



HAL
open science

Comparative study of machine learning algorithms for face recognition

Harouna Naroua, Kadri Chaibou, Atsu Alagah Komlavi

► **To cite this version:**

Harouna Naroua, Kadri Chaibou, Atsu Alagah Komlavi. Comparative study of machine learning algorithms for face recognition. 2022. hal-03620410v2

HAL Id: hal-03620410

<https://hal.science/hal-03620410v2>

Preprint submitted on 27 Sep 2022 (v2), last revised 1 Feb 2024 (v7)

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

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

Comparative study of machine learning algorithms for face recognition

Harouna NAROUA¹, Kadri CHAIBOU², Atsu ALAGAH KOMLAVI^{*1,2}

¹Department of Mathematics and Informatic, Abdou Moumouni University, Niamey, Niger

²Department of Mathematics and Informatic, Abdou Moumouni University, Niamey, Niger

^{*1,2}Department of Mathematics and Informatic, Abdou Moumouni University, Niamey, Niger

*E-mail : alagahdejesus@yahoo.fr

Abstract

Background: The fundamental need for authentication and identification of humans through their physiological, behavioral or biological characteristics, continues to be used extensively to monitor and secure localities, property, financial transactions, etc. These biometric systems, especially those based on face characteristics, continue to attract the attention of researchers, major public and private services. In the literature, a multitude of methods have been deployed by different authors. Naturally an order of performance must be defined in order to be able to recommend the most effective method, because an error in an attempt to identify a person is of serious consequence. So, the main objective of the work carried out in this article is to make a comparative study of different existing techniques.

Methods: A biometric system is generally composed of four stages: acquisition of facial images, preprocessing, extraction of characteristics and finally classification. In this comparison the focus is on the performance of traditional machine learning algorithms used in the last stage. These algorithms are: Support Vector Machines (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), Random Forests (RF), Logistic Regression (LR) and Naive Bayesian Classification (NB: Naive Bayes' Classifiers). The comparison criterion is the average performance, calculated using three performance measures: recognition rate, f-measure of the confusion matrix, and the Area Under the Receiver Operating Characteristic (ROC) curve.

Results: Based on this criterion, the performance comparison of these traditional machine learning algorithms, ranked SVM the best, with an average performance of 99.05%.

Discussion: Consideration of performance measure of all six classification techniques such as SVM, LR, ANN, RF, KNN, NB classifiers, the following average performances are obtained respectively : 99.05%, 98.91%, 98.89%, 97.14%, 96.82% and 95.89%. According to these results of which the worst is 95.89%, the traditional machine learning algorithms for classification can be one key step for face recognition. The comparison of these results with existing state-of-the-art methods is nearly the same. Authors achieved for the following classifiers: SVM, ANN, LR, RF, KNN and NB respectively the following accuracies: 98.52%, 97.27%, 96.44%, 96.15%, 95.79%, 94.82%. Finally, in depth discussion, we concluded that between all these approaches which are useful in face recognition, the SVM is the best classification algorithm.

Keywords

Face Recognition; Comparative Study; Machine Learning Algorithms.

I INTRODUCTION

We can all notice that our world is characterized by increasing technological development in all sectors. For example, the means of transport and communication are very fast and very cheaper. That is why the world has become a global village. Some of the benefits of science and technology are unfortunately used to endanger people and their wealth. Due to this continuous development, the necessity to secure, to protect, and to control people and their wealth (bank accounts, cash) is increasing day after day. Because of the failures of traditional solutions such as passwords, extra badges, ID cards, PINs and passwords, keys, which were once effective, currently there is increased interest in biometrics systems, with many applications for access control using these systems being developed and implemented. Among the most used biometric modalities, the face has attracted many researchers for many reasons. First, a human face recognition system is non-intrusive. Second, the growth of technology, digital cameras and storage devices allow the management of enormous face databases. Finally, the evolution of facial recognition systems is mainly caused by scientific advances in machine learning with powerful data analysis algorithms. Consequently, facial recognition is becoming a reliable technology for identity verification. For this panoply of facial technologies presented and which is growing exponentially, it is more than necessary to perform a comparative study to assess the performances of these techniques. Moreover, it is necessary to choose one system that is the best. In this study, the question is to know which classification technique is the most efficient in face recognition, using a performance measure based on three specific indicators: accuracy rate, confusion matrix and area under the Receive Operating Characteristic (ROC) curve.

II BIOMETRIC SYSTEM TESTING

Any biometric system, such as facial recognition, has two phases: training and testing. The effectiveness of these systems depends on the quality of four additional stages within these two phases. These four additional stages are as follows: image acquisition, preprocessing, feature extraction and classification. There are many recent researches that use traditional machine learning algorithms for the classification stage. This section presents some of the most relevant recent works presented by other researchers.

In 2003, Xiaoou Tang and Xiaogang Wang [4], proposed a face recognition approach combining a Bayesian probabilistic classifier and Gabor filter responses. To evaluate this system, two image databases were used: the XM2VTS database on which the authors found an accuracy rate of 97.10% and the AR database which gave an accuracy rate of 93.30%. In 2006 and 2010 authors [5] [7] have put into practice the use of artificial neural networks in face recognition on feature vectors obtained with the Principal Component Analysis (PCA) and then with the Linear Discriminant Analysis (LDA). The authors obtained satisfactory results on the ORL dataset. Alaa Eleyan et al, got accuracy rates of 95.00% with the PCA and 97.00% with the LDA. Mayank Agarwal et al, in their works obtained an accuracy rate of 97.01% combining PCA and ANN like Alaa Eleyan et al. In 2012, some approaches were implemented using random forests and two facial feature extractors: Wavelet Gabor and Histogram of Oriented Gradients (HOG) [8]. The results obtained on the ORL database were: 95.10% and 95.70% for HOG and Gabor, respectively.

In 2015, work presented by Vaishali et al, [15] showed another way to use random forests. On the front end, they used a combination of PCA, DCT, and DWT for preprocessing and feature vector extraction on the ORL database. They obtained an acceptable accuracy of 96%. A study has achieved good result when using Support Vector Machine (SVM) as a classifier to implement face recognition system [12], as assessed by PCA. The recognition rate achieved by this study was, 98.75% [12]. Other works [13] have clearly reported good results by applying the Naive Bayes' Classifiers. With a face representation that is done using PCA a recognition rate of 94.35% on the Yale B is obtained; against a recognition rate of 94.56% using LDA as features extractor.

In 2016, certain approaches [18] have effectively proposed a combination of PCA and LDA as a features extractor. But before that, They carried out the image preprocessing stage by histogram equalization, image size normalization, and conversion rgb or coloured images into grey scale images. The evaluation of these systems based on Artificial Neural Network (ANN) gave an accuracy rate of 100.00% on the ORL database. In [20] the authors have performed face recognition using PCA, Local Binary Pattern (LBP) and SVM to achieve better recognition. For data acquisition Sumita Nainan et al, have chosen Yale and ORL databases. The face region from each image is extracted using Viola-Jones algorithm followed by image resizing and cropping into 70×70 pixel. Authors applied contrast limiting adaptive histogram equalisation for image enhancement. The feature extraction is performed separately with PCA and LBP. LBP is a method that used to describe texture characteristics of the surfaces. For classification, the feature vectors of training and testing images are input to the SVM classifier to create and evaluate the models. However, the method PCA+SVM obtained a precision rate of 86.67% on Yale database and a precision rate of 100.00% on ORL database. And, the method LBP+SVM gives 100% accuracy for both; Yale as well as ORL database [20].

In 2017, Vanlalhrauaia et al, [24] have proposed a face recognition by applying Logistic regression algorithm. In fact, they converted color images to gray scale, and removed the noise using LWSD and Adaptive Thresholding. Then, researchers found Face Region by horizontal and vertical profiling and extracted features from them to make the classification using Logistic Regression technique. Author achieved better recognition rate of 100% on the used dataset. Authors C. Zhou et al., [10] also exemplified the use of Logistic Regression classifier. They extracted the signatures using PCA, and obtained good recognition rates of 93.33% on Yale database [2] and 96% on ORL database.

In 2018, Lahaw Zied Bannour et al, [27] have presented methods using PCA, LDA, Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT). The Two-Dimensional Principal Component Analysis (2D-DWT) is a multi level decomposition technique used to preprocess images. It converts images from spatial domain to frequency domain. This operation decomposes the original image into four sub-bands low-low (LL), low-high (LH), high-low (HL) and high-high (HH). The low-low (LL) sub-band is used as input image for feature extraction process based on ICA, PCA and LDA algorithms. These features were then classified by using the SVM. The models DWT+PCA+SVM, DWT+LDA+SVM, DWT+ICA+SVM obtained on the ORL database the following rates of recognition respectively: 96%, 96%, 94.50%.

Kak et al, [29] applied DWT and PCA techniques for preprocessing and features extraction, and a K-Nearest Neighbors (KNN) classifier for classification. They obtained a 99.25% recognition rate on the ORL database. Other approaches using KNN have been tested on face features obtained with PCA [14] and using the ORL dataset. That previous researcher found a 92.47% recognition rate. In another study [19], Bala et al., also studies facial recognition using KNN classifiers. Those authors developed the extraction stage with LDA, and this approach provided an interesting recognition rate of 93.70% on the ORL database of pictures.

Huda Mady et al, [28] proposed a method using Random Forest classifier (RF). First of all, the system utilized Viola Jones Algorithm in detecting faces from a given image. Then, to extract the feature vectors, computation of LBP and HOG descriptors is done simultaneously. Authors combined both HOG and LBP features for classification done by Random Forest classifier (RF). At last, the system obtained a recognition precision of 97.82% Mediu staff database (Mediu-S-DB).

in 2019, three groups of researchers implemented facial recognition systems using SVM methods. Authors Laith R. Fleah et al, [30], started with image preprocessing using two techniques: Normalization size of the image and Contrast stretching. Then, they applied the DWT and the PCA to decrease the size of the image to the half, remove the noise, reduce the processing time and extract features for classification. Pranati Rakshit et al, [36], also used PCA technique to extract the features and reduce dimensionality of the images. In the same way, Putta Sujitha et al [32], applied Gabor Wavelet to extract rotation and scale invariant features from the normalized face image and also used PCA for feature reduction. These three authors have done the classification with SVM. Thus, on three data sets: ORL, AR and Grimace, Putta Sujitha et al, obtained the following accuracy rates respectively: 97.65%, 92.31% and 100.00%. The models Laith R. Fleah et al, achieved an accuracy of 95.20% for the Yale database and 96.25% for the ORL database. Pranati Rakshit et al, also performed a good accuracy rate of 92% on ORL database.

