FEA-ACCU CASCADE FOR FACE DETECTION

Shengye Yan[1, 2, 3], *Shiguang Shan*[1, 2], *Xilin Chen*[1, 2], *Wen Gao*[4, 2]

[1]Key Lab of Intelligent Information Processing, Chinese Academy of Sciences (CAS) , Beijing, China
[2]Digital Media Research Center, Institute of Computing Technology, CAS, Beijing, 100190, China
[3]Graduate School of the Chinese Academy of Sciences, Beijing, 100039, China
[4]School of EE&CS, Peking University, Beijing, 100871, China
{syyan, sgshan, xlchen, wgao}@jdl.ac.cn

ABSTRACT

Aiming at unloading the high training time burden of the popular cascaded classifier, in this paper, a novel cascade structure called Fea-Accu cascade is proposed. In Fea-Accu cascade training, the times of feature selection are largely reduced by enhancing the correlation among different stage classifiers of the cascaded classifier. In detail, for each stage classifier, before selecting new features out, the features selected out by previous stage classifiers are reused through creating new corresponding weak classifiers. To verify the efficiency and effectiveness of the proposed method, experiment is designed on frontal face detection problem. The experimental results show that it can largely reduce the training time. A frontal face detector with state-of-the-art classification performance can be learned in less than 10 hours.

***Index Terms*—**face detection, cascade, Haar-like feature

1. INTRODUCTION

Face detection in images has been studied in many aspects ranging from representative features and classification methods to whole architecture of the face detector. The challenge of face detection is to make the machine having the ability to classify between face windows and enormous non-face windows in a very efficient way, regardless of the expression, pose, scale, lighting and camera condition changes in face pattern.

During the past decade, lots of methods have been proposed for robust and fast face detection such as [1-17]. Among these methods, Boosting cascade using simple Haar-like features proposed by Viola and Jones received large amounts of attention for its high computation efficiency [14]. Through using thousands of simple Haar-like features with a pre-computed "integral image" and cascade architecture, the detector can detect faces in an image with size 384*288 in nearly real time. However, with respect to training time, because of thousands of Haar-like features to be selected out from a even larger feature pool by Adaboost, the training time of the cascade is very long and counted on the order of weeks. So long a time has made it boring to produce a face detector.

In this paper, we proposed a novel cascade structure, called Fea-Accu cascade to reduce the training time of the cascaded classifier proposed by Viola and Jones. With respect to the capacity of further reducing the training time, the basic insight is that the correlation between different stage classifiers in the cascade of Viola and Jones is so weak that it can be promoted further, and so the times of feature selection can be decreased. In "Fea-Accu" cascade, the correlation is enhanced by reusing the features learnt out by previous stage classifiers in an accumulative way, and it is just in this sense, "Fea-Accu" is named, where "Fea" means feature, "Accu" means accumulated.

In detail, Fea-Accu cascade reuses the features of the previous stage classifiers to constructing the current stage classifier. Provided with different non-face training set, the features previously learned are used to make new correspondent weak classifiers and are added in to the current strong classifier. Then after all previous features are added, the "Fea-Accu" strong classifier picks out new features from feature pool and adds it into the strong classifier further. The strong classifier constructed by only using these weak classifiers created from previous stage features is already so discriminative that only a small new features are needed to be selected out in current training strong classifier. This can reduce the training time because most train time is consumed on the feature selection process.

Finally, to verify the proposed method, experiment is designed on frontal face detection problem. The experimental results show that it can largely reduce the training time. A frontal face detector with state-of-the-art classification performance can be learned in less than 10 hours.

The paper is organized as follows: section 2 give the introduction of related work of nested cascade; in section 3, Fea-Accu cascade method is described in detail and the difference between it and nested cascade is presented; experiments are showed in section 4.

2. RELATED WORKS

Cascade designed by Viola and Jones expresses the classifier as a cascade of sub-classifiers as shown in Fig.1. Each sub-classifier makes a decision to either reject the input candidate, i.e. classifying it as non-face, or continue evaluation using the next sub-classifier. Those window candidates that survive the last sub-classifier are classified as an object window. Such strategy facilitates the efficient removal of the vast majority of non-object candidates within a minimal amount of computation.

