Deep Semantic Image Retargeting
Supplementary Material

This supplementary material contains three parts. First, we give qualitative comparisons between the importance maps generated by our method and other baseline methods. Second, we show more results of our retargeting systems. Third, we present more information on our experimental results over the Amazon Mechanical Turk (AMT) platform. In all of our experiments, the original images are resized to half of their widths while keeping their heights unchanged.

1. Comparisons on importance maps

In this section, we provide several qualitative comparisons of importance maps. These importance maps contain the maps generated by our method and 6 other methods, including MC [13], GC [2], RCC [3], importance map used in the original IF (oriIF) [4], eDN [10], and DNEF [9]. We selected 30 Images from the new proposed Semantic Retarget dataset, specifically 5 images from each of the 6 categories, i.e. single person, multiple people, single object, multiple objects, indoor scene, and outdoor scene. Comparisons are listed below.

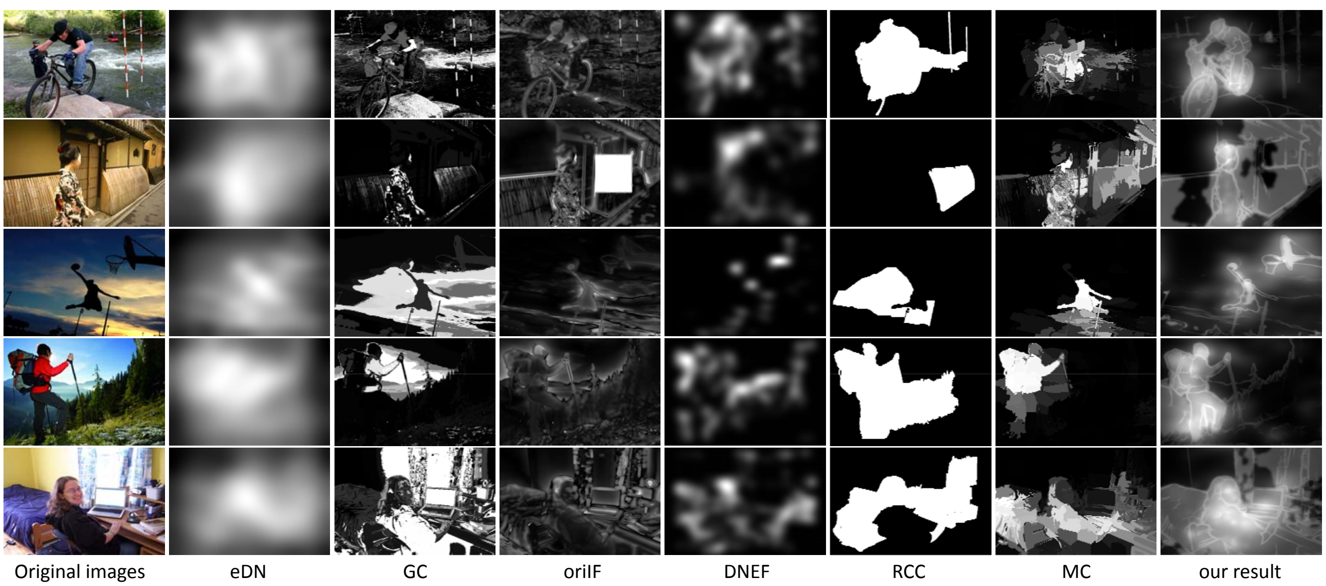

Figure 1. Comparisons between our map and 6 baseline maps. Images are selected from the category of single person.

2. More comparisons with retargeting systems

We feed our importance map to the carrier IF [4] which is selected because of its excellent performance and fast speed. We present more comparisons between our retarget results and 6 other retarget systems. These baseline retarget systems are listed below.

SOAT [5]: SOAT is a method to effectively obtain image thumbnails based on cropping and warping. It models the human perception of thumbnail by visual acuity theory, and gets the *Scale and Object Aware Saliency* (SOAS) first, and then it uses SOAS to do image thumbnailing.

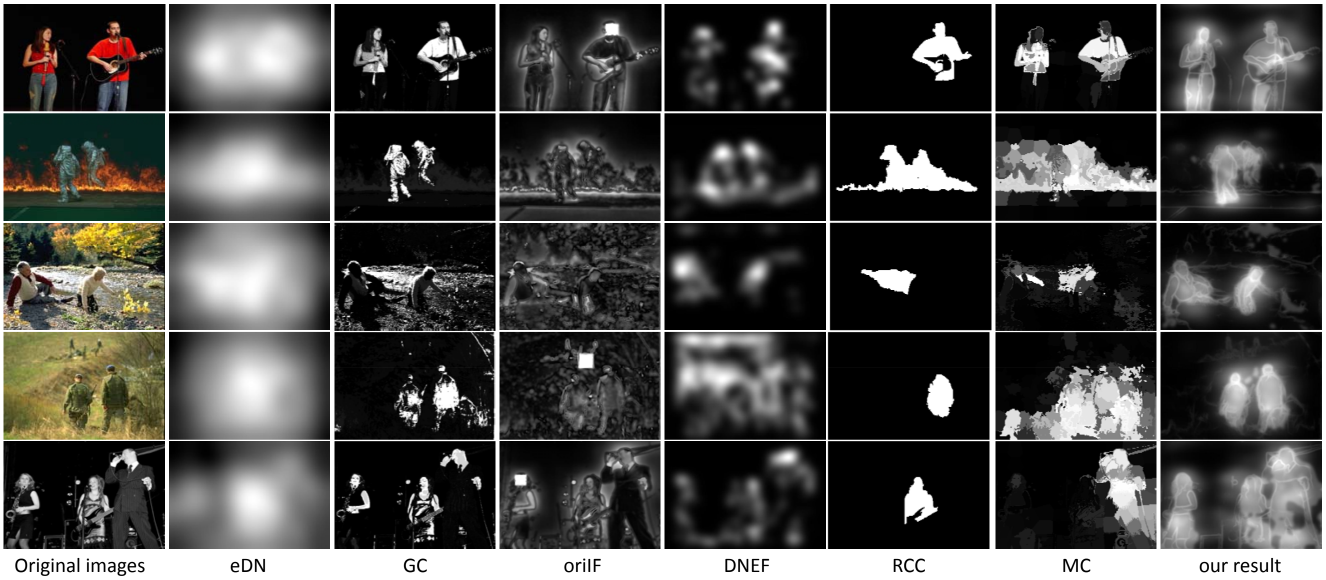

Figure 2. Comparisons between our map and 6 baseline maps. Images are selected from the category of multiple people.

