**Fusion Effects: Utilizing Multi-Classifier System with PCA and LDA for Face Recognition**

**by Paul Charles *[Paul Charles claims sole authorship of this paper and all the content within it]***

**Abstract**

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two of the most common methods utilized for dimensionality reduction. Both practices have gained widespread acclaim for yielding high accurate results for pattern recognition. Scientists and researches 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 attained with combining multiple classifiers. After researching the effects of multi-classifier fusion for image recognition of individuals people, it revealed that multi-classifier fusion yielded better results over individual classifiers[1]. There were two experiments conducted, oen with multi-classifier system (MCS) and multi-instance based system. In the results, it showed that for multi-instance based system at score level, that PCA and LDA produced an area of 0.967 and 0.955, while individually they produced 0.940 and 0.945.

**Introduction**

One of the largest trends in industry is face recognition with biometrics. Several new technologies include 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. [Anil K Jain]. Much successful research has proven that recognition with biometrics has been successful [anil k jain]. As mentioned previously, multi-classifier systems have been very successful, however there are a number of challenges associated with them. One includes merging classifiers which have different representations of patterns to be classified. As a result, coming up with a standard methodology where multiple pattern representations are used was a tedious process [Kittler]. 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 creates an overlap of results. This was also shown in previous research in which for both max and min, and the resolution was to normalize the results through normal distribution [Kuncheva].

**Method**

There were a number of methods to execute the experiments:

* Create LDA and PCA vector space: This was completed in Matlab using matrix calculations
* Create labels: Create labels that relate to imposter vs genuine
* EZRoc: Utilize Ezroc to obtain results.
* MCS Experiment A: Combine scores for PCA and LDA using minimum, maximum, and average rules.
* Multi-instance system: Average the scores by comparing the test sample with all the templates of a particular identity.

**Results and Discussion**

The first experiment conducted incorporated combining LDA and PCA using the minimum, maximum, and average computations. This is commonly used and these simple fusion methods a great for developing a good multiple classifier system [Kuncheva]. The minimum and maximum methods are great for 2 classes [Kuncheva] [Figure C.1]. The way this was computed was by using the min, max, and mean for each score between the LDA and PCA vectors [Appendix Equation C.2, C.3, C.4]. The results that were produced for this was fairly reasonable, with an ROC area of 0.94478 (average), 0.9453 (min), and 0.9400 (max) [Appendex Table B.1]. While the single classifier for LDA and PCA was for LDA and 0.94013 for PCA. As a result, it can be assumed that the max value score when combining PCA and LDA scores produced the better results. This is also supported from research by Kuncheva in which the minimum and maximum classifier produced the lowest classifer error than that of the single classifier [Kuncheva].

As for the second experiment, which was multi-instance system for average scores, it was revealed that this procedure both LDA and PCA produced higher scores than that of single classifier. This makes sense in light of the research by Cano, in which multi-instance methods produce better performance than single classifiers. As for the results, the multi-instance for LDA produced 0.9549 and PCA produced 0.9671 for ROC area [Appendix, Table B.2]. As mentioned previously these results were better than what was produced for single classifiers [Figure A.2, figure with multiple roc curves]. As a result, this shows success of multi-instance classifiers producing better than single classifiers.

This entire procedure of combining classifiers using simple methods is known as voting or ensembles (Textbook, p.492). In this practice the base-learners use their outputs and combine them based on the result (p. 492). The sum rule is the most widely used. The median rule is most powerful to noise, and the minimum/maximum rules will either give trends that are negative or positive (p. 493). As a result, this is becoming a more popular practice that is gaining widespread attention.

**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 because of the varied data. As a result, one take away point is to use min-max normalization in order to make the data more consistent. As a result, utilizing multi-classifier system is important and something that should be utilized more in image recognition.

**Appendix A**

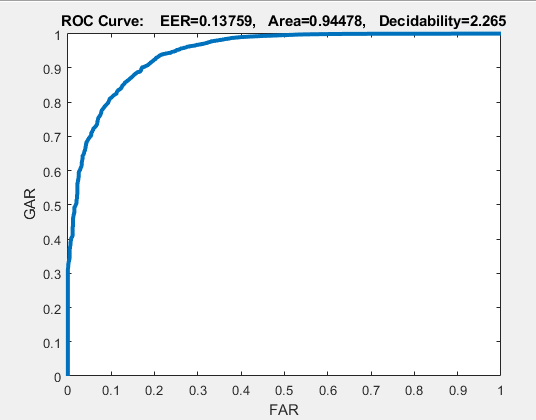
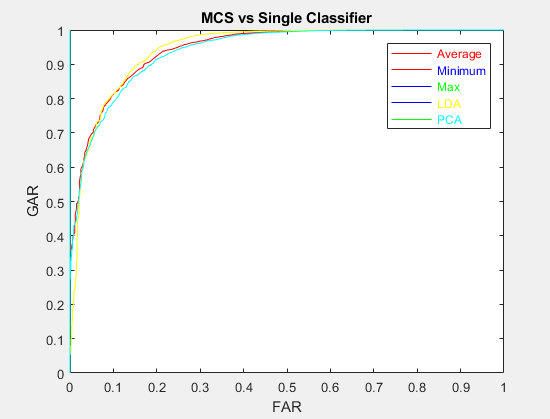