Sri Sutarti et al [33], put forward a method for extracting features from face images. This new technique called Two-Dimensional Principal Component Analysis (2D-PCA) is developed for face representation and recognition, and overcame weakness of PCA. However, this PCA method was applied by Ni Kadek Ayu Wirdiani et al, [34], to extract characteristic features. In addition, conversion to grayscale, region of interest (ROI) and haar cascade segmentation were applied for image preprocessing by Ni Kadek Ayu Wirdiani et al. After this, the classification is done by using KNN method. Result testing with 2D-PCA+KNN method obtained an accuracy rate of 96.88% on ORL database [33], while PCA+KNN method performed a recognition rate of 81% on a dataset research contains 790 faces from 158 people taken from several angles [34].

In 2021, a group of researchers named Muhammad Shakeel Faridi et al, through their experimental work [37], performed a comparative study between four traditional machine learning algorithms using the ORL database. These four machine learning techniques being compared are: PCA, 1-Nearest Neighbor (1-NN), LDA, and SVM. Then, the researchers extracted features from the datasets, trained the classifiers and created the models. Finally, they evaluated the performance of these models with the 5-fold cross validation (n=5) technique. The systems based on LDA, 1-NN, PCA and SVM achieved 96%, 96.25%, 96.75% and 98% accuracy rates respectively [37]. These results demonstrate the effectiveness of SVM method in face recognition.

In 2022, ZM Nabat et al [40], work on the fusion of the two algorithms to implement a face recognition system. the first one is the SVM method. And the second is a new metaheuristic method called Rain optimization algorithm (ROA) that is inspired by the raindrops. This method can find global extremum as well as local extremums if its parameters are correctly tuned. In this study, the goal of ROA is the optimization of parameters such as: C and σ of the Radial Basis Function (RBF) kernel SVM. Thus, to evaluate the proposed system, they used the Yale face dataset. by performing an n-fold validation ($n = 10$), the authors obtained a recognition rate of 86% [40].

Benradi hicham et al, [39], performed a comparative study of machine learning algorithms in the field of face recognition. For image acquisition, these researchers worked with two databases: the ORL database and the Sheffield face database which contains 564 images of 20 individuals where each image has an identical size of 220×220 pixels and a 256-bit grayscale. On these databases divided the two datasets into train and test dataset, they applied the feature extraction using the Scale-Invariant Feature Transform (SIFT), the Speeded Up Robust Features (SURF), the Features from Accelerated Segment Test (FAST) and LBP. The features vectors extracted from the face images, are used in classification by applying SVM, KNN, PCA, and 2D-PCA methods to create the models of prediction. The prediction models were tested. The results showed that the proposed techniques such as: SIFT+SVM, LDA+KNN, PCA, 2D-PCA performed respectively in ORL Dataset the following accuracy rates: 99.16%, 96%, 92.50% and 96.25% [39]. And, On Shieffiled Dataset, the accuracy rates of the methods SIFT+SVM, LDA+KNN, PCA, 2D-PCA are as follows in order : 99.44%, 96%, 27.11%, 43.10%. Definitively, Benradi Hicham et al, concluded that SVM perform better than the other methods tested.

Nowadays, many researchers of facial recognition use the new techniques of Deep Learning (DL) and more specifically Convolutional Neural Networks (CNN) [35] [38]. In fact, before the emergence of Deep Learning, feature extraction was carried out for traditional Machine learning techniques. The difficulty is that it is necessary to choose which features are important in each given image and most of the applied features need to be identified by an domain expert. The biggest advantages of CNN techniques are that they typically are composed of two basic parts such as feature extraction and classification. Feature extraction includes several convolution layers followed by Max-Pooling and an Activation Function. The classification usually consists of fully connected layers. LeNet-5 was one among the earliest CNN which promoted the event of deep learning [3].

Some works [21] of face recognition are implemented by simplifying a CNN structure to get a new accurate version. This new version of CNN has fused together the convolution and sub sampling to reduce the number of layers required. The proposed architecture of Syafeeza, A.R, et al, shown in Figure 3 has four layers named : C1, C2, C3 and an output layer F4. C1 layer has 5 feature maps, 14 feature maps in C2 layer and 60 feature maps in C3 layer. The F4 layer has 40 feature maps because of the classification of 40 different labels contained in the ORL database. The design has a reduced size of feature maps, so as not to require padding in the convolution process. After some preprocessing steps, authors Syafeeza, A.R, et al, created the model and evaluated it using the images of 40 subjects from AT&T database and 10 subjects from JAFFE database. On test dataset the classification accuracy produced 100%.

S. Meenakshi et al, [31] proposed a new CNN architecture in order to nullify the effect of pose and illumination changes, any occlusions, the facial expressions, etc. The developed method is built by varying feature maps in convolutional layers C1, C2 and C3, with the aim of finding the most efficient architecture. This architecture has also an input layer with image size of 32×32 , subsampling layer and fully connected layer. To evaluate the efficacy of the model, several tests are conducted on ORL database after image resizing to 32×32 with pixels. Among different architectures, the best accuracy of 98.75% is obtained with 15-90-150 architecture [31].

A model for face recognition was proposed by Zhiming Xie et al, [35]. Figure 1 shows the architecture of this model based on two main points: the number of hidden layer neurons and the number of convolutional layer feature maps. This has made the accuracy of face recognition further improved. This CNN architecture includes the input layer, convolution layer 1, pooling layer 1, convolution layer 2, pooling layer 2, fully connected layer and Softmax regression classification layer. So, they defined the structure C1-C2-H, where C1 is the number of feature maps in the first convolutional layer, C2 is the number of feature maps in the second convolutional layer, and H represents the number of hidden layer neurons. Through multiple sets of experimental tests, Zhiming Xie et al, found the optimal model 36-76-1024. And they obtained a face recognition rate reaches 100% [35] on ORL database.

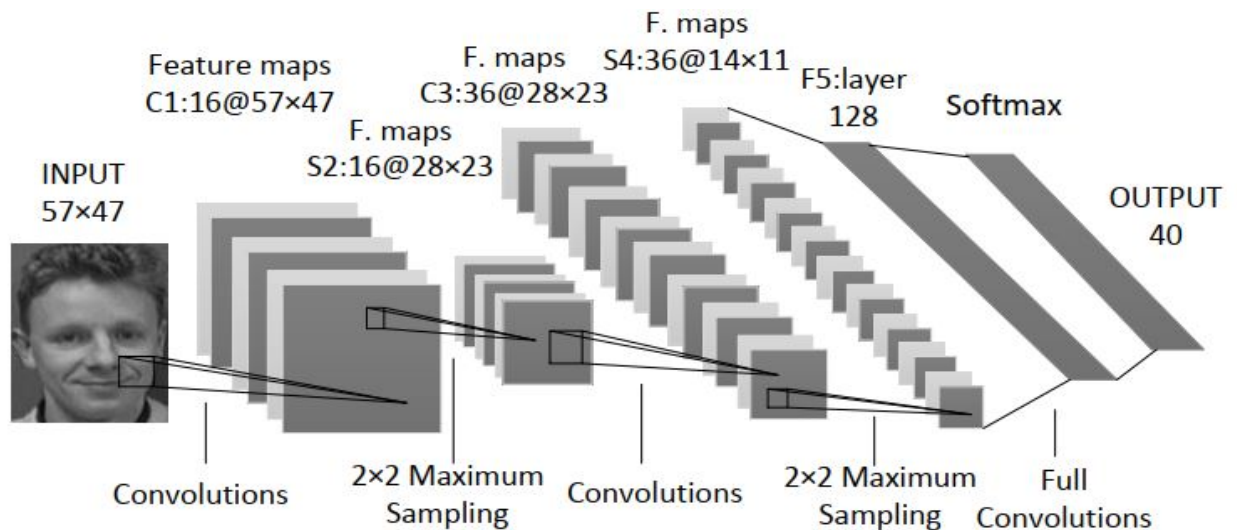


Figure 1: CNN infrastructure model designed for experiment. [35]

Yohanssen Pratama et al, in their works [38], used one of the CNN architectures called Residual Networks-50 to perform a face recognition system. These Residual Networks achieves 3.57% error on the ImageNet test set and won the first place on the ILSVRC 2015 classification task [17]. The contribution of this research paper is to determine effectiveness ResNet architecture using different configuration of hyperparameters such as the number of hidden layers, the number of units in the hidden layer, batch size, and learning rate. To evaluate the model they used a dataset of 1050 images divided into a train and test datasets, with a share of 80% for the data train and 20% for the test data. In the process of training models, several hyperparameters are used. Based on this, Yohanssen Pratama et al, have found that learning rate of 0.0001, epoch of 100 and step per epoch of 150 give a model with accuracy rate of 99%.

In this day and age, Deep Learning techniques have achieved accuracy that are far beyond that of classical Machine Learning methods in many domains. However, a good and outstanding feature extraction and a traditional Machine Learning algorithm can be combined to develop an efficient system for face recognition in terms accuracy rate. That is why in this paper the focus is on these classical techniques of Machine Learning. And, Tables 1, 2, 3, 4, 5 and 6 show averages of some good performances of each one of them.

Authors	Methods	Database	Accuracy (%)
Sumita Nainan et al, 2016	PCA+SVM	ORL	100
Sumita Nainan et al, 2016	LBP+SVM	ORL	100
Putta Sujitha et al, 2019	GABOR+PCA+SVM	ORL	100
Benradi hicham et al, 2022	SIFT+SVM	Sheffield	99.44
Benradi hicham et al, 2022	SIFT+SVM	ORL	99.16
Hadi Seyedarabi et al, 2015	PCA+SVM	ORL	98.75
Muhammad Shakeel et al, 2021	SVM	ORL	98
Putta Sujitha et al, 2019	GABOR+PCA+SVM	ORL	97.65
Laith R. Fleah et al, 2019	DWT+PCA+SVM	ORL	96.25
Lahaw Zied Bannour et al, 2018	DWT+PCA+SVM	ORL	96
Average Accuracy	—	—	98.52

Table 1: : Average accuracy of some works using SVM in the recent literature.