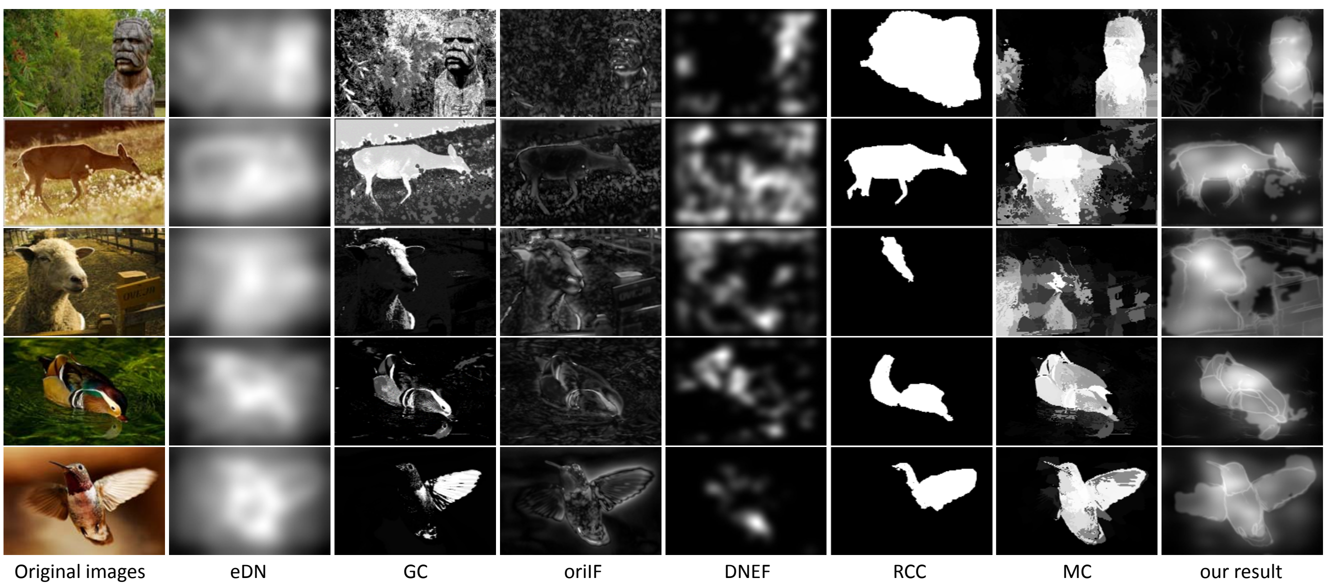

Figure 3. Comparisons between our map and 6 baseline maps. Images are selected from the category of single object.

ISC [7]: Improved Seam Carving (ISC) is the improved version of the famous Seam Carving [1] approach for retargeting. For the improvement, instead of using the dynamic programming approach of seam carving, ISC prefers to graph cuts. A novel energy criterion is also used in ISC to improve the visual quality of the retargeted images.
Multi-Op [8]: Multi-Operator (Multi-Op) approach provides a hybrid method of doing image retargeting. Multiple operators, including seam carving, cropping and scaling, are used alternatively to produce the results. An image similarity measure, named Bi-Directional Warping, is used to find the optimal path in the retargeting space.
Warp [12]: Warp is a warping approach for retargeting. It first analyzes the importance of each region, and then it applies a transformation which shrinks less important regions more than important ones. This method can work both on images and videos.
AAD [6]: Axis-Aligned Deformation (AAD) is a robust image retargeting method. To avoid harmful visual distortions,

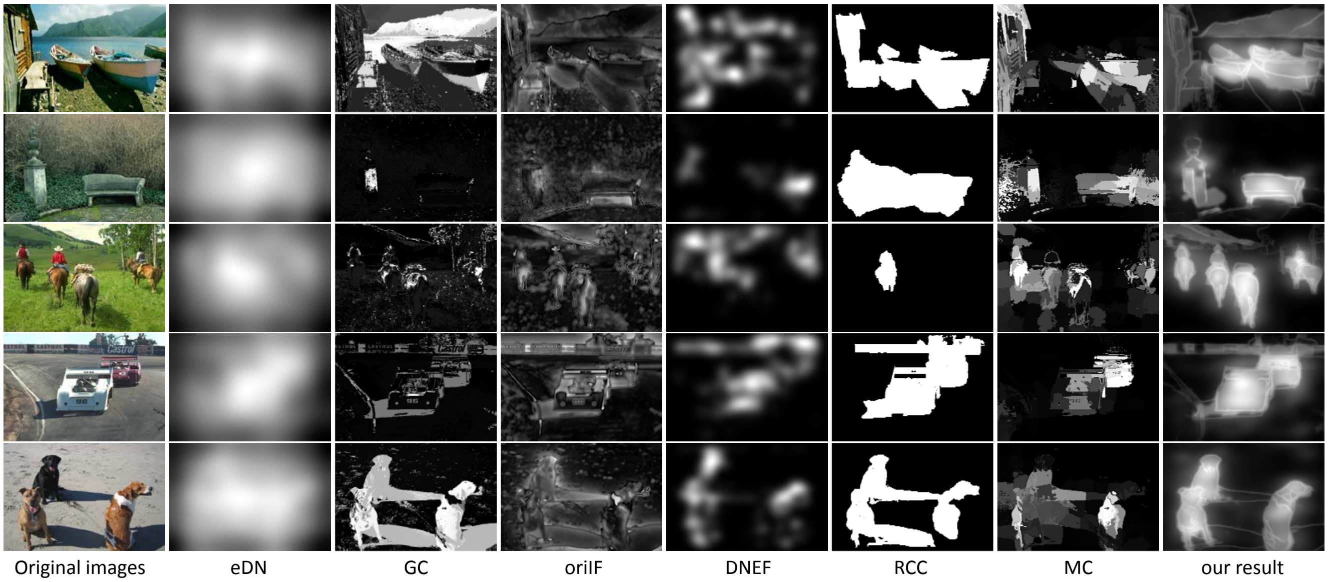

Figure 4. Comparisons between our map and 6 baseline maps. Images are selected from the category of multiple objects.

