Supplementary Material for "Self-Supervised Difference Detectionfor Weakly-Supervised Semantic Segmentation"

A.1 Details of the simple decision

In the proposed method, we select *advice* by inference results of difference detection. The confidence score is calculated from the viewpoint of how close the value of d^K to d^A. In the proposed method, if this difference is large enough, we ignore the *advice*. Therefore, if the inferences of the difference detection are too easy, the values of d^K for *advice* that is not true become close to d^A, and the proposed method does not work effectively. In particular, if the inference results of the difference detection are ($d^K = 1, d^A = 1, d^A = d^K$), we cannot distinguish whether the *advice* belongs to the set of true values $|S^{A,T}|$ or the set of false values $|S^{A,F}|$ based on the results of the difference detection. Therefore, we judge the typical failure examples of $advice$ and excluded them from the training sample so that the differences between d^K and d^A were large in the inference of the bad *advice*. To be concrete, when the number of differences in the pixels in each class of mask is obviously large, we assume that the *advice* has failed. We define the bad training samples as the pair of the masks for the difference detection that satisfies the following equation:

$$\forall c \in \mathcal{C}, \frac{|S_c^{m^A}|}{|S_c^{m^K}|} < 0.5, \tag{16}$$

where $\mathcal{C}$ is a set of image-level label of the input image. We decide the threshold 0.5 empirically.

A.2 Details of the bias in Eq.(3)

In Eq.(3), we use $bias$, which is a kind of hyperparameter. In this section, we discuss this $bias$. We define the $bias$ as follows:

$$bias_u = \begin{cases} b_{dd} \pm b_{class} & \text{if } m_u^A \textit{ or } m_u^k \textit{ belongs to } \hat{\mathcal{C}} \\ b_{dd} & \text{if } \textit{otherwise} \end{cases}, \tag{17}$$

where $\forall c \in \hat{\mathcal{C}}$ satisfy $\frac{|S_c^{m^A}|}{|S_c^{m^K}|} < 0.5$ and $c \in \mathcal{C}$. b_{dd} is a bias for the difference between *knowledge* and *advice*, and b_{class} is a bias for the class category. When the number of differences in the pixels in each class of mask is obviously large, it is assumed that the *advice* has failed, and to prioritize the label of that class over the results of the difference detection, we use the bias b_{class}. We defined the values of b_{dd} and b_{class} by using the grid search.

A.3 Values of hyperparameters

We explore good hyperparameters by a grid search and verify the effect of the hyperparameters. We change the values of the hyperparameters and measure the mean IoU scores. Table A-1 shows the hyper parameter values and the mean IoU scores. The hyperparameters (b_{dd}, b_{class}) are used in Eq.(17) as the bias values. In $b_{dd} = 0.4$, the mean IoU score becomes the maximum value. We also set the bias b_{class} for the missing categories. We observe that the setting $b_{class} = 1.0$ achieved a maximum mean IoU. It is expected that the class biases for the missing categories help to the train for robustness. In addition, we also verify the effect of hyperparameters for coefficients of losses in Eq.(11). Though we had expected that the value of α would affect the performance, the hyper parameter was not critical for the change of the mean IoU. The balanced setting, that is, $\alpha = 0.5$ showed the best score.

Table A-1. Experimental results with different parameters.

b_{dd}	0.0	0.1	0.2	0.3	**0.4**	0.5
mIoU	62.2	63.9	64.6	64.2	64.9	62.7
b_{class}	0.0	0.5	**1.0**	1.5	2.0	
mIoU	64.3	63.0	64.9	64.5	63.7	
α	1.0	0.75	**0.5**	0.25	0.0	
mIoU	63.1	64.4	64.9	64.3	63.2	

A.4 Detailed comparison with existing works on the PASCAL VOC 2012 *val* and *test* sets

Table A-2. Results on PASCAL VOC 2012 *val set* without additional supervision.

methods	bg	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv	**mIoU**
MIL-FCN [24]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	24.9
CCNN [23]	68.5	25.5	18.0	25.4	20.2	36.3	46.8	47.1	48.0	15.8	37.9	21.0	44.5	34.5	46.2	40.7	30.4	36.3	22.2	38.8	36.9	35.3
EM-Adapt [22]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	38.2
DCSM [32]	76.7	45.1	24.6	40.8	23.0	34.8	61.0	51.9	52.4	15.5	45.9	32.7	54.9	48.6	57.4	51.8	38.2	55.4	32.2	42.6	39.6	44.1
BFBP [29]	79.2	60.1	20.4	50.7	41.2	46.3	62.6	49.2	62.3	13.3	49.7	38.1	58.4	49.0	57.0	48.2	27.8	55.1	29.6	54.6	26.6	46.6
SEC [16]	82.4	62.9	26.4	61.6	27.6	38.1	66.6	62.7	75.2	22.1	53.5	28.3	65.8	57.8	62.3	52.5	32.5	62.6	32.1	45.4	45.3	50.7
CBTS [28]	85.8	65.2	29.4	63.8	31.2	37.2	69.6	64.3	76.2	21.4	56.3	29.8	68.2	60.6	66.2	55.8	30.8	66.1	34.9	48.8	47.1	52.8
TPL [15]	82.8	62.2	23.1	65.8	21.1	43.1	71.1	66.2	76.1	21.3	59.6	35.1	70.2	58.8	62.3	66.1	35.8	69.9	33.4	45.9	45.6	53.1
MEFF [7]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
PSA [1]	88.2	68.2	30.6	81.1	49.6	61.0	77.8	66.1	75.1	29.0	66.0	40.2	80.4	62.0	70.4	73.7	42.5	70.7	42.6	68.1	51.6	61.7
SSDD (ours)	89.0	62.5	28.9	83.7	52.9	59.5	77.6	73.7	87.0	34.0	83.7	47.6	84.1	77.0	73.9	69.6	29.8	84.0	43.2	68.0	53.4	64.9

Table A-3. Results on PASCAL VOC 2012 *test set* without additional supervision.

methods	bg	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv	**mIoU**
MIL-FCN [24]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	25.7
CCNN [23]	68.5	25.5	18.0	25.4	20.2	36.3	46.8	47.1	48.0	15.8	37.9	21.0	44.5	34.5	46.2	40.7	30.4	36.3	22.2	38.8	36.9	35.3
EM-Adapt [22]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	39.6
DCSM [32]	78.1	43.8	26.3	49.8	19.5	40.3	61.6	53.9	52.7	13.7	47.3	34.8	50.3	48.9	69.0	49.7	38.4	57.1	34.0	38.0	40.0	45.1
BFBP [29]	80.3	57.5	24.1	66.9	31.7	43.0	67.5	48.6	56.7	12.6	50.9	42.6	59.4	52.9	65.0	44.8	41.3	51.1	33.7	44.4	33.2	48.0
SEC [16]	83.5	56.4	28.5	64.1	23.6	46.5	70.6	58.5	71.3	23.2	54.0	28.0	68.1	62.1	70.0	55.0	38.4	58.0	39.9	38.4	48.3	51.7
CBTS [28]	85.7	58.8	30.5	67.6	24.7	44.7	74.8	61.8	73.7	22.9	57.4	27.5	71.3	64.8	72.4	57.3	37.0	60.4	42.8	42.2	50.6	53.7
TPL [15]	83.4	62.2	26.4	71.8	18.2	49.5	66.5	63.8	73.4	19.0	56.6	35.7	69.3	61.3	71.7	69.2	39.1	66.3	44.8	35.9	45.5	53.8
MEFF [7]	86.6	72.0	30.6	68.0	44.8	46.2	73.4	56.6	73.0	18.9	63.3	32.0	70.1	72.2	68.2	56.1	34.5	67.5	29.6	60.2	43.6	55.6
PSA [1]	89.1	70.6	31.6	77.2	42.2	68.9	79.1	66.5	74.9	29.6	68.7	56.1	82.1	64.8	78.6	73.5	50.8	70.7	47.7	63.9	51.1	63.7
SSDD (ours)	89.5	71.8	31.4	79.3	47.3	64.2	79.9	74.6	84.9	30.8	73.5	58.2	82.7	73.4	76.4	69.9	37.4	80.5	54.5	65.7	50.3	65.5

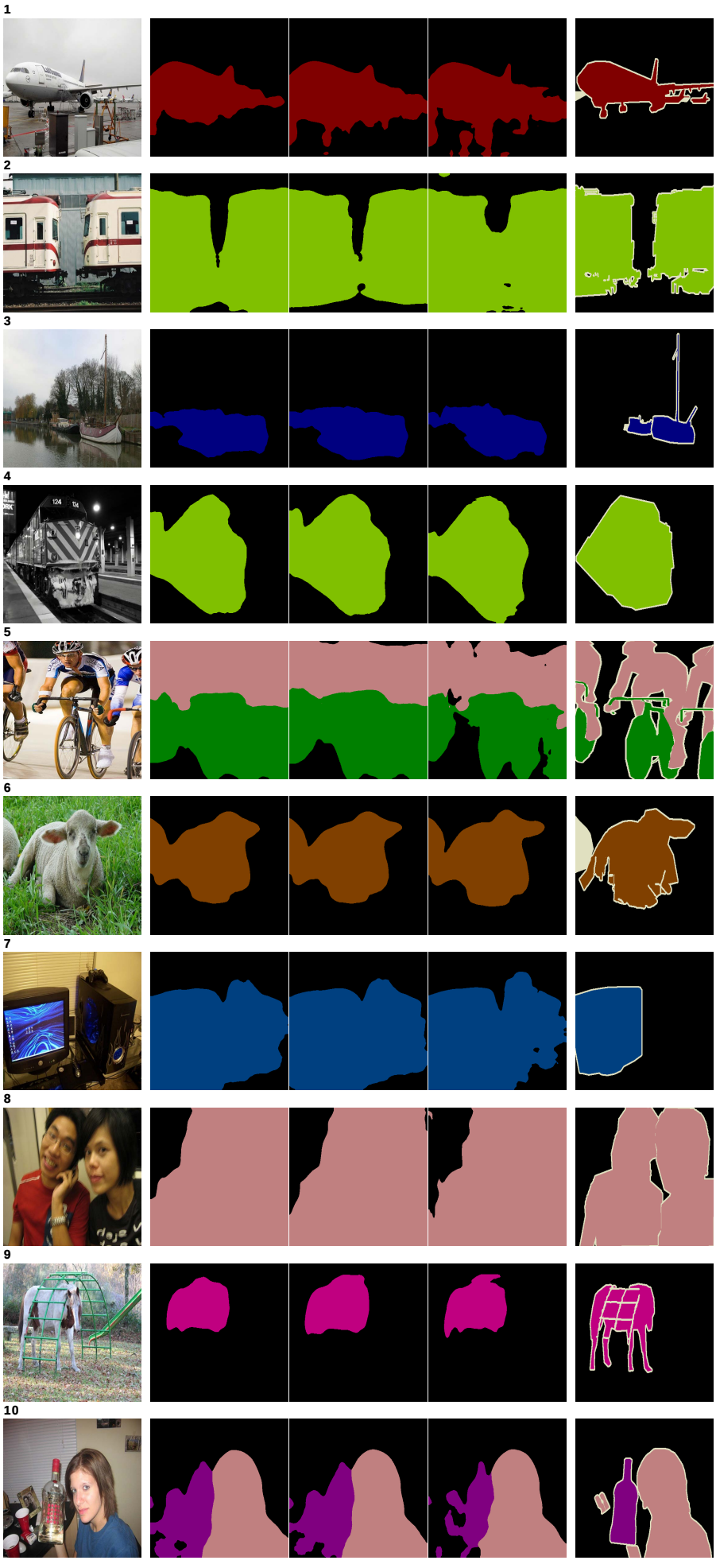
Table A-4. Results on PASCAL VOC 2012 *val set* with additional supervision.

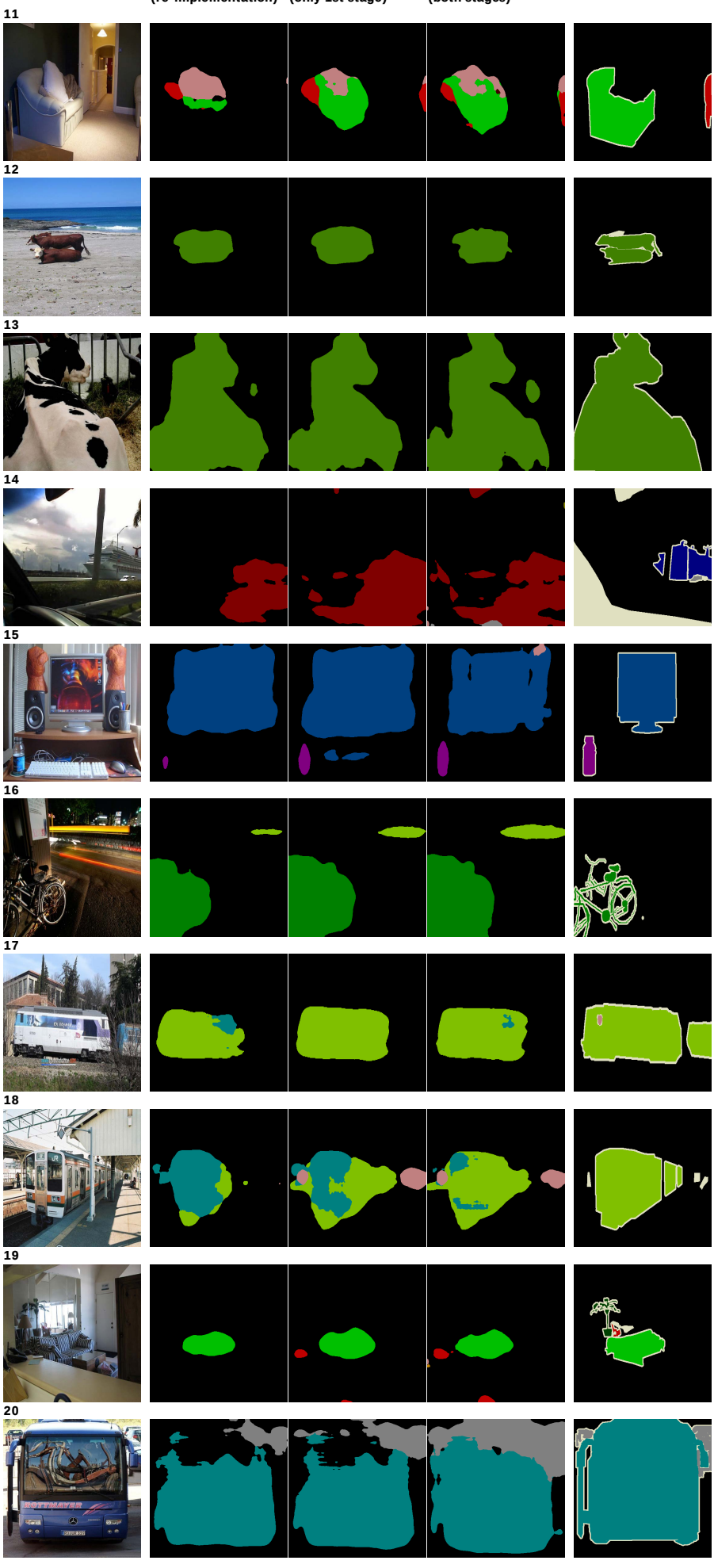
methods	info type†	bg	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv	**mIoU**
MIL-seg [25]	S	79.6	50.2	21.6	40.6	34.9	40.5	45.9	51.5	60.6	12.6	51.2	11.6	56.8	52.9	44.8	42.7	31.2	55.4	21.5	38.8	36.9	42.0
MCNN [34]	WV	77.5	47.9	17.2	39.4	28.0	25.6	52.7	47.0	57.8	10.4	38.0	24.3	49.9	40.8	48.2	42.0	21.6	35.2	19.6	52.5	24.7	38.1
AFF [27]	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	54.3
STC [37]	S	84.5	68.0	19.5	60.5	42.5	44.8	68.4	64.0	64.8	14.5	52.0	22.8	58.0	55.3	57.8	60.5	40.6	56.7	23.0	57.1	31.2	49.8
Oh et al. [30]	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	55.7
AE-PSL [36]	S	83.4	71.1	30.5	72.9	41.6	55.9	63.1	60.2	74.0	18.0	66.5	32.4	71.7	56.3	64.8	52.4	37.4	69.1	31.4	58.9	43.9	55.0
Hong et al. [9]	WV	87.0	69.3	32.2	70.2	31.2	58.4	73.6	68.5	76.5	26.8	63.8	29.1	73.5	69.5	66.5	70.4	46.8	72.1	27.3	57.4	50.2	58.1
WebS-i2 [14]	WI	84.3	65.3	27.4	65.4	53.9	46.3	70.1	69.8	79.4	13.8	61.1	17.4	73.8	58.1	57.8	56.2	35.7	66.5	22.0	50.1	46.2	53.4
DCSP [3]	S	88.9	77.7	31.3	73.2	59.8	71.0	79.2	74.5	80.0	15.1	73.3	10.2	76.1	72.2	69.1	72.1	39.9	73.9	14.6	70.3	53.1	60.8
GAIN [18]	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56.8
MDC [38]	S	89.5	85.6	34.6	75.8	61.9	65.8	67.1	73.3	80.2	15.1	69.9	8.1	75.0	68.4	70.9	71.5	32.6	74.9	24.8	73.2	50.8	60.4
MCOF [35]	S	87.0	78.4	29.4	68.0	44.0	67.3	80.3	74.1	82.2	21.1	70.7	28.2	73.2	71.5	67.2	53.0	47.7	74.5	32.4	71.0	45.8	60.3
DSRG [11]	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	61.4
Shen et al. [31]	WI	86.8	71.2	32.4	77.0	24.4	69.8	85.3	71.9	86.5	27.6	78.9	40.7	78.5	79.1	72.7	73.1	49.6	74.8	36.1	48.1	59.2	63.0
SeeNet [10]	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	63.1
AISI [6]	IS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64.5
SSDD (ours)	-	89.0	62.5	28.9	83.7	52.9	59.5	77.6	73.7	87.0	34.0	83.7	47.6	84.1	77.0	73.9	69.6	29.8	84.0	43.2	68.0	53.4	64.9

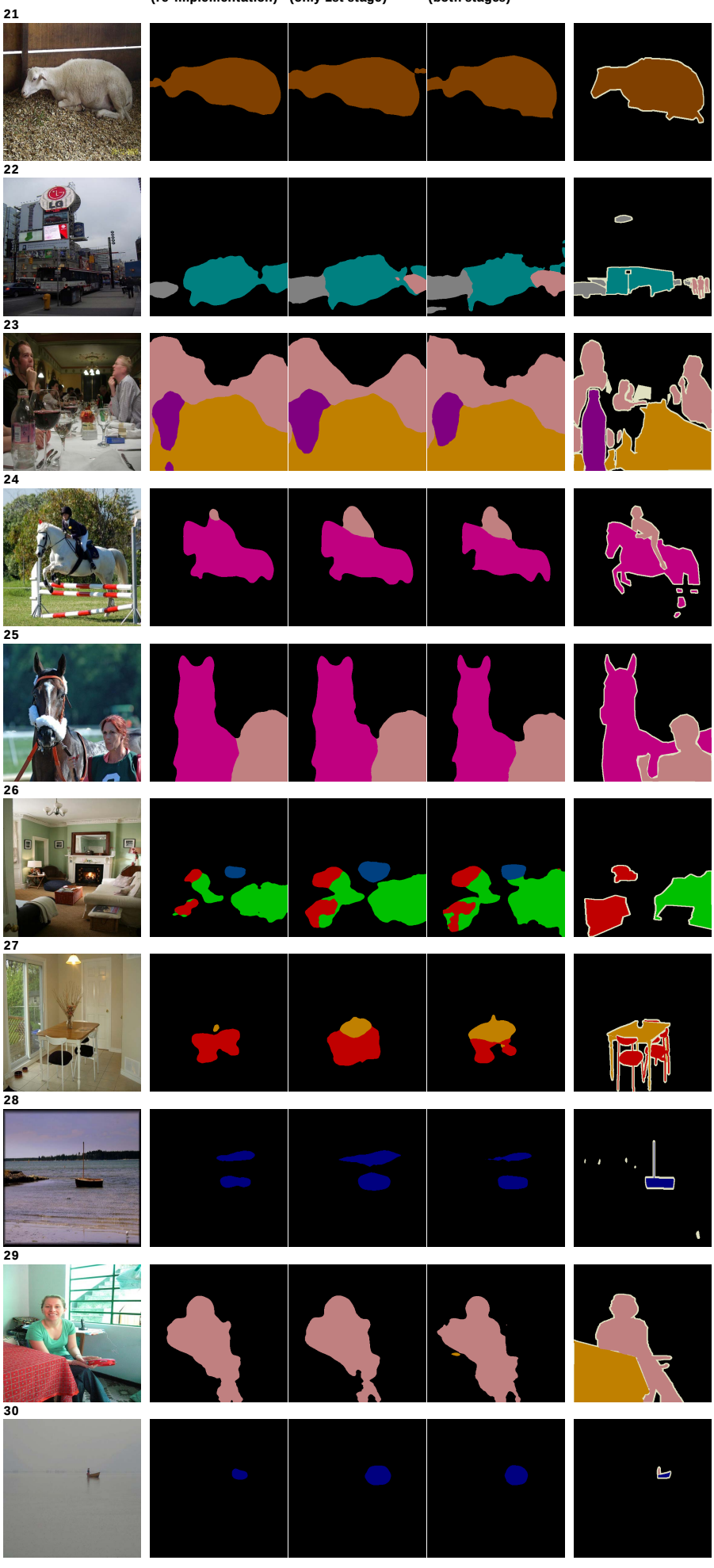
(† AS:Saliency mask, WV:web videos. WI Web images. IS Instance saliency mask.)

