

000
001 **Supplementary Material**
002
003 **Deep Semantic Matching with Foreground**
004 **Detection and Cycle-Consistency**
005
006
007
008

Anonymous ACCV 2018 submission

Paper ID 632

009 In this supplementary document, we present additional visual examples for
010 comparing our method with the state-of-the-art semantic matching algorithm [1].
011
012

013 **1 Visual Comparisons**
014
015

016 **1.1 Forward-Backward Consistency**
017
018
019
020
021
022
023
024

In Fig. 1 ~ Fig. 4, we demonstrate the effectiveness of the proposed forward-backward consistency loss by visualizing the reprojection errors using the estimated forward and backward transformations. We uniformly sample 100 pixel coordinates in the source image, i.e., a 10×10 grid \mathbf{v} , and mark them with red circles. The geometric transformation T_{AB} is learned to match the source image to the target image, while another transformation T_{BA} is derived to match it back. We then apply T_{AB} and T_{BA} in order to each of the sampled points to obtain the reprojected points \mathbf{u} marked with green circles, namely,

$$\mathbf{u} = T_{BA}(T_{AB}(\mathbf{v})). \quad (1)$$

025 The yellow line links the red and green circles represents the distance (loss)
026
027 between the linked points. It can be observed that introducing the proposed
028 forward-backward consistency loss greatly reduces the reprojection errors, hence
029 improving semantic matching.
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044

Fig. 1: **Visualization of the effect of the forward-backward consistency.**

075
076
077
078
079
080
081
082
083
084
085
086
087
088
089



Fig. 2: Visualization of the effect of the forward-backward consistency.

