Supplementary Material
Deep Semantic Matching with Foreground Detection and Cycle-Consistency

Anonymous ACCV 2018 submission

Paper ID 632

In this supplementary document, we present additional visual examples for comparing our method with the state-of-the-art semantic matching algorithm [1].

1 Visual Comparisons

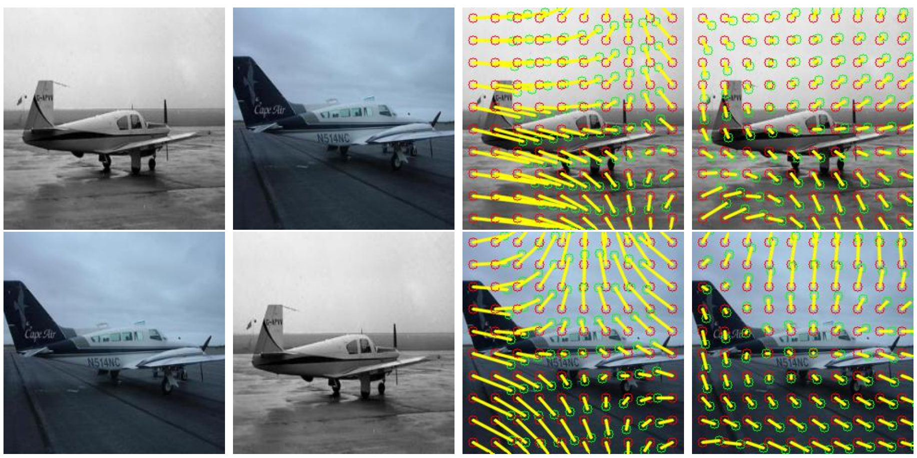
1.1 Forward-Backward Consistency

In Fig. 1 $\sim$ Fig. 4, we demonstrate the effectiveness of the proposed forward-backward consistency loss by visualizing the reprojection errors using the estimated forward and backward transformations. We uniformly sample 100 pixel coordinates in the source image, i.e., a 10×10 grid $\mathbf{v}$, and mark them with red circles. The geometric transformation T_{AB} is learned to match the source image to the target image, while another transformation T_{BA} is derived to match it back. We then apply T_{AB} and T_{BA} in order to each of the sampled points to obtain the reprojected points $\mathbf{u}$ marked with green circles, namely,

$$\mathbf{u} = T_{BA}(T_{AB}(\mathbf{v})). \tag{1}$$

The yellow line links the red and green circles represents the distance (loss) between the linked points. It can be observed that introducing the proposed forward-backward consistency loss greatly reduces the reprojection errors, hence improving semantic matching.