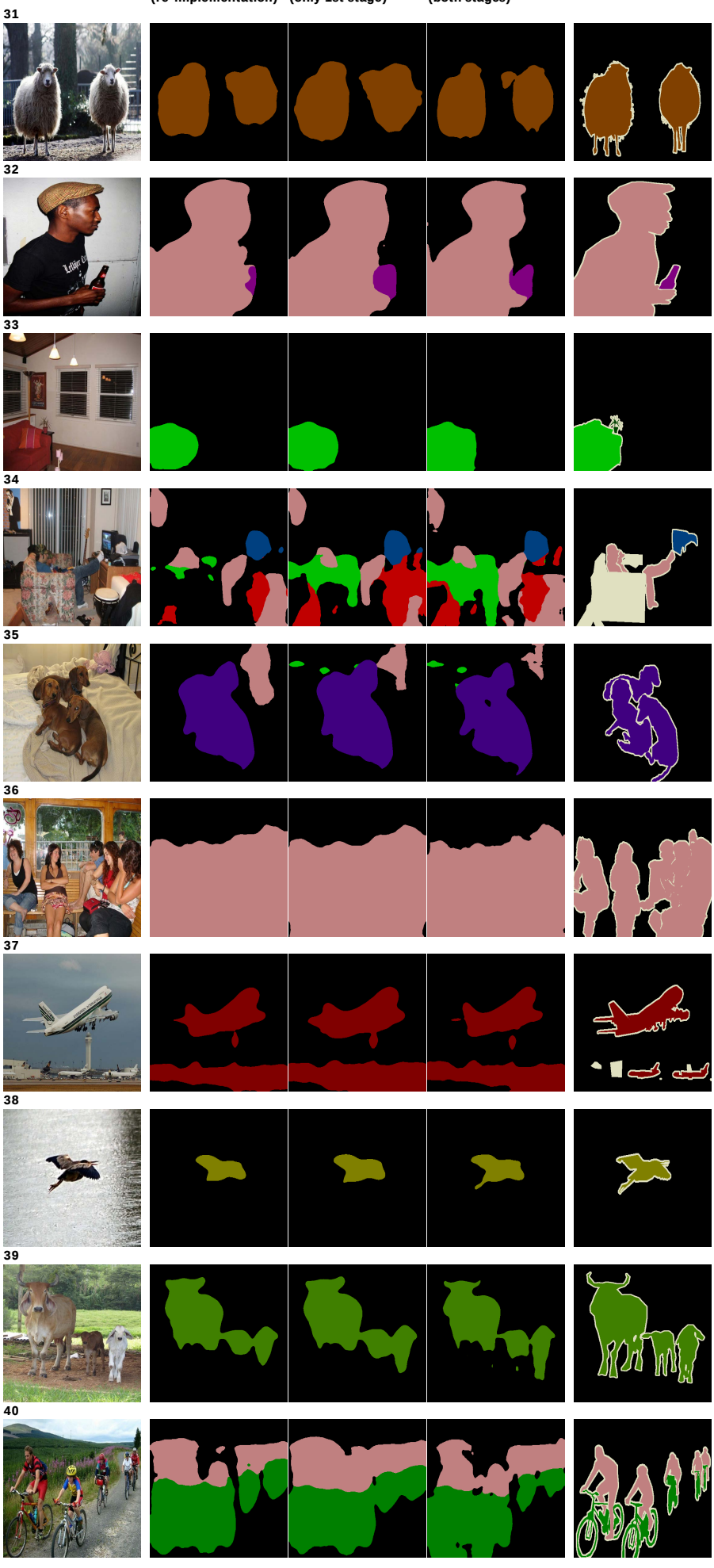
B. Five hundred results on the validation set.

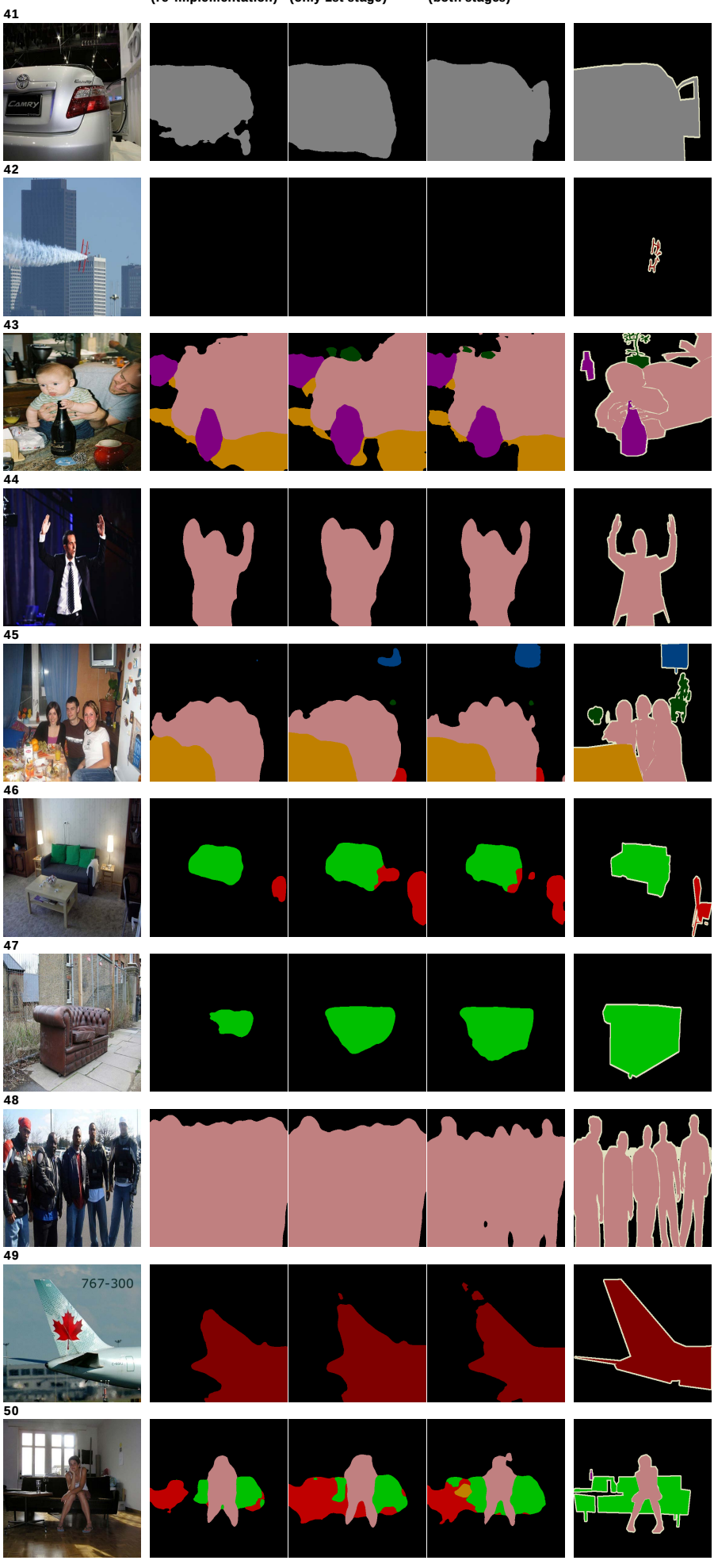
The results are given in Table A-2. In each row, an input image, a result by PSA [1], a result by SSDD with only first stage, a result by SSDD with both the first and second stage (full results), and ground truth are shown in the leftmost column.

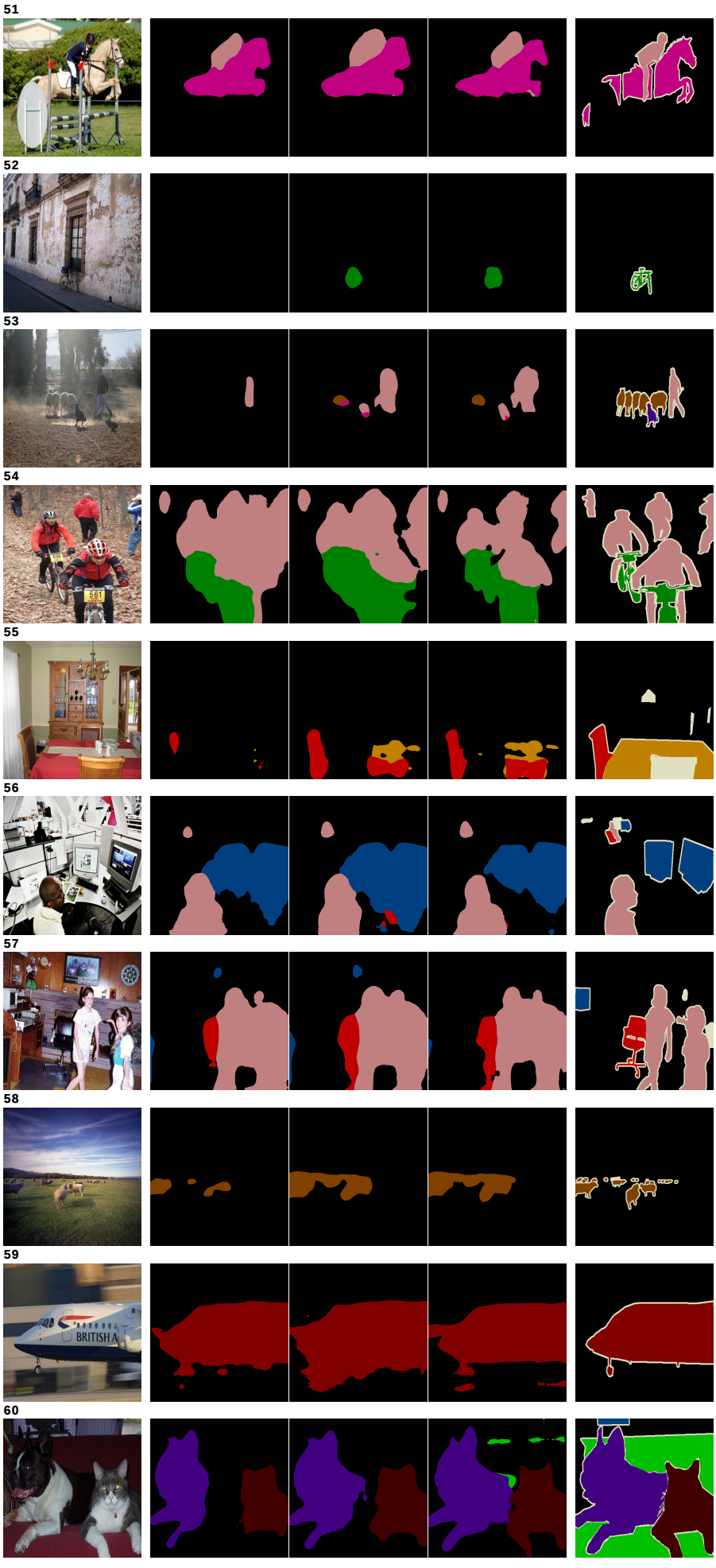

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
1
2
3
4
5
6
7
8
9
10

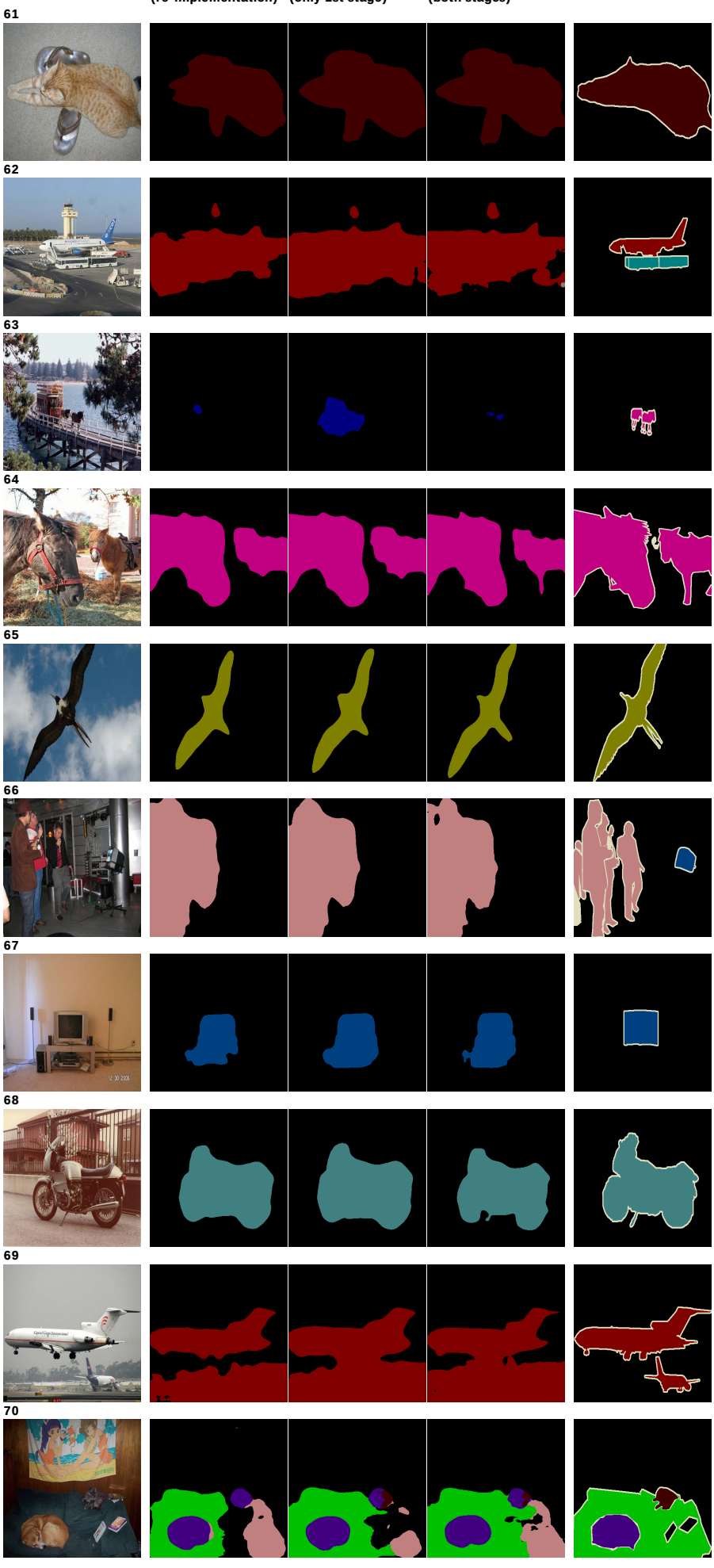

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
11
12
13
14
15
16
17
18
19
20

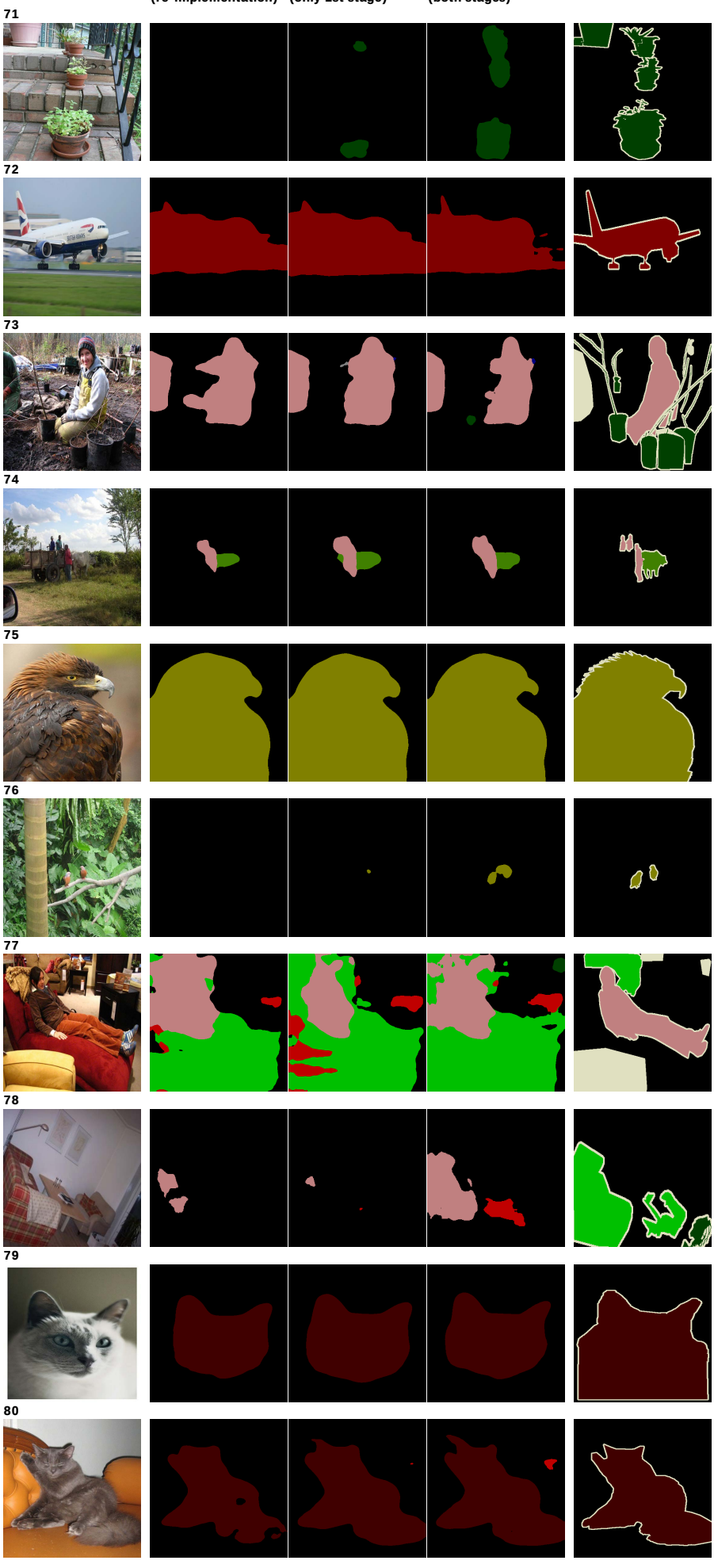

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
21
22
23
24
25
26
27
28
29
30

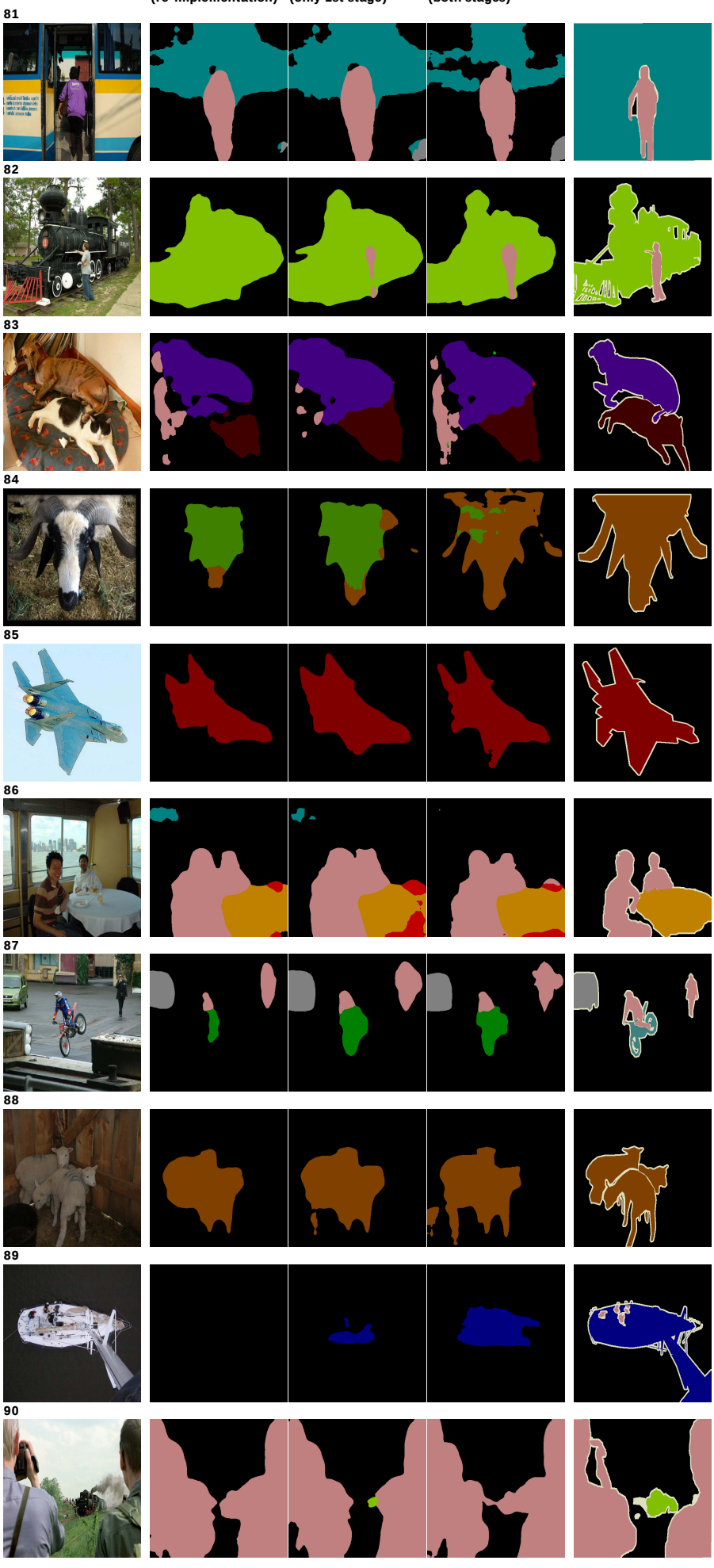

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
31
32
33
34
35
36
37
38
39
40

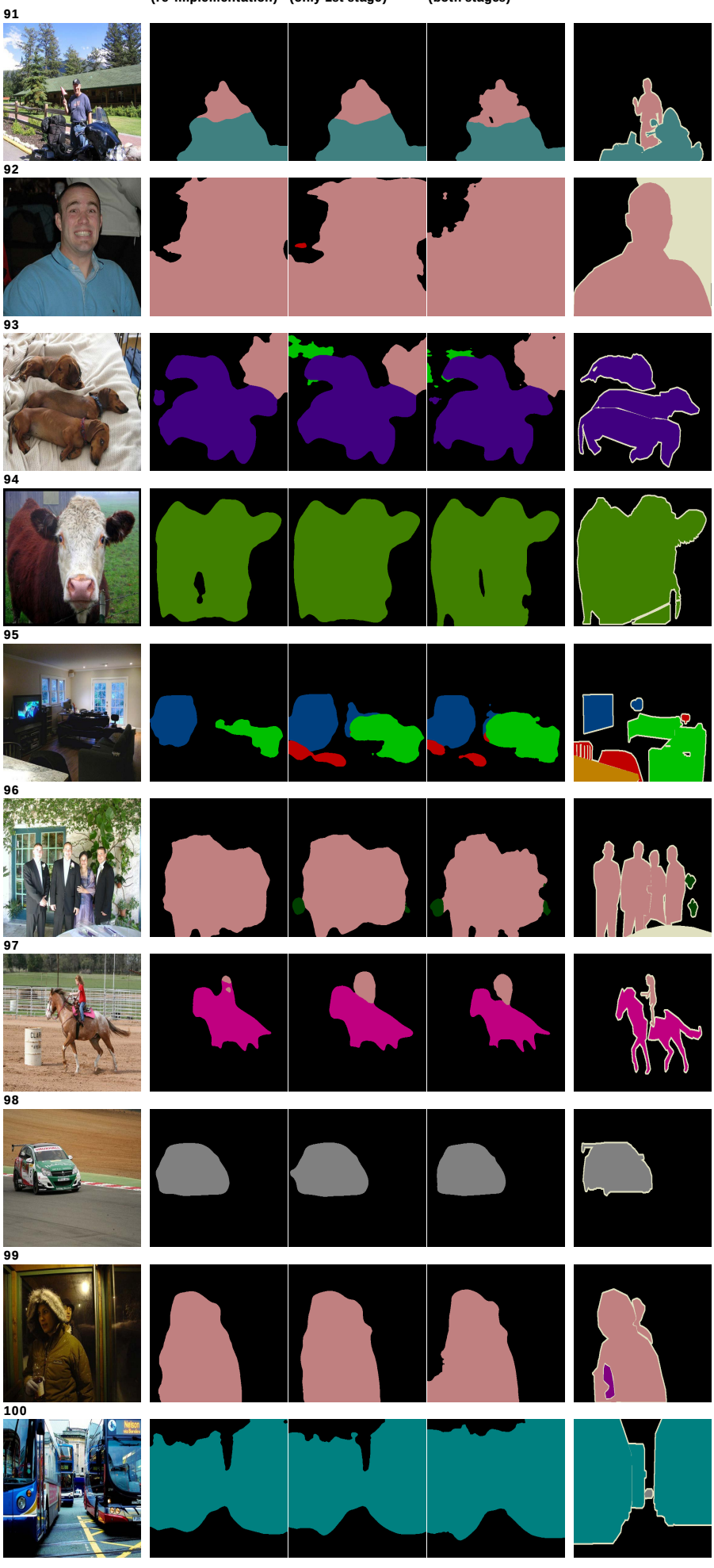

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
41
42
43
44
45
46
47
48
49
50
CAMRY
767-300