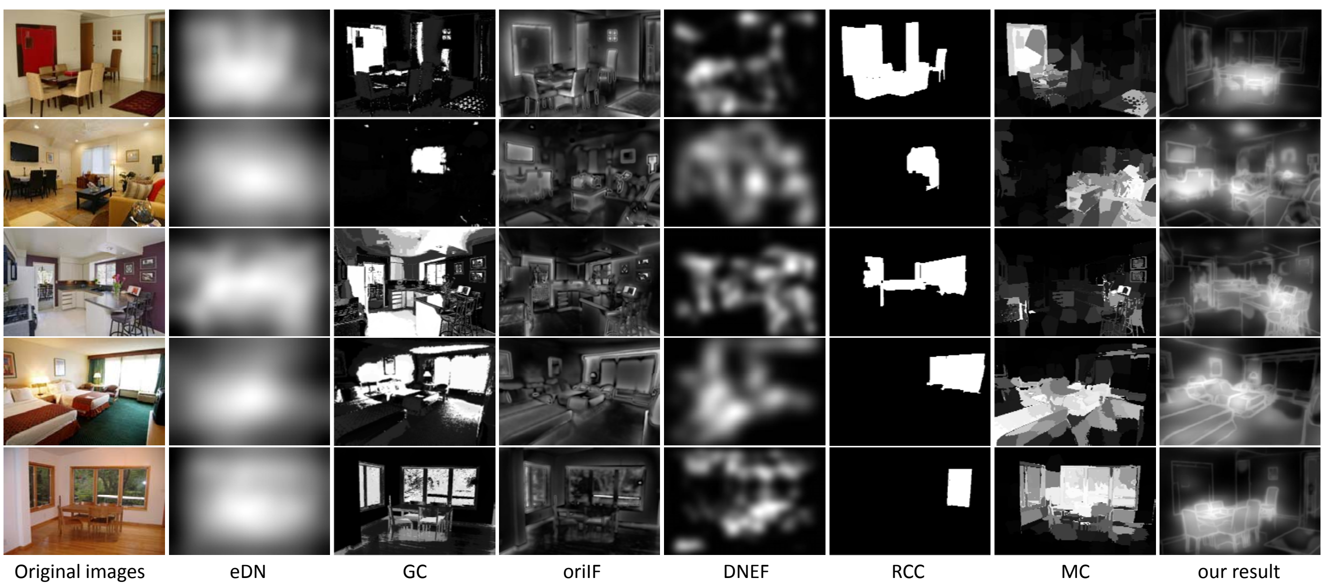

Figure 5. Comparisons between our map and 6 baseline maps. Images are selected from the category of indoor scene.

deformations in AAD are parameterized in 1-dimension. Due to this 1-dimension parameterization, AAD only needs solving a small quadratic program, so AAD method is very efficient.

OSS [11]: Optimized Scale-and-Stretch (OSS) is a warping method which can retarget images into any aspect ratios without harmful visual distortions. OSS works through iteratively computing optimal local scaling factors for each localized region and then updating warped images to match these scaling factors as closely as possible. An efficient formulation for the nonlinear optimization is also developed to do the warping function computations.

Results are given below.

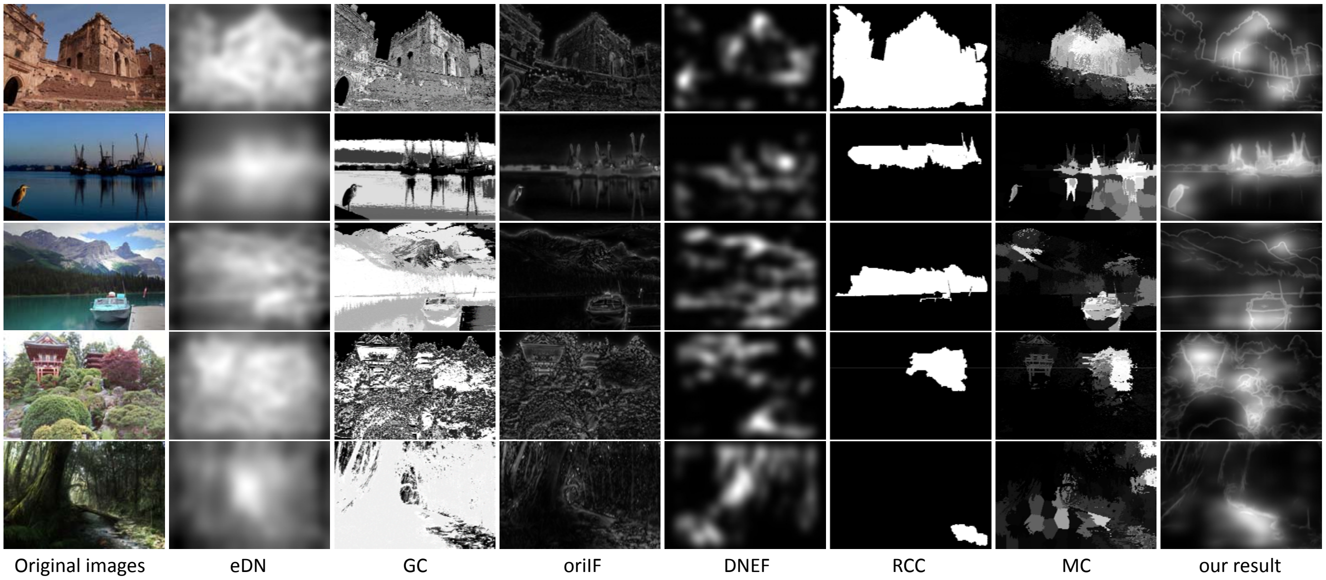

Figure 6. Comparisons between our map and 6 baseline maps. Images are selected from the category of ourdoor scene.

