Research Proposal

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. Introduction

Pedestrian detection is a key problem in computer vision, with several applications that have the potential to positively impact quality of life.

Pedestrian detection is a difficult task from a machine vision perspective. The lack of explicit models leads to the use of machine learning techniques, where an implicit representation is learned from examples. In this proposal, we summarize the different perspectivs of studying Pedestrian Detection problem with the help of several well written survey [1]–[3]. And illustrate what we could do during the work.

II. Previous Study

There are numerous papers studying the problems from different perspectives, which includes:

ROI Selection: The simplest technique to obtain initial object location hypotheses is the sliding window technique, where detector windows at various scales and locations are shifted over the image.

Classification: After a set of initial object hypotheses has been acquired, further verification (classification) involves pedestrian appearance models, using various spatial and temporal cues

Tracking: There has been extensive work on the tracking of pedestrians to infer trajectory-level information. One line of research has formulated tracking as frame-by-frame association of detections based on geometry and dynamics without particular pedestrian appearance models

III. Current Resources

We currently have over 50 papers on the Pedestrian Detection Problem, all of them from the ICCV, ECCV, CVPR after 2011. We have access to over 6 large data bases and more than 5 open source source codes on the problem.

A. Related Work

1) Haar Wavelet-Based Cascade: The Haar wavelet-based cascade framework provides an efficient extension to the sliding window approach by introducing a degenerate decision tree of increasingly complex detector layers.

2) Neural Network Using Local Receptive Fields: Adaptive local receptive fields (LRF) have been shown to be powerful features in the domain of pedestrian detection, in combination with a multilayer feed-forward neural network architecture (NN/LRF).

3) Histograms of Oriented Gradients with Linear SVM (HOG/linSVM): We follow the approach of Dalal and Triggs to model local shape and appearance using well-normalized dense histograms of gradient orientation

4) Combined Shape-Texture-Based Pedestrian Detection: We consider a monocular version of the real-time PROTECTOR system by cascading shape-based pedestrian detection with texture-based pedestrian classification.

B. Data Set

Multiple public pedestrian datasets have been collected over the years, which includes: INRIA, ETH, TUD-Brussels, Daimler (Daimler stereo), Caltech-USA, and KITTI.

IV. What Could We Do

A. Groups to Follow

To efficiently get started and not to be drowned by the numerous papers, we focus on several groups that are consistently researching the problems, which includes: Dr. Xiaogang Wang [4]–[10] and Dr. Rodrigo Benenson from Max-Planck-Institut für Informatik [11]–[17]

B. Time Table

We sort out and finish the literature review by Week 9. We read and revise by week 11, after which we will focus on the developing of our won algorithms. The research point could be applying deep learning, but no conclusion could be made before a full understanding of literature is accomplished.

References

[1] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 12, pp. 2179–2195, 2009.

[2] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 4, pp. 743–761, 2012.

[3] R. Benenson, M. Omran, J. Hosang, and B. Schiele, "Ten years of pedestrian detection, what have we learned?" in *Computer Vision - ECCV 2014 Workshops*, ser. Lecture Notes in Computer Science, L. Agapito, M. M. Bronstein, and C. Rother, Eds. Springer International Publishing, 2015, vol. 8926, pp. 613–627.

[4] 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.

[5] 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.

[6] 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.

[7] 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.
[8] 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.
[9] W. Ouyang and X. Wang, "A discriminative deep model for pedestrian detection with occlusion handling," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 3258–3265.
[10] P. Luo, X. Wang, and X. Tang, "Pedestrian parsing via deep decompositional network," in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2648–2655.
[11] 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.
[12] 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.
[13] J. Hosang, R. Benenson, P. Dollár, and B. Schiele, "What makes for effective detection proposals?" *arXiv*, 2015.
[14] 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.
[15] M. Mathias, R. Benenson, R. Timofte, and L. Van Gool, "Handling occlusions with franken-classifiers," in *ICCV*, 2013.
[16] R. Benenson, M. Mathias, R. Timofte, and L. Van Gool, "Fast stixels estimation for fast pedestrian detection," in *ECCV, CVVT workshop*, 2012.
[17] ——, "Pedestrian detection at 100 frames per second," in *CVPR*, 2012.