



Novel Descriptors for Effective Recognition of Face and Facial Expressions

Farid Ayeche¹, Adel Alti^{2,3*}

¹ LMETR Laboratory, E1764200 Optics and Precision Mechanics Institute, Ferhat Abbas University Setif 1, Setif 19000, Algeria

² Department of Management Information Systems, College of Business & Economics Qassim University, P.O. Box 6633, Buraidah 51452, Saudi Arabia

³ LRSD Laboratory Computer Science Department, University of SETIF-1, Setif 19000, Algeria

Corresponding Author Email: a.alti@qu.edu.sa

<https://doi.org/10.18280/ria.340501>

ABSTRACT

Received: 5 August 2020

Accepted: 12 October 2020

Keywords:

directional gradient descriptor, texture feature analysis, SVM classifier, face recognition

In this paper, we present a face recognition approach based on extended Histogram Oriented Gradient (HOG) descriptors to extract the facial expressions features allowing classifying the faces and facial expressions. The approach is based on determining the different directional codes on the face image based on edge response values to define the feature vector from the face image. Its size is reduced to improve the performance of the SVM (Support Vector Machine) classifier. Experiments are conducted using two public datasets: JAFFE for facial expression recognition and YALE for face recognition. Experimental results show that the proposed descriptor achieves recognition rate of 92.12% and execution time ranging from 0.4s to 0.7s in all evaluated databases compared with existing works. Experiments demonstrate and confirm both the effectiveness and the efficiency of the proposed descriptor.

1. INTRODUCTION

Face recognition is becoming an attractive research field for the development of Human-Machine Interaction (HMI). There is a rapid evolution of its scientific and socio-economic challenges [1]. In particular, machines have allowed identifying facial features through Facial Recognition (FR) and Facial Expression Recognition (FER) with an accurate identification rate and deep understanding of the dynamic behavior of speech and facial expressions. The recognition of facial expression is largely used in various social domains, describing emotional identification of expressions, human activity, such as fatigue eating, etc. [2]. Facial recognition represents a great instrument to track feature of a person [3]. Thus, one major goal of the automatic facial recognition is to help verify people identify to highlight their verbal message in “face to face” as well as their spoked facial expression [4]. Additionally, to persons with special needs, analyzing and modeling facial expressions comes with a multitude of benefits for all social domains. It helps to reduce learning costs and enhance user’s safety [4-7].

The FR and FER systems are mainly classified into two categories: geometric feature-based approaches and appearance-based approaches. Geometric feature-based approaches [8] use the facial location and distance features extraction (eyes, nose, and mouth). The major problem with these approaches is the incorrect and inaccurate location of different features extraction, with wrong matching results. The appearance-based approaches [9] utilize various features extraction methods such as color differences, texture gradient direction, Gabor features, presenting as advantages of good effectiveness against the environmental change and lower computational complexity. However, this approach was a leak

from the good features extraction of facial images and poor recognition rates.

Many recognition systems were the results of existing technologies. How to use it and how to combine it can lead to novel sophisticated system. In this work, the tradeoff between accuracy and computational complexity problem was addressed by proposing new descriptors HDG (Histogram of Directional Gradient) and HDGG (Histogram of Directional Gradient Generalized). Indeed, we think that the following points constitute the major contributions of the current work:

- Two new descriptors for extracting more discriminative features by using magnitude and orientation maps, and eight directional edge responses values between the same region of facial images,
- Providing a new feature vector reduction (i.e. size of 8 for each regions of image) for fast face recognition.
- Computing the feature vector by a new equation based on gradient magnitude and orientation.
- To improve the generalization performance of the model, 11 standard descriptors among the most common algorithms train-tests validation process has been employed.
- Implementing and re-testing all standard and proposed descriptors for facial expressions recognition using same datasets on the same platform.
- A detailed qualitative and quantitative analysis for the performance of the proposed descriptors have carried out using standard AFFE [10] and YALE [11] datasets which come with a variety of images in multiple databases through SVM (Support Vector Machine) classification.

Section 2 presents related works of feature descriptions.

Section 3 presents existing descriptors. Section 4 describes the proposed descriptors for feature extractions. Section 5 presents the proposed framework. Section 6 exposes details of the implementation and experimental result as well as discussions. Section 7 concludes the paper.

2. RELATED WORK

Several descriptors are proposed for facial recognition systems. Ahonen et al. [12] use a well-known Local Binary Pattern (LBP). In the literature, LBP is considered one of the most powerful discriminative descriptors that was applied to diverse classification problems including face recognition. In fact, LBP is known to be particularly efficient to use with several other descriptors. However, the problem with LBP is the fact of being feature-based classifiers i.e. high-dimensional feature vector leads to high cost. The obtained result is affected by rotation changes, which is confounded with the classification process.

Additionally, high computational complexity required in training phase may in some cases limit their applicability. Coupled with the Gabor wavelet for feature extraction, the histograms of Local Gabor Binary Pattern (LGBPHS) [8] and Gabor Phase Pattern (GPP) [8], LBP become the best discrimination technique for face recognition. Local Phase Quantization (LPQ) [13] is another well-known feature-based technique for face identification. It builds on the blur invariance property of the Fourier phase spectrum. This technique ends up with more accurate and stable classification results. However, one limitation of LPQ is the slowness of the model in some cases especially with a high number of features.

Chen et al. [14] use Weber Local Descriptor (WLD), which is inspired from Weber's law. Other discriminative techniques were also explored in facial recognition systems such as Local Gradient Pattern (LGP) [15] and Gradient Direction Pattern (GDP) [15]. Jabid et al. [16] propose Local Directional Pattern (LDP). This work suffers from the self-imposed restriction of always having a leading 1, which immediately reduces the number of available combinations in the binary word by half. The problem of fixed number of 0 and 1 continue from the original LDP, depending on threshold value k .

