International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.39, October 2022– www.ijais.org

A Comparison between PCA and SIFT Algorithms in Face Recognition Accuracy

Israa Abdulrauof Othman
SE Program
Elnasr Technical College
Omdurman-Sudan

Sallam Osman Fageeri
Dept. IS, College of EMIS
University of Nizwa
Nizwa, Sultanate of Oman

Ashraf Osman Ibrahim
Dept. CS, Faculty of CSIT
Al-zaiem Al-azhari University
Khartoum-Sudan

ABSTRACT
The systems that use facial recognition have contributed to the identification of criminals or wanted persons or even to control attendance and departure. However, the problem is whether the error rate is large. This may cause many problems. There are many useful aid such as SIFT and PCA. The purpose of this research is to analyze and compare the performance i.e. accuracy of principles of Comparison Analysis (PCA) and (SIFT) algorithms. The evaluation method used is a confusion matrix for measuring accuracy in precision, recall, F-measure, and success rate. Based on the comparative analysis, the (SIFT) algorithm gains the accuracy better by variation of compared to (PCA) algorithm in the implementation for 5 research data sets.

General Terms
Recognition, Accuracy, Algorithms, Classification

Keywords
Principles of Comparison Analysis, PCA, SIFT, ConfusionMatrix

1. INTRODUCTION
Human sees so many people's face repeatedly in his life. Whenever they meet someone, he remembers peculiar facial features of that person with the feature extraction process rather than whole face. So they will recognize facial image naturally. Of course, this feature extraction process is unconscious activity and is unknown process to us. In human face profiles, the shape and size of eyes, nose, mouth and their relationship have been commonly used as feature. With correctly extracted features it can be easily to recognize humane face. However, shadow hair, glasses, and noise or rotation of a face may distort the face shape [1].A number of studies using deep learning methods have claimed high performance in a significant number of tasks. These include image classification [2], natural language processing [3]. There are many algorithms that perform this purpose, but the problem lies in the accuracy of these algorithms. The percentage of accuracy may lead to the injustice of some people for example: when a criminal's identification system, a person on the basis that he is a criminal and in fact is not so, But the lack of precision of the system led to this big mistake for you to imagine how critical. It is clear that the main reason is not to test the accuracy of the algorithm in the form required so this paper aims to measure the accuracy of PCA and SIFT algorithms. The evaluation method used is a confusion matrix for measuring accuracy in precision, recall, F-measure, and success rate, Based on the comparative analysis.

2. RELATED WORK
Yau and Othman [4] have done the research by comparing different classification techniques using WEKAfor Breast Cancer. The model evaluation method is split percentage. Only 75% of the overall data is used for training and the rest is used for testing the accuracy of the classification. The measurements of accuracy are based on, incorrectly classified instances, correctly classified instances, and time consuming. The result is that the highest accuracy belongs to the Bayes network classifier. This research used as the data set only Breast Cancer and percentage split as the evaluation method. Confusion matrix has not been discussed in this research. Laterally, Fontana et al. [5] presented a big study that analyze and experiments different configurations of 6 ML-algorithms on detecting 4 smell types. For training, the authors considered a set of oracles composed of several examples of code smells manually validated by different programmers. Although, these oracles did not identify these programmers. As results, the authors reported that all evaluated techniques present a high accuracy. The highest one was obtained by two algorithms based on Decision Trees. The authors also affirmed that were necessary a hundred training examples to the techniques reach an accuracy of, at least, 0.95.

3. METHODOLOGY
SIFT and PCA algorithms were trained for the 5 groups of 20 people Photos. Two different images to each person were added, Confusion Matrix was also used for evaluation.

3.1 Principal Component Analysis (PCA)
The GLCM texture features resulted in a huge matrix of data. The PCA was developed to reduce this matrix. Moreover, PCA is able to find the optimum features in which it minimizes the execution time for the classification process [6].Generally, the first few principal components are accepted while the last few principal components are removed. For that reason a large data matrix can be reduced.

3.1.1 How PCA Works
Principal Component Analysis (PCA) is a learning algorithm that reduces the dimensionality (features number) within a dataset while still retaining as much information as possible. PCA reduces dimensionality by finding a new set of features which is called components, they are composites of the original features, but are uncorrelated with each other. The first component accounts for the largest possible variability in the data, the second component the second most variability, and so on.It is an unsupervised dimensionality reduction algorithm. It means that: labels that might be associated with the objects in the training dataset aren't used [6].Given the input of a matrix with rows each of dimension 1 * d, the data

is fractionate into mini-parts of rows and distributed among the training nodes (workers). Each worker then computes a summary of its data. The summaries of the different workers are then unified into a single solution at the end of the computation.

