The Golden Retrieber

Cameron Davidson-Pilon Jennifer Smith

December 21, 2011

Abstract

The *Golden Retrieber* is an algorithm that collects Twitter profile pictures and performs a binary classification on the images, those containing Canadian pop-idol Justin Bieber, and those not. Through much experimentation, we found that kernelized supervised PCA combined with kernalized Support Vector Machine performed the best authentication on our images with an error rate of 35.33%.

1 Introduction

Facial authentication is a much studied area of machine learning and has broad application in commercial and security fields. Currently, photos of poor quality, varying facial expressions, or containing large illumination and pose variation still present a large challenge for machines. In programs initiated as recently as 2006 there were efforts to stimulate research in new algorithms and improve techniques in this area, such as the Facial Recognition Grand Challenge [7]. In this paper, we endeavor to face this challenge in the form of a binary classification of randomly collected images sourced from Twitter.

1.1 Motivation and Accomplishments

Justin Bieber is one of the most prominent figures on the social media site Twitter. With over 15 million followers on Twitter, an enormous fraction of his followers feature Justin Bieber (henceforth called JB) in their profiles' display picture. With our project, we tried to find these devoted fans by classifying randomly selected profile pictures as either containing JB or not containing JB.

We used a Python Twitter API [6] to collect the 40 by 40 pixel profile images from Twitter at random, then hand-selected the images not containing humans to be discarded. We then seperated JB images and non-JB images. In total we used 47 JB images and 28 non-JB images. The images were pre-processed in MATLAB: we used a facial detection algorithm [3] to locate the faces in the images to crop them and turn the photos to greyscale. This was done to help center the faces within the image and reduce the dimensionality of the data, respectively.

After pre-processing the images and further dimensionality reduction using kernelized supervised PCA (KSPCA) [4], we used kernelized Support Vector Machine (KSVM) for classification. Originally, we wanted to use Gabor filters, KNN (K-Nearest Neighbour) or Un-

Figure 1: A sample of Twitter profile pictures. The images on the left are Twitter profile pictures of Justin Bieber, and images on the right are a random sample of non-Bieber Twitter profile pictures. The *Golden Retrieber* tries to classify these images.

correlated Local FDA (ULFDA)[2] in our algorithm, but encountered problems with these methods. Thus, our best error rate using a combination of KSPCA and KVM was .3533.

1.2 Background

The Facial Recognition Grand Challenge (FRGC) was a series of progressively difficult challenges in facial recognition promoted by the U.S. Government and academia to further research in the various problems that were still associated with facial recognition. Each challenge involved a data set and a predefined set of experiments. One such challenge was to perform facial recognition on uncontrolled images: these images had varying facial expressions, pose and illumination variations, presence of eye-glasses, etc. [7] In this paper, we also aim to classify randomly collected images taken from Twitter. Since the challenge with uncontrolled images presented by the FRGC is a similar problem to the one we are facing, much of our early research overlapped with techniques developed for this challenge. For instance [8] developed a technique to deal wtih this problem involving using Correlation Feature Analysis to reduce dimensionality along with Support Vector Machine (SVM). [9] used class-dependent kernel discrete cosine transform features. This approach borrows from traditional approaches for classification.

The rest of the paper is as follows: Section 2 contains a discussion of our methodologies and further describes the tools and techniques we used in our project, Section 3 summarizes our results, Section 4 explores errors that may have negatively affected our results, and Section 5 contains some concluding remarks.

2 Methodology

In our first attempt to classify our data, we applied different Gabor filters [1, p. 62-81] to each image. Gabor filters are reasonably tolerant to variation in background and rotation of the face and have had very good results when applied to the problem of classifying facial expression [1]. The 2-D Gabor filter is defined as follows:

$$G(x, y) = ke^{-\pi S^2(x^2+y^2)}\left(e^{j(2\pi F(x\cos(W)+y\sin(W))+P)} - e^{-\pi(F/S)^2+jP)}\right)$$

where x, y are values of pixels in the image, S is the variance, (F, W) is the polar frequency, P is the phase, k is a normalizing constant, and j is the imaginary unit $\sqrt{-1}$. See [5] for more

