Dynamic Context Correspondence Network for Semantic Alignment Supplement

Shuaiyi Huang, Qiuyue Wang, Songyang Zhang, Shipeng Yan, Xuming He
ShanghaiTech University

A. More Results on PF-Pascal

In order to have a better understanding of our model DCCNet, we provide detailed analysis on PF-PASCAL [3] here. Apart from the PCK ($\alpha = 0.1$) results reported in the main paper, we additionally include PCK results with $\alpha = 0.05$ and $\alpha = 0.15$, as shown in Table S1. Our DCCNet surpasses the state-of-the-art correspondence method by 1.3%, 3.4% and 3.5%, and reaches 55.6%, 82.3% and 90.5% at $\alpha = 0.05$, 0.10 and 0.15, respectively. This demonstrates the effectiveness of DCCNet for matching at both fine and coarse level.

More qualitative comparisons with NC-Net [10] are presented in Fig. S1 and Fig. S2. To better understand our features, we visualize the raw correlation score maps of the local conv feature and our context-aware semantic feature in the third columns of Fig. S2 (a,b). Compared with the local conv feature ((a) col.3), our proposed context-aware semantic feature gives more informative correlation score maps ((b) col.3)). The fourth columns of Fig. S2 (a,b) shows the final correlation score maps from the NC-Net and our DCCNet. It is evident that our model is more robust to repetitive patterns and background clutters.

B. Qualitative Results on PF-WILLOW and TSS

More qualitative results on PF-WILLOW [2] from our DCCNet are shown in Fig. S3. Our model produces reasonable matching results despite heavy background clutters, large viewpoint changes and scale variations.

More qualitative results from our DCCNet on TSS [11] are shown in Fig. S4.

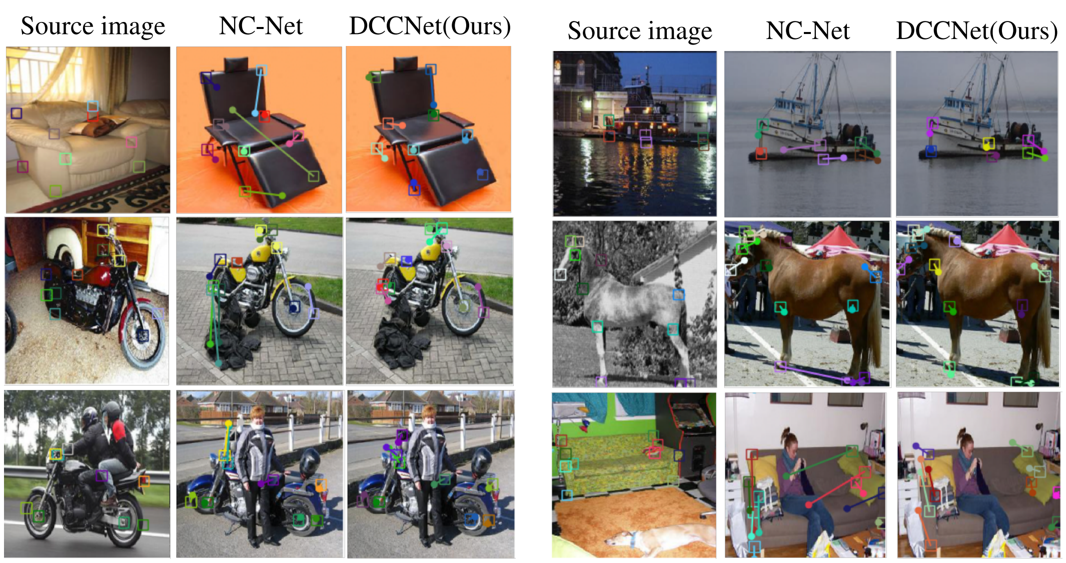

Figure S1: **Qualitative comparisons on PF-PASCAL.** From left to right shows **source images**, **predictions from NC-Net** [10] and **from ours** respectively. Ground truth keypoints are shown in squares and predicted keypoints in dots, with their distance in target images depicting the matching error. It is clear that our model is robust to repetitive patterns and background clutters.

Methods	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.15$
HOG+PF-LOM [3]	31.4	62.5	79.5
DCTM [7]	34.2	69.6	80.2
UCN-ST [1]	29.9	55.6	74.0
CAT-FCSS [6]	33.6	68.9	79.2
SCNet [4]	36.2	72.2	82.0
ResNet-101+CNNGeo [8]	41.1	71.9	84.9
ResNet-101+CNNGeo(W) [9]	46.0	75.8	88.4
RTN [5]	55.2	75.9	85.2
NC-Net [10]	54.3	78.9	86.0
Our Method	**55.6**	**82.3**	**90.5**

Table S1: Detailed results on the PF-PASCAL benchmark [3].