Authors	Methods	Database	Accuracy (%)
Alaa Eleyan et al, 2006	PCA+ANN	ORL	95
Alaa Eleyan et al, 2006	LDA+ANN	ORL	97
Mayank Agarwal et al, 2010	PCA+ANN	ORL	97.1
Gurleen Kaur et al, 2016	PCA+LDA+ANN	ORL	100
Average Accuracy	—	—	97.27

Table 2: : Average accuracy of some works using ANN in the recent literature.

Authors	Methods	Database	Accuracy (%)
Kak et al, 2018	DWT+PCA+KNN	ORL	99.25
K.S.Maheswari et al, 2015	PCA+KNN	ORL	92.47
Bala et al, 2016	LDA+KNN	ORL	93.7
Sri Sutarti et al, 2019	2D-PCA+KNN	Sheffield	96.88
Muhammad Shakeel et al, 2021	KNN	ORL	96.25
Benradi hicham et al, 2022	LDA+KNN	ORL	96
Benradi hicham et al, 2022	LDA+KNN	Sheffield	96
Average Accuracy	—	—	95.79

Table 3: : Average accuracy of some works using KNN in the recent literature.

Authors	Methods	Database	Accuracy (%)
C. Zhou et al, 2015	PCA+Logistic Regression	ORL	96
C. Zhou et al, 2015	PCA+Logistic Regression	ORL	93.33
Vanlalhraia et al, 2017	Logistic Regression	FEI image dataset	100
Average Accuracy	---	---	96.44

Table 4: :Average accuracy of some works using Logistic Regression in the recent literature.

Authors	Methods	Database	Accuracy (%)
Xiaoou Tang et al, 2003	Gabor Filter+Naive Bayes	XM2VTS	97.1
Xiaoou Tang et al, 2003	Gabor Filter+Naive Bayes	AR	93.30
Telgaonkar Archana et al, 2015	PCA+Naive Bayes	Yale	94.35
Telgaonkar Archana et al, 2015	LDA+Naive Bayes	Yale	94.56
Average Accuracy	---	---	94.82

Table 5: : Average accuracy of some works using Naive Bayes Classifier (NB) in the recent literature.

Authors	Methods	Database	Accuracy (%)
Abdel Ilah Salhi et al, 2012	HOG+RF	ORL	95.1
Abdel Ilah Salhi et al, 2012	Gabor+RF	ORL	95.70
Vaishali et al, 2015	PCA+DCT+DWT+RF	ORL	96
Huda Mady et al, 2018	LBP+HOG+RF	Mediu staff	97.82
Average Accuracy	---	---	96.15

Table 6: : Average accuracy of some works using Random Forest Classifier (RF) in the recent literature.

A summary of the results produced by these authors in the recent literature gives the following histogram of Figure 2 which contains the average recognition rates per traditional machine learning algorithm. This analysis ranks Support vector machines (SVM) classifiers at the top, with a satisfactory average recognition rate of 98.52%. They are followed, in order, by Artificial Neural Networks (ANN), Logistic Regression (LR), Random Forests (RF), K-Nearest Neighbors (KNN), and Naive Baye's Classifiers (NB) which provided 97.27%, 96.44%, 96.15%, 95.79%, 94.82% as average recognition rate respectively. This ranking will be compared to the one that will be established during the experiments.

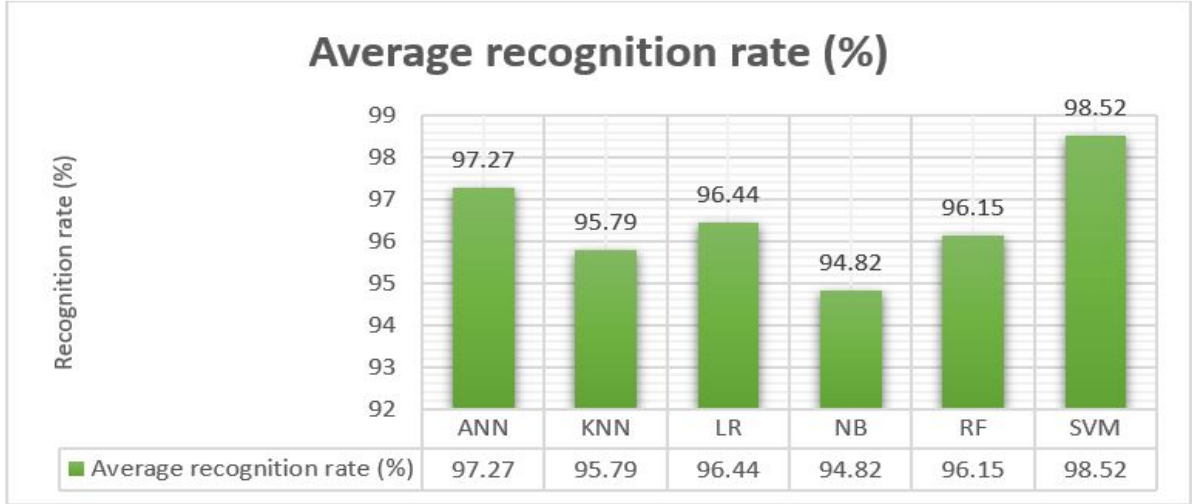


Figure 2: Performance of some traditional machine learning algorithms in recent literature.

III MACHINE LEARNING ALGORITHMS

Machine Learning is defined as the science of programming computers to learn from data [23]. It can be subdivided into supervised and unsupervised learning. Supervised learning is defined by its use of labeled datasets to train models and predict outcomes accurately. It can be separated into two types of problems: classification and regression. In this paper of comparative study, the focus is on classification algorithms. There are different types of classification methods as such: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression (RF), Random forest (RF), Naive Baye’s Classifiers (NB), Artificial Neural Network (ANN).

3.1 Support vector machines (SVM)

Support vector machines are powerful supervised machine learning algorithms, capable of performing binary classifications for linear or non-linear problems. These very effective methods for general purpose pattern recognition [20], were introduced by Vladimir Vapnik in 1965. In addition to performing linear classification, SVM can efficiently perform a non-linear classification using the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. This extension to non-linear problems was discovered in 1995 by Vapnik in association with Corinna Cortes [12] [26]. SVMs were designed for binary classification and do not natively support classification tasks with more than two classes. However, one other extension that has made SVMs very popular is the ability to solve multiclass classification problems. [22]. Given training samples (x_i, y_i) where x_i are observations and $y_i \in \{-1, +1\}$ are labels; this dataset is linearly separable when there exists a hyperplane whose equation is a linear function in x as:

$$f(x) = w^* \cdot x + b^* \tag{1}$$

where w is a weight vector, b is the bias and x is the problem variable. To decide to which class a data item x_i belongs, we just take the sign of the decision function [22], thus:

- if $w^* \cdot x_i + b^* > 0$ then x_i belongs to the +1 positive label class,
- if $w^* \cdot x_i + b^* < 0$ then x_i belongs to the negative label class -1.

There have been many various hyperplanes that are able to separate the data. The basic idea of SVM methods is to obtain the best function $f(x)$ that maximizes the margins of the two classes. Learning data nearest to the optimal hyperplane are called support vectors. The margin is the distance between the hyperplane and its support vectors. The optimal hyperplane is therefore selected so as to maximize the margin. And, to obtain the largest possible margin between the hyperplane and the support vectors, we need to minimize the following function :

$$L = \frac{1}{2} \|w\|^2 \quad (2)$$

under the constraints $y_i f(x_i) \geq 0$ which verify that the data (x_i, y_i) are well classified.

3.2 K-Nearest Neighbours (KNN)

The K-Nearest Neighbors algorithm (KNN) is one of the simplest yet most commonly used classifiers in supervised machine learning [16] [14]. This lazy learner technique doesn't technically train a model [25] [11]. So, the training stage doesn't require any computation. It consists of storing the training dataset with little or no processing. However, the test stage requires an intensive distance calculation between a test data point and all training data points, in order to find the k nearest neighbors of this test data. The distance is so important for the KNN technique. The most often used distance, is the Minkowski distance of order q where q is an integer. Between two data points $U = (U_1, \dots, U_n)$ and $V = (V_1, \dots, V_n)$, this distance is defined by the following formula :

$$d_q(U, V) = \left(\sum_{i=1}^n |U_i - V_i|^q \right)^{\frac{1}{q}} \quad (3)$$

It is a generalization of both the Manhattan distance (order $q = 1$) and the Euclidean distance (order $q = 2$). The formulas of these two particular distances are defined by:

- Manhattan distance is the distance a car would drive in a city (Manhattan) :

$$d_1(U, V) = \sum_{i=1}^n |U_i - V_i| \quad (4)$$

- Euclidean distance is the length of line segment between the two points :

$$d_2(U, V) = \sqrt{\sum_{i=1}^n (U_i - V_i)^2} \quad (5)$$

The KNN algorithm is widely used to determine to which class a new data belongs to. As shown in Figure 3, the predicted class is typically the class that is the most voted in the k nearest neighbors of this data (majority vote of its neighbors) [26].

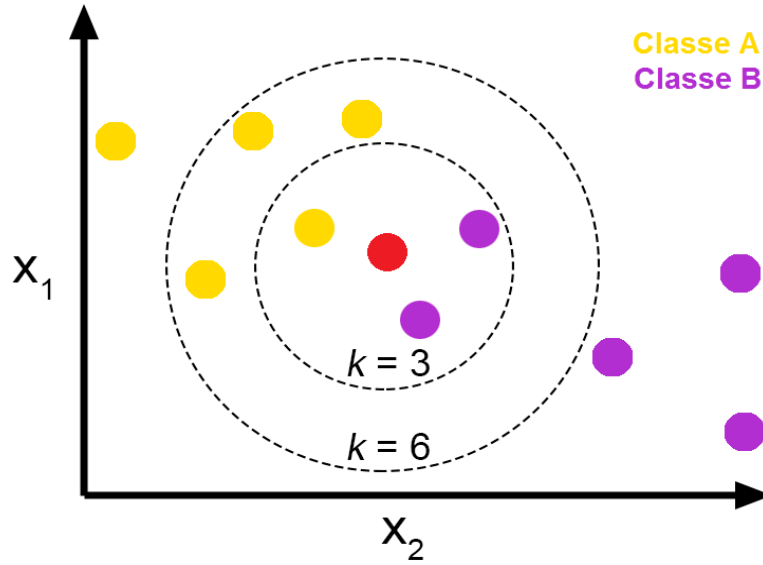


Figure 3: K-Nearest Neighbours Algorithm.

