

Image Inpainting usingg Block-wise Procedural Training with Annealed Adversarial Counterpart

Supplementary Material

Paper ID 1283

1 More Experiments

1.1 Comparing Patch Perceptual Losses with ℓ_2

We compare the quality of inpainting that is trained with Patch Perceptual Loss and ℓ_2 loss. ℓ_2 is used to train CE and GLI, but it does not correspond well to human perception of similarity. We investigate how our Patch Perceptual Loss compares with ℓ_2 loss by training under the same condition but using different reconstruction losses (PPL vs ℓ_2). From the test cases, we see that our results are overwhelmingly better than ℓ_2 results in terms of sharpness and coherence with context. Fig. 1 shows two examples of comparison.



Fig. 1. Effects of different types of reconstruction losses. Zoom in for best quality.

1.2 Data Acquisition for Jointly Training an Inpainting and Harmonization Network

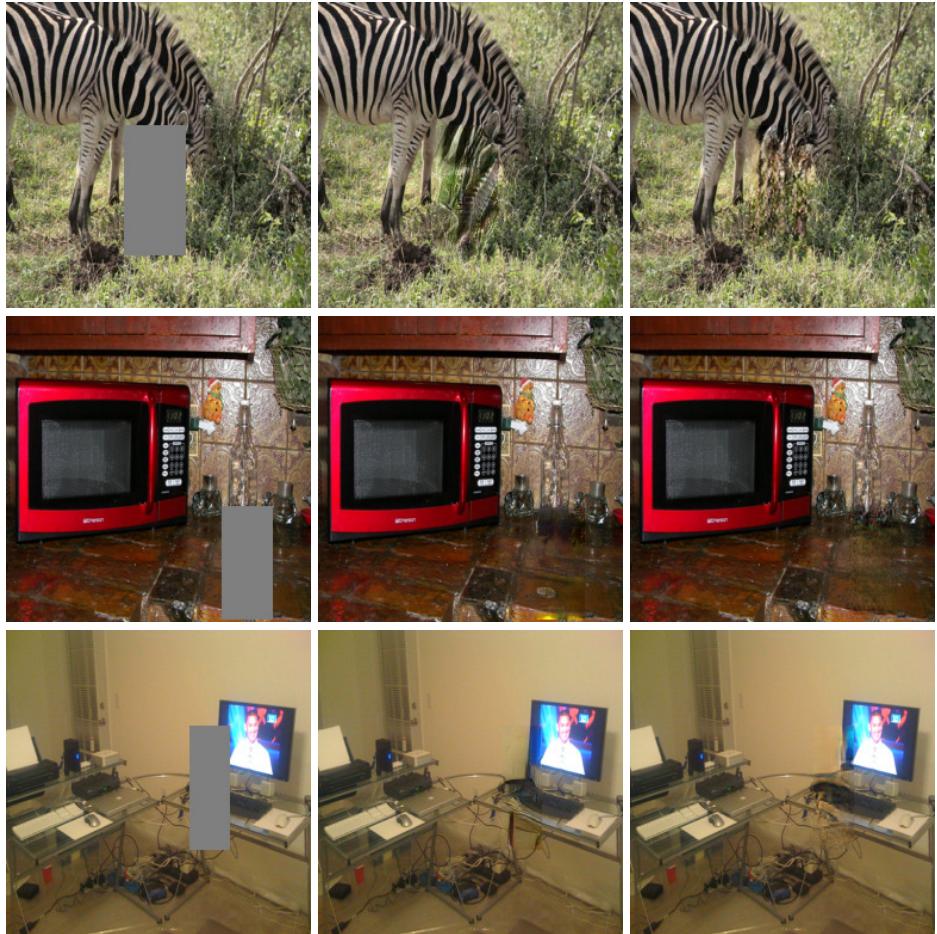
To accomplish guided inpainting, we jointly train our inpainting and harmonization using the following data acquisition procedure: at each iteration given the input image I , we randomly select another image I_g from the dataset. We then select and segment an object I_o from I based on the segmentation mask of COCO. We then transfer the color from I_g to I_o , and paste I_o onto I_g at a random location. Finally, we crop a bounding box I_b from I_g that contains I_o , and paste I_b back to I . The result, together with the foreground/background segmentation mask of I_b , is given to the network as input. The model is modified to output two images, one for inpainting result and the other for harmonization result. We use the original I as the ground truth for both tasks.

