

KOALAnet: Blind Super-Resolution using Kernel-Oriented Adaptive Local Adjustment – Supplementary Material –

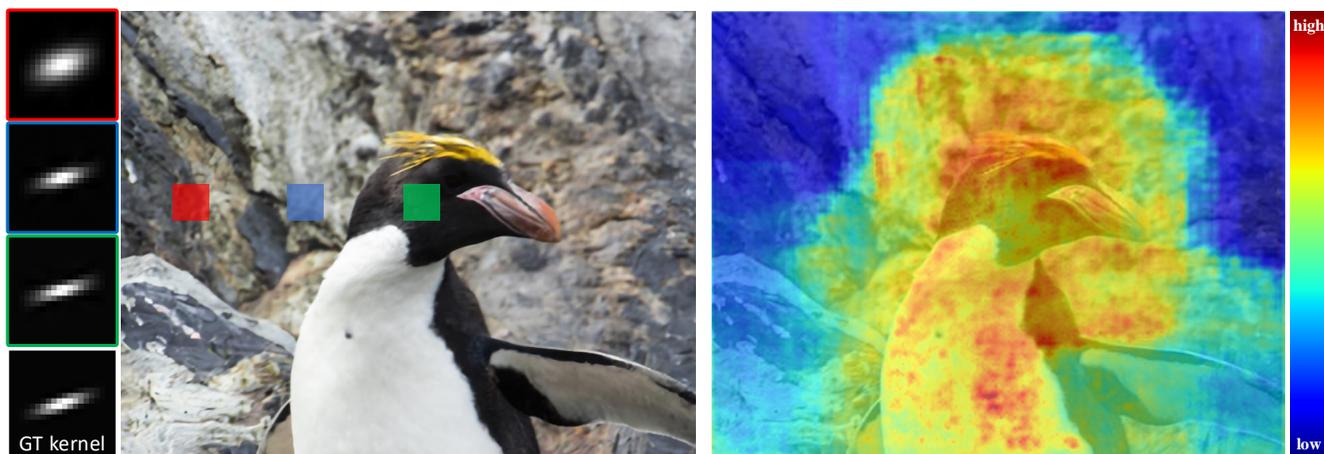
Soo Ye Kim*

Hyeonjun Sim*

Munchurl Kim[†]

Korea Advanced Institute of Science and Technology

{sooyekim, flhy5836, mkimee}@kaist.ac.kr



(a) Estimated degradation kernels at different spatial locations of the low resolution image.

(b) Visualization of the cosine similarity of the estimated per-pixel degradation kernels and the ground truth kernel.

Figure 1: Degradation kernel visualization on different spatial locations of the same image. In this example, the penguin is the object in focus and the background is blurry (out of focus). The estimated degradation kernels differ depending on the spatial location as shown in (a). The cosine similarity of each per-pixel kernel is visualized in (b). Accurate degradation kernels are estimated near the focused region, especially in the boundary areas between the in-focus and out-of-focus areas.

