A Comparative Analysis Using Different Machine Learning: An Efficient Approach for Measuring Accuracy of Face Recognition


Abstract—Feature extracting and training module can be done by using face recognition neural learning techniques. Moreover, these techniques are widely employed to extract features from human images. Some detection systems are capable to scan the full body, iris detection, and finger print detection systems. These systems have deployed for safety and security intensification. In this research work, we compare different machine learning algorithms for face recognition. Four supervised face recognition machine-learning classifiers such as Principal Component Analysis (PCA), 1-nearest neighbor (1-NN), Linear Discriminant Analysis (LDA), and Support Vector Machine (SVM) are considered. The efficiency of multiple classification systems is also demonstrated and tested in terms of their ability to identify a face correctly. Face Recognition is a technique to identify faces of people whose images are stored in some databases and available in the form of datasets. Extensive experiments conducted on these datasets. The comparative analysis clearly shows that which machine-learning algorithm is the best in terms of accuracy of image detection. Despite the fact, other identification methods are also very effective; face recognition has remained a major focus of research due to its non-meddling nature and being the easy method of personal identification for people. The findings of this work would be useful identification of a suitable machine-learning algorithm in order to achieve better face recognition accuracy.

Index Terms—Feature extraction, face recognition, linear discriminant analysis, 1-nearest neighbor, support vector machine, principal component analysis.

I. INTRODUCTION

Detection of face is assumed to be an active research area to spanning few rules i.e. picture processing, pattern identification, computer vision, subjective science, psychology & physiology and neuro science [1]. In explanation, a face identification system, a researcher put any random picture as a query in the database then it finds the people’s identification with respect to that random image. The human facial image shows a complicated, meaningful and multidimensional, visual analytic. Developing a computational model to recognize facial recognition is very difficult. It is complex of face detection because of many possibilities [2] that is present in human faces like: pose and expression of human face, position and orientation of image, face complexion, a wearing glasses, facial hairs, variability in camera, condition of light, and image pixel. Many enterprise systems using face identification are now obtained [3]. Statistical research work has been spotlight on video-image modeling and processing. Over 20 years, artificial neural networks were mostly used for resolving signal-processing theorems. Researchers determine many possible ideas of artificial neural networks [4]. Now the problem is to face the most suitable neural system models, which helps us in resolving sensible theorems. Face recognition is the undertaking to recognize as a recognized or unknown face an entity already identified. The question of facial popularity is often confused with the issue of facial recognition, however it is important to decide whether the “face” is known to use the facial database as a way to verify the face of this entry. Aim of this research work is to determine best possible approach to recognize image out of set of chosen approaches. Each algorithm recognizes the images based on open source face images dataset. Algorithm processes all the images and converts into required formats and then recognizes the images accordingly which is then measure by comparing results to original data set.

Face recognition systems may include many steps feature extraction, detection and image identification. Apart of these, recognition may also contain image border detection; position and analysis namely get the pre-processed excess face image from a query, whether the query is simple or complex, determining its position and dissolve the image. Face recognition describes the final output of the given face image. It is clear in face identification systems to evaluate the algorithms. Evaluation offers observation on the recognition algorithms and systems performed already. Important facts that read in old feature extraction, detection and image identification evaluation:

• Huge volume of test pictures are important for proper evaluation.
• The case must be significantly as much as possible to the picture which came in the properties being deliberate.
• Scoring must be completed in a way that’s follow the costs and other requirement of the system that changes the result from erroneous in identification.
• Error behavior should be rejected by studio, not only just the forced identification.
• Most useful way of evaluation is based as closely as possible on a specific theorem.
The objective of this work is to achieve better accuracy by comparing the performance of LDA, 1-NN, PCA and SVM machine learning algorithms for face recognition. The findings of this work would be useful for identification of a suitable machine-learning algorithm to achieve the better accuracy of face recognition. This paper includes the following sections: Section I, and Section II comprise a discussion of related work in terms of previous work related to face identification and detection. Section III related to proposed work. A popular reputation algorithm are studied in Section IV related to face recognition techniques, which are used in our proposed work. Section V related to result and analysis on face recognition algorithm during research work. This is followed by conclusion in Section VI and list of references present at the end of this work.