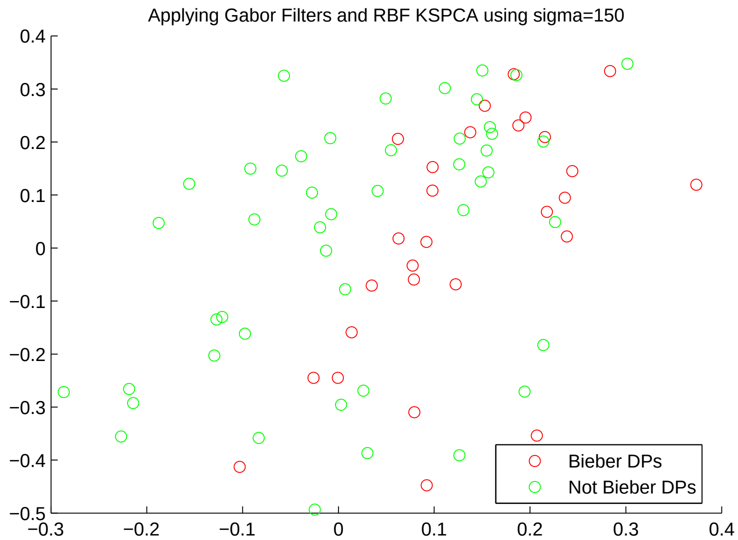

Figure 2: The results of KSPCA applied to Gabor filtered data, projected onto a 2D plane. The 1NN 10-fold cross validated error rate was over 0.5.

information on the implementation of the filter.

To each post-processed image, we applied the Gabor filter with varying parameters. The matrix output of this was very high dimensional (about 7,200 to 36,000 times the number of images, depending on image size and number of filters applied). Hence we used a radial-basis function kernel (see Appendix A) and SPCA to reduce the dimensionality of our data, and 1NN to classify. We observed poor performance of this technique after cross-validating for optimal dimension projection and kernel parameter. This is potentially due to the challenges that our data presents with the varying illumination, facial expression, and drastic changes in pose. Figure 2 displays a sample output of KSCPA applied to Gabor-filtered data.

Uncorrelated Local FDA (ULFDA) is a promising new method for facial recognition [2]. ULFDA aims to use subspace learning to reduce the dimensions of the feature space by mapping the images onto the whole new data space without redundancy in the basis vectors. It has promising results in classifying multi-modal labeled data. This is relevant to our project as we are attempting a 'one versus many' classification, and our non-JB data set, which is comprised of a many unique individuals, will exhibit a multi-model structure. Unfortunately, we unsuccessfully implemented the algorithm in [2]: our *local between-manifold scatter matrix* and *local within-manifold scatter matrix* converged too close to a singular matrix and our results were not significant.

Through our origingal research, we discovered that PCA and PCA-based techniques [11] are among the most common for facial recognition. After many setbacks, we decided that KSPCA was a strong option for our data set. Since we have labelled data, traditional PCA methods were not a good candidate for classification, but KSPCA is attractive because it is

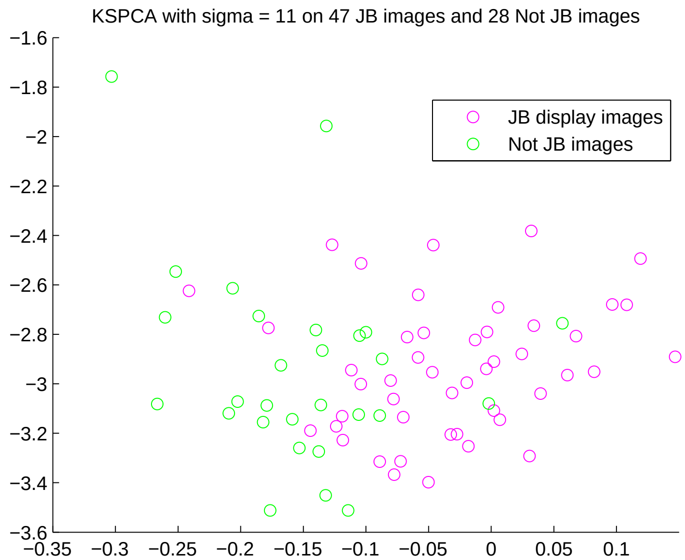

Figure 3: The 2-dimensional results of KSPCA with $\sigma_{SPCA} = 11$ to unfiltered data. The data is almost linearly separable, but requires a non-linear classifier.