Rivera et al. [17] uses Local directional number (LDN) pattern based in compass mask to represent the image into different directions, and then encodes it using the directional sign and numbers. Tong et al. [18] introduced a new local gradient Code (LGC), which was a variant on the LBP results by describing the gradient of the horizontal, vertical, and diagonal detail information of the facial image. Independent Component Analysis (ICA) [19], Eigenfaces using Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [20-22], Histogram of Orientation Gradient (HOG) [23] and Wavelets [24] are also widely used for FR and FER systems. The different forms of Eigenfaces are used as a base for other face recognition techniques. In general, the Eigenfaces find the similarities between faces with minimal controlled environments. However, PCA suffers from a low recognition rate. PCA is much more reduction of data dimension than other existing recognition methods.

Several classification techniques span over a broad range of machine learning techniques from classic classifiers such as Artificial Neural Networks (ANNs) and Hidden Markov Models (HMMs) [25], to more advanced approaches such as Support Vectors Machines (SVMs) [4, 5]. These methods aim

to develop and evaluate the performance of a statistical classifier based on a new generation of neural networks using pattern code faces. However, SVMs are the most used approach for FR and FER systems. Recently deep learning with deep networks is being widely used for FER. It integrates both feature extraction and learning phases [25]. In this context, SVM classifier has been used to improve classification performance [26], elucidating issues related to overfitting and local minima that occur with more conventional neural network algorithms. These characteristics are important in pattern recognition applications such as human face recognition. More recently, Ayeche et al. [27] evaluate the performance of machine learning for recognizing facial emotions in terms of precision and execution time.

In our work, we propose a new face recognition operator based on the gradient directions. The approach is based on determining the different directional codes on facial images based on edge response values allowing improving the classification accuracy rate for FR and FER systems. Therefore, we validate and compare it with several existing descriptors designed for face recognition and conducted several experiments on the widely used benchmarks. Our results are produced very quickly and automatically based on face parameters. Due to the low time cost, the proposed algorithm can also be applied to smart digital camera.

3. STANDARD DESCRIPTORS FOR FEATURES EXTRACTION

Several standard descriptors are used for facial recognition in the literature. The most common descriptors are:

- Local Binary Pattern (LBP) [2] is a simple effective descriptor. The general principle of this descriptor is to compare the gray level of a pixel with the gray levels of its neighbors. It proposes to assign a binary code to a pixel according to its neighborhood. This binary code somehow describes the local texture of a region.
- Local Phase Quantization (LPQ) [13] is a texture descriptor based on the blur invariance property of Fourier phase spectrum. It is calculated using the short-term Fourier transform (STFT) within the neighborhood. LPQ is invariant to centrally symmetric image blurring.
- Weber Local Descriptor (WLD) [14] is a well-known descriptor based on Weber's Law to stimulus changes depends on the original intensity of the original image. WLD encodes the difference between the intensity of center pixel and its neighbors within a local differential excitation and orientation.
- Local Directional Pattern (LDP) [16] is an enhancement of LBP descriptor, which uses the directional responses by using Kirsch compass kernels. It offers more resistance against noise and illumination variations. Hence, LDP represents a robust image descriptor for face recognition.
- Local Directional Number Pattern (LDN) [17] is a micro-pattern descriptor computed from the edge response of the neighborhood using compass mask. It is a six-bit binary code for each pixel of the input image obtained from the intensity transitions along the boundary of the texture. LDN considers the top directional numbers, which is the most positive and negative directions of those edge responses.

- Local Gradient Code (LGC) [18] Comparing to the LBP descriptor, the LGC descriptor describes the distribution of the gray levels in the neighborhood of the center pixel. However, it uses the horizontal, vertical, and diagonal gradients instead of only the central pixel value.
- Histogram of Oriented Gradient (HOG) [23] is a successful feature descriptor based on the gradient orientation that has been used widely used for face recognition. HOG divides the face image into a set of blocks; each block was represented by a histogram describing the local distribution of orientations and amplitudes. The extracted feature vector was created by concatenating all normalized histograms to prepare a face image data input to different classifiers models, an illumination normalization is applied. HOG is often used with SVM classifier. Each HOG descriptor that is computed is fed to SVM classifier to identify the face. HOG is efficient under difficult lighting conditions.
- Gradient Direction Pattern (GDP) [28] is a more invariant feature description to noise while using edge response value instead of the intensity of pixel. It considers only uniform pattern, which cuts the feature vector length by half. Therefore, GDP codes provide cost-effective significant features to represent facial images.
- Local Ternary Pattern (LTP) [29, 30] is a potential enhancement descriptor of the original LBP. LTP extends LBP code by using three values of encoding to provide uniform and near-uniform consistency regions.
- Gradient Local Ternary Pattern (GLTP) [31] is a content-based pattern that profits from the advantages of Sobel-LBP and LTP operators. GLTP uses more robust gradient magnitude values instead of gray levels with a three-level encoding scheme to discriminate between smooth and high textured facial regions. GLTP is a robust texture pattern against noise and illumination variations.
- Local Gradient Patterns (LGP) [32] is computed based on local gradient flow from one side to another side through the center pixel in a 3x3 pixels region. The center pixel of that region is represented by two bits binary patterns, named as Local Gradient Pattern for the pixel.

4. PROPOSED DESCRIPTORS

The direction is one of the most powerful factors for its high performance to capture texture contained into a face image. Different descriptors for face image recognition have been used to investigate the reduction of information contained within an image, HOD is considered as one of the most promising descriptors among those based on the representation of an image by gradient transforms [23]. HOG (Histogram of Oriented Directional) and HDGG (Histogram of Directional Gradient Generalized) present two extensions of features edge response value and magnitude and orientation maps. The edge response value allows using different directional codes to improve classification accuracy. The magnitude and orientation maps values on the horizontal and vertical coordinates allow capturing more efficiently a set of oriented directional gradient features. They are based on gradient directions. HDG and HDGG encode the directional

information of face's texture in a reduced way for producing a more discriminative code than existing methods. While the classification is performed using SVM classifier through the multi-class implementation. The first descriptor allows the definition of a new notion, which describes boundaries for an object. The second determines a new attribute that can be specified to describe long-distance relationships. Both descriptors HDG and HDGG use unique feature vector and reduce its size, which enhances the classification accuracy and execution time of FR and FER systems. They provide faster, potentially more stable computation and express more clearly object boundaries in a long-distance relationship. In the next session, we present in detail the proposed descriptors and then we detail all steps of face classification and recognition.