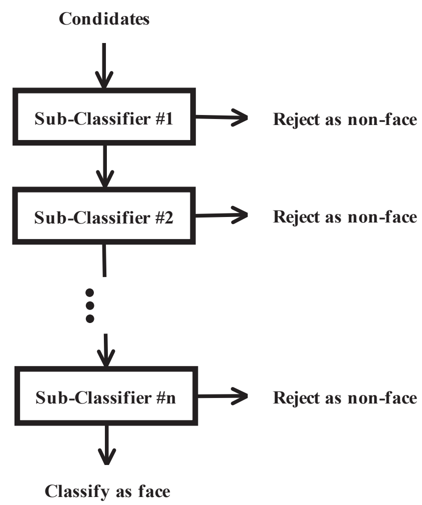


Fig. 1. Structure of the cascaded classifier

In the cascade learning method proposed by Viola and Jones, after a strong stage classifier H(x) is learned, bootstrap strategy is adopted to collecting new training samples for next stage classifier, i.e., only samples with *H(x)=1* are collected to construct the training set of the coming stage. In this way, the neighboring stages classifiers in the cascaded classifier are correlated. However, the correlation is rather loose since all the samples passed the previous stage are treated equally and given the same distribution weight in boosting learning of the coming stage. In fact, the previous learned features and classifiers still have some discriminative power on the new training data and can be used further. In consideration of this, Huang etc have proposed nested cascade in [4] after Viola and Jones' work. They enhanced the use of the classification information of previous stages by introduce a new nested weak classifier as the first weak classifier to encode the previous strong classifier for each stage except the first stage which have no previous classifier. To encoding the confidence learned by previous strong classifier, real value weak classifier is used in nested classifier. The real-value weak classifier output a confidence other than the Boolean classification used by Viola and Jones. The feature value space is divided into equal sub-space as showed in Fig.2. A piece-wise function is learned by RealBoost [18] to approximate complex distribution of training samples. If *f(x)* has been normalized to [0, 1], the range is divided into n sub-regions, the characteristic function is:

$$B_n^{(j)}(u) = \begin{cases} 1 & u \in [j/n, (j+1)/n) \\ 0 & u \notin [j/n, (j+1)/n) \end{cases}, j = 0, \ldots, n-1 \quad (1)$$

The piece-wise function is formally expressed as:

$$h(x) = \sum_{j=0}^{n-1} h^{(j)} B_n^{(j)}(f(x)) \quad (2)$$

The final output classifier of RealBoost is:

$$H(x) = sign[\sum_{t=1}^{T} h_t(x) - b], \quad (3)$$

Where b is a threshold whose default is 0. The confidence of H(x) is:

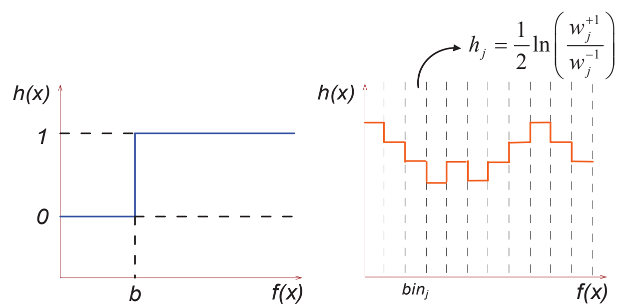
$$Conf(x) = \sum_{t=1}^{T} h_t(x) - b \quad (4)$$

Derived from the piece-wise weak classifier, a novel nested weak classifier $h_{nested}(conf(x))$ is defined based on *conf(x)*. With the help of nested weak classifiers, the confidence value is inherited by the coming stage as:

$$conf_n(x) = h_{nested}(conf_{n-1}(x)) + \sum_{t=1}^{T} h_t^{(n)}(x) - b \quad (5)$$

where $h_t^{(n)}$ is the t-th weak classifier of stage n.

$$H_n(x) = sign(conf_n(x)). \quad (6)$$



(a) Threshold-type function (b)Piece-wise function

Fig.2. Two types of weak classifiers

3. FEA-ACCU CASCADE

As described in section 2, in nested cascade, except for the 1st stage classifier, it adds a nested weak classifier to encode the classification information which has been learnt in the previous classifier. Different from it, we propose alternative way to reuse previously learned information which can promote train speed largely. We propose to only use the features learned previously, rather than the previously learned weak classifier and the strong classifier. This is the key idea of Fea-Accu cascade. we argue that it can enhance the correlation better. An illustration of the method is shown in Fig.3. K_i is the feature number of the i-th stage, C_i is the i-th strong classifier. For i-th stage, the features selected by previous stages are renewed provided with current newly collected non-face samples. Boosting is used to create new weak classifier for previously selected features. In following we will give explanations on how to use boosting to create a weak classifier for a specific feature and why our method can promote the training speed better.

Analytically, Boosting can be considered as consisting of two phases, feature selection and using the features in sequence to create each weak classifiers and combining them by an additive model to build the final strong classifier as shown in equation (3). Empirically, the first phase, feature selection, is an iterative process and every iteration is time-consuming, for each iteration of one feature selection it must check all features in a large-scale training feature pool to test the classification performance on current training sample distribution. But the second step is very fast, it only needs to test the classification performance and compute the weak classifier corresponding for each feature. So Boosting can be used to create the weak classifiers and combining them into a linear strong classifier for previously selected Haar-like features.

