Fusion Effects:

Combining PCA and LDA to Optimize Face Recognition

by Paul Charles ***[Paul Charles claims sole authorship of this paper and all the content within it, any external information or knowledge contained in this paper is thoroughly referenced.]***

Abstract

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two of the most commonly used methods for dimensionality reduction. Both practices have gained widespread acclaim for yielding fairly accurate results for pattern recognition. Scientists and researchers are continuously trying to discover and implement newer methods and practices to yield better results for LDA and PCA. One method that is gaining extensive appeal is fusion, or multi-classifier system, in which better results are obtained by combining multiple classifiers. After researching the effects of multi-classifier fusion for image recognition of various individuals, it is concluded that multi-classifier fusion yields better results than individual classifiers[1]. There were two experiments conducted, Experiment A was a multi-classifier system (MCS) and Experiment B was the multi-instance system. As for the results, it showed that for multi-instance based system, at score level, PCA and LDA produced a ROC curve area (C.6) of 0.967 and 0.955, while individually they produced 0.940 and 0.945, which hints to success of fusion classifiers.

Introduction

One of the largest trends in industry is face recognition with biometrics [2]. There are several new technologies that utilize this, such as for identification fraud. A biometric pattern recognition system is one that extracts or parses unique features from an individual and identifies them based on those features such as facial, eyes, fingerprints, voice, etc. [2]. Much successful research has proven that recognition with biometrics has been successful [2]. As mentioned previously, multi-classifier systems have been very effective, and has proven to produce better results than that of single classifiers [3][4], however there are a number of challenges associated with them. One includes merging classifiers which have different representations of patterns to be classified [1][4]. As a result, coming up with a standard methodology where multiple pattern representations are used was a tedious process [1]. Another challenge was overlap in certain combination schemes. For example, in max classifier it would tend to show the results of one of the classifiers if the results were too high. This was displayed in the LDA results for max. This therefore created an overlap of results in the comparison figures [Figure A.1].

Method

There were a number of methods to execute the experiments:

- Import LDA and PCA vector space from previous project: This was completed in Matlab using matrix calculations
- Create labels: Implement label matrix that relates to imposter vs genuine
- EZRoc: Utilize the provided ezroc3 function in Matlab to obtain results.
- Experiment A: Combine scores for PCA and LDA using minimum (C.3), maximum (C.2), and average rules (C.4).

- Experiment B: Average the scores by comparing the test sample with all the templates of a particular identity (C.5).

Results and Discussion

Experiment A incorporated combining LDA and PCA using the minimum, maximum, and average computations. The results that were produced for this were fairly reasonable, with the ROC area (C.6) of 0.94478 (average), 0.9453 (min), and 0.9400 (max) [Table B.1]. While the single classifier for LDA and PCA was 0.9449 for LDA and 0.94013 for PCA. The same could be said about the EER. As a result, it can be assumed that the minimum value score, when combining PCA and LDA scores, produced the most ideal results. This is also supported from research by Kuncheva in which the minimum and maximum classifiers produced the lowest classifier error than that of the single classifier [3]. This is commonly used and these simple fusion methods were great for developing a good multiple classifier system, also known as an ensemble [3].

As for Experiment B, which was the multi-instance system for average scores, it was revealed that this procedure, for both LDA and PCA, gave better performance than that of a single classifier. The multi-instance for LDA produced 0.9549 and PCA produced 0.9671 for ROC area [Table B.2]. This makes sense in light of the research by Cano, in which multi-instance methods produce better performance than single classifiers [4]. As mentioned previously, these results were better than what was produced for single classifiers [Figure A.5]. As a result, this shows success of multi-instance classifiers performing better than single classifiers.

For Experiment A, the procedure of combining classifiers using simple methods is known as voting or ensembles [5, pp.492]. In this practice the base-learners use their outputs and combine them based on the result [5, pp.492]. The sum rule is the most widely used. The median rule is most resistant to noise, and the minimum/maximum rules will either give trends that are negative or positive [5, pp.493]. For the second experiment, it mimics a practice called boosting, in which multiple classifiers are compounded by increasing the probability of accuracy with the individual classifiers [5, pp.499]. This is also known to decrease the error significantly, which is what it did with the results.

Conclusion

Overall, the experiments were rather successful. Based on previous research, it seemed that the multi-classifier and multi-instance systems tended to produce highly accurate results when in comparison to the single classifiers. One interesting facet that was noticed is that the LDA produced higher results for Experiment A because of the fluctuation of varied data. As a result, one good implementation for future is to use min-max normalization in order to make the data more consistent. This was also shown in previous research for both max and min, and the resolution was to normalize the results through normal distribution [3]. As a result, utilizing multi-classifier systems is important and something that could be implemented more for image recognition.

