Supplementary material for "Contextual Part Rescoring for Human Pose Estimation"

BMVC 2014 Submission # 456

This supplementary material accompanies the paper "Contextual Part Rescoring for Human Pose Estimation". First, we show the outputs of our proposed basic part detection rescoring mechanism for some sample images. Secondly, we include an extended set of human pose estimation qualitative results on the UIUC Sports [2] and LSP [1] datasets.

1 Contextual part rescoring

In order to illustrate the benefits of our proposed contextual rescoring mechanism in human pose estimation, Figures 1- 3 show the outputs from our contextual rescoring mechanism for different basic parts (body joints) in different sample images with different pose configurations. More specifically, we compare our rescored score maps against the input "original" score maps for different body joints, obtained by scanning the corresponding set of HOG filters over the image. For clarity, we only show the original and rescored score maps from a certain mixture component for each basic part. The mid-level representation used to generate the rescored score maps was defined by our proposed poselet selection method, maximizing precision and enforcing a covering of a validation set of images.

While not perfect, the output from our contextual rescoring method, produces score maps which are more spatially consistent with the actual locations of the body parts. Our rescoring is able to "fade out" false-positive high-scoring original detections (head detections are a clear example of this). In the same way, our method also hallucinates the actual location of a body part when the original scores are poor (see left wrist detection in Figure 3).

2 Qualitative results

Figures 4- 7 show qualitative human pose estimation results for the UIUC Sports dataset [2]. We compare the results from our proposed extended pictorial structure model in contrast to those from Yang and Ramanan [3]. When using our extended pictorial structure formulation, we define the mid-level representation by means of our proposed poselet selection method. Figures 8- 11 show qualitative results for the LSP dataset [1]. We separate cases where our rescoring method achieves better results in the upper and lower body.

References

[1] Sam Johnson and Mark Everingham. Clustered pose and nonlinear appearance models for human pose estimation. In *BMVC*, pages 12.1–11, 2010.

[2] Yang Wang, Duan Tran, and Zicheng Liao. Learning hierarchical poselets for human parsing. In *CVPR*, pages 1705–1712, 2011.

[3] Yi Yang and D. Ramanan. Articulated human detection with flexible mixtures of parts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(12):2878–2890, Dec 2013.

