

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.

092

093

094

095

096

097

098

099

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

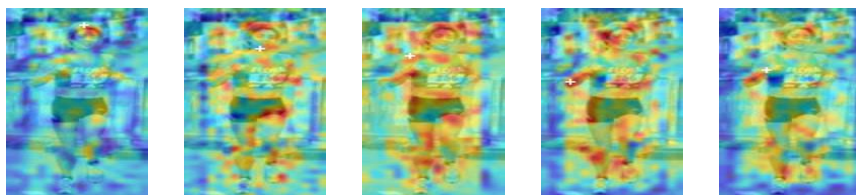
134

135

136

137

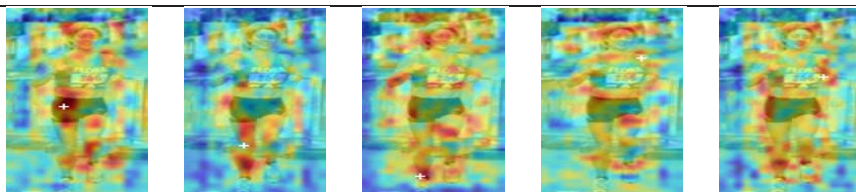
Original score



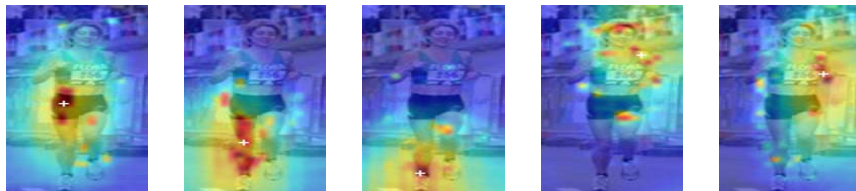
Contextual rescoring



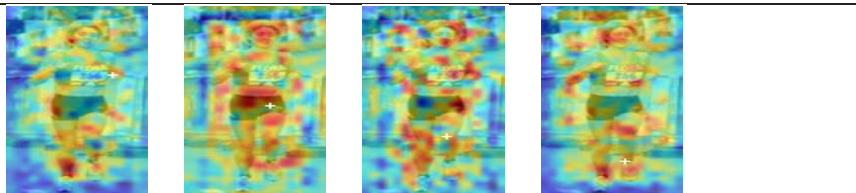
Original score



Contextual rescoring



Original score



Contextual rescoring

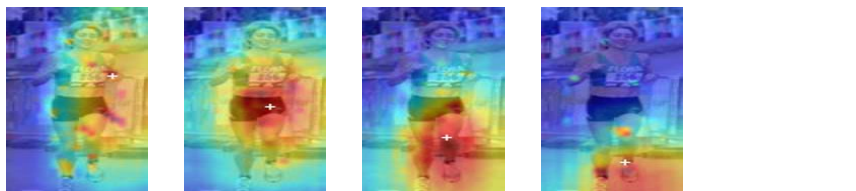


Figure 1: Contextual rescoring in a frontal pose. Odd rows show original detection score maps for different basic body parts. Even rows show the output from our contextual rescoring method. Bluish pixels correspond to low scoring locations, while reddish pixels correspond to high scoring ones. White crosses show ground-truth annotations for each basic part (body joints).

