Improved Identical Twins Face Recognition with Kin Images

Ivan Chit Hin Ho
Stanford University
Stanford, CA 94305
Ivanho1@stanford.edu
Source Code via Github: https://github.com/alrightyi/stanford_cs231a

Abstract

Recent additions of face recognition technology to mobile devices have prompted discussions on its accuracy when user has identical siblings. There have also been numerous studies on algorithms and techniques to face recognition of identical twins. This paper evaluates accuracy of face recognition on identical twins using recent technologies such as CNN detection and FaceNet encoding, as well as devises a new algorithm to improve accuracy of identification by correlating features with twins' family images.

1. Introduction

Recent additions of face recognition technology to mobile devices has prompted discussions on its accuracy when user has identical siblings. See https://9to5mac.com/2017/10/31/face-id-twins for example. As rate of new-born babies who are twins increases [1] (3% of every new born in United States are twins), ability to recognize twins' faces becomes statistically important.

In academia, there have also been numerous studies [2] [3] [4] [5] on algorithms and techniques to recognize faces of identical twins.

Personally, I have always had a hard time recognizing my friend's twin daughters ever since they were born. As they live in a different city, I rely on glancing at their images on social network without name correspondences. Significant amount of guess-work need to be done to correctly identify them.

Conventional wisdom suggests that children often resembles their parents and family members. Phrases like "you have your mother's eyes" often rings true from the human eye but has not been mathematically explored in computer vision field. In this paper, I attempt to compute the resemblance factor among twins and family to see if accuracy of twin identification could be improved.

Two major tasks have been completed as part of the project:

1) implement a face recognition pipeline consisting of face detection, estimation and morphing, encoding, and classification, and

2) propose and implement a two-stage support vector classifier by calculating the Relative Euclidean Distance among image samples of identical twins and their family members.

With limited dataset, results from 2) show an improvement of twin identification accuracy over 1) by a margin of 5 to 30 percent when varying family size from 2 to 8, respectively. I also evaluated the algorithm against changing the number of classes in stage-1 and 2 of the classifiers.

1.1. Baseline Techniques (1-Stage SVC)

The following steps and techniques are used to construct a baseline of face recognition pipeline (from hereon called 1-stage classifier):

1) Resize sample and test images using Python Image Library [6] to 1000x500x3 pixels
2) Apply face detection to images (using Convolutional Neural Network from Open-Face [7])
3) Estimate face using 68-point face landmark estimator from Yusuf *et al*. [8].
4) Encode face into a 128-element vector using a Convolutional Neural Network implementation of FaceNet [9] by Dlib [10].
5) Implement a Support Vector Classifier using scikit-learn [11], varying number of classes from 3 to 15. 3 is the minimum number of classes, representing the twins as well as a class for unknown faces. The data used to train the SVC comes from sample images of various known and unknown faces, as well as the twins and their family members. The SVC uses a balanced approach to classification, with tolerance = 1e-6 and maximum iterations = 20,000.

The steps outlined above follow recommendations from https://github.com/ageitgey/face_recognition#face-recognition

1.2. 2-Stage SVC with RED

Multi-stage Support Vector Classifiers (SVCs) have been proposed for face recognition in a number of studies [12] [13] and have offered improvements compared with existing single-stage approach due to the coarse-to-fine nature of machine-learned classifiers. Kuo et al.'s multi-stage classifier for face recognition uses a combination of Support Vector Machine (SVM), Eigenface, and RANSAC techniques to achieve nearly 100 percent accuracy, though the dataset does not specify the percentage of data representing twin faces.

The approach taken in this paper is to leverage family relationship information to construct a second-stage SVC to improve accuracy of identifying twin images.

By calculating the distance in Euclidean space between the encodings we can represent how close the two images resemble one another.

1.4. Stage-2 SVC Algorithm

Below describes the detailed algorithm for implementing the 2-stage SVC pipeline.

1) Follow steps from 1.1
2) Train a second-stage SVC using the Relative Euclidean Distance of the sample images of the family against images of the individuals (twins) in question. Family size ranges from 2 to 8, which is the number of classes in the second-stage SVC.

