Improved Face and Facial Expression Recognition Based on a Novel Local Gradient Neighborhood

Farid AYECHÉ, Adel ALTİ*, Abdallah BOUKERRAM
Department of Computer Science, Faculty of Sciences, University of Ferhat Abbas Setif-1 El-Bez City, Setif 19000, Algeria
{ayeche_farid@yahoo.fr} {altidel2002@yahoo.fr}

ABSTRACT: Computing efficiency is a key in biometric identification systems for automatic facial expression recognition. It was integrated within advanced pattern recognition as an excellent paradigm while users shifted towards underlying patterns. Most existing face recognition models suffer from a low recognition rate and significant execution time. To overcome these drawbacks, we propose a new Local Gradient Neighborhood (LGN) descriptor for effective face and facial expression recognition. Firstly, the LGN components obtained by applying LGN for each block of the face image which is represented by 9-size vector. Secondly, the system concatenates features vectors of different blocks to obtain the final feature vector for the face image. Finally, it applies SVM and KNN techniques to classify the input images. Unlike other similar works, the new proposed descriptor is evaluated on two benchmarks, for face recognition and facial expression recognition respectively. The experimental results show an excellent recognition rate and fast execution time. The recognition rate for the ORL face database is 98.50% and the recognition rate for the JAFEE database is 84.28%.

Subject Categories and Descriptors: [I.4.7 Feature Measurement]; [I.5 PATTERN RECOGNITION]: Neural nets

General Terms: Local Gradient Neighborhood, Face Expression Recognition, Classification, SVM, Feature Extraction

Keywords: Local Gradient Neighborhood, Face Expression Recognition, Classification, SVM, Feature Extraction and Reduction

Received: 26 September 2019, Revised 14 December 2019, Accepted 24 December 2019

Review Metrics: Review Scale: 0-6, Review Score: 4.75, Inter-reviewer consistency: 82.2%

DOI: 10.6025/jdim/2020/18/1/33-42

1. Introduction

With the massive growing utilization of digital technologies in everyday life, the efficient processing of multimedia information becomes major concerns. Due to the urgency of efficiently processing the multimedia information, face recognition is incorporated within advanced pattern recognition, which brings new solutions for people recognition. Such potential benefits are related with difficulties while guaranteeing facial contrasts such as emotions analysis, facial expression, smart human-computer interface, etc. [35].

This work focuses on the emotional recognition from facial images. The main steps of the emotion’s classification system are feature extraction: the extraction of distinctive facial features, classification: categorizing the extracted data (patterns) through learning process. In the literature, there are several learning techniques that used for facial recognition [13]. Most existing techniques aim to reduce the execution time of the computational tasks in varying faces images. They offer high-performance solutions for facial expression recognition and emotion analysis of the underlying feature patterns in biometric
Despite the recent efforts regarding the performance improvement of facial recognition techniques. It remains a fundamental challenge for improving the effectiveness of rate accuracy. Whereas, biometric recognition systems used standard descriptors for facial features to build their detection/recognition model: the Local Binary Pattern (LBP) [1] [2] and the Local Gradient Code (LGC) [3] [10]. The LBP considers only the relationship between the center pixel and its neighbor, but the LGC descriptor uses the relation between gray levels of the neighbourhood’s pixels, which is far from the center pixel. These descriptors do not reduce the information of facial features, as well as they possess specific problems in less classification accuracy.

We propose a novel descriptor known as LGN (Local Gradient Neighborhood) that combines both LBP with LGC qualities applied on well-known benchmarks system. This descriptor is applied on the face image to obtain the LGN image that divided into sub-blocks. Each block is mapped on a 9-size vector. Then, all blocks vectors are concatenated to obtain the image features vector. Finally, Support Vector Machines (SVM) [6] and K-Nearest Neighbors (KNN) [9] are applied to classify the input images based on features vectors. Some benefits consist of faster processing and support reduced information [6]. The proposed method improves facial recognition accuracy and execution time using reduced feature of vector size.

The rest of the paper is organized as follows. Section 2 presents the previous works related to face recognition. Section 3 details the proposed solution and description of the algorithm used. Section 4 presents the experimental results and comparative study. Finally, Section 5 concludes the work by proposing some future perspectives.