3. More results from Amazon Mechanical Turk

All quantitative comparisons in the paper are carried on the Amazon Mechanical Turk (AMT) [1]. Our target image and the result by a baseline are shown in randomly order to the AMT workers, who are asked to select the better one. Each pair is compared by 3 different AMT workers, and we record the numbers of votes preferring our result than the baseline. The comparison results are shown in the form of "A(B)" which means that in the total (A+B) comparisons, our method wins A times while baseline wins B times. The larger gap between A and B means more advantage.

Baseline map	Single person	Multiple people	Single object	Multiple objects	Indoor scene	Outdoor scene	All
SOAT	503(37)	497(43)	494(46)	491(49)	498(42)	502(38)	1980(180)
ISC	480(60)	479(61)	485(55)	486(54)	479(61)	481(59)	1929(231)
Multi-Op	466(74)	459(81)	422(118)	440(100)	450(90)	411(129)	1771(389)
Warp	512(28)	507(33)	511(29)	522(18)	504(36)	516(24)	2044(116)
AAD	446(94)	459(81)	438(102)	433(107)	430(110)	446(94)	1760(588)
OSS	474(66)	479(61)	474(66)	481(59)	482(58)	488(52)	1916(244)

Table 1. Comparion between our importance map and 6 baseline maps when combined with 3 different carriers.

We also collected information about the workers in our experiments on AMT. Statistical information are given below.

[1] https://www.mturk.com/mturk

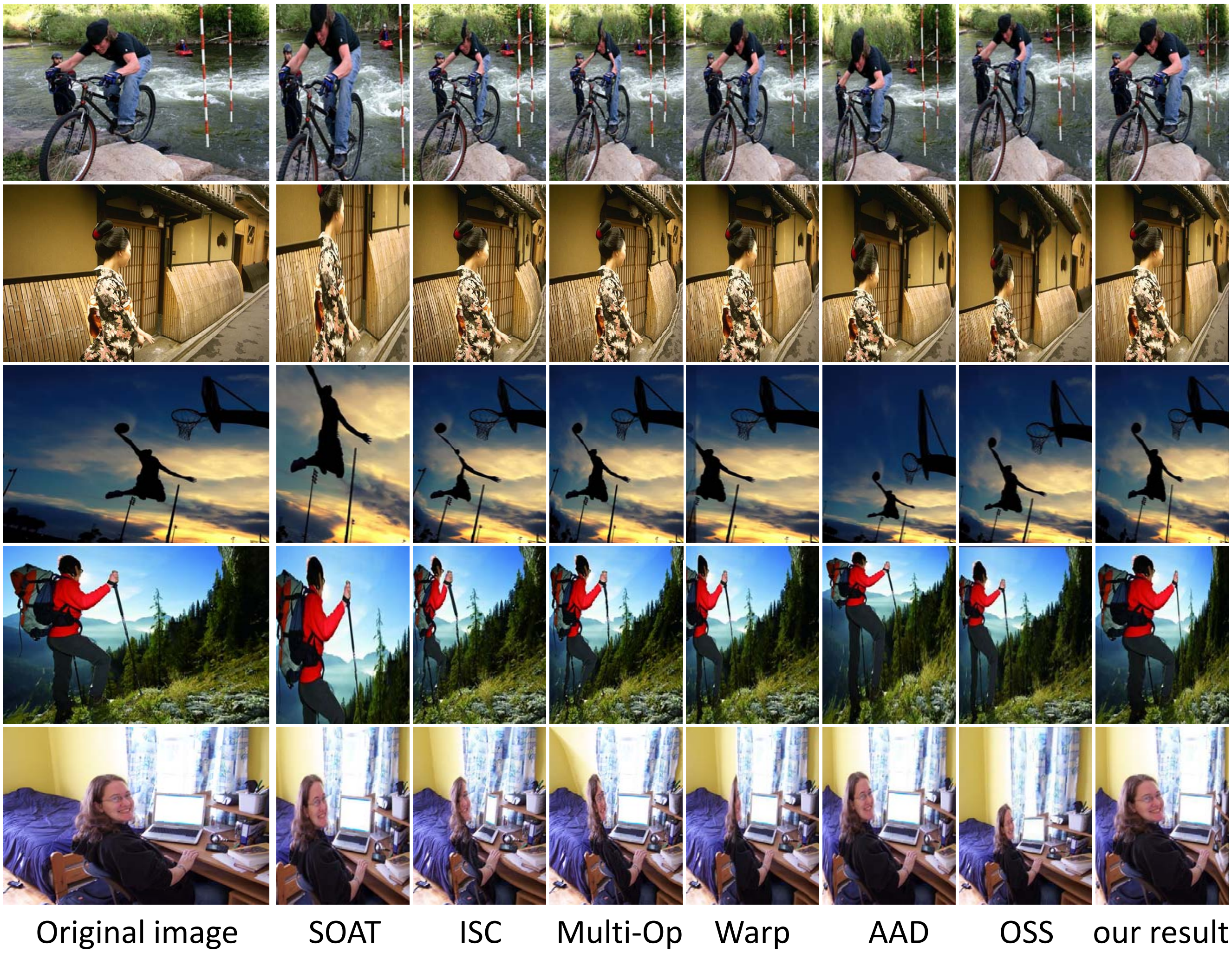

Figure 7. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of single person.