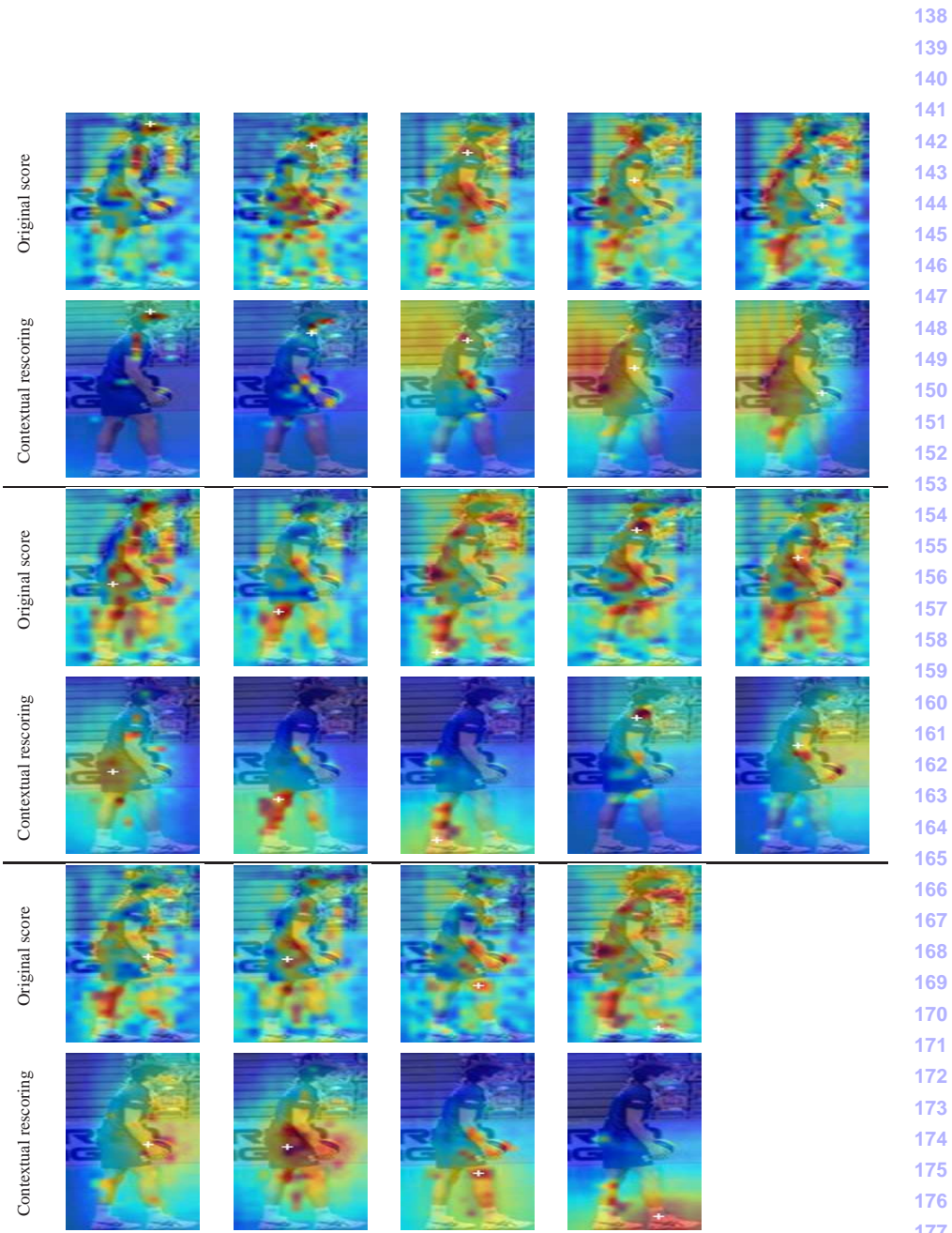


Figure 2: Contextual rescaling in a right-facing profile pose. See Figure 1 for an explanation of the visualization legend.

