaging*, *27*(2), 023001.
- Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal Of Computer Vision, 60*(2), 91-110.
- Martinez, A. R., & Benavente, R. (1998). *The AR face database*. Barcelona: ComputerVision Center(CVC).
- *Mathworks*. (n.d.). Graycoprops. Retrieved from www.mathworks.com
- Mittal, A., Soundararajan, R., & Bovik, A. C. (2012). Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*, *20*(3), 209-212.
- Nasrollahi, K., & Moeslund, T. B. (2008). Face Quality Assessment System in Video Sequences, Biometrics and Identity Management: First European Workshop, BIOID 2008, Roskilde, Denmark, May 7-9, 2008. Revised Selected Papers.
- Nikitin, M., Konushin, A., & Konushin, V. (2014). Face Quality Assessment for Face Verification in Video. *Proceedings of GraphiCon'2014*, (pp. 111–114).
- Patel, S. (2017). Random Forest Classifier. In *Machine Learning 101.* Retrieved from www. medium.com
- Ratyal, N., Taj, I. A., Sajid, M., Mahmood, A., Razzaq, S., Dar, S. H., . . . Mussadiq, U. (2019). Deeply Learned Pose Invariant Image Analysis with Applications in 3D Face Recognition. *Mathematical Problems in Engineering*.
- Saad, M. A., Bovik, A. C., & Charrier, C. (2011, September). DCT statistics model-based blind image quality assessment. In *2011 18th IEEE International Conference on Image Processing* (pp. 3093-3096). IEEE.
- Saad, M., Bovik , A., & Charrier, C. (2010). A DCT statistics-based blind image quality index. *IEEE Signal Processing Letters, 17*(6), 583-586.
- Sang, G., Li, J., & Zhao, Q. (2016). Pose-Invariant Face Recognition via RGB-D Images. *Computational Intelligence and Neuroscience*.
- Solomon, C., & Breckon, T. (2011). *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. John Wiley & Sons.

- Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer Society Conference On Computer Vision And Pattern Recognition. CVPR 2001* (Vol. 1, pp. I-I). IEEE.
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions On Image Processing*, *13*(4), 600-612.
- Wong, Y., Chen, S., Mau, S., Sanderson, C., & Lovell, B. C. (2011, June). Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition. In *CVPR 2011 WORKSHOPS* (pp. 74-81). IEEE.
- Zhang, G., & Wang, Y. (2009, November). Asymmetry-based quality assessment of face images. In International Symposium on Visual Computing (pp. 499-508). Springer, Berlin, Heidelberg.