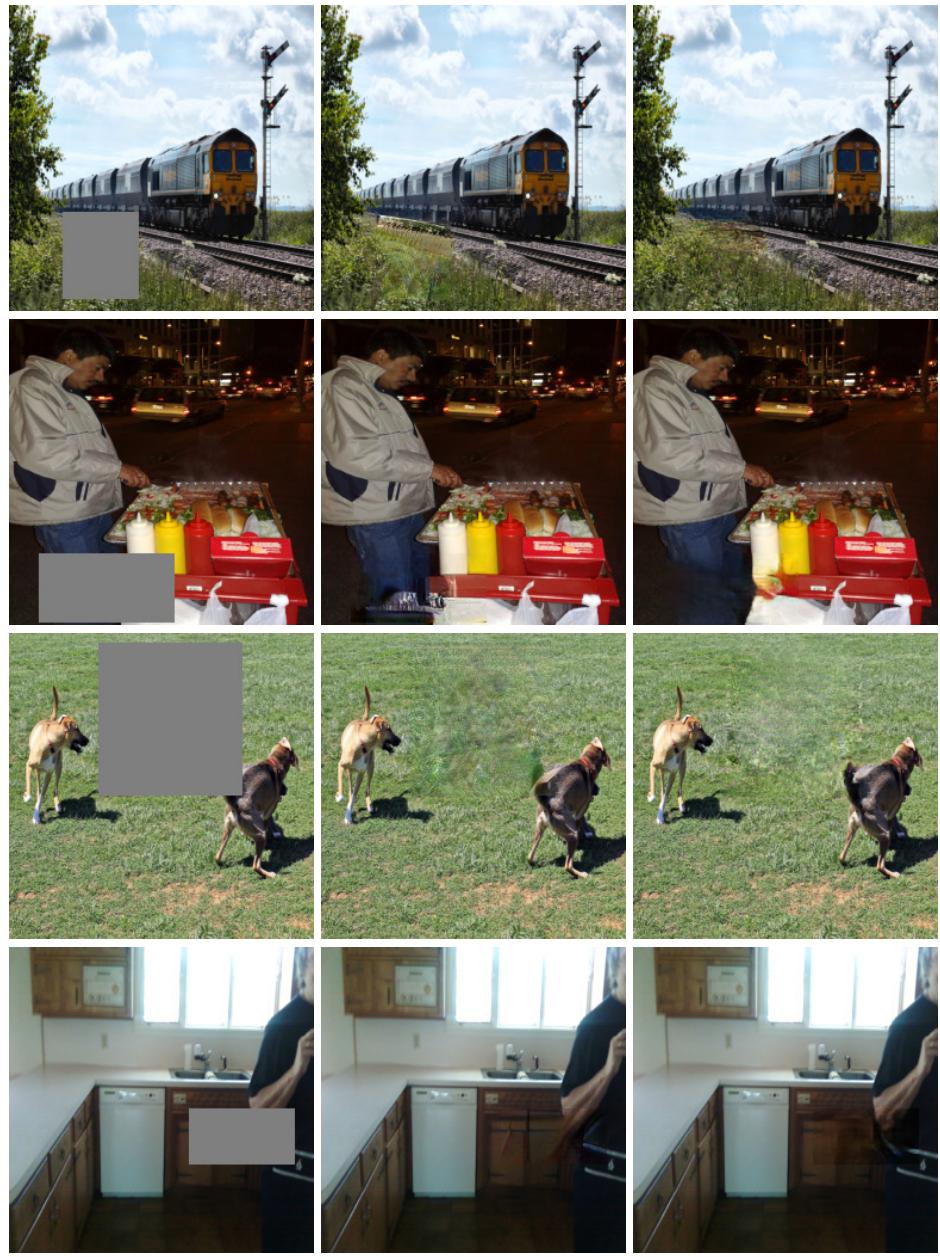
2 More Results

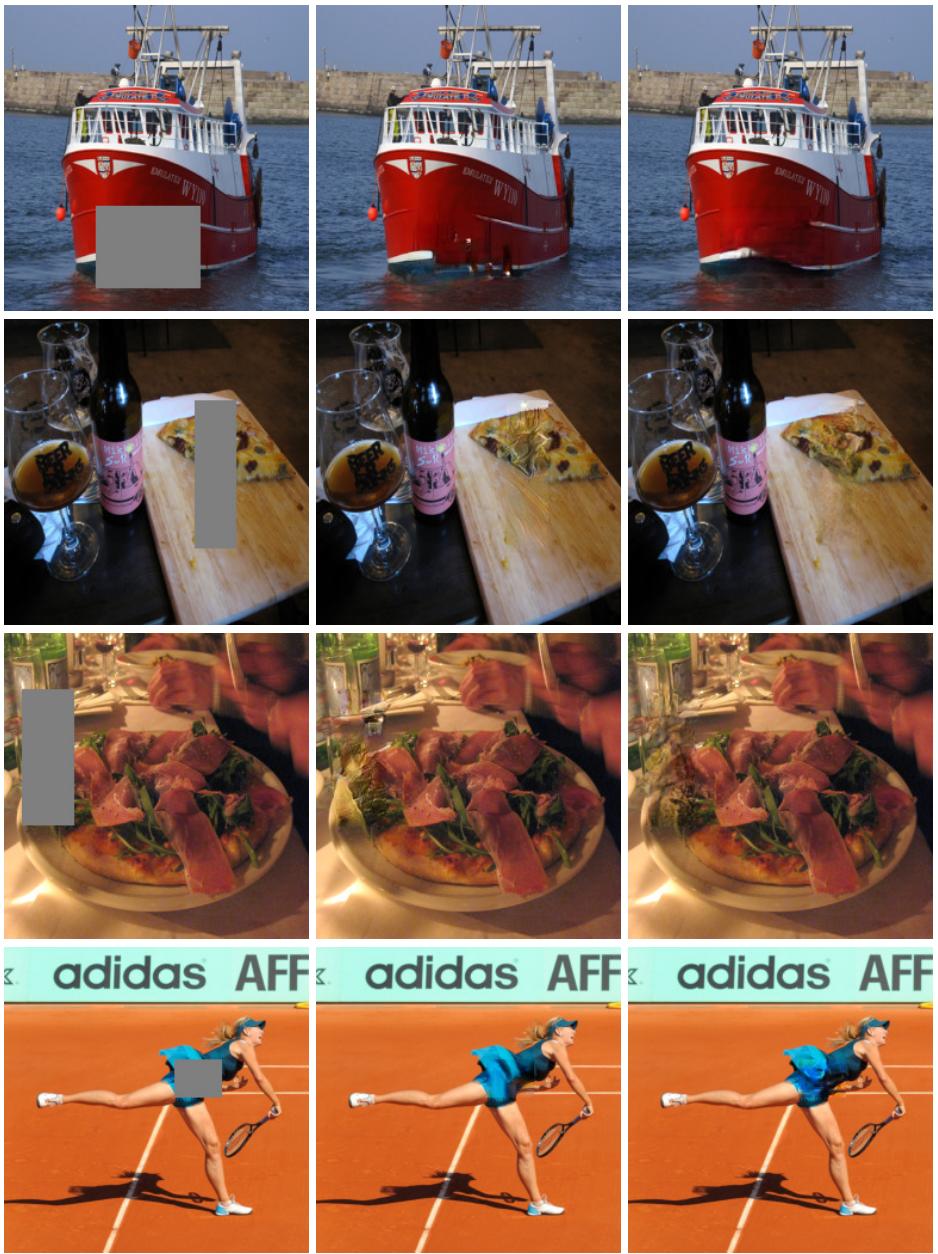
We show more results on 256x256 images with random holes and compare with GLI. We also show more 512x512 results with random holes. We then show more face results. Finally, we show more harmoinzation (and compare with deep image harmonization) and interactive guided inpainting result. All results are randomly sampled from test set.