Appendix A

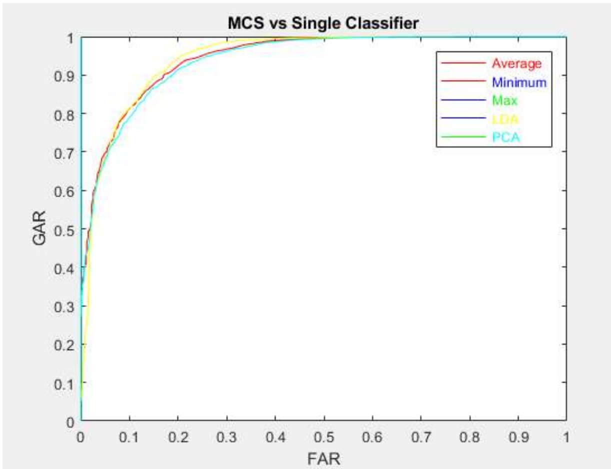

Figure A.1. Exp A: MCS Comparison

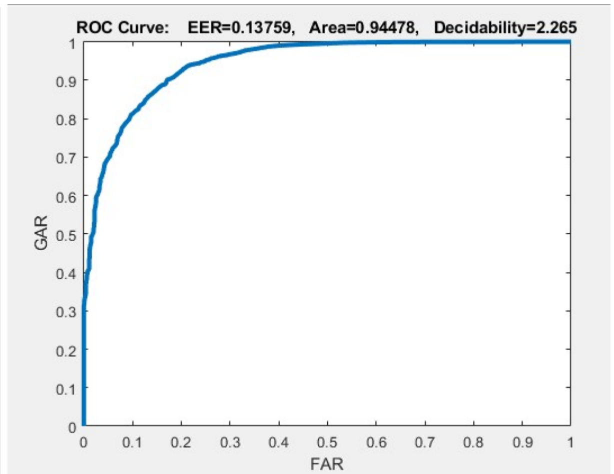

Figure A.2. Exp A: MCS Average Results

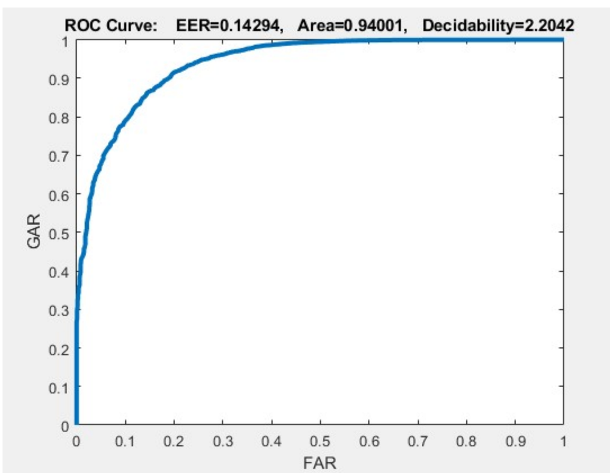

Figure A.3. Exp A: MCS Max Results

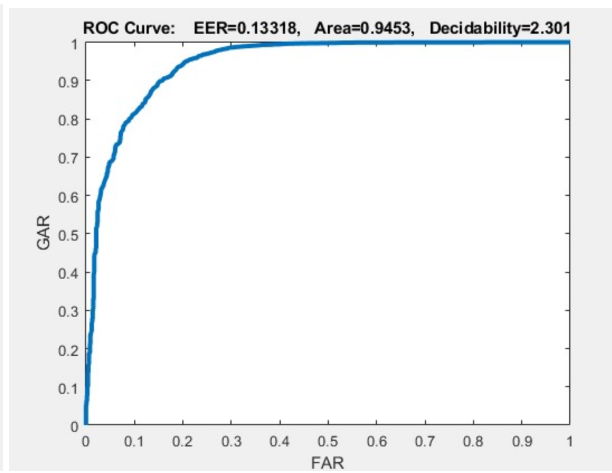

Figure A.4. Exp A: MCS Min Results

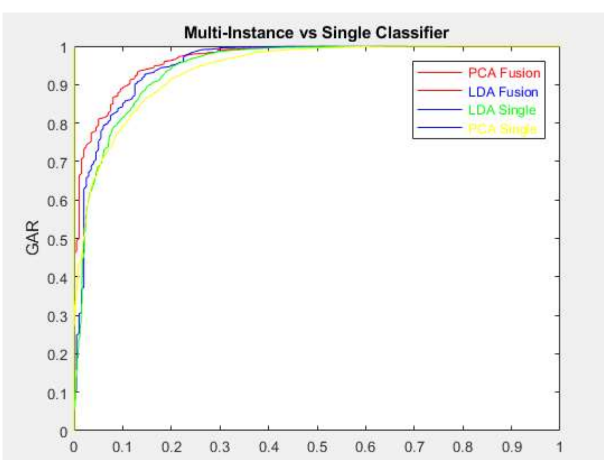

Figure A.5. Exp B: Multi-Inst Comparison

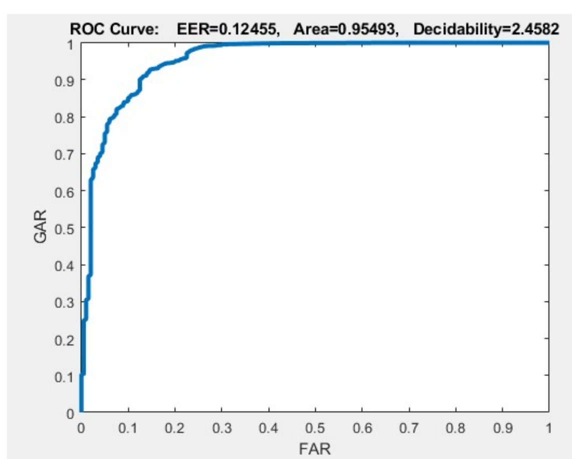

Figure A.6. Exp B: Multi-Inst LDA Fusion

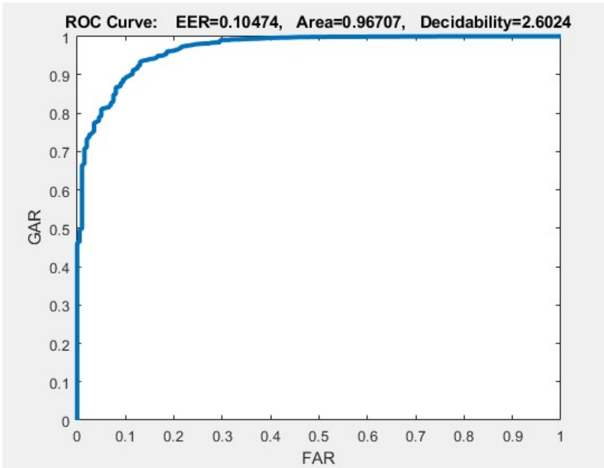

Figure A.7. Exp B: Multi-Inst PCA Fusion

***Note: For comparison graphs (Figure A.1 and Figure A.5), the legend color is noted by the color of the font not the line to the left of it**

Appendix B

Table B.1. Experiment A: Classifier Statistics

Classifier	ROC Area (FRR)	EER
Min	0.9453	0.13318
Avg	0.94478	0.13759
Max	0.9400	0.14294
PCA	0.94013	0.1421
LDA	0.9449	0.13351

Table B.2. Experiment B: Classifier Statistics

Classifier	ROC Area (FRR)	EER
LDA Fusion	0.95493	0.12455
PCA Fusion	0.96707	0.10474
LDA NO Fusion	0.9449	0.1421
PCA NO Fusion	0.94013	0.13351

Appendix C

$$di(x) = F\big(d1, i(x), \ldots, dL, i(x)\big), i = 1{,}2 \qquad (C.1)$$

Simple Fusion Model Diagram. F is fusion model from Kuncheva (see [3])

$$maxValue = \max(LDA_{Value}, PCA_{Value}) \qquad (C.2)$$

Equation that returns max value for LDA and PCA

$$minValue = \min(LDA_{Value}, PCA_{Value}) \qquad (C.3)$$