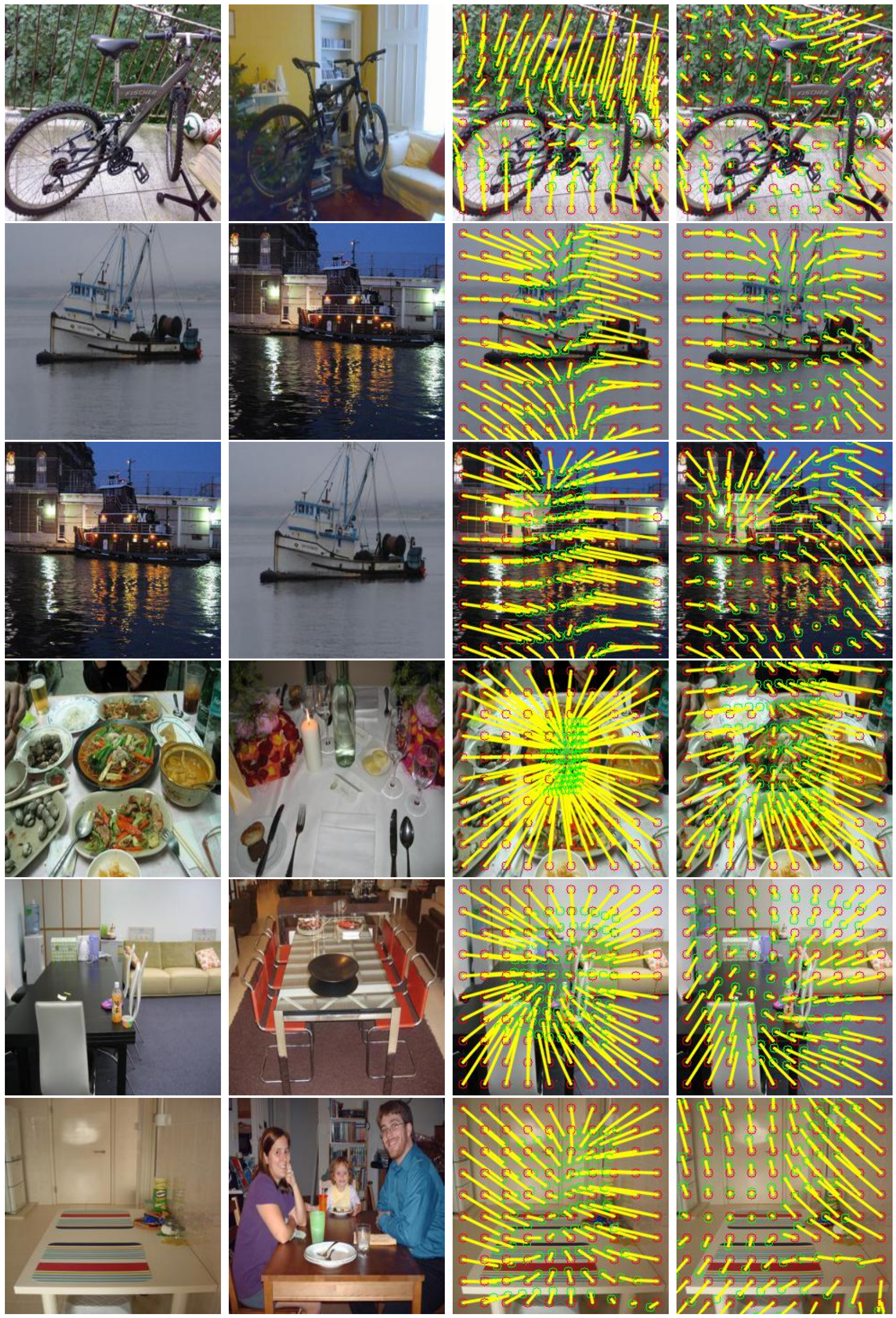
ACCV2018
#632



(a) Source (b) Target (c) Baseline [1] (d) Ours

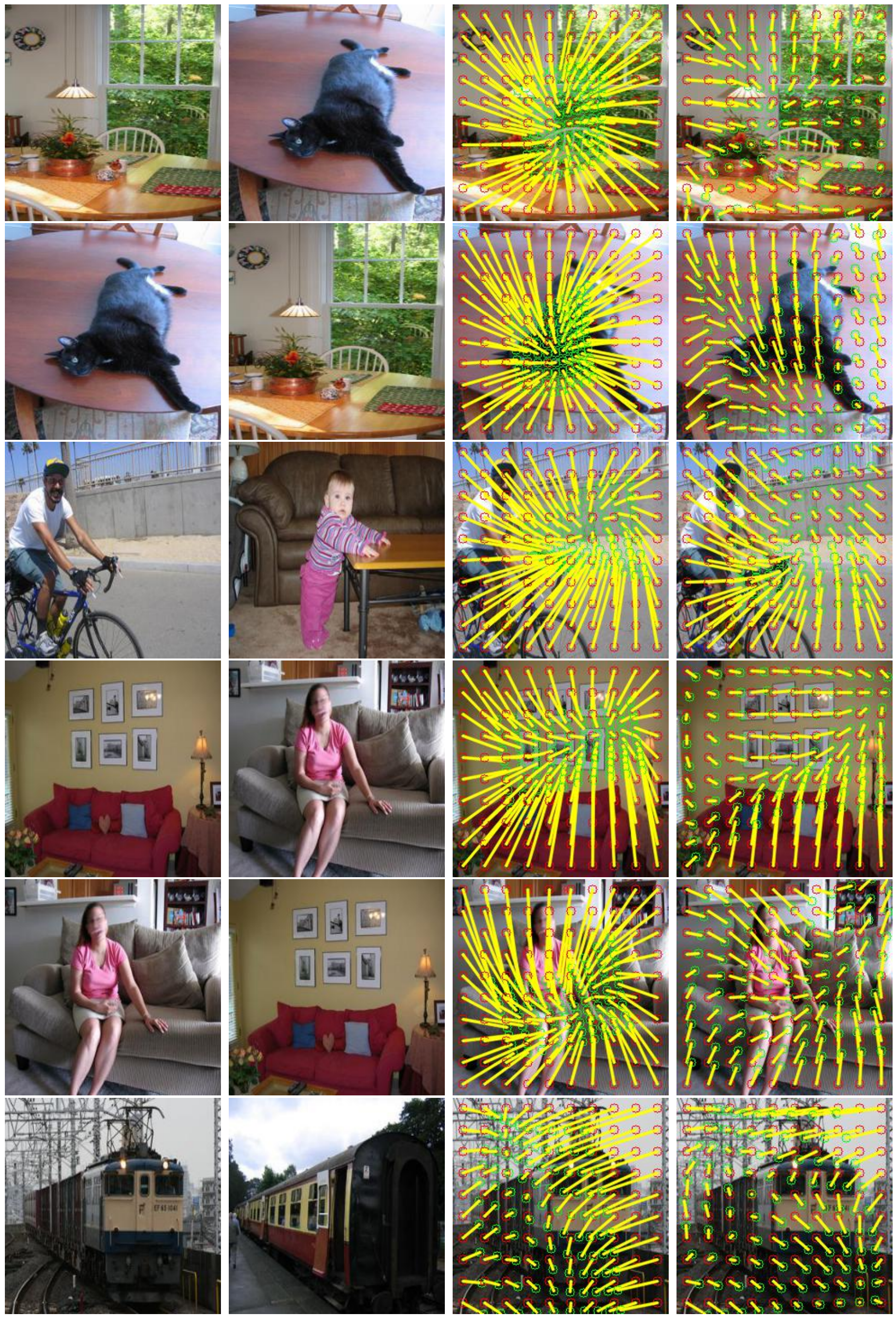
Fig. 1: **Visualization of the effect of the forward-backward consistency.**

