Learning to Compose Hyperpixels for Visual Correspondence
-Supplementary material-

Juhong Min[1,2] Jongmin Lee[1,2] Jean Ponce[3,4] Minsu Cho[1,2]

[1]POSTECH[*] [2]NPRC[†] [3]Inria [4]ENS[‡]
http://cvlab.postech.ac.kr/research/DHPF/

In this supplementary material, we present additional qualitative comparison with recent state of the arts [3, 5, 6] on PF-PASCAL [2] and SPair-71k [4] datasets. **The qualitative comparisons** are shown in Fig. 1, 2, 3 and 4. **Representative failure cases** of our model are shown in Fig. 5. **Example results of animals and vehicle categories** with selected layer distributions are shown in Fig. 6 and 7; the animal category tends to select fewer layers compared to vehicle category in our experiments. **Examples with the least and the most numbers of selected layers** are shown in Fig. 8, 9 and 10; easy examples tend to select fewer layers compared to hard examples with background clutters, occlusion, and truncation. All the results are visualized by warping each source image to its target image with TPS [1] transformation according to correspondences predicted by the models.

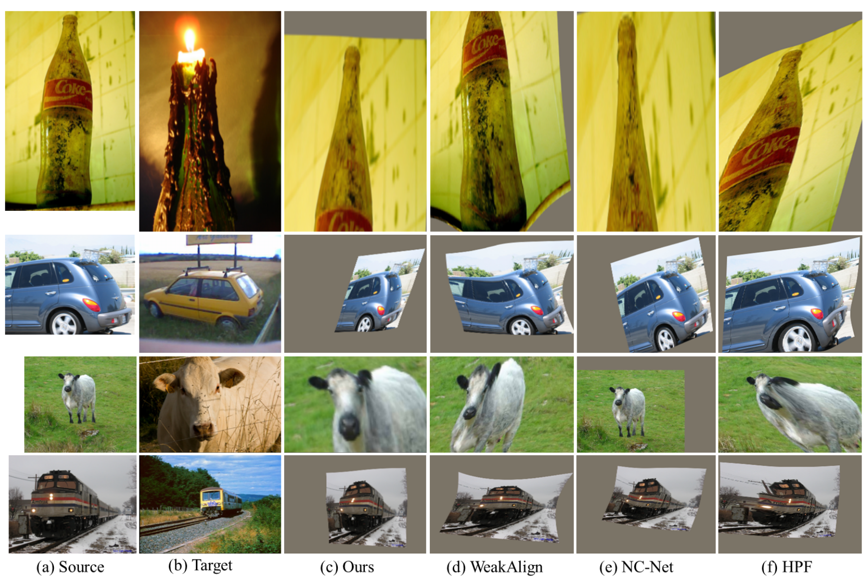

Fig. 1: Examples with truncation from SPair-71k dataset [4]: (a) source image, (b) target image (c) DHPF (ours), (d) WeakAlign [5], (e) NC-Net [6] and (f) HPF [3].