2. Related Work

This section presents some of the related approaches for face and facial expression recognition. In [13] proposed Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA) for reducing the attributes of large datasets, increasing interpretability and minimizing information loss. The model creates new uncorrelated variables that successively maximize variance. There are two-dimensional methods called 2D PCA [15] and 2D LDA [17]. These techniques enhanced recognition rates compared to 1D PCA methods [13]. We need to improve the results performance by stabilizing the images.

Another remarkable work [14] on extracting the feature of image and data reduction using the Iterated Function System (IFS) was carried out on several biometric applications such as IRIS recognition [20] and face recognition [21]. Then research on extracting feature vector to face images was done using IFS [36] [15]. However, the major drawback of IFS is the high computational complexity involved in the Partitioned Iterated Function System (PIFS) process.

The Local Binary Pattern (LBP) [2] is the most known feature extraction and texture analysis techniques due to its simplicity and efficacy. This model utilized the gray value variations of eight local vicinity pixels located in dissimilar regions to build a binary pattern. It defines a micro-edge information of the texture images such as edges, spots, or corners. Several variants of the LBP have been appearing such as Local Ternary Pattern (LTP), which is an extension of LBP [31]. This extension uses three levels of discrimination instead of two levels used in LBP. Authors in [18] proposed the Sobel-LBP operator based on gradient magnitude to increase the stability instead of gray-level intensity values for LBP. [33] introduced Gradient-based Ternary Texture Pattern (GLTP) that combines the advantages of Sobel-LBP and LTP operators. The robustness of GLTP against noise and brightness is significantly remarkable.

An advanced approach based on neural networks is to use a Convolutional Neural Network (CNN) to develop an online facial expression recognition system (Pato and Millett, 2010). Recently in [9] proposed a hybrid of machine learning algorithms, which performed inaccurate identification of human emotion. [30] proposed 3D-HOG temporal difference method for the facial micro-expression detection method. They applied Chi-square distance to find facial motion on 25 ROIs (Region of Interest) and detected the motion with an automatic peak detector.

Like it seen bellow, most of proposed works focused on recognizing correctly the facial and emotional expressions that require features extractions and facial image classification. However, there is still a lack of high-accuracy at low speed of face and facial expression recognition. Therefore, we need a more reduced feature vector for quality recognition. The methods and contributions of this paper are different from above related works. Firstly, a new Local Gradient Neighborhood (LGN) descriptor for effective face and facial expression recognition applied on well-known benchmark system. This descriptor is applied on face image to obtain the LGN image that divided into sub-blocks. Each block is mapped to 9-size vector. Then, all block vectors are concatenated to obtain the image feature vector. Finally, SVM and KNN techniques are applied to classify the input images based on features vectors. Unlike other similar works, we evaluated the
proposed descriptor on two different research directions—face recognition and facial expression recognition.

3. Proposed Technique

The proposed approach in this paper used a new descriptor called Local Gradient Neighborhood (LGN) for feature extraction and reducing. The proposed approach mainly consists of two main processes: the extraction feature and recognition processes based on mapped interval. Fig. 1 gives the comprehensive design of the proposed approach. The recognition processes apply SVM and KNN methods to classify the input images based on extracted feature vectors. Our goal is to improve a more accurate recognition rate and fast execution time without affecting the cost.

3.1 Gradient Feature Extraction and Reduction

Features extraction is one of the main steps in the face and facial expression recognition process (Figure 1). In this step, we extract unique features from an image, these features called descriptors that used to describe the image. The Local Binary Patterns (LBP) [1] [2] considers only the relationship between the center pixels and neighboring pixels, whereas the Local Gradient Code (LGC) [3] uses the relationship of the gray levels between the neighboring pixels. In this work, we propose a new descriptor called Local Gradient Neighborhood (LGN) that combines both LBP with LGC qualities. It considers both the relation of the gray level of the center pixel with those of the neighborhood as well as the relation of the pixels of the neighborhood between them. The approach is described in steps, including respectively the calculation of the LGN components of each block, building the final feature vector for the entire image and finally finding the requested face image.