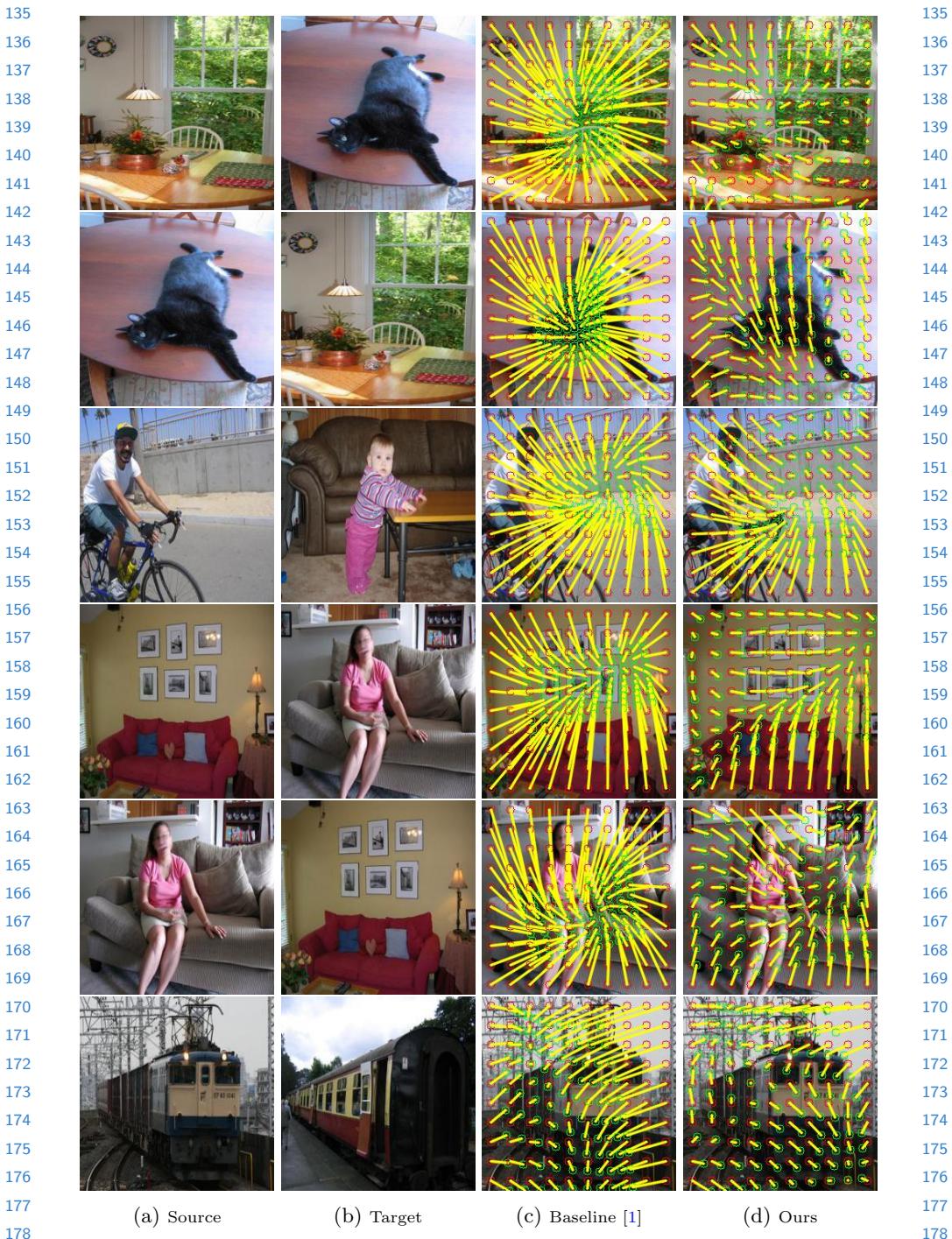


Fig. 3: Visualization of the effect of the forward-backward consistency.