able to handle labelled data, and it reduces the high dimensions of our data by picking out the principal components of maximum variability. [4]

We observed that KSCPA applied to post-processed data without any filters produced good results (see figure 3 for results). As the projected data was very clustered, we chose to compare two non-linear classifiers: KNN and kernelized SVM [1]. We set the SPCA RBF parameter σ_{SPCA} equal to 11 and the SVM RBF parameter $\sigma_{SVM} = 0.4$, and performed 23 trials of randomized 8-fold cross validation to compare the two classifiers. (We also tried other SVM kernels, but they worked either about the same, or worse). Figure 4 show the results. The main result is that for higher dimensional projection, KSVM does significantly better than 1NN.

3 Results

Observing that SVM does better than 1NN, we decided to pursue the 21-dimensional KSVM and use cross-validation to find the best σ_{SVM}. Again, using 23 trials of 8-fold cross validation, the optimal parameter is $\sigma_{SVM} = .15$, and gives an error rate of 0.3533. Figure 5 displays the resulting error rates for different values of the parameter.

[1] for all SVM trials, we set the soft margin parameter $\gamma = 0.5$. The hard-margin case was almost universally worse.

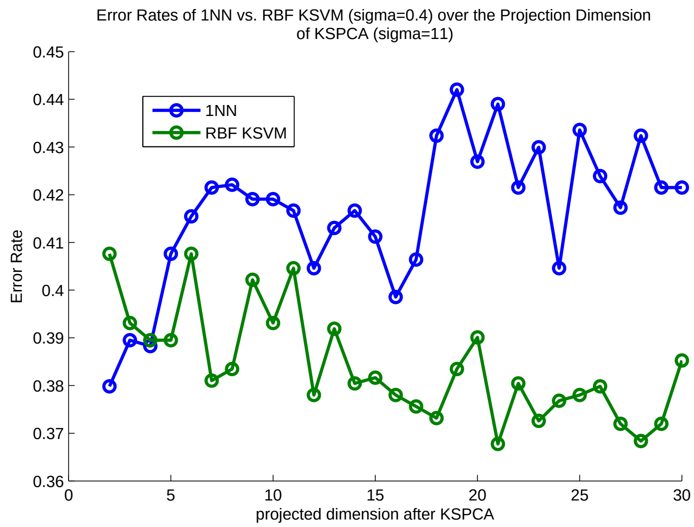

Figure 4: The results of 23 trials of 8-fold cross validation comparing RBF KSVM and 1NN over different projection dimensions ($\sigma_{SVM} = 0.4, \sigma_{SPCA} = 11$). Initially KSVM does poor but eventually outperforms 1NN.

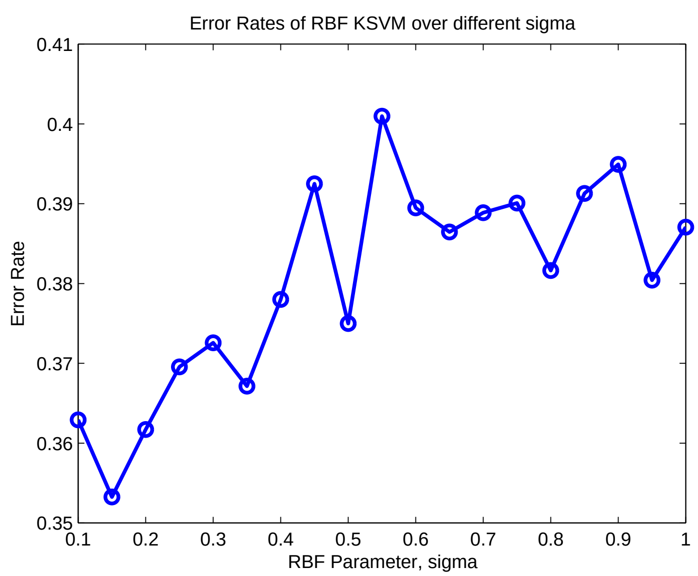

Figure 5: The results of 23 trials of 8-fold cross validation over different values of σ_{SVM}.