046
047
048
049
050
051
052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
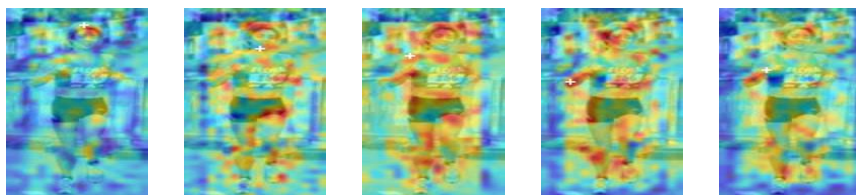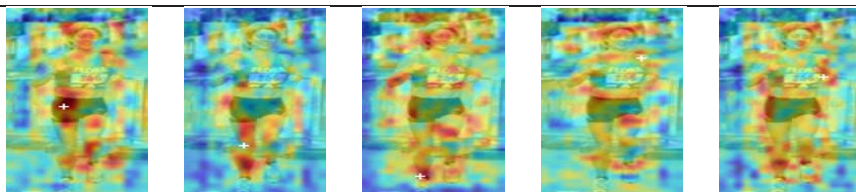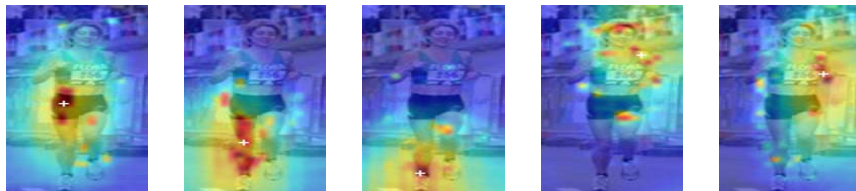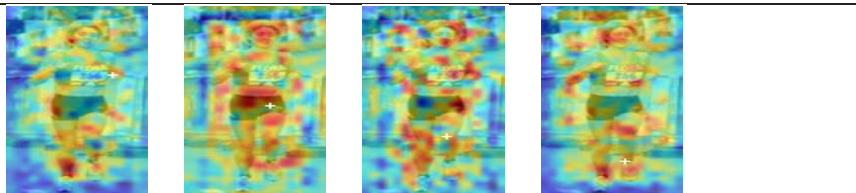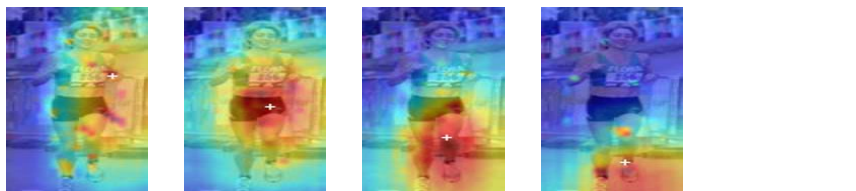
Figure 1: Contextual rescoring in a frontal pose. Odd rows show original