Philipp Omenitsch 30.08.2015

Gender classification with Fisher vector faces on the Adience dataset

Method

The task was to use fisher vector faces [1] to calculate features of the faces provided in the adience dataset [2] and train a Support vector machine (SVM) for a gender classifier.

Dataset

The dataset contains 26580 images. I used the in 2D-plain aligned version of the dataset for my experiments. Also I used only frontal images of faces. So the dataset contained only 13560 images.

Preparation

First all faces are cropped from the images with a static mask (114x114 pixels of a 250x250 image obtained by a Viola Jones face detector).

Fisher Vector Faces

I used the provided MATLAB implementation of the fisher vector faces [1] and adopted it for my needs. For Fisher vector faces, first dense SIFT features are computed on the input image. In the next step they are then encoded into one feature vector using fisher vectors, which use Gaussian Mixture Models to fit the data. (This tries to keep as much information while lowering the number of need features). Next a dimensionality reduction is applied by searching for a linear projection which minimizes distances between samples of the same group and maximizing samples of different groups (in our case faces).

In order to compute features for the adience data set I used an already trained set of features which were trained on the unrestricted Labeled Faces in the Wild (LFW) data set. The provided data by the authors made it only possible to use feature sets learned on a tenth of the whole dataset without a major code rewrite.

Machine learning

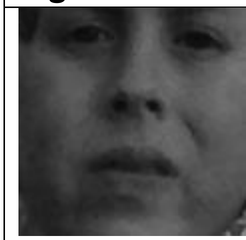
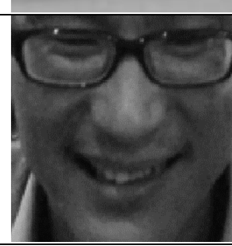
For training I used a SVM using k-fold cross validation with k=5 using the provided folds by the adience dataset. I optimized the C parameter (or lambda = 1/C) to find the value that provides the smallest error C = 40 (lambda = 0.025) by using exhaustive grid search on the averaged error rates of the 5 possible folds.

Results

On evaluation with 5-fold cross validation I get the following results as compared to other methods:

	Mean	Std. Dev
Best from [2]	77.8	1.3
Best from [3]	79.3	0.0
Best from [4]	86.8	1.4
With FVF [1]	68.0	2.4

Qualitative results

Right detections	**Wrong detections**
 The image is nicely aligned	 The image is very difficult even for me to properly classify (ground truth is female). The out of plane face and the big face deforming open mouth, make this a very difficult (as in probably few similar images in trainingset), image to classify.
 Since the first image is recognized correctly, it is logical a similar image is also properly classified	 Again the males mouth is an unusual deformation of the face. The feature extractor should be robust against this, nevertheless the quality of the image is also poor (a lot of noise) which might be another reason for the wrong classification.
 This images is nice aligned and shows many hard contours and beard, I can imagine that these may be features more often found in men than women	 This woman's image is available in low resolution with large noise, this might be a reason for wrong classification.
 The face is soft with smooth contours and big eyes, also it is nicely aligned.	 Here the first reason for wrong classification are the glasses, which might produce features often found in images of women.

Possible next steps

It would be good to calculate the feature codebook on both the LFW and adience trainingsset together to come up with a pretrained featureset based on more training data.

Another good idea might be to try different machine learning algorithms, I think a random forest might work well, because there are only two classes, of course a pruning technique should be used to reduce overfitting.

References

[1] K. Simonyan, O. Parkhi, A. Vedaldi and A. Zisserman, "Fisher Vector Faces in the Wild," in *Proceedings of the British Machine Vision Conference*, 2013.

[2] E. Eidinger, R. Enbar and T. Hassner, "Age and gender estimation of unfiltered faces," *Information Forensics and Security, IEEE Transactions on,* vol. 9, no. 12, pp. 2170-2179, 2014.

[3] T. Hassner, S. Harel, e. Paz and R. Enbar, Effective face frontalization in unconstrained images, arXiv preprint arXiv:1411.7964, 2014.

[4] G. Levi and T. Hassner, "Age and Gender Classification using Convolutional Neural Networks," 2014.