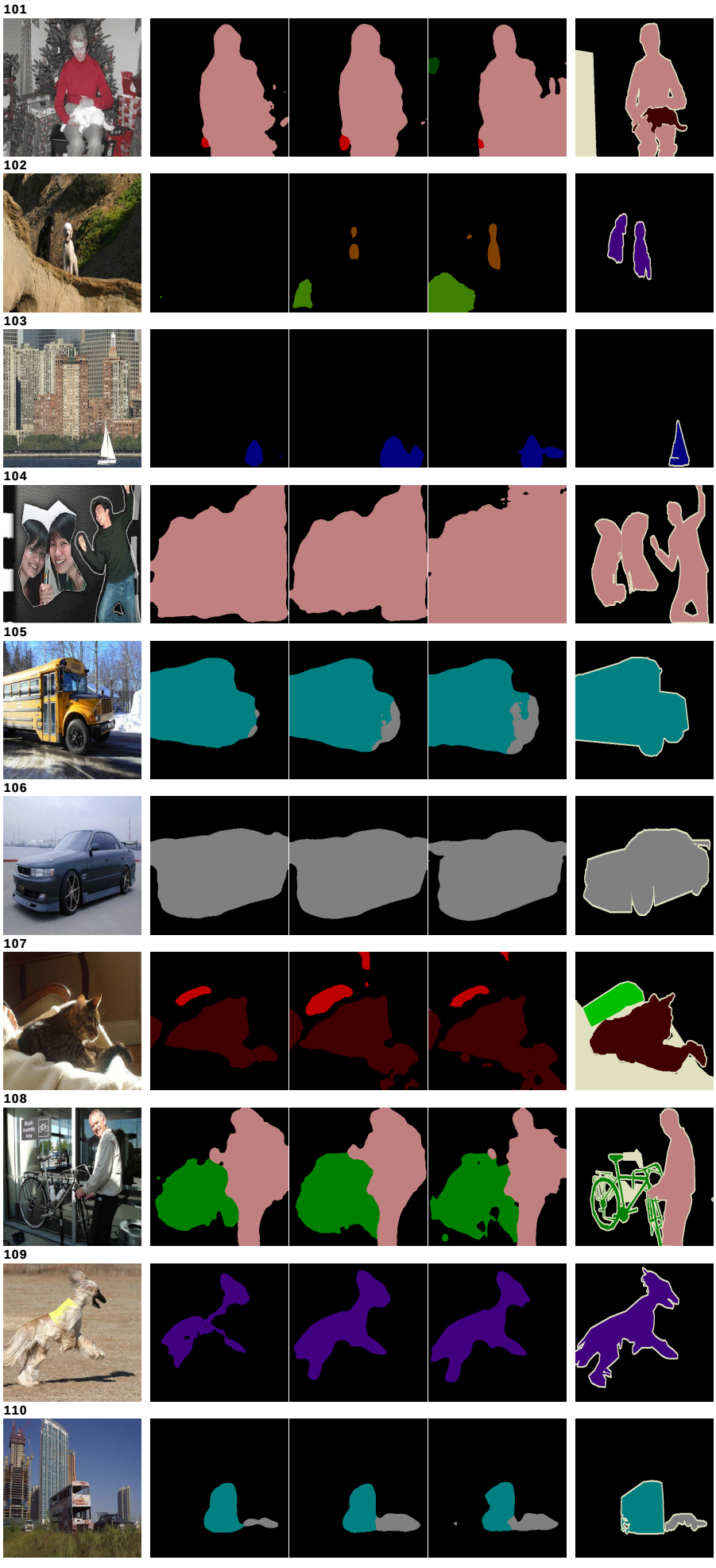

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
51
52
53
54
55
56
57
58
59
60

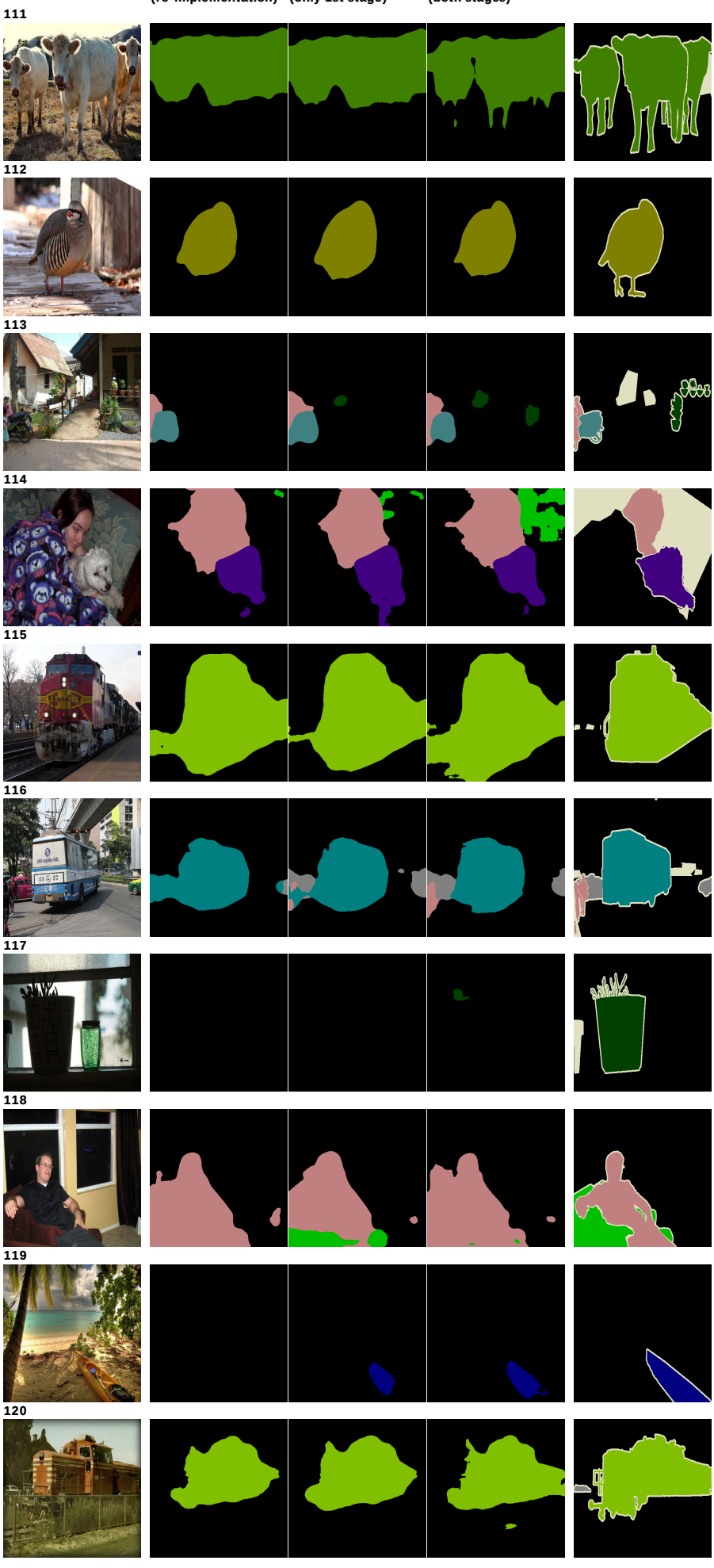

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
61
62
63
64
65
66
67
68
69
70

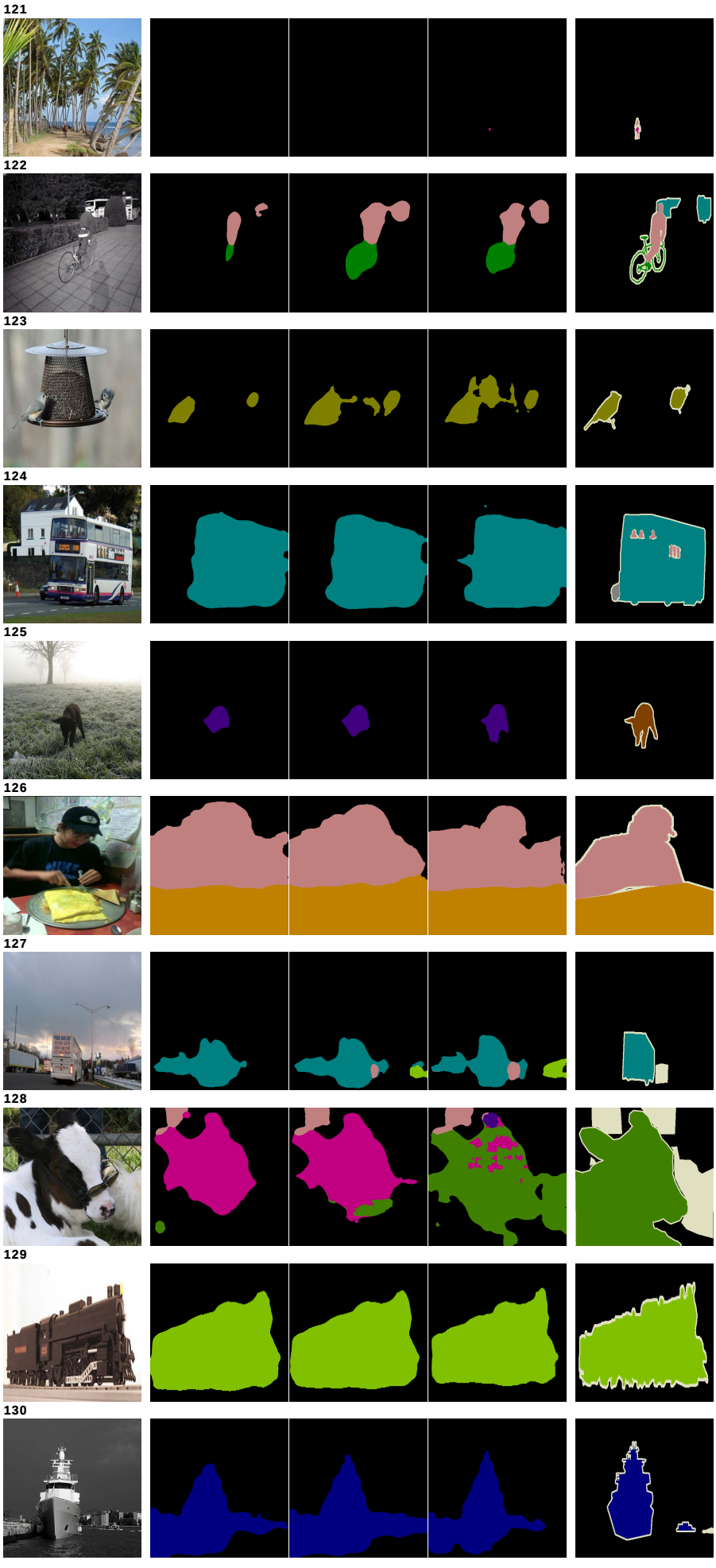

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
71
72
73
74
75
76
77
78
79
80

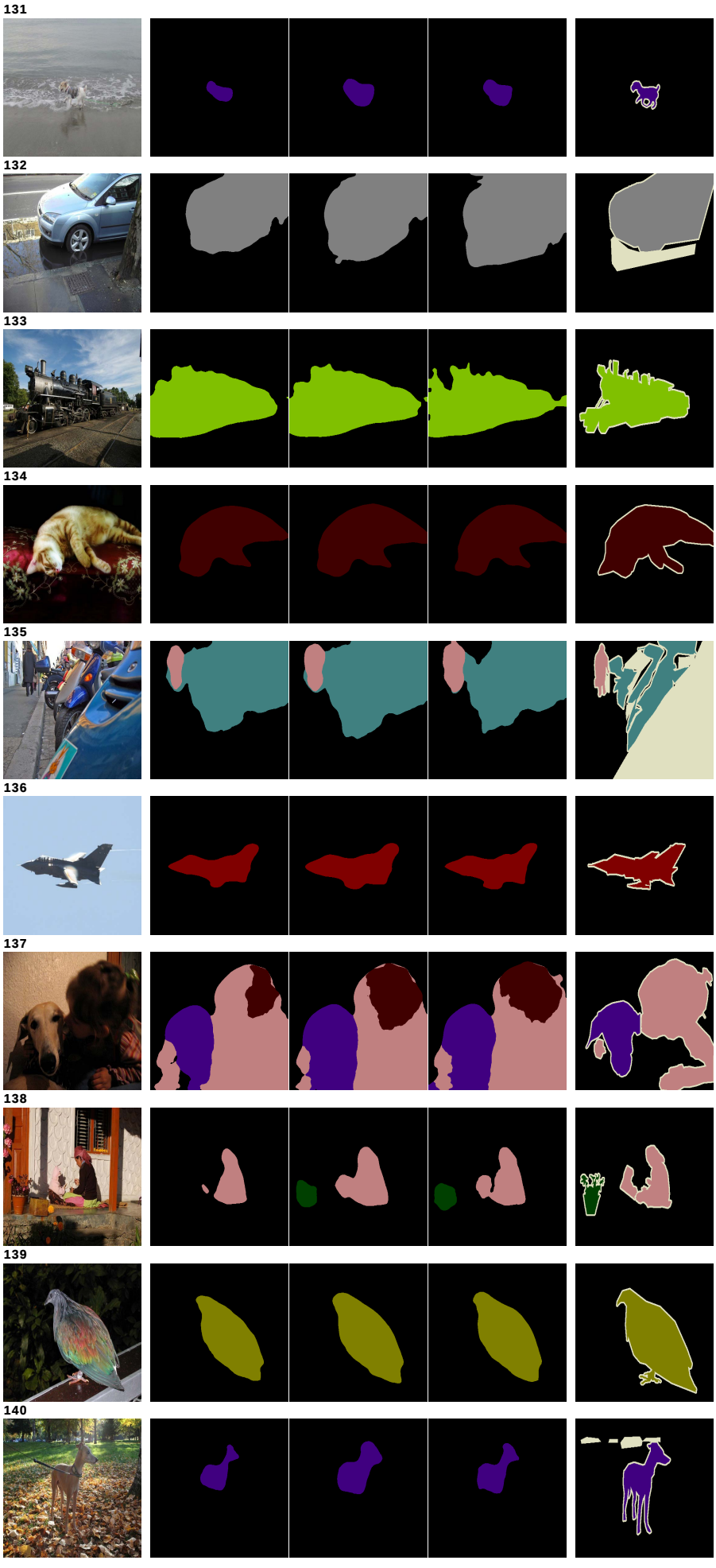

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
81
82
83
84
85
86
87
88
89
90

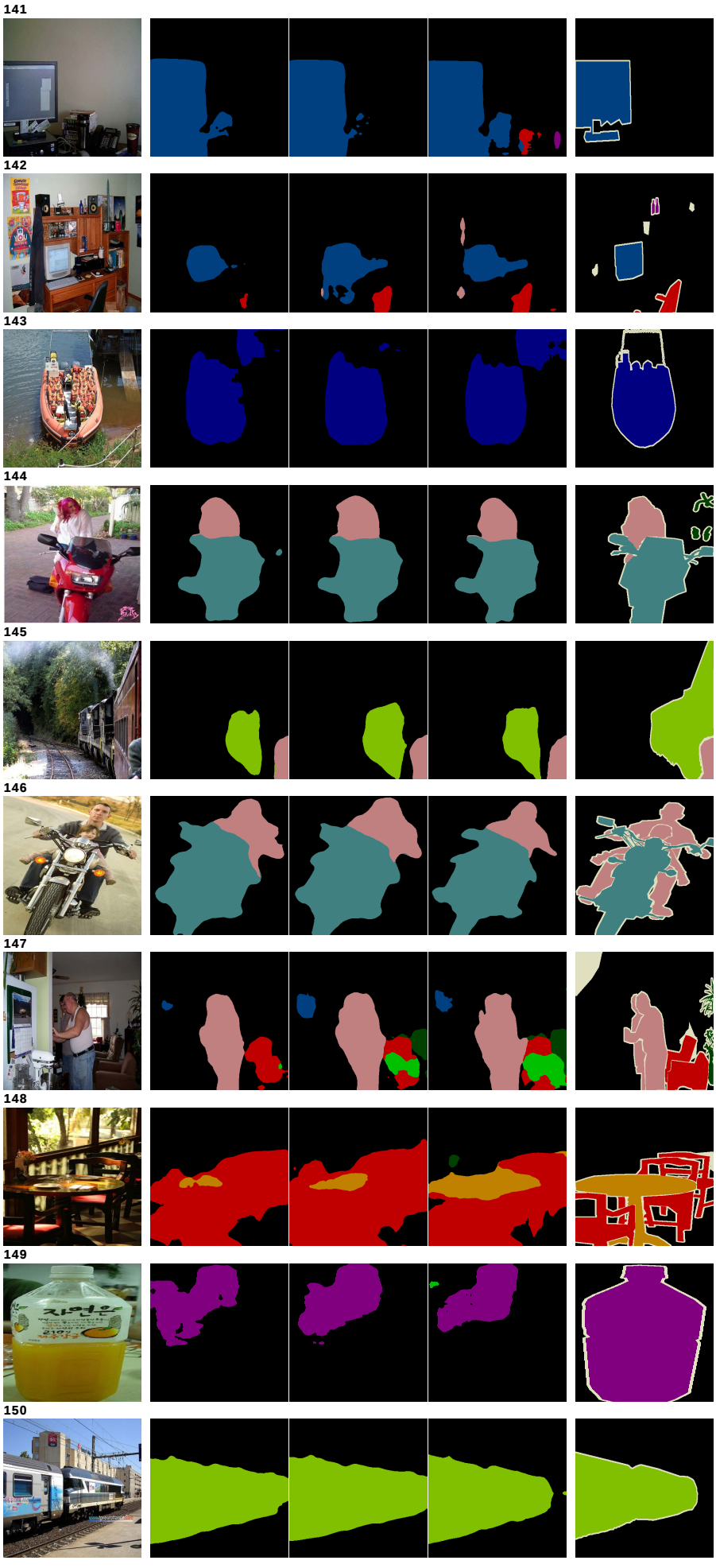

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
91
92
93
94
95
96
97
98
99
100

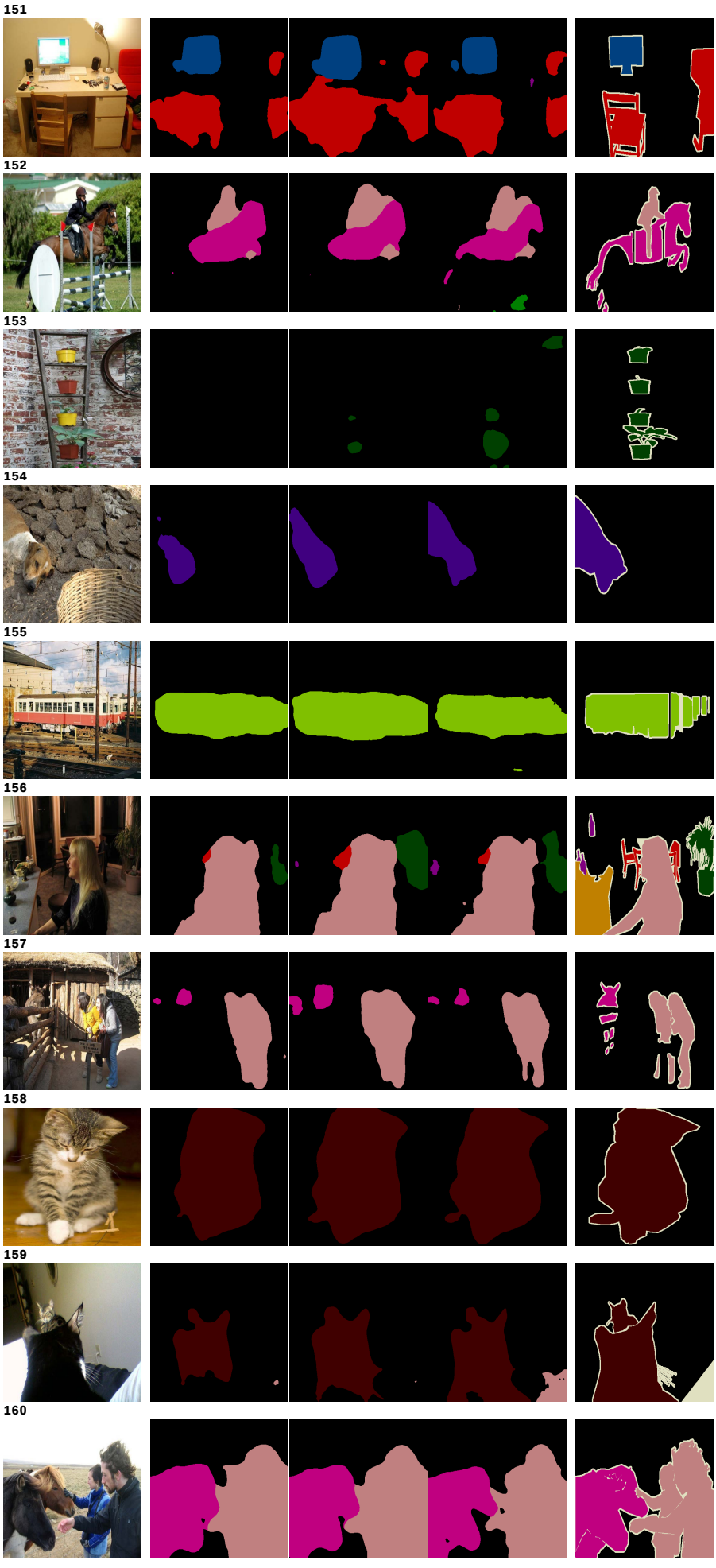

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
101
102
103
104
105
106
107
108
109
110