4.1 HDG for face and facial expressions recognition

Due to high computation cost of HOG because of its repeated and complex computation during the feature extraction phase or recognition phase, we propose a directional factor based on the fusion of eight different gradient values of HOG for each pixel of a face image. The Histogram of Directional Gradient (HDG) is a new efficient operator based on the gradient direction. The features extracted are the distribution of directions of oriented gradients of the image. Gradients are typically micro-pattern structures extracted from face image to clearly express object boundaries by considering the directional information. We encode such information using the prominent direction indices and sign it. This allows distinguishing among similar structural patterns that have different gradient transitions. We also include histogram reduction algorithms to enhance the execution time.

We compute the edge response value using Kirsch masks, as shown in Figure 1, in different directions by eight different gradients for each pixel of an image. Next, we subdivide the image into $n \times m$ blocks. Each block will be represented by a histogram of eight values. Each value is a cumulative sum of information in each direction. Finally, the HDG descriptor is a concatenation of all histograms. Figure 2 shows the original image and the HDG results after applying Kirsh.

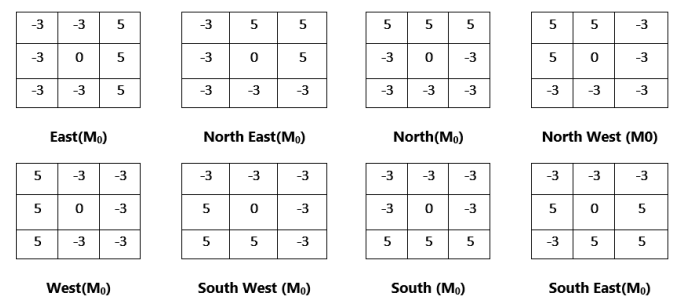


Figure 1. Example of edge response value

After extracting the features, we represent the intensity I_c at pixel x_c, y_c of the face image using the eight directional masks of Kirsch $M_i, M_0, M_1, \dots, M_7$ in this pixel, we obtain eight edge response values in a respective direction denoted by $m_i, i=0, 1, \dots, 7$, the eight edge response values are used to represent each pixel. As shown in Figure 3, we compute the sum of all edge response values bit to each block X in eight directions independently (one by direction) as follows:

$$b_i = \sum_{m_i \in X} |m_i| \quad i = 0, 1, \dots, 7 \quad (1)$$

where,

i : represents a direction,

X : a block of the image and

m_i : the response value of pixel for direction i .

Each part of the image is represented by a histogram $\langle B \rangle$.

$$B = \{b_i\} i = 0, 1, \dots, 7 \quad (2)$$

All histograms will be concatenated to form a feature vector of size $n \times m \times 8$. This vector will use it as a face descriptor.

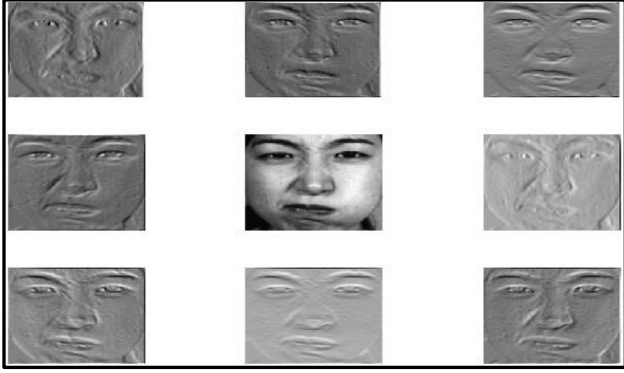


Figure 2. Original image and filtered images resulted from Kirsh masks

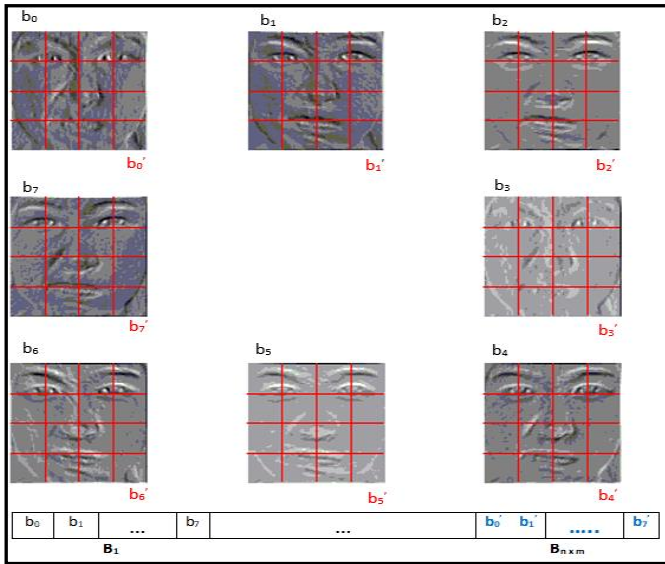


Figure 3. Building features vector after applying HDG operator

4.2 HDGG for face and facial expressions recognition

The HDGG extends the HOD feature extraction approach to magnitudes and orientation maps that requiring proper descriptive vector. HDGG consists of summing all gradient values of image pixels referring to 8-pixels using Kirsch filter, which will be mapped on magnitudes and orientation maps. HDGG consists of the following steps:

Step 1. Apply Kirsch filter on each block as HOD then compute for each pixel, 8-gradient feature. HDGG considers these eight values as eight oriented vectors (see Figure 4).

Step 2. Compute the sum of gradient vectors on a new pixel's vector, as shown in Figure 5, by using the following Eqns. (3) and (4):

$$x = \sum_{i=0}^7 (m_i \cos(i * \pi/4)) \quad (3)$$

$$y = \sum_{i=0}^7 (m_i \sin(i * \pi/4)) \quad (4)$$

Step 3. Perform the magnitude and orientation maps values on the horizontal and vertical coordinates of each pixel's vector according to G and θ values

$$G = \sqrt{x^2 + y^2} \quad (5)$$

$$\theta = \tan^{-1}(y/x) \quad (6)$$

Step 4. Decompose the whole image into $n \times m$ non-overlapping blocks, then quantizes orientation values of each block in the histogram with 9-orientation bins, where the magnitude values are the votes.