138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183

184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229

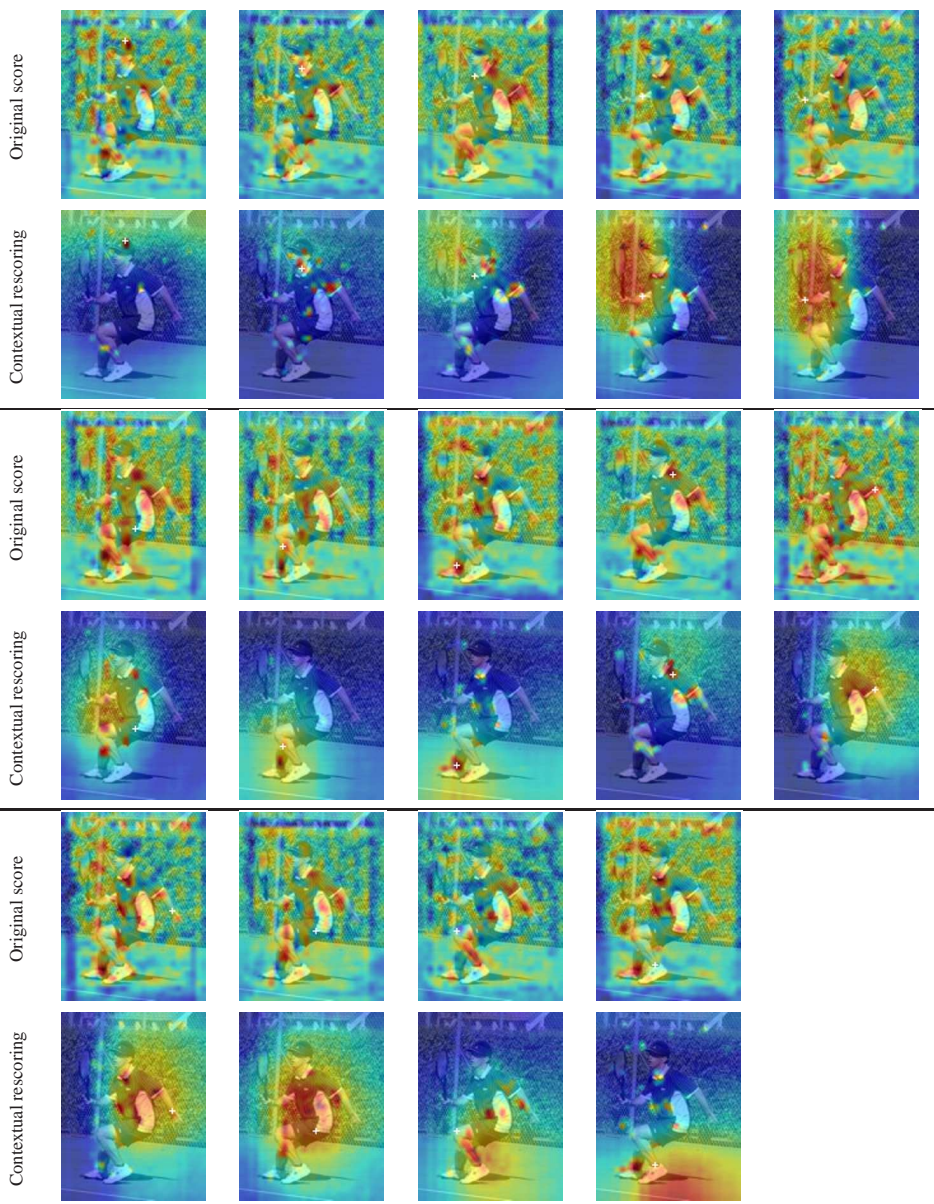


Figure 3: Contextual rescoreing in a left-facing profile pose. See Figure 1 for an explanation of the visualization legend.

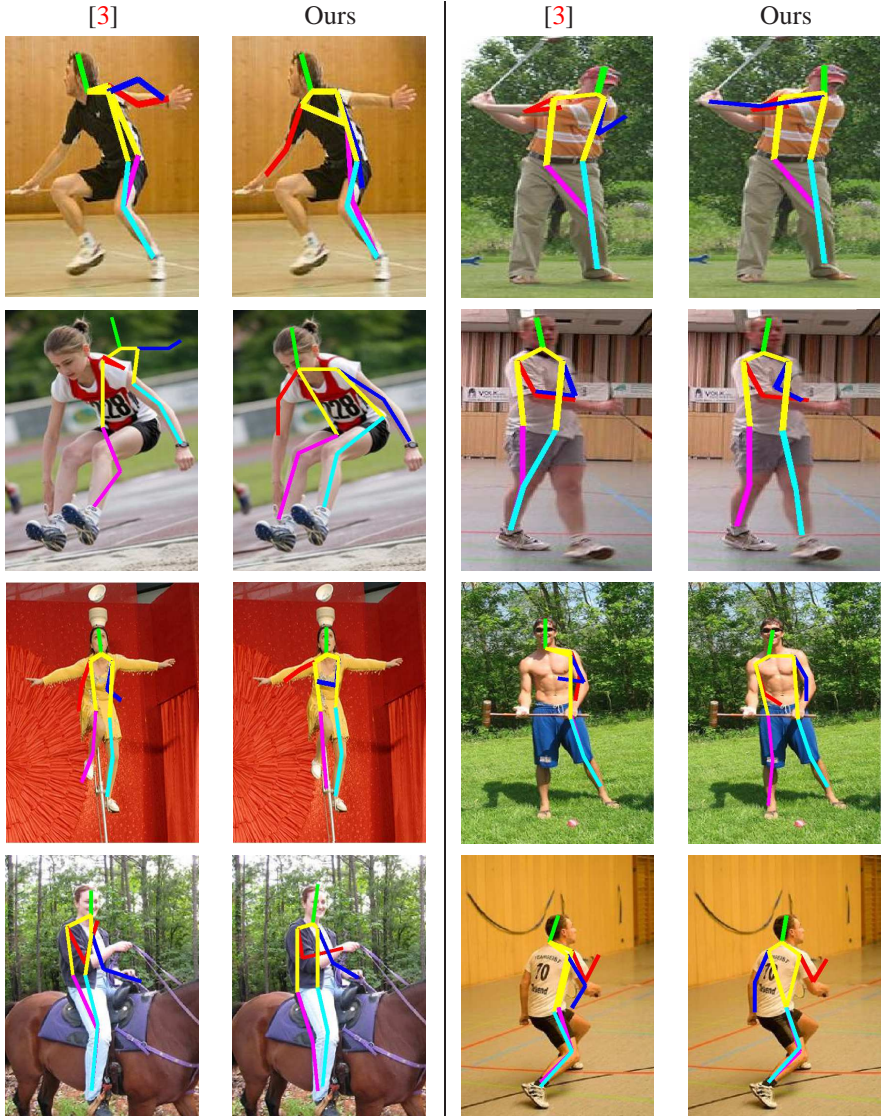


Figure 4: Qualitative results for the UIUC Sports dataset. Left image from each pair of images shows the result from [3] and right image shows our results.