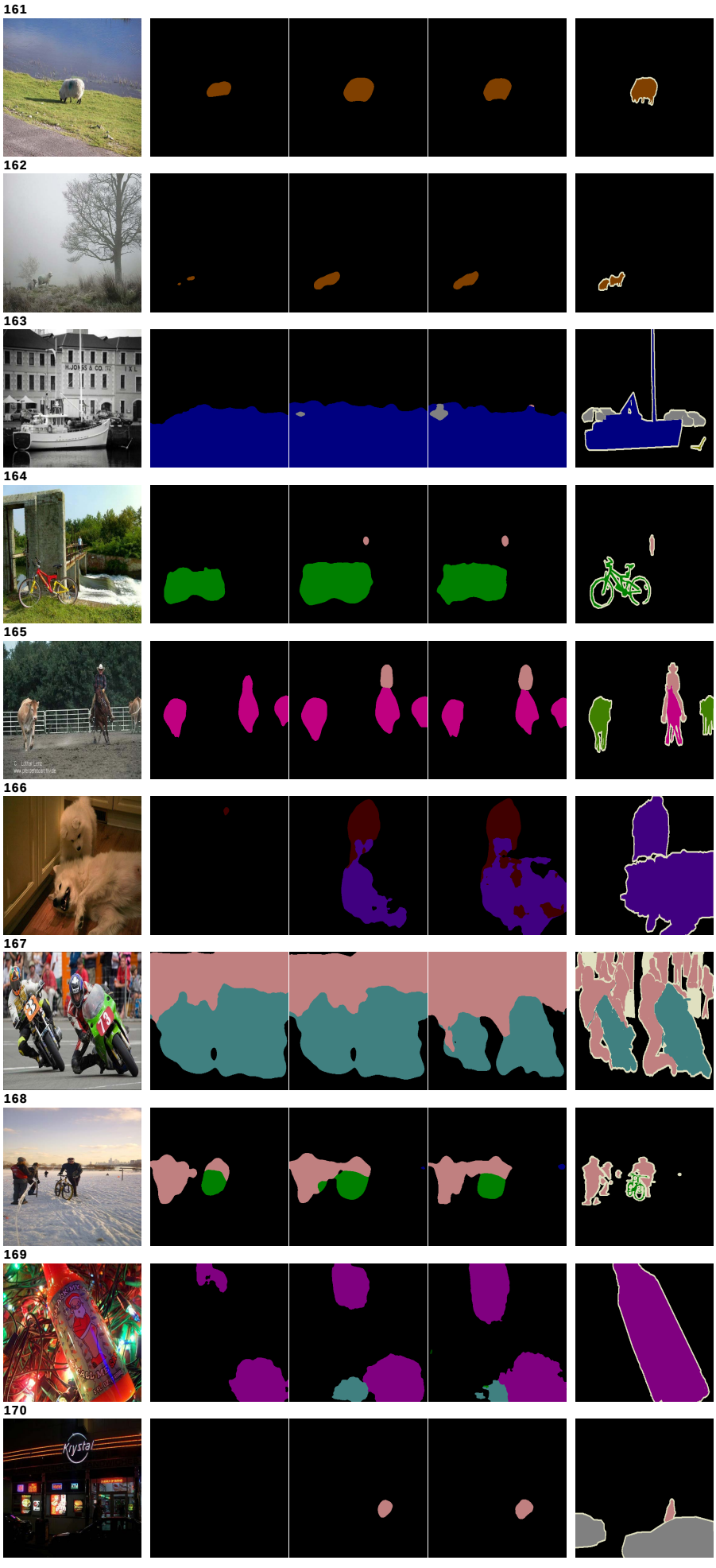

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
111
112
113
114
115
116
117
118
119
120

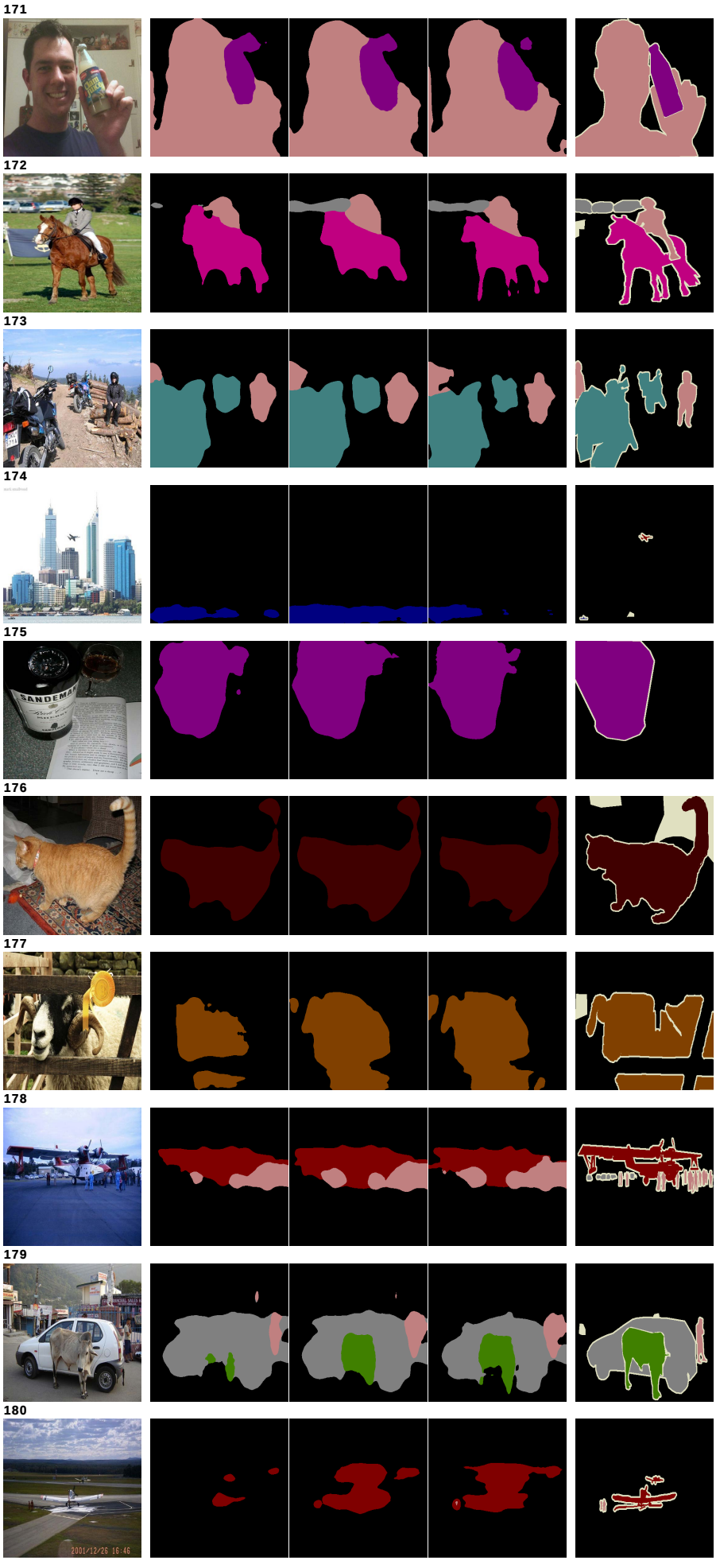

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
121
122
123
124
125
126
127
128
129
130

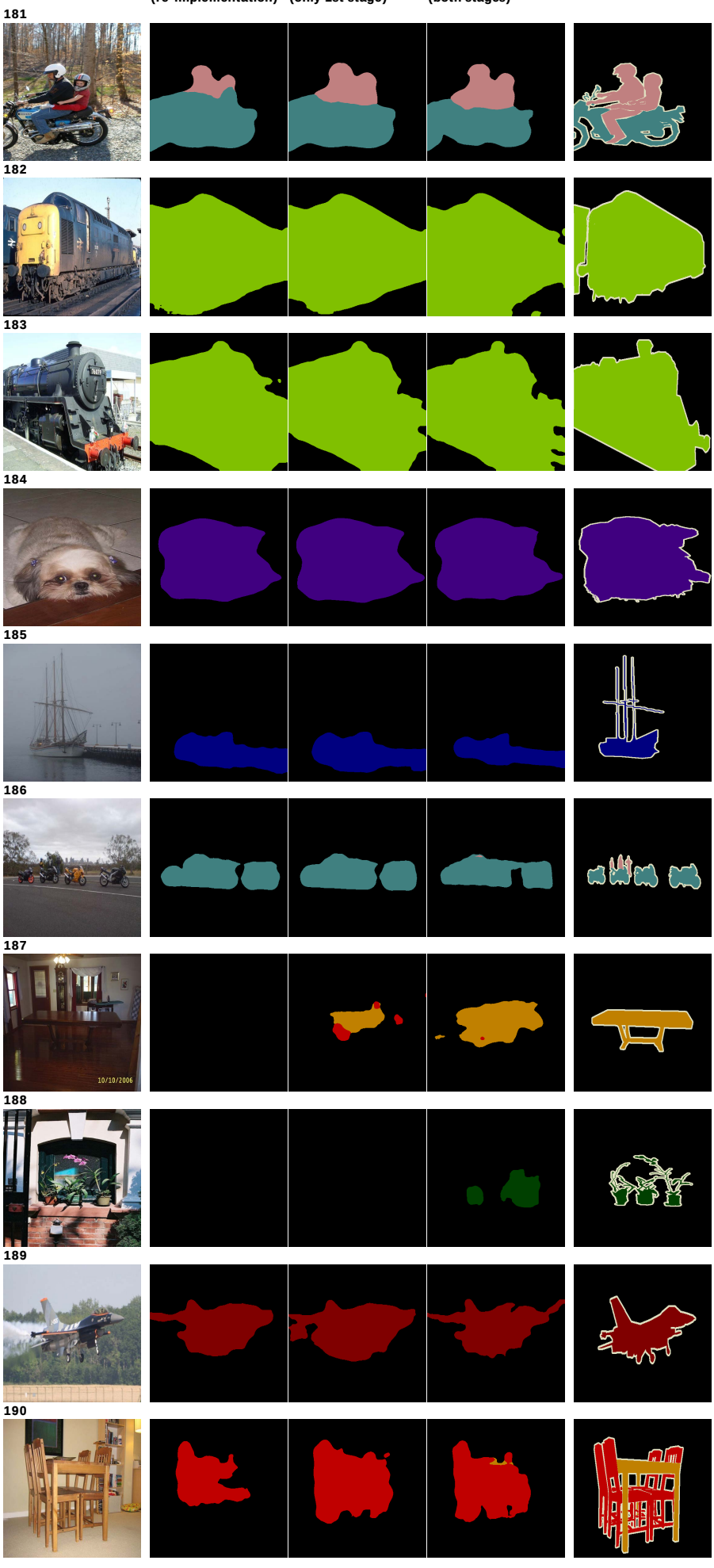

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
131
132
133
134
135
136
137
138
139
140

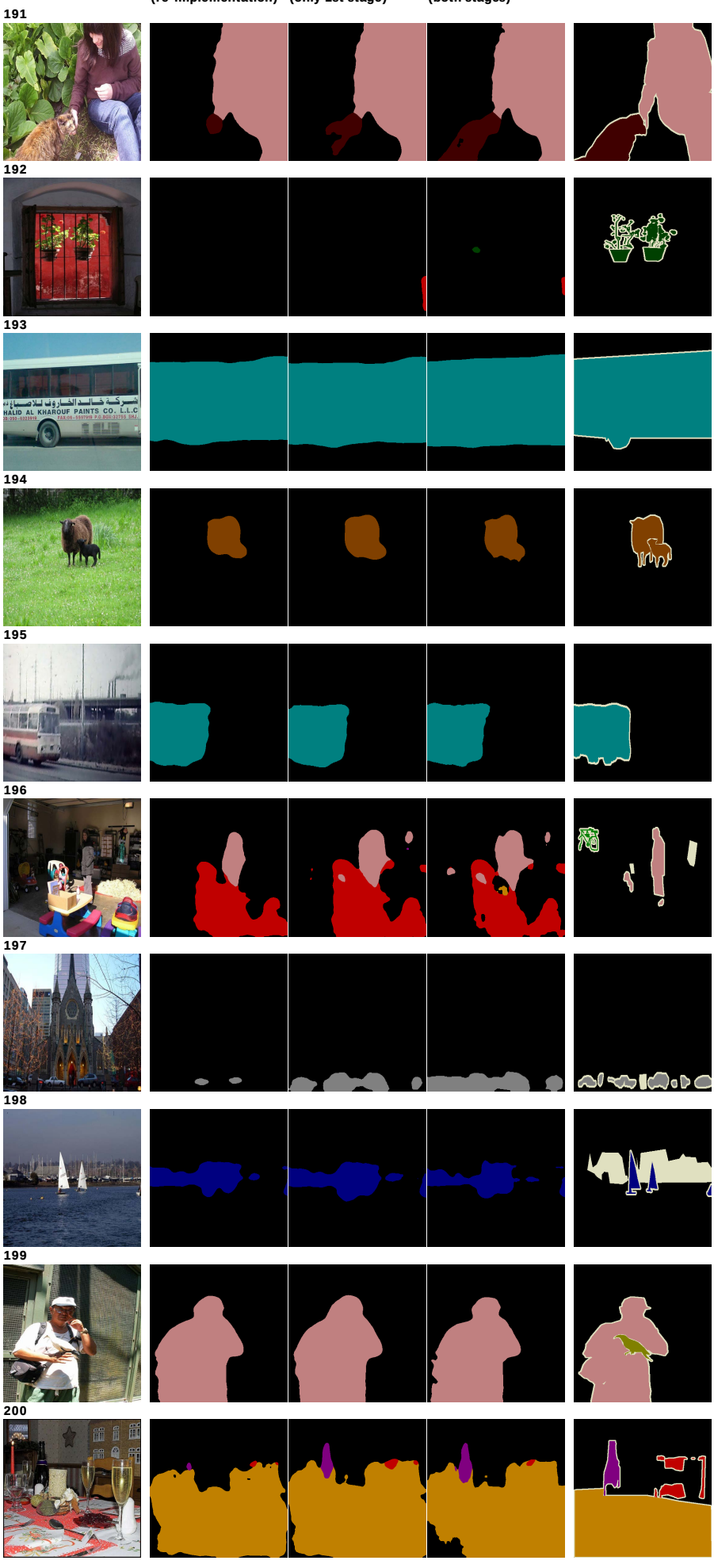

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
141
142
143
144
145
146
147
148
149
150

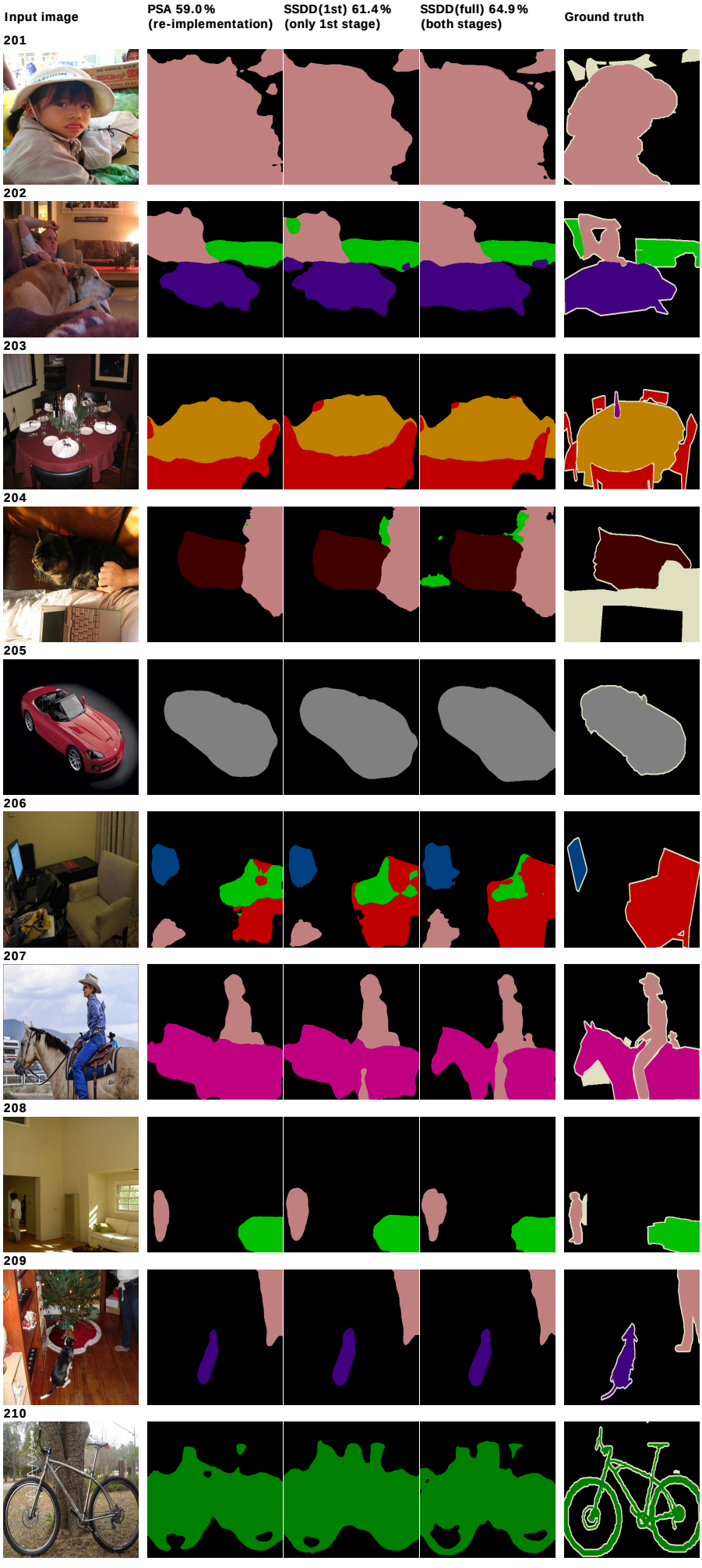

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
151
152
153
154
155
156
157
158
159
160

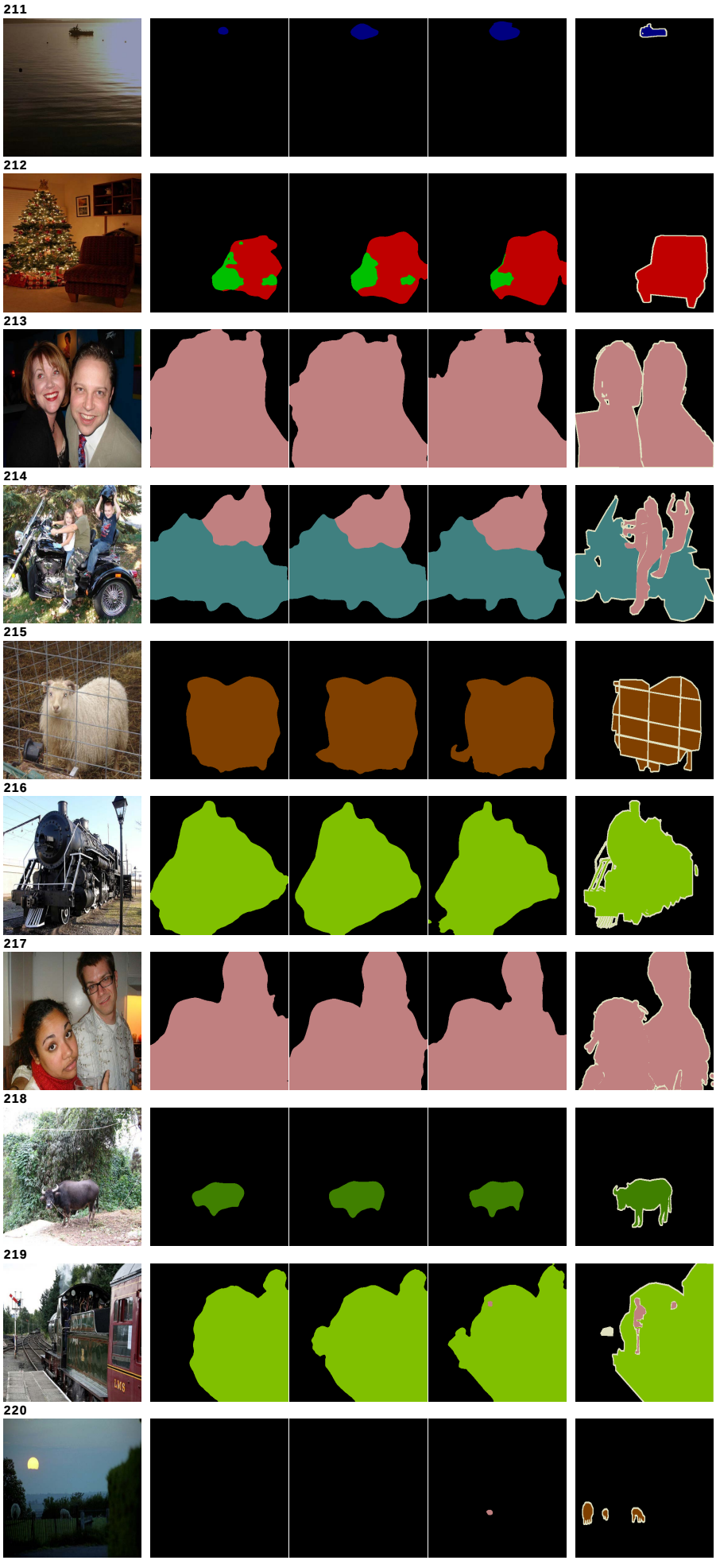

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
161
162
163
164
165
166
167
168
169
170

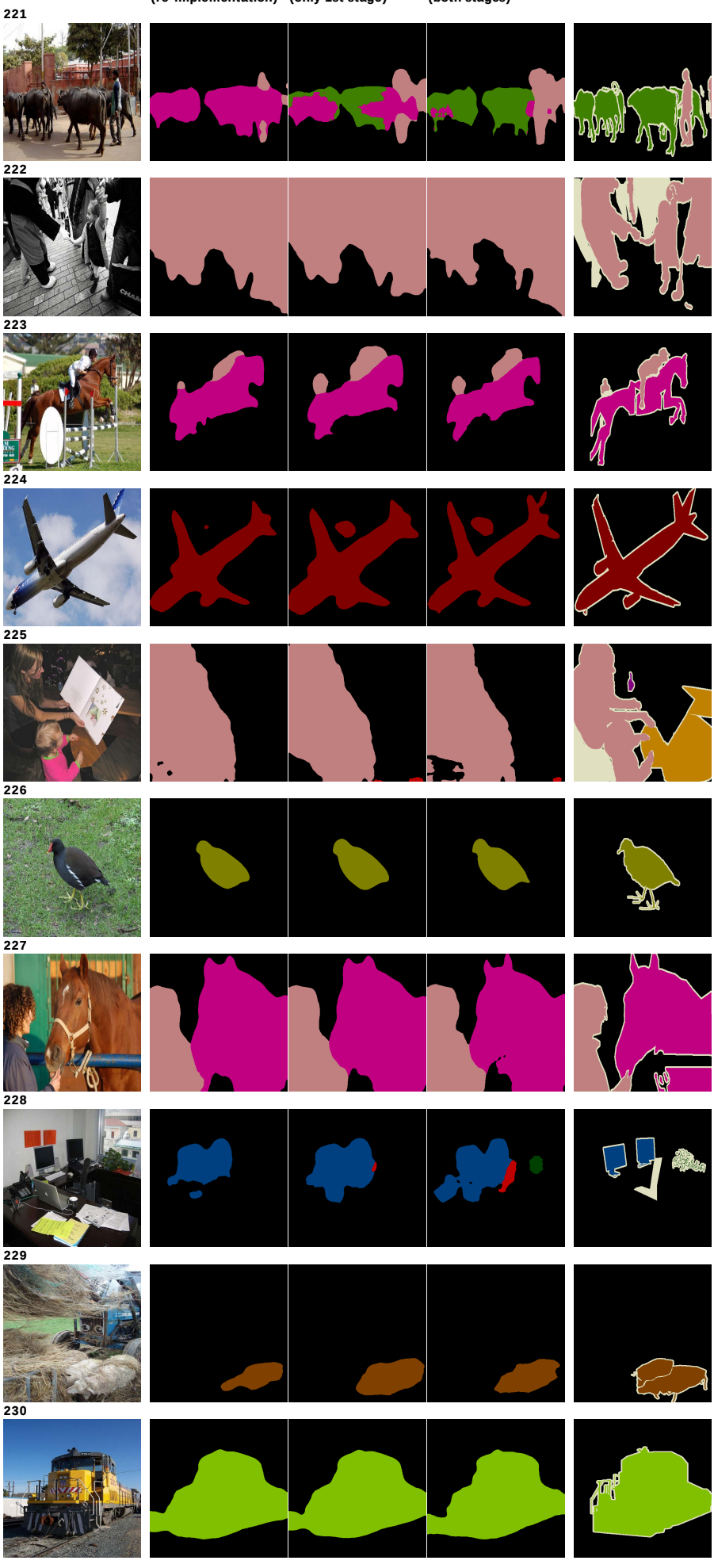

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
171
172
173
174
175
176
177
178
179
180

