

Supplementary Materials for Deep Image Comparator: Learning to Visualize Editorial Change

Alex Black¹ Tu Bui¹ Hailin Jin² Vishy Swaminathan² John Collomosse^{1,2}
¹CVSSP, University of Surrey ²Adobe Research
{alex.black,t.v.bui}@surrey.ac.uk {vishy,hljjin,collomos}@adobe.com

1. Additional results

1.1. Near Duplicate Search

We provide retrieval examples for our methods and baselines in Fig. 1-2 (complementing Fig.3 in our main submission). While the baselines might perform well for clean manipulated queries (PSBat-Ret Manip set), the presence of in- and out-place transformations significantly drop its retrieval performance. Our 2-stage method of training a retrieval model in a contrastive metric learning fashion, followed by a reordering stage with a novel image alignment and comparison network (ICN), successfully retrieves the images of interest even under severe transformations subjected on the queries.

1.2. Heatmap Visualization and Classification

Fig. 3 shows ICN classification and heatmap prediction performance when the manipulated query undergoes an increasingly severe ‘benign’ i.e. non-editorial transformation. Overall the prediction accuracy gradually drops as the benign transformation becomes more destructive. However, even under the most severe transformation, ICN still predicts the category of the query-candidate image pair correctly in most cases, as well as rough estimation of the manipulated areas. Fig. 4 shows a similar behaviour, but for the benign query.

Fig. 5 show examples of rare cases in the PSBat-Pair where the queries are subjected to multiple transformations. Ambiguous prediction results can arise in the presence of very severe benign transformations, as depicted in the last row of Fig. 5. It is interesting that even in these sparsely occurring cases, ICN predicts higher probability for the query-candidate pair being *manipulated* than being *distinct*. We note this is consistent to our findings in Fig.7 (right) in our main paper, where ICN achieves overall 99.6% accuracy if the classification results are reorganized to binary mode i.e. *distinct* versus *non-distinct*. This enables a possible further correction stage e.g. via manual inspection of the non-distinct results.

2. Heatmap Interpretability Task

We provide examples of the heatmaps evaluated in the Mechanical Turk (MTurk) user study to assess heatmap interpretability. Fig 6 provides examples of 10 tasks of the 200 presented to participants for the thresholded heatmap assessment. The manipulated image along with its ground truth (manipulated region) was shown to participants along with heatmaps from several methods (see main paper for experimental detail). The participant picked the heatmap that best described the manipulation. As methods were randomized in order, a numeric key (private from participants) is used to decode letters (A-I) to methods. Please see the figure caption for the key.

3. Video demo: web-based prototype

We include a screen-capture video demonstration of a web app built using the technology presented in this paper. The user drags and drops a query image (that may be manipulated and/or have benign transformations) and the system will find a near-duplicate match to an original within its trusted database. The image comparator is then used to visualize any differences due to manipulation (and ignore any due to benign transformations that may have occurred during image distribution online). The video shows examples of exact and manipulation variants of originals, and also examples of manipulated images that have been subjected to benign transformation (degradation and warping). The image comparator correctly ignores the latter when visualizing the differences between the query and retrieved original.

The video also includes an experiment applying the image comparator to individual frames of a video. The comparator correctly identifies regions of video manipulation and ignores benign transformations.



Figure 1. Top-10 returned results when querying a manipulated (top) and its transformed version (bottom) against the 2M PSBat-Ret dataset. Green box indicates relevant image. Our proposed models successfully retrieve the image of interest in both cases.



Figure 2. Top-10 returned results when querying two benign transformed images against the 2M PSBat-Ret dataset. Green box indicates relevant image. Our models demonstrate the robustness against severe (benign) transformations.