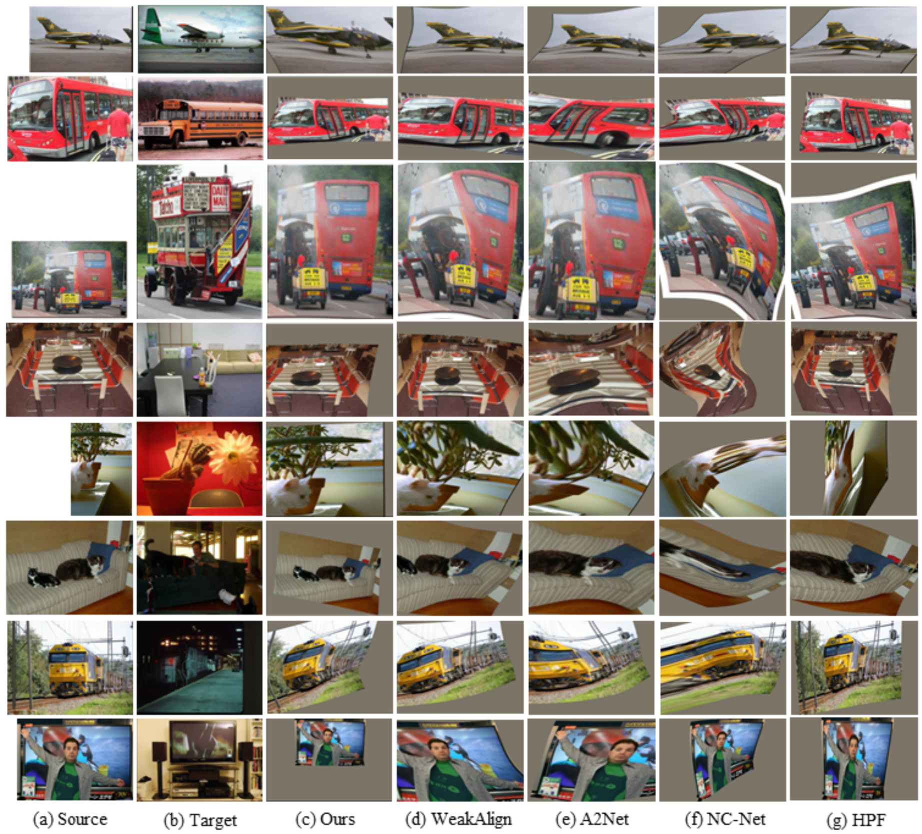

Fig. 2: Example results on PF-PASCAL dataset [2]: (a) source image, (b) target image (c) DHPF (ours), (d) WeakAlign [5], (e) NC-Net [6] and (f) HPF [3].

Fig. 3: Examples with large view-point differences from SPair-71k dataset [4]: (a) source image, (b) target image (c) DHPF (ours), (d) WeakAlign [5], (e) NC-Net [6] and (f) HPF [3].

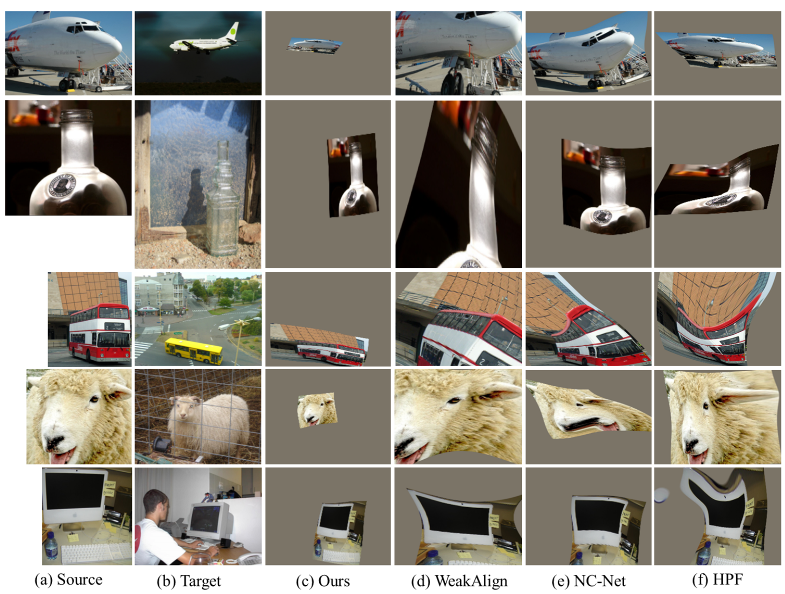

Fig. 4: Examples with large scale differences from SPair-71k benchmark [4]: (a) source image, (b) target image (c) DHPF (ours), (d) WeakAlign [5], (e) NC-Net [6] and (f) HPF [3].