276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321

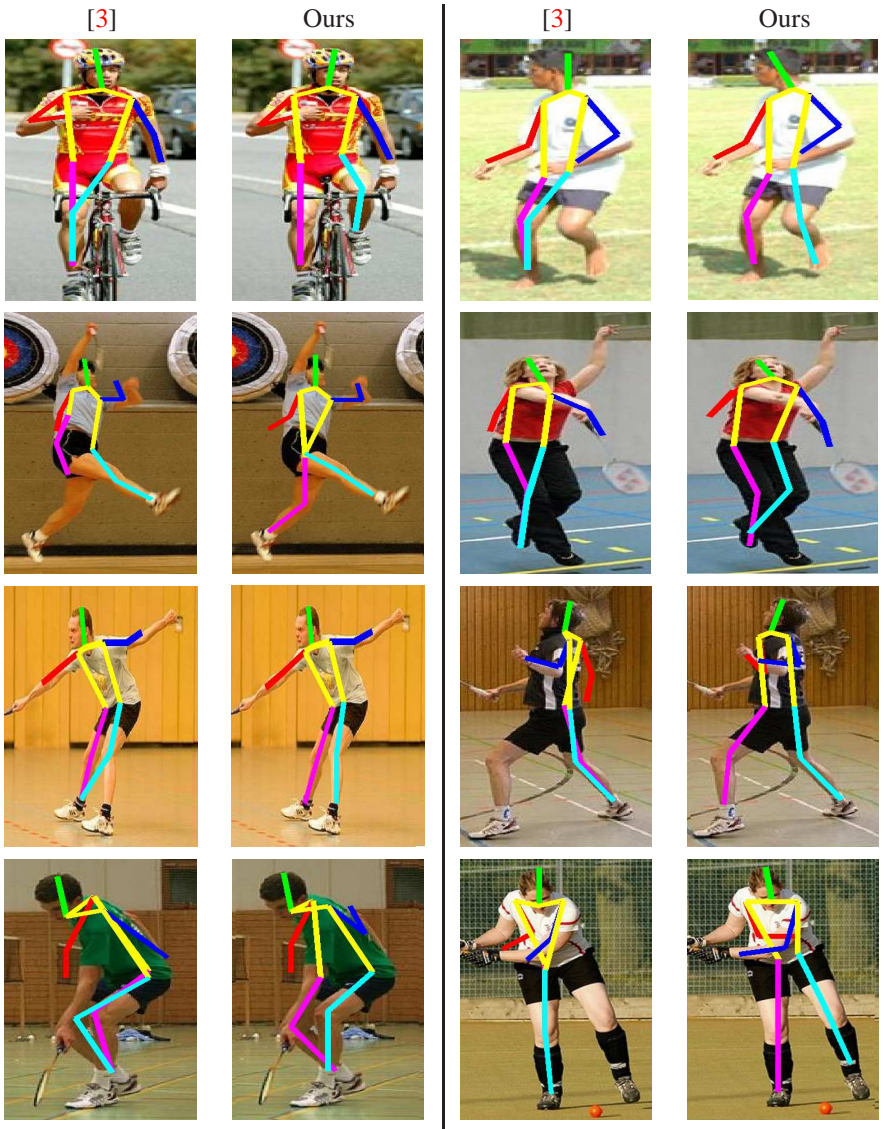


Figure 5: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.