1. Additional Kernel Analyses

1.1. Spatially-variant Kernel Visualizations

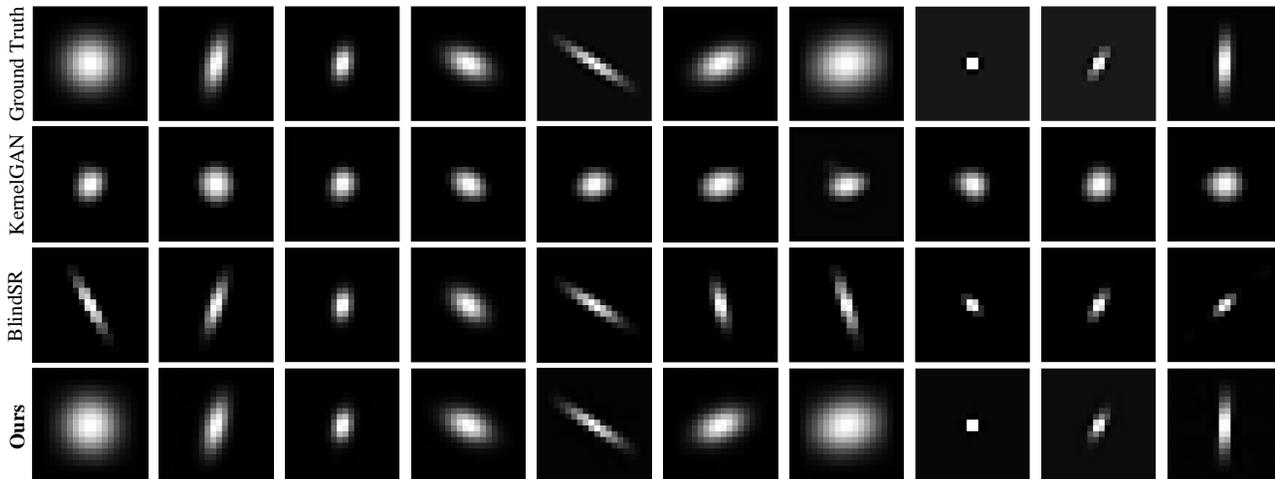
Photography enthusiasts tend to take pictures with intentionally blurry (out-of-focus) areas, in order to emphasize objects or regions of interest in the depth-of-field (DoF) by controlling the aperture size or the focal length of the camera lens to focus on the areas of interest. In the main paper, we showed that existing blind super-resolution (SR) methods tend to generate over-sharpened or blurry results for these types of artistic images (Fig. 1 in main paper). We further showed that a vanilla SR network that does not consider the

degradation information over-sharpens even the *intentionally* out-of-focus area, resulting in images with a deeper DoF (Fig. 5 in main paper). The over-sharpening tends to happen in the boundary regions between in-focus and completely out-of-focus areas. This is an important observation that was not previously dealt with in literature, which can be useful in handling images with intentionally blurry regions.

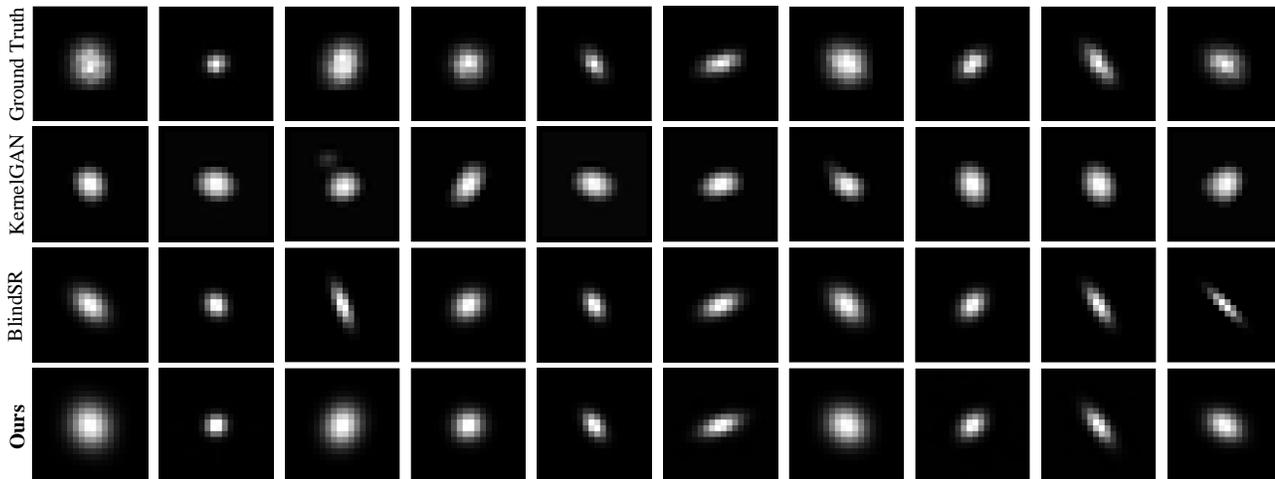
To further analyze this aspect, we show the spatially-variant kernel estimations in an image mixed with in-focus and out-of-focus areas in Fig. 1. In 1 (a), the estimated degradation kernel in the *green box*, which is the in-focus area where the high frequency details must be restored, is highly similar to the ground truth degradation kernel. Comparing the *red box* and *blue box* in the smooth area, more accurate kernels are predicted in the region closer to the penguin (*blue*

*Both authors contributed equally to this work.