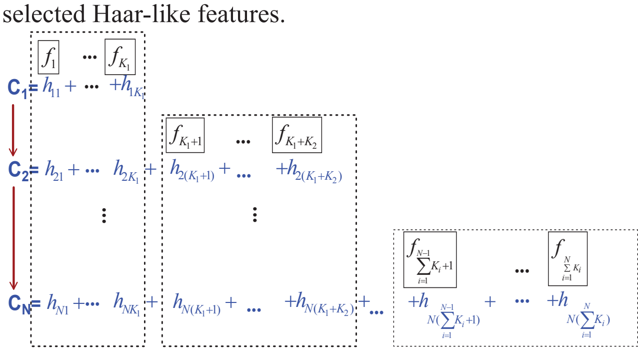


Fig.3. The stage classifier of Fea-Accu cascade and their components

Until now, we have described the training process of the Fea-Accu cascade. In detection, when we apply the trained Fea-Accu cascade, there are less feature evaluation for each strong classifier except the first classifier in Fea-Accu cascade, the only extra computation introduced is that a look-up-table caused by the weak classifiers of previous Haar-like features and an addition for confidence accumulation of strong classifier function.

4. EXPERIMENT

The proposed Fea-Accu cascade is investigated on frontal face detection. There are 20,000 face samples and 10,000 non-face samples for each stage learning. All samples are resized to 24*24. With respect to Haar-like feature pool, there are totally 31,728 Haar-like features in it.

To investigate the previously information reusing extent, the "Fea-Accu" strong classifier and nested weak classifier without picking out any new feature are compared. adding new Haar-like features than using nested strong classifier.. The experiment is done on the 10-th stage of the cascaded classifier. Fig.4 shows confidence VS sample number. From which one can see that Fea-Accu cascade learn a strong classifier more discriminative than nested cascade. To compare the discriminative extents numerically, the Bhattacharyya distances [19] of them are computed for nested weak classifier and Fea-Accu strong classifier. The Bhattacharyya distance is an important criterion evaluating the diversity or the similarity of two random variables based on their probabilistic distributions. For details of Bhattacharyya distance, refer to [19]. By quantify the real value confidence into 41 bins, we get the approximate Bhattacharyya distances. For nested cascade the Bhattacharyya distance is 201.56. While for the proposed method, the Bhattacharyya distance is just 243.8.

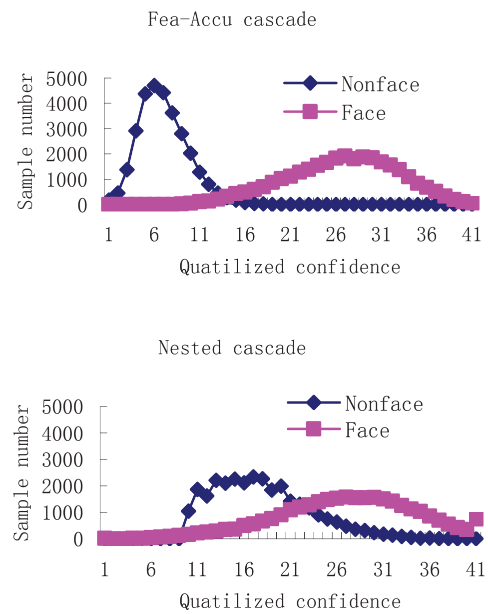


Fig.4. Confidence VS sample number

Further, with respect to newly selected out Haar-like features, there are less features for each stage of Fea-Accu cascade than nested cascade. To verify this, two detectors are learned by using nested cascade and Fea-Accu cascade separately. The feature numbers for each stage of both cascaded detectors are shown in Fig.5. Only nearly 10 hours are needed to train a Fea-Accu cascaded detector with false alarm rate of 1/1,000,000, which is only half of that of the nested cascade, which needs 40 hours to meet the same false alarm rate.

To get a sense on the classification performance of both method, we test the detector on CMU Frontal face test set which including 130 images with 507 faces. The ROC curves of both methods are shown in Fig.6. In the figure, to compare with other methods reported, the ROC curves of them are shown in Fig.6 too. It can be seen that our method gets the best performance.

