Midterm Report

Quan Wang*, Tingwu Wang*, Wenxin Wang*, and Tianpen Li*
*Department of Electronic Engineering, Tsinghua University, Beijing, 100084, P. R. China
E-mail: wtw12@mails.tsinghua.edu.cn, wangquan.thu@aliyun.com, stieizc.33@gmail.com, 370756595@qq.com

I. Paper Reading and coding

During the past months, we read the following papers:

1) Benenson's Group: Benenson's group provides the state-of-art performance of real time pedestrian detection [1]–[5]. But it is also clear that the gpu required codes are not suitable for reproducing. It is difficult to implement, as none of us are equipped with the required hardware.

2) Subhransu Maji's Group: The paper [6] discuss straightforward classification using kernelized SVMs They evaluate the kernel for a test vector and each of the support vectors. But the algorithm is kind of out of date. And although we succeeded in rerunning it, but we choose to drop it.

3) Ouyang's Group: After reading [7]–[11], we believe it is the right group to follow. In their work, a multi-pedestrian detector is learned with a mixture of deformable part-basedmodels to effectively capture the unique visual patterns appearing in multiple nearby pedestrians. The training data is labeled as usual, i.e. a bounding box for each pedestrian. The spatial configuration patterns of multiple nearby pedestrians are learned and clustered into mixture component. In the multi-pedestrian detector, each single pedestrian is specifically designed as a part, called pedestrian-part. A new probabilistic framework is proposed to model the configuration relationship between results of multi-pedestrian detection and 1-pedestrian detection. With this framework, multi-pedestrian detection results are used to refine 1-pedestrian detection results.

II. A github homepage

We set up a github homepage https://github.com/WilsonWangTHU/pedestrian-detection-thu2015.

And currently we are revising the code for the first stable commit.

III. What next?

There are several few more points possible to be revised.

A. A Convolutional implementation?

Clearly the pedestrian detection could do more, like pipline a Fast-RCNN to further detect the pose and clothes of the pedestrians. It is clear the results of [11] could be further used as a bounding box or ROI. I have experience of using Fast-RCNN and caffe pipeline, and clearly the results of the [11] could be used as the input of a new deep network.

B. Contextual information

Convolutional network could further improved if we could use some additional contextual information. The work of [11] did not, and it could possible be helpful.

References

[1] S. Zhang, R. Benenson, and B. Schiele, "Filtered channel features for pedestrian detection," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[2] J. Hosang, R. Benenson, and B. Schiele, "How good are detection proposals, really?" in *British Machine Vision Conference (BMVC)*, British Machine Vision Association. British Machine Vision Association, September 2014.

[3] J. Hosang, R. Benenson, P. Dollár, and B. Schiele, "What makes for effective detection proposals?" *arXiv*, 2015.

[4] J. Hosang, M. Omran, R. Benenson, and B. Schiele, "Taking a deeper look at pedestrians," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[5] M. Mathias, R. Benenson, R. Timofte, and L. Van Gool, "Handling occlusions with franken-classifiers," in *ICCV*, 2013.

[6] S. Maji, A. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, June 2008, pp. 1–8.

[7] X. Zeng, W. Ouyang, M. Wang, and X. Wang, "Deep learning of scene-specific classifier for pedestrian detection," in *Computer Vision–ECCV 2014*. Springer, 2014, pp. 472–487.

[8] W. Li, R. Zhao, T. Xiao, and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 152–159.

[9] P. Luo, Y. Tian, X. Wang, and X. Tang, "Switchable deep network for pedestrian detection," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 899–906.

[10] W. Ouyang and X. Wang, "Joint deep learning for pedestrian detection," in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2056–2063.

[11] X. Zeng, W. Ouyang, and X. Wang, "Multi-stage contextual deep learning for pedestrian detection," in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 121–128.