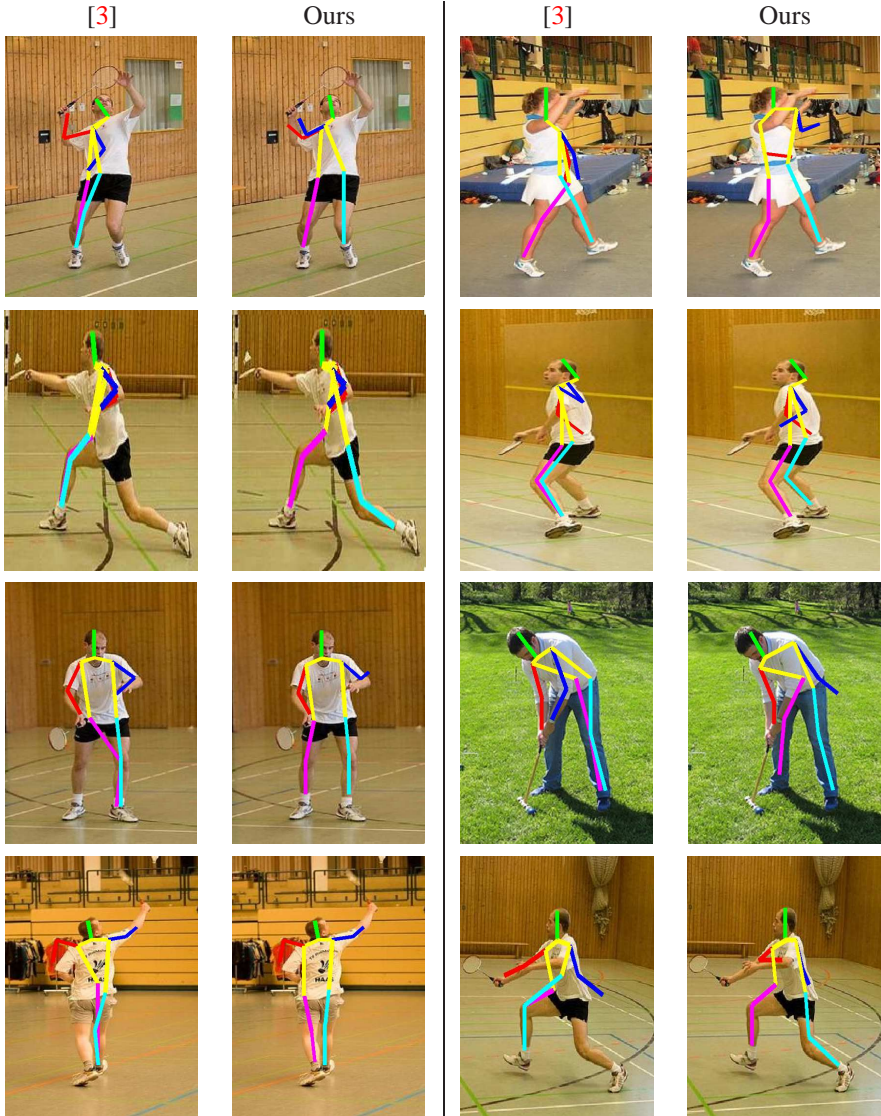


Figure 6: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.

322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367

368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413

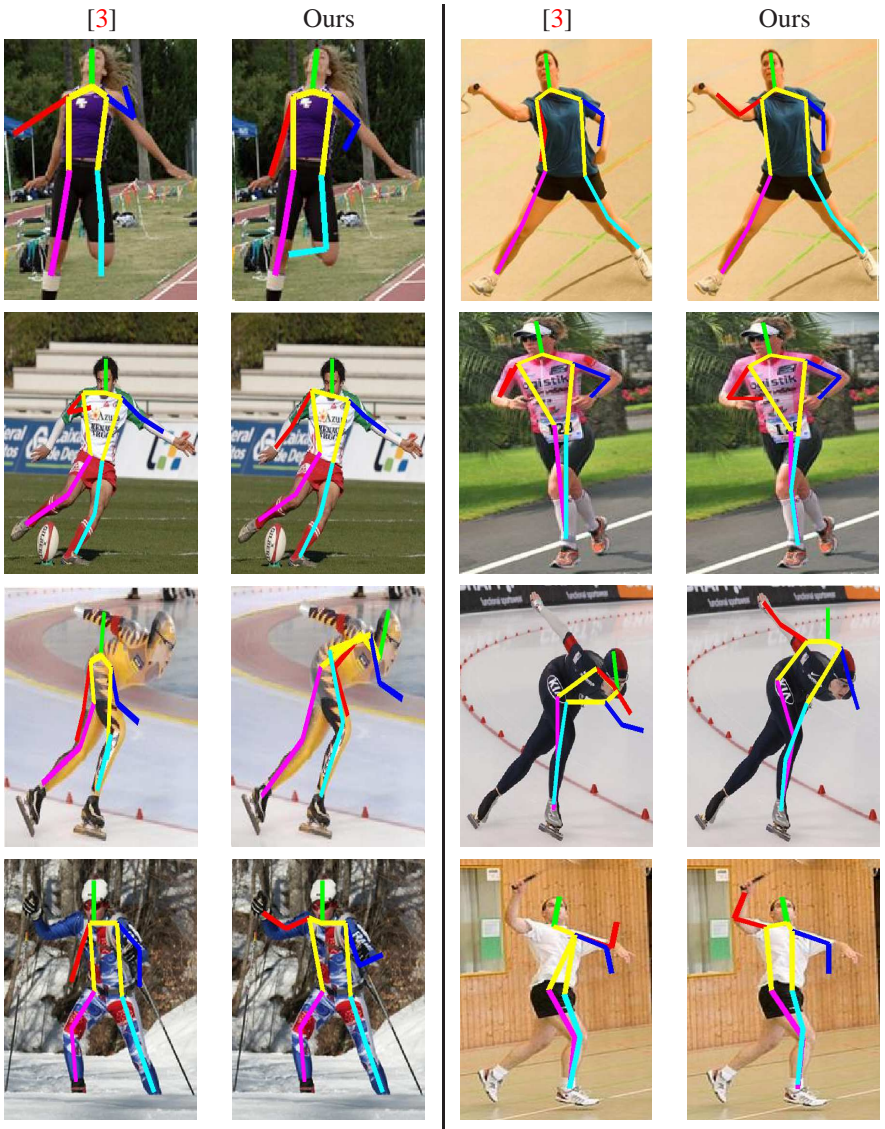


Figure 7: Qualitative results for the UIUC Sports dataset (continued). Example images where our method yields a notable improvement in the pose of the upper body.