detection score maps for different basic body parts. Even rows show the output from our contextual rescoring method. Bluish pixels correspond to low scoring locations, while reddish pixels correspond to high scoring ones. White crosses show ground-truth annotations for each basic part (body joints).
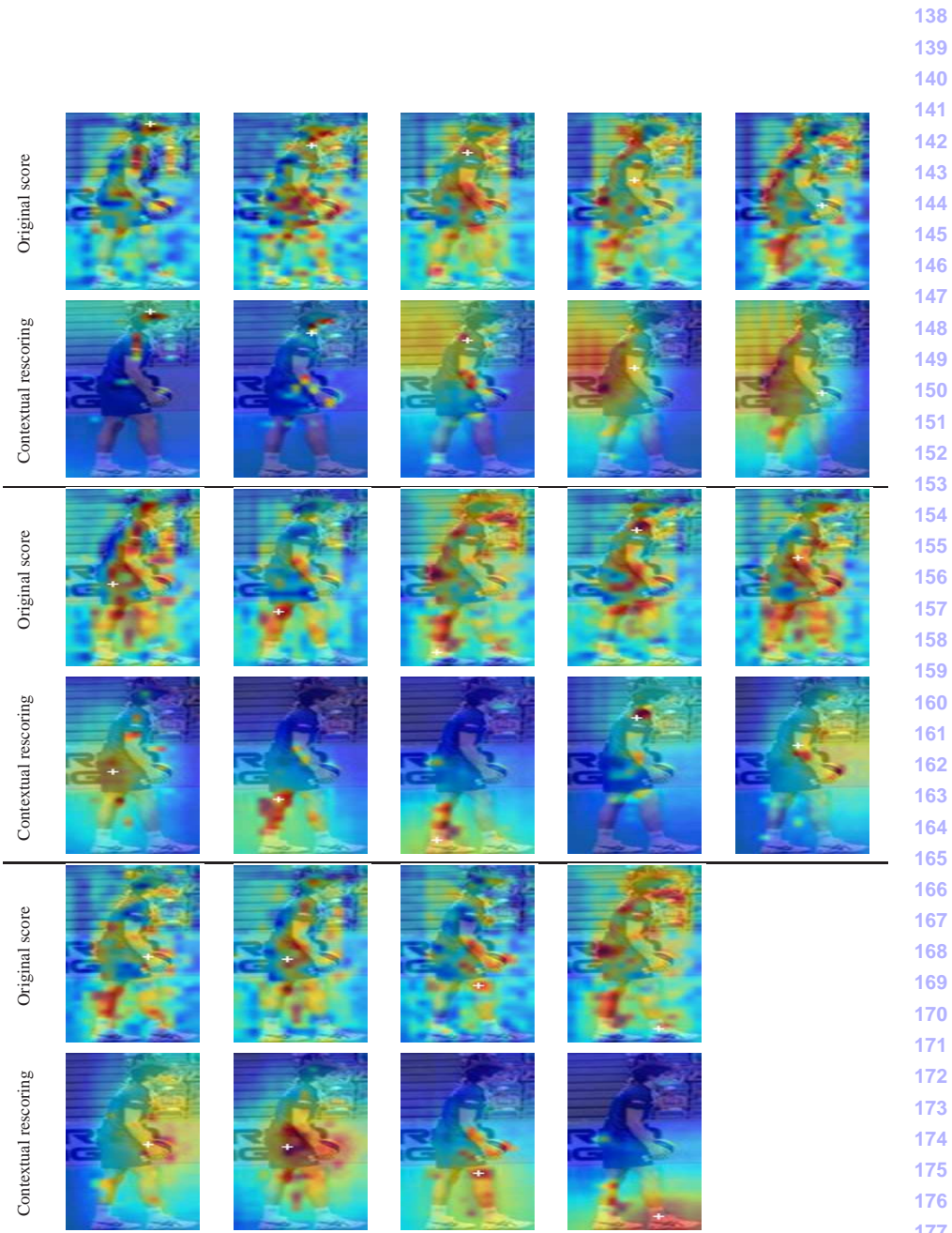