3.3 Logistic regression (LR)

The logistic regression is actually a widely used supervised classification technique [25]. This originally binary, and later multiclass classification algorithm began its emergence in 1989 with Hosmer and Lemeshow [10]. Logistic regression uses a function that takes as argument a linear combination of the input variables and models the probability of belonging to a class. This logistic regression model computes a weighted sum of the input features, and a linear model like $u(x)$ is included in a logistic function $\rho(u)$, sigmoid function that returns values between 0 and 1 [26] [11]. These functions are defined by these following formulas:

$$U(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (6)$$

$$\rho(u) = \frac{1}{1 + e^{-u}} \quad (7)$$

$$\rho(u) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)}} \quad (8)$$

Instead of providing the result directly, the logistic function $\rho(u)$, provides the logitics of the linear model $u(x)$.

3.4 Random Forest (RF)

As shown in the figure 4, Random Forests is a collection of trees-based models trained on random subsets of the training data. This algorithm is an ensemble of decision trees, usually trained with the bagging method (or sometimes pasting). A method is called bagging when sampling is performed with replacement. The first version of this technique was created in 1995 by Tin Kam Ho. Then in 2001 [[26], an extension of the decision tree forests was developed by Leo Breiman and Adele Cutler who registered the Random Forests as a trademark in 2006. The Random forests can be used for both classification and regression tasks. And, for both supervised learning settings, a collection of labeled observations known as the training set like $D = (x_i, y_i), \dots, (x_n, y_n)$ was provided, each x_i is a vector and y_i its true target variable.

- **Classification:**

For classification tasks, the prediction of the random forest is the most dominant class among the predictions made by the individual trees. Consider T trees in the forest, and consider that a class called m has received a number of votes called V_m defined by the following formula :

$$V_m = \sum_{t=1}^T I(\hat{y}_t == m) \quad (9)$$

where, \hat{y}_t is the prediction of the t -th tree on a particular random subsets of the training data. The function $I(\hat{y}_t == m)$ takes on the value 1 if the condition is met, else it is zero. The final class predicted by the algorithm is the class with the most votes as shown in the following formula where $m \in \{1, \dots, M\}$:

$$\hat{y} = \arg \max_{m \in \{1, \dots, M\}} V_m \quad (10)$$

- **Regression:**

For regression tasks, the prediction of the random forest is the mean or average of the predictions made by the individual trees. Consider T trees in the forest, and each predict \hat{y}_t , then the final prediction \hat{y} is :

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t \quad (11)$$

Algorithm 1 Random Forest for Classification (RFC)

- 1: **for** $b \leftarrow 1, B$ **do**
 - 2: (a) Draw a bootstrap sample Z^* of size N from the training data.
 - 3: (b) Grow a random forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached.
 - 4: (I) Select m variables at random from the p variables.
 - 5: (II) Pick the best variable/split-point among the m .
 - 6: (III) Split the node into two daughter nodes.
 - 7: **end for**
 - 8: Output the ensemble of trees $\{T_b\}_1^B$
 - 9: Make prediction at new point x :
 - 10: Let $\hat{C}_b(x)$ be the class prediction be the class prediction of the b th random forest tree. Then $\hat{C}_r f^B(x) = \text{majority vote } \{\hat{C}_b(x)\}_1^B$.
-

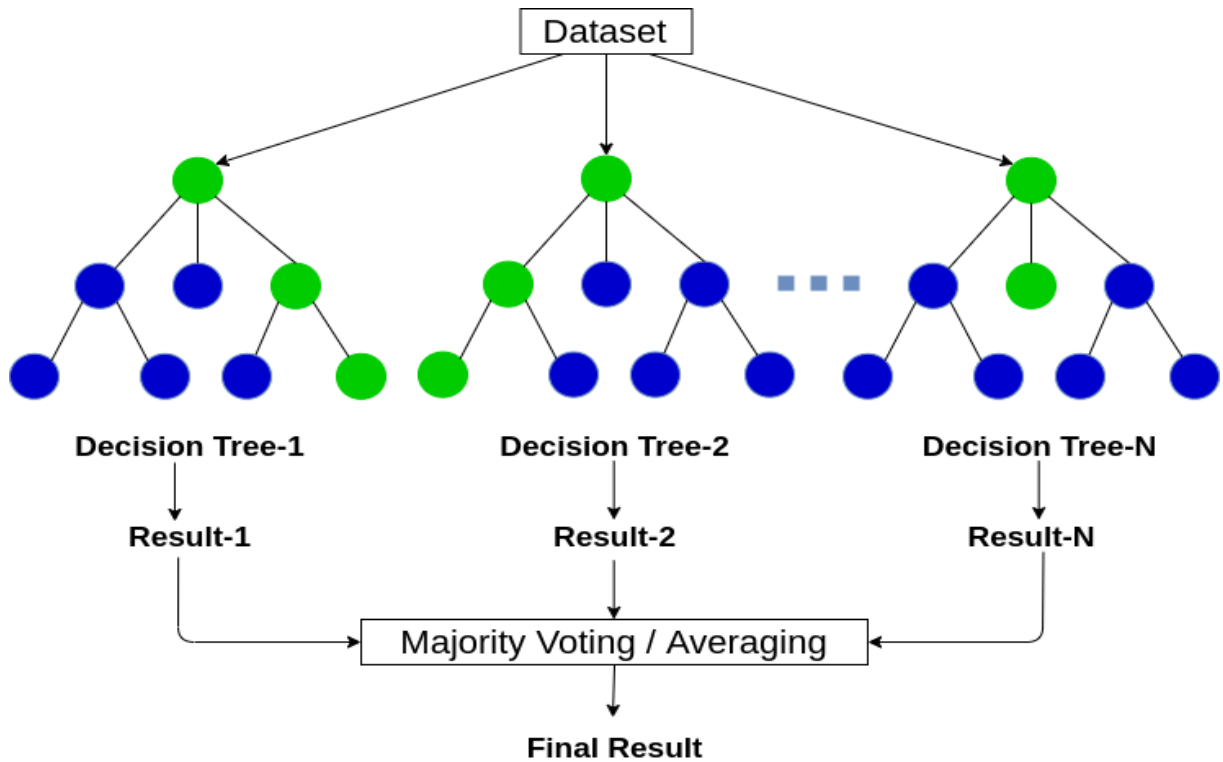


Figure 4: Random Forest Algorithm.

3.5 Naive Baye's Classifiers (NB)

Naive Baye's Classifier is a supervised classification algorithm based on two fundamental ideas: the first is the naive assumption that all features are conditionally independent of each other, and the second is the use of Bayes' theorem which is discovered by Thomas Bayes in the 18th century and defined by [26] [11]:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (12)$$

A and B are events; $P(A)$ is the probability of observing event A, and $P(B)$ is the probability of observing event B. $P(A|B)$ is the conditional probability of observing A given that B was observed [11]. In machine learning, the Naive Baye's Classifier is based on Bayes' theorem that is reworded and given the following more natural formula for a classification task [25] :

$$P(Y|X_1, X_2, \dots, X_j) = \frac{P(X_1, X_2, \dots, X_j|Y)P(Y)}{P(X_1, X_2, \dots, X_j)} \quad (13)$$

Where, $P(Y|X_1, X_2, \dots, X_j)$ is the posterior, the probability that an observation is class y given the observation's values for the j features, X_1, X_2, \dots, X_j . $P(X_1, X_2, \dots, X_j|Y)$ is the likelihood of an observation's values for features, X_1, X_2, \dots, X_j , given their class, y . $P(y)$ is the prior, the probability of class y before looking at the data. $P(X_1, X_2, \dots, X_j)$ is the marginal probability and is the probability of observing a particular feature in the training set; it is constant for all dataset. That is why the comparison of an observation's posterior values for each possible class, is focused on the numerators of the posterior for each class. For each observation, the class with the greatest posterior numerator becomes the predicted class, \hat{y} .

3.6 Artificial neural networks (ANN)

Artificial Neural Networks (ANN) are a family of machine learning techniques which are inspired by the biological neural networks that constitute the human brain [18]. By using a simple node called formal neuron carrying a value between 0 and 1, called activation, we can create a network composed of neurons organized in three different layers: the input layer, the hidden layers and the output layer shown in the figure 5. The neurons of a layer are connected to the neurons of the adjacent layers by links carrying weights which play an important role in the training of the network. There is a diversity of architectures and training algorithms for neural networks. The Multi-Layer Perceptron (MLP) invented by Frank Rosenblatt at the Cornell Aeronautical Laboratory in the late 1950s [11], is considered as the first simple and complete neural network [26]. This network, based on the concept of gradient backpropagation as a training algorithm, becomes one of the most used and productive model in various domains, such as image recognition, data classification.

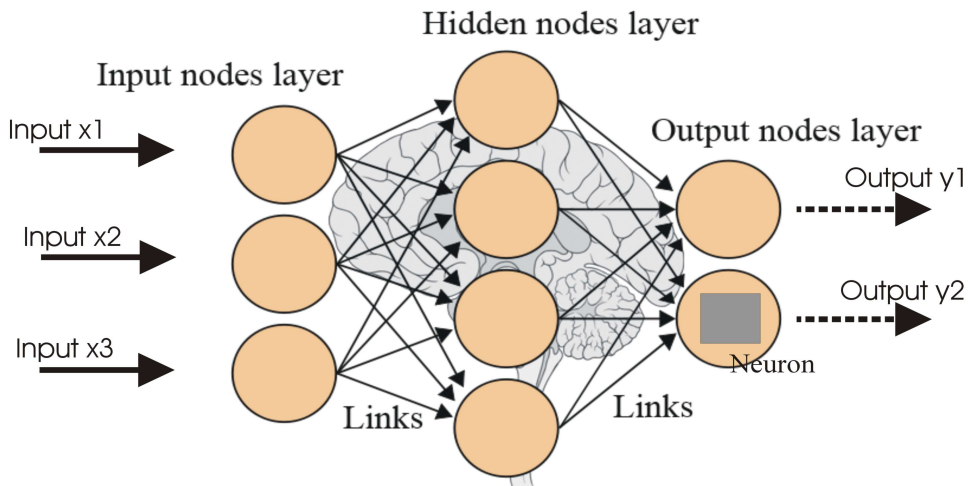


Figure 5: Artificial Neural Network Structure.