ACCV2018
#632



(a) Source (b) Target (c) Baseline [1] (d) Ours

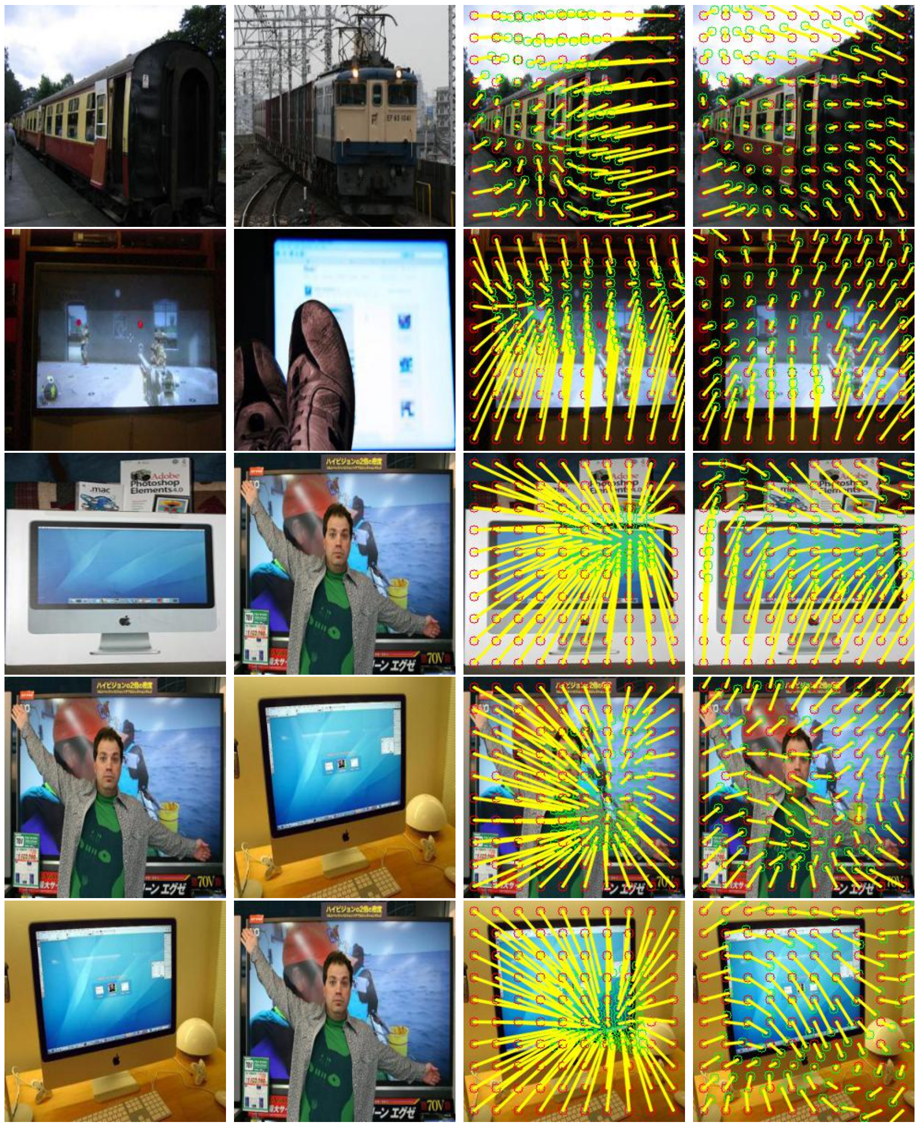
Fig. 2: **Visualization of the effect of the forward-backward consistency.**