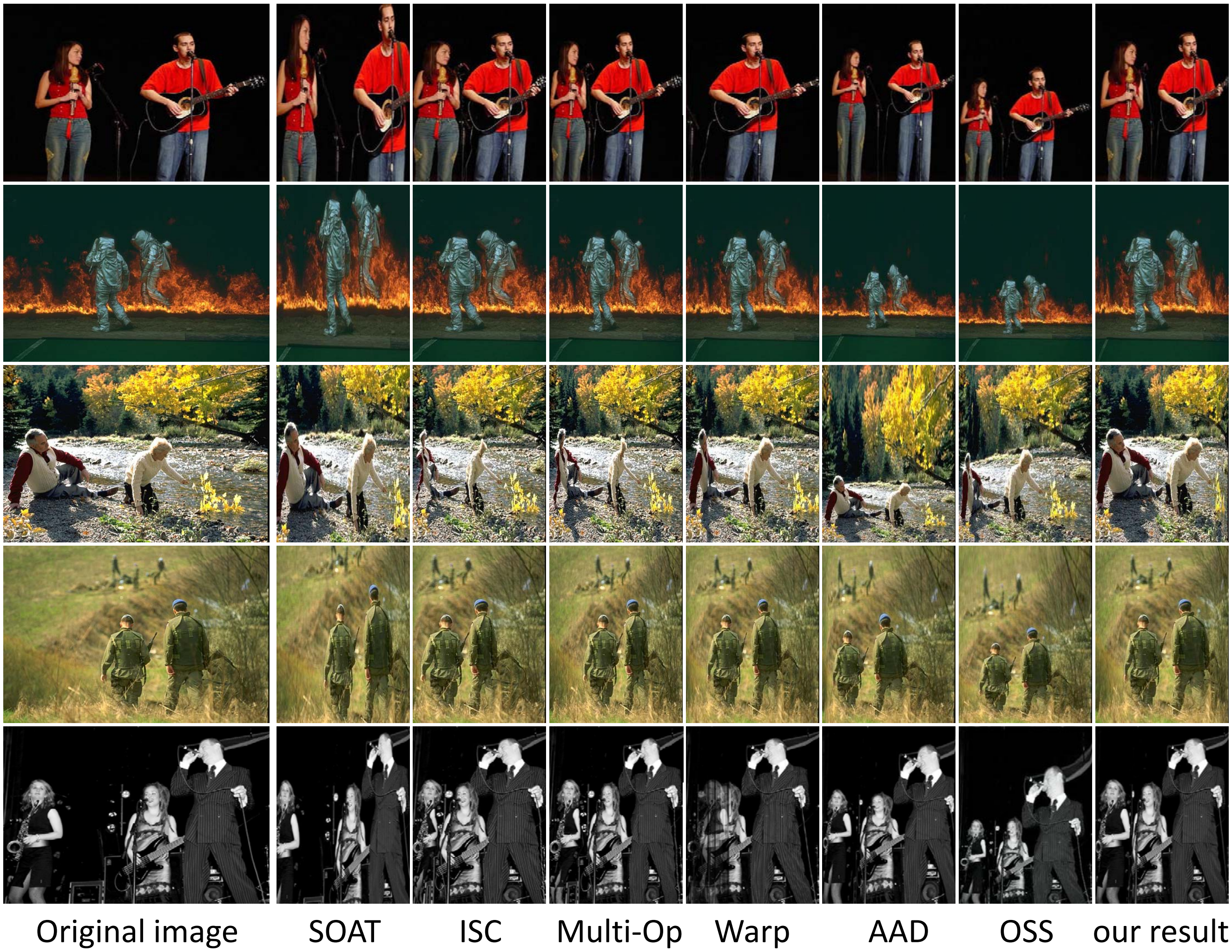

Figure 8. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of multiple people.