4 Identifying Sources of Error

Most research in facial recognition and authentication uses data sets that are close to ideal: perfectly centered eyes and faces, uniform background and illumination, simple facial expressions etc. The added difficulty of our data set is the lack of such control. For example, some of our images contained text over the face of an individual, or the head was severely misaligned in the image resulting in poor face recognition and cropping.

Along with the varying conditions, there was another serious problem, one that has never been encountered in traditional facial data sets. Our target individual physically changed! JB become popular when he was very young, and there exist many images on Twitter of pre-pubescent JB. He currently looks more physically mature than he did three or so years ago. Our algorithm makes no distinction between this.

Because of the already small image sizes (40 by 40 pixels) and the face not being immediately centered, our post-processed, cropped images suffered from a lack of data. To compensate for this, we scaled all our images to 30 by 30 pixels before appending it to our data matrix.

5 Conclusion

The automated task of facial authentication with noisy, or varying conditions, is still a challenging problem. The report here confirms that challenge. We compared our algorithm with the abilities of another previously successful Justin Bieber detection algorithm: a little sister of one of the researchers in this project. She scored 100% and rolled her eyes.

Further research, which we hope to pursue with the goal of creating a web application, will involve more advanced pre-processing of images, more robustness to a maturing JB and a classifier that will distinguish between images containing humans (and thus possibly JB) and images that do not. Also, according to [10], there exists a range of parameters (σ_{SVM}, γ) that yield optimal classification performance. We did not explore this, as it requires a grid search over a large space and was too computationally expensive.

References

[1] Zhang, Yu-Jin. "Advances in Face Image Analysis". IGI Global. Hershey PA, 2011.

[2] Huang, H. J. Liu, and H. Feng. "Uncorrelated Local Fisher Discriminant Analysis for Face Recognition". IJPRAI Vol 25 .6 (2011): 863-887.

[3] Chatterjee, Arunima. April 14, 2009. *DETECTION OF FACE PORTION FROM A SINGLE FACE IMAGE.* [Matlab Program] http://www.mathworks.com/matlabcentral/fileexchange/23772-detection-of-face-portion-from-a-single-face-image.

[4] Barshan, Elnaz , Ali Ghodsi, Zohreh Azimifar, and Mansoor Zolghadri Jahromi. "Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds". Pattern Recognition 44.7 (2011): 1357-1371. Elsevier.

[5] Schwanz Dias, Stiven. January 2007. *IMPROVED 2D GABOR FILTER*. Matlab Program. http://www.mathworks.com/matlabcentral/fileexchange/13776-improved-2d-gabor-filter

[6] The Python-Twitter Developers. *PYTHON TWITTER*. [Python Program] http://code.google.com/p/python-twitter/

[7] National Institute of Standards and Technology (NIST). "Face Recognition Grand Challenge (FRGC)". NIST.gov. February 15, 2011. http://www.nist.gov/itl/iad/ig/frgc.cfm.

[8] Savvides, M., R. Abiantun, J. Heo, S. Park, C.Xie, and B.V.K. Vijayakumar. "Partial & Holistic Face Recognition on FRGC-II data using Support Vector Machine Kernel Correlation Feature Analysis". *IEEE Conf. Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2006.

[9] Savvides, Marios, Jingu Heo, Ramzi Abiantun, Chunyan Xie, and B.V.K. Vijaya Kumar. "Class Dependent Kernel Discrete Cosine Transform Features for Enhanced Holistic Face Recognition in FRGC-II". *IEEE Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2006.

[10] Weston, J., Ben-Hur, A. "A User's Guide to Support Vector Machines". http://pyml.sourceforge.net/doc/howto.pdf

[11] Lemieux, Alexandre and Marc Parizeau. "Experiments on eigenfaces robustness". *16th International Conf. on Pattern Recognition (ICPR)*. Vol. 1 (2002): 421-424.