[†]Corresponding author.



(a) Comparison of estimated degradation kernels on DIV2K-val



(b) Comparison of estimated degradation kernels on DIV2KRRK[1]

Figure 2: Visualizations of ground truth kernels and estimated degradation kernels by KernelGAN[1], BlindSR[2] and Ours on two datasets, (a) DIV2K-val and (b) DIV2KRRK [1]. The downsampling network of our KOALANet is able to predict accurate degradation kernels on DIV2KRRK as well as DIV2K-val.

box). In Fig. 1 (b), we have visualized the cosine similarity map between the vectorized predicted per-pixel kernels and the ground truth blur kernel (*red* denotes high similarity and *blue* denotes low similarity). It can be clearly seen that the predicted degradation kernels are indeed *spatially-variant* depending on the location in the image. Also, thanks to the large receptive field of our U-Net-based downsampling network, accurate blur kernels are predicted even in the smooth regions near the in-focus area (Fig. 1 (b)). This helps to effectively handle the boundary areas between in-focus and completely out-of-focus areas so that these regions are not over-sharpened after SR. Note that conventional SR methods are unable to disentangle the degradation blur and the intended blur, and thus generates over-sharpened results in

boundary regions (Fig. 1 in main paper).

1.2. Comparison of Estimated Degradation Kernels

Details on the kernel accuracy measurement. In Table 4 of the main paper, we compared the estimation accuracy of the degradation kernels generated by KernelGAN [1], BlindSR [2] and the downsampling network of our KOALANet on the DIV2K-val testset with random anisotropic Gaussian degradations, and DIV2KRRK [1]. For each method, the degradation kernels were estimated from the input LR images degraded via the corresponding ground truth kernels. BlindSR[2] predicts three values, standard deviations σ_{11} and σ_{22} and the rotation angle θ , of a bivariate Gaussian distribution to parametrize a Gaussian kernel.

Since the peak of the estimated Gaussian kernel is located at the center pixel due to the size of the kernel (15×15) being an odd number in the original implementation provided by the authors of BlindSR [2], we calibrated the estimated kernels by convolving them with a 2×2 mean filter with 0.25 values. This yielded lower l_2 error. Furthermore, since the center of the ground truth kernels, k_{gt} , and the estimated kernels, k , can be different, we shifted k and k_{gt} in x and y directions to find the minimum l_2 error, as follows:

$$l_2 \text{ error} = \min_{\Delta h, \Delta w} \sum_{h,w} \|k_{gt}(h, w) - k(h + \Delta h, w + \Delta w)\|^2, \quad (1)$$

where h, w represent the locations in x, y dimensions and $\Delta h, \Delta w$ is the shift along the x, y dimensions.

Visualization of the estimated kernels. For the qualitative comparison of the estimated kernels, we visualized some examples of the ground truth kernels and the estimated kernels by KernelGAN [1], BlindSR [2] and the downsampling network of our KOALAnet on the DIV2K-val testset and DIV2KRRK [1] in Fig. 2. As shown, our KOALAnet robustly estimates the latent degradation kernels from the LR images compared to other methods, even in DIV2KRRK [1] where noise is injected to the Gaussian kernels.

2. Details of Complexity Evaluation

2.1. KOALAnet

Our proposed framework is implemented on Python 3.6 with Tensorflow 1.13, and we used an NVIDIA Titan RTX for our experiments. The total number of filter parameters