3.1.2 Modes

The Amazon SageMaker PCA algorithm uses either of two modes to calculate these summaries, depending on the situation:Regular: for datasets with moderate number of observations and features and sparse data.Un Regular: for datasets with both a large number of features and observations. This mode uses an approximation algorithm. In the last step of algorithms, it performs the singular value decomposition on the unified solution, from which the principal components are then obtained.

Mode 1: Regular

The workers jointly compute both because are 1 * d row vectors, is a matrix (not a scalar). Using row vectors within the code allows to derive efficient caching. The covariance matrix is computed as, and its top num_components singular vectors form the model.Moreover, computing and subtracting is avoided, If subtract mean is False. Use this algorithm when the dimension d of the vectors is small enough so that can fit in memory.

Mode 2: UN Regular

I all mini-batch of dimension b * d, (extra_components +num_components)*b matrixmust beinitialized randomly, then multiply by each mini-batch, to create a (extra_components + num_components) * d matrix. The sum of these matrices is computed by the workers, and the servers perform SVD on the final (num_components + extra_components) *d matrix.Thetop right num_components singular vectors of it are the approximation of the top singular vectors of the input matrix. Let = num_components + extra_components. It gives a mini- batch of dimension b * d, the worker draws a random matrix of dimension relies on whether the environment uses a dimension size and CPU or GPU, the matrix is either a random sign matrix where each entry is +-1 or a FJLT (fast Johnson Linden Strauss transform). The worker then computes and maintains. Also the worker maintains, the sum of columns of (T being the total number of mini-batches), and s, the sum of all input rows. Then, the worker sends the server B, h, s, and n (the number of input rows), after processing the entire shard of data.Denote the different inputs to the server as the server computes B, s, h, n the sums of the competent inputs. It then computes after that, and finds its singular decomposition value. The singular values and top-right singular vectors of C are used as the approximate solution to the problem.

3.2 SIFT Algorithm

SIFT analysis [7]. The SIFT algorithm greatest characteristic is scale invariance. In order to achieve scale invariance, SIFT uses a DoG (Gaussian Difference) function, to do convolution on an image. It achieves different scale images by changing σ. After that, it subtracts the images which are adjacent in the same resolution to get a DoG pyramid. The DoG function is a kind of a Gauss-Laplace algorithmimprovement. SIFT also compares each point with its adjacent 26 pixels, that is the sum of eight adjacent pixels nine pixels in the upper and lower adjacent layers and in the same layer. If the point is minimum or maximum, the location and scale of this point are recorded. Moreover, SIFT locates extreme points exactly and gets all extreme points of DoG scale-space. Then, it removes low contrast and unstable edge points. It further removes interference points, using 2 2 × Hessian matrix achieved from adjacent difference images. Next, in the scale of each key point, SIFT computes the strength of gradient and direction of every neighborhood. According to gradient directions, SIFT uses the summations as the gradient strengths of a keypoint and votes in histogram for every neighborhood. Actually, the main direction of this keypoint is defined as the direction whose gradient strength is maximal. Then, SIFT uses the keypoint as a center to choose an adjacent 16 16 × region. SIFT divides this region into 4 4 × sub-regions, and sums the gradient strength in each sub-region after the region is chosen. Then, SIFT uses eight directions in each sub-region to obtain an eight-dimensional vector. Thereby, SIFT gets a 128-dimensional feature description from 16 sub-regions, SIFT uses eight directions in each sub-region to generate an eight-dimensional vector. Thereby, SIFT gets a 128- dimensional feature description from 16 sub-regions, according to a certain order[8].

3.3 Dataset

Labeled Faces in the Wild is used in this research, On October 29th at ICCV 2019 in Seoul, the creators of LFW were honored with the Mark Everingham Award for service to the Computer Vision Community. Labeled Faces in the Wild is a public benchmark for face verification, also known as pair matching. No matter what the performance of an algorithm on LFW, it should not be used to conclude that an algorithm is suitable for any commercial purpose. There are many reasons for this. Here is a non-exhaustive list:

- Face verification and other forms of face recognition are very different problems. For example, it is very difficult to extrapolate from performance on verification to performance on 1:N recognition.
- Many groups are not well represented in LFW. For example, there are very few children, no babies, very few people over the age of 80, and a relatively small proportion of women. In addition, many ethnicities have very minor representation or none at all.
- While theoretically LFW could be used to assess performance for certain subgroups, the database was not designed to have enough data for strong statistical conclusions about subgroups. Simply put, LFW is not large enough to provide evidence that a particular piece of software has been thoroughly tested.
- Additional conditions, such as poor lighting, extreme pose, strong occlusions, low resolution, and other important factors do not constitute a major part of LFW. These are important areas of evaluation, especially for algorithms designed to recognize images “in the wild”.