Step 5. Normalize all of the histogram blocks to obtain the feature vector. For each face image in the training set, we have calculated and stored the associated feature vector.

To extract facial features vector in facial image, we divide it into $n \times m$ blocks. We use 8-equally spaced intervals in the interval $[0, \pi]$. The 8-equally spaced intervals are the most significant intervals of the face image since it consists on reflecting regions with different discriminative information contrarily to other descriptors in which regions with small degrees are considered. For each block, a local histogram is automatically generated and normalized. These normalized histograms are concatenated to form the image's global histogram, which may be used to comparison of face recognition method. Figure 5 shows an example of HDGG magnitude and HDGG orientation results.

Note that the normalization of the image histogram orientation and magnitude is uniform and the environmental impact is reduced, after processed by Krish filter, which is helpful to the restoration of details and edge information.

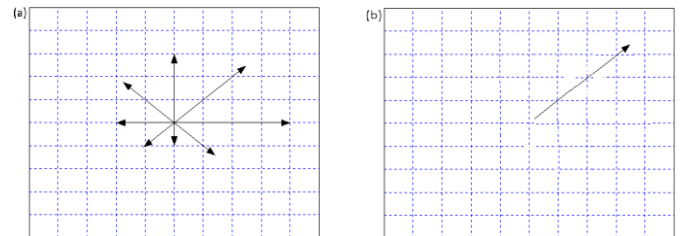


Figure 4. (a) Pixel represented by eight vectors (b) Pixel represented by a single vector



Figure 5. (a) Original image; (b) HDGG magnitude Result and (c) HDGG Orientation Result

5. THE PROPOSED FRAMEWORK

The proposed framework finds perfect facial features of the face image and provides directional gradients face recognition technique for identifying faces in images that ensures high accuracy and good effectiveness. The proposed face

recognition framework consists of two main processes: face training and face classification. Figure 6 illustrates the proposed framework model. This framework finds and manipulates facial features.

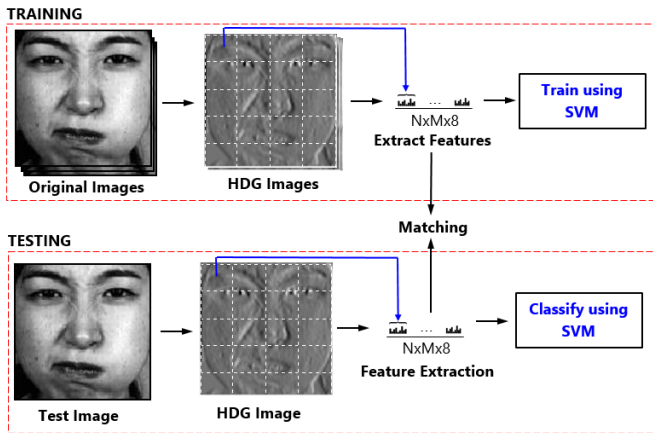


Figure 6. Proposed face recognition framework

5.1 Training process

The training process is an essential step as it extracts the facial features of human faces from images. Due to fact that the SVM is the most popular and most used approach for FR and FER systems, we use SVM with a multi-class kernel for the face image training and classification. The SVM takes a dataset as input and produces train set. The face-training algorithm goes through face images and identifies gradient directional pattern by applying HDG and HDGG operators in the whole face. We divide the image into $n \times m$ block to offer a compact representation and make our HDG and HDGG more robust to noise. Then, we compute HDG and HDGG for each of those blocks. We build a histogram among the 9-values of the gradient directions and their magnitude inside each block. Finally, all the 9-vectors are normalized and concatenated into a final feature vector, which is stored with face image in a database.

5.2 The classification process

Classification of the face image is the last step in the proposed recognition system. Firstly, we convert the input image to a face vector. Next, we normalize the face vector as the training phase. We compare it to those of the normalized vector to those of the train faces and calculate the distance between them. The face image is classified into the class with minimum distance. Develop, evaluate, and compare SVM machine learning approaches belonging using proposed descriptors using several metrics such as accuracy and execution time.

6. EXPERIMENTAL RESULTS AND EVALUATION

6.1 Experimental setup

All our experiments were performed on an Intel Core i3-5005U CPU 2.00 GHz, 4 GB RAM Laptop, using Windows 8 (64 bit) system. During the implementation process, facial images are collected from two image datasets that consist of different facial images that help to assess the efficiency of the

proposed operator's based gradient direction. For that purpose, we have used The Japanese Female Facial Expression (JAFFE) [10] and YALE database [11] benchmarks. After collecting the images, various features are extracted which are trained and classified using a Library for Multiclass Support Vector Machine (LIB-SVM) [33], pairwise approach (one vs one) applied with Linear Kernel function. LIB-SVM is a modern toolkit that contains several machine-learning algorithms that help in writing sophisticated C++ based face recognition applications. This library creates feature vectors out of faces and identifying a specific face across trained faces.

The performance of the system is examined using the classification accuracy with the traditional feature extraction operators such as LBP [2], LPQ [13], WLD [14], GDP [28], LDP [16], LDN [17], LGC [18], HOG [23], LGP [31], LTP [30], and GLTP [31]. All descriptors are implemented and evaluated on the same platform used for our proposed operators HDG and HDGG. We conducted many experimental evaluations in terms of block's dimension that contains small features information. The block's dimension defines the size of block of face image to the system. We evaluate the performance and effectiveness in automatically classifying face and facial expressions in terms of

- Execution time (T) is the time needed to accomplish the classification process of a specific face of faces dataset.
- Precision (P) to identify the number of relevant faces among the classified ones.
- Recall (R): to identify the number of relevant faces among the total number of expected relevant images.
- F-score to evaluate a weighted average of P and R. It is an important factor based on weighted recall. The F-score is computed by the and ARR values as follows:
- Accuracy (A): is the ratio of the correct predictions to the total predictions.