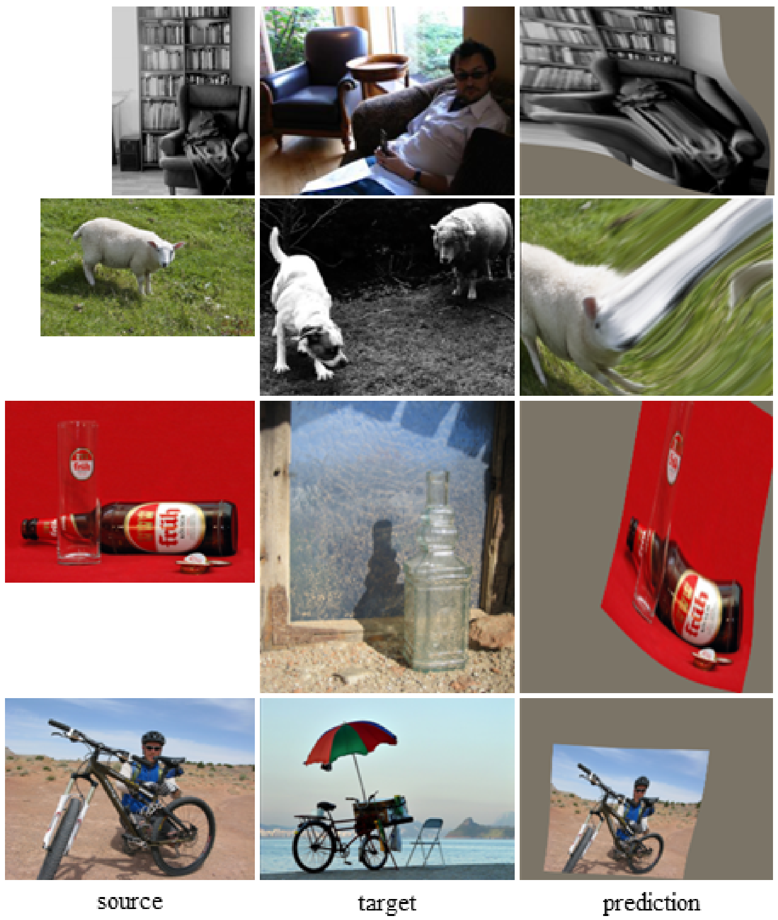

Fig. 5: Failure cases on PF-PASCAL [2] and SPair-71k [4] dataset in presence of multiple instances with similar appearance, and large rotation or view-point differences.

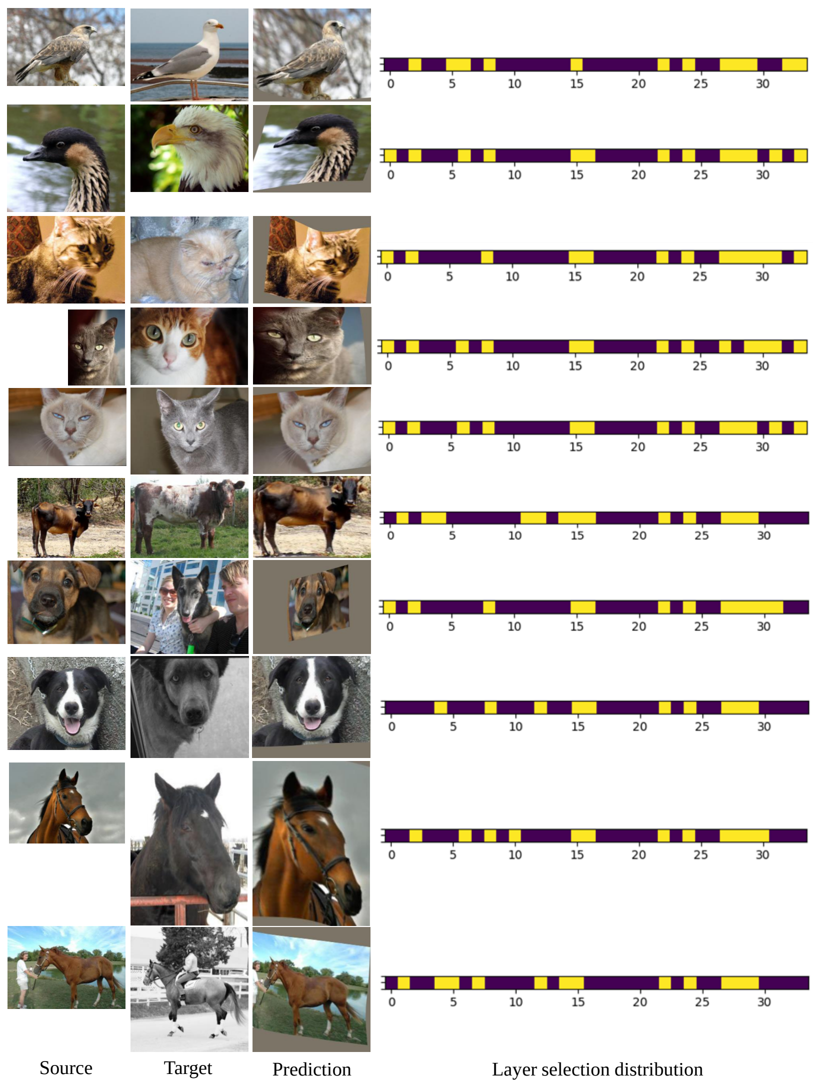

Fig. 6: (Left) Results of animal category from PF-PASCAL [2] and (Right) their corresponding layer selection distribution (selected layers are colored in yellow).