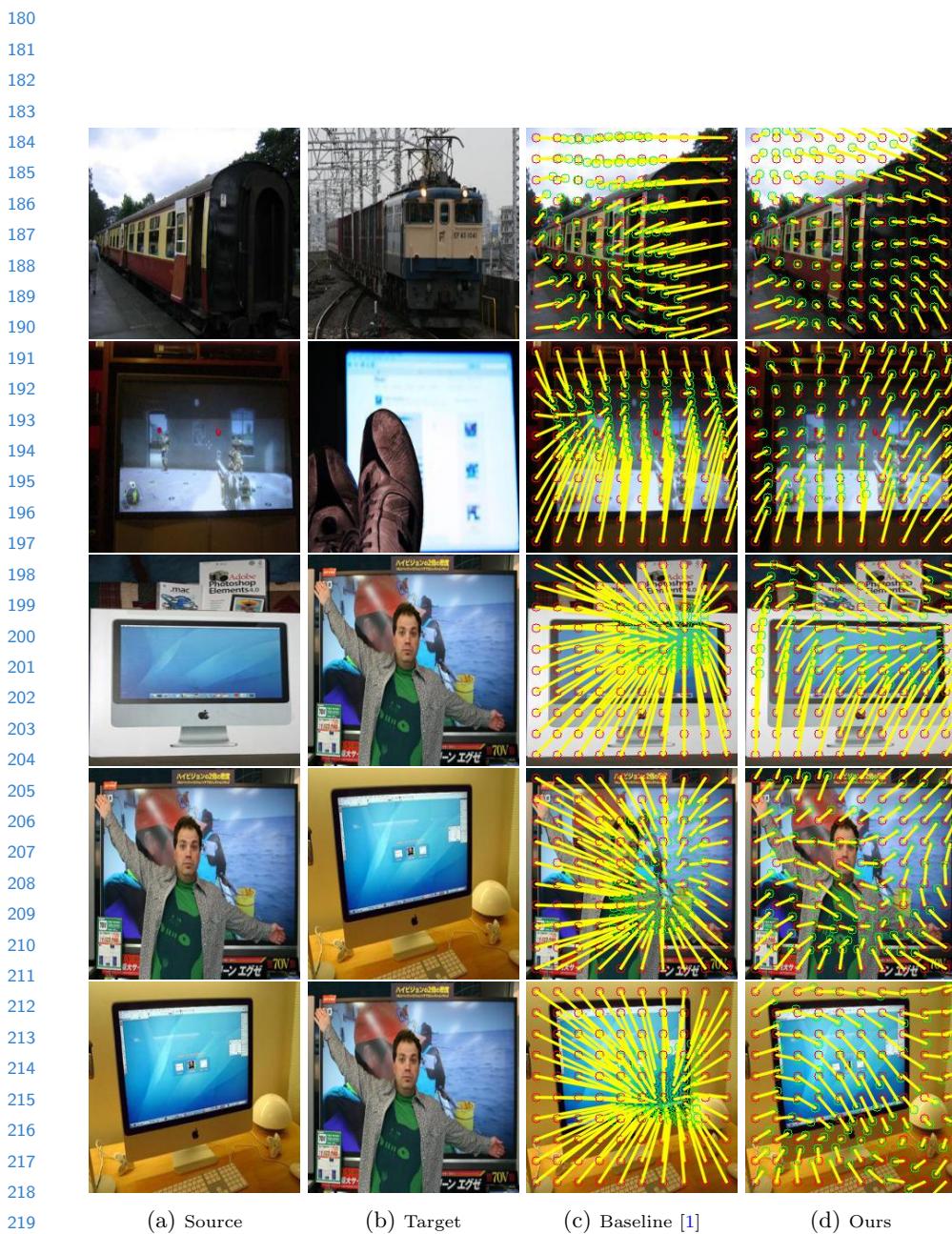


Fig. 4: Visualization of the effect of the forward-backward consistency.

221

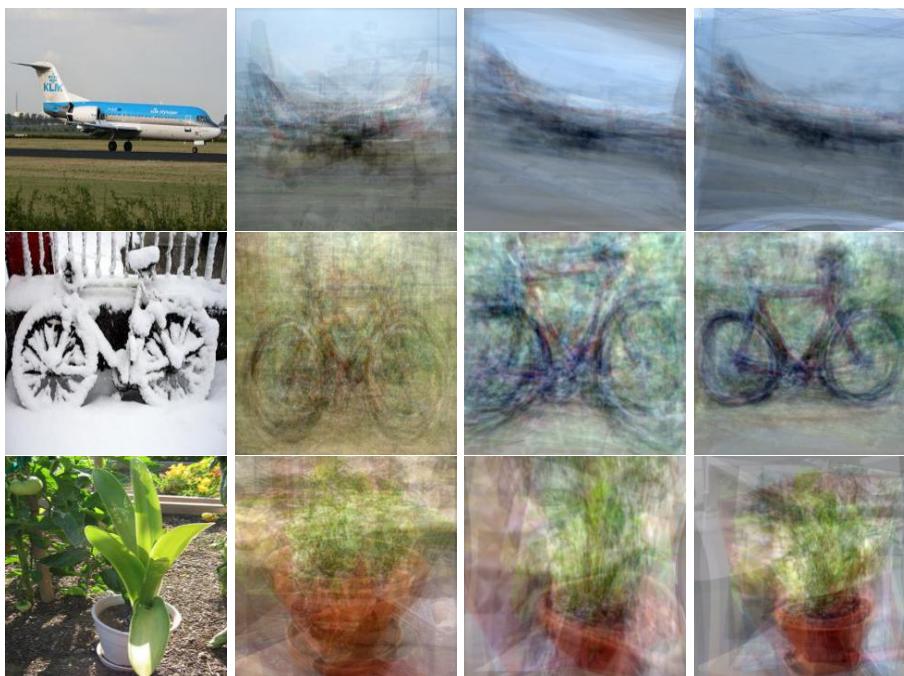
222

223

224

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	1.2 Image Alignment	225
226	We compare the method in [1] and the proposed method for image alignment. For	226
227	each image pair (I_A, I_B), each method estimates a geometric transformation T_{AB}	227
228	which warps image I_A to \tilde{I}_A so that \tilde{I}_A and I_B are well aligned. The alignment	228
229	results are displayed with the GIF files collected by an html file. Please refer to	229
230	“warp.html” for the visual results.	230
231		231
232		232
233	1.3 Average Images within Object Category	233
234	Let $\mathcal{D} = \{I_i\}_{i=1}^N$ denote a set of images covering object instances of a specific	234
235	category, where I_i is the i^{th} image and N is the number of images. For each	235
236	source image $I_A \in \mathcal{D}$, the geometric transformation T_{BA} can be estimated for	236
237	each other image $I_B \in \mathcal{D}$. We warp each I_B based on the geometric transforma-	237
238	tion T_{BA} and then compute the average warped image I_{avg} by averaging all	238
239	the warped images. We present the visual comparisons among three different	239
240	methods: Average, Baseline [1], and Ours. For the “Average” method, we di-	240
241	rectly compute the average image over all images $\{I_B\}$ without applying any	241
242	geometric transformations. As can be seen in Fig. 5 and Fig. 6, the average im-	242
243	ages by method “Average” do not align the source images. Compared with the	243
244	competing method [1], our method can better align the given source images, and	244
245	the resultant average images look sharper.	245
246		246
247		247
248		248
249		249
250		250
251		251
252		252
253		253
254		254
255		255
256		256
257		257
258		258
259		259
260		260
261		261
262		262
263		263
264		264
265		265
266		266
267		267
268		268
269		269