2.1 More 256x256 random hole results and comparing with GLI

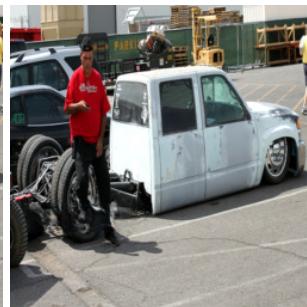
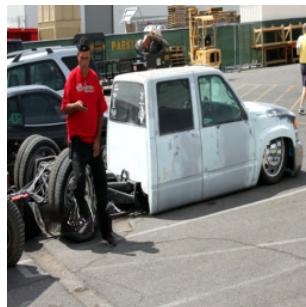
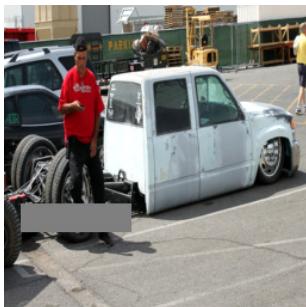
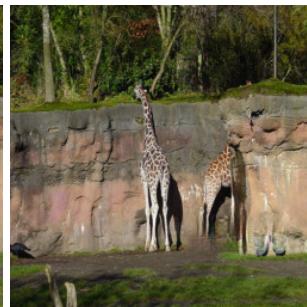
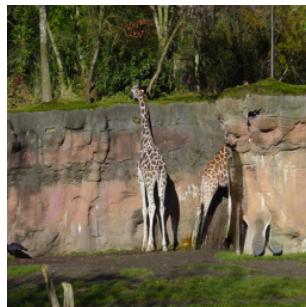
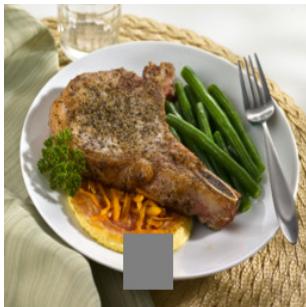
From left to right: input, GLI result, our result.