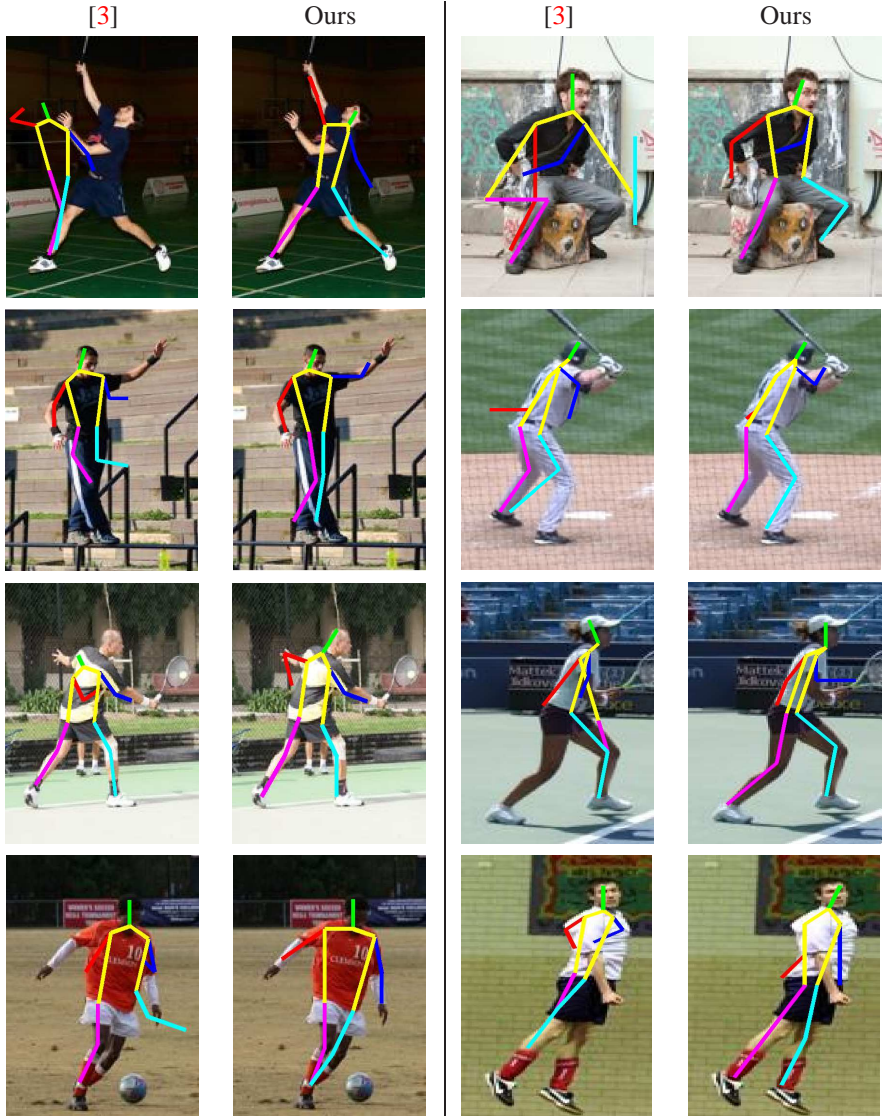


Figure 8: Qualitative results for the LSP dataset. Left image from each pair of images shows the result from [3] and right image shows our results.