These metrics are computed as follow:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$F_score = 2 \times \frac{P \times R}{P + R} \quad (9)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

where,

- TP is the number of relevant faces (true positives);
- FP is the number of non-relevant faces (true negatives);
- TN is the number of relevant faces that are not classified by the given approach (true negatives);
- FN is the number of non-relevant faces that are not classified by the given approach (false negatives);

Based on the above metrics, the excellence of the proposed descriptors is evaluated while recognizing faces and facial expressions images from the dataset. Therefore, the main objective of an approach is to maximize both the precision and accuracy to minimize execution time. We consider an approach to work well if it can verify correctly as many faces in a very specific time with a very high ratio of accuracy. That is, the approach will be evaluated using four metrics, accuracy, precision, recall, and execution time. The effectiveness of HDG and HOG based classification process is compared with

all traditional feature extraction descriptors: LBP [2], LPQ [13], WLD [14], GDP [28], LDP [16], LDN [17], LGC [18], HOG [23], LGP [31], LTP [30], and GLTP [31], which is described in section 3. All these descriptors were implemented and integrated into our system. We compare the proposed descriptors HDG and HDGG with the most standard descriptors in recognizing facial expressions using the same datasets and considered performance metrics to enhance the objectivity and quality analysis. We will illustrate how SVM work with HDG and HDGG descriptors for enhancing accuracy rate and execution time. Then, we will show how to perform classification with both HDG and HDGG and how block dimension can affect the resulting classification. Finally, we will compare the execution time of this work with some other existing.

6.2 Evaluations on JAFFE dataset

The proposed descriptors HDG and HDGG are applied on the JAFFE dataset [10] to evaluate its effectiveness. The JAFFE dataset contains 213 peak facial expressions from 10 subjects, seven emotion categories are considered. They are happiness, sadness, surprise, anger, disgust, and naturel. The gray level images are of size 256 X 256. We used the fdlbmex library, free code available for MATLAB for face detection. We normalized all the evaluated face images before the experimentation in size of 128 X 128 pixels. In the experiment, we used an image as a test image, and remains images are used as training samples for the SVM classifier. The input face is divided into several n X m equal-sized blocks because a block can give more location information. In the experiment, the 128 X 128 sized face is equally divided into 8 X 8=64 blocks. A feature vector is extracted from each block. Concatenating feature histograms of all the blocks produces a unique feature vector for the whole image. Figure 7 presents samples of JAFFE. Table 1 shows the feature vector size and feature dimension for 64 blocks for the proposed HDG operator compared to other existing methods. In cases of LBP or LGC descriptor, the feature vector size is $8 \times 8 \times 256 = 16384$ for LBP, LGC and LPQ, $8 \times 8 \times 56 = 3584$ for LDP, $8 \times 8 \times 8 = 512$ for HDG and $8 \times 8 \times 9 = 432$ for HDGG.

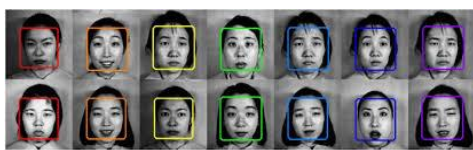


Figure 7. Sample images from JAFFE used in experimentation

Table 1. The feature vector sizes using different methods

Method	Features length	Features vector length
LTP	512	$8 \times 8 \times 512 = 32768$
GLTP	512	$8 \times 8 \times 512 = 32768$
LBP	256	$8 \times 8 \times 256 = 16384$
LDP	56	$8 \times 8 \times 56 = 3584$
LDN	56	$8 \times 8 \times 56 = 3584$
LGC	256	$8 \times 8 \times 256 = 16384$
LPQ	256	$8 \times 8 \times 256 = 16384$
WLD	32	$8 \times 8 \times 32 = 2048$
GDP	8	$8 \times 8 \times 8 = 512$
LGP	7	$8 \times 8 \times 8 = 448$
HOG	9	$8 \times 8 \times 9 = 576$
HDG	8	$8 \times 8 \times 8 = 512$
HDGG	9	$8 \times 8 \times 9 = 576$

6.2.1 Effectiveness evaluation

We evaluate the effectiveness of HDG and HDGG operator in classifying faces in terms of accuracy, precision F-score and execution time. The first experiment evaluates the impact of block-size on the recognition rate, HDG and HDGG operators are applied at different sizes: 1 X 1, 2 X 2, 4 X 4, 8 X 8, and 16 X 16. The results are shown in Table 2. As can be seen from Table 2, an 8 X 8-block combination is the best for the 128 X 128 original image size for both HDG and HDGG descriptors. It is observed that the small block dimension ignored the object boundaries. Therefore, a high block dimension will provide better recognition accuracy.

Table 2. Accuracy of HDG and HDGG using different block sizes on JAFFE database

Block Dimension	% Accuracy using HDG	% Accuracy using HDGG
1 x 1	38.57	39.52
2 x 2	53.33	66.67
4 x 4	73.80	82.38
8 x 8	90.00	91.43
16 x 16	89.04	88.57

Table 3. Accuracy of different methods on JAFFE database

Method	% Accuracy
LTP	40.95
GLTP	38.09
LBP	42.00
LDP	52.38
LDN	68.57
LGC	78.57
LPQ	69.52
WLD	81.90
GDP	87.14
LGP	83.80
HOG	88.57
HDG	90.00
HDGG	91.43

Table 4. Confusion matrix using HDG -SVM on JAFFE database

	AN	CO	DI	FE	HA	SA	SU
AN	100	0.00	0.00	0.00	0.00	0.00	0.00
CO	0.00	89.65	0.00	3.22	0.00	6.45	0.00
DI	0.00	0.00	84.38	0.00	3.33	9.68	3.70
FE	0.00	0.00	0.00	80.65	0.00	16.12	3.7
HA	0.00	0.00	3.13	6.45	90.00	0.00	0.0
SA	0.00	0.00	6.25	3.20	0.00	87.09	0.0
SU	0.00	0.00	0.00	0.00	0.00	0.00	100

