Replication Studies on a State-of-the-art Part-based Human Detector

Yu-Wei Chao
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109, USA
ywchao@umich

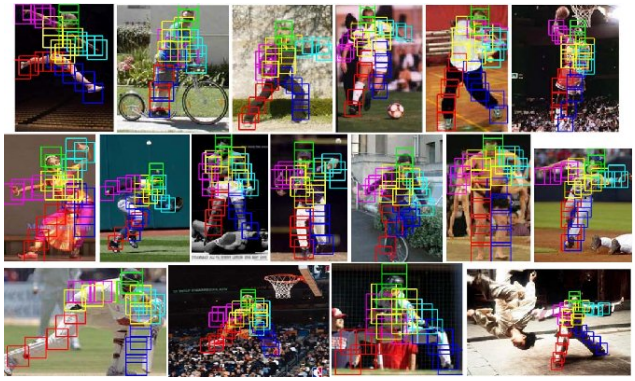

Figure 1: Illustration of Yang and Ramanan's work on human detection and pose estimation [5]. Different color of boxes identify different local parts of human body.

1. Introduction

Human detection and pose estimation is a challenging problem in computer vision studies. Recently there has been many outstanding works published in addressing the these problems [2, 1, 3, 5, 4]. A very interesting line of work is the application of part-based models [3, 5]. Part-based models can be viewed an extension of the rigid template models in the way that the target objects are represented by local parts, and the locations of these local parts have some amount of flexibility to capture the uncertainly in real-world data.

In this work, we proposed to replicate Yang and Ramanan's paper [5] on human detection and pose estimation. Unlike the famous deformable part-based model [3], which addressed generic object detection, Yang and Ramanan focus specifically on human detection in RGB images. They carefully designed a system that can locate the body parts of the human, as illustrated in Fig. 1. They used a mixtures of templates for each parts so they can effetively modeled human poses.

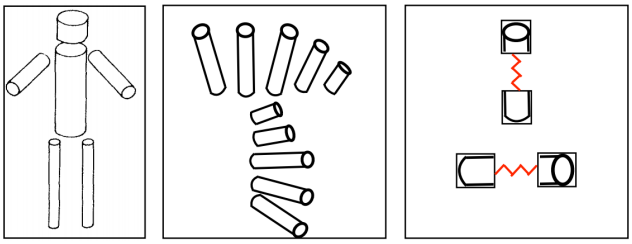

Figure 2: Illustration of the pictorial structure and mixture of local parts.

2. Replication Plan

We list our plans for this replication study as following:

- Re-do the training process. The paper uses a fixed training set provided by the Image Parse dataset and the Buffy Stickman dataset. We plan to re-train the model with different settings, either using a subset or on a new dataset. The goal is to better understand how the trained model changes by different training settings.
- Re-implement the inference algorithm. We plan to investigate the efficiency of the inference algorithm, which is not presented in the paper.
- Build on own human dataset. We can build a small but more challenging dataset, and see how far can the target method go.

3. Milestone

The milestones are shown in Table 1

References

[1] L. Bourdev, S. Maji, T. Brox, and J. Malik. Detecting people using mutually consistent poselet activations. In *Proceedings of the 11th European Conference on Computer Vision*, 2010. 1

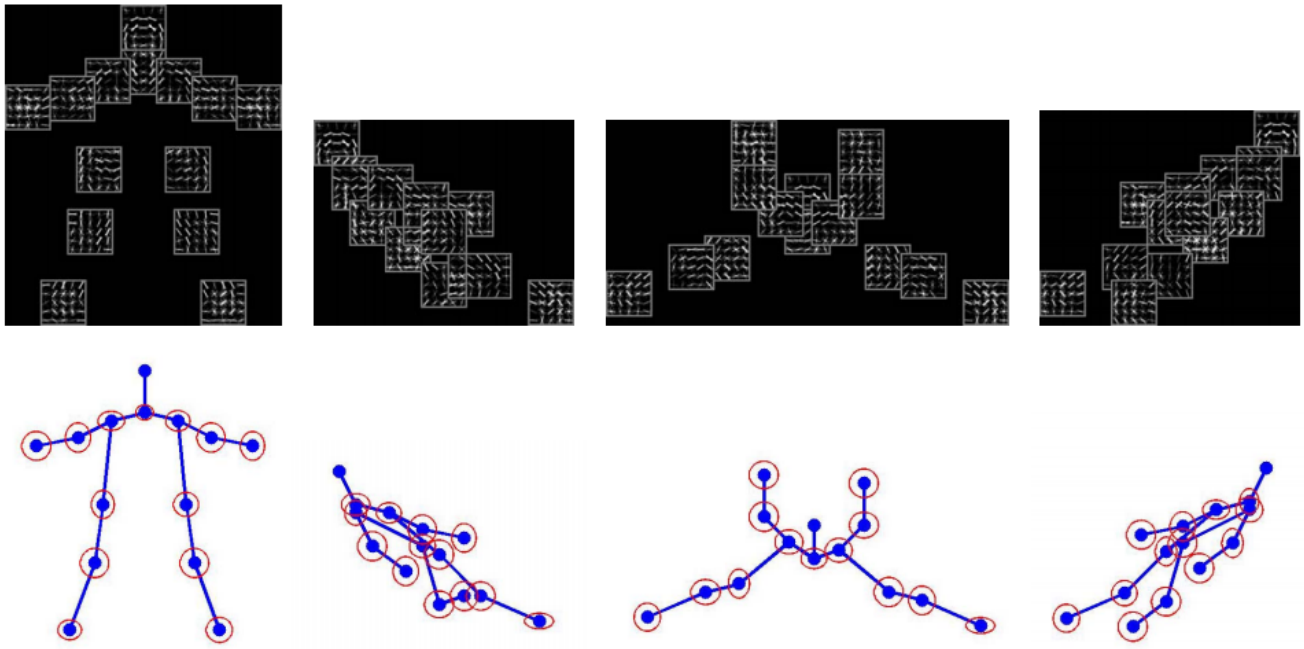

Figure 3: Visualization of the learned model.

Date	Plan
02.05 - 02.11	re-implement the inference algorithm
02.12 - 02.18	re-implement the inference algorithm
02.19 - 02.25	re-implement the inference algorithm compare the result with the paper analyze the result
02.25 - 03.04	re-do training
03.05 - 03.11	re-do training
03.12 - 03.18	collect new dataset
03.19 - 03.25	collect new dataset
03.26 - 04.01	run experiment on new dataset
04.02 - 04.08	run experiment on new dataset
04.09 - 04.15	Buffer time
04.16 - 04.21	Prepare for presentation

Table 1: Project milestones

[2] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In *Proceedings of the IEEE International Conference on Computer Vision*, 2009. 1

[3] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1627–1645, 2010. 1

[4] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011. 1

[5] Y. Yang and D. Ramanan. Articulated human detection with flexible mixtures of parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2878–2890, 2011. 1