II. LITERATURE OVERVIEW

In the past, the researchers have developed several algorithms and strategies for facial recognition. These are discussed in the section below:

A. Independent Component Analysis Based on Face Recognition

Many current face popularity algorithms use face representation, which is determined by unsigned non-surveyed statistical strategies. These techniques typically find a hard and fast base pics and constitute of faces based on combination of linear snap shots [5]. Analysis of the main issue (PCA) is a well-known example of such techniques. PCA’s foundation pictures are most dependent on pairwise relationships within the picture database of pixels. Together with face credibility, in an undertaking where critical records can be stored in high order pixel relationships, it is advisable to expect appropriate foundation images to be judged by techniques that are sensitive to these excessive order statistics. One such approach is the independent aspect assessment (ICA), a generalization of PCA [5].

B. Eigenspaces

An Eigenspace, based primarily on facial recognition, is certainly one of the most successful methods for computational identification of faces in digital pixels [6].

Starting with the Eigen face Algorithm, different Eigen space-based totally, techniques for the recognition of faces have been proposed. This research work intends to provide an independent comparative have a take a look at amongst a number of the primary Eigen space-based definitely procedures. We consider that carrying out unbiased studies is relevant, because of the fact comparisons are commonly accomplished the usage of the operations of the research groups which have projected each method, which does no longer take into account equal operating situations for the algorithms. Very regularly, a completion among the capabilities of the studies corporations in preference to an evaluation among methods is achieve.

C. Elastic Bunch Graph Matching (EBGM)

EBGM is a popular proposed approach of face recognition.

In this painting, we support a wide range of flexible group graph matching and its existing variations in landmark model matching.

This work used information from FREET database in order to evaluate experiments and aforementioned strategies. We observe particle swarm optimization to beautify and show utility of facial graphs within flexible group graph matching techniques. The matching of the landmark model is mainly based on the Gabor wavelets, which feature excavation to find the landmarks [7]. We show that improvement can be achieved by combining the gray surface pro face. In addition, we leverage the well-known costs of hybridizing Gabor Violet's hybridizing resources with Egan Facial functions through principal components analysis, which can provide the complete information related to the appearance of face [4], [7].

D. Face Recognition Approaches

There are two important techniques for facial credibility problems: photometric (a full view of a face) and geometric (a feature-based). As the researcher's passion for facial recognition persists. Many excellent algorithms have been developed for face recognition. These identification techniques are divided into two basic methods:

1) Geometric

It is based entirely on the geometric court between facial features, or we can say dimensional arrangements about facial features. In this way, predictions of face, including eyes, mouth and nostrils, are placed first, the face is labeled with geographical distances and angles whose characteristics are described in Fig. 1.

2) Photometric Stereo

First-rate lighting fixtures are used to improve the appearance of a product from some of the pictures taken under the conditions as clearly indicated in Fig. 2. The shape of the recovered item is illustrated with the help of a gradient
map, made up of the usual rows of earth [9].

III. PROPOSED WORK

We used open set data set of Olivetti Research Laboratory, UK. Dataset contains 40 folders having 10 different images. For some articles, photos were captured of variant times, the light was slightly different, facial expressions (open eyes, closed eyes, smiles) and facial in details (glasses and without glasses) were all against the dark. Photos has been taken similar backgrounds and articles are in the right, front glasses) were all against the dark. Photos has been taken closed eyes, smiles) and facial in details (glasses and without light was slightly different, facial expressions (open eyes, For some articles, photos were captured of variant times, the UK. Dataset contains 40 folders having 10 different images. Cross-validation is usually applied in applied gadget studying to observe and decide a version for a given prognostic modelling trouble because of this: it is easy to perceive and quite easy to region into impact. Its ends in skills estimates unremarkably that have a remittent bias than extraordinary methods. Cross-validation can be a resampling manner wont to assess the system approximately to variance combinations of the exclusive variables.” perhaps PCA’s best strengths rectangular measure in its functionality for statistics cut price and interpretation [13].

A. A Step by Step PCA Clarification

1) Standardization

This can be done by changing each one and dividing by the standard deviation for each value mathematically of each variable.

\[ z = \frac{\text{value} - \text{mean}}{\text{standard deviation}} \]

Once the normalization done, the same scale will be transformed to all the variables.