Firstly, the gray levels of neighborhood pixels, \( g_i \), \( i = 1...8 \), are centralized with regard to the intensity of center pixel \( g_c \) according to Eq. 1.

\[
g_i = |g_i - g_c| \quad i = 1...8
\]

The second step is to compute gray levels difference between every two adjacent pixels. The new gray level of center pixels is computed using Eq. 2.

\[
\text{LGN}(x, y) = S(g_1 - g_2) * 2^7 + S(g_2 - g_3) * 2^6 + \\
S(g_3 - g_4) * 2^5 + S(g_4 - g_5) * 2^4 + \\
S(g_5 - g_6) * 2^3 + S(g_6 - g_7) * 2^2 + \\
S(g_7 - g_8) * 2^1 + S(g_8 - g_c) * 2^0
\]

Where LGN operator is defined using a function \( s(x) \) of variable \( x = 0, 1, ..., 255 \) \( s(x) \) represents gray levels difference between two adjacent pixels.

\[
s(x) = \begin{cases} 
1 & \text{if } x \geq 0 \\
0 & \text{else} 
\end{cases}
\]

Figure 2 shows a 3 x 3 template for LGN descriptor and Figure 3 shows the original image and the LGN results.

---

**Figure 1.** The proposed approach

![Original Image](image1)

![LGN Result](image2)

![Decomposition on nxm Blocks](image3)

**Figure 2.** 3*3 template for LGN descriptor

| \( g_1' \) | \( g_2' \) | \( g_3' \) |
| \( g_8' \) | \( g_c' \) | \( g_4' \) |
| \( g_7' \) | \( g_6' \) | \( g_5' \) |
The LGN operator is applied on the entire face image and then subdividing it into $N \times M$ blocks. For each block, a statistical histogram is extracted, reduced and normalized in the same way as the Histograms of Oriented Gradients (HOG) [25]. The use of a statistical histogram aims to improve the execution time and achieve maximum accuracy by reducing processed information. From each block, a feature vector of size $B$ is extracted such that $B \ll 256$. In this regard, we map each block in $B$ intervals (bins) [25]. The value of the $b^{th}$ interval is given by:

$$
\phi^b(x, y) = \begin{cases} 
\text{LGN}(x, y) & \text{if } \text{LGN}(x, y) \in \text{bin}_b \\
0 & \text{else} 
\end{cases} 
$$  

(4)

Where $\phi^b(x, y)$ is calculated on a block and therefore be performed efficiently using the integral image.

Each vector is normalized by the L2-Norm according to eq.5:

$$
U = \frac{U}{\sqrt{||U||^2}} 
$$  

(5)

The final feature vector is obtained by concatenating all the histograms vectors. It encoded as a unique vector of length $n \times m \times 9$. The extracted vectors are classified to determine their facial expression using supervised machine-learning algorithms. An unknown feature vector is compared to the trained feature vectors. We use two supervised machine-learning algorithms: KNN and SVM.

3.2 Face Images Classification

K-Nearest Neighbour (KNN): The KNN [9] is a very simple supervised classification algorithm that consists of electing the nearest neighbors of a new image based on a distance function. The KNN [9] is used in both statistical estimation and pattern recognition. The whole dataset is classified into either the training or the testing sample data. From the training sample point, the distance is calculated using the following Euclidean Distance.

$$
d = \sqrt{\sum_{i=1}^{n} (x_i - q_i)^2} 
$$  

(6)

Where $n$ is the size of the data, $x_i$ is an element in the dataset and $q_i$ is a central point. If the result is less than the neighbours around that data, then it is considered the neighbour.

Support Vector Machine (SVM): The SVM [5] is a trained algorithm for learning classification problems proposed by Vladimir Vapnik. SVM is based on structural risk minimization and is related to regularization theory [5] [7] [8]. The parameters are found by solving a quadratic programming problem with linear equality and inequality constraints. The defined algorithm searches for an optimal separating surface, known as the hyperplane. All the data is then separated using the hyperplane. If there are too many outliers using the calculated hyperplane, a new hyperplane is calculated until the simplest hyperplane is formulated.

4. Experiments

All our experiments have been performed on a PC using an Intel Core i5 2.67 GHz, 4 GB RAM, Windows 7 (32 bits) system and MATLAB. We evaluate the performance of the proposed approach on two benchmark databases for two research directions, the first is ORL database used.
for face recognition while the second is JAFFEE data-
base used for facial expression recognition. The ORL
dataset contains 400 face images of size $112\times92$ and the
JAFFEE dataset contains 210 facial expression images
of size $100\times100$. Firstly, we apply the LGN operator for all
images of a dataset, the LGN image result is segmented
into $N\times M$ blocks. Then, the LGN descriptor of each block
is represented as a 9-size vector, instead of using a fea-
ture vector with size 251 in LBP and LGC descriptors, a
feature vector with size 512 in LTP and GLTP descriptors.
Each facial image is represented by a vector of size
$N\times M\times 9$ instead of using a size of $N\times M\times 256$ in order to
reduce the execution time up to 96.5%.