414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459

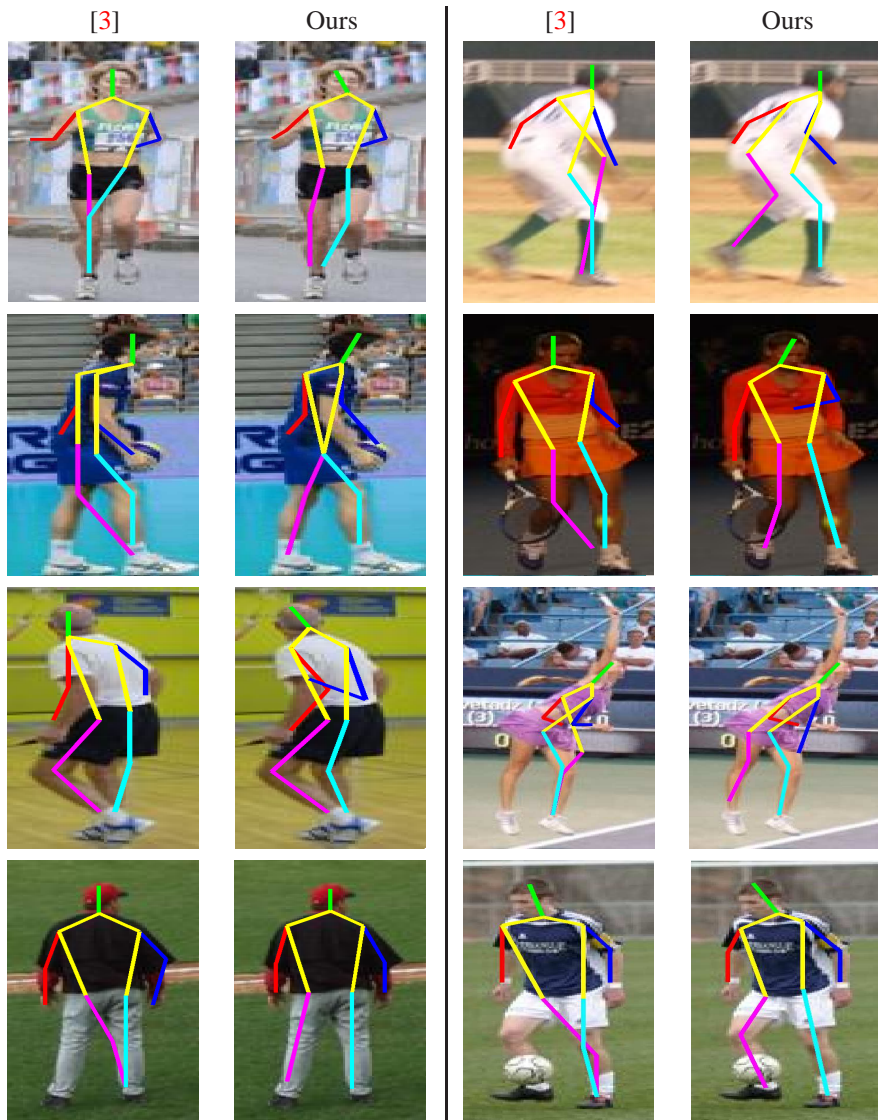
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505

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.