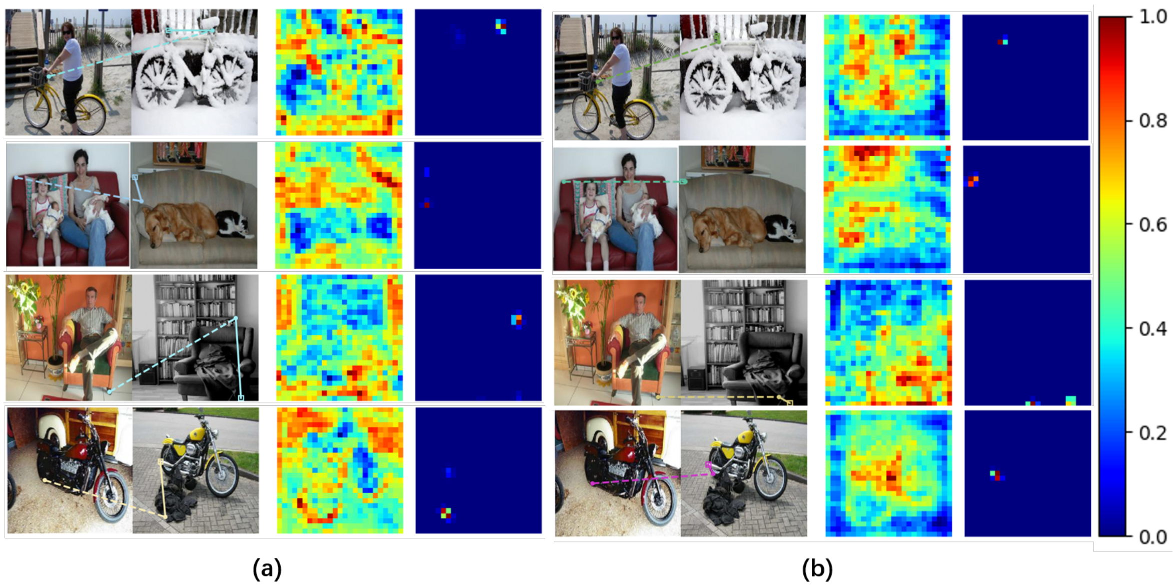

Figure S2: **Prediction results from *(a) NC-Net* [10] and *(b) DCCNet (ours)* on PF-PASCAL [3].** Given a source point (dot in column 1), the goal is to match the target point (square in column 2). Raw correlation score maps from local conv features and our context-aware semantic features are shown in column 3 *(a)* and *(b)* respectively. Column 4 *(a)* and *(b)* shows the final correlation score maps of the NC-Net and our DCCNet while their predictions are shown in column 2 (dot) of *(a)* and *(b)* respectively. Our model generates more accurate correlation maps and final predictions.

References

[1] C. B. Choy, J. Gwak, S. Savarese, and M. Chandraker. Universal correspondence network. In *Advances in Neural Information Processing Systems(NeurIPS)*, 2016.

[2] B. Ham, M. Cho, C. Schmid, and J. Ponce. Proposal flow. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2016.

[3] B. Ham, M. Cho, C. Schmid, and J. Ponce. Proposal flow: Semantic correspondences from object proposals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[4] B. H. K.-Y. K. W. M. C. C. S. Kai Han, Rafael S. Rezende and J. Ponce. Scnet: Learning semantic correspondence. In *Proceedings of the International Conference on Computer Vision(ICCV)*, 2017.

[5] S. Kim, S. Lin, S. R. JEON, D. Min, and K. Sohn. Recurrent transformer networks for semantic correspondence. In *Advances in Neural Information Processing Systems(NeurIPS)*, 2018.

[6] S. Kim, D. Min, B. Ham, S. Lin, and K. Sohn. Fcss: Fully convolutional self-similarity for dense semantic correspondence. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[7] S. Kim, D. Min, S. Lin, and K. Sohn. Dctm: Discrete-continuous transformation matching for semantic flow. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2017.

Figure S3: **Qualitative results on PF-WILLOW** [2] for semantic correspondence.

Figure S4: **Qualitative results on the TSS benchmark** [11]. From left to right are source image, target image, results from WeakAlign [9], NC-Net [10] and our model, respectively.

[8] I. Rocco, R. Arandjelović, and J. Sivic. Convolutional neural network architecture for geometric matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2017.

[9] I. Rocco, R. Arandjelović, and J. Sivic. End-to-end weakly-supervised semantic alignment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2018.

[10] I. Rocco, M. Cimpoi, R. Arandjelović, A. Torii, T. Pajdla, and J. Sivic. Neighbourhood consensus networks. In *Advances in Neural Information Processing Systems(NeurIPS)*, 2018.

[11] T. Taniai, S. N. Sinha, and Y. Sato. Joint recovery of dense correspondence and cosegmentation in two images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, 2016.