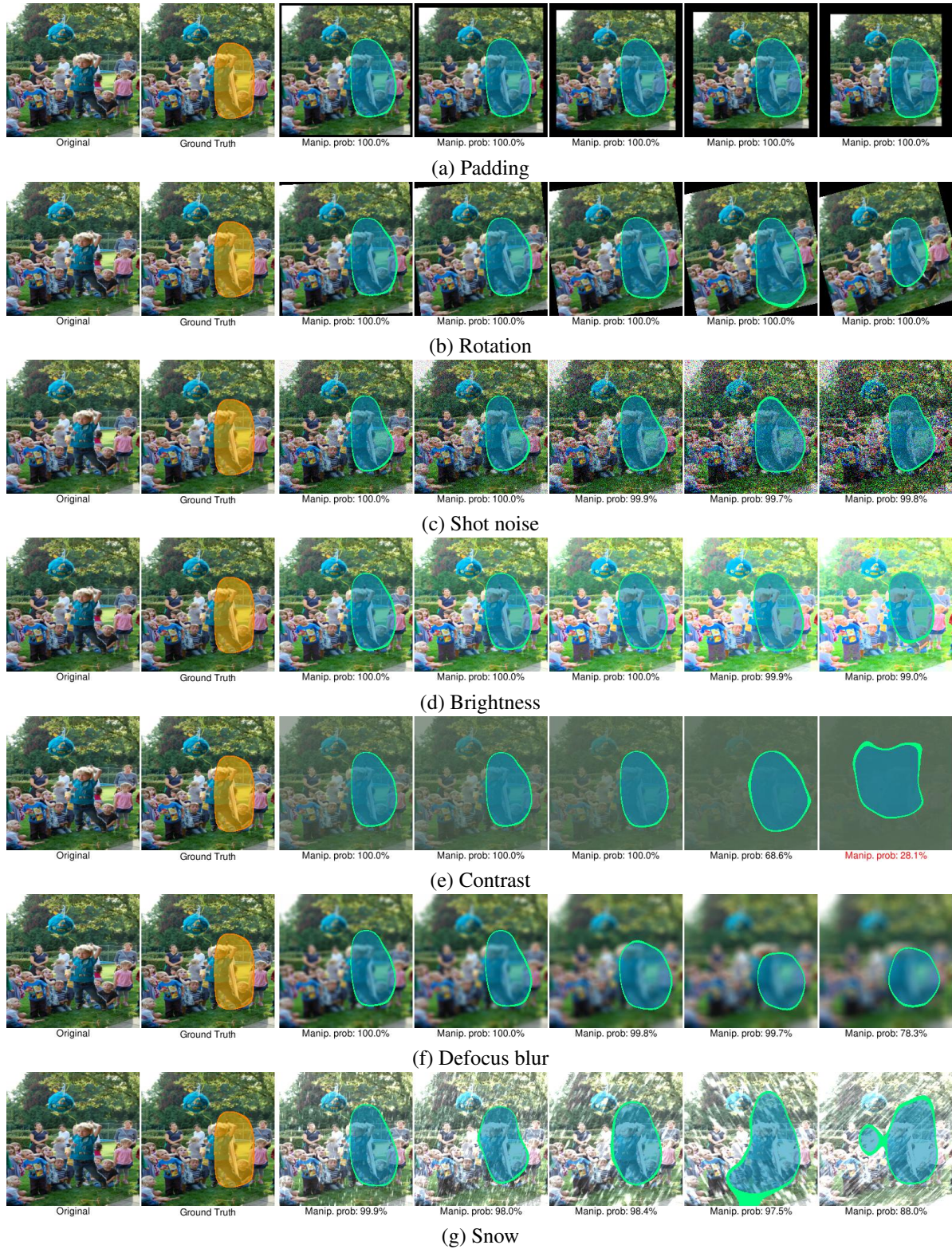


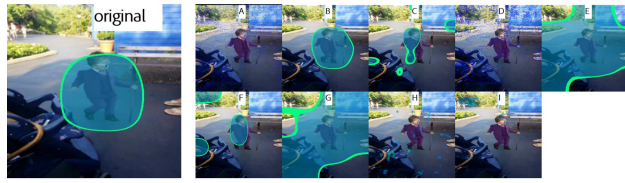
Figure 3. Effects of transformation severity on ICN predicted heatmap and classification performance, for a manipulated query. Red text indicates incorrect classification.



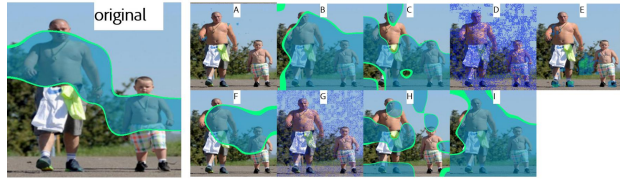
Figure 4. Effects of transformation severity on ICN classification performance, for a benign query.