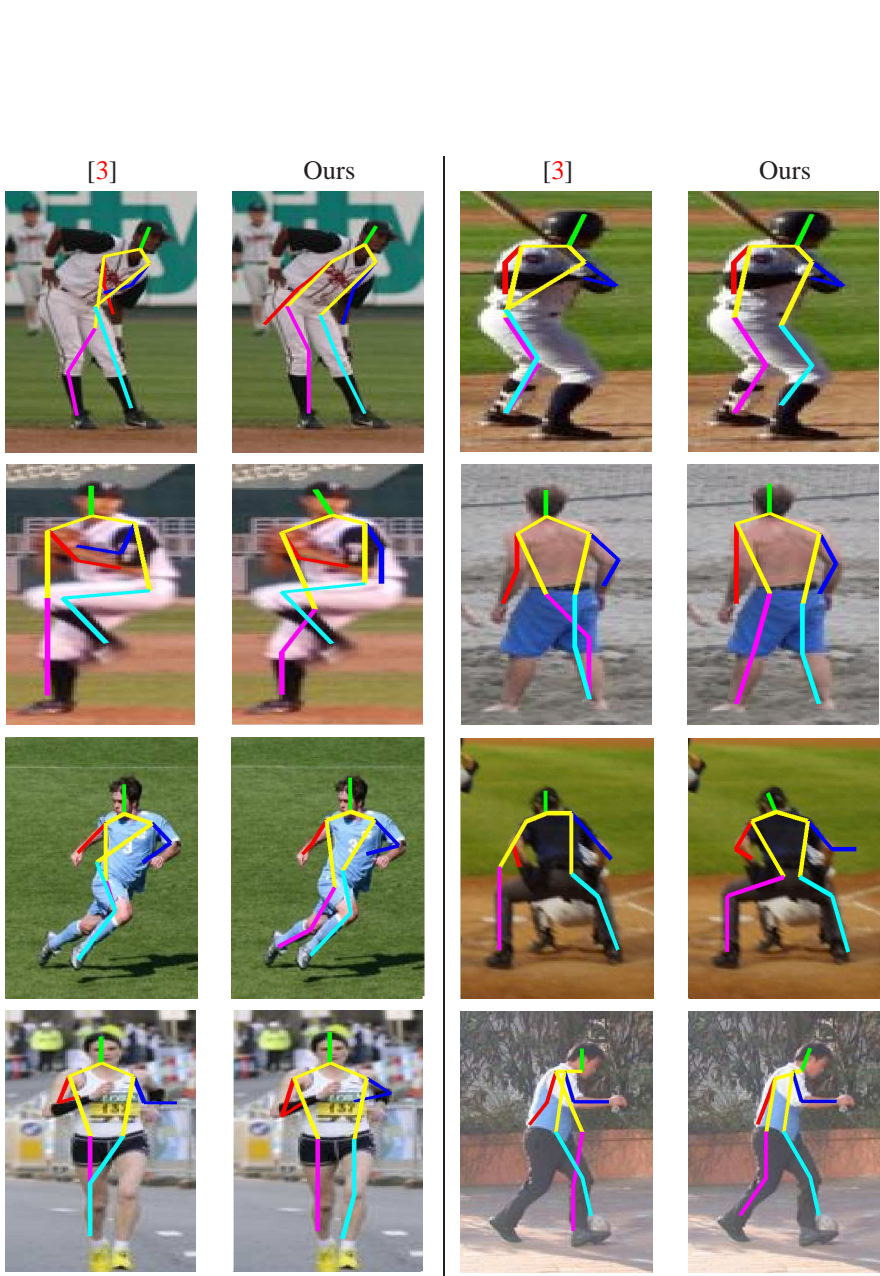


Figure 10: Qualitative results for the LSP dataset (continued). Example images where our method yields a notable improvement in the pose of the lower body.

545
546
547
548
549
550
551

552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597

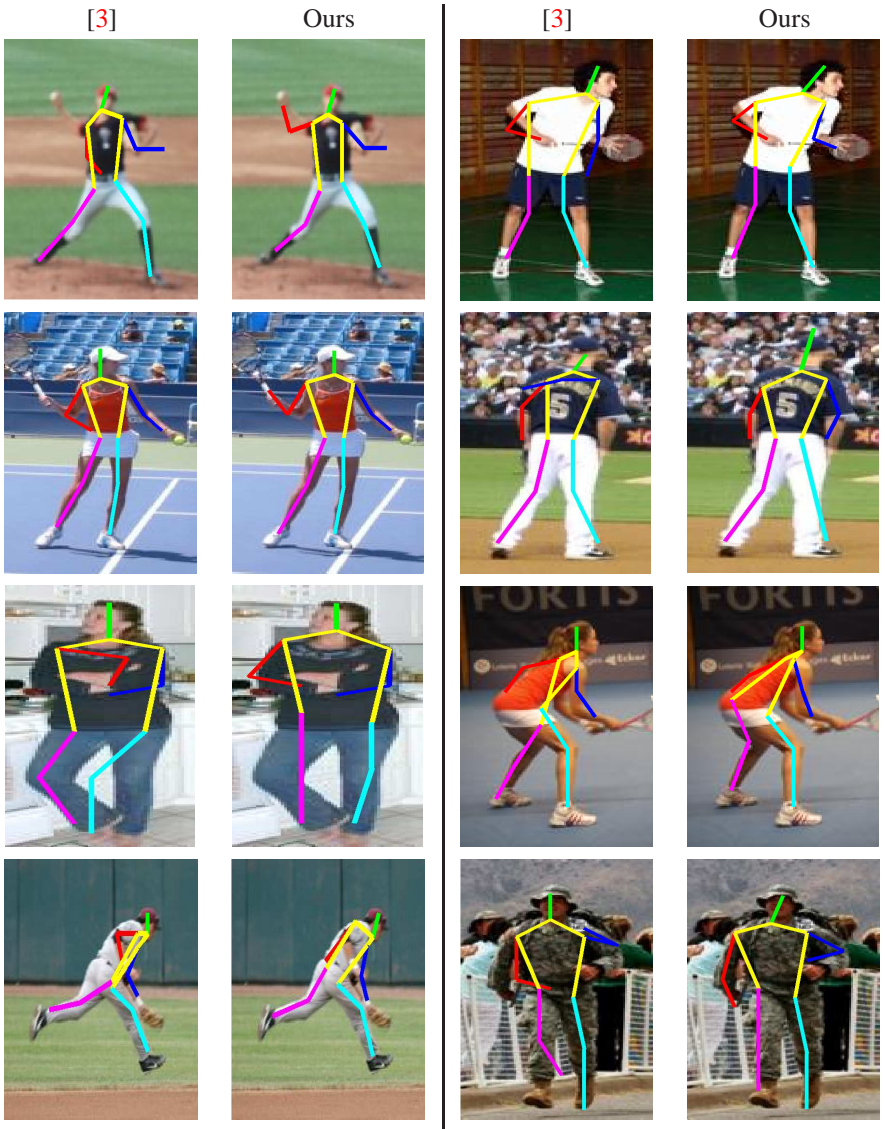


Figure 11: Qualitative results for the LSP dataset (continued). Example images where our method yields a notable improvement in the pose of the upper body.