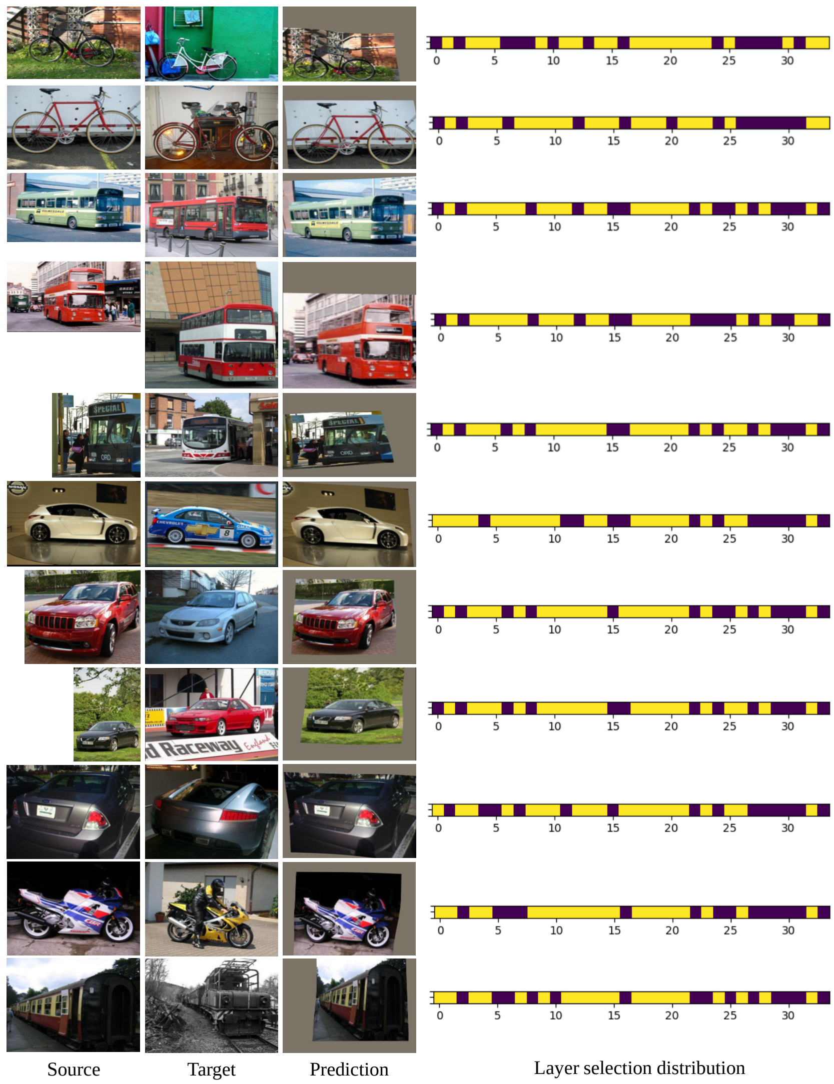

Fig. 7: (Left) Results of vehicle category from PF-PASCAL [2] and (Right) their corresponding layer selection distribution (selected layers are colored in yellow).

