

# Robust Optimization for Deep Regression (Supplementary Material)

Vasileios Belagiannis<sup>1,2</sup>, Christian Rupprecht<sup>1,3</sup>, Gustavo Carneiro<sup>4</sup>, and Nassir Navab<sup>1,3</sup>

<sup>1</sup>Computer Aided Medical Procedures, Technische Universität München

<sup>2</sup>Visual Geometry Group, Department of Engineering Science, University of Oxford

<sup>3</sup>Johns Hopkins University

<sup>4</sup>Australian Centre for Visual Technologies, University of Adelaide

*vb@robots.ox.ac.uk, {christian.rupprecht, navab}@in.tum.de, gustavo.carneiro@adelaide.edu.au*

## 1. Introduction

In this next sections, we provide additional numerical results as well as more visual results of our method applied on 2D human pose estimation. In particular, an additional comparison between  $L_2$  and *Tukey's biweight* loss functions follows in Sec. 2 and more visual results are presented in Sec. 3. Finally, we provide visualization of the dataset for the task of age estimation from face images [2].



Figure 1: **Additional results on model refinement - PARSE:** Additional results for the coarse-to-fine model using the cascade of ConvNets are presented from the PARSE dataset [6]. On the top row the result of a single ConvNet is presented, while on the bottom row the refined result using the cascade of ConvNets.

## 2. Additional Comparisons

In the comparison of  $L2$  and *Tukey's biweight* loss functions, we have presented the final result of the full body using the mean squared error (MSE) and PCP (percentage of correctly estimated parts) performance metrics. In this section, we report, in addition, the scores for each body part for the PCP metric. We distinguish two variants of the PCP score according to the literature [5], as described in our paper. The results are summarized in Fig. 2.

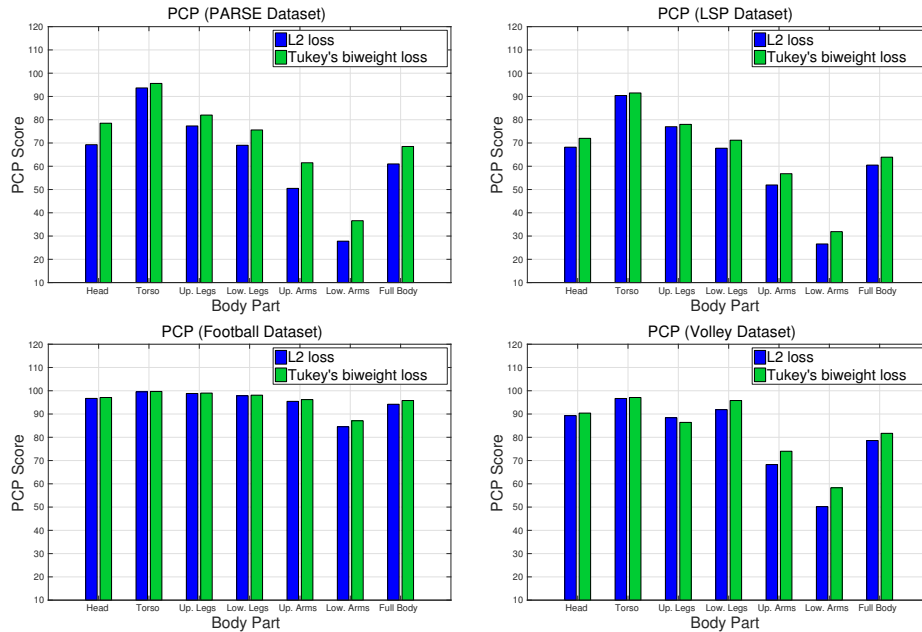


Figure 2: **Further Comparison of  $L2$  and *Tukey's biweight* loss functions:** We compare our results (*Tukey's biweight* loss) with  $L2$  loss for each body part on PARSE [6], LSP [3], Football [4] and Volleyball [1] datasets. In PARSE [6] and LSP [3], the evaluation has been performed using the *strict* PCP, while in Football [4] and Volleyball [1] using the *loose* PCP.

## 3. More Results

In this section, additional visual results using our coarse-to-fine model are presented for the task of 2D human pose estimation. In Fig. 1 and 3, the refined results using our coarse-to-fine model are illustrated. Moreover, a few cases of failures are presented in Fig. 5. Finally, we provide additional results for each dataset in Fig. 4, 6, 7 and 8.