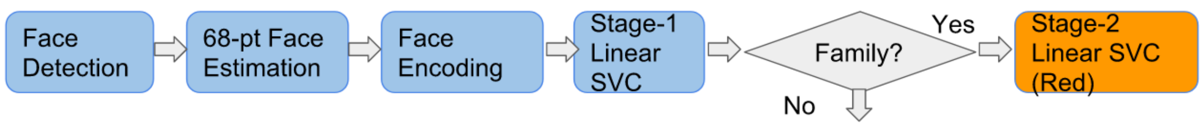

Figure 1: 2-Stage SVC using RED to classify twin image encodings

Relative Euclidean Distance (RED) is calculated between 128-element vector encodings of all family sample images against each of the member's images as input to the second-stage classifier. The intent is to correlate the encoding information of all the family members with that of the twins, so the stage-2 classifier could learn to distinguish the twins better than with just the encodings.

1.3. Relative Euclidean Distance (RED)

The calculation of RED of a given family member encoding h against another family member's encoding i is defined by:

Equation 1: Relative Euclidean Distance

$$RED_{ih} = \sqrt{\sum_{j=1}^{p} \left[\left(\frac{a_{ij}}{\sqrt{\sum_{j=1}^{p} a_{ij}^{2}}} \right) - \left(\frac{a_{hj}}{\sqrt{\sum_{j=1}^{p} a_{hj}^{2}}} \right) \right]^{2}}$$

where a_{ij} is one of the 128 elements in the family sample image encoding. And a_{hj} is equivalent element from the classifying family member encoding.

p is the number of the elements in the encoding, which is 128.

Mathematically, RED represents the chord distance between two points on the surface of a unit hypersphere.

The SVC uses a balanced approach, with tolerance = 1e-6 and maximum iterations = 20,000.

3) Feed a mix of twins, their family, non-family members, and unknown image encodings into the face recognition pipeline.
4) If the output of the stage-1 SVC predicts a family member, then it goes to the stage-2 SVC. Otherwise the output is recorded and compared to actual for accuracy.
5) Before going to stage-2 SVC, the RED of the test image encoding is calculated against the same set of known family images to train the stage-2 SVC.
6) Results from 5) are then fed into the stage-2 SVC for prediction. Output is compared with actual for accuracy calculation.

Figure 1 shows a high-level block diagram of the 2-stage SVC with RED face recognition pipeline.

1.5. Datasets and Models

Due to limited datasets found on existing twin images, the project uses existing open-source models to first form the face detection, estimation, and encoding components of the pipeline. Details are listed below:

- CNN Max-Margin Object Detection (MMOD) model from Dlib [10], an open-source toolkit for machine learning.
- 68-point face detector from Kazemi and Sullivan [8]. This is implemented using library provided by

Dlib
- FaceNet Deep Neural-Network Encoding from Google.

Image datasets from Kinship Face in the Wild [14] are used to train the 2-stage SVCs to identify unknown faces from images.

Sample images from friends, including images from twins Eman and Eden Shen, and their family are used to train known classes in the SVCs.

In total, over 200 images are used to train the SVCs and another 150 images are used for testing accuracy, 50 of which are from the twins.

2. Results

Two sets of graphs are obtained, comparing the accuracy of classifying the images using results from stage-1 SVC versus stage-2 SVC.

The following accuracies are measured for each stage of the SVC:

1) Accuracy of all test images in percentage points
2) Accuracy of non-twin images in percentage points
3) Accuracy of twin images in percentage points

The accuracy is plotted against family size (representing the number of classes in the second-stage SVC), which is set to 2, 4, 6, and 8.