IV METHODS

4.1 Face recognition databases

For the implementation of a facial recognition system and its evaluation, face images databases are needed. Among those available, some have aroused a growing interest among many researchers. The GRIMACE database contains 360 pictures of 18 different persons. All the images have a resolution of 200×180 pixels and same lighting conditions [32]. The ORL database (Olivetti Research Laboratory) has been gathered from 1992 to 1994 in AT & T laboratory of Cambridge University in England [12] [32]. It includes 400 pictures of 40 distinct persons, with a resolution of 112×92 pixels and contains slight variations in illumination, in pose, and facial details [25]. The YALE image dataset is designed at the CVC (Center for Computational Vision and Control) of Yale University in a fully controlled environment [20]. It is composed of 165 gray-scale frontal images of 15 persons under varying illumination. All the images of these databases are taken in different facial expression situations.

4.2 Image processing and feature extraction techniques

These steps of preprocessing and feature extraction are very important because they directly influence the performance of the systems. For these tasks there are several efficient techniques among which the following are used. The Discrete Wavelet Transform (DWT) which is used to compress the images by eliminating the redundancies which mobilize a great quantity of computing resources. Principal Component Analysis (PCA) is generally considered as the main approach to reduce the dimension of dataset. It was introduced by M. A. Turk and M. P. Pentland in 1991 [1] [6]. Linear Discriminant Analysis (LDA), which appeared with the work of Belhumeur in 1997 at Yale University in the USA [6]. It is a good dimension reduction technique always used before a classification algorithm. The GABOR filter was introduced by Dennis Gabor in 1946, then improved in 1980 by John G. Daugman [9]. This filter is used to extract useful local facial features from an image. It is efficient and robust to illumination variations and geometric transformations such as facial expression.

V METHODOLOGY

To perform the experiments and evaluate the performance of the face recognition systems, an experimental methodology in different steps is used.

In the first step, we prepare the data on the ORL, YALE and GRIMACE image databases by dividing them into two distinct sets: the training set and the testing set. The first one will be used to train the model and the second to evaluate it.

In the second step, algorithms such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Gabor filter associated with PCA, Discrete Wavelet Transform (DWT) associated with PCA, are used to preprocess and extract the face feature vectors.

In the third step, classification is done through the following algorithms: SVM, KNN, RF, LR, NB, and ANN. At this step, the model is well trained, saved and able to make the prediction.

At last, models need to be evaluated using the testing dataset by following the same process of image preprocessing and feature extraction. Then, created models take as arguments these feature vectors and predict classes from them. The diagram of figure 6 explains the experimental methodology graphically.

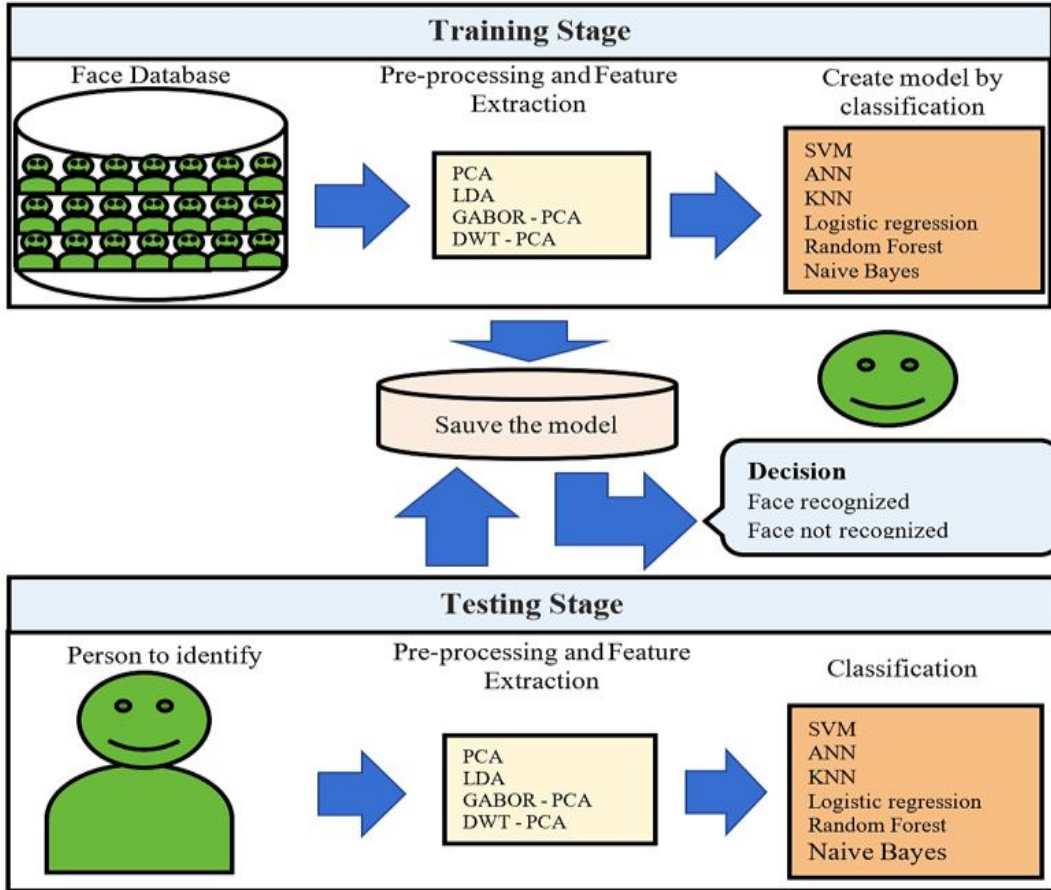


Figure 6: Diagram of the experimental methodology.

VI PERFORMANCE MEASURES

Three performance measures will be simultaneously used to evaluate the quality of the machine learning models. First, the recognition rate which is the quotient of the number of correct classifications by the total number of face images tested [14]. Second, the confusion matrix that counts how many times an observation from a class A was classified into a class B [26]. Several measures are defined based on the confusion matrix. One is the recall, also called sensitivity, which is the True Positive rate. Another measure often associated with this matrix, the precision, which is the rate of correct predictions among the positive predictions. The simultaneous use of these two indicators gives a good measure called F-measure: harmonic mean of precision and recall. It is defined as:

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (14)$$

The ROC curve plots graphically the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings [26]. In machine learning for model comparison, researchers focus, most often, on the use of the Area Under the Curve (AUC). In this paper, the average value of these three performance measures is used.

VII RESULTS

Each system is evaluated on the three databases. Its performance is the average value of the accuracy rate, the f-measure and the area under the ROC curve (AUC), defined by the following formula:

$$Performance = \frac{Accuracy + F_measure + AUC}{3} \quad (15)$$

7.1 Classification methods based on ORL Database

The experimental results on the ORL database are shown in Figures 7, 8, 9, 10 and the average rates in Table 1.

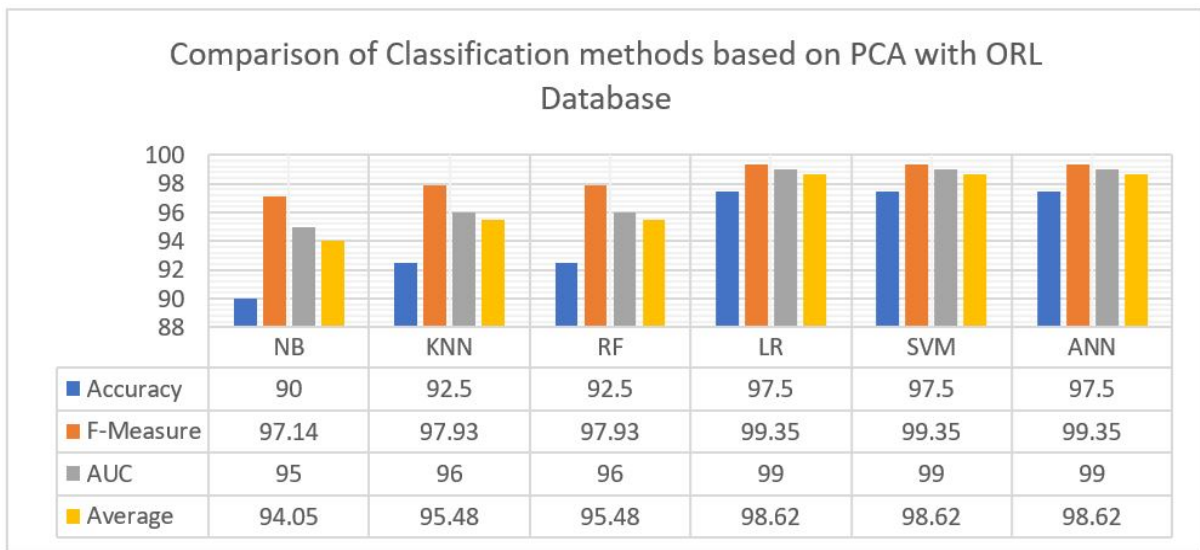


Figure 7: Performance of some classifiers on ORL database using PCA.

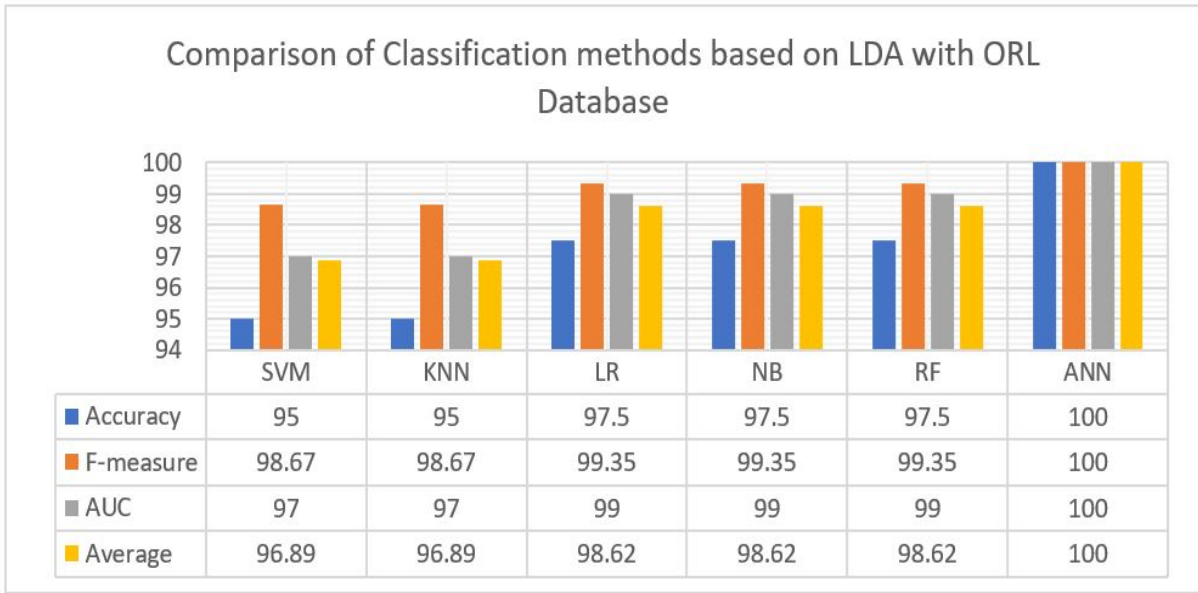


Figure 8: Performance of some classifiers on ORL database using LDA.