(a) Source (b) Target (c) Baseline [1] (d) Ours

Fig. 3: **Visualization of the effect of the forward-backward consistency.**

ACCV2018
#632



(a) Source (b) Target (c) Baseline [1] (d) Ours

Fig. 4: **Visualization of the effect of the forward-backward consistency.**

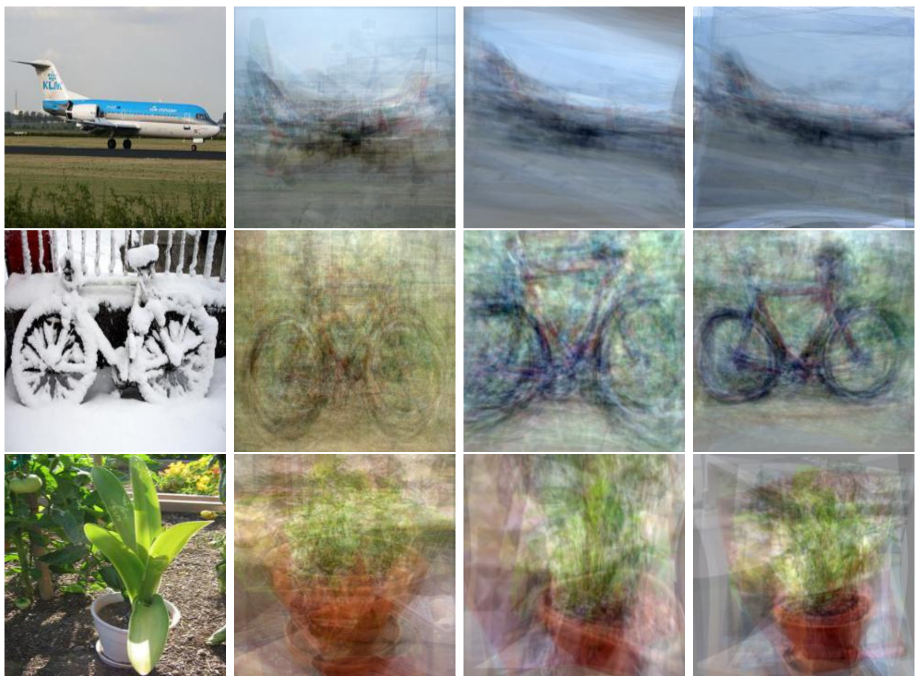
ACCV2018 #632

1.2 Image Alignment

We compare the method in [1] and the proposed method for image alignment. For each image pair (I_A, I_B), each method estimates a geometric transformation T_{AB} which warps image I_A to $\tilde{I}_A$ so that $\tilde{I}_A$ and I_B are well aligned. The alignment results are displayed with the GIF files collected by an html file. Please refer to "warp.html" for the visual results.