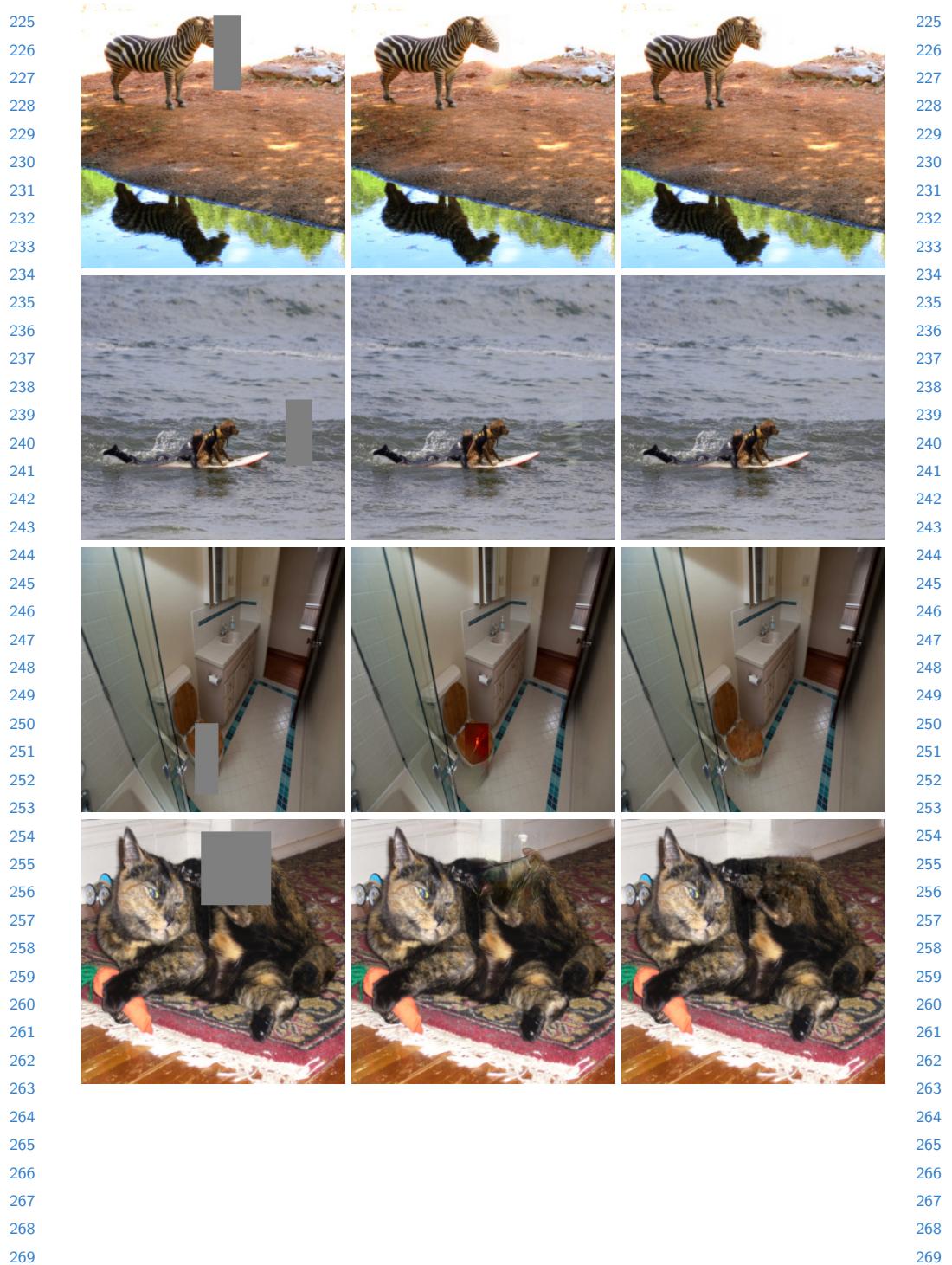


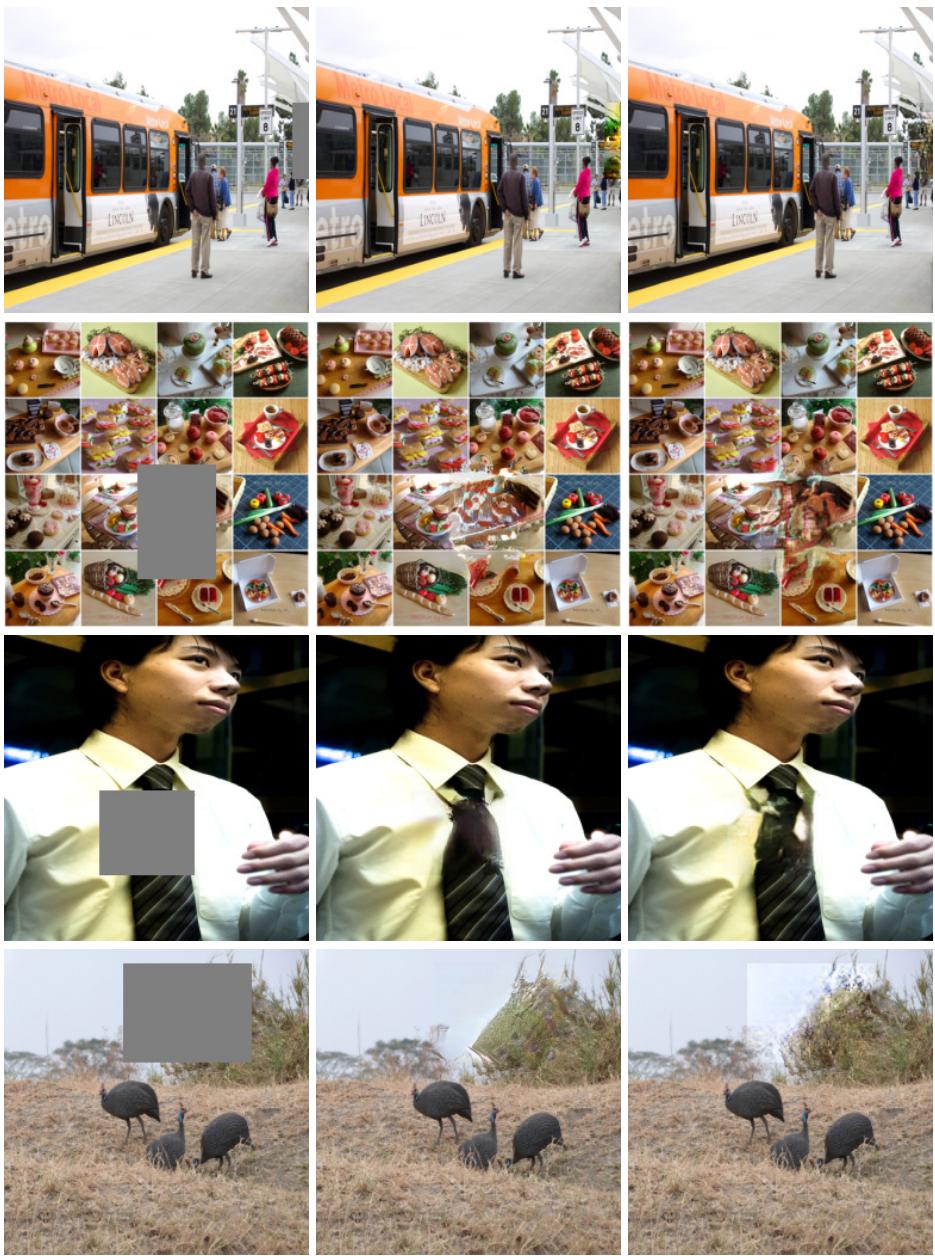




135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179

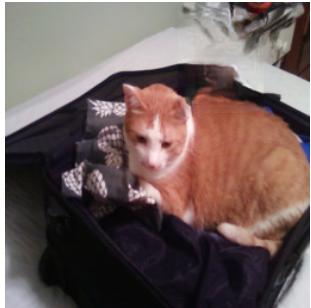
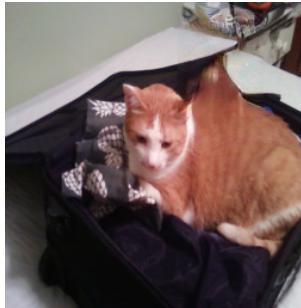
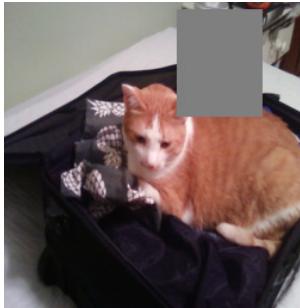