3.4 Confusion Matrix

A confusion matrix states the accuracy of the solution to a classification problem. Given m classes, a confusion matrix is an m x m matrix where entry ci,j represents the tuples from D that were assigned to class Cj but where the correct class is Ci. Definitely, the best solutions will have only zero values outside the diagonal [9]. The confusion matrix is a useful tool for analyzing how well the classifier can recognize tuples of different classes. classes. Given m classes, a confusion matrix is a table of at least size m by m [6].Table 1 shows a confusion matrix for height classification. In confusion matrix, the columns represent the predicted classifications, and the rows represent the actual (true) classifications [10]. In a multiclass classification, a confusion matrix is necessary to be observed and changed into table of confusion as shown in Table 1.

Table 1. Table of Confusion

TRUE POSITIVE (TP)	FALSE NEGATIVE (FN)
FALSE POSITIVE (FP)	TRUE NEGATIVE (TN)

Precision= TP/TP+FP
Recall(TP Rate)
=TP/TP+FN
F-
Measure=2*(Precision*Recall/Precision+Recall)
Success Rate = TP+TN/P+N
Where: P = TP + FN and N = FP + TN

In this research, a system of ―Confusion Matrix for Accuracyis made by researchers based ontheformulasof(1), (2), (3), and (4). The system is built using Visual C# 2010. The purposes are to observe the confusion matrix, investigate the table of confusion, and measure the accuracy of each algorithm in precision, recall, F-measure, and success rate.

4. EXPERIMENT & RESULT

In this section, the results of the research on 5 data sets are presented in some tables and figures

Table 2. Precision Measurement

Data Sets	Algorithm	
	PCA	*SIFT*
1	3	1
2	3	1
3	4	2
4	3	1
5	2	1
Precision Average	3(%)	12 (%)

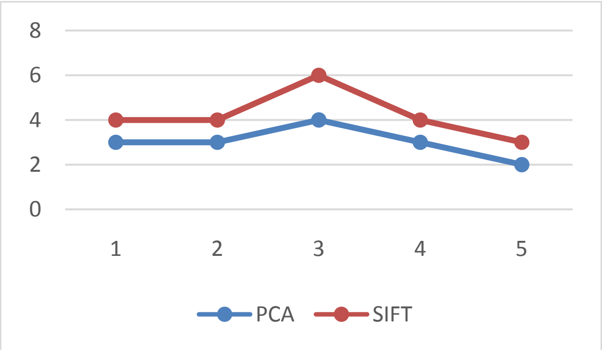

Fig 1: Comparison of Accuracy in Precision

From the Table 2 and Figure 1, it is clearly seen that SIFT gives better precision compared to PCA in all Evaluation data set. The deference in Precision Average between PCA and SIFT is 9%.

Table 3. Recall Measurement

Data Sets	Algorithm	
	PCA	*SIFT*
1	2	4
2	1	3
3	1	5
4	1	4
5	2	3
Recall Average	1.4 14(%)	3.8 38(%)

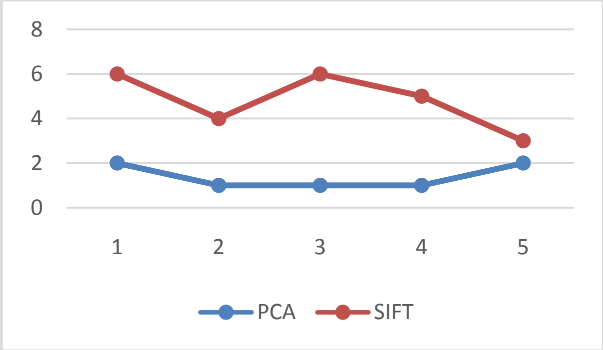

Fig 1: Comparison of Accuracy in Recall

From the Table 3 and Figure 2, it is clearly seen that SIFT gives better precision compared to PCA in all Evaluation data set.

The deference in Recall Average between PCA and SIFT is 24%.

Table 4. F-Measure Measurement

Data Sets	Algorithm	
	PCA	*SIFT*
1	1.3	3.4
2	1	3
3	1.3	4.4
4	1	3.4
5	1.3	2.4
F-Measure Average	1.18 11.8(%)	3.32 33.2(%)

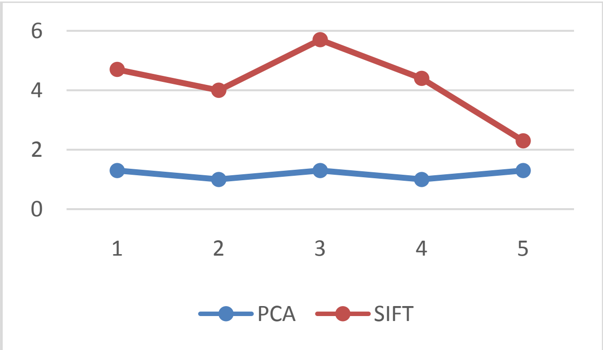

Fig 3: Comparison of Accuracy in F-Measure

From the Table 4 and Figure 3, it is clearly seen that SIFT gives better precision compared to PCA in all Evaluation data set.The deference in F-MEASURE Average between PCA and SIFT is 21.4%.