Figure 2: Contextual rescoring in a right-facing profile pose. See Figure 1 for an explanation of the visualization legend.
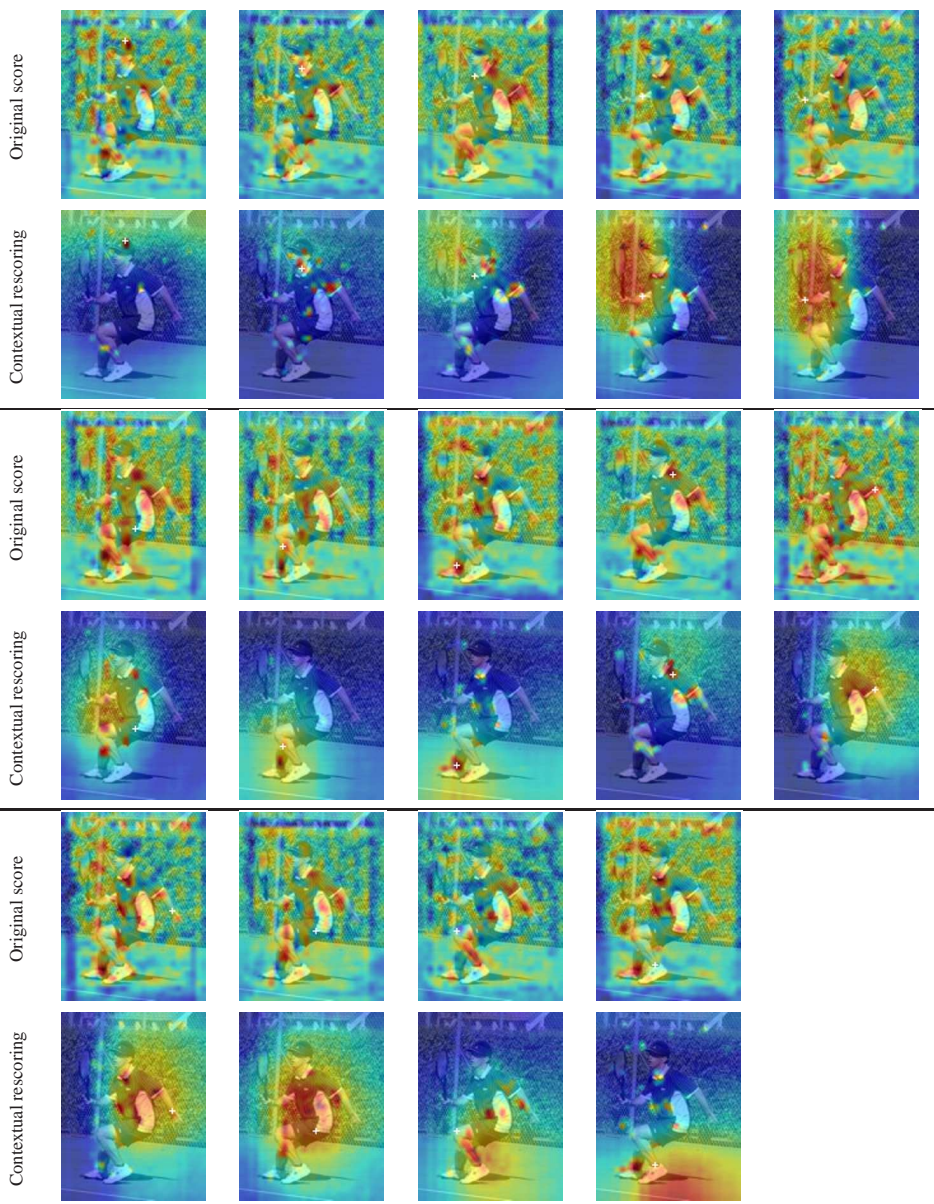
Figure 3: Contextual rescoring in a left-facing profile pose. See Figure 1 for an explanation of the visualization legend.
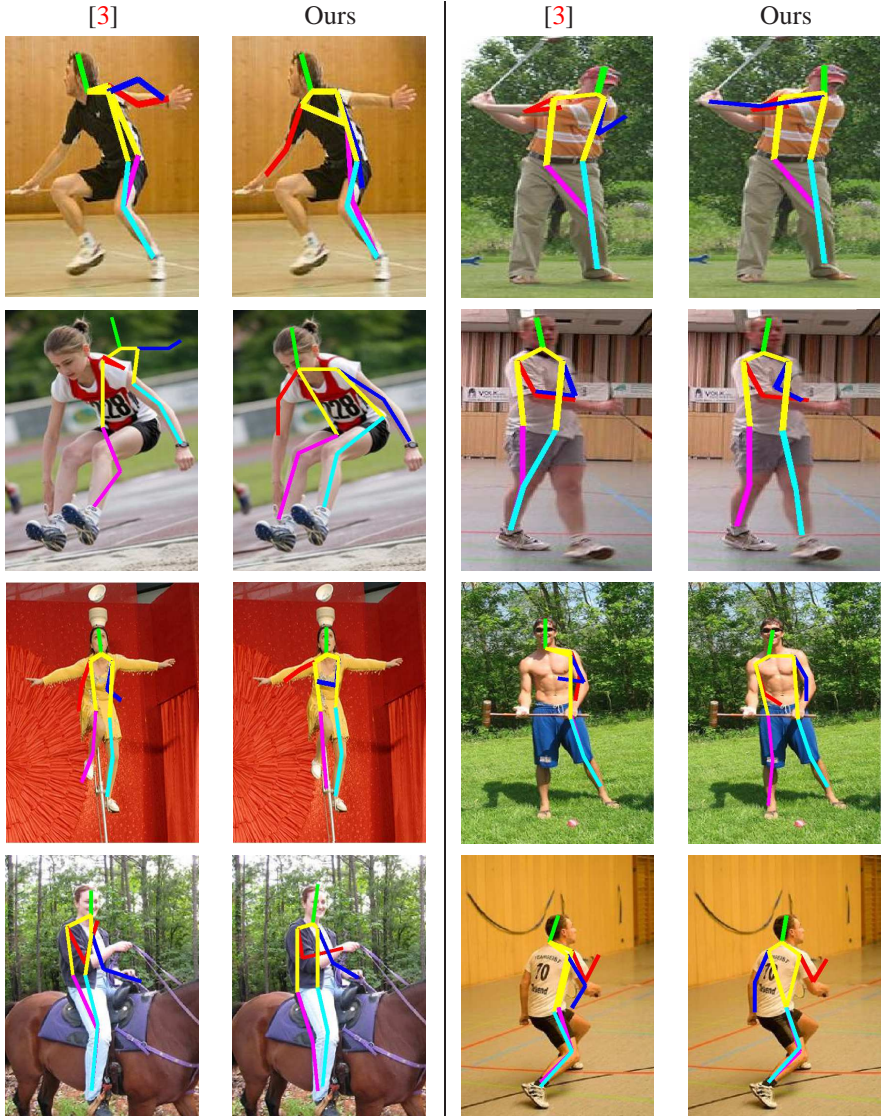