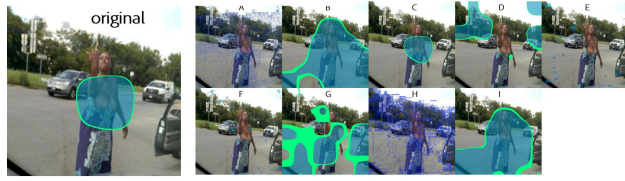
Figure 5. Classification and heatmap visualisation of several challenging examples in a 8x2 grid. For each example, the order of images are: far left - original/candidate image, middle left - query image, middle right - ground truth (yellow) and predicted (green) heat maps, far right - prediction scores. The orange tint on some images is due to both predicted and target heatmaps covering the whole image. Red boxes indicate failure cases.



091__7_4_0_8_2_1_3_5_6



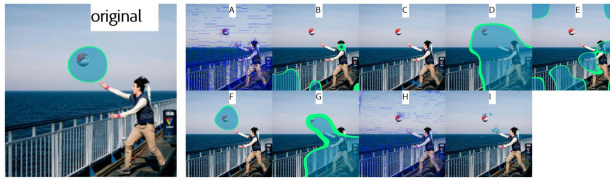
121__5_3_1_8_6_4_7_0_2



129__7_2_4_0_6_5_1_8_3



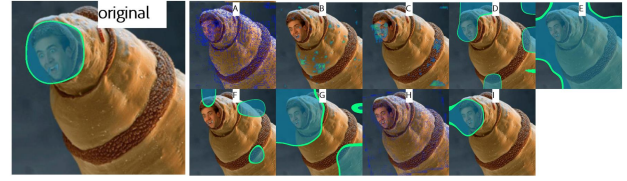
133__5_7_8_0_6_1_2_4_3



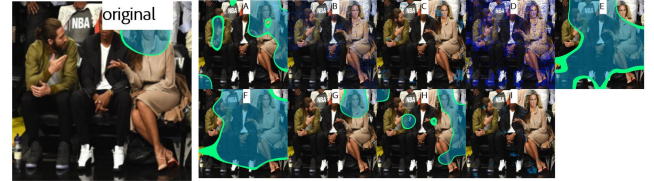
194__8_1_6_2_0_4_3_7_5



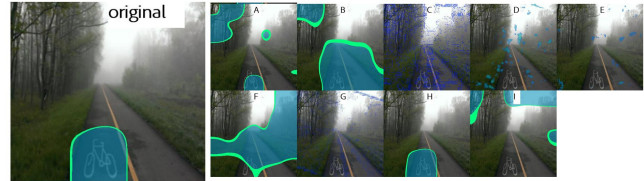
134__7_6_8_0_5_4_2_1_3



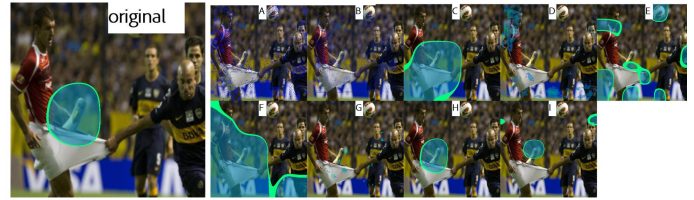
122__8_5_6_1_2_0_3_7_4



028__1_7_6_8_2_3_4_0_5



135__1_2_8_5_6_3_7_4_0



195__8_7_2_6_0_3_5_4_1

Figure 6. Heatmap interpretability study: sample of data shown to Mechanical Turk participants: manipulated original and heatmaps. The heatmaps are presented as options A-I in randomized order. Disregarding the initial number (task ID), the remaining numeric sequence is a key (not shown to participants) decoding A-I to the method producing the heatmap: 0) SSD; 1) SSD+DWU; 2) ResNet; 3) ResNet+DWU; 4) Our ICN Method; 5) MantraNet; 6) MantraNet+DWU; 7) ErrAnalysis; 8) ErrAnalysis+DWU.