Table 5. Confusion matrix using HDGG -SVM on JAFFE database

	AN	CO	DI	FE	HA	SA	SU
AN	96.6	3.45	0.00	0.00	0.00	0.00	0.00
CO	0.00	93.1	3.25	0.00	0.00	3.23	0.00
DI	0.00	3.45	81.2	0.00	3.33	9.68	3.70
FE	0.00	0.00	0.00	87.09	0.00	9.68%	3.70
HA	0.00	0.00	6.25	3.22	86.67	3.23	0.00
SA	0.00	0.00	0.00	3.22	0.00	96.77	0.00
SU	0.00	0.00	0.00	0.00	0.00	0.00%	100

To demonstrate the superiority of the proposed descriptors, we compare it with some standard descriptors techniques in terms of precision. Table 3 shows the comparison of accuracy

rates of the proposed descriptors with some other existing descriptors using JAFFE dataset where the HDG exceeds 90% and the HDGG exceeds 91%. The proposed approach performs better precision with JAFFE dataset. The confusion matrix is a summary of prediction results on a classification problem. The highest confusion value between different expressions means the good classification of the presented approach. The results in Table 4 and Table 5 show respectively the confusion matrix of 7-class expression, accuracy using template matching applied with SVM on JAFFE dataset using HDG and HDGG operators. The proposed operators' HDG or HDGG yields the best accuracy on all images involved in the experiments. The results in Table 4 and Table 5 prove the effectiveness of HDG and HDGG to identify facial expressions.

We compared the precision, recall, and F-score respectively of HDG and HDGG operators with performance of other traditional descriptors. As shown in Figure 8, we can notice that the results show that the proposed descriptors HDG and HDGG achieve higher precision over other traditional descriptors are 90.25% and 91.65% respectively in the JAFFE database. The proposed descriptors achieve the 90.83% Recall and 92.03% Recall for HDG and HDGG respectively. In addition, HDGG operator has higher F-score rate 91.84% and 90.54% for HDG descriptor.

The results in Figure 8 depicted that HDGG descriptor with SVM achieves 91.84% F-score, 91.65% precision, and 92.03% recall, which shows that the proposed HDGG-SVM effectively helps to recognize the face successfully.

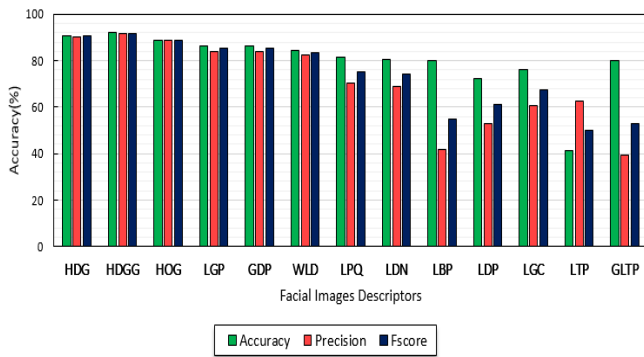


Figure 8. Performance comparison on JAFFE database using different descriptors

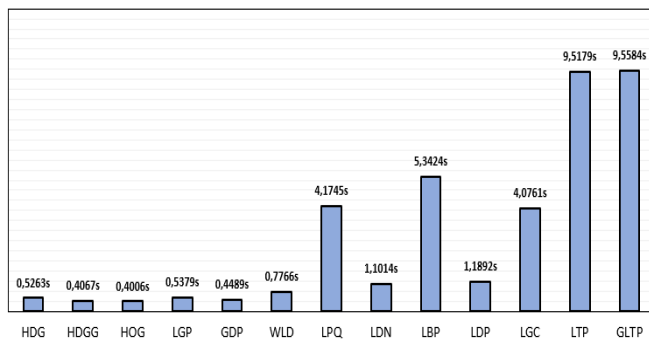


Figure 9. Execution time comparison using different descriptors on the JAFFE database

6.2.2 Execution time comparison

Based on the above descriptors, the extracted faces features are realized into the SVM method, which successfully

classifies the faces features with minimum time. The comparison of the execution time in seconds (see Figure 9) demonstrates that HDG and HDGG effectively classifies the extracted facial features with minimum time when compared to other traditional descriptors. We observe that the execution time of HDG and HDGG is 0.526s and 0.4067ms respectively. HDG and HDGG have a less vector size compared to other descriptors, which may decrease the classification-processing time and the learning time.

6.3 Evaluations on YALE dataset

The YALE Face dataset [11] contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, 1 per different facial expression or configuration: center light, no glasses, happy, left light, and normal, right light, sad, sleepy, surprised, and wink. All images are of size 64 X 64 divided into blocks with 8 X 8. In our work, we have only considered the frontal images. All images are 128 X 128 size and divided into 8 X 8 equally blocks. Figure 10 presents samples of YALE dataset.



Figure 10. Sample of YALE images used in experimentation

6.3.1 Effectiveness evaluation

Table 6 presents the accuracy rates of different operators on YALE database. Based on the mentioned results in Table 6, we can note that the proposed HDG and HDGG successfully recognize the facial expression features from the extracted ones with 92.12, due to the effective block segmentation and detection method. These results prove the efficiency of the proposed face descriptors in achieving a higher accuracy rate compared to other methods on YALE.

We have evaluated the precision with a different number of faces images using a palette of features extraction operators (HDG, HDGG, HOG, LGP, GDP, WLD, LPQ, LDN, LBP, LGC, LTP, and GLTP). As shown in Figure 11, we observe that the precisions rates of HDG and HDGG are very high which shows better performance of the proposed face recognition descriptors. The precision rate for HDG on YALE face database is 90.30% and the recognition rate for HDGG on YALE database is 92.12%. We compare the recall of proposed descriptors with recall results of other standard descriptors. As shown in Figure 11, we can notice that the recall rates of proposed descriptors are much better than other traditional descriptors. In addition, the recall of HDGG for YALE facial expression images compared to other descriptors where the JAFFE exceeds 92.16%. Low Recall rate in the case of GLTP explains the significant effect of light conditions on YALE images, where the covariance between YALE lighted image and JAFFE original image becomes higher. This improves the effectiveness of the proposed feature extraction operators. The F-score results of the proposed descriptors were compared to the F-score results of the other standard facial features extraction descriptors. As shown in Figure 11, the obtained results of HDG and HDGG are much better than other related

descriptors. In YALE database, SVM with HDG operator has better F-score of 91.25% while SVM with HDGG can achieve 92.67%. Thus, the proposed two face recognition descriptors achieve the 92.67% accuracy, 92.12% of precision, and 92.67% of recall, which shows that the proposed segmentation method effectively identifies the affected region from the dataset that helps to retrieve the features successfully.