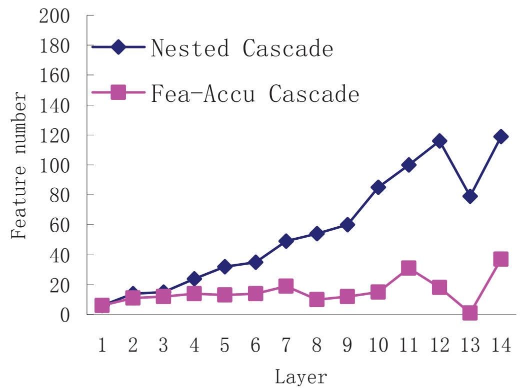


Fig.5. stage vs. number of features

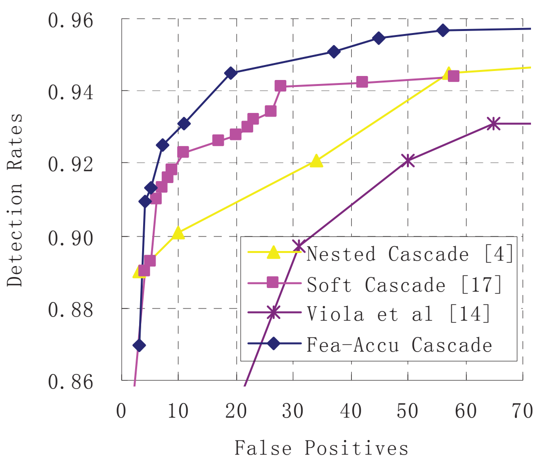


Fig.6. ROC curves of different methods

5. CONCLUSION

This paper proposes a novel cascade structure called Fea-Accu cascade for fast training of cascaded classifier. Fea-Accu cascade enhances the correlation among different stage classifiers by reusing features learned in preceding stage classifiers. Experimental results on frontal face detection problem show its efficiency and effectiveness. Based on Fea-Accu cascade, a robust face detection system is developed quickly.

ACKNOWLEDGEMENTS

This paper is partially supported by Natural Science Foundation of China under contracts No.60832004 and No.60872124; Grand Program of International S&T Cooperation of Zhejiang Province S&T Department under contract No. 2008C14063; Co-building Program of Beijing Municipal Education Commission; and ISVISION Technology Co. Ltd.

REFERENCES

1.B. Heisele, T. Poggio, and M. Pontil. Face Detection in Still Gray Images. CBCL Paper #187. MIT, Cambridge, MA, 2000.

2.R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, Face detection in color images, IEEE Trans. Pattern Analysis and Machine Intelligence, pp.696–706, 2002.

3.C. J. Liu. A Bayesian Discriminating Features Method for Face Detection, IEEE Trans. Pattern Anal. and Machine Intel., pp. 725-740. 2003.

4.C. Huang, H. Z. Ai, B. Wu, S. H. Lao: Boosting Nested Cascade Detector for Multi-View Face Detection. International Conference on Pattern Recognition (2) 2004: 415-418.

5.E. Osuna, R. Freund, and F. Girosi, Training support vector machines: An application to face detection, IEEE Conf. on Computer Vision and Pattern Recognition, pp. 130–136. 1997,

6.C. P. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in Proc. 6th Int. Conf. Computer Vision, pp.555–562. 1998,

7.H. A. Rowley, S. Baluja, and T. Kanade. Neural Network-Based Face Detection. IEEE Trans. Pattern Analysis and Machine Intelligence, pp. 23-38. 1998.

8.H. Schneiderman and T. Kanade. A Statistical Method for 3D Object Detection Applied to Faces. Comp. Vision and Pattern Recognition, pp. 746-751. 2000.

9.H. A. Rowley, S. Baluja, and T. Kanade. Rotation Invariant Neural Network-Based Face Detection. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 38-44. 1998.

10.K. K. Sung, and T. Poggio. Example-Based Learning for View-Based Human Face Detection. IEEE Trans. Pattern Analysis and Machine Intelligence. pp. 39-51. 1998.

11.M. H. Yang, D. Roth, and N. Ahuja. A SNoW-Based Face Detector. Advances in Neural Information Processing Systems 12, MIT Press, pp. 855-861. 2000.

12.S. Li and Z. Zhang. Floatboost learning and statistical face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp.1112 – 1s123, 2004.

13.C. Liu, H. Y. Shum. Kullback-Leibler Boosting. IEEE Conf. on Computer Vision and Pattern Recognition. 2003.

14.P. Viola and M. Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 511-518. 2001.

15.R. Xiao, M. J. Li, H. J. Zhang. Robust Multipose Face Detection in Images, IEEE Trans on Circuits and Systems for Video Technology, Vol.14, No.1 pp. 31-41. 2004,

16.H. Schneiderman, Feature-centric evaluation for efficient cascaded object detection, in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 29—36, 2004.

17.L. Bourdev, J. Brandt, Robust object detection via soft cascade. IEEE Conf. on Computer Vision and Pattern Recognition, vol.2, pp 236- 243, 2005.

18.R. E. Schapire and Y. Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions", Machine Learning, 37, 1999, 297-336.

19.S. Theodoridis and K. Koutroumbas, Pattern Recognition, Elsevier Science, 2003, 177-179.