2.1. Varying Number of Classes in Stage-1 SVC

The following graph shows the results when family size is set to 2, 4, 6, 8, and varying the number of classes in stage-1 SVC from family size to 15:

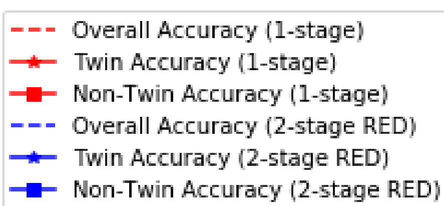

Figure 2: Legend for Figure 3

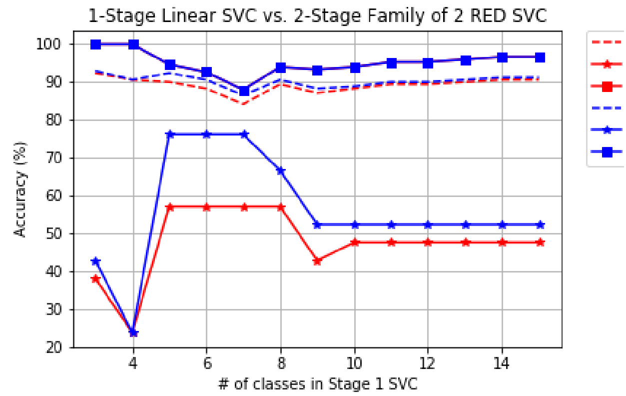

Figure 3: 1-Stage Linear SVC vs. 2-Stage Family of 2 RED SVC

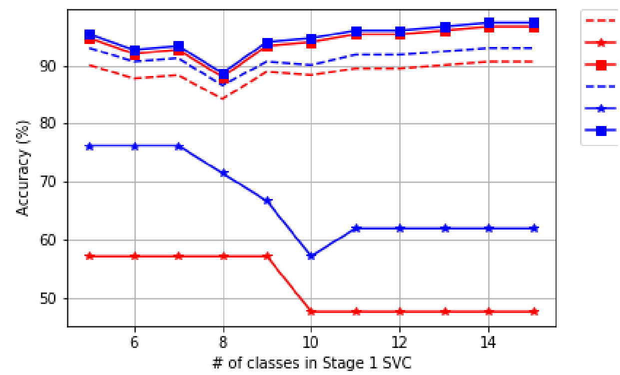

Figure 4: 1-Stage Linear SVC vs. 2-Stage Family of 4 RED SVC

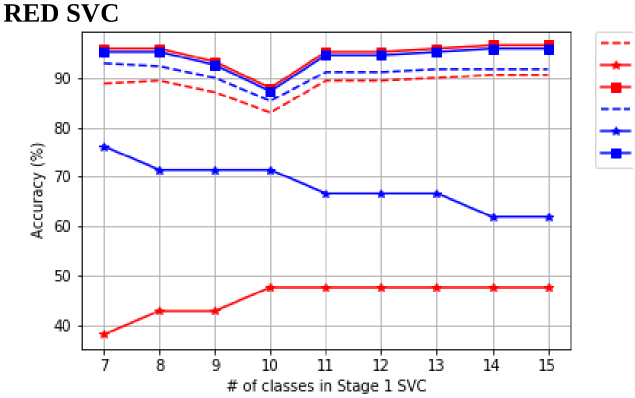

Figure 5: 1-Stage Linear SVC vs. 2-Stage Family of 6 RED SVC

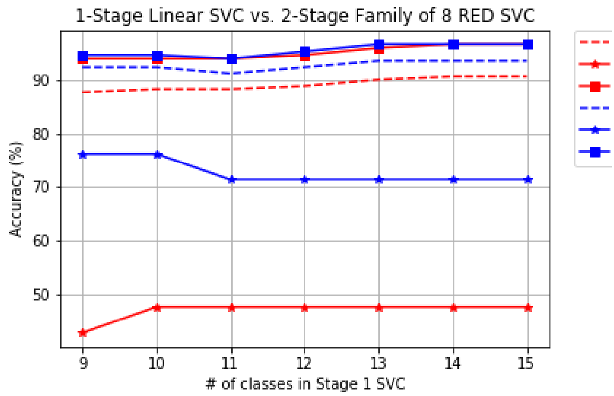

Figure 6: 1-Stage Linear SVC vs. 2-Stage Family of 8 RED SVC

Results from Figure 3 to Figure 6 show that accuracy gain for identifying twin images is highest when number of classes in stage-1 SVC is equal to the family size plus unknown. This is expected because the 2nd stage RED classifier further distinguished the twin images without filtering false negatives from the first stage.

Non-twin accuracy remains relatively the same between 1-stage and 2-stage results.