315
316
317
318
319
320
321
322
323



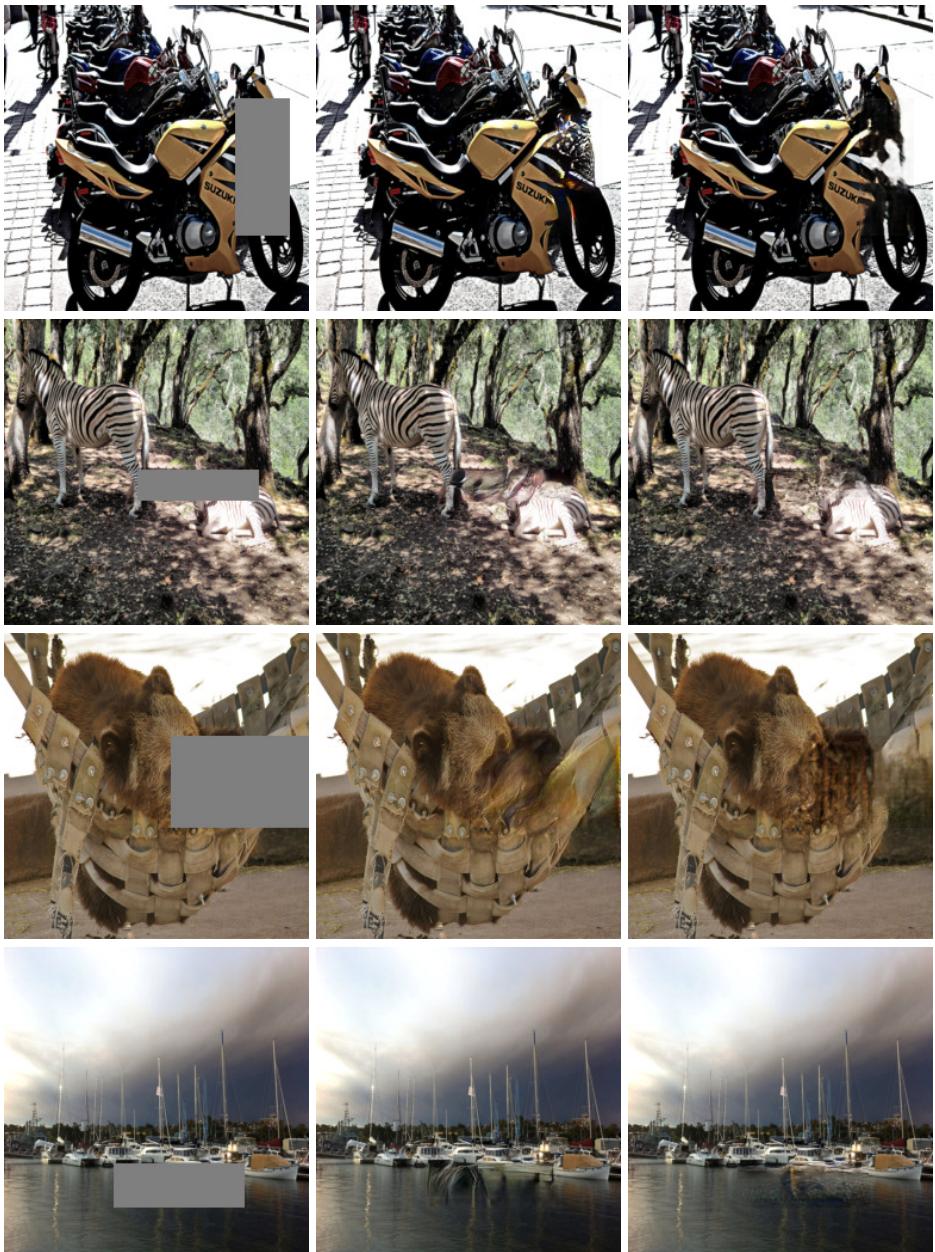
334
335
336
337
338
339
340
341
342
343



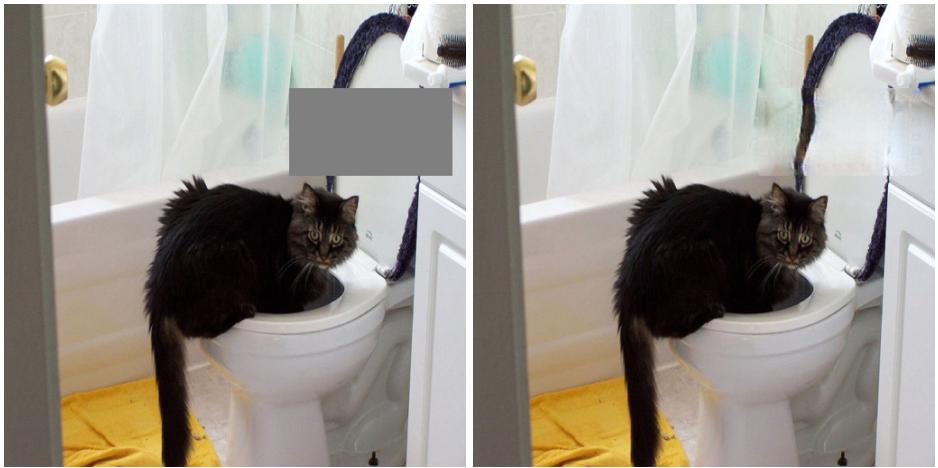
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359



354
355
356
357
358
359



2.2 More 512x512 random hole results



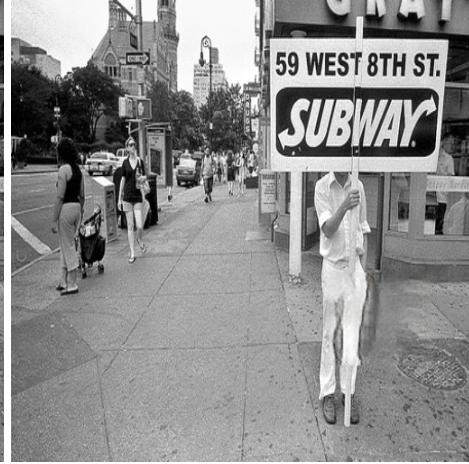
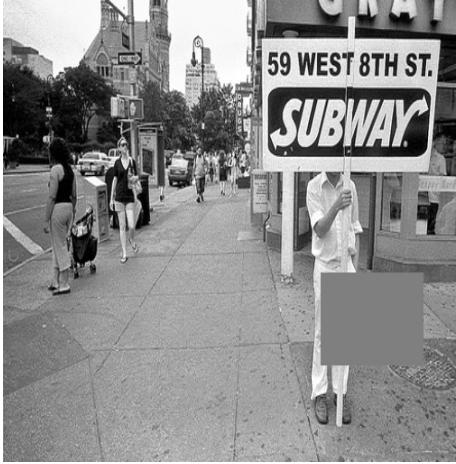