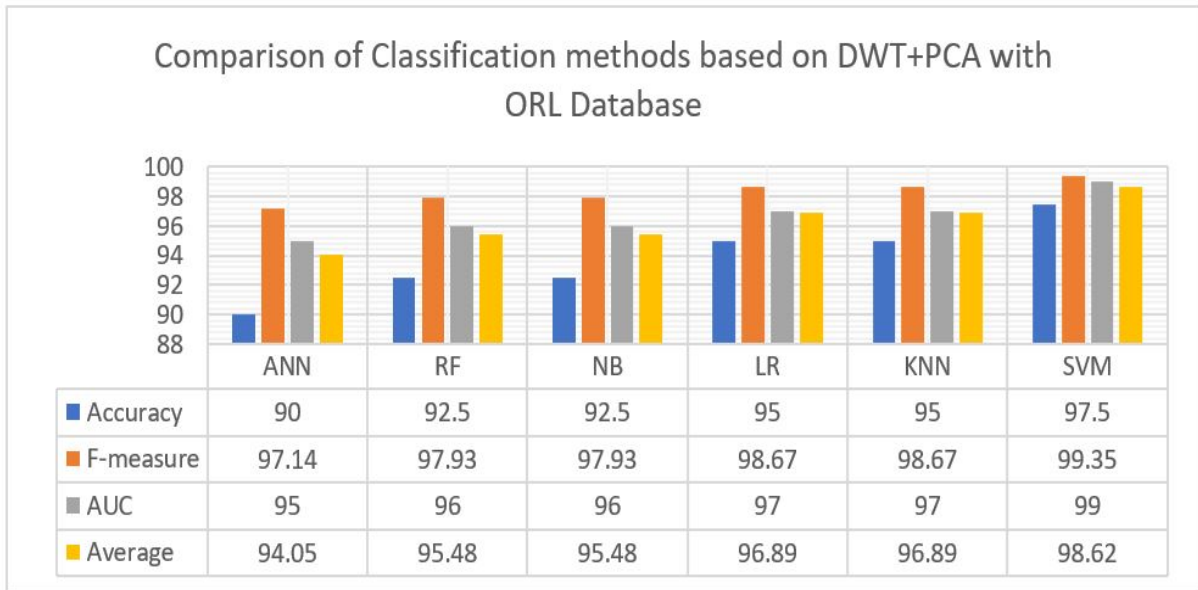


Figure 9: Performance of some classifiers on ORL database using DWT and PCA.

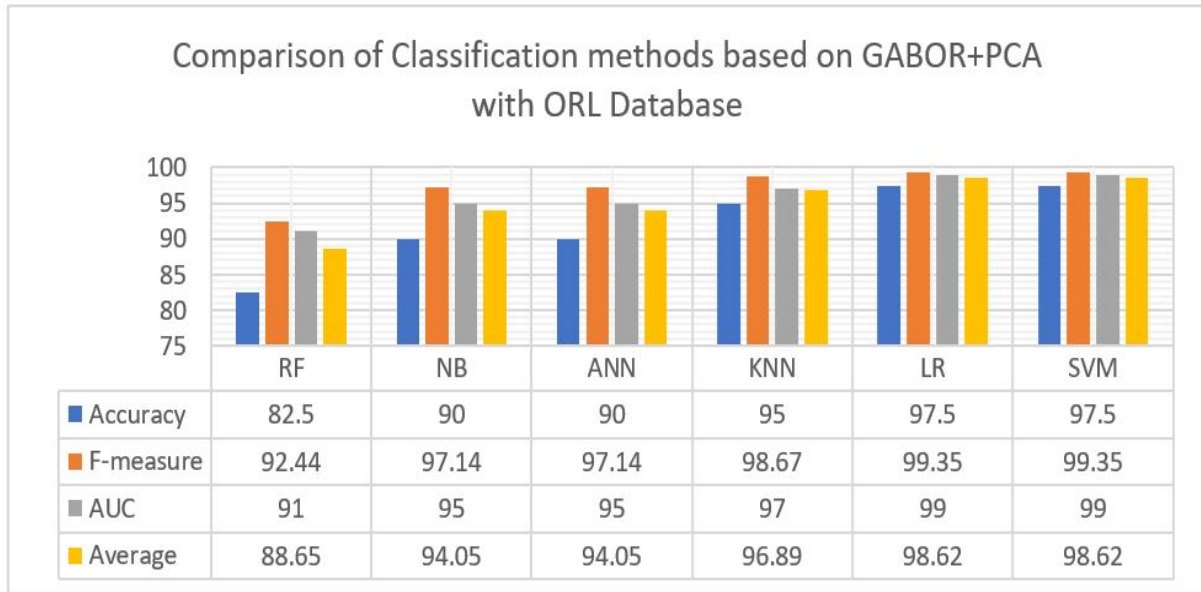


Figure 10: Performance of some classifiers on ORL database using GABOR and PCA.

The table 7 summarizes the average of the three performance measures for different classifiers associated with the different preprocessing and extraction techniques.

ALGORITHM	PCA	LDA	DWT+PCA	GABOR+PCA	Average (%)
RF	95.48	98.62	95.48	88.65	94.56
NB	94.05	96.89	95.48	94.05	95.12
KNN	95.48	96.89	96.89	96.89	96.54
ANN	98.62	100	94.05	94.05	96.68
LR	98.62	96.89	96.89	98.62	97.76
SVM	98.62	96.89	98.62	98.62	98.19

Table 7: : Comparative Analysis of Face Recognition Approaches on ORL Database using different classifiers with an average value of the F- measure, the accuracy and the AUC.

Support vector machines (SVM) outperform the others, with an overall average performance of 98.19%. It is immediately followed by logistic regression (LR) with an average of 97.76%, then artificial neural networks (ANN) with an average of 96.68%. In this work Random Forest algorithm achieved the lowest performance of 94.56%. From the discussion it is clear that, the SVM classifier has comparatively higher performance than the other classifiers.

7.2 Classification methods based on YALE Database

The comparison of the performance of the six classifiers are represented in Figures 11, 12, 13 and 14. It is clearly observed from these Figures that the techniques namely NB, KNN, LR, ANN, SVM, and RF give significant results.

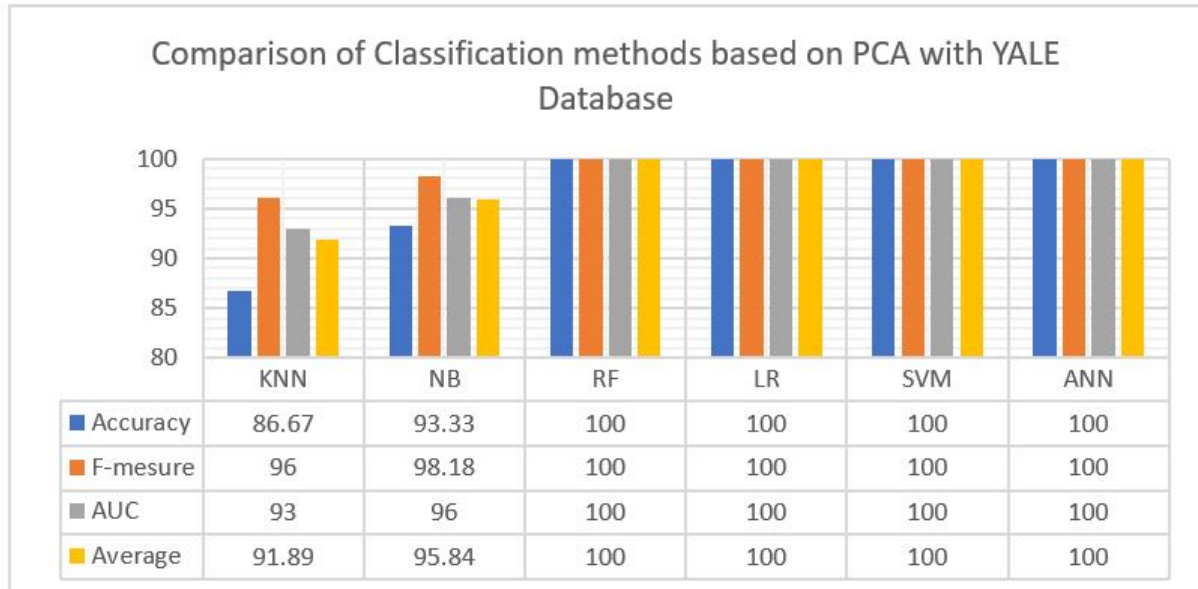


Figure 11: Performance of some classifiers on YALE database using PCA.

