딥 러닝을 통한 얼굴 크기 탐지

바트홍고르 체뎅*, 이해연*
*금오공과대학교 소프트웨어공학과
e-mail:batkhongop@gmail.com

Face Size Detection using Deep Learning

Batkhongor Tseden*, Hae-Yeoun Lee*
*Dept of Software Engineering, Kumoh National Institute of Technology

Abstract

Many deep learning approaches are studied for face detection in these days. However, there is still a performance problem to run efficiently on devices with limited resources. Our method can enhance the detection speed by decreasing the number of scaling for detection methods that use many different scaling per image to detect the different size of faces. Also, we keep our deep learning model easy to implement and small as possible. Moreover, it can be used for other special object detection problems but not only for face detection.

1. Introduction

Detecting faces in different sizes is one of the most important part in every single face detection methods. Even there is only one big sized face, current some methods check a number of different sizes of faces [1]. It spends resource for unnecessary operations and increase detection time. As shown in Figure 1, DDFD [2], for example, scales image 10 times from bigger to smaller.

Figure 1. Detected faces using DDFD [2]

Our method is trying to solve that problem by finding face size and it is possible to use minimal number of scaling for face detection. Moreover, Figure 2 shows heat-map on each face detection scales, here we can see that there is higher probability to detect wrong face at some scales. Therefore, if we find right sizes, it is able to decrease false positive face detection. First of all, we proposed to find minimum face size in this paper.
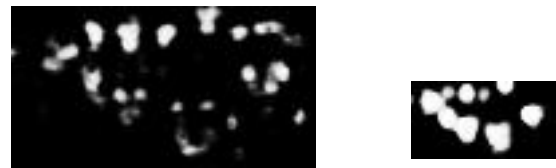

Figure 2. heap-map at scale 6.2 and 3.15, respectively on Figure 1

2. Proposed algorithm to detect face size

We created training data using AFLW dataset [3] images for image scaling and labeled images between 0 – 4 according to proportion of minimum face size and image size as shown in Figure 3.
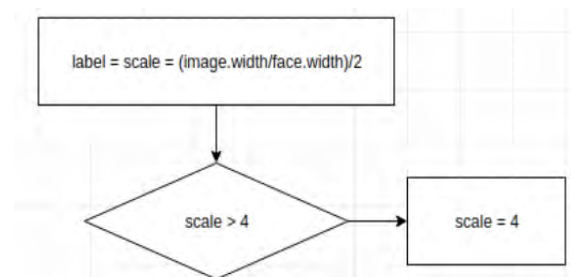

Figure 3. Label calculating for training images

We used Alexnet [4] as deep learning model and Caffe [5] as library for our face size detection method. Then, we trained Alexnet and modified Alexnet which has same structure but less parameters.

3. Experimental results and conclusion

Using PASCAL Face [6], AFW [7] datasets, both models were tested on nVIDIA GeForce GTX 1080 TI. The accuracy analysis result is shown in Table 1.

Table 1. Comparison of 2 trained models

Model Name	Model size	Images per second	Accuracy on Pascal [7]	Accuracy on AFW [8]
Alexnet	222Mb	5	86%	48%
Modified Alexnet	40Mb	12	85%	48%

To use this method as part of face detection, its accuracy must be more than 90 percent, but it could not perform like we expected. However, there were only 18,000 images for training which is not enough to perform excellent. We need to increase number of images on training by combining AFLW [3] and FDDB [8] datasets to improve accuracy. Furthermore, there is need to use other methods like subsampling to increase the number of images.

Acknowledgement

References

[1] P. Viola and M. J. Jones. Robust real-time face detection. International Journal of Computer Vision, 2004.

[2] S. Farfade, M. Saberian, Li-Jia Li. Multi-view Face Detection Using Deep Convolutional Neural Networks 2015

[3] Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization Martin Koestinger, Paul Wohlhart, Peter M. Roth and Horst Bischof In Proc. First IEEE International Workshop on Benchmarking Facial Image Analysis Technologies, 2011

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. 2012.

[5] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell. Caffe: Convolutional architecture for fast feature embedding. 2014

[6] J. Yan, X. Zhang, Z. Lei, and S. Li. Face detection by structural models.

[7] D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In Proceedings of CVPR, 2012

[8] Vidit Jain and Erik Learned-Miller. FDDB: A Benchmark for Face Detection in Unconstrained Settings.Technical Report UM-CS-2010-009, Dept. of Computer Science, University of Massachusetts, Amherst. 2010.