Table 6. Accuracy of the different methods on YALE database

Method	% Accuracy
LTP	48.48
GLTP	36.36
LBP	87.88
LDP	56.36
LDN	79.39
LGC	70.00
LPQ	73.93
WLD	92.12
GDP	90.91
LGP	90.91
HOG	90.30
HDG	92.12
HDGG	92.12

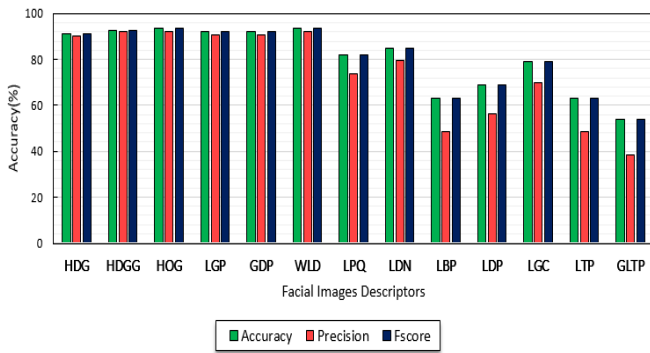


Figure 11. Performance comparison using different descriptors on YALE database

6.3.2 Execution time comparison

We evaluate the execution time of our approach using YALE dataset. Figure 12 presents the results of variation of the execution time according to applied feature extraction descriptor. In this figure, we can see that the response time significantly decreases according to varied tested facial images. As the results of execution time in cases of YALE face images, Figure 12 ensures that the proposed features extraction operators model provides fast execution in case of YALE face images. While comparing two face recognition operators HDG and HDGG will serve the fastest execution time when compared to the descriptors LBP and HOG.

6.4 Lessons learned and discussion

As observed from the experiments above, HDGG is much better than other traditional operators for classifying and recognizing faces images. HDG and HDGG are more efficient and increases the precision ratio, the recall ratio and the considerable execution time, but also make it practicable in spite of using very large-size faces images database. However, the execution time must be improved in future works with large-size dataset using advanced Face Deep Learning.

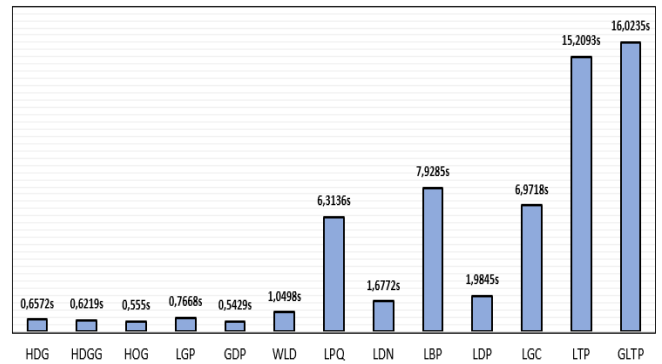


Figure 12. Execution time comparison using different descriptors on YALE database

7. CONCLUSION

Two new texture domain-based feature extraction descriptors HDG and HDGG based on the gradient directions have been proposed in this paper. The proposed features extraction descriptors benefit from the effectiveness of HDG and edge response value features. The recognition rate and effectiveness of two benchmarks are analyzed in terms of feature vector size and error rates. Where the feature vector size not exceeding 512, the recognition rate reaching 92.12%, and the error rate ranged 0.08-0.1. We have conducted many experimental evaluations with several datasets using HDG and HDGG. First, we show how the block dimension affected the accuracy of the model. Essentially, a high block dimension will generate better accuracy than a smaller one. Afterward, we use HDG and HDGG to find the best facial expression on JAFFE dataset and face recognition in YALE dataset. Finally, we trained a face classifier with 92.12% accuracy. Our future work will focus on advanced machine learning using OpenCV and Face Deep Learning.

REFERENCES

- [1] Gustavsson, A., Jonsson, L., Rapp, T., Reynish, E., Ousset, P.J., Andrieu, S., Wimo, A. (2010). Differences in resource use and costs of dementia care between European countries: Baseline data from the ICTUS study. *The Journal of Nutrition, Health & Aging*, 14(8): 648-654. <https://doi.org/10.1007/s12603-010-0311-7>
- [2] Bartlett, M.S., Littlewort, G., Fasel, I., Movellan, J.R. (2003). Real time face detection and facial expression recognition: development and applications to human computer interaction. 2003 Conference on Computer Vision and Pattern Recognition Workshop, Madison, Wisconsin, USA, pp. 53-53. <https://doi.org/10.1109/CVPRW.2003.10057>
- [3] Mehrabian, A., Russell, J.A. (1974). *An Approach to Environmental Psychology*. The MIT Press.
- [4] Zhang S., Hu B., Li T., Zheng X. (2018). A study on emotion recognition based on hierarchical adaboost multi-class algorithm. In: Vaidya J., Li J. (eds) *Algorithms and Architectures for Parallel Processing*. ICA3PP 2018. Lecture Notes in Computer Science, vol 11335. Springer, Cham. https://doi.org/10.1007/978-3-030-05054-2_8