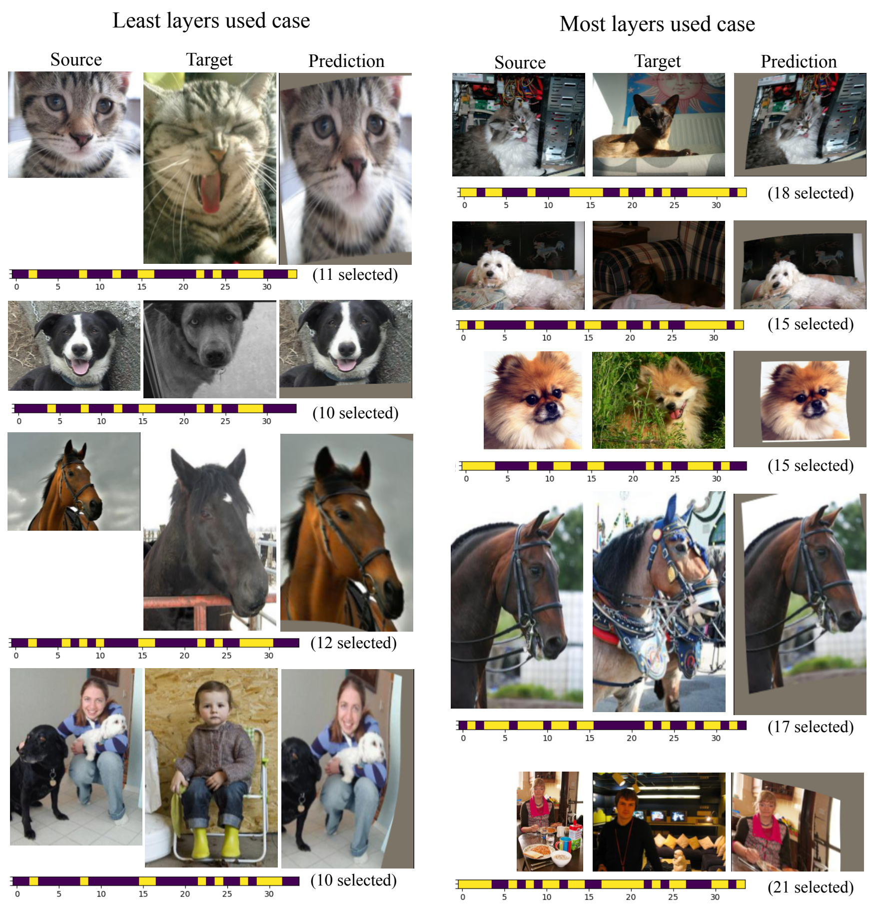

Fig. 8: Examples of the least and the most numbers of selected layers on PF-PASCAL [2] animal category with corresponding layer distributions.

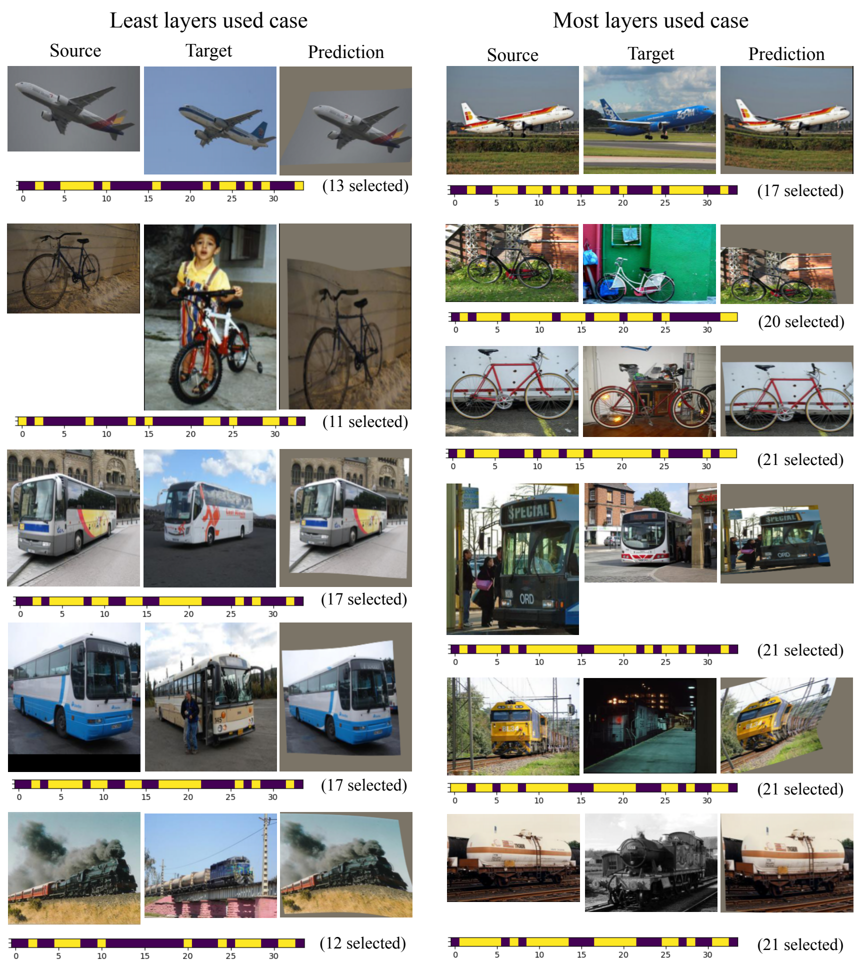

Fig. 9: Examples of the least and the most numbers of selected layers on PF-PASCAL [2] vehicle category with corresponding layer distributions.