270
271
272
273
274
275
276
277
278
279
280

(a) Source

(b) Average

(c) Baseline [1]

(d) Ours

302
303
304
305
306
307
308
309
310
311
312
313
314270
271
272
273
274
275
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

Fig. 5: Class-wise average warping.

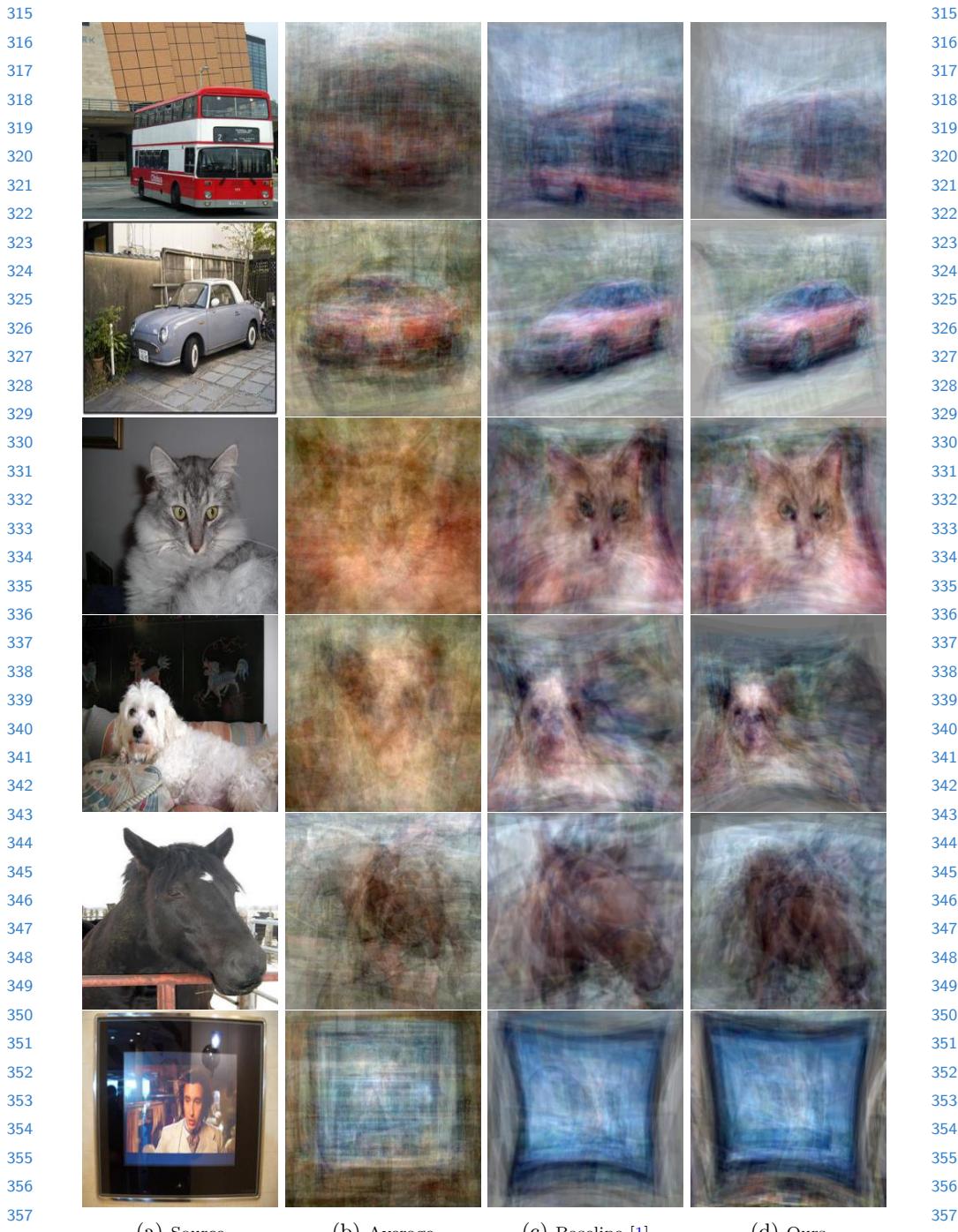
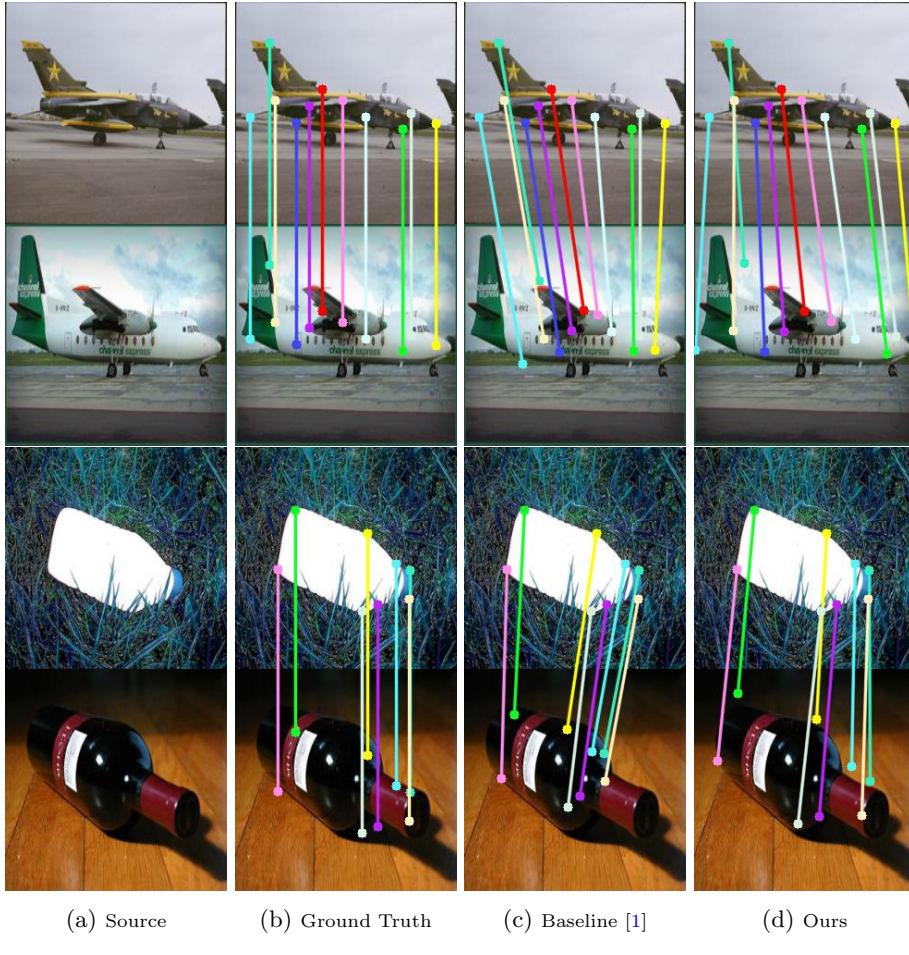


Fig. 6: Class-wise average warping.

360 1.4 More Results of Semantic Correspondence

361 We display more results of semantic matching on the PF-PASCAL dataset in
 362 Fig. 7 and Fig. 8, where our method establishes more accurate correspondences
 363 than the competing method [1].