A Appendix

The *radial basis function*, $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is defined as

$$k(x, y, \sigma) = e^{\frac{-||x-y||^2}{2\sigma^2}}$$

and is a commonly used kernel in machine learning. One drawback with using this kernel, and most kernels for that matter, is determining appropriate values for the parameters. We used cross validation to get a sense of what magnitude σ should be.

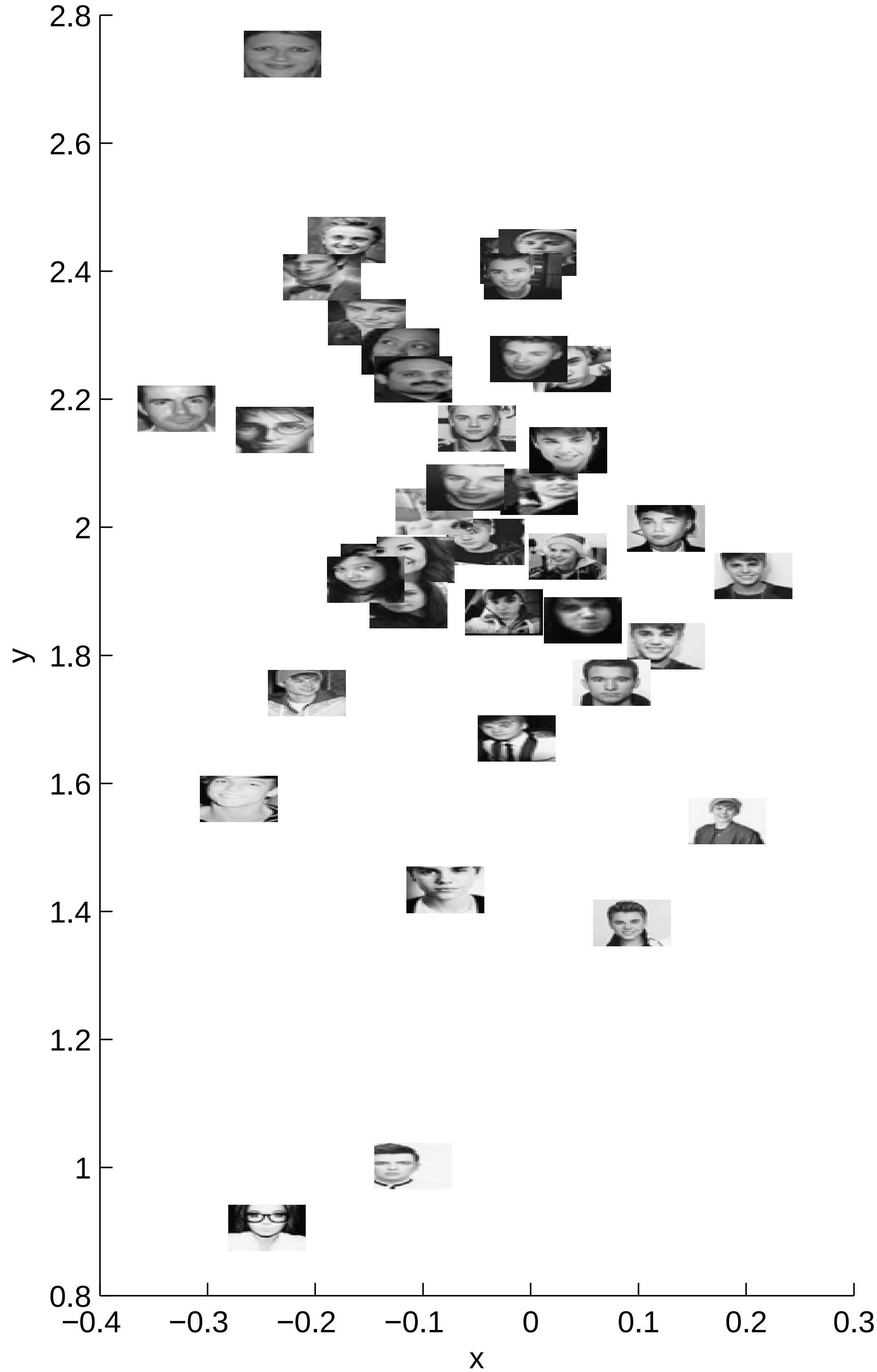

Figure 6: JBiebs vs the world!