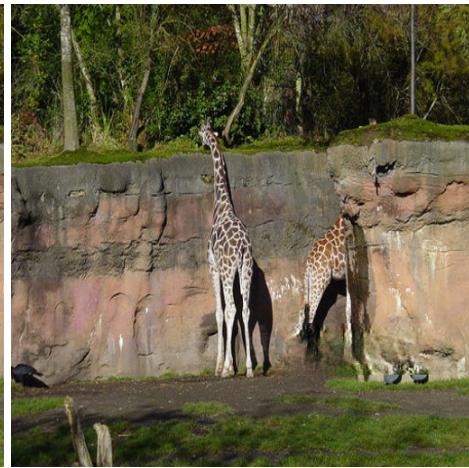
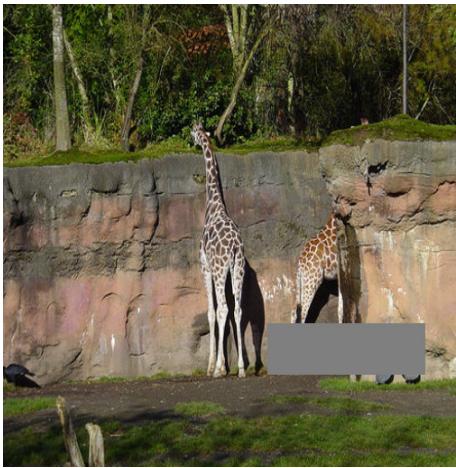
405
406
407
408
409
410
411
412
413
414
415
416
417
418



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



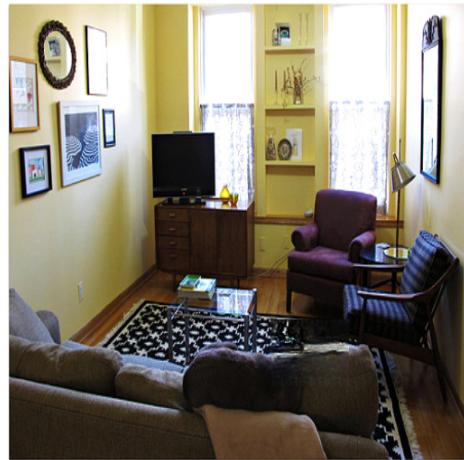
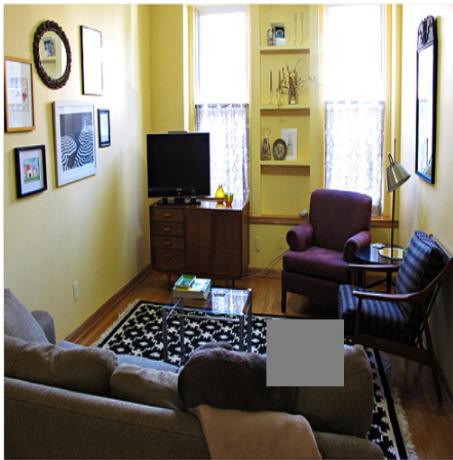
495
496
497
498
499
500
501
502
503
504
505
506
507
508