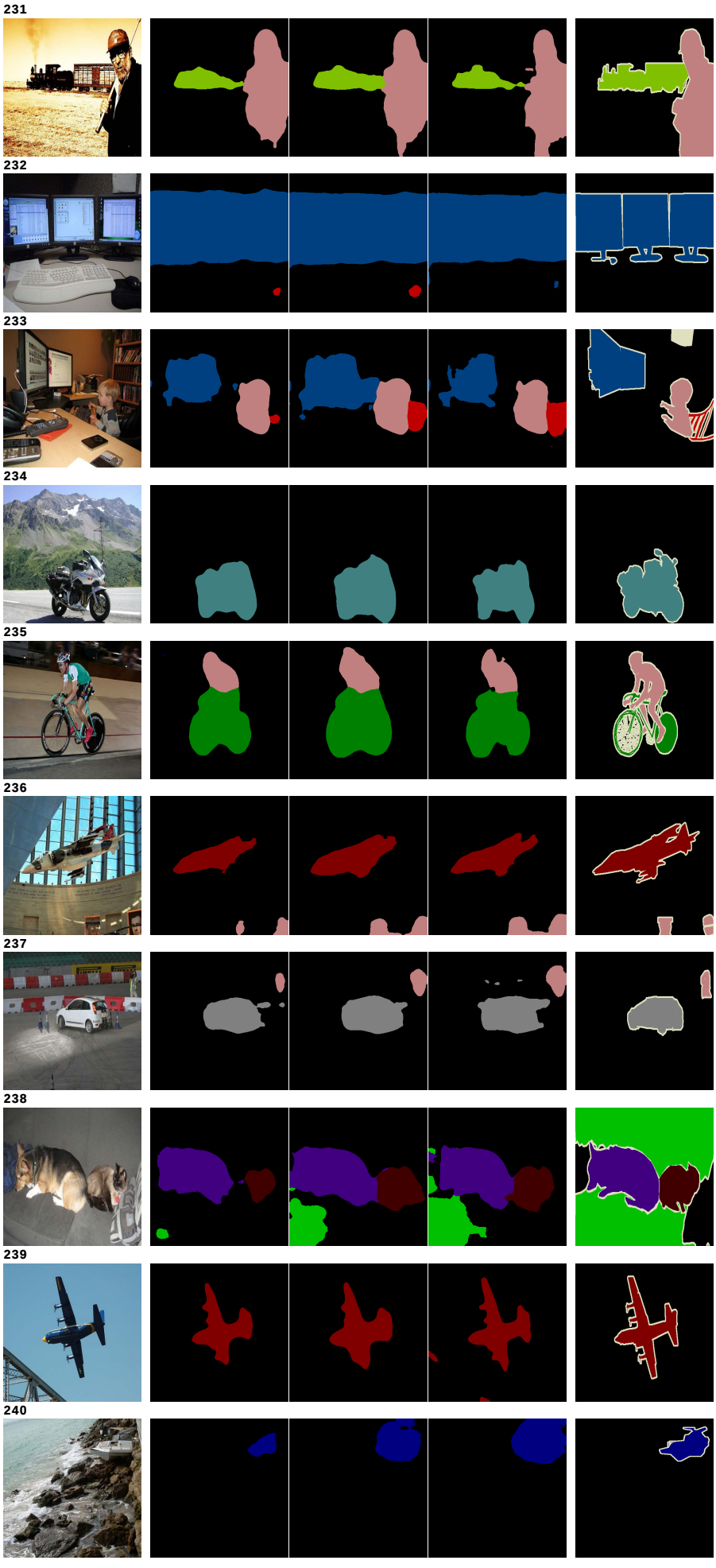

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
181
182
183
184
185
186
187
188
189
190

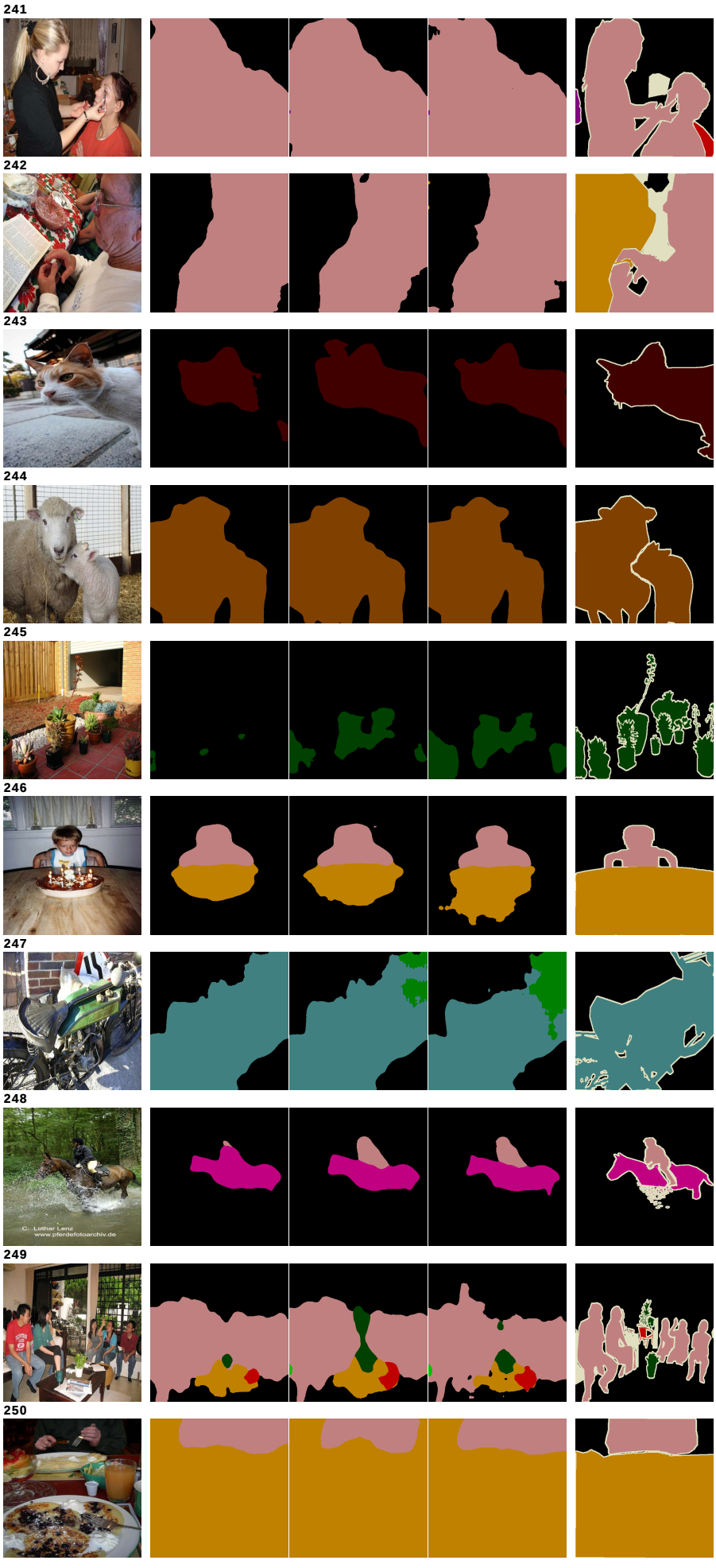

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
191
192
193
194
195
196
197
198
199
200

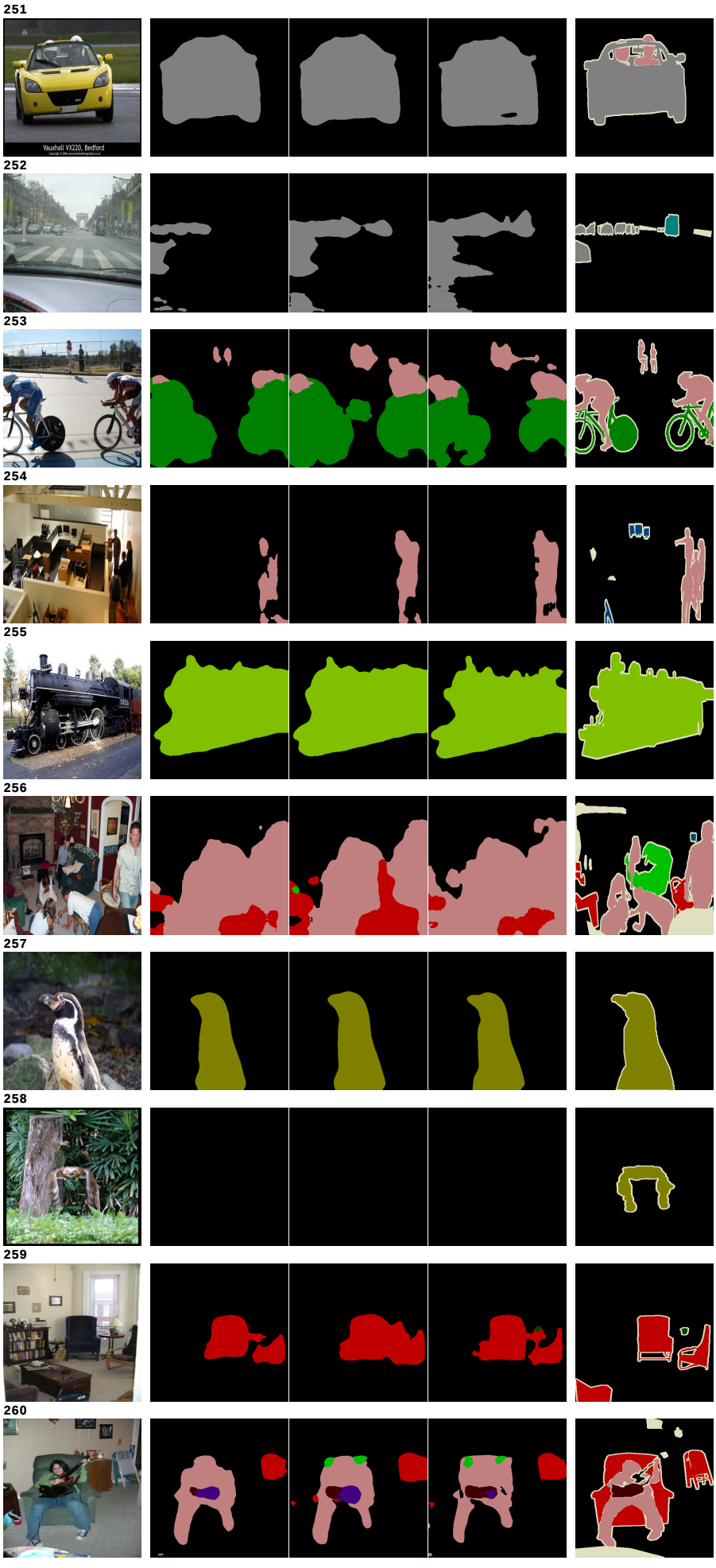

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
201
202
203
204
205
206
207
208
209
210

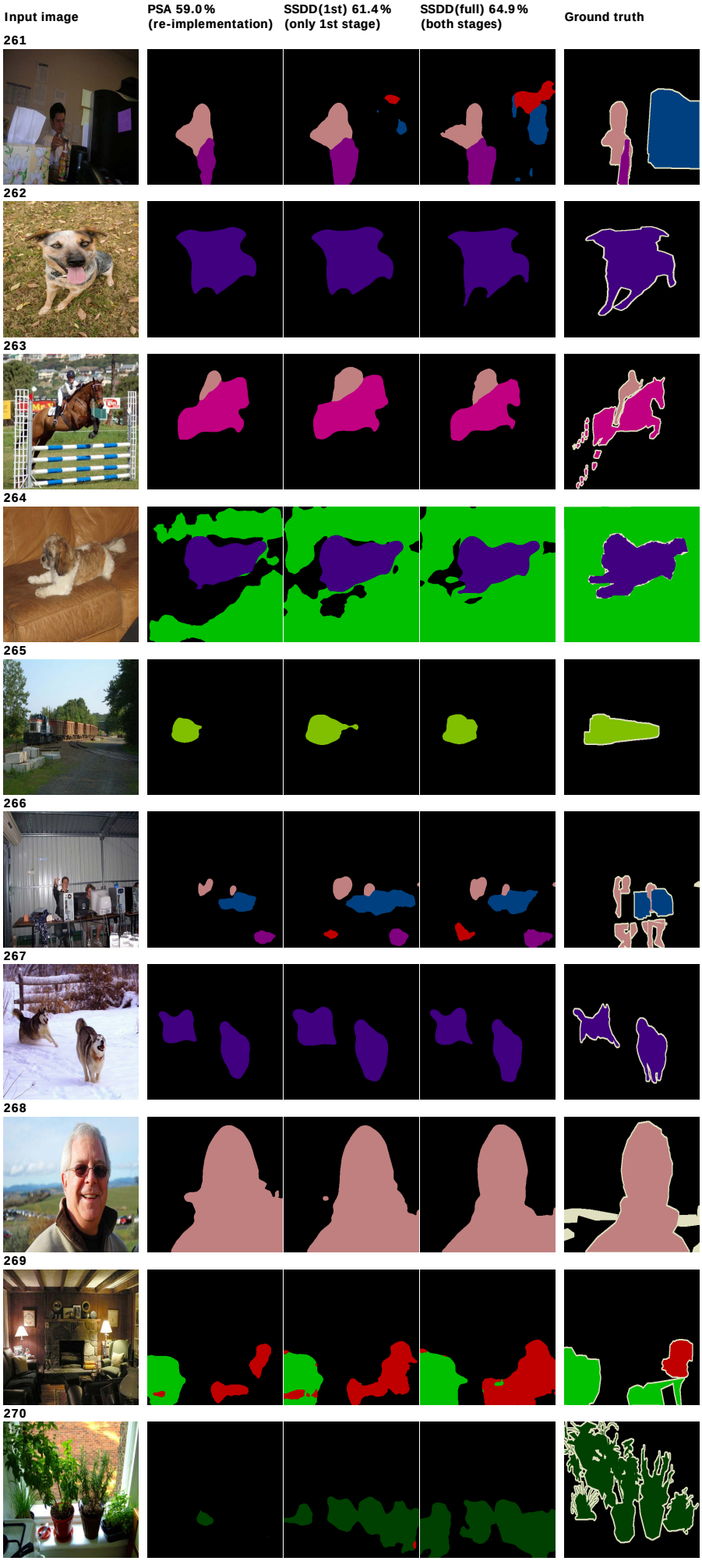

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
211
212
213
214
215
216
217
218
219
220


Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
221
222
223
224
225
226
227
228
229
230

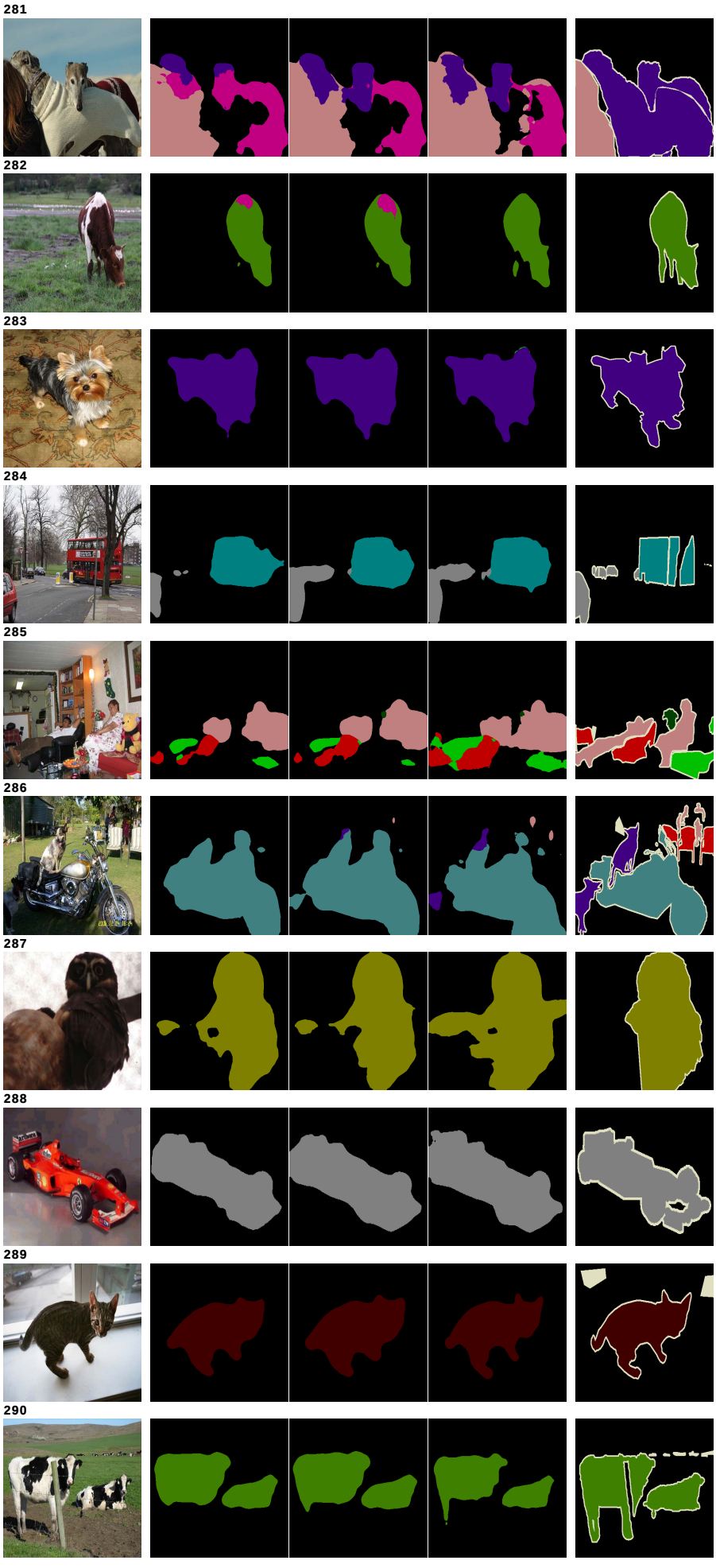

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
231
232
233
234
235
236
237
238
239
240

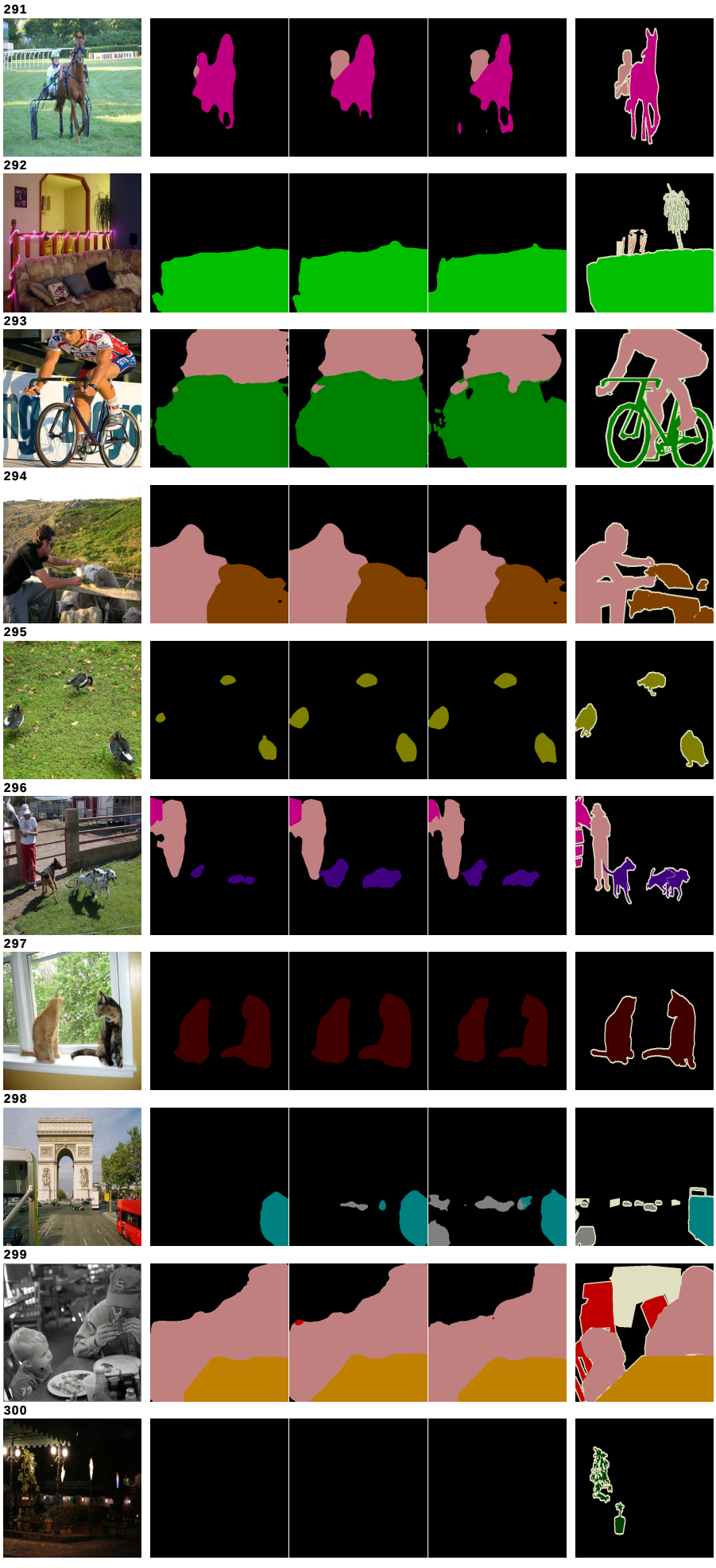

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
241
242
243
244
245
246
247
248
249
250