- [5] Cruz, A.C., Bhanu, B., Thakoor, N.S. (2014). Vision and attention theory based sampling for continuous facial emotion recognition. *IEEE Transactions on Affective Computing*, 5(4): 418-431. <https://doi.org/10.1109/TAFFC.2014.2316151>
- [6] Moses, D.A., Leonard, M.K., Makin, J.G., Chang, E.F. (2019). Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nature Communications*, 10(1): 1-14. <https://doi.org/10.1038/s41467-019-10994-4>
- [7] Surie, D., Partonia, S., Lindgren, H. (2013). Human sensing using computer vision for personalized smart spaces. 2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing, Vietri sul Mare, pp. 487-494. <https://doi.org/10.1109/UIC-ATC.2013.24>
- [8] Zhang, X., Guan, Y., Wang, S., Liang, J., Quan, T. (2006). Face recognition in color images using principal component analysis and fuzzy support vector machines. 2006 1st International Symposium on Systems and Control in Aerospace and Astronautics, Harbin. <https://doi.org/10.1109/ISSCAA.2006.1627699>
- [9] Valstar, M.F., Pantic, M. (2011). Fully automatic recognition of the temporal phases of facial actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(1): 28-43. <https://doi.org/10.1109/TSMCB.2011.2163710>
- [10] JAFFE dataset. (2007). <http://www.kasrl.org/jaffe.html>, accessed on March 10, 2020.
- [11] YALE dataset. (2005). <http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html>, accessed on March 10, 2020.
- [12] Ahonen, T., Hadid, A., Pietikäinen, M. (2004). Face recognition with local binary patterns. In *European conference on computer vision*, (pp. 469-481). Springer, Berlin, Heidelberg, 2004.
- [13] Ojansivu, V., Heikkilä, J. (2008). Blur insensitive texture classification using local phase quantization. In: Elmoataz A., Lezoray O., Nouboud F., Mammass D. (eds) *Image and Signal Processing. ICISP 2008. Lecture Notes in Computer Science*, vol 5099. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69905-7_27
- [14] Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., Gao, W. (2009). WLD: A robust local image descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9): 1705-1720. <https://doi.org/10.1109/TPAMI.2009.155>
- [15] Islam, M.S. (2013). Gender classification using gradient direction pattern. *Sci. Int (Lahore)*, 797-799. arXiv:1310.6808
- [16] Jabid, T., Kabir, M.H., Chae, O. (2010). Local directional pattern (LDP) for face recognition. 2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE), Las Vegas, NV, pp. 329-330. <https://doi.org/10.1109/ICCE.2010.5418801>
- [17] Rivera, A.R., Castillo, J.R., Chae, O.O. (2012). Local directional number pattern for face analysis: Face and expression recognition. *IEEE Transactions on Image Processing*, 22(5): 1740-1752. <https://doi.org/10.1109/TIP.2012.2235848>
- [18] Tong, Y., Chen, R., Cheng, Y. (2014). Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle. *Optik*, 125(16): 4186-4189. <https://doi.org/10.1016/j.ijleo.2014.04.062>
- [19] Liu, Y.H., Chen, Y.T. (2007). Face recognition using total margin-based adaptive fuzzy support vector machines. *IEEE Transactions on Neural Networks*, 18(1): 178-192. <https://doi.org/10.1109/TNN.2006.883013>
- [20] Mahmud, F., Khatun, M.T., Zuhori, S.T., Afroge, S., Aktar, M., Pal, B. (2015). Face recognition using principle component analysis and linear discriminant analysis. 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, pp. 1-4. <https://doi.org/10.1109/ICEEICT.2015.7307518>
- [21] Meher, S.S., Maben, P. (2014). Face recognition and facial expression identification using PCA. 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, pp. 1093-1098. <https://doi.org/10.1109/IAdCC.2014.6779478>
- [22] Ksantini, R., Boufama, B., Ziou, D., Colin, B. (2010). A novel Bayesian logistic discriminant model: An application to face recognition. *Pattern Recognition*, 43(4): 1421-1430. <https://doi.org/10.1016/j.patcog.2009.08.021>
- [23] Dalal, N., Triggs, B. (2005). Histograms of oriented gradients for human detection. 005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, pp. 886-893. <https://doi.org/10.1109/CVPR.2005.177>
- [24] Kar, A., Bhattacharjee, D., Nasipuri, M., Basu, D.K., Kundu, M. (2013). High performance human face recognition using Gabor based pseudo hidden Markov model. *International Journal of Applied Evolutionary Computation (IJAEC)*, 4(1): 81-102. arXiv:1312.1684.
- [25] Tan, X., Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. In: Zhou S.K., Zhao W., Tang X., Gong S. (eds) *Analysis and Modeling of Faces and Gestures. AMFG 2007. Lecture Notes in Computer Science*, vol 4778. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75690-3_13
- [26] Siddiqi, M.H., Lee, S., Lee, Y.K., Khan, A.M., Truc, P.T.H. (2013). Hierarchical recognition scheme for human facial expression recognition systems. *Sensors*, 13(12): 16682-16713. <https://doi.org/10.3390/s131216682>
- [27] Ayeche, F., Alti, A. (2020). Performance evaluation of machine learning for recognizing human facial emotions. *Revue d'Intelligence Artificielle*, 34(3): 267-278. <https://doi.org/10.18280/ria.340304>
- [28] Islam, M.S. (2013). Gender classification using gradient direction pattern. *Sci. Int (Lahore)*, 25(4): 797-799.
- [29] Kotsia, I., Zafeiriou, S., Pitas, I. (2008). Texture and shape information fusion for facial expression and facial action unit recognition. *Pattern Recognition*, 41(3): 833-851. <https://doi.org/10.1016/j.patcog.2007.06.026>
- [30] Tan, X., Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, 19(6): 1635-1650. <https://doi.org/10.1109/TIP.2010.2042645>
- [31] Ahmed, F., Hossain, E. (2013). Automated facial expression recognition using gradient-based ternary texture patterns. *Chinese Journal of Engineering*, 2013: 1-8. <https://doi.org/10.1155/2013/831747>

[32] Islam, M.S. (2014). Local gradient pattern-a novel feature representation for facial expression recognition. *Journal of AI and Data Mining*, 2(1): 33-38. <https://doi.org/10.22044/JADM.2014.147>

[33] Chang, C.C., Lin, C.J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3): 1-27. <https://doi.org/10.1145/1961189.1961199>