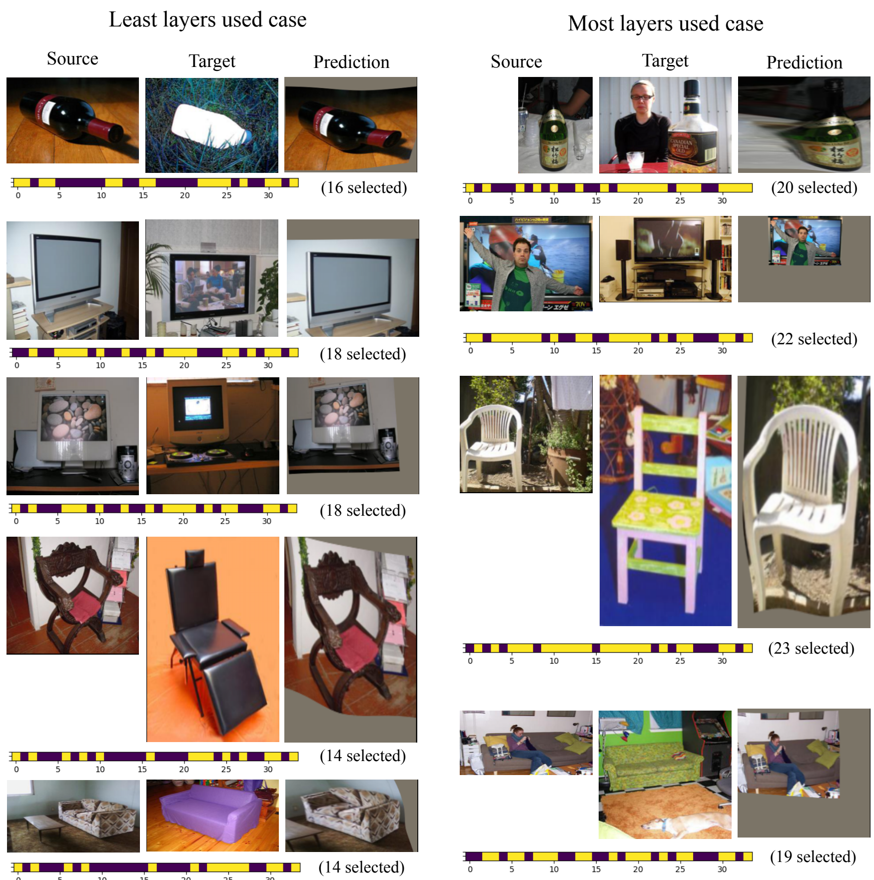

Fig. 10: Examples of the least and the most numbers of selected layers on PF-PASCAL [2] man-made object category with corresponding layer distributions.

References

1. Donato, G., Belongie, S.: Approximate thin plate spline mappings. In: Proc. European Conference on Computer Vision (ECCV) (2002) 1
2. Ham, B., Cho, M., Schmid, C., Ponce, J.: Proposal flow: Semantic correspondences from object proposals. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) **40**, 1711–1725 (2018) 1, 2, 4, 5, 6, 7, 8, 9
3. Min, J., Lee, J., Ponce, J., Cho, M.: Hyperpixel flow: Semantic correspondence with multi-layer neural features. In: Proc. IEEE International Conference on Computer Vision (ICCV) (2019) 1, 2, 3
4. Min, J., Lee, J., Ponce, J., Cho, M.: SPair-71k: A large-scale benchmark for semantic correspondence. arXiv prepreint arXiv:1908.10543 (2019) 1, 3, 4
5. Rocco, I., Arandjelović, R., Sivic, J.: End-to-end weakly-supervised semantic alignment. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018) 1, 2, 3
6. Rocco, I., Cimpoi, M., Arandjelović, R., Torii, A., Pajdla, T., Sivic, J.: Neighbourhood consensus networks. In: Proc. Neural Information Processing Systems (NeurIPS) (2018) 1, 2, 3