1.3 Average Images within Object Category

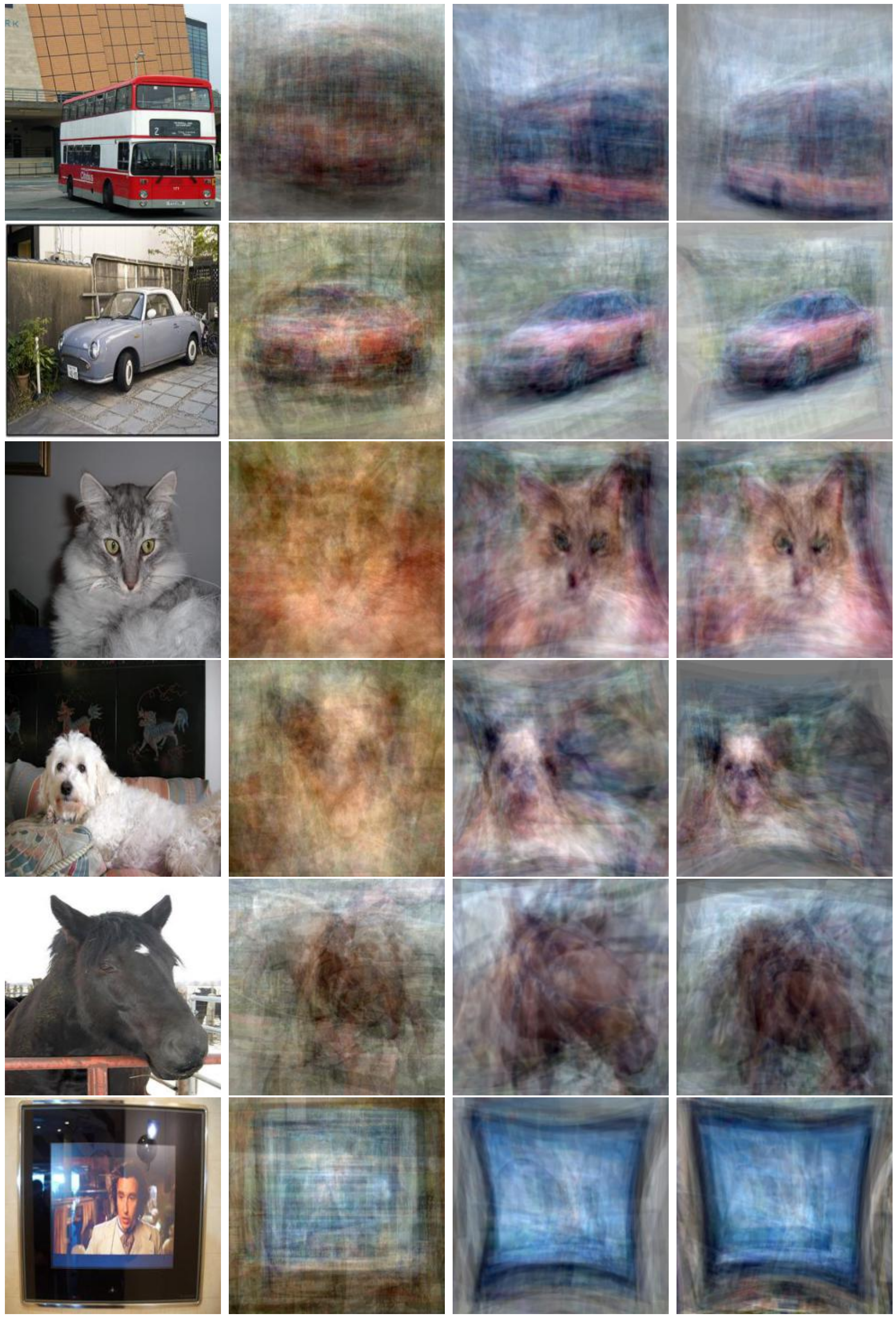
Let $\mathcal{D} = \{I_i\}_{i=1}^{N}$ denote a set of images covering object instances of a specific category, where I_i is the i^{th} image and N is the number of images. For each source image $I_A \in \mathcal{D}$, the geometric transformation T_{BA} can be estimated for each other image $I_B \in \mathcal{D}$. We warp each I_B based on the geometric transformation T_{BA} and then compute the average warped image I_{avg} by averaging all the warped images. We present the visual comparisons among three different methods: Average, Baseline [1], and Ours. For the "Average" method, we directly compute the average image over all images $\{I_B\}$ without applying any geometric transformations. As can be seen in Fig. 5 and Fig. 6, the average images by method "Average" do not align the source images. Compared with the competing method [1], our method can better align the given source images, and the resultant average images look sharper.