Twin accuracy drops when number of classes in stage-1 SVC increases to 10 and then levels off when increasing further. There are two possible explanations:

1) This may be due to limited dataset available causing skewed results.
2) It is also possible that the first-stage SVC correctly predicted the family member but subsequent prediction in stage-2 overrides it to an incorrect identification. This negative effect reaches a local minimum at number of classes = 10.

2.2. Varying Number of Classes in Stage-2 SVC

The following graph compares the accuracy of identifying the twins when number of classes in stage-1 SVC is fixed at 15, then varying the number of classes (the family size) in stage-2 SVC from 2 to 8:

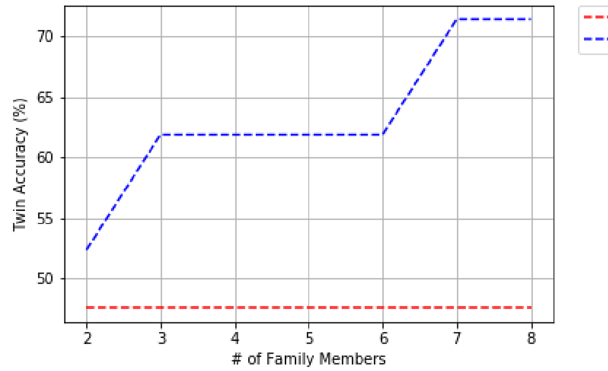

Figure 7: 1-Stage Linear SVC vs. 2-Stage Family RED SVC (blue: 2-stage, red: 1-stage)

Results from Figure 7 show that the accuracy of identifying the twins with 2-stage SVC increases in a step-wise fashion when the number of classes (family size) in stage-2 SVC increases. This is because if stage-1 classifer incorrectly predicts an actual family member image to non-family member, then it will not enter the second stage for possible correction.

Figure 8 illustrates a sample output of a side-by-side image of the twins, showing the results of stage-2 SVC correcting an incorrect prediction from stage-1 SVC.

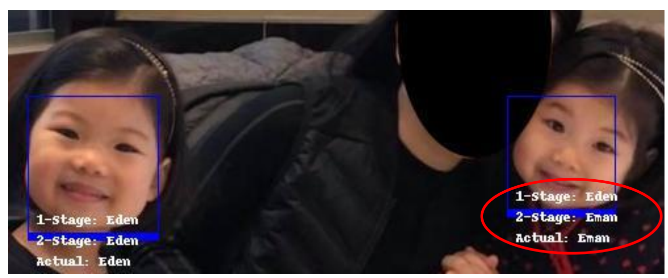

Figure 8: Test image comparing 2-stage SVC with RED prediction over 1-stage SVC and actual

2.3. Interpretations

The idea of using family members' image encodings to correlate with twins' encodings to achieve higher accuracy can be generalized to correlation among encodings with closest resemblance to one another, mathematically represented by their Relative Euclidean Distance.

For example, resemblance could be grouped based on age, gender, ethnicity, or a combination of.

It can also be mathematically grouped based on a preset threshold of their RED. In this case, one could set the criteria for adding the class to the 2-stage SVC without prior information about the subjects such as family relationship.

2.4. Future Work

Several future tasks can be considered beyond the scope of this project:

1) Rerun algorithm with larger datasets, including:
    a. More samples of other twins and their family members
    b. Varying images with age, gender, and diversity, representing general population distribution.
2) Apply preset RED threshold to second-stage classifier instead of using family members as criteria.
3) Compare twin accuracy results of the 2-stage RED SVC with other multi-stage classifiers mentioned in this paper.
4) Increase number of stages of SVC to evaluate improvement gains.
5) Explore other face recognition techniques based on family resemblance (e.g. correlation of facial landmarks)

3. Conclusion

Two major tasks have been completed as part of the project:

1) implement a face recognition pipeline consisting of face detection, estimation and morphing, encoding, and classification, and

2) propose and implement a two-stage support vector classifier by calculating the Relative Euclidean Distance among image samples of identical twins and their family members.

With limited dataset, results from 2) show an improvement of twin identification accuracy over 1) by a margin of 5 to 30 percent when varying family size from 2 to 8, respectively. Evaluation of the algorithm against changing the number of classes in stage-1 and 2 of the classifiers also shows consistent behavior.