Figure 4: Qualitative results for the UIUC Sports dataset. Left image from each pair of images shows the result from [3] and right image shows our results.
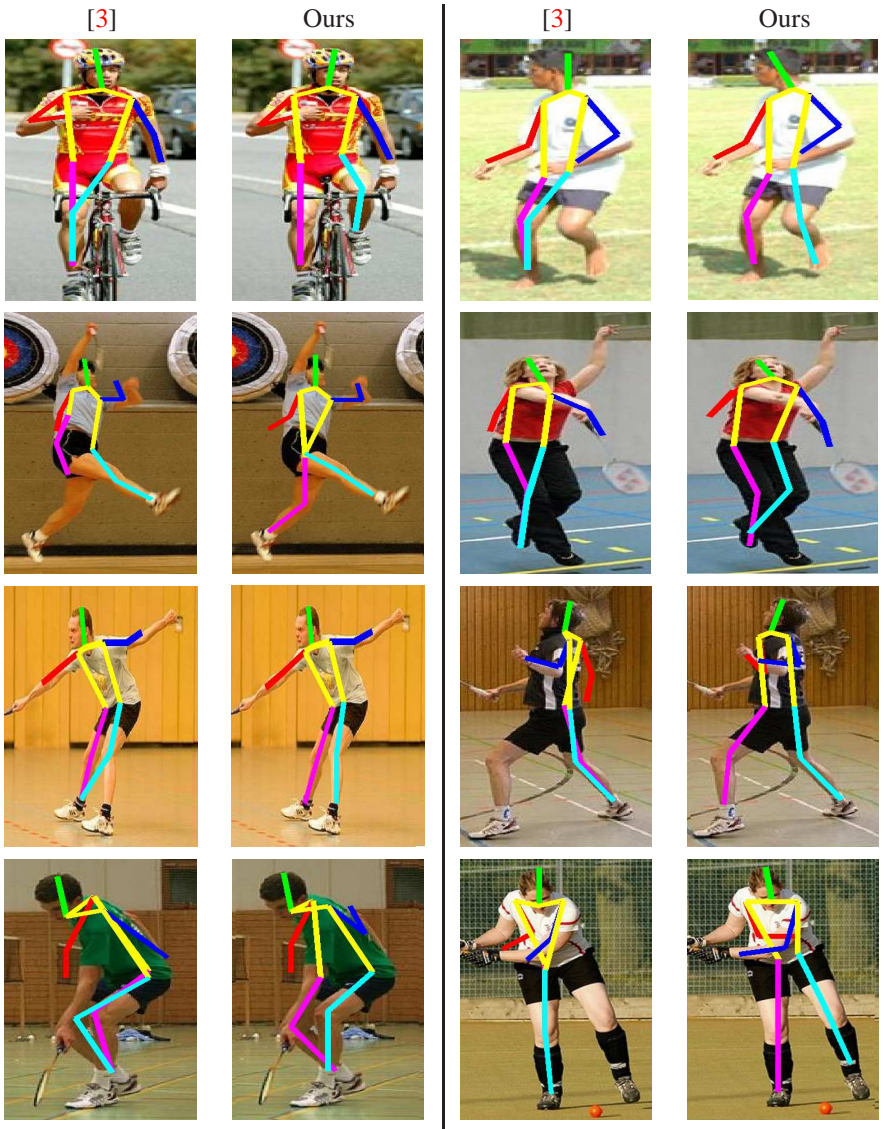
Figure 5: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.
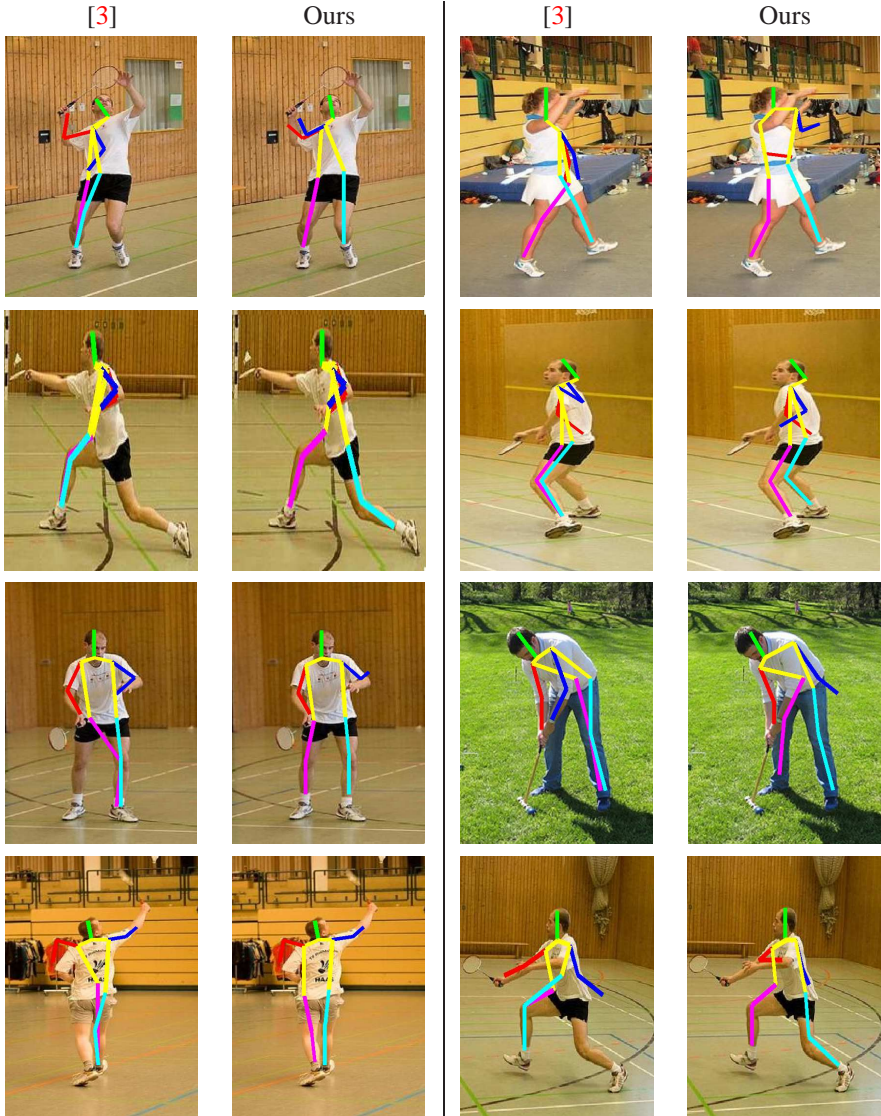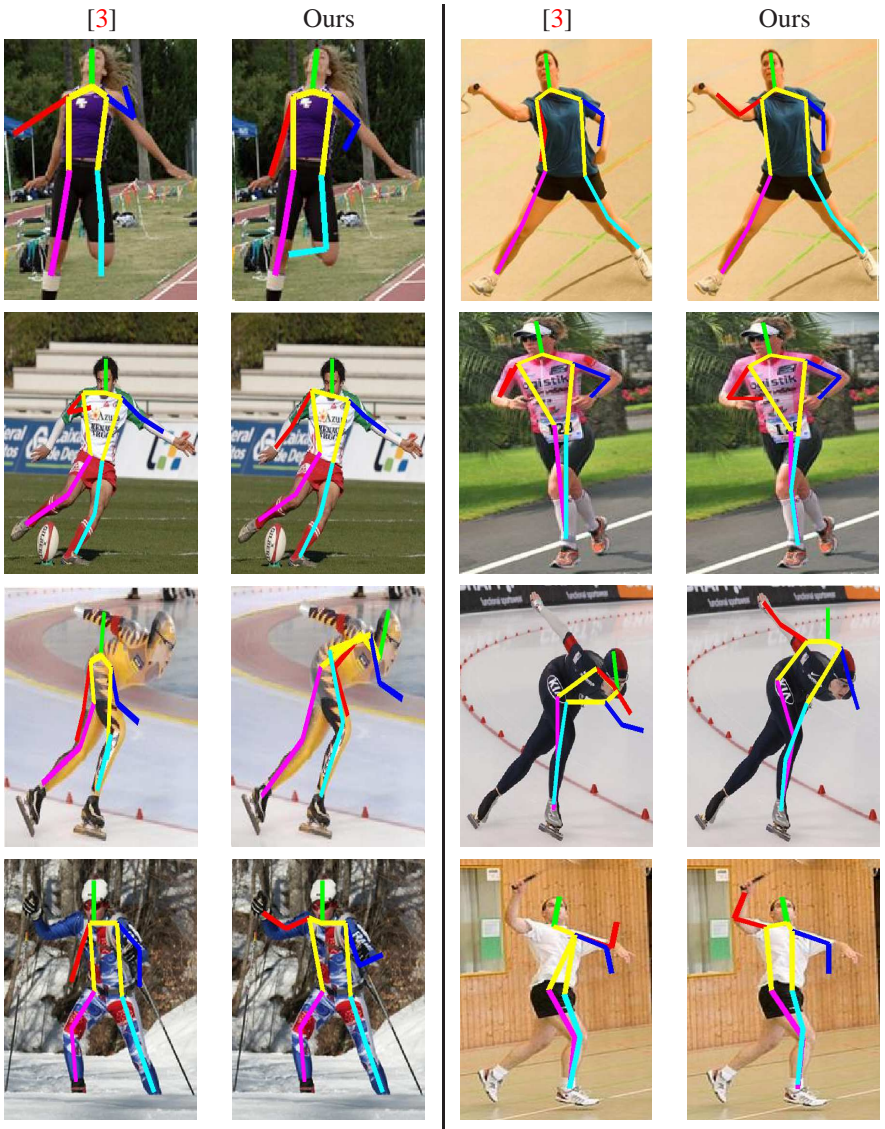
Figure 6: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.