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
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







675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703



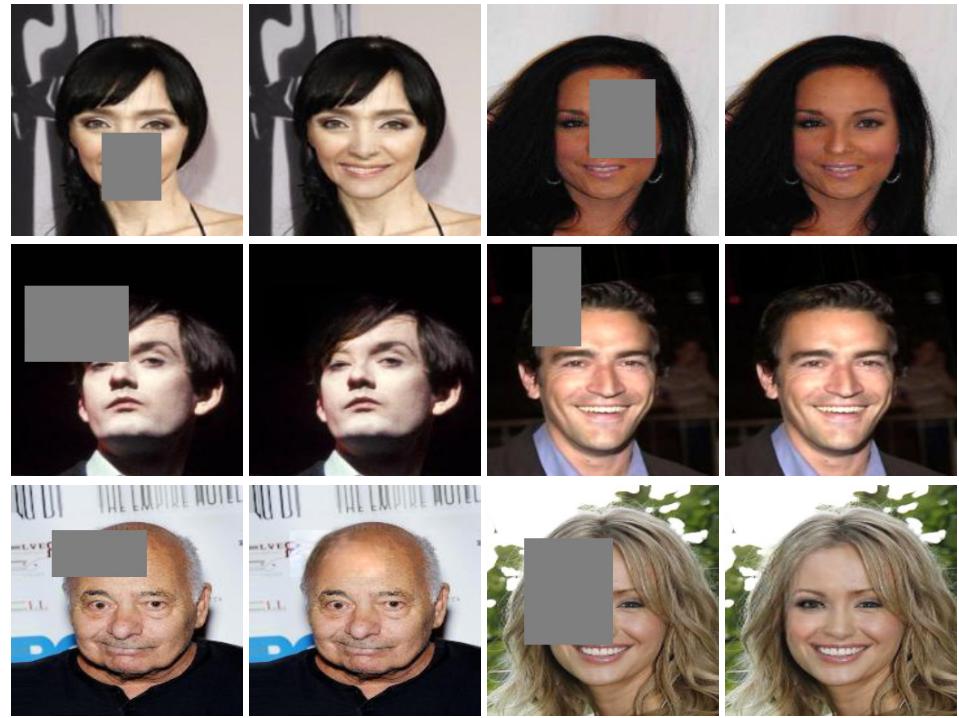
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719