2) Covariance matrix computation

The covariance matrix is a symmetric matrix that includes covariance with all possible pairs of initial variables. A 2-dimensional data set with 2 variables x and y covariance matrix, there is a 2x2 matrix as shown below:

\[
\begin{bmatrix}
\text{Cov}(x, x) & \text{Cov}(x, y) \\
\text{Cov}(y, x) & \text{Cov}(y, y)
\end{bmatrix}
\]

Covariance Matrix for 2-Dimensional Data

3) Compute the eigen vectors and values of above matrix to classify principal components

Eigen values and Eigen vectors are linear algebraic concepts that we need to calculate with a variable matrix to determine the basic components of the data. For example, for a 2-dimensional matrix, there are two variables and two eigen vectors with two corresponding eigen values.

4) Feature vector

What we do, is to choose whether to put all these components or subtract the least significance (lower eigen values), and to become a vector matrix with the rest of what we call a feature vector.

5) Reorganize the axes data along the principal components

\[
\text{FinalDataSet} = \text{FeatureVector}^T \ast \text{StandardizedOriginalDataSet}^T
\]

For instance, a 100x100 picture detail place containing a face might be terribly befittingly painted through merely 40 eigenvalues. Every Eigen value describes the significance each Manfred Eigen face in every photograph. Moreover, all interpretation operations will currently be completed the use of merely the forty eigenvalues to represent a face in preference to manipulating the 10 thousand values contained in an extremely 100x100 photograph. Not great is that this computationally less nerve-racking but the actual reality that the popularity records of numerous thousand.

6) Linear discriminant analysis (LDA)

Linear discrimination analysis (LDA) detects a maximum
linear transformation whereby the original data is transformed into a very small dimensional space. The goal of the LDA is to seek a linear transformation that maximizes class separation in low dimensional space. From the n independent variable of your dataset, LDA extracts \( p < n \) new independent variable that separate the most the classes of the dependent variable.

**B. A Step by Step Explanation of LDA**

a) To evaluate the separability between variant classes
b) To evaluate distance between mean and sample of every class, that is named variance within class
c) To build lower dimensional space in order to maximize between class variance and to reduce the variance within class

Several face class figures are used by both PCA and ICA. LDA uses the elegance records to find a green way to represent the face vector field (Fig. 4). This distinguished faces of people, but recognizes faces of the same character [14], [15].

1) **Support vector machine (SVM)**

The Support Vector Machine (SVM) was first heard in 1992, which was incorporated by the Boozer, Gavin, and Vapnik into the COLT-92. SVMs are a defining task to perform related monitoring to learn about classes and techniques used for regression [16]. SVM may be a supervised tool gaining knowledge of formulation that may be used for every category or regression challenge [17]. However, it is loosely used in classification troubles. during this formula, we have a tendency to plot every statistics item as a degree in n-dimensional area (wherein n is that the wide selection of capabilities you’ve got got) with the worth of every perform being the worth of a particular coordinate. Then, we have a tendency to perform category by means that of finding the hyper-aircraft that differentiates the 2 coaching totally. Support Vectors are just the coordinates of distinct observation. Support Vector Machine may be a boundary that glorious separates the 2 classes (hyperplane/ line) (Fig. 3).

![Fig. 3. Distinguishes the two classes in SVM.](image)

2) **Nearest neighbor algorithm (1-NN)**

In this algorithm , “the learner observes a sample S of labelled points \((X, Y) = (X_i, Y_i) \in [\sigma_i]\) , where \(X_i\) is a point in some metric space \(X\) and some countable \(Y_i \in Y\) is its label, for some finite labels set \(Y\). Being a metric space, \(X\) is equipped with a distance function \(\rho: X \times X \rightarrow \mathbb{R}\). Given a new unlabelled point \(x \in X\) to be classified, \(x\) is assigned the same label as its nearest neighbor in \(S\), which is \(\text{argmin}_{Y_i \epsilon Y} \rho(X_i, X)\). The 1-NN is a simple case of the k-nearest-neighbors (k-NN) algorithm with \(k = 1\) [18], [19]. In pattern recognition, the nearest neighbor's set of rules (k-NN) is a non-parametric technique used for sophistication and regression. In each cases, the enter includes the k closest training examples internal the feature area. The output is predicated upon on whether or not k-NN is used for classification or regression [20]. The k-NN, for example, is a form of basic study, in which the function is regionally closest and all counts are suspended until the classification.

“The kNN rule classifies each unlabeled example by the majority label among its k-nearest neighbors in the training set. Its performance thus depends crucially on the distance metric used to identify nearest neighbors. In the absence of prior knowledge, most kNN classifiers use simple Euclidean metric to measure the dissimilarities between examples represented as vector inputs” [21].