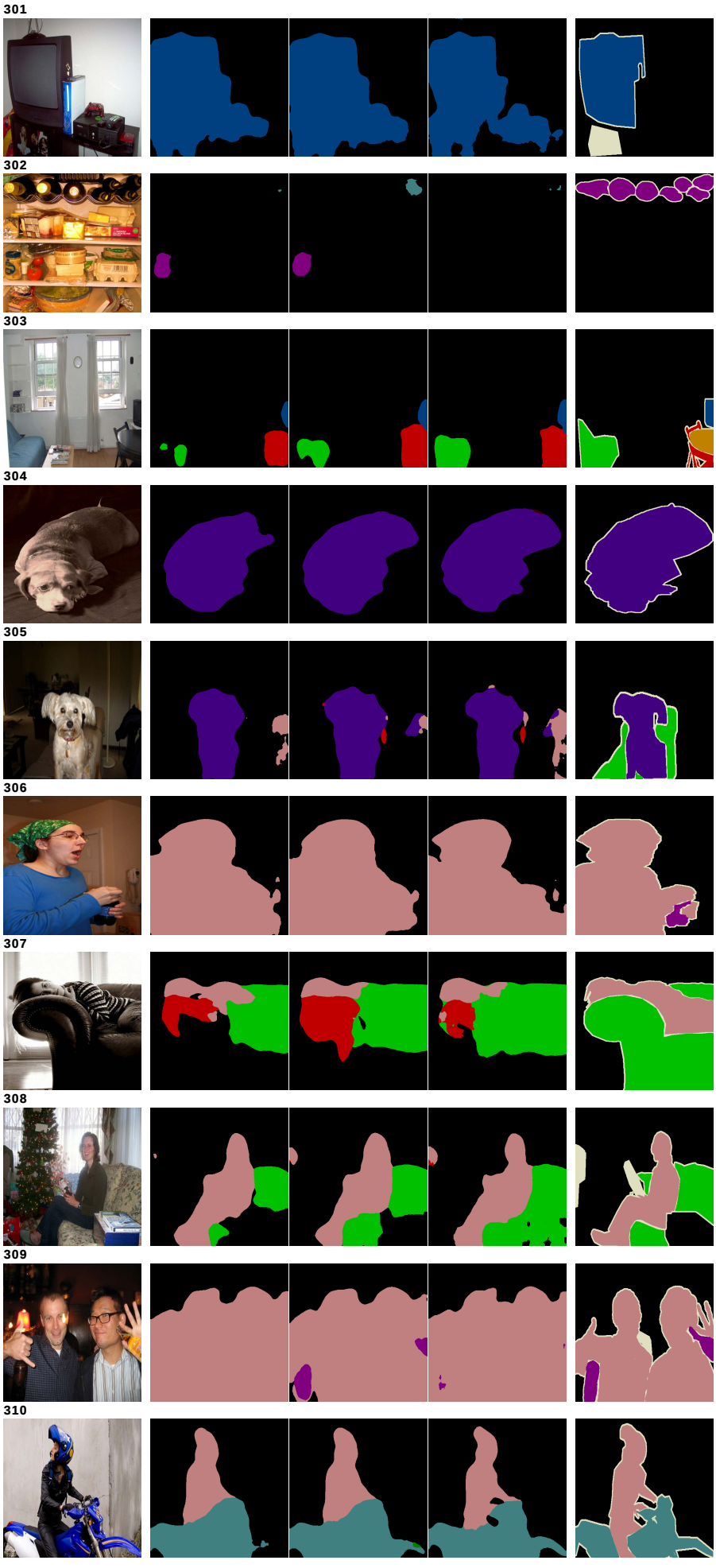

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
251
252
253
254
255
256
257
258
259
260

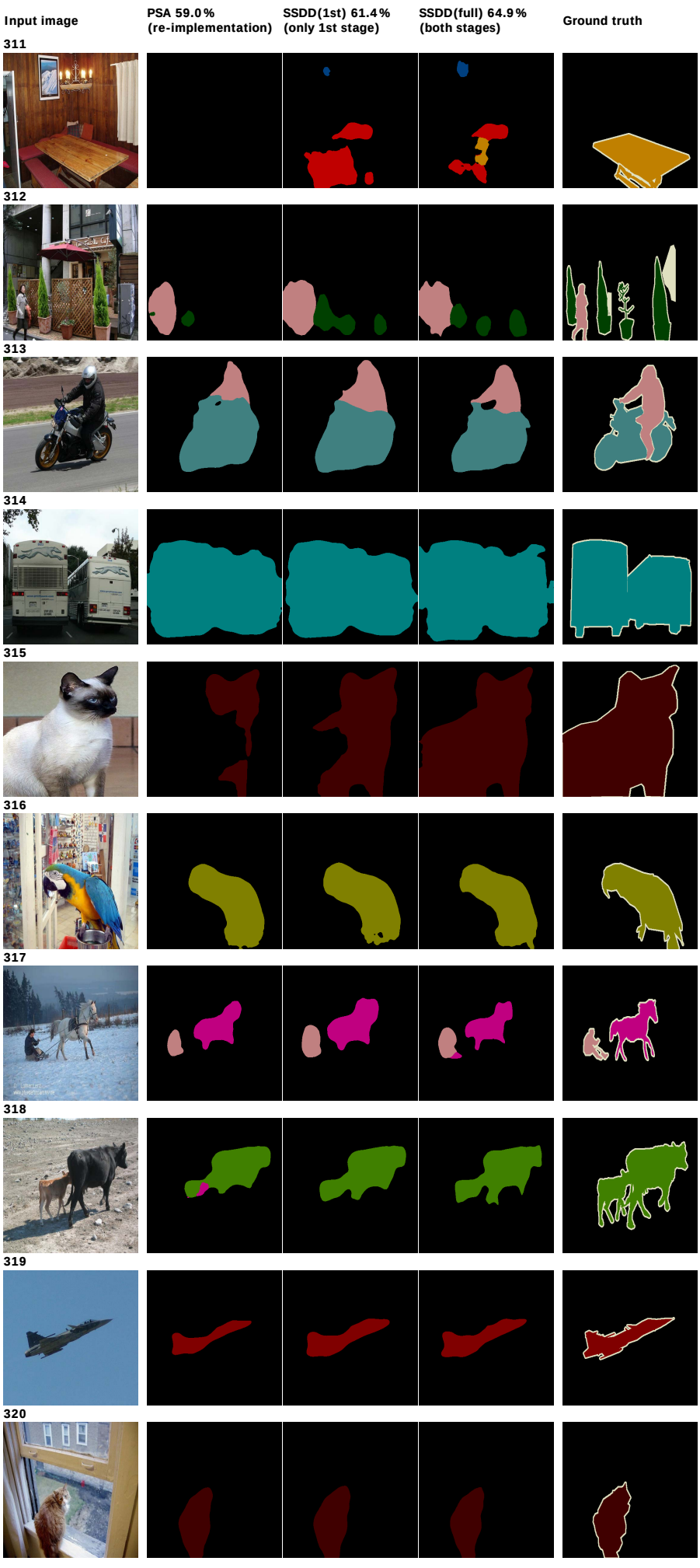

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
261
262
263
264
265
266
267
268
269
270

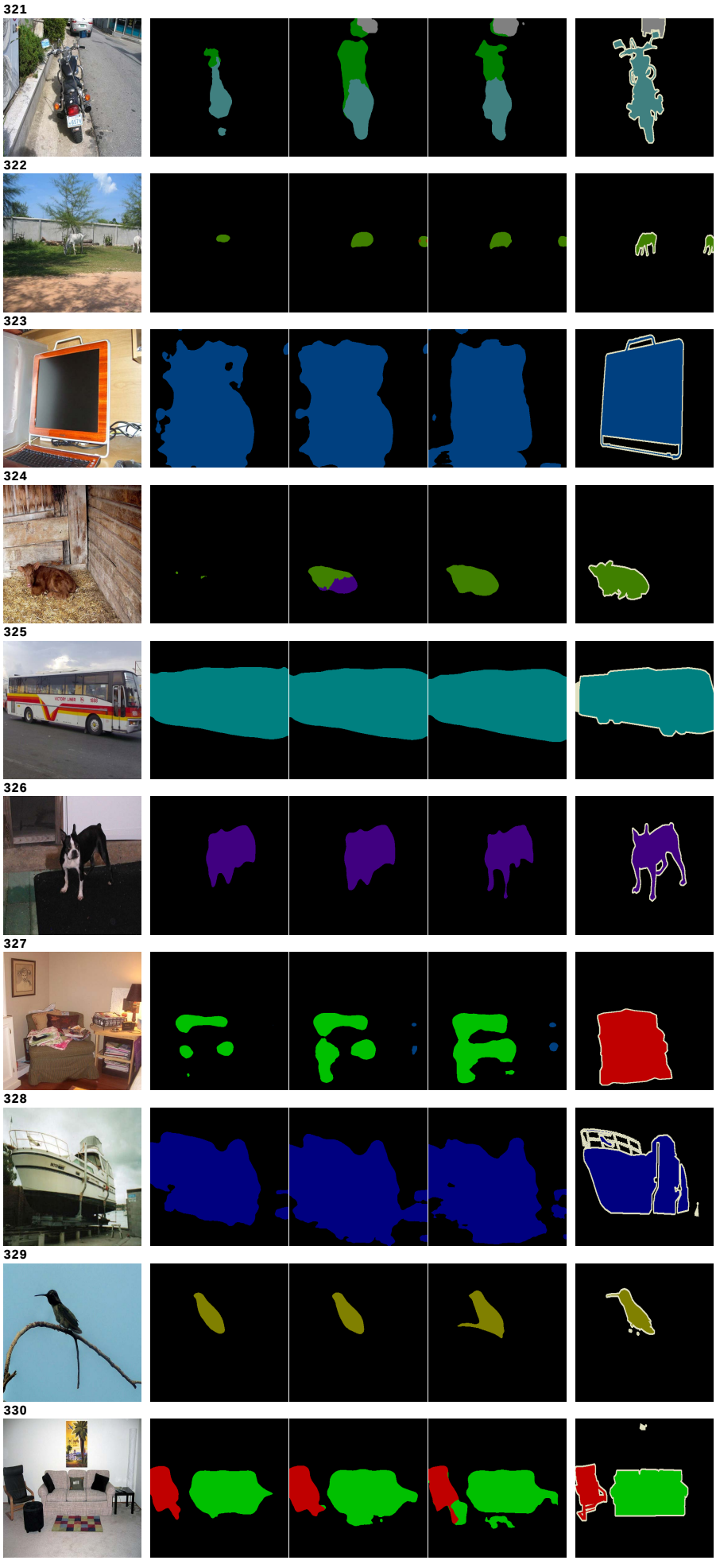

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
271
272
273
274
275
276
277
278
279
280

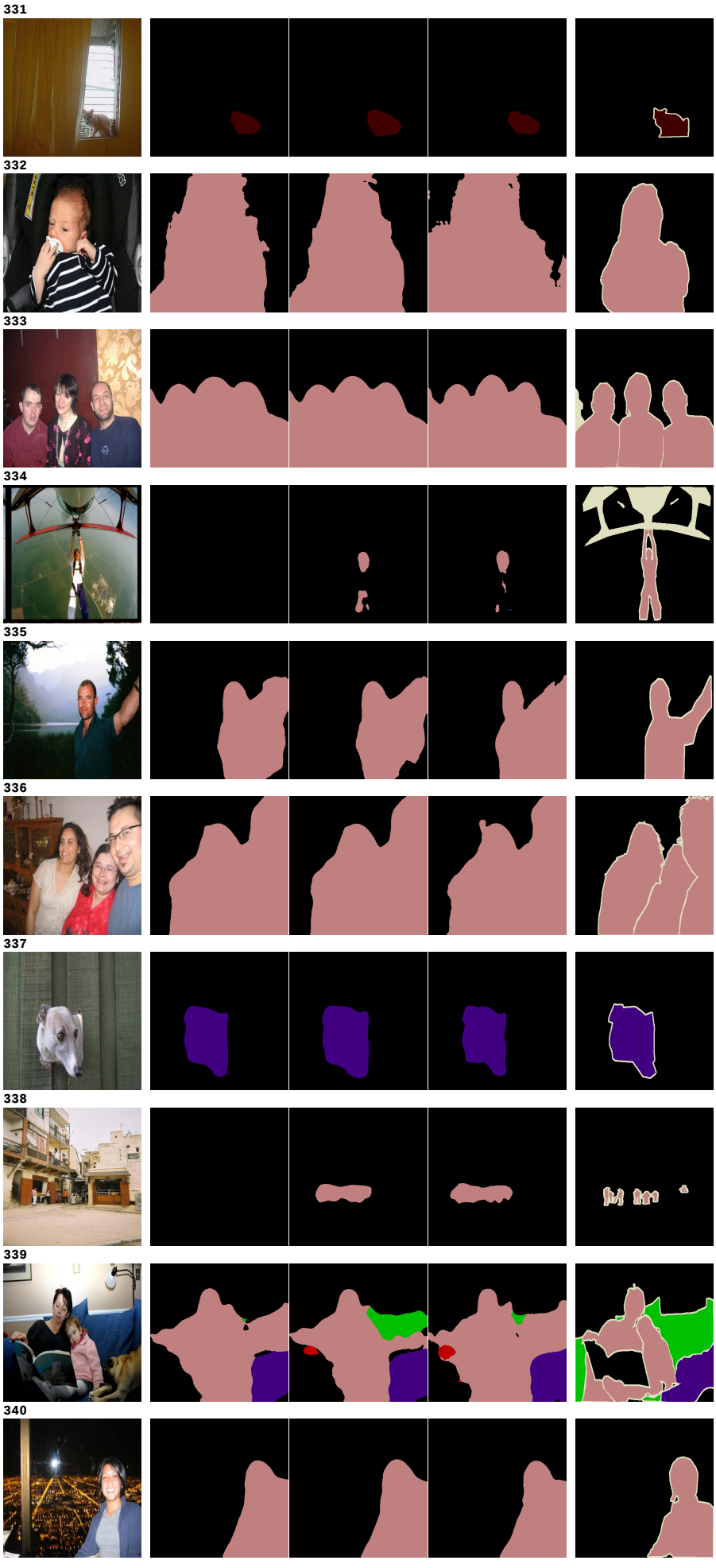

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
281
282
283
284
285
286
287
288
289
290

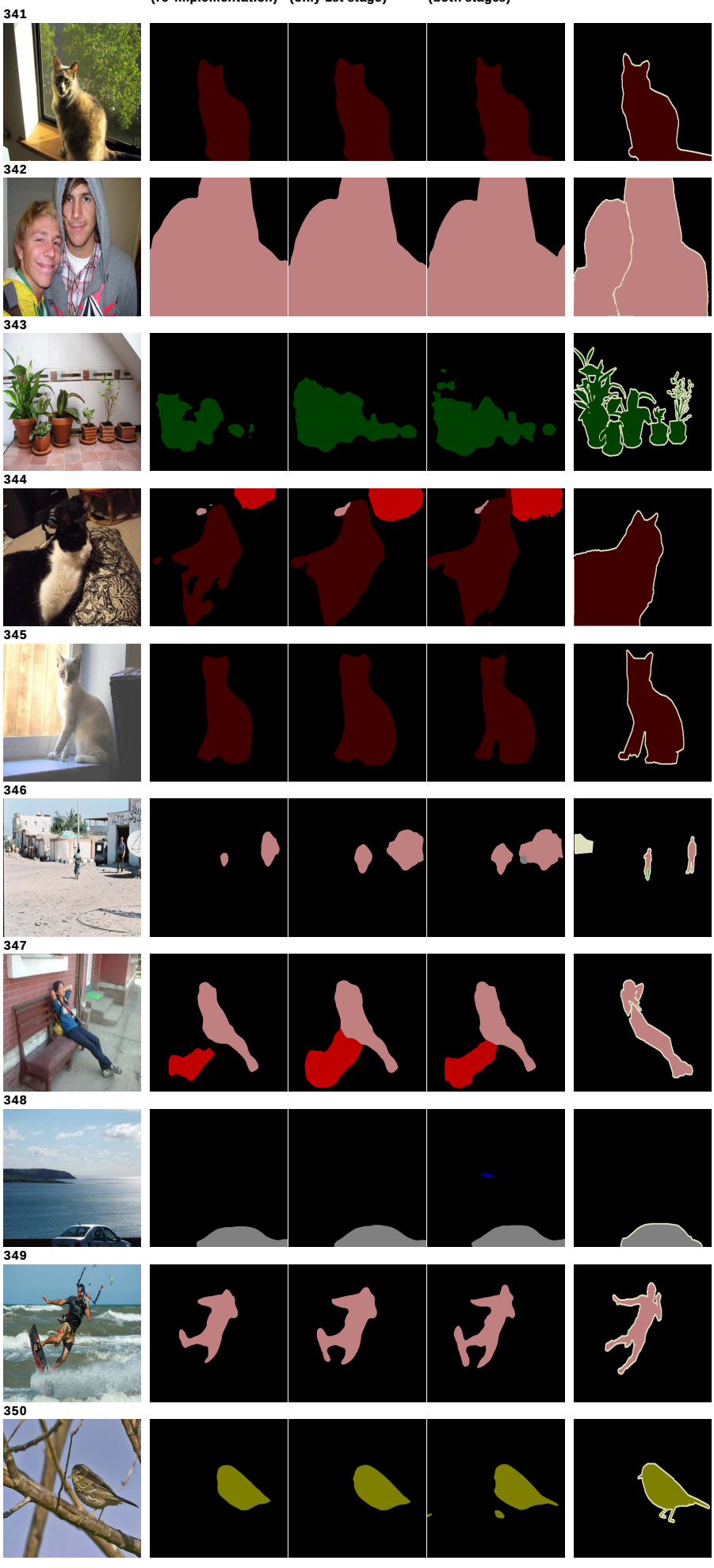

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
291
292
293
294
295
296
297
298
299
300

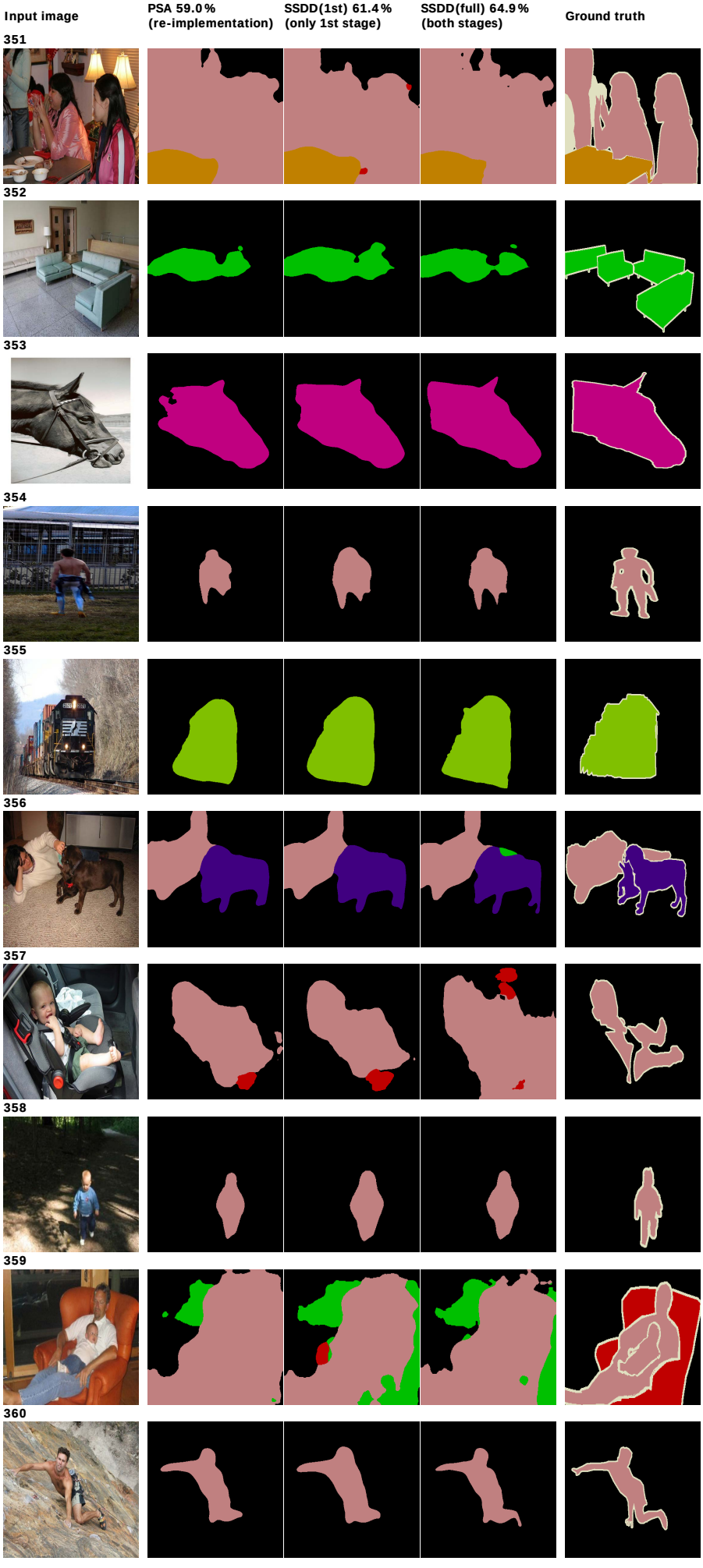

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
301
302
303
304
305
306
307
308
309
310