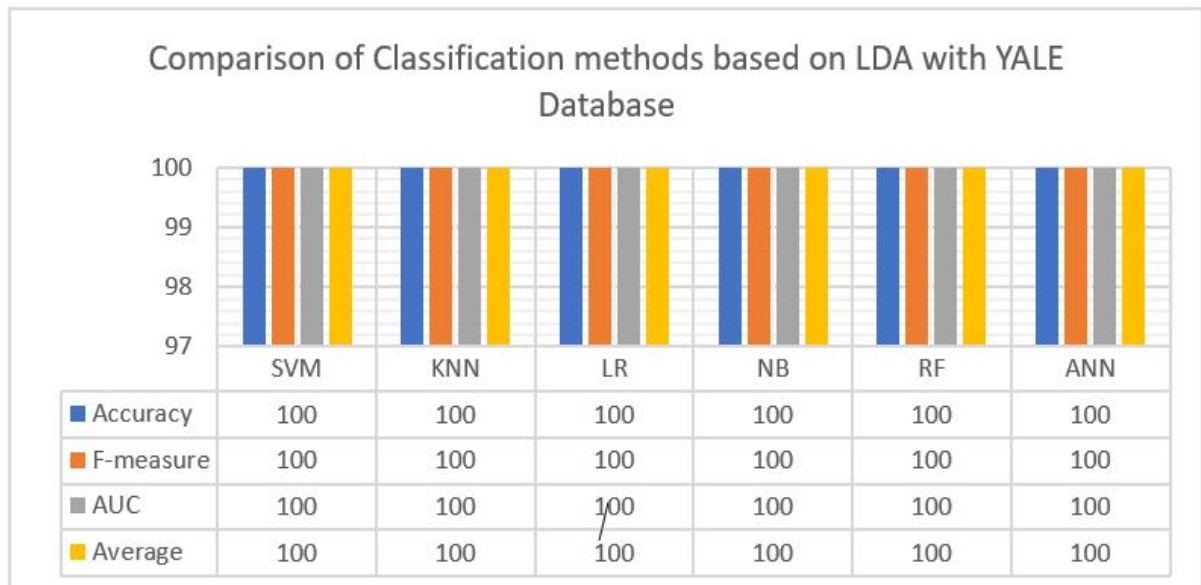


Figure 12: Performance of some classifiers on YALE database using LDA.

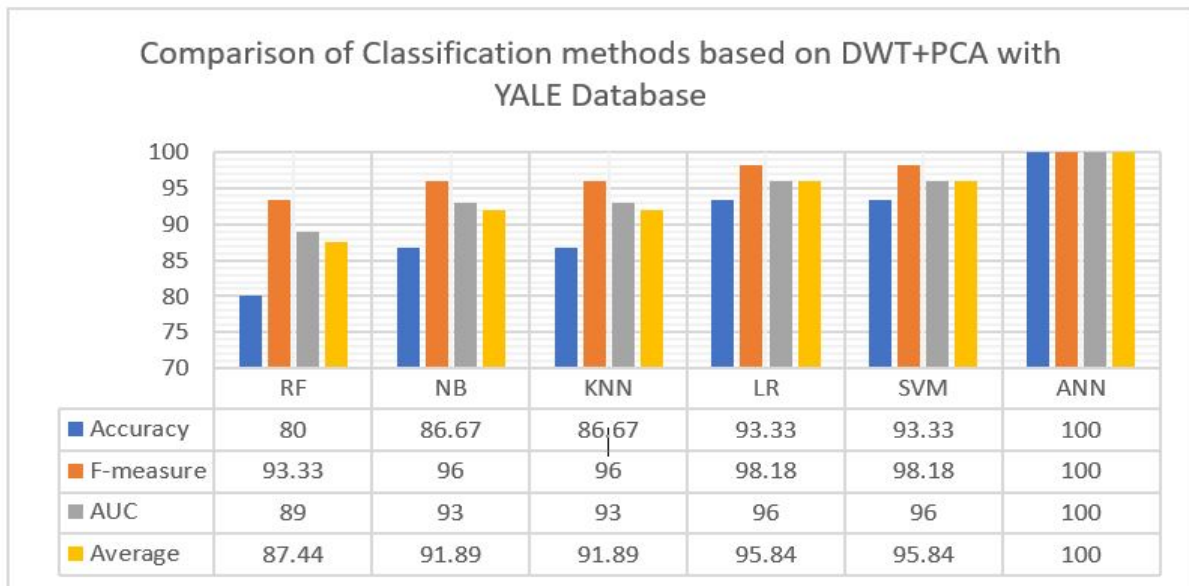


Figure 13: Performance of some classifiers on YALE database using DWT and PCA.

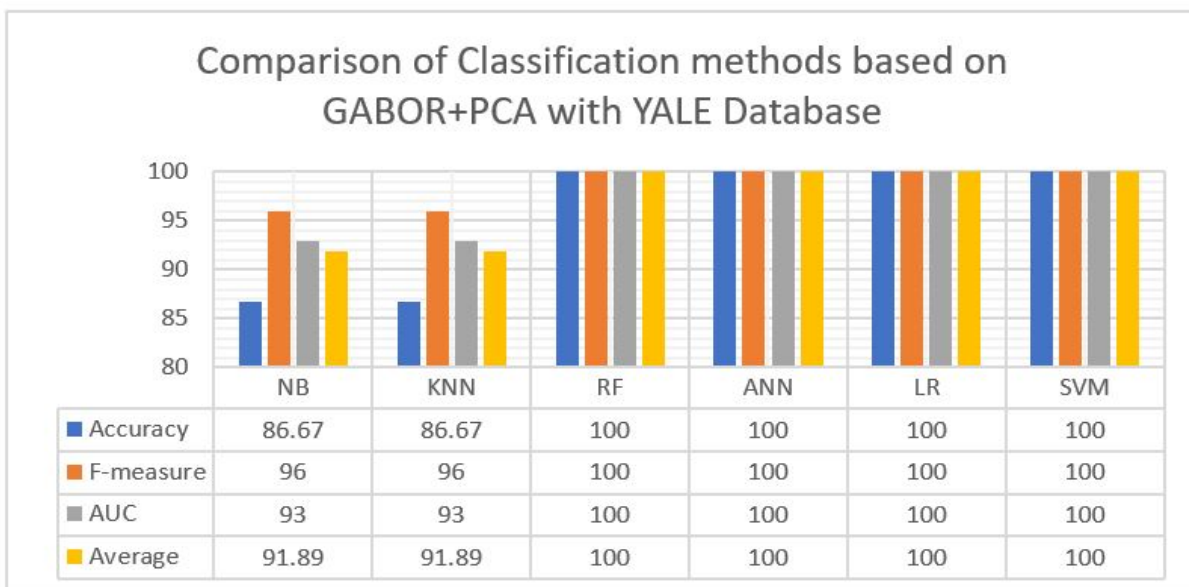


Figure 14: Performance of some classifiers on YALE database using GABOR and PCA.

The table 8 recapitulates the average classification accuracy of each classifier associated with the different preprocessing and extraction techniques. Then, these methods are compared with each other.

ALGORITHM	PCA	LDA	DWT+PCA	GABOR+PCA	Average (%)
RF	91.89	100	91.89	91.89	93.92
NB	95.84	100	91.89	91.89	94.91
KNN	100	100	87.44	100	96.86
ANN	100	100	95.84	100	98.96
LR	100	100	95.84	100	98.96
SVM	100	100	100	100	100

Table 8: : Comparative Analysis of Face Recognition Approaches on YALE Database using different classifiers with an average value of the F- measure, the accuracy and the AUC.

On the YALE database, artificial neural networks ANN outperform all other classification algorithms by achieving a perfect average score of 100%, as shown in Table 8.

7.3 Classification methods based on GRIMACE Database

For all algorithms, the average performance rate is 100%, except the Naive Bayes' Classifier associated to the combination of DWT and PCA which gives 90.59%, as illustrated in the histogram of the Table 9.

ALGORITHM	PCA	LDA	DWT+PCA	GABOR+PCA	Average (%)
RF	100	100	100	100	100
NB	100	100	95.59	100	97.65
KNN	100	100	100	100	100
ANN	100	100	100	100	100
LR	100	100	100	100	100
SVM	100	100	100	100	100

Table 9: : Comparative Analysis of Face Recognition Approaches on GRIMACE Database using different classifiers with an average value of the F- measure, the accuracy and the AUC.

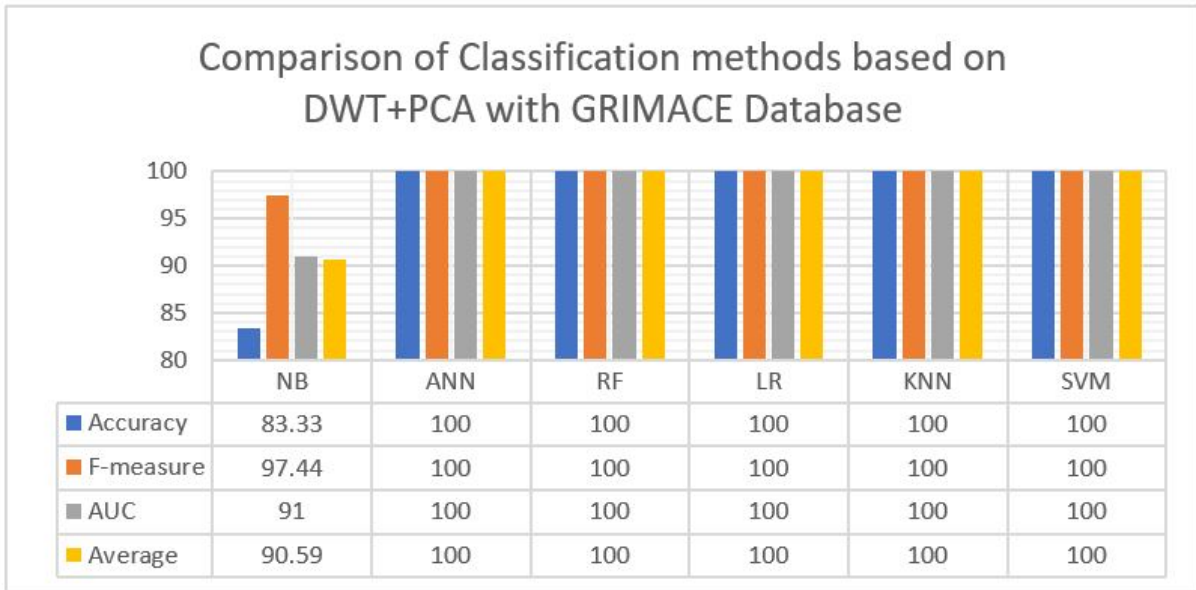


Figure 15: Performance of some classifiers on GRIMACE database using DWT and PCA.

With the Grimace database, the summary in the table 9 shows the importance of the image acquisition stage. The Grimace database does not bring us any difference from the performance point of view, necessary in establishing a real comparison of the machine learning algorithms.

7.4 Comparison of classifications methods and discussion

For the three databases studied: ORL, YALE and GRIMACE, the final results obtained are shown in the Table 10.

ALGORITHM	ORL	YALE	GRIMACE	Average (%)
NB	95.12	94.91	97.65	95.89
KNN	96.54	93.92	100	96.82
RF	94.56	96.86	100	97.14
ANN	96.68	100	100	98.89
LR	97.76	98.96	100	98.91
SVM	98.19	98.96	100	99.05

Table 10: : Comparative Analysis of Face Recognition Approaches on ORL, YALE, GRIMACE Databases using different classifiers with an average value of the F- measure, the accuracy and the AUC of the ROC curve..