The algorithm can be generalized based on a preset threshold of the images' Relative Euclidean Distance. In this case, one could set the criteria for adding the class to the 2-stage SVC without prior information about the subjects (such as family relationship).

4. Acknowledgement

I would like to thank Dr. Savarese and his Teaching Assistants in CS231A for the opportunity and their positive feedback on the project.

I would also like to thank the Shen family for providing critical sample images of themselves, including their twin daughters Eman and Eden, and brother Enoch.

This project would not complete without the support of my wife, who has stayed next to me at coffee shops and libraries every weekend for the last two and half months, patiently waiting for opportunity for us to spend time together.

5. References

[1] F. . Miró, E. . Vidal and J. . Balasch, "Increased live birth rate in twin pregnancies resulting from embryo assistance.," *Obstetrics & Gynecology,* vol. 119, no. 1, pp. 44-49, 2012.

[2] F. . Juefei-Xu and M. . Savvides, "An Augmented Linear Discriminant Analysis Approach for Identifying Identical Twins with the Aid of Facial Asymmetry Features," , 2013. [Online]. Available: http://users.ece.cmu.edu/~juefeix/felix_cvpr13_twin.pdf. [Accessed 20 3 2018].

[3] J. R. Paone, P. J. Flynn, P. J. Philips, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn, M. T. Pruitt and J. M. Grant, "Double Trouble: Differentiating Identical Twins by Face Recognition," *IEEE Transactions on Information Forensics and Security,* vol. 9, no. 2, pp. 285-295, 2014.

[4] S. . Biswas, K. W. Bowyer and P. J. Flynn, "A study of face recognition of identical twins by humans," , 2011. [Online]. Available: http://dl.acm.org/citation.cfm?id=2360519. [Accessed 21 3 2018].

[5] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn and M. T. Pruitt, "Distinguishing identical twins by face recognition," , 2011. [Online]. Available: https://nist.gov/publications/distinguishing-identical-twins-face-recognition. [Accessed 20 3 2018].

[6] "Python Imaging Library," , . [Online]. Available: http://www.pythonware.com/products/pil/. [Accessed 21 3 2018].

[7] T. . Baltrusaitis, P. . Robinson and L.-P. . Morency, "OpenFace: An open source facial behavior analysis toolkit," , 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7477553. [Accessed 21 3 2018].

[8] A. A. Yusuf, Z. . Sufyanu, F. S. Mohamad and K. . Nanaa, "Facial Landmark Detection and Estimation under Various Expressions and Occlusions," *International Journal of Sciences: Basic and Applied Research,* vol. 18, no. 2, pp. 40-50, 2014.

[9] F. . Schroff, D. . Kalenichenko and J. . Philbin, "FaceNet: A unified embedding for face recognition

and clustering," *arXiv: Computer Vision and Pattern Recognition*, vol. , no. , pp. 815-823, 2015.

[10] D. E. King, "Dlib-ml: A Machine Learning Toolkit," *Journal of Machine Learning Research,* vol. 10, no. , pp. 1755-1758, 2009.

[11] "2.3. Clustering — scikit-learn 0.18.1 documentation," , . [Online]. Available: http://scikit-learn.org/stable/modules/clustering.html#affinity-propagation. [Accessed 21 3 2018].

[12] C.-H. . Kuo, J.-D. . Lee and T.-J. . Chan, "A novel multi-stage classifier for face recognition," , 2007. [Online]. Available: http://dblp.uni-trier.de/db/conf/accv/accv2007-2.html. [Accessed 20 3 2018].

[13] J.-D. . Lee and C.-H. . Kuo, "A Multi-Stage Classifier for Face Recognition Undertaken by Coarse-to-fine Strategy," , 2009. [Online]. Available: https://intechopen.com/books/howtoreference/state_of_the_art_in_face_recognition/a_multi-stage_classifier_for_face_recognition_undertaken_by_coarse-to-fine_strategy. [Accessed 21 3 2018].

[14] M. B. López, E. . Boutellaa and A. . Hadid, "Comments on the “Kinship Face in the Wild” Data Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 38, no. 11, pp. 2342-2344, 2016.