**V. RESULT AND ANALYSIS**

We conducted the performance of different machine-learning algorithms and found that LDA, 1-NN, SVM and PCA are providing the best accuracy result comparing to others. For this study, we used open set data set of Olivetti Research Laboratory in Cambridge, UK. There are 40 folders having 10 different images. For some articles, photos were taken at different times, the light being slightly different, facial expressions (open / closed eyes, smiles / smiles) and face details (glasses / no glasses) all against the dark. Photos is been taken similar backgrounds and articles of different, facial expressions (open / closed eyes, smiles / smiles) all against the dark. Photos is been taken similar backgrounds and articles.

A Step by Step Explanation of LDA

a) To evaluate the separability between variant classes
b) To evaluate distance between mean and sample of every class, that is named variance within class
c) To build lower dimensional space in order to maximize between class variance and to reduce the variance within class

Several face class figures are used by both PCA and ICA. LDA uses the elegance records to find a green way to represent the face vector field (Fig. 4). This distinguished faces of people, but recognizes faces of the same character [14], [15].

1) **Support vector machine (SVM)**

The Support Vector Machine (SVM) was first heard in 1992, which was incorporated by the Boozer, Gavin, and Vapnik into the COLT-92. SVMs are a defining task to perform related monitoring to learn about classes and techniques used for regression [16]. SVM may be a supervised tool gaining knowledge of formulation that may be used for every category or regression challenge [17]. However, it is loosely used in classification troubles. during this formula, we have a tendency to plot every statistics item as a degree in n-dimensional area (wherein n is that the wide selection of capabilities you’ve got got) with the worth of every perform being the worth of a particular coordinate. Then, we have a tendency to perform category by means that of finding the hyper-aircraft that differentiates the 2 coaching totally. Support Vectors are just the coordinates of distinct observation. Support Vector Machine may be a boundary that glorious separates the 2 classes (hyperplane/ line) (Fig. 3).

![Fig. 3. Distinguishes the two classes in SVM.](image)

2) **Nearest neighbor algorithm (1-NN)**

In this algorithm , “the learner observes a sample S of labelled points \((X, Y) = (X_i, Y_i) \in [\sigma_i]\) , where \(X_i\) is a point in some metric space \(X\) and some countable \(Y_i \in Y\) is its label, for some finite labels set \(Y\). Being a metric space, \(X\) is equipped with a distance function \(\rho: X \times X \rightarrow \mathbb{R}\). Given a new unlabelled point \(x \in X\) to be classified, \(x\) is assigned the same label as its nearest neighbor in \(S\), which is \(\text{argmin}_{Y_i \epsilon Y} \rho(X_i, X)\). The 1-NN is a simple case of the k-nearest-neighbors (k-NN) algorithm with \(k = 1\) [18], [19]. In pattern recognition, the nearest neighbor's set of rules (k-NN) is a non-parametric technique used for sophistication and regression. In each cases, the enter includes the k closest training examples internal the feature area. The output is predicated upon on whether or not k-NN is used for classification or regression [20]. The k-NN, for example, is a form of basic study, in which the function is regionally closest and all counts are suspended until the classification.

“The kNN rule classifies each unlabeled example by the majority label among its k-nearest neighbors in the training set. Its performance thus depends crucially on the distance metric used to identify nearest neighbors. In the absence of prior knowledge, most kNN classifiers use simple Euclidean metric to measure the dissimilarities between examples represented as vector inputs” [21].

**V. RESULT AND ANALYSIS**

We conducted the performance of different machine-learning algorithms and found that LDA, 1-NN, SVM and PCA are providing the best accuracy result comparing to others. For this study, we used open set data set of Olivetti Research Laboratory in Cambridge, UK. There are 40 folders having 10 different images. For some articles, photos were taken at different times, the light being slightly different, facial expressions (open / closed eyes, smiles / smiles) and face details (glasses / no glasses) all against the dark. Photos is been taken similar backgrounds and articles are in the right, front position. Results and discussions are as following:

1) **Classification using PCA then 1-NN with 5-fold cross validation**

For PCA we took the following steps:

a) Centerised the images.
b) Calculated Covariance matrix.
c) Calculated the Eigen Values & Eigen Vectors.
d) We sorted Eigen Values and Eigen Vectors in descending orders.
e) Took the first 70 principal components and plotted the data into new dimensions.