## 4. Age Estimation Dataset

In Fig. 9, we provide visual samples from the Apparent Age Estimation dataset [2] for age estimation from face images. We have used this dataset for demonstrating the generalization of our robust loss into different regression tasks.

## References

- [1] V. Belagiannis, C. Amann, N. Navab, and S. Ilic. Holistic human pose estimation with regression forests. In *AMDO*, 2014. 2, 7
- [2] S. Escalera, X. Gonzalez, Jordi Baró, and I. Guyon. Chalearn looking at people challenge 2015: Dataset and results. In *ICCV*. 1, 2, 7
- [3] S. Johnson and M. Everingham. Clustered pose and nonlinear appearance models for human pose estimation. In *BMVC*, 2010. 2, 3, 5
- [4] V. Kazemi, M. Burenius, H. Azizpour, and J. Sullivan. Multi-view body part recognition with random forests. In *BMVC*, 2013. 2, 6
- [5] L. Pishchulin, A. Jain, M. Andriluka, T. Thormahlen, and B. Schiele. Articulated people detection and pose estimation: Reshaping the future. In *CVPR*, 2012. 2
- [6] Y. Yang and D. Ramanan. Articulated human detection with flexible mixtures of parts. *TPAMI*, 2013. 1, 2, 4

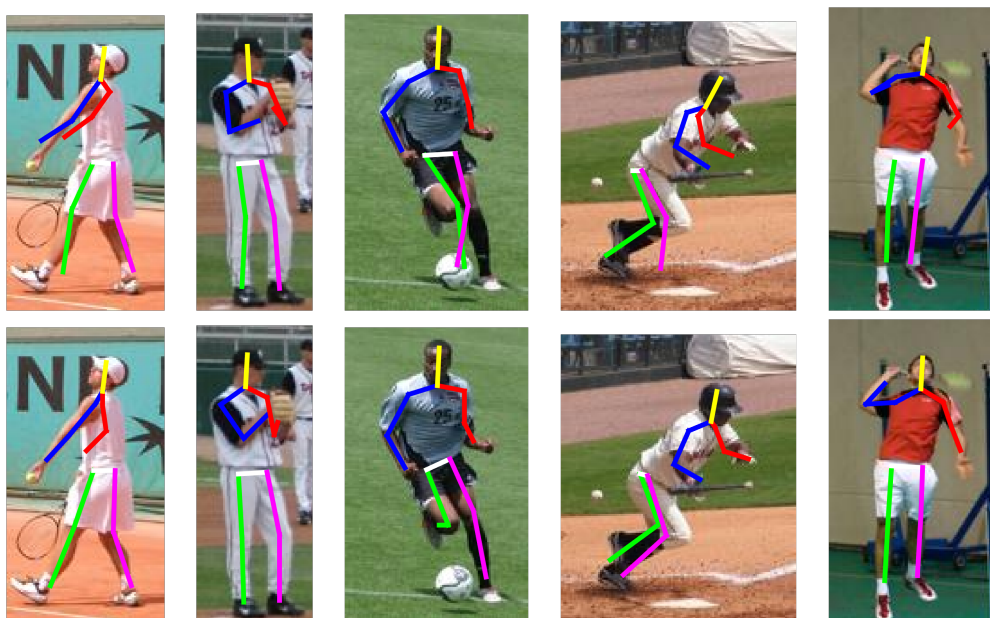


Figure 3: **Additional results on model refinement - LSP:** Additional results for the coarse-to-fine model using the cascade of ConvNets are presented from the LSP dataset [3]. On the top row the result of a single ConvNet is presented, while on the bottom row the refined result using the cascade of ConvNets.