Table 5. Success Rate Measurement

Data Sets	Algorithm	
	PCA	*SIFT*
1	0.81	0.97
2	0.78	0.86
3	1.08	0.99
4	0.88	1
5	0.93	0.98
Success Rate Average	0.89 89(%)	0.96 96 (%)

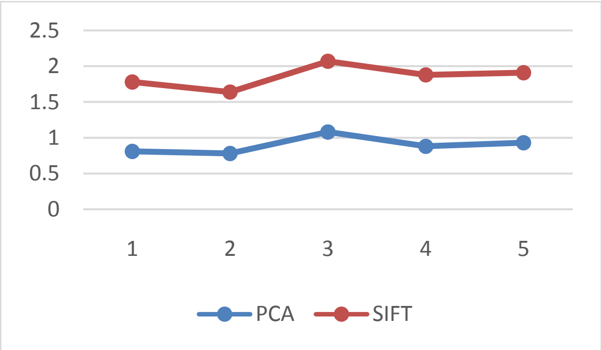

Fig 4: Comparison of Accuracy in Success Rate

From the Table 5 and Figure.4, it is clearly seen that SIFT gives better precision compared to PCA in all Evaluation data set. The deference in Success Rate Average between PCA and SIFT is 7%.

5. CONCLUSION

In conclusion, this paper explained the superiority of SIFT in terms of accuracy but there are still some efforts that can be made to compare the recognition algorithms of faces in other aspects such as speed and synchronization. In this research, the success rate of PCA and SIFT is above 80%. In comparative analysis, SIFT is having quite better results compared to PCA. It is giving accuracy with a variation of 7%-21.4% in 5 out of 5 datasets. Generally, the image size is not important for a PCA based face recognition system as long as the number of signatures before PCA-projection is more than the total number of sample images. Pose and Expression have minimal effect to the recognition rate while illumination has great impact on the recognition accuracy. As afuture work this research recommend to making an enhancement in Sift algorithm to be a real-time facial recognize algorithm. And Merging Sift and PCA algorithms to achieve high qualityin facialrecognitions, moreover dealing with illuminations on images in order to achieve satisfactory results on face recognition accuracy, can provide useful performance evaluation criteria for optimal design and testing of human face recognition systems.

6. REFERRENCS

[1] Mattew Turk and Alex Pentland," Eigenfaces for Recognition," SPIE Vol.1192 IRCVVIn (i989),22-32

[2] Wu, R.; Yan, S.; Shan, Y.; Dang, Q.; Sun, G. Deep image: scaling up image recognition. arXiv 2015, arXiv:1501.02876.

[3] . Torfi, A.; Shirvani, R.; Keneshloo, Y.; Fox, E. Natural language processing advancements by deep learning: A survey. arXiv 2020, arXiv:2003.01200.

[4] M.F. Othman, and T.M.S. Yau, ―Comparison of Different Classification Techniques using WEKA for Breast Cancer‖.F.Ibrahim,N.A.AbuOsman,J.Usman,andN.A. Kadri (Eds.). IFMBE Proceedings 15, 2007, pp. 520-523

[5] Fontana,F.A.,Mäntylä,M.V.,Zanoni,M.,andMarino, A. (2015). Comparing and experimenting machine learning techniques for code smell detection. Empirical Software Engineering.

[6] Amazon SageMaker-DeveloperGuide https://docs.aws.amazon.com/sagemaker/latest/dg/how-pca-works.html .10/11/2018 -3:50 PM

[7] MEASUREMENT SCIENCE REVIEW, Volume 13, No. 3, 2013 122 A Comparative Study of SIFT and its Variants Jian Wu1 , Zhiming Cui1 , Victor S. Sheng2 , Pengpeng Zhao1 , Dongliang Su1 , Shengrong Gong1 1 The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China, jianwu@suda.edu.cn 2 Department of Computer Science, University of Central Arkansas, Conway 72035,USA

[8] I.H. Witten and E. Frank, Practical Machine Learning Tools anda Techniques, Second Edition, USA: Elsevier Inc. & Morgan Kaufmann Publishers,2005.

[9] Huliman, Analisis Akurasi Algoritma Pohon Keputusandan k-Nearest Neighbor (k-NN), Tesis, Universitas Sumatera Utara, 2013.

[10] M.H. Dunham, DATA MINING Introductory and Advanced Topics, USA: Prentice Hall, 2003.