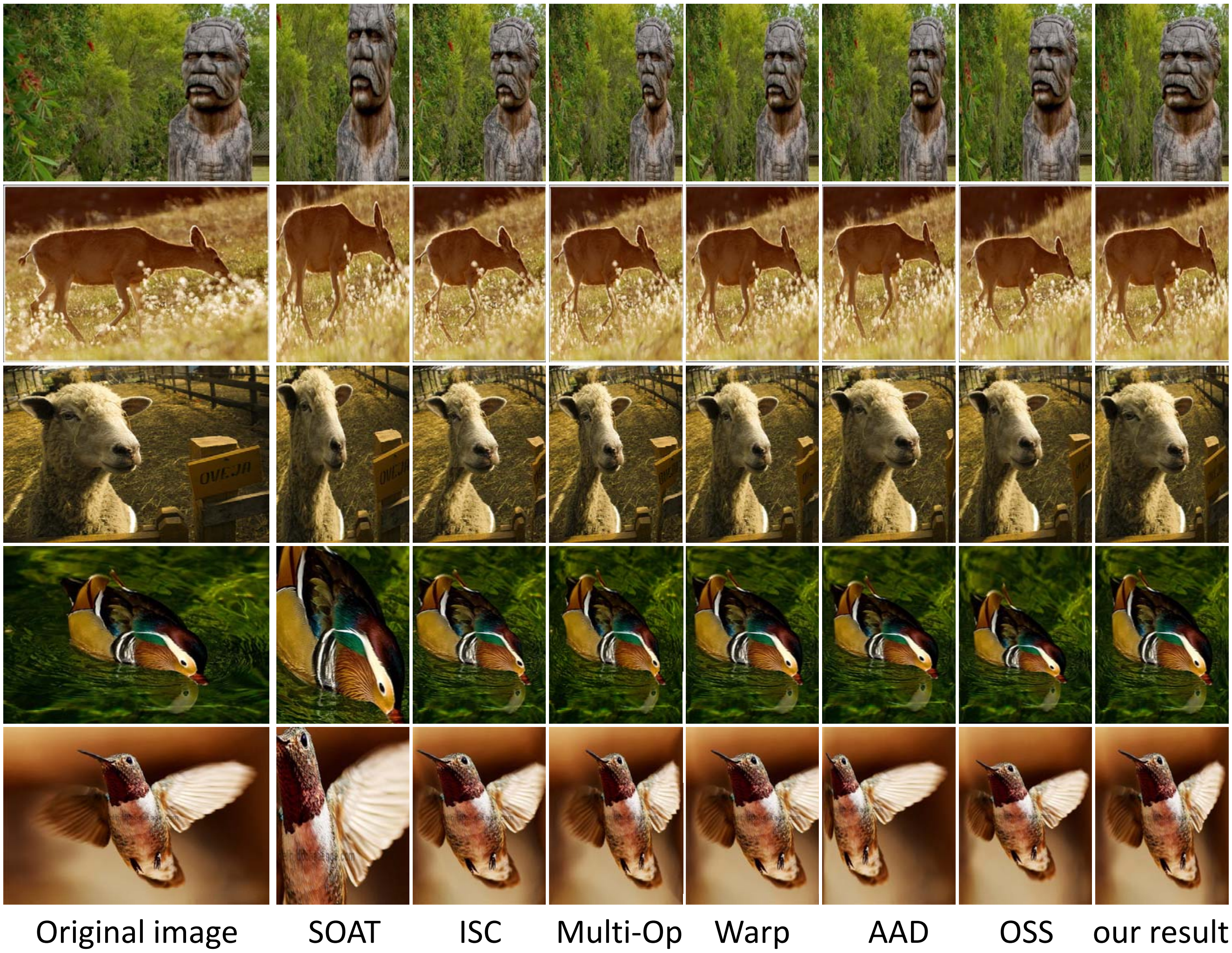

Figure 9. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of single object.

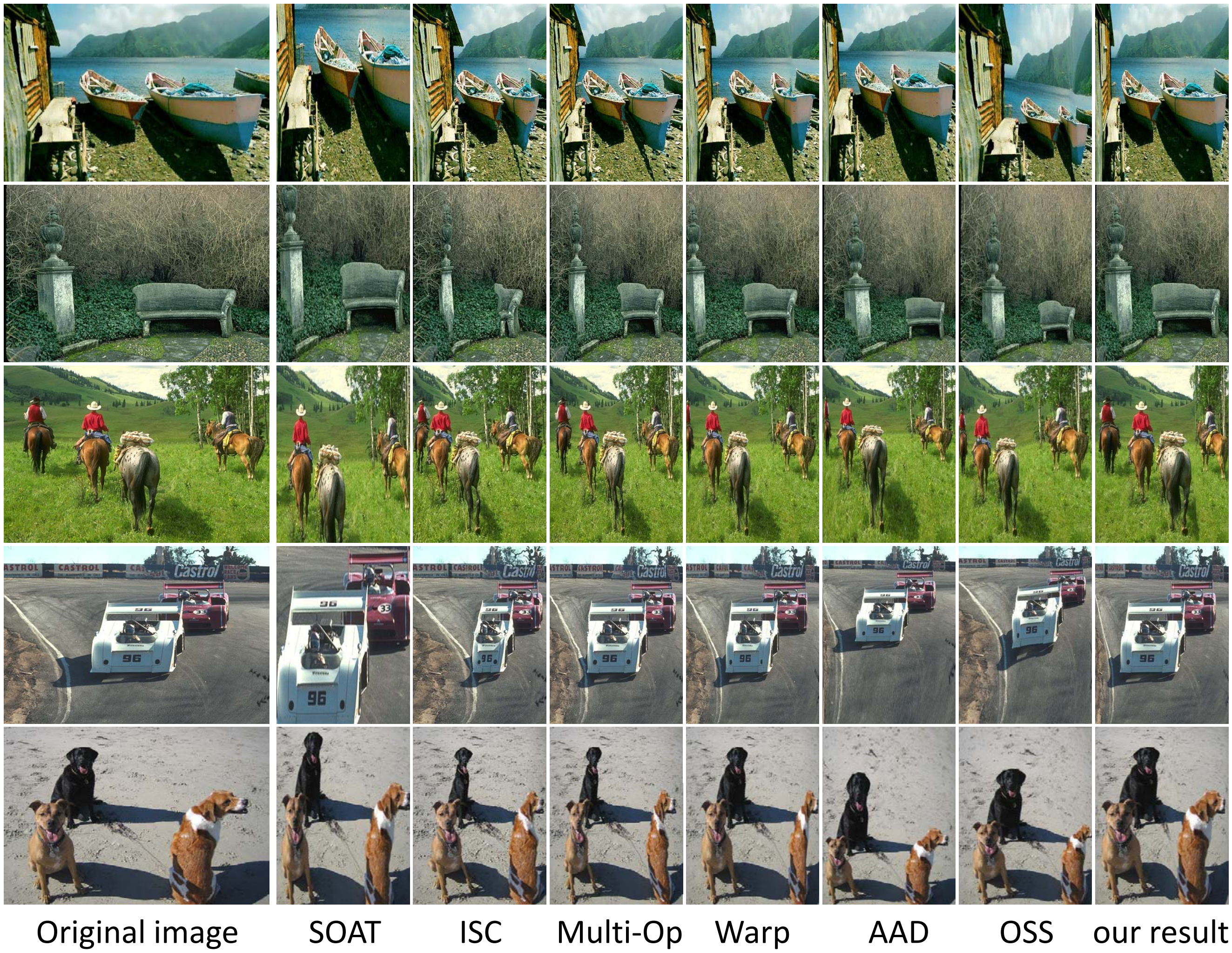

Figure 10. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of multiple objects.