2.3 More face results



705
706
707
708
709
710
711
712
713
714
715
716
717
718
719



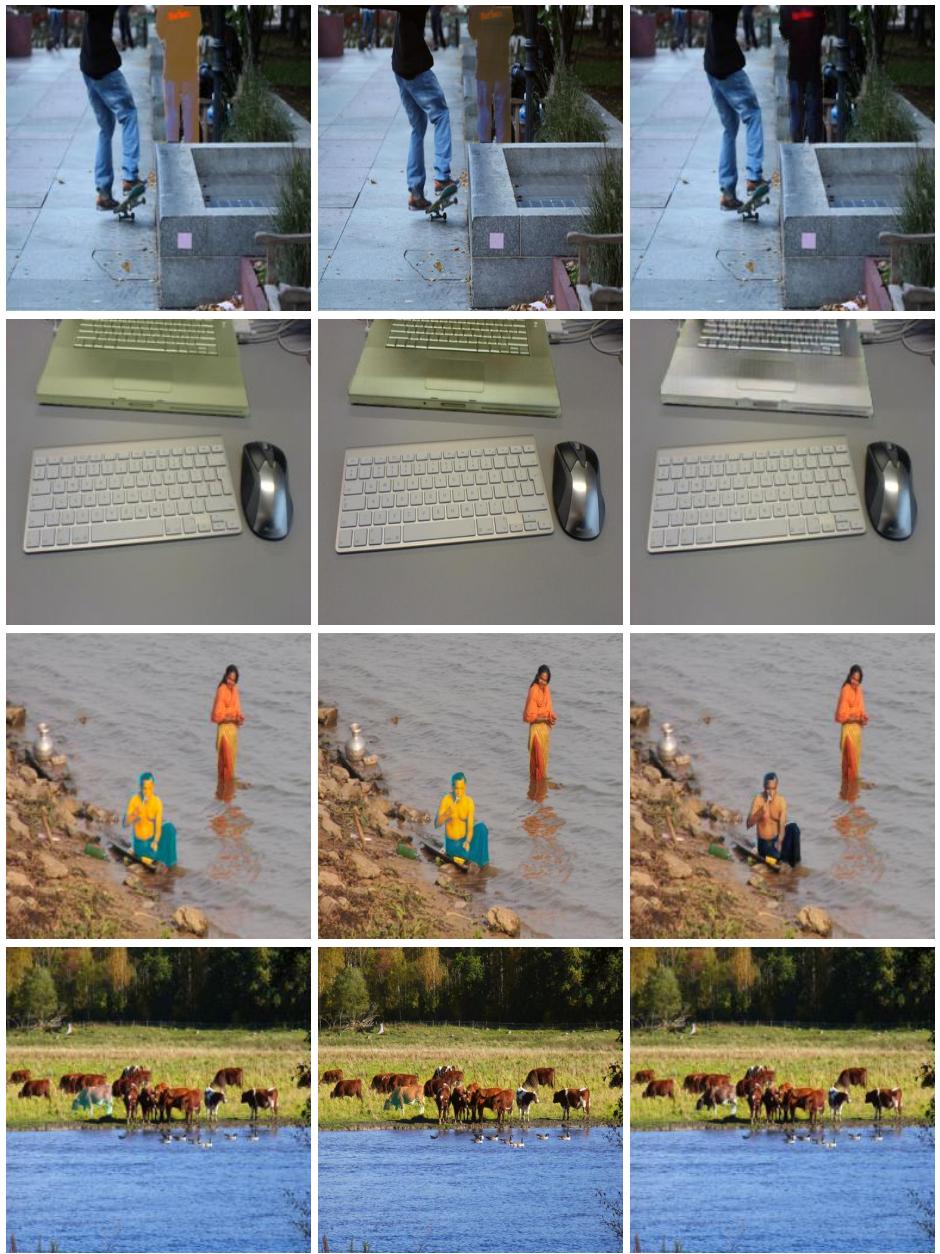


2.4 More Harmonization Results

From left to right: input (where the color of one object is altered), deep image harmonization result, our result.









938
939
940
941
942
943
944

900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944



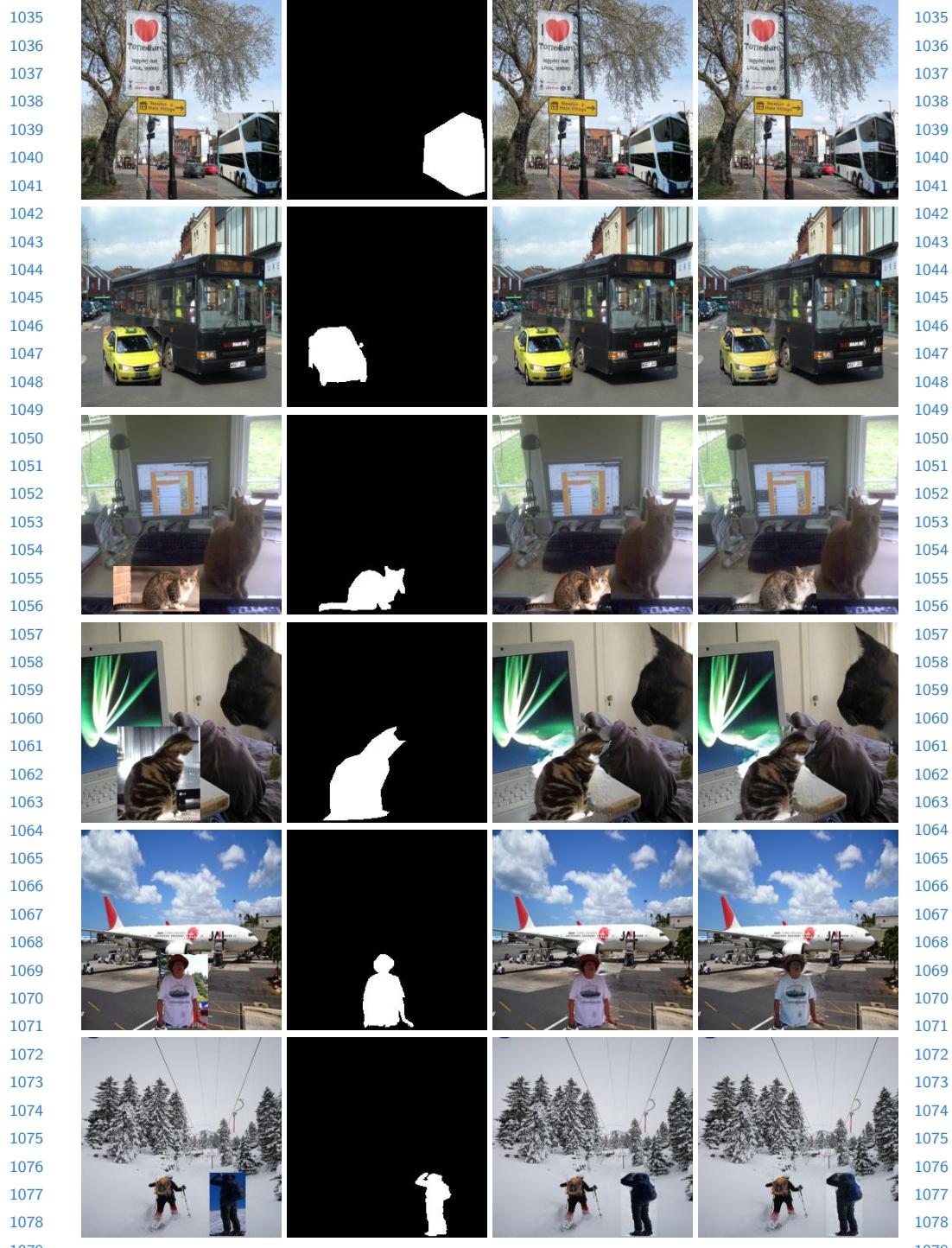
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989

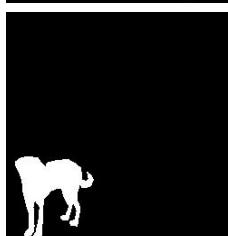
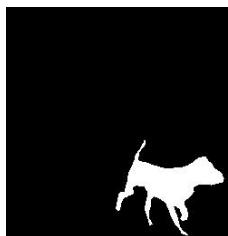
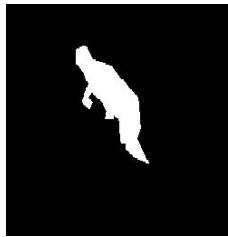


2.5 More Guided Inpainting Results

From left to right: input, segmentation mask, inpainting result, final result (with harmonization).







1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124