Figure 7: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the upper body.
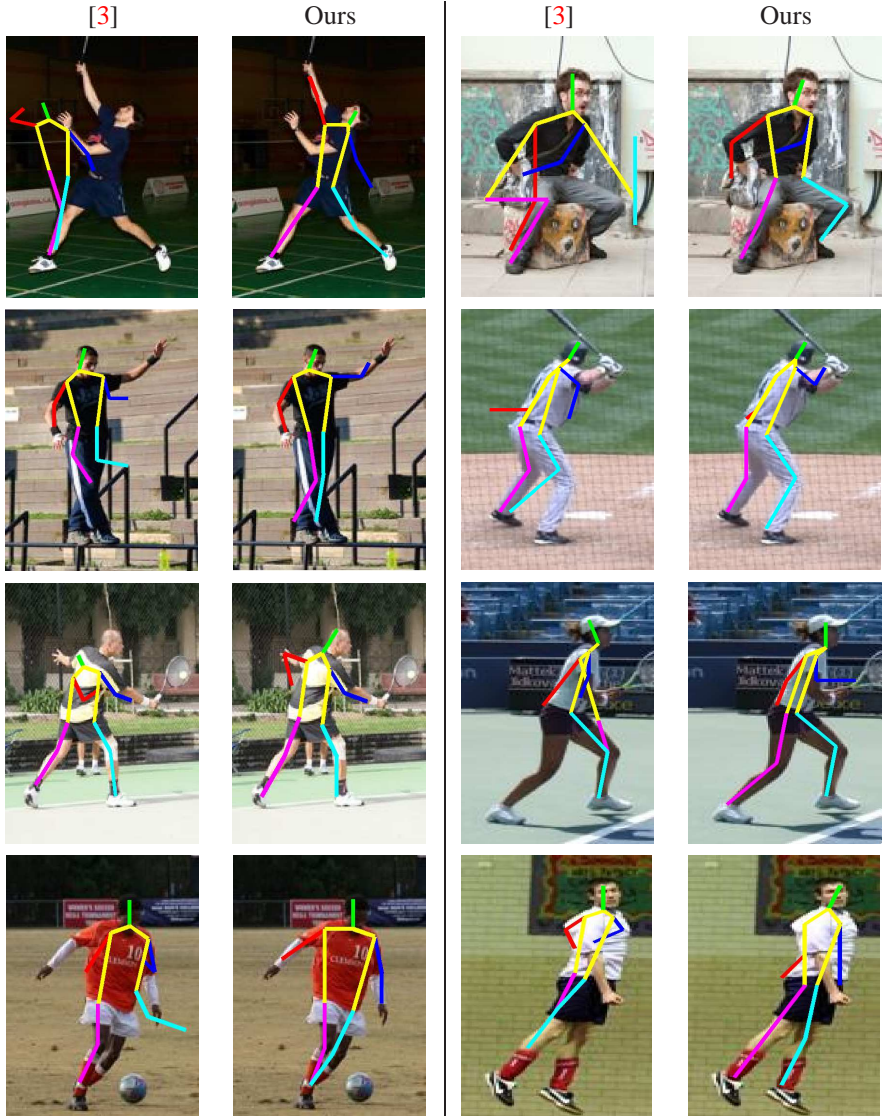
Figure 8: Qualitative results for the LSP dataset. Left image from each pair of images shows the result from [3] and right image shows our results.
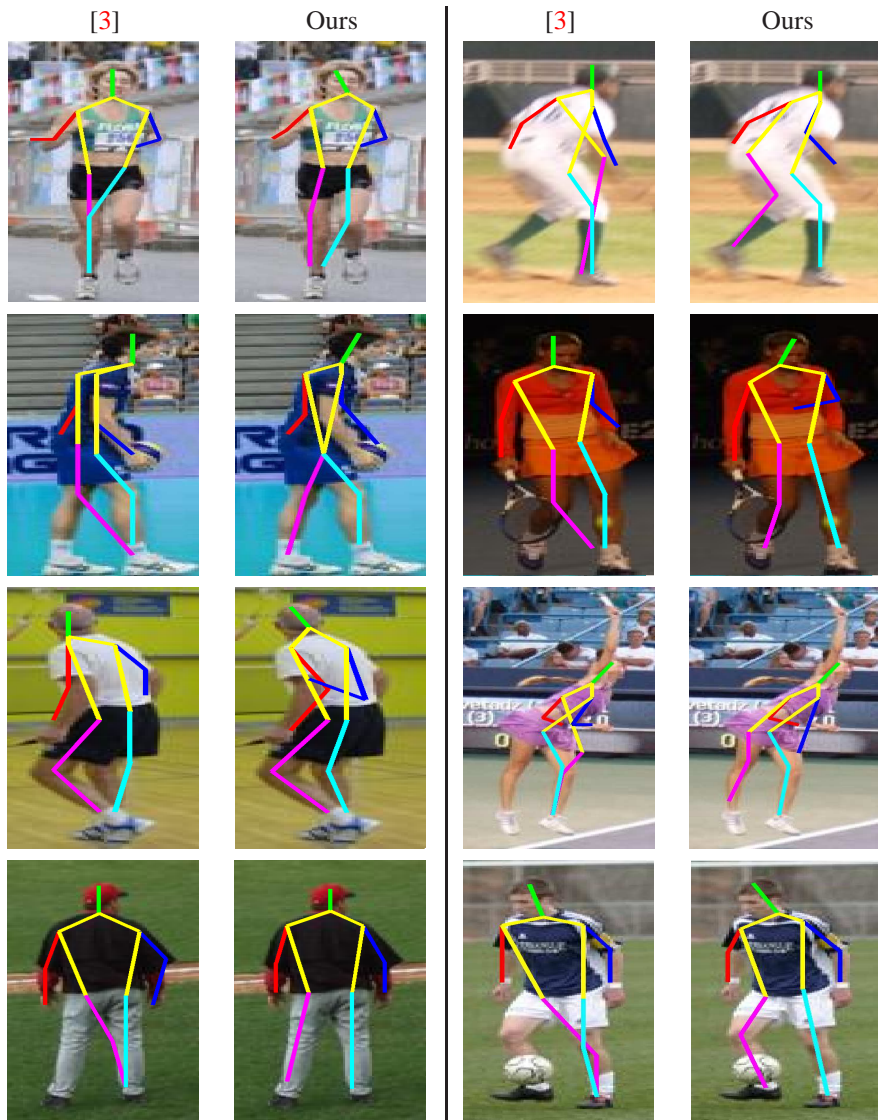
Figure 9: Qualitative results for the LSP dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.