Figure A.1. Part A: MCS Comparison Figure A.2. Part A: MCS Average Results

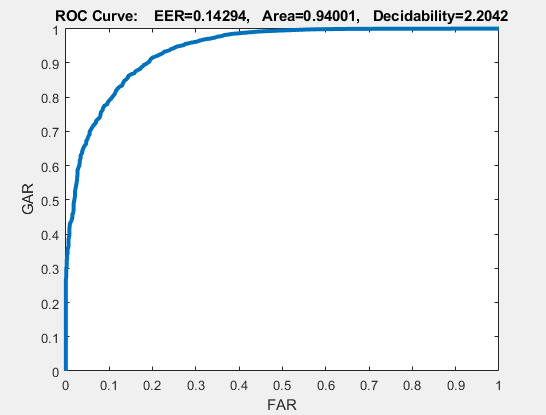
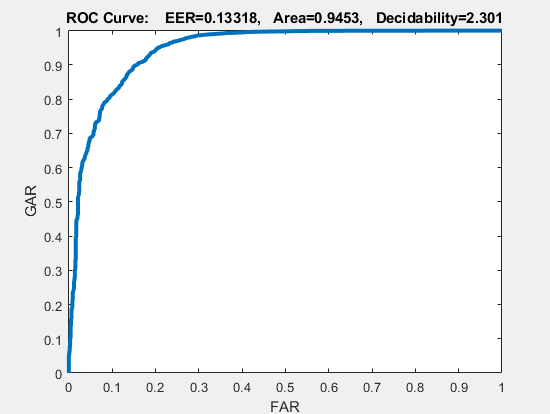
 

Figure A.3. Part A: MCS Max Results Figure A.4. Part A: MCS Min Results

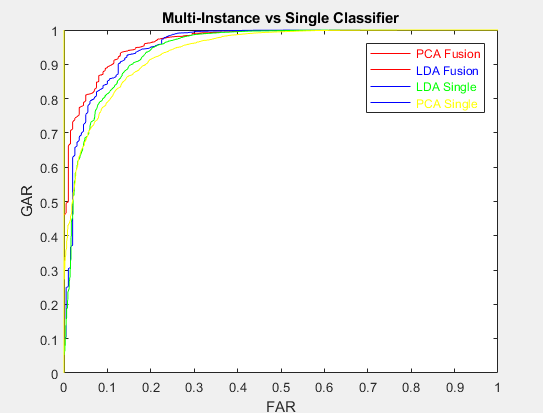
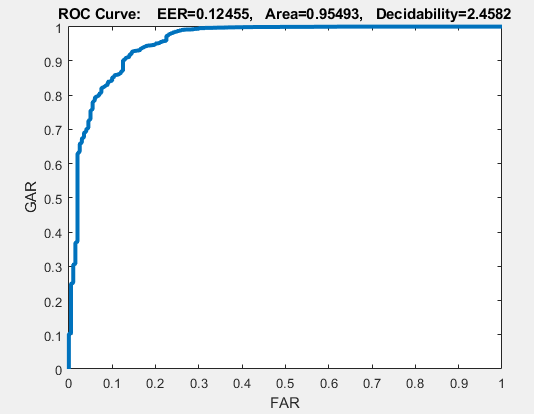
 

Figure A.5. Part B: Multi-Inst Comparison Figure A.6. Part B: Multi-Inst LDA Fusion

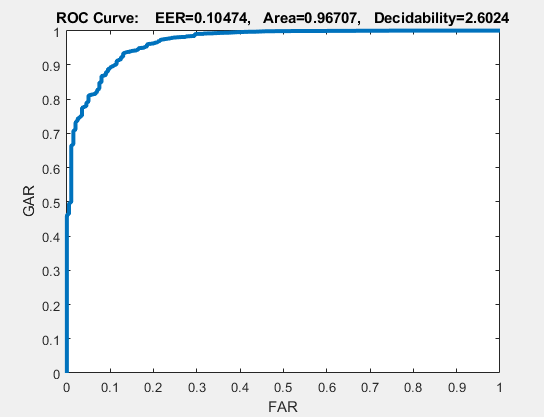


Figure A.7. Part B: Multi-Inst PCA Fusion

**Appendix B**

**Model B.1: Simple Fusion Model Diagram. F is fusion model [Kuncheva]**

**Model B.2: Equation that returns max value for LDA and PCA**

**Model B.3: Equation that returns min value for LDA and PCA**

**Model B.4: Equation that returns mean value for LDA and PCA**

|  |  |  |
| --- | --- | --- |
| **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 |

**Appendix C**

**Table C.1. Part A: 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 |

**Table C.2. Part B: Classifier Statistics**

**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, pp 492-494