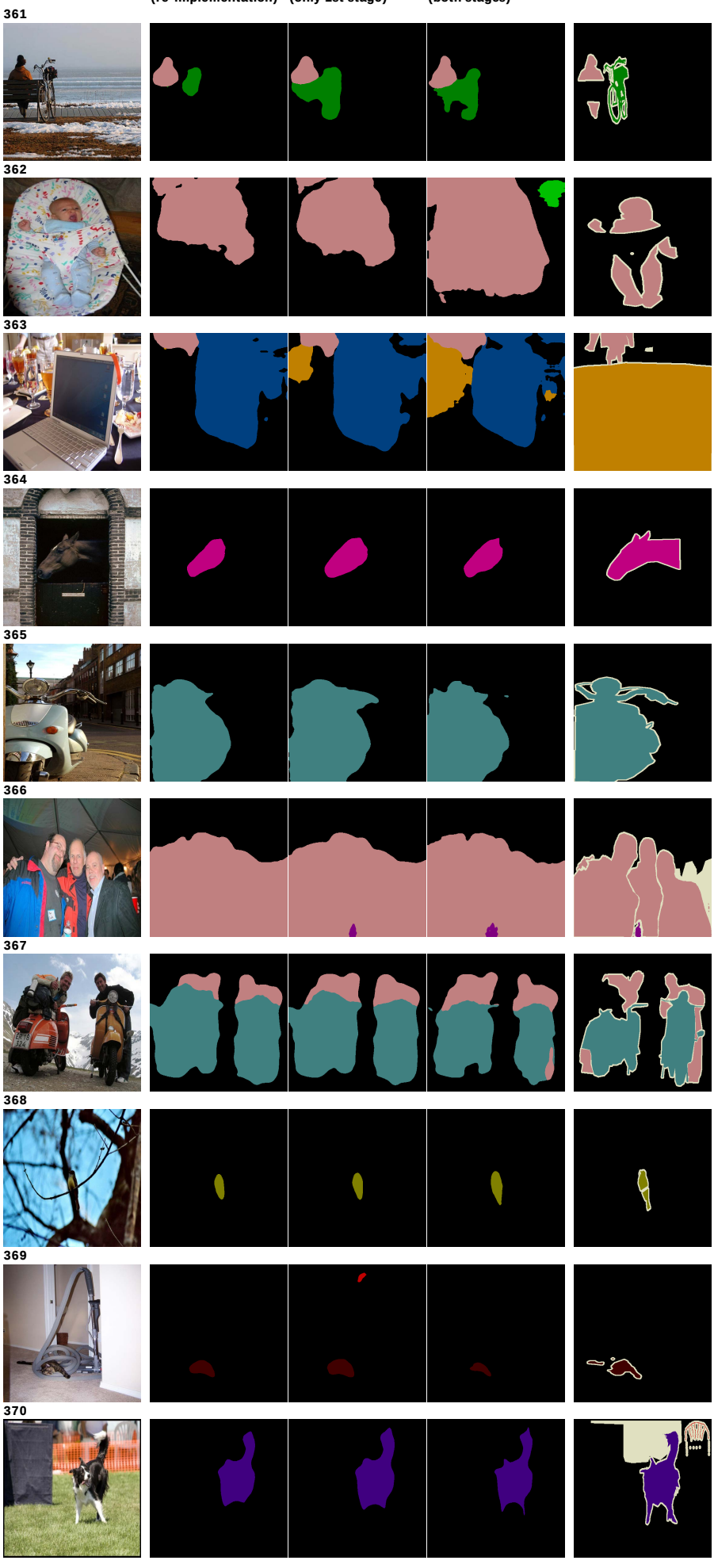

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
311
312
313
314
315
316
317
318
319
320

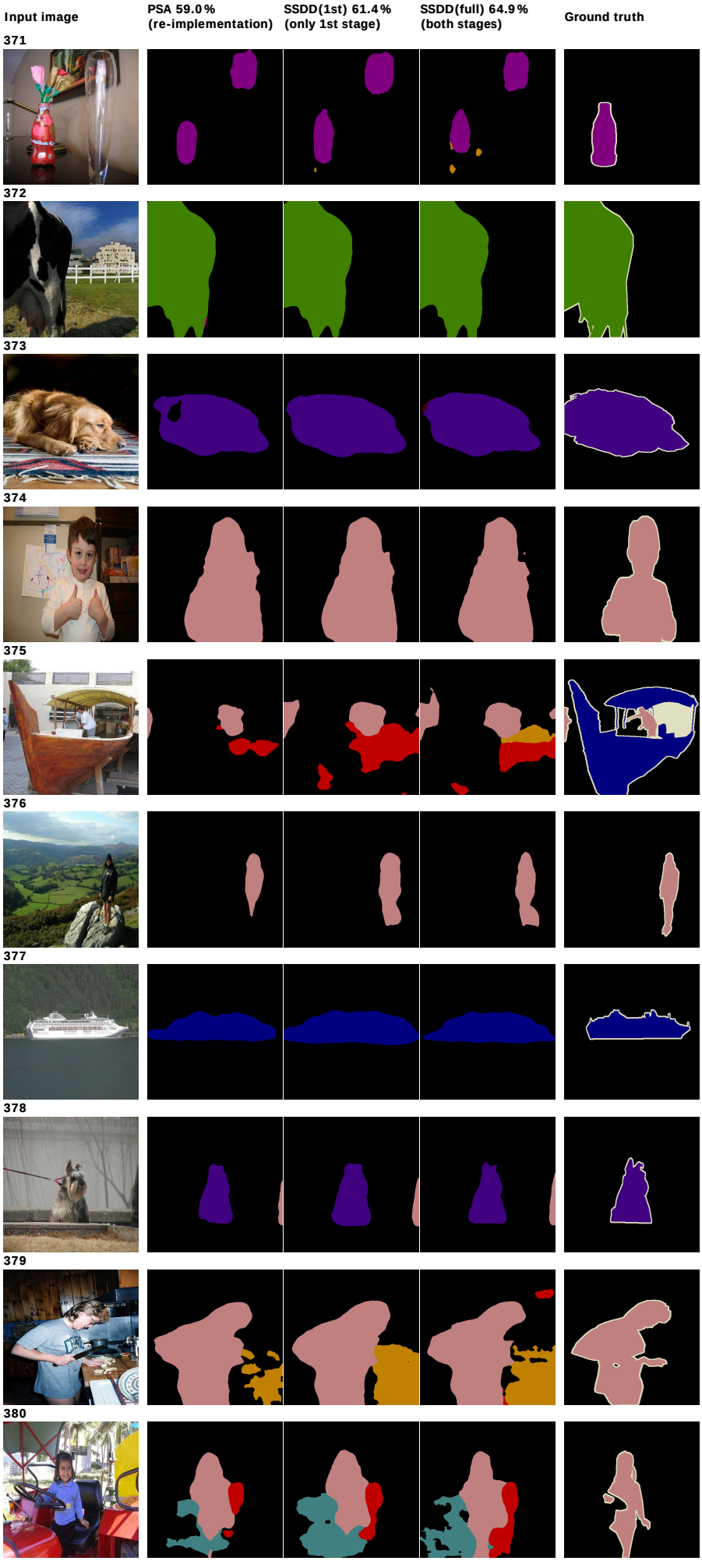

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
321
322
323
324
325
326
327
328
329
330

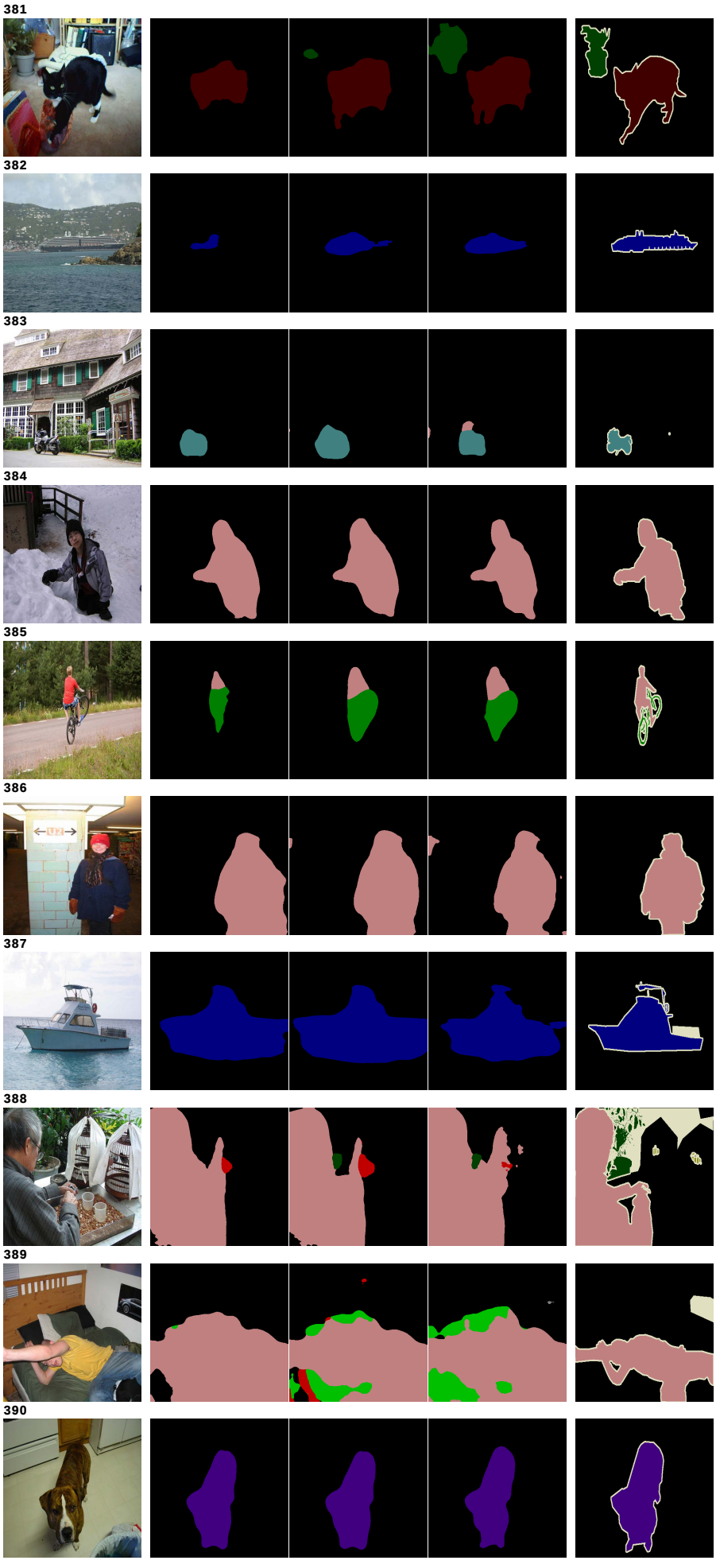

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
331
332
333
334
335
336
337
338
339
340

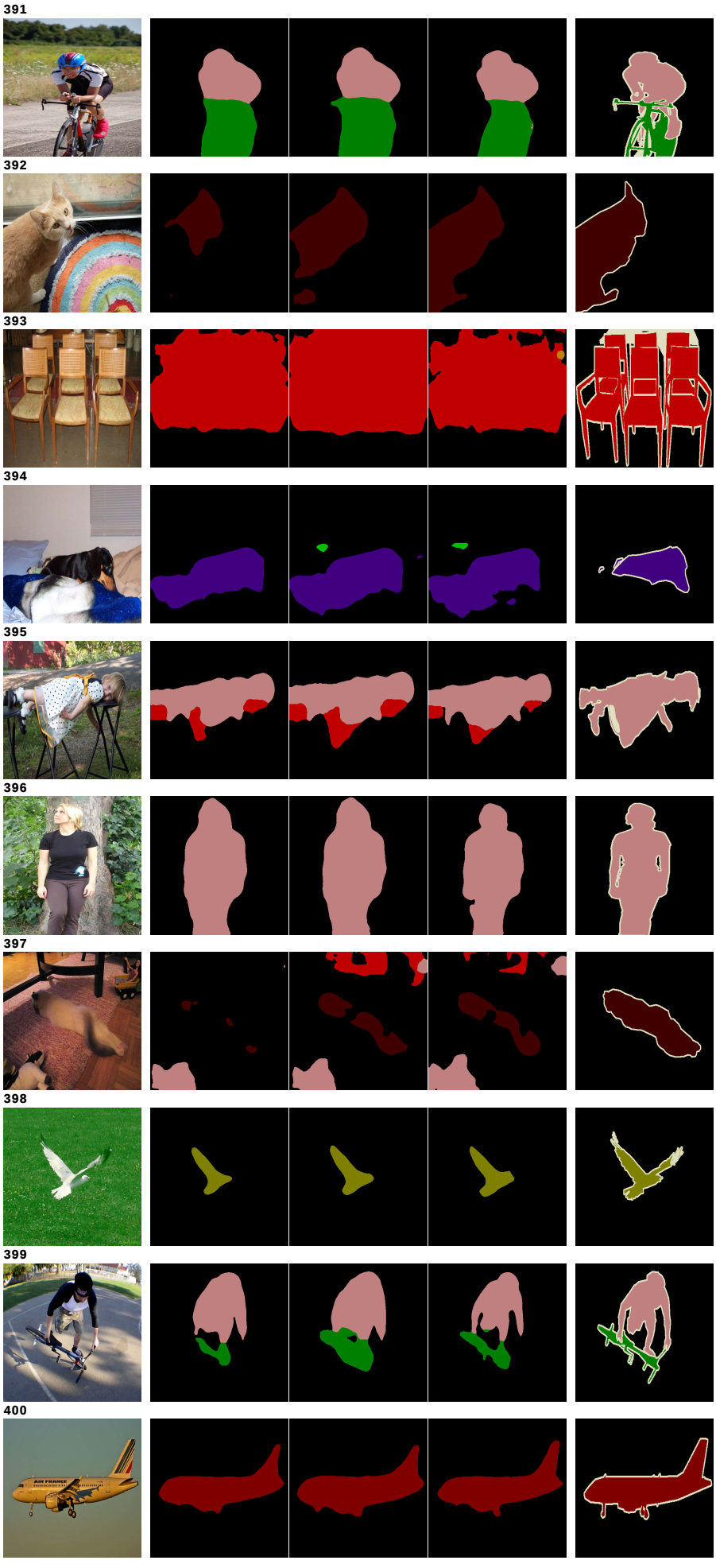

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
341
342
343
344
345
346
347
348
349
350

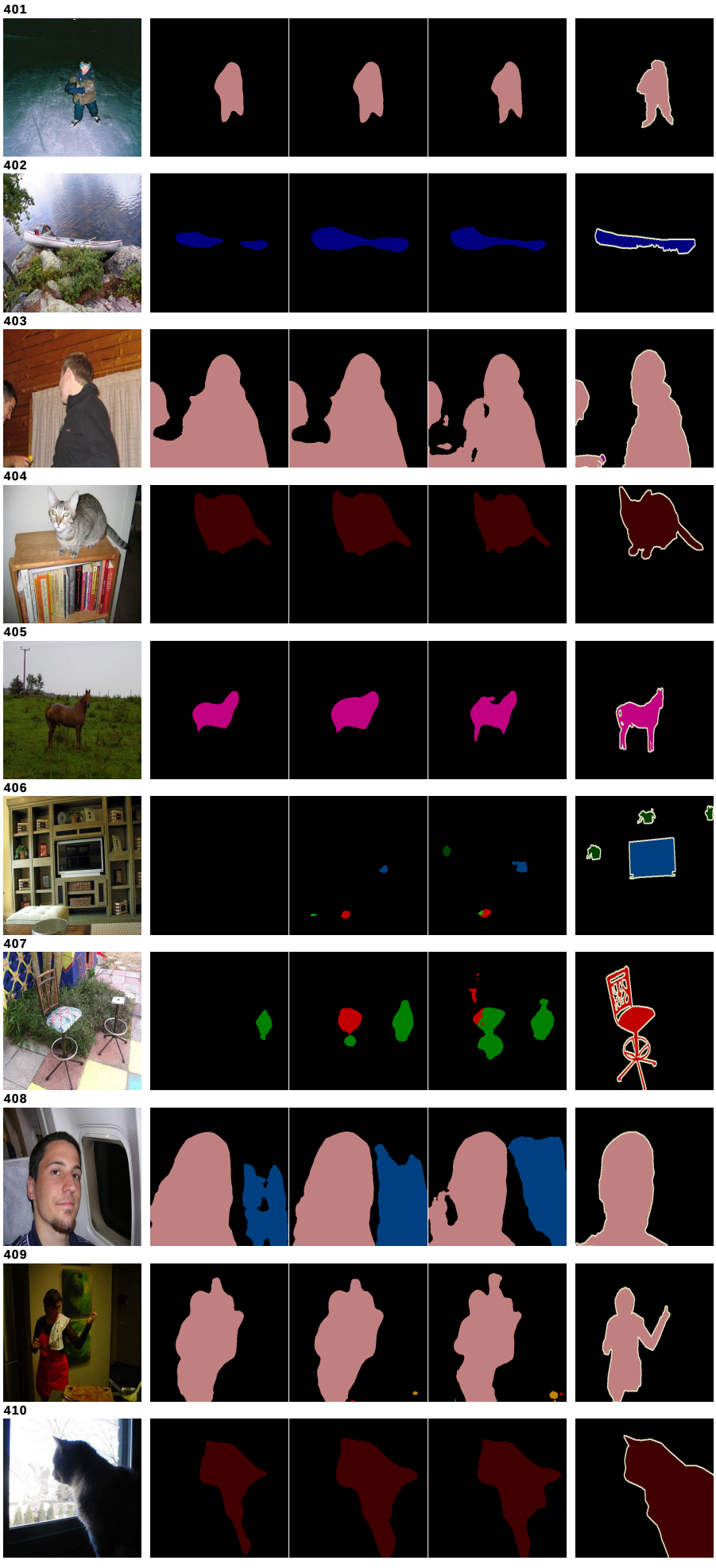

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
351
352
353
354
355
356
357
358
359
360

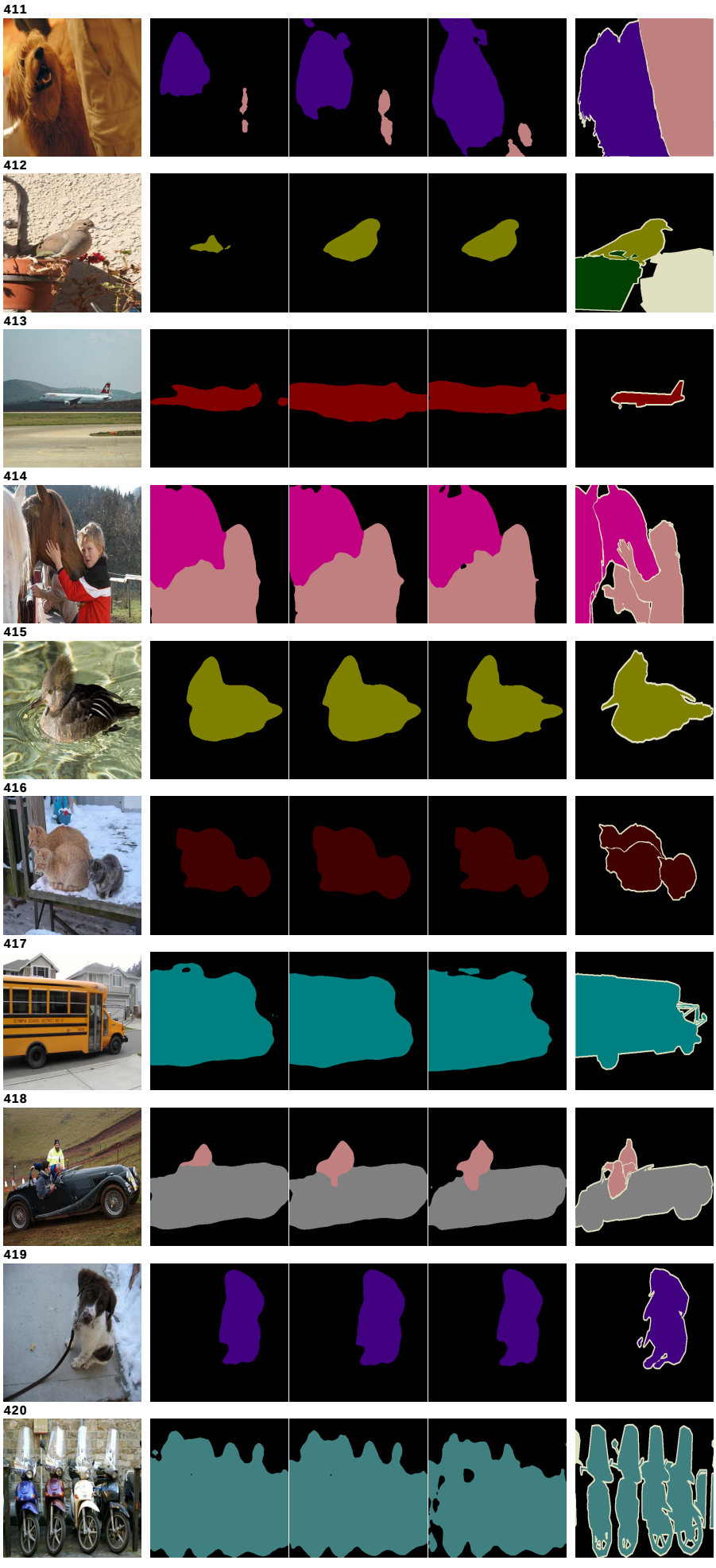

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
361
362
363
364
365
366
367
368
369
370

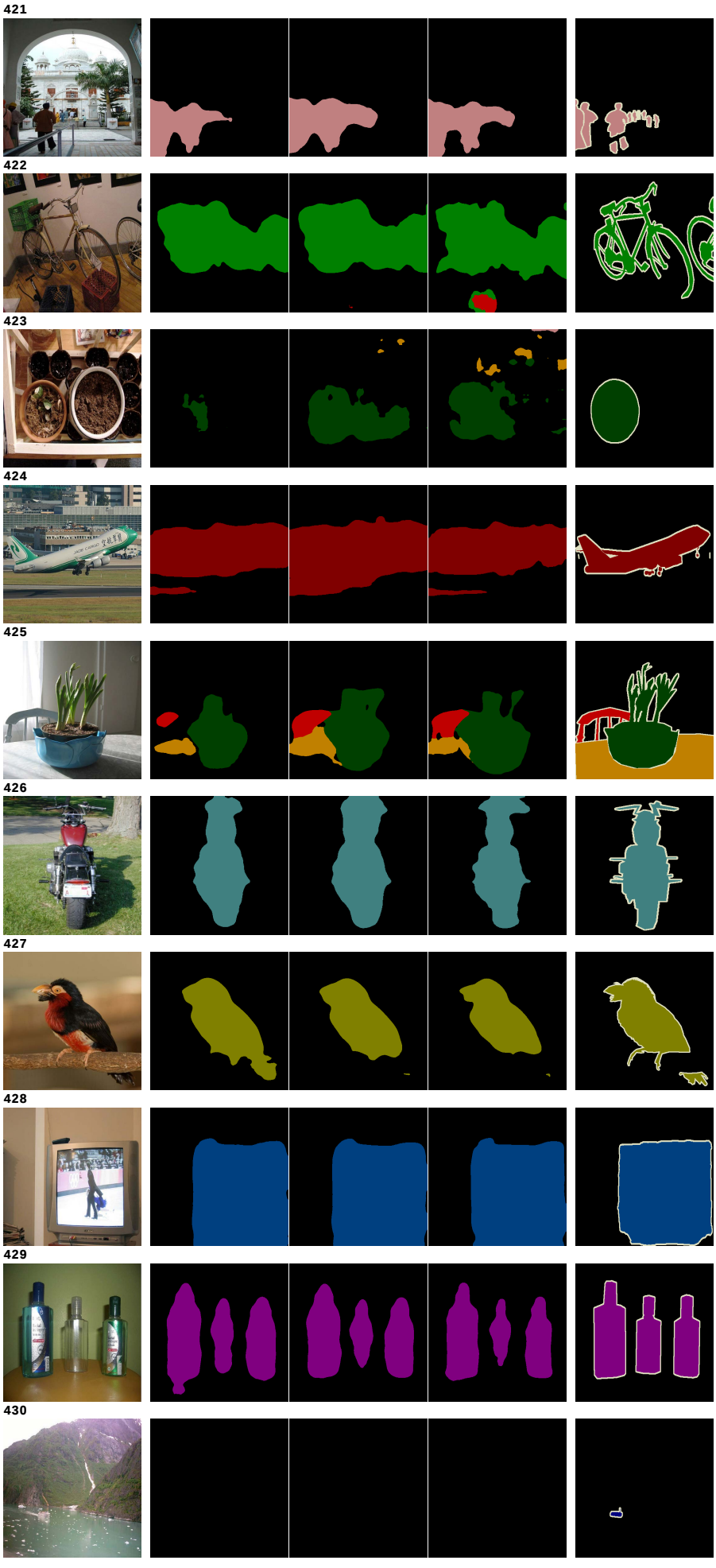

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
371
372
373
374
375
376
377
378
379
380