395 Fig. 7: Semantic correspondence results.
 396

397
 398
 399
 400
 401
 402
 403
 404

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

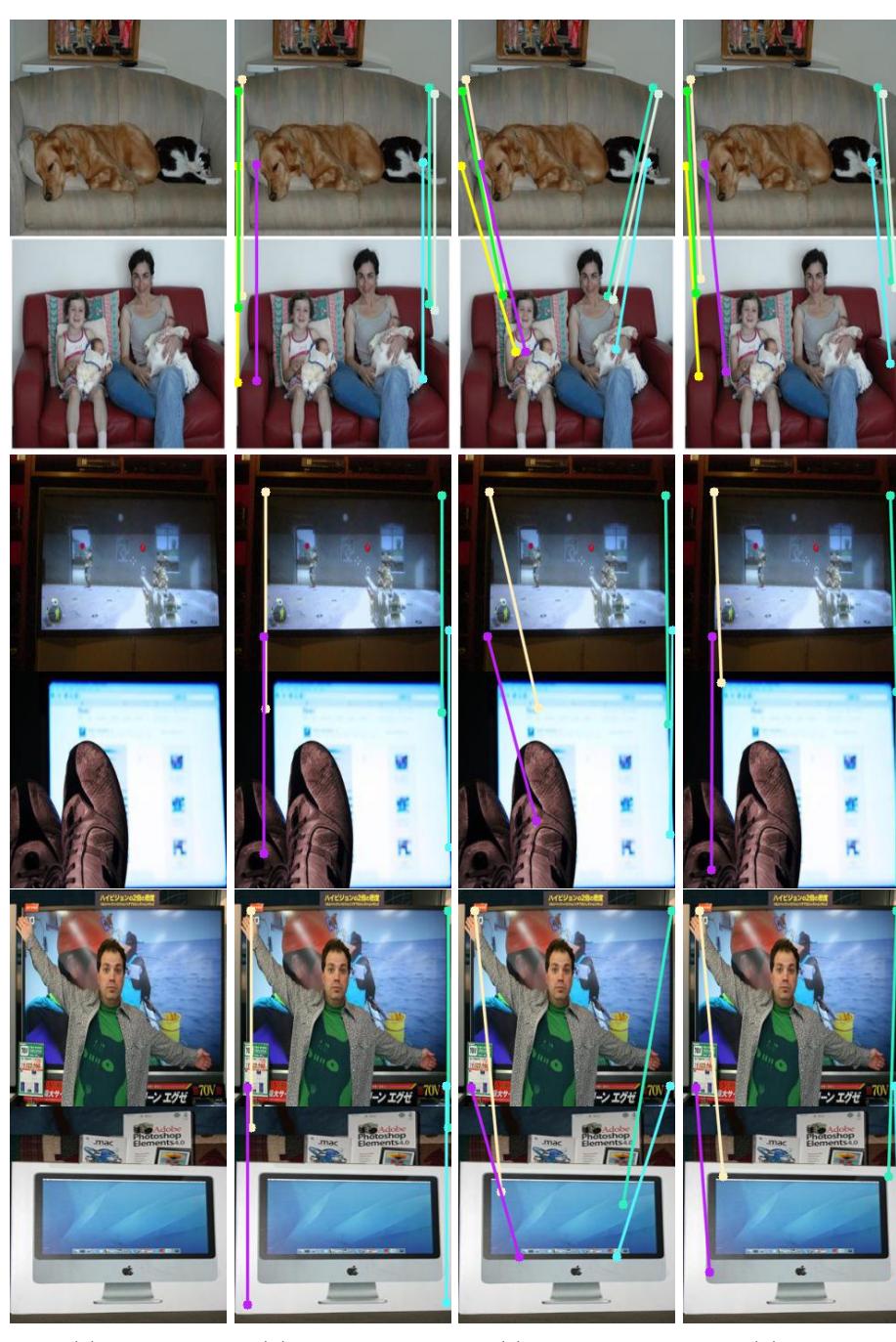


Fig. 8: Semantic correspondence results.

References

- 450 **References** 450
451
452 1. I. Rocco, R. Arandjelović, and J. Sivic. End-to-end weakly-supervised semantic
453 alignment. In *CVPR*, 2018.
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