In this paper, our study performed a validation of the performance of the machine learning techniques applied in face recognition, as presented in Table 4. The experimental results show that support machine vector model achieved 99.05% average performance. We note that this best result corroborates that of the literature in which SVMs also give the best classification accuracy rate as 98.80%. Also, LR, ANN, RF, KNN and NB produced respectively 98.91%, 98.89%, 97.14%, 96.82% and 95.89% as average performance rates. This ranking is roughly equal to what we found in the literature. In the latter, after SVM, we found that ANN, LR, RF, KNN and NB provided the following results in order: 97.27%, 96.44%, 96.15%, 95.79%, 94.82%. This small difference is probably due to the imbalance in the dataset used but also to the unequal conditions and criteria for comparison in the literature.

VIII CONCLUSION

We are in a favorable context for the development of facial recognition systems. The technological evolution, the apparition of powerful machines in calculation and treatment, the advent of artificial intelligence, are the important factors which allowed the researchers to direct their work towards this topic. Thus, many approaches using machine learning algorithms and giving different results have emerged in the recent literature [14] [10] [8] [4] [22]. That is why, this paper found it necessary to search for the best performing method by a comparative study of classification algorithms. Using a specific methodology, many systems are implemented with different supervised machine learning techniques such as: SVM, LR, ANN, RF, KNN, NB. For the evaluation of each technique on the ORL, Yale and Grimace databases, three performance measures are used simultaneously by calculating their average value: recognition rate, f-measure of the confusion matrix, and area under the ROC curve. The results obtained are good. And, the SVM, LR, ANN, RF, KNN, NB classifiers produced respectively the following average performances: 99.05%, 98.91%, 98.89%, 97.14%, 96.82% and 95.89%. SVM take the first place in the ranking. Thus, the future perspectives are the research of the best parameterization of the SVMs, and the research of the best methods of preprocessing and extraction of the features, in order to increase the performances of this classifier.

REFERENCES

- [1] M. A. Turk and A. P. Pentland. "Face Recognition Using Eigenfaces". In: *IEEE, Vision and Modeling Group, The Media Laboratory Massachusetts Institute of Technology* (1991), pages 586–591.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection". In: *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* 19.7 (1997), pages 711–720.
- [3] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11 (1998), pages 2278–2324.
- [4] X. Wang and X. Tang. "Bayesian Face Recognition Using Gabor Features". In: *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications* (2003), 70–73.
- [5] A. Eleyan and H. Demirel. "PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier". In: *Springer-Verlag* 4105 (2006), pages 199–206.
- [6] N. MORIZET, F. A. Thomas EA Florence ROSSANT, and A. AMARA. "Revue des algorithmes PCA, LDA et EBGM utilises en reconnaissance 2D du visage pour la biometrie". In: *Institut Supérieur d'Electronique de Paris (ISEP), Departement d'Electronique* (2006), pages 1–11.

- [7] M. M. K. Mayank Agarwal Nikunj Jain and H. Agrawal. “Face Recognition Using Eigen Faces and Artificial Neural Network”. In: *International Journal of Computer Theory and Engineering* 2.4 (2010), pages 624–629.
- [8] A. I. Salhi, M. Kardouchi, and N. Belacel. “Fast and efficient face recognition system using random forest and histograms of oriented gradients”. In: *2012 BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*. Darmstadt, Germany: IEEE, 2012, pages 1–11.
- [9] F. Bellakhdhar, K. Loukil, and M. ABID. “Face recognition approach using Gabor Wavelets, PCA and SVM”. In: *International Journal of Computer Science Issues (IJCSI)* 10.3 (2013), pages 201–207.
- [10] Q. Z. Changjun Zhou Lan Wang and X. Wei. “Face recognition based on PCA and logistic regression analysis”. In: *Elsevier* 125.20 (2014), 5916–5919.
- [11] G. Hackeling. *Mastering Machine Learning with scikit-learn*. Birmingham - Mumbai (UK): Packt Publishing Ltd., 2014.
- [12] P. S. Gharamaleki and H. Seyedarabi. “Face recognition using Eigen faces, PCA and support vector machines”. In: *European Journal of Applied Engineering and Scientific Research* 4.3 (2015), pages 24–30.
- [13] T. A. H and D. Sachin. “Dimensionality Reduction and Classification through PCA and LDA”. In: *International Journal of Computer Applications (0975 – 8887)* 122.17 (2015), pages 100–104.
- [14] K.S.Maheswari and C. Babu. “A Color Face Recognition Using PCA and KNN Classifier”. In: *International Journal and Magazine of engineering, Technology, Management and research* 2.9 (2015), pages 1110–1116.
- [15] V. W. Parate and P. Patel. “PCA, DCT and DWT based Face Recognition System using Random Forest Classifier”. In: *International Journal of Digital Application and Contemporary research* 3.6 (2015), pages 1–7.
- [16] E. Setiawan and A. Muttaqin. “Implementation of K-Nearest Neighbors Face Recognition on Low-power Processor”. In: *TELKOMNIKA* 13.3 (2015), pages 949–954.
- [17] K. He, X. Zhang, S. Ren, and J. Sun. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pages 770–778.
- [18] G. Kaur and H. Kaur. “Efficient Facial Recognition Using PCA-LDA Combination Feature Extraction with ANN Classification”. In: *International Journal of Advanced Research in Computer Science and Software Engineering* 6.7 (2016), pages 258–263.
- [19] P. S. Manju Bala and M. S. Meena. “FACE RECOGNITION USING LINEAR DISCRIMINANT ANALYSIS”. In: *International Journal of Electrical and Electronics Research* 4.2 (2016), pages 96–103.
- [20] B. Pathya and S. Nainan. “Performance Evaluation of Face Recognition using LBP, PCA and SVM”. In: *SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE)* 3.4 (2016), pages 85–88.
- [21] S. A. Syazana-Itqan K and S. N.M. “A MATLAB-Based Convolutional Neural Network Approach for Face Recognition System”. In: *Bioinformatics, Proteomics and Immaging Analysis* 2.1 (2016), pages 71–75.
- [22] S. S. H. Ab Waheed Lone and M. A. Ahad. “Face Recognition by SVM Using Local Binary Patterns”. In: *International Journal on Computer Science and Engineering* 9.5 (2017), pages 22–26.
- [23] A. Géron. *Machine Learning avec Scikit-Learn*. Paris: Dunod, 2017.

- [24] Vanlalhrauaia, Y. K. Singh, and N. D. Singh. “Binary Face Image Recognition using Logistic Regression and Neural Network”. In: *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*. Chennai, India: IEEE, 2017, pages 3883–3888.
- [25] C. Albon. *Machine Learning with Python Cookbook*. Gravenstein Highway North, Sebastopol: O’Reilly Media, Inc., 2018.
- [26] C.-A. Azencott. *Introduction au Machine Learning*. Malakoff (France): Dunod, 2018.
- [27] Z. B. Lahaw, D. Essaidani, and H. Seddik. “Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms”. In: *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*. 2018, pages 1–5.
- [28] H. Mady and S. M. S. Hilles. “Face recognition and detection using Random forest and combination of LBP and HOG features”. In: *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*. 2018, pages 1–7.
- [29] F. M. M. Shakir Fattah Kak and P. R. Valente. “Discrete Wavelet Transform with Eigenface to Enhance Face Recognition Rate”. In: *Academic Journal of Nawroz University (AJNU)* 7.4 (2018), pages 9–17.
- [30] L. R. Fleah and S. A. Al-Aubi. “A Face Recognition System Based on Principal Component Analysis-Wavelet and Support Vector Machines”. In: *Cihan University-Erbil Scientific Journal* 3.2 (2019), pages 14–20.
- [31] D. M. S. Meenakshi M. Siva Jothi. “Face Recognition using Deep Neural Network Across Variations in Pose and Illumination”. In: *International Journal of Recent Technology and Engineering (IJRTE)* 8.1S4 (2019), pages 289–292.
- [32] V. P. Sujitha and K. V. K. Kishore. “Scale Invariant Face Recognition with Gabor Wavelets and SVM”. In: 2019.
- [33] S. Sutarti, A. T. Putra, and E. Sugiharti. “Comparison of PCA and 2DPCA Accuracy with K-Nearest Neighbor Classification in Face Image Recognition”. In: *Scientific Journal of Informatics* (2019).
- [34] N. K. A. Wirdiani, P. Hridayami, N. P. A. Widiari, K. D. Rismawan, P. B. Candradinata, and I. P. D. Jayantha. “Face Identification Based on K-Nearest Neighbor”. In: 2019.
- [35] Z. Xie, J. Li, and H. Shi. “A Face Recognition Method Based on CNN”. In: *Journal of Physics: Conference Series* 1395.1 (2019), page 012006.
- [36] P. Rakshit, R. Basu, S. Paul, S. Bhattacharyya, J. Mistri, and I. Nath. “Face Detection using Support Vector Machine with PCA”. In: *2nd International Conference on Non-Conventional Energy: Nanotechnology & Nanomaterials for Energy & Environment (IC-NNEE) 2019 (Archive)* (2020).
- [37] M. S. Faridi, M. A. Zia, Z. Javed, I. Mumtaz, and S. Ali. “A Comparative Analysis Using Different Machine Learning: An Efficient Approach for Measuring Accuracy of Face Recognition”. In: *International Journal of Machine Learning and Computing* 11.2 (2021), pages 115–120.
- [38] E. H. L. N. Yohanssen Pratama Lit Malem Ginting and A. E. Rismanda. “Face recognition for presence system by using residual networks-50 architecture”. In: *International Journal of Electrical and Computer Engineering (IJECE)* 11.6 (2021), pages 5488–5496.
- [39] C. A. Benradi hicham and L. Abdelali. “Face recognition method combining SVM machine learning and scale invariant feature transform”. In: *10th International Conference on Innovation, Modern Applied Science and Environmental Studies (ICIES 2022)*. Volume 351. 2022, pages 1–5.

- [40] M. T. M. Zahraa Modher Nabat and S. A. H. Shnain. “Face Recognition Method based on Support Vector Machine and Rain Optimization Algorithm (ROA)”. In: *Webology* 19.1 (2022), pages 2170–2181.