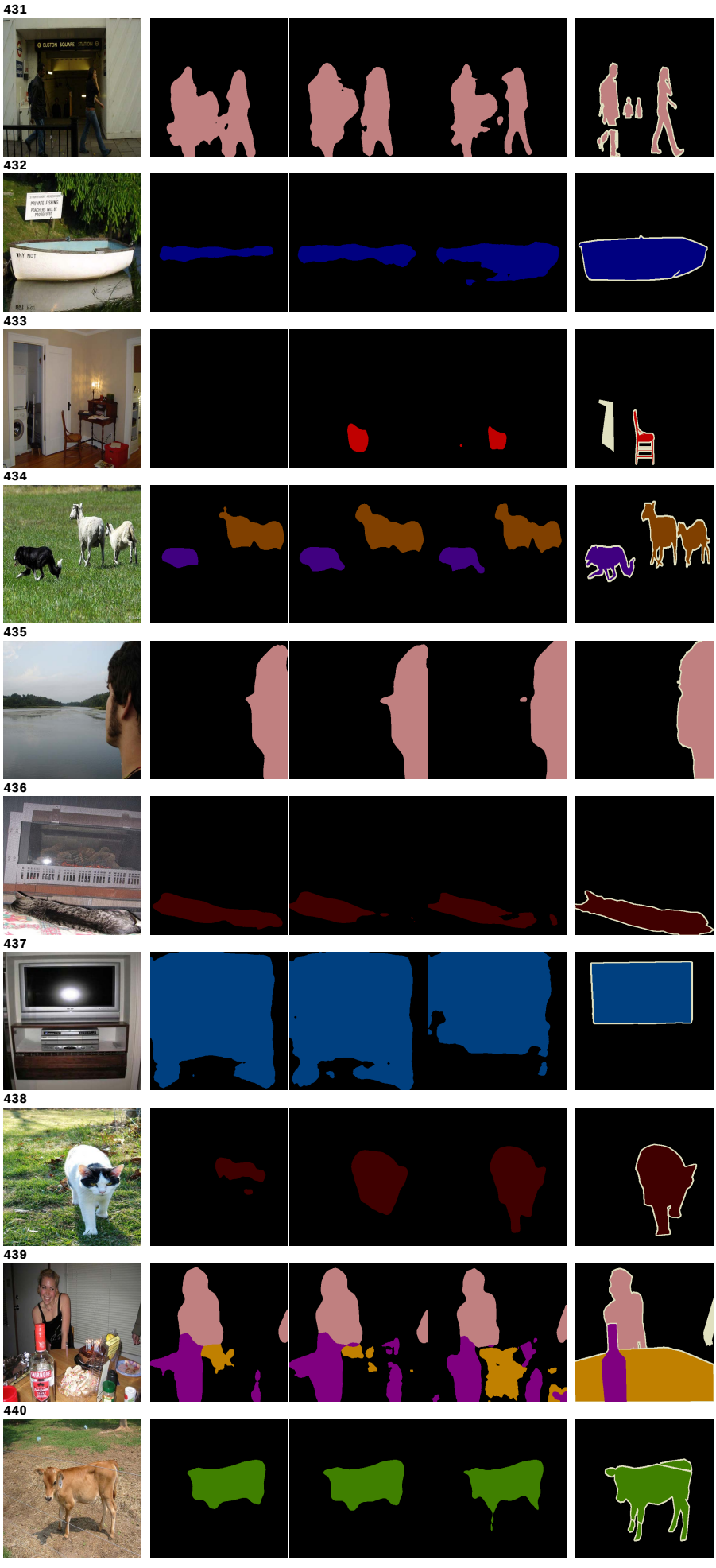

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
381
382
383
384
385
386
387
388
389
390

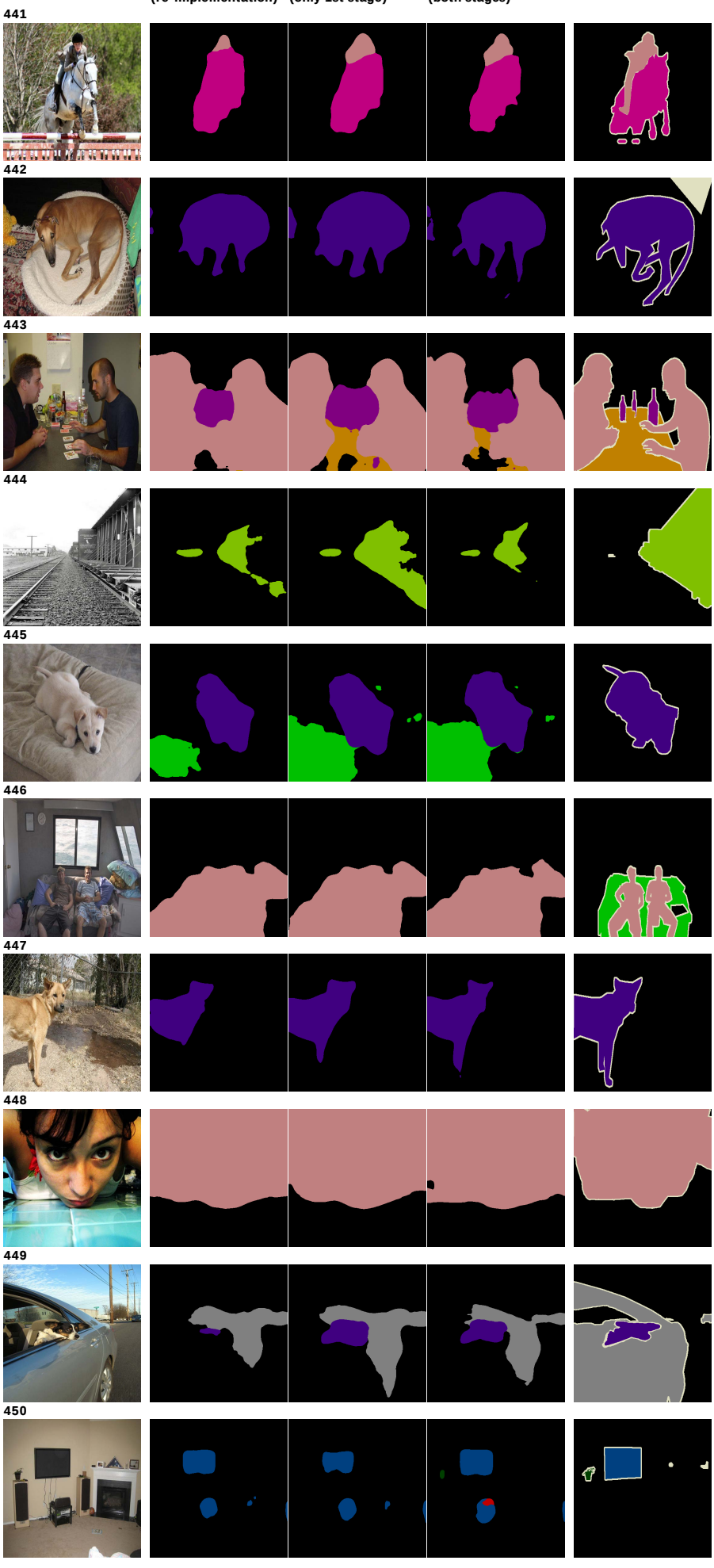

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
391
392
393
394
395
396
397
398
399
400

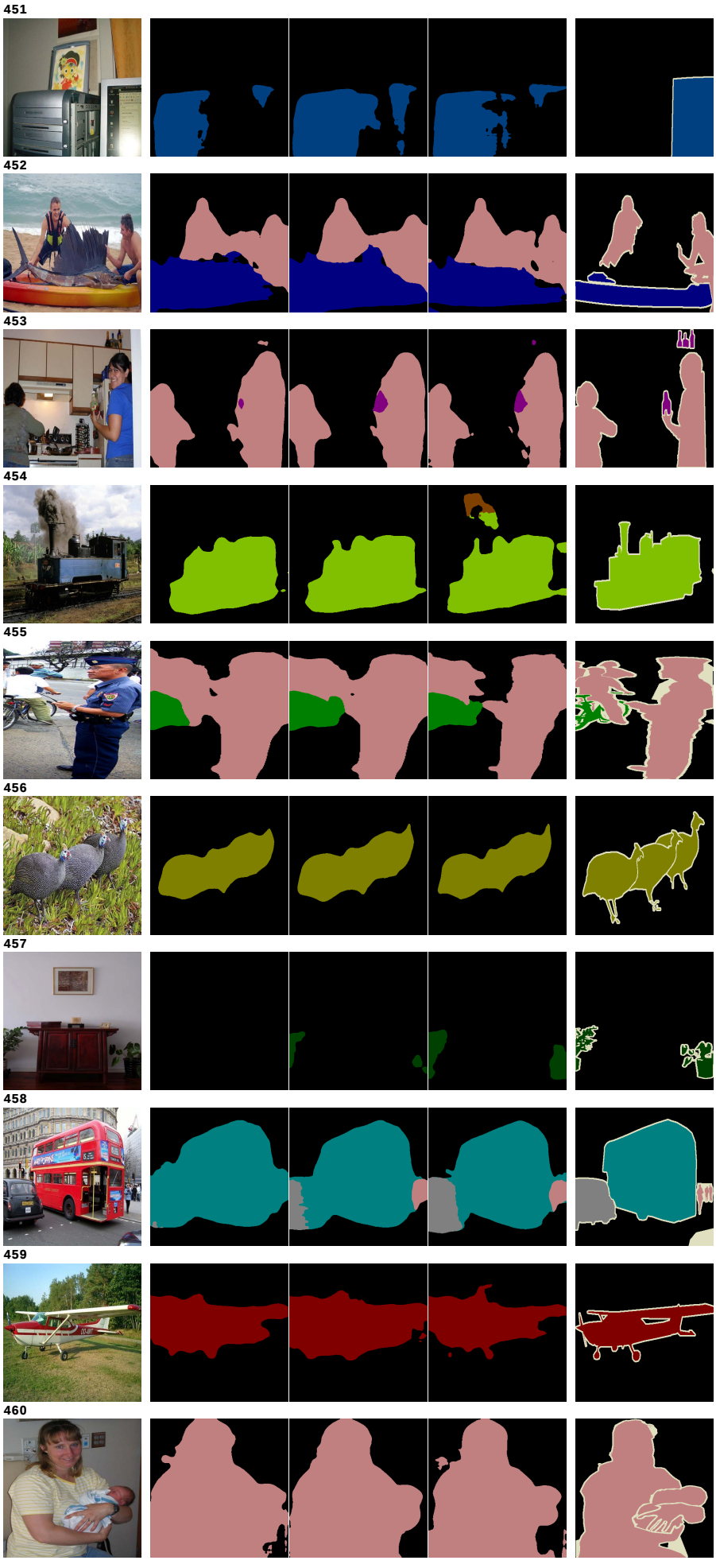

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
401
402
403
404
405
406
407
408
409
410


Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
411
412
413
414
415
416
417
418
419
420

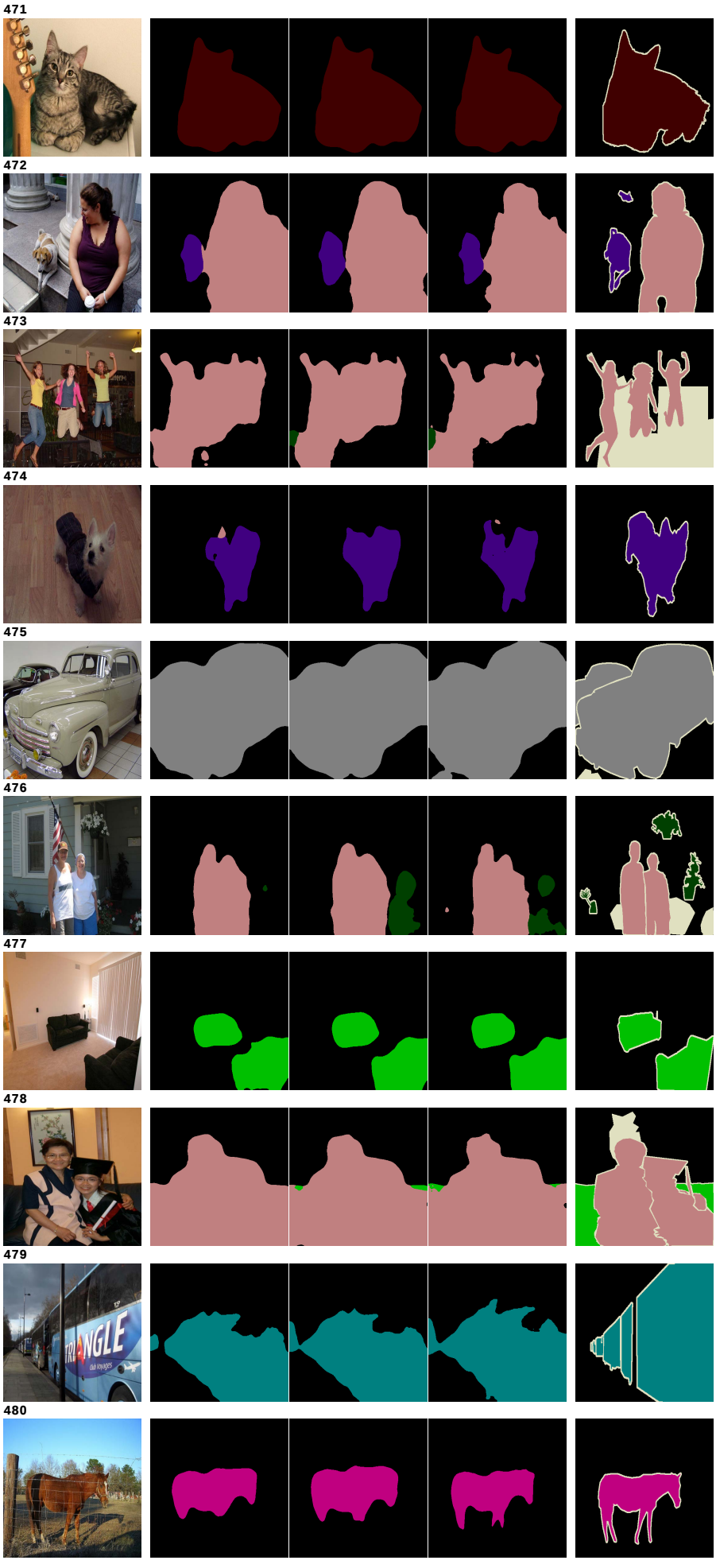

Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
421
422
423
424
425
426
427
428
429
430

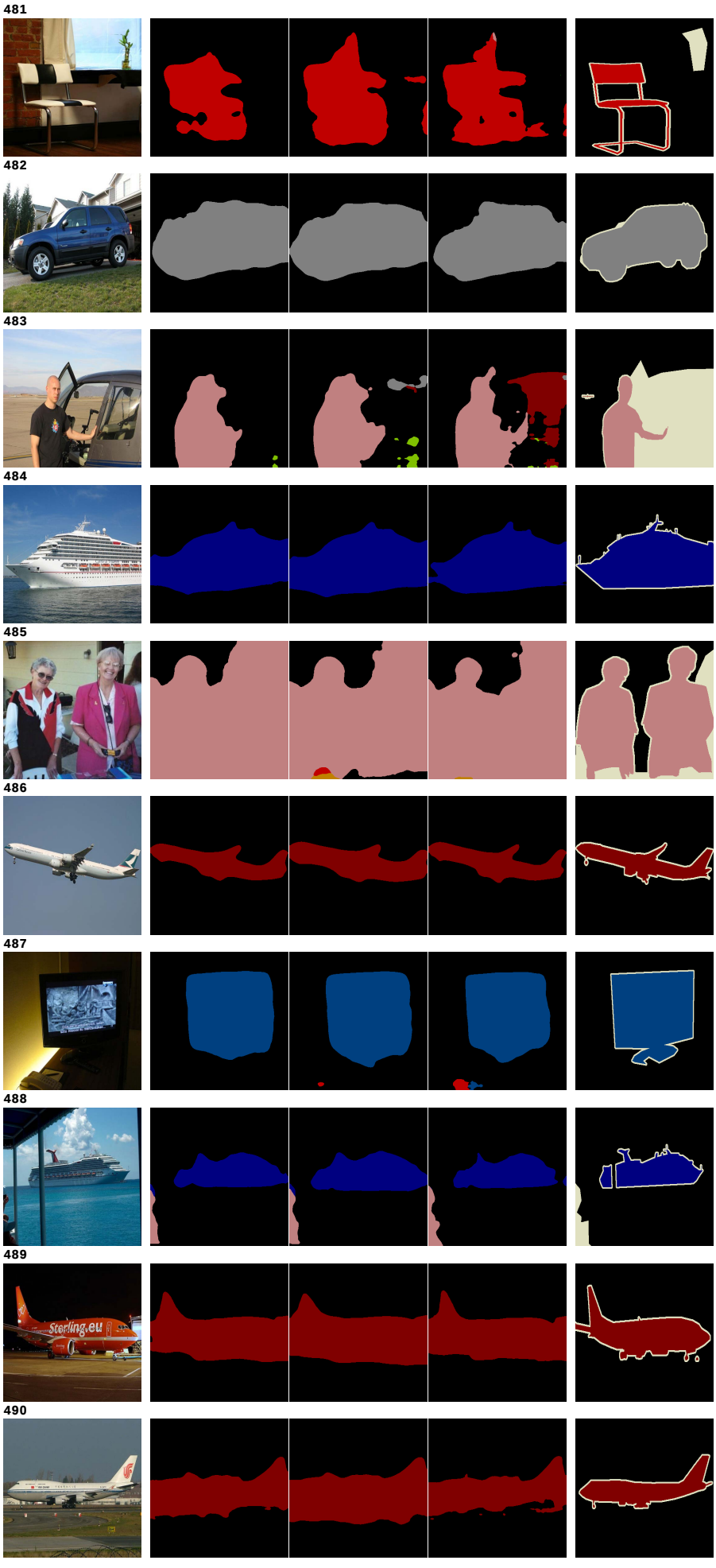

Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
431
432
433
434
435
436
437
438
439
440

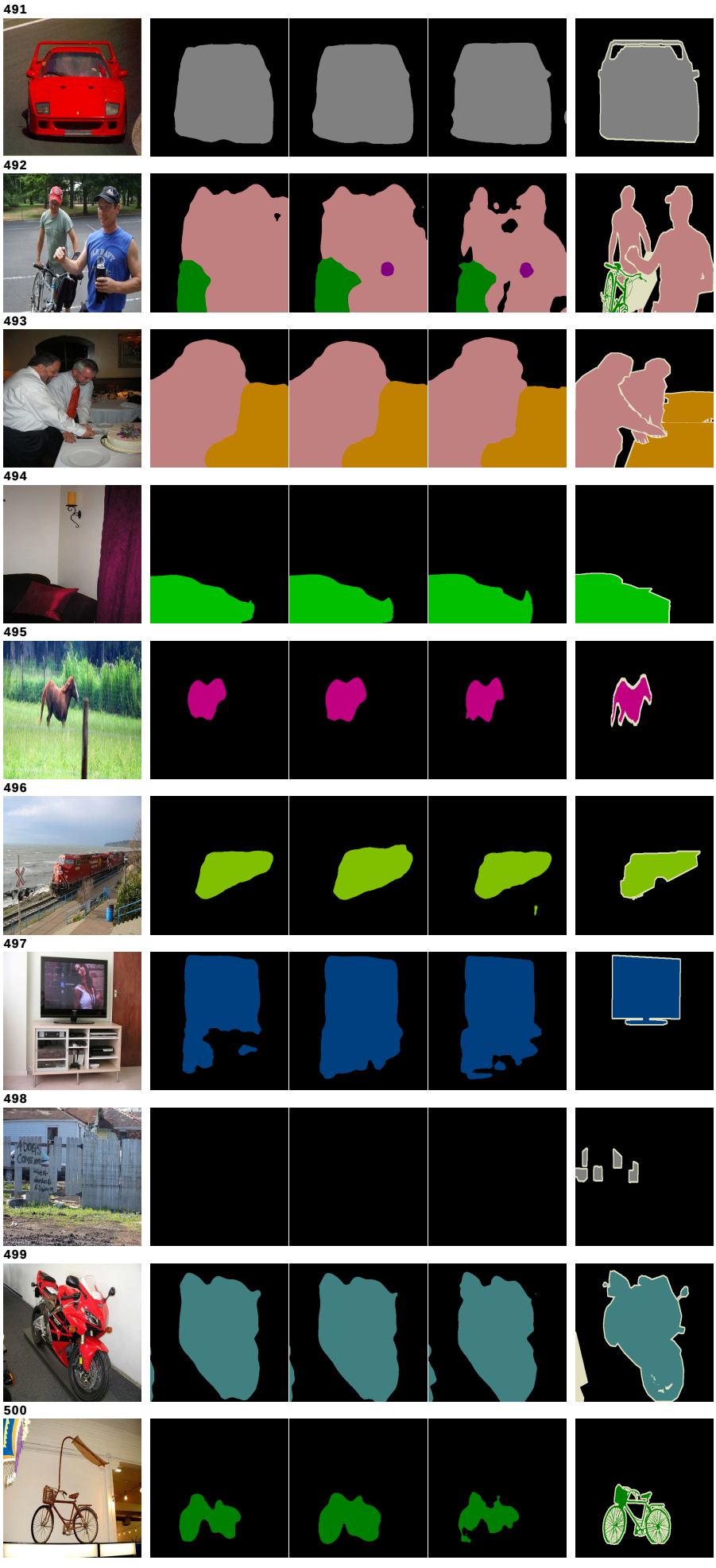

Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
441
442
443
444
445
446
447
448
449
450


Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
451
452
453
454
455
456
457
458
459
460


Input image
PSA 59.0 %
(re-implementation)
SSDD(1st) 61.4 %
(only 1st stage)
SSDD(full) 64.9 %
(both stages)
Ground truth
461
462
463
464
465
466
467
468
469
470


Input image
PSA 59.0% (re-implementation)
SSDD(1st) 61.4% (only 1st stage)
SSDD(full) 64.9% (both stages)
Ground truth
471
472
473
474
475
476
477
478
479
480


Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
481
482
483
484
485
486
487
488
489
490


Input image
PSA 59.0%
(re-implementation)
SSDD(1st) 61.4%
(only 1st stage)
SSDD(full) 64.9%
(both stages)
Ground truth
491
492
493
494
495
496
497
498
499
500