506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
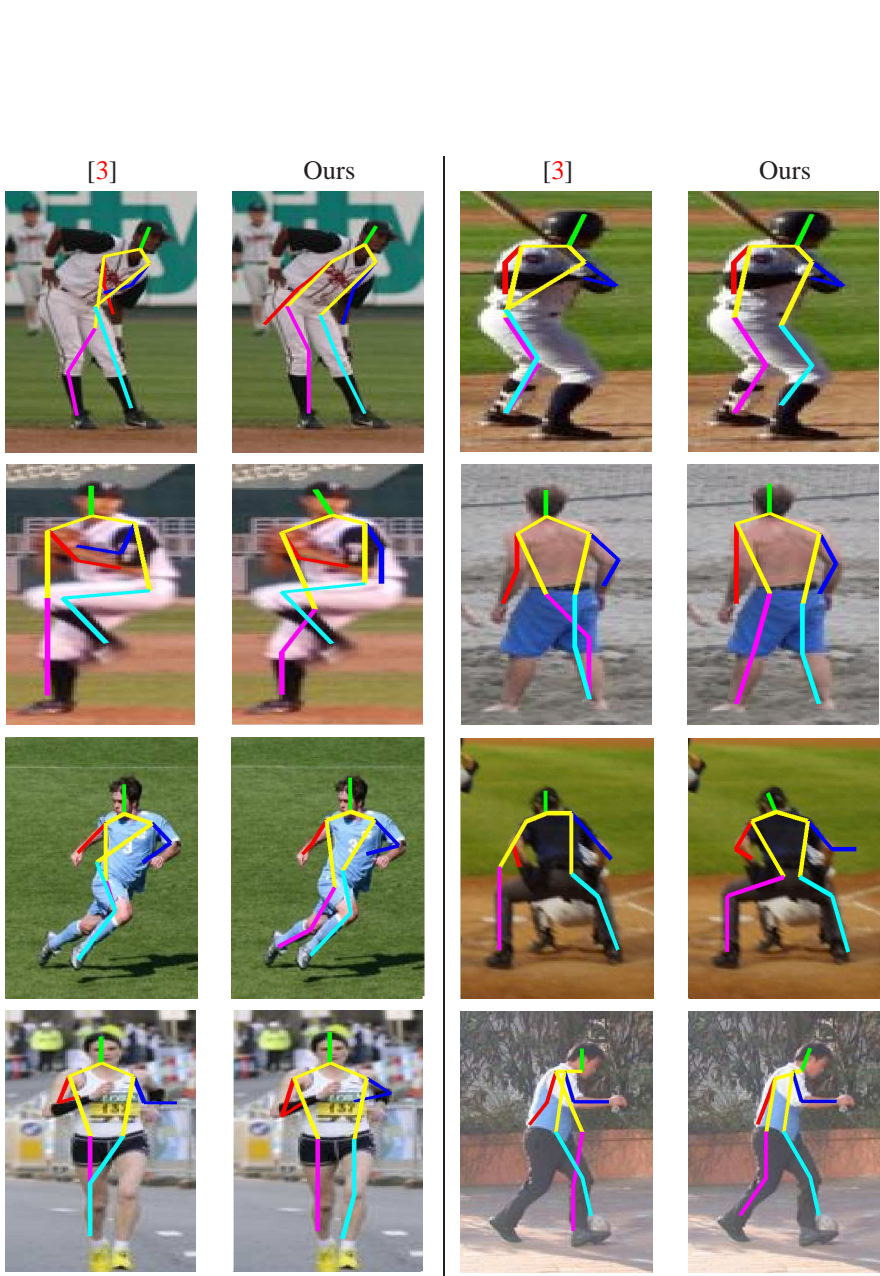541
542
543
544
545
546
547
548
549
550
551