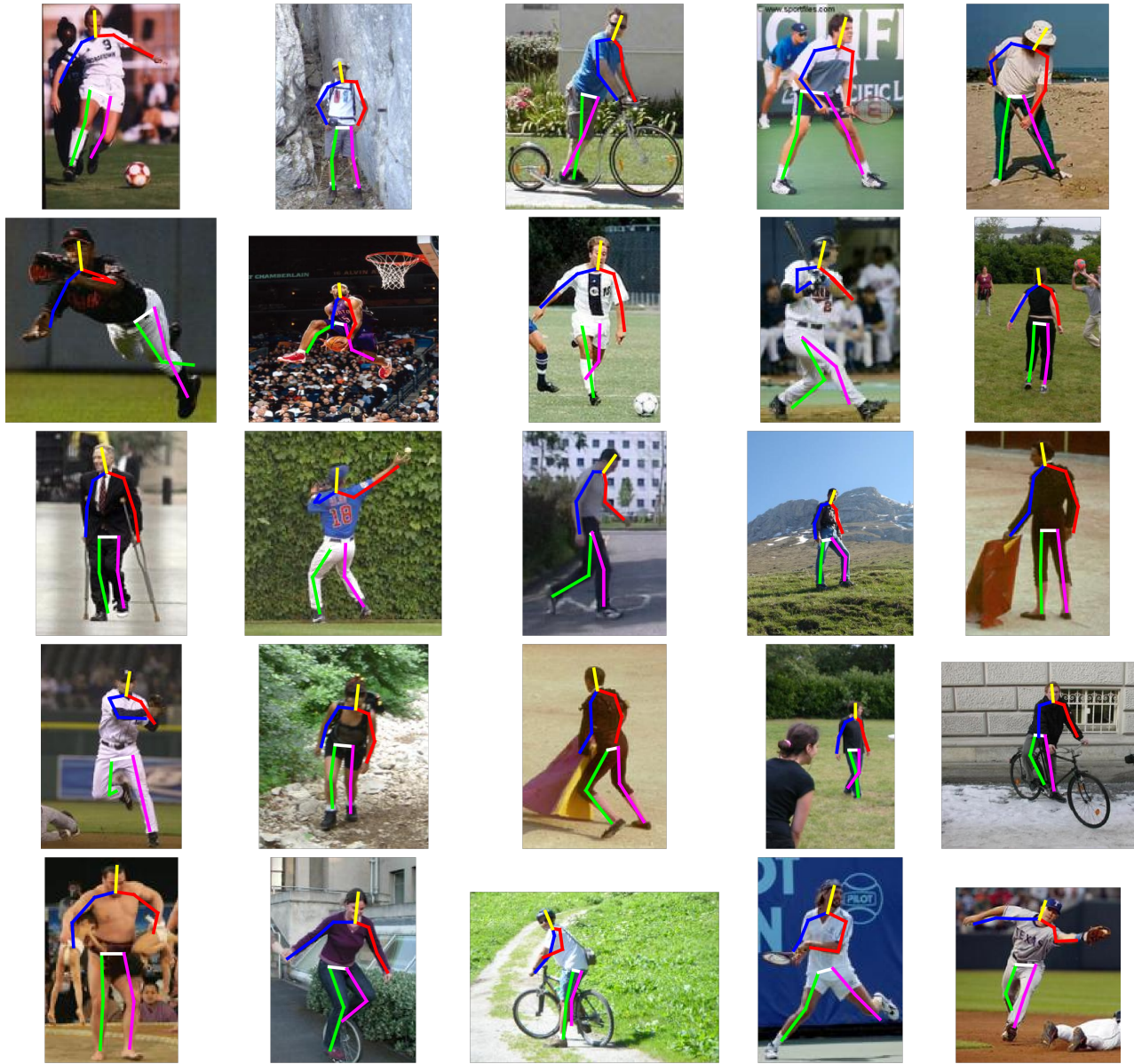


Figure 4: **Additional results on PARSE [6]:** Samples of our results using the coarse-to-fine model on 2D human pose estimation are presented.

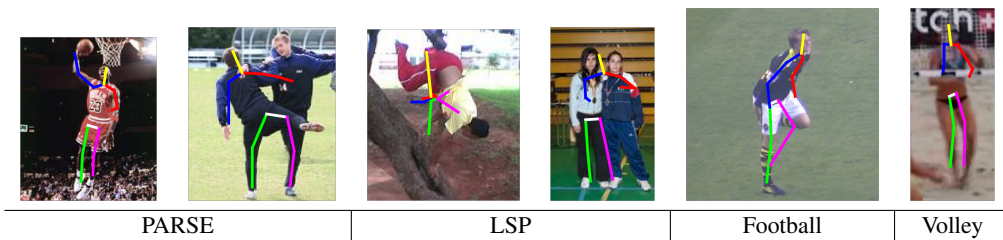


Figure 5: **Failure cases:** There have been some cases where our coarse-to-fine model did not successfully regress the body joints.