(a) Source (b) Average (c) Baseline [1] (d) Ours

Fig. 5: **Class-wise average warping.**

ACCV2018 #632

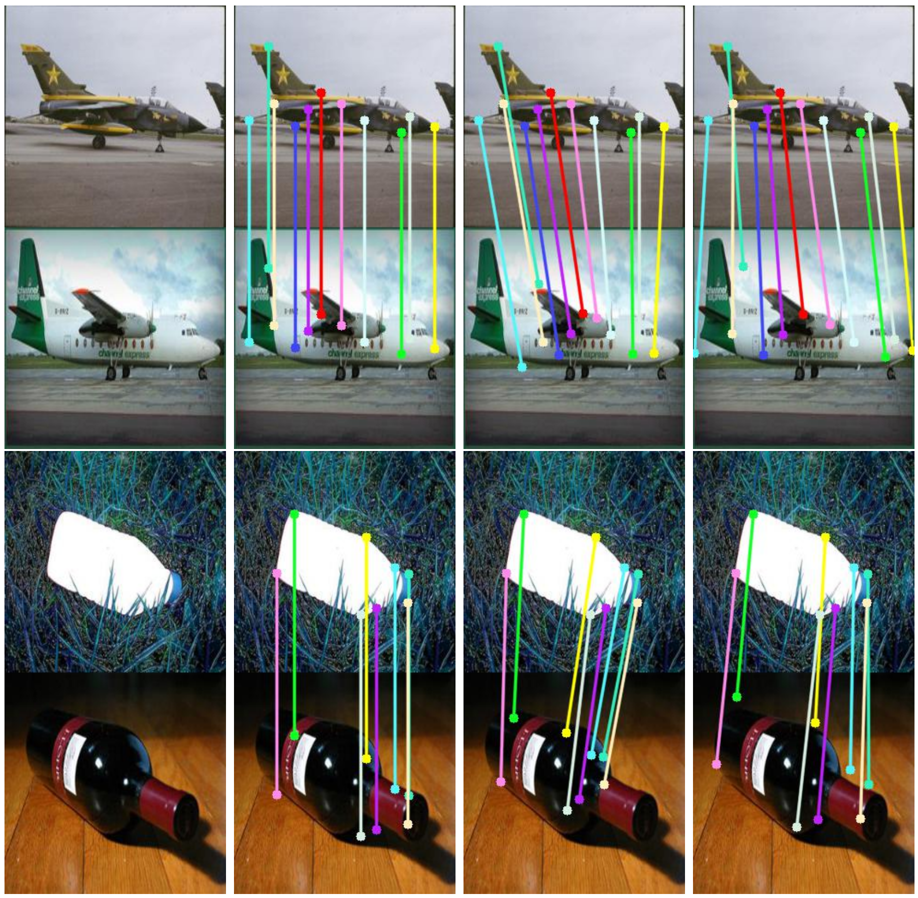


(a) Source (b) Average (c) Baseline [1] (d) Ours

Fig. 6: **Class-wise average warping.**

1.4 More Results of Semantic Correspondence

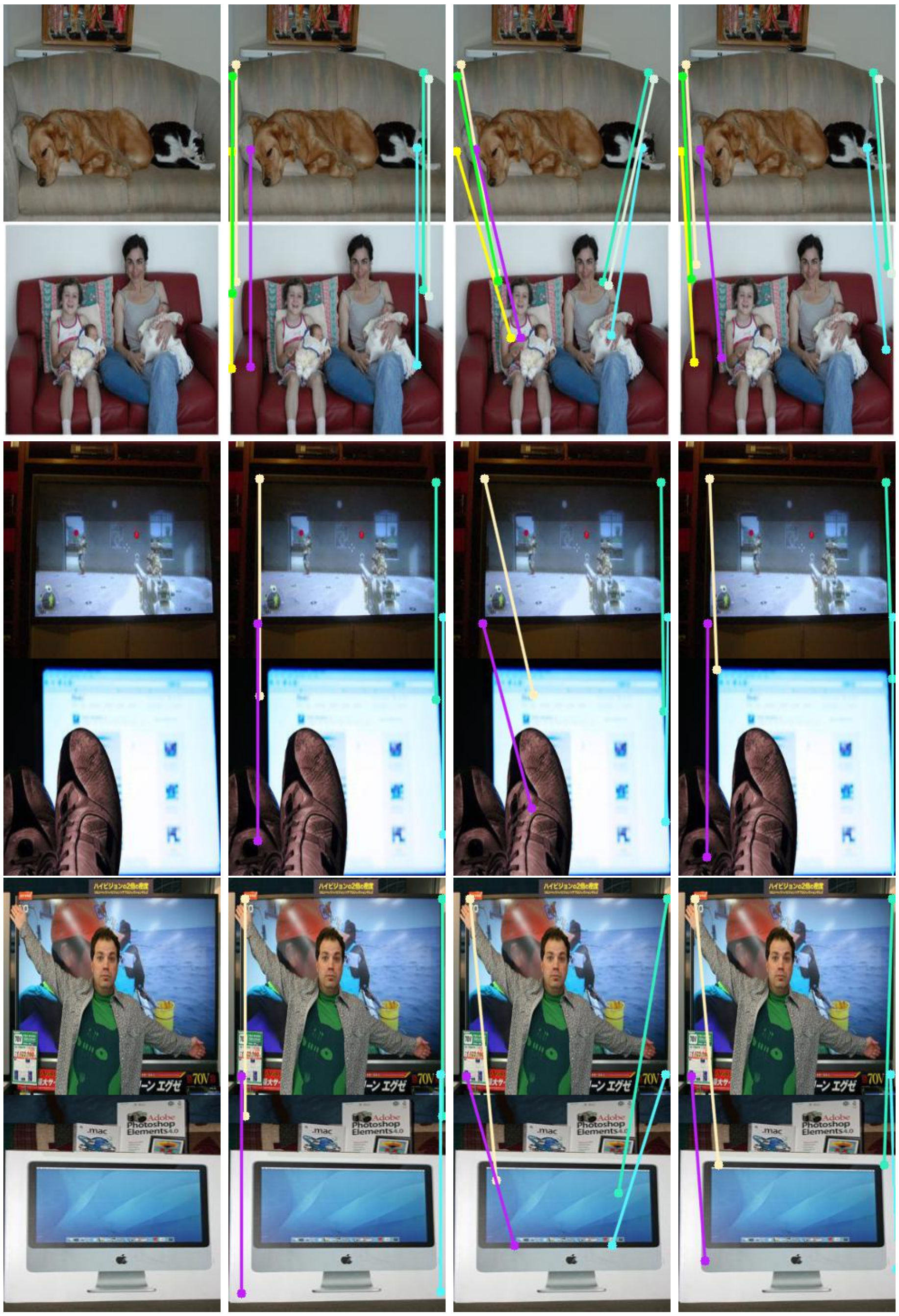
We display more results of semantic matching on the PF-PASCAL dataset in Fig. 7 and Fig. 8, where our method establishes more accurate correspondences than the competing method [1].



(a) Source (b) Ground Truth (c) Baseline [1] (d) Ours

Fig. 7: **Semantic correspondence results.**

ACCV2018
#632



(a) Source (b) Ground Truth (c) Baseline [1] (d) Ours

Fig. 8: **Semantic correspondence results.**

ACCV2018 #632

References

1. I. Rocco, R. Arandjelović, and J. Sivic. End-to-end weakly-supervised semantic alignment. In *CVPR*, 2018.