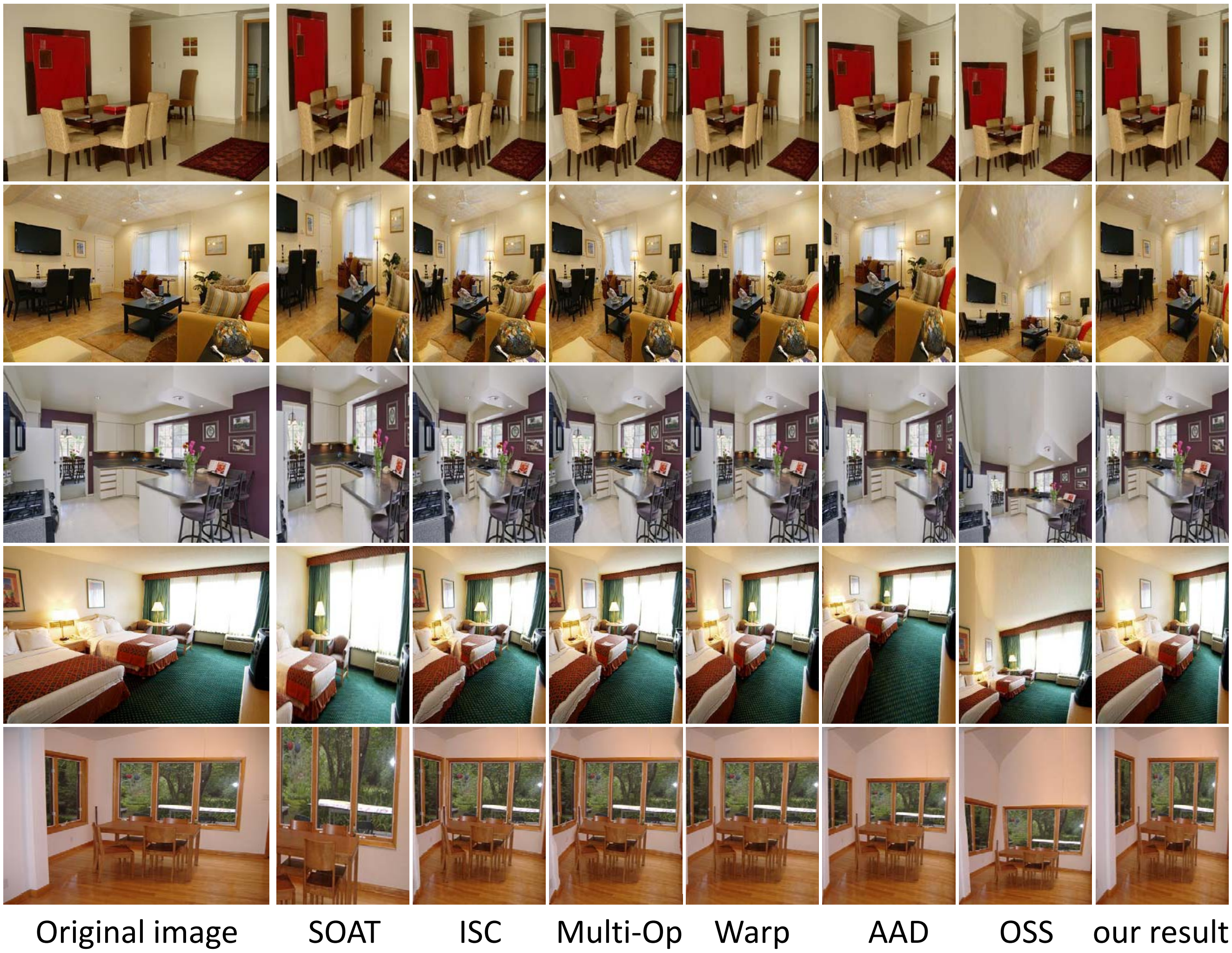

Figure 11. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of indoor scene.

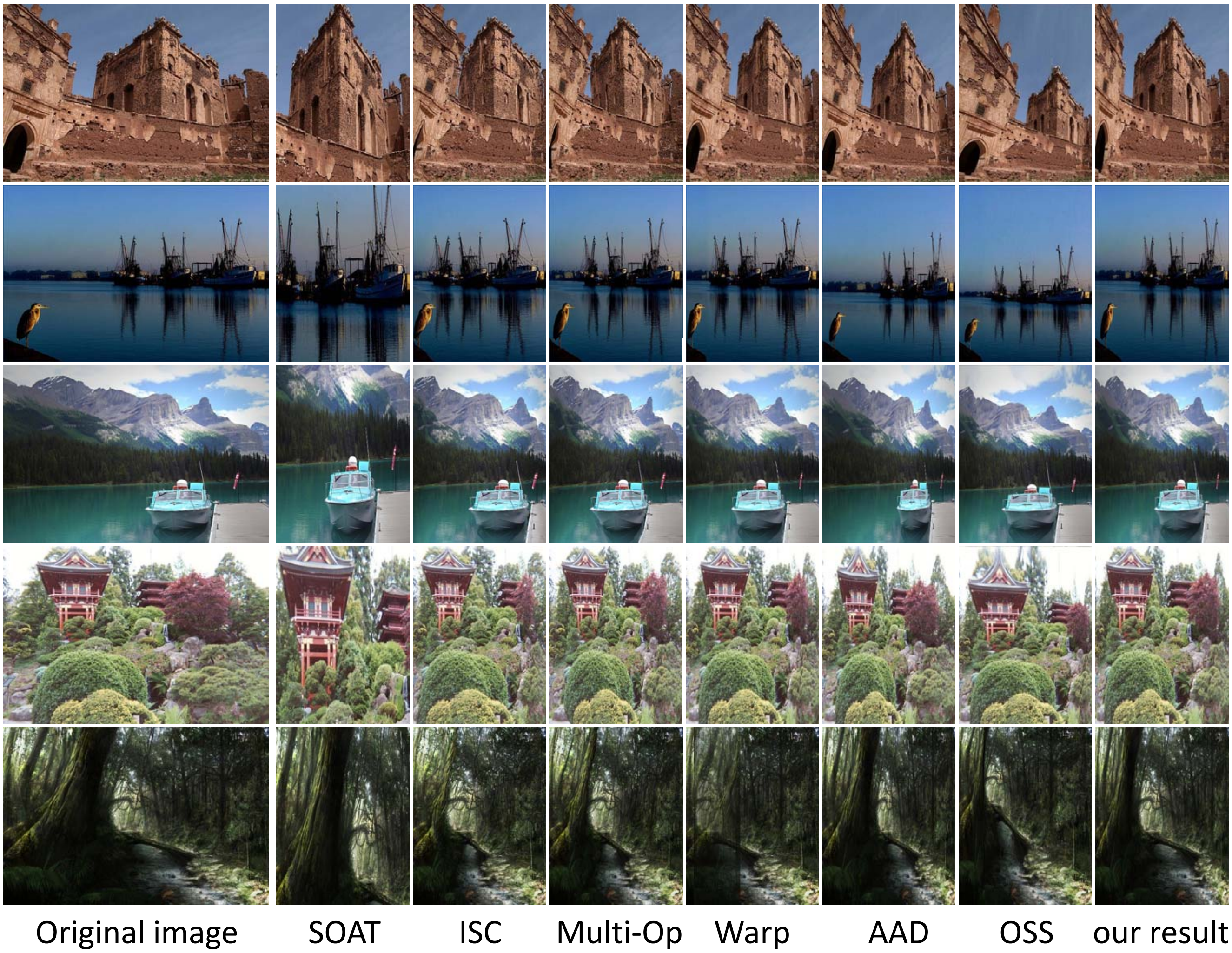

Figure 12. Comparisons between our result, which generated by our importance map and the IF carrier, and 6 baseline retargeting systems. Images are selected from the category of outdoor scene.