We used value for \(k=70\) as we saw that as the best value by experimenting and by looking at the Fig. 4 (Eigen values graph).
For 1-NN we took the following steps in our research work:

a) Calculated the Euclidean distance for each test data image with every train data image.
b) Set the distance in ascending order and take the first number as a result.
c) Compare the result with original class and computed the accuracy.

After the 5-fold cross validation, the average accuracy for 1-NN is **96.50%** (Fig. 5).

2) Classification using PCA with 5-fold cross validation

For PCA we took the following steps in our research work:

a) Calculated the Euclidean distance for each test data image with every train data image.
b) Set the distance in ascending order and take the first number as a result.
c) Compare the result with original class and computed the accuracy.

After the 5-fold cross validation, the average accuracy for PCA is **96.25%** in Fig. 6.

3) Classification using LDA with 5-fold cross validation

For LDA we took the following steps:

a) Calculated class wise mean vectors
b) Calculated in-Class Scatter Matrix
c) Calculated between-class scatter matrix
d) Calculated Eigen Values & Eigen Vectors
e) We sorted Eigen Values and Eigen Vectors in descending orders.
f) Took the first 70 principal components and plotted the data into new dimensions.

By using the 5-fold Cross-Validation method. We find the average accuracy of LDA **96.00%** (Fig. 7).

4) Classification using SVM with 5-fold cross validation

For SVM we took the following steps:

a) Generated Gramian Matrix
b) Calculated values of and solved quadratic equation
c) Determined support vectors
d) Calculated bias and weights
e) Prediction was done using one-vs-rest approach

After the 5-Fold Cross Validation, average accuracy for SVM was **98.00%** (Fig. 8).

5) Fold cross validation execution results

After execution of our program, we select the option for algorithm execution. By default, it comes in first fold. At this stage, it mix up data/shuffle data and make it more complex for the algorithm to train and test. The program do training in its first fold execution. After that it test the data and then it move towards the second fold. Then the same way second, third and fourth fold works. When it reach the fifth fold. It combine the result the all folds and calculate average of all folds.

During research, it is observe that accuracy rate of the algorithm SVM is **98.0%** obtained on face database (Fig. 9). It is the higher accuracy percentage while compare to PCA, LDA, 1-NN algorithm applied on dataset.

**VI. CONCLUSION**

This research work was a contribution to achieve the better accuracy among machine learning algorithms for face recognition. We compared the performance of four
machine-learning algorithms. These algorithms are LDA, 1-NN, SVM and PCA. For this study, we used open set data set of Olivetti Research Laboratory in Cambridge, UK. There are 40 folders having 10 different images. For some articles, photos were taken at different times, the light being slightly different, facial expressions (open / closed eyes, smiles / smiles) and face details (glasses / no glasses) all against the dark. Photos is been taken similar backgrounds and articles are in the right, front position. By using these algorithms, we extracted features from training sets and trained the classifiers. After that, we extracted features from testing sets and find the accuracy in percentage by using 5-fold cross validation on these four algorithms. The experimental result showed that SVM Algorithm achieved higher accuracy rate 98% of image recognition by comparing to other machine-learning algorithms such as 1-NN, LDA and PCA 96.25, 96 and 96.75 respectively. In future, we will consider machine-learning algorithms such as 1-NN, LDA and PCA. For this study, we used open set data set of Olivetti Research Laboratory in Cambridge, UK. There are 40 folders having 10 different images. For some articles, photos were taken at different times, the light being slightly different, facial expressions (open / closed eyes, smiles / smiles) and face details (glasses / no glasses) all against the dark. Photos is been taken similar backgrounds and articles are in the right, front position. By using these algorithms, we extracted features from training sets and trained the classifiers. After that, we extracted features from testing sets and find the accuracy in percentage by using 5-fold cross validation on these four algorithms. The experimental result showed that SVM Algorithm achieved higher accuracy rate 98% of image recognition by comparing to other machine-learning algorithms such as 1-NN, LDA and PCA 96.25, 96 and 96.75 respectively. In future, we will consider our work on 3-D face recognition in order to recognize faces from variant directions.

**CONFLICT OF INTEREST**

The authors DECLARE NO CONFLICT OF INTEREST.

**AUTHOR CONTRIBUTIONS**

Muhammad Shakeel Faridi wrote the paper, analyzed and tested the data. Azam Zia and Saqib Ali corrected the error and conducted the research. Zahid Javed and Imran Mumtaz collected the data. All authors approved the final version.

**REFERENCES**


Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).