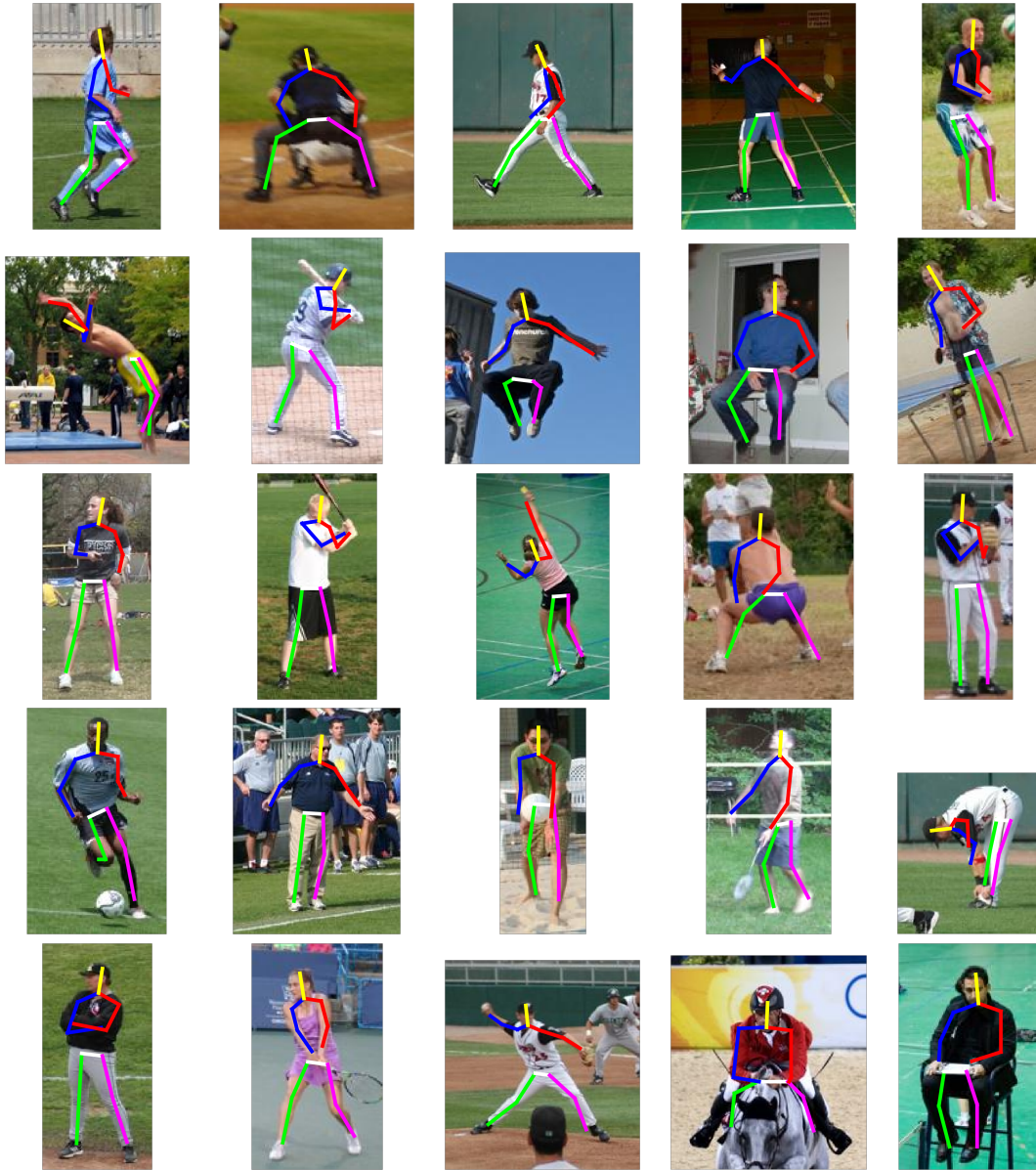


Figure 6: **Additional results on LSP [3]:** Samples of our results using the coarse-to-fine model on 2D human pose estimation are presented.



Figure 7: **Additional results on Football [4]:** Samples of our results using the coarse-to-fine model on 2D human pose estimation are presented.



Figure 8: **Additional results on Volleyball [1]**: Samples of our results using the coarse-to-fine model on 2D human pose estimation are presented.



Figure 9: **Age estimation dataset**: Samples from the apparent age estimation dataset [2], used for the regression task of age estimation from face images.