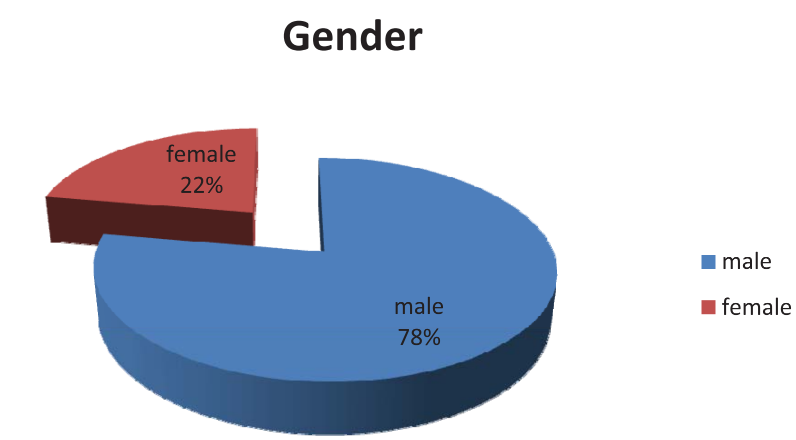

Figure 13. Statistics of workers on AMT: Gender

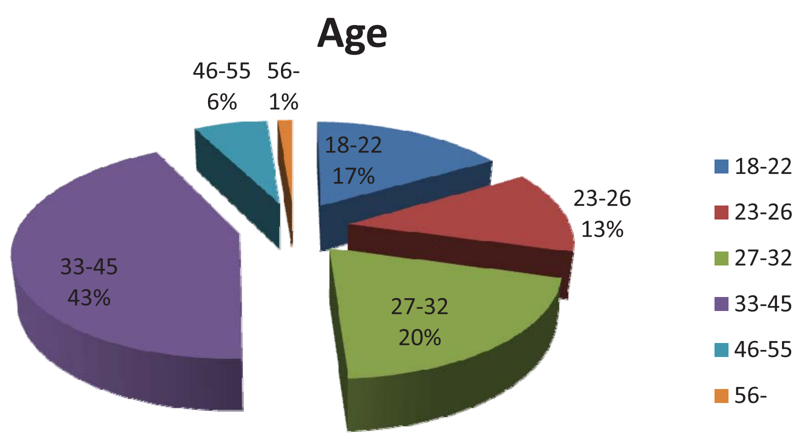

Figure 14. Statistics of workers on AMT: Age

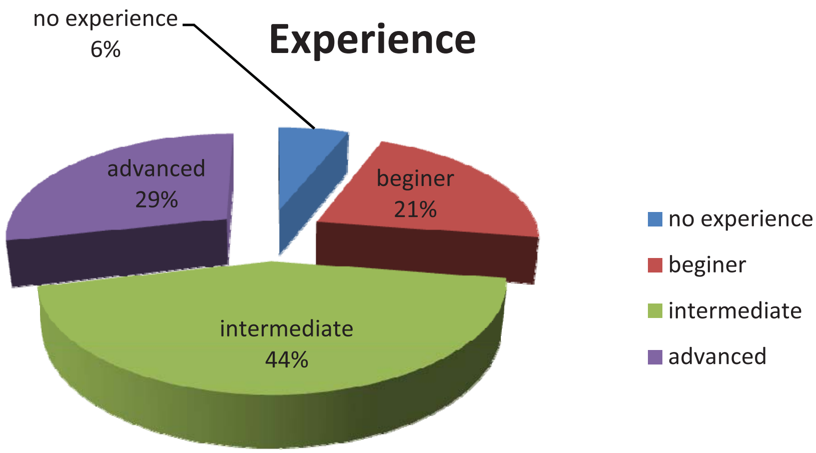

Figure 15. Statistics of workers on AMT: Experience

References

[1] S. Avidan and A. Shamir. Seam carving for content aware image resizing. In *TOG*, 2007. 2

[2] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S.-M. Hu. Global contrast based salient region detection. *TPAMI*, 2015. 1

[3] M.-M. Cheng, J. Warrell, L. Wen-Yan, S. Zheng, V. Vineet, and N. Crook. Efficient salient region detection with soft image abstraction. 2013. 1

[4] Y. Ding, J. Xiao, and J. Yu. Importance filtering for image retargeting. 2011. CVPR. 1

[5] S. Jin and L. haibin. Scale and object aware image thumbnailing. *IJCV*, 2013. 1

[6] D. Panozzo, O. Weber, and O. Sorkine. Robust image retargeting via axis-aligned deformation. In *EUROGRAPHICS*, 2012. 2

[7] M. Rubinstein, A. Shamir, and S. Avidan. Improved seam carving for video retargeting. In *TOG*, 2008. 2

[8] M. Rubinstein, A. Shamir, and S. Avidan. Multi-operator media retargeting. *TOG*, 2009. 2

[9] C. Shen, M. Song, and Q. Zhao. Learning high-level concepts by training a deep network on eye fixations. *DLUFL Workshop, in conjunction with NIPS*, 2012. 1

[10] E. Vig, M. Dorr, and D. Cox. Large-scale optimization of hierarchical features for saliency prediction in natural images. In *CVPR*, 2014. 1

[11] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee. Optimized scale-and-stretch for image resizing. *TOG*, 2008. 3

[12] L. Wolf, M. Guttmann, and D. Cohen-Or. Non-homogeneous content-driven video-retargeting. 2007. ICCV. 2

[13] R. Zhao, W. Ouyang, H. Li, and X. Wang. Saliency detection by multi-context deep learning. In *CVPR*, 2015. 1