Equation that returns min value for LDA and PCA

$$meanValue = \frac{LDA_{Value} + \ PCA_{Value}}{2} \qquad (C.4)$$

Equation that returns mean (average) value for LDA and PCA

$$(LDA, PCA)Identity_{AverageFeatures} = \frac{(LDA, PCA)_1 + \cdots . + \ (LDA, PCA)_5}{5} \qquad (C.5)$$

Calculates average of all test samples for an identity for either LDA or PCA

$$ROC\ Area[FAR] = \frac{accepted\ attempts}{total\ attempts} \qquad (C.6)$$

Evaluation model to evaluate classifier performance

References

[1] Kittler, Josef and Hatef, Mohamad and Duin, Robert P.W., and Matas, Jiri, "On Combining Classifiers", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 20, no. 3, March 1998. [Online].

[2] Jain, Anil K. and Ross, Arun and Prabhakar, Salil, "An Introduction to Biometric Recognition". *IEEE Transactions on Circuits and Systems For Video Technology*, vol 14, no. 1, January 2004. [Online].

[3] Kuncheva, Ludmila I., "A Theoretical Study on Six Classifier Fusion Strategies". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 24, no. 2, February 2002. [Online].

[4] Cano, Albert, "An ensemble approach to multi-view multi-instance learning". *Knowledge-Based Systems*, vol 136, January 2017. [Online]. Available: http://www.elsevier.com/locate/knosys.

[5] Alpaydin, Ethem. *Introduction to Machine Learning*. Cambridge, Massachusetts: The MIT Press, Third Edition, 2014, Sections 17.4-17.7. pp 492-502