The experiments for this paper are performed using two
different classification methods SVM and KNN whether
for face recognition or facial expression recognition. We
used the SVM pair wise approach (one vs one) with radial
basis kernel function and KNN-standard approach with $k$
equal to 3. A comparison with four existing standard de-
scriptors (e.g. Local ternary pattern (LTP), Local Direc-
tional Pattern (LDP), Local Binary Pattern (LBP) and Lo-
cal Gradient Code (LGC)) is conducted to evaluate the
performance of the proposed approach.

4.1 Evaluations on ORL Face Database
The ORL database was created by AT & T laboratories
Cambridge [27]. The database contains 40 people with
10 views each person. All the images are sized $112\times92$
pixels and collected on a dark background. Some im-
ages are collected on different dates with variations in
facial expressions (neutral expression, smile and closed
eyes) and partial occultation by the glasses. The head
poses have some variations in depth compared to the
frontal pose. However, these variations relate only to
certain people and are not as systematic as well. Each
face image is divided into $6\times 8$ blocks. Figure 4 shows a
sample of images of the ORL face database used in ex-
perimentation.

4.1.1 Evaluation of the Recognition Accuracy
We evaluate the proposed descriptor in terms of recogni-
tion accuracy, using two classifiers SVM and KNN. Table
1 shows the recognition accuracy of each classifier with
the reduced size of features vector for the ORL database.
Results show that SVM-LGN gives a recognition accu-
racy equal to 98.50%, although the recognition accuracy
with KNN is still good, 94.50%. It is observed that the
highest accuracy for face recognition is recorded for SVM-
LGN.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-LGN</td>
<td>98.50</td>
</tr>
<tr>
<td>KNN-LGN</td>
<td>94.50</td>
</tr>
</tbody>
</table>

Table 1. Recognition accuracy using different classifiers on
the ORL database

4.1.2 Recognition Accuracy Comparison
The aim of this study is to compare the recognition accu-
racy of the proposed descriptor with other proposed de-
scriptors LBP, LGC, LTP and GLTP. Classification using
the SVM and KNN methods. Table 2 shows the recogni-
tion accuracies of different descriptors on the ORL face
database. It can be observed from Table 2 that the pro-
posed descriptor yields a higher recognition accuracy us-
ing both classification methods.
In comparison to LTP [31] and GLTP [33] our proposed approach (i.e. SVM-LGN) yields superior performances on the ORL database, 98.5% over SVM-LTP (90.25%) and SVM-GLTP (83%).

The same conclusions were drawn from the recognition accuracy on the ORL face database for recognizing face images using our proposed approach KNN-LGN (94.50%) being slightly higher than using KNN-LTP (92.50%) and KNN-GLTP (93%); These results show that KNN-LGN method succeeds in finding high quality and accurate solutions.

### 4.1.3 Comparison of the Execution Time

Figure 5 presents a comparison between the execution time of the proposed approach with the related descriptors LBP, LGC, LTP, and GLTP. The SVM classifier takes significant execution time compared to KNN classifier for all images of the ORL database. However, the KNN classification algorithm runs 208 times faster than SVM classifier for the ORL database. The results in Figure 6 prove the ability of the proposed descriptor using KNN to respond to new image greatly quicker than SVM. In cases of LBP or LGC descriptor, the feature vector size is $6 \times 8 \times 256 = 1228$, $6 \times 8 \times 512 = 24576$ for LTP and GLTP and $6 \times 8 \times 9 = 432$ for LGN.

The SVM-LGN and KNN-LGN classifiers achieved fast execution time comparing with other proposed descriptors in all tested images. It can be seen clearly that the SVM’s execution time on the ORL database is only 1.33 seconds using our descriptor compared to LBP, LGC, LTP and GLTP descriptors. Additionally, the achieved KNN’s execution time on the ORL database using the proposed descriptor is only 0.03 seconds, which is more than 16 times lower than KNNS execution time using LBP and LGC. This is due to the reduced vector size of face features.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Descriptor</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>LBP</td>
<td>81.50</td>
</tr>
<tr>
<td></td>
<td>LTP</td>
<td>90.25</td>
</tr>
<tr>
<td></td>
<td>GLTP</td>
<td>83.00</td>
</tr>
<tr>
<td></td>
<td>LGC</td>
<td>94.50</td>
</tr>
<tr>
<td></td>
<td>LGN</td>
<td>98.50</td>
</tr>
<tr>
<td>KNN</td>
<td>LBP</td>
<td>93.75</td>
</tr>
<tr>
<td></td>
<td>LTP</td>
<td>92.50</td>
</tr>
<tr>
<td></td>
<td>GLTP</td>
<td>93.00</td>
</tr>
<tr>
<td></td>
<td>LGC</td>
<td>94.25</td>
</tr>
<tr>
<td></td>
<td>LGN</td>
<td>94.50</td>
</tr>
</tbody>
</table>

Table 2. Recognition accuracy using different classifiers and descriptors on the ORL database

![Figure 5. Execution time comparison using different feature extraction descriptors on the ORL database](image)
4.2 Evaluations on JAFFE Facial Expression Database
The JAFFE database consists of 210 images collected from 10 female expressed with 2 to 4 images for each expression [28]. The grayscale images resolutions are 256×256 pixels. In our work, fdlibmex library, free code available for Matlab is used for face detection. The used face has uniform size of 100×100 pixels. In order to improve accuracy, each image is divided into 11×11 blocks. Figure 5 shows a sample of images used in experimentation on the JAFFE database.

4.2.1 Evaluation of the Recognition Accuracy
To evaluate the recognition accuracy of the proposed technique, we use facial expression confusion matrix [26]. Table 3 reports our results on JAFFEE facial expressions database.

The confusion matrix for all classes of expressions is an important performance factor in facial expression recognition, which reflects the ability of the system to classify correctly facial expressions. The lowest confusion value between different expressions means the good classification of the presented approach. Table 4 shows the confusion matrix of 7-class expression, accuracy using template matching applied with SVM on JAFFEE facial expressions database. Clearly, the presented method yields better accuracy on all images involved in the experiments.

![Figure 4. Samples of JAFFE facial expression images used in the experimentation](image-url)

![Table 3. Recognition accuracy using different classifiers on the JAFFE database](table3)

<table>
<thead>
<tr>
<th>Classifier</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-LGN</td>
<td>84.28</td>
</tr>
<tr>
<td>KNN-LGN</td>
<td>72.38</td>
</tr>
</tbody>
</table>

![Table 4. Recognition accuracy using SVM classifier on the JAFFE database](table4)

<table>
<thead>
<tr>
<th></th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Joy</th>
<th>Surprise</th>
<th>Sad</th>
<th>Naturel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>96.70</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Disgust</td>
<td>3.30%</td>
<td>75.80%</td>
<td>9.30%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>6.50%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Fear</td>
<td>0.00%</td>
<td>3.40%</td>
<td>65.60%</td>
<td>3.20%</td>
<td>6.70%</td>
<td>19.3%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Joy</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>90.00%</td>
<td>0.00%</td>
<td>6.40%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.20%</td>
<td>90.00%</td>
<td>0.00%</td>
<td>7.40%</td>
</tr>
<tr>
<td>Sad</td>
<td>0.00%</td>
<td>3.40%</td>
<td>6.20%</td>
<td>3.20%</td>
<td>3.30%</td>
<td>80.60%</td>
<td>3.70%</td>
</tr>
<tr>
<td>Naturel</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>6.40%</td>
<td>92.50%</td>
</tr>
<tr>
<td></td>
<td>Anger</td>
<td>Disgust</td>
<td>Fear</td>
<td>Joy</td>
<td>Surprise</td>
<td>Sad</td>
<td>Naturel</td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>----------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>Anger</td>
<td>80.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>12.90</td>
<td>7.40%</td>
</tr>
<tr>
<td>Disgust</td>
<td>10.00%</td>
<td>75.80%</td>
<td>9.30%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.20%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Fear</td>
<td>0.00%</td>
<td>3.40%</td>
<td>71.80%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>19.30</td>
<td>7.40%</td>
</tr>
<tr>
<td>Joy</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>80.00%</td>
<td>3.30%</td>
<td>3.20%</td>
<td>14.80%</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.10%</td>
<td>3.20%</td>
<td>83.30%</td>
<td>0.00%</td>
<td>11.10%</td>
</tr>
<tr>
<td>Sad</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.10%</td>
<td>3.20%</td>
<td>0.00%</td>
<td>74.10</td>
<td>22.20%</td>
</tr>
<tr>
<td>Naturel</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.10%</td>
<td>0.00%</td>
<td>13.30%</td>
<td>3.20%</td>
<td>77.70%</td>
</tr>
</tbody>
</table>

Table 5. Recognition accuracy using KNN classifier on the JAFFE database

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Descriptor</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>LBP</td>
<td>52.85</td>
</tr>
<tr>
<td></td>
<td>LTP</td>
<td>66.19</td>
</tr>
<tr>
<td></td>
<td>GLTP</td>
<td>52.85</td>
</tr>
<tr>
<td></td>
<td>LGC</td>
<td>82.38</td>
</tr>
<tr>
<td></td>
<td>LGN</td>
<td>84.28</td>
</tr>
<tr>
<td>KNN</td>
<td>LBP</td>
<td>71.42</td>
</tr>
<tr>
<td></td>
<td>LTP</td>
<td>69.04</td>
</tr>
<tr>
<td></td>
<td>GLTP</td>
<td>68.57</td>
</tr>
<tr>
<td></td>
<td>LGC</td>
<td>72.85</td>
</tr>
<tr>
<td></td>
<td>LGN</td>
<td>94.50</td>
</tr>
</tbody>
</table>

Table 6. Recognition accuracy using different classifiers and descriptors on the JAFFE database

Figure 6. Execution time comparison using different feature extraction descriptors on the JAFFE database
In addition, we also evaluate the classification quality of KNN with LGN on JAFFEE facial expressions database. The results are detailed in Table 5. We can notice from Table 5 that the proposed method gives high recognition accuracy. This caused by a small change detection in the LGN descriptor of the facial expression images. Accordingly, the obtained results in both tables 4 and 5, the proposed approach in case of SVM classifier is more accurate comparing to KNN classifier.

4.2.2 Recognition Accuracy Comparison

Based on the obtained recognition accuracies as shown in Table 6, we not in most cases that the KNN method with the proposed descriptor achieved a high accuracy, 72.38%, on JAFFEE facial expression database compared to have applied the KNN method with LBP, LGC, LTP and GLTP descriptors, which shows better performance of the proposed descriptor. However, SVM with the proposed descriptor achieves an interesting accuracy, 84.28%, on JAFFEE facial expression database compared to have applied the SVM method using LBP, LGC, LTP and GLTP descriptors.

The achieved accuracy in the proposed descriptors LTP and GLTP is interesting, where the accuracy did not exceed 67%. Classification using our descriptor produced higher accuracy than using LTP and GLTP with lower recognition accuracy.

4.1.3 Comparison of the Execution Time

Figure 6 presents classification results of SVM with the proposed descriptor and the related proposed descriptors LBP, LGC, LTP, and GLTP. It takes an extensive classification time compared to the KNN classifier for all images of the JAFFE database. However, KNN classifier runs 208 times faster than SVM classifier for the JAFFE database. This signifies that after training and testing the KNN, it can respond to a new image greatly quicker than SVM in such a database. The size of a feature vector using LBP or LGC descriptor is $11 \times 11 \times 512 = 61952$ for LTP and GLTP while our feature vector size $11 \times 11 \times 9 = 1089$. Figure 6 shows that the proposed descriptor LGN achieved fast execution time comparing with other proposed descriptors LBP, LGC, LTP, and GLTP in all tested images. The total execution time of KNN classifier with the proposed descriptor on the ORL database is only 0.55 seconds, which is more than 16 times lower than KNNs execution time using LBP, LGC, LTP and GLTP. This is due for dimension reduction with LGN descriptor for recognition phase.

6. Conclusion and Future Works

This paper introduces a novel facial expression recognition method through a new Local Gradient Neighborhood (LGN) descriptor. The proposed descriptor benefit from both methods of feature extraction, Local Binary Pattern (LBP) and Local Gradient Code (LGC). Our goal is to improve the recognition accuracy and reduce computational cost. It is used to classify and recognize the facial Expression using SVM and KNN in two types of datasets benchmarks. The numerical results are found encouraging and demonstrate the efficiency and the effectiveness of SVM-LGN and KNN-LGN classification methods. All experiments illustrate that KNN can classify faster than SVM by a factor of thousands. As indicated in our experiments in two types of datasets, the recognition rate of SVM is ample greater than KNN because SVM methods usually make more of support vectors, whereas KNN requires few vectors for the same datasets. For future directions, one proposal is to investigate other classification methods for solving the problem of lighting property in order to evaluate the possible benefits in the recognition process.

References


Acoustics, Speech and Signal Processing (ICASSP) (p. 1298-1302).