Figure 10: Qualitative results for the LSP dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.
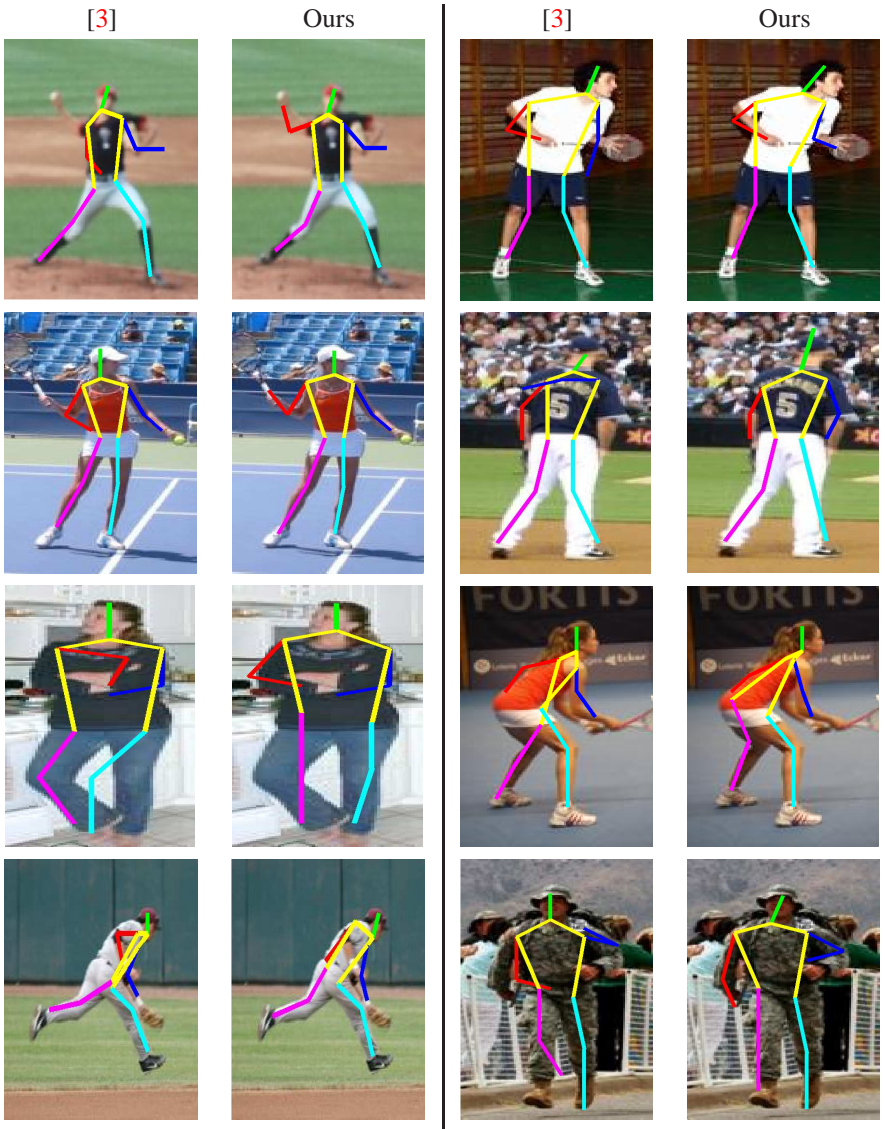
Figure 11: Qualitative results for the LSP dataset (continued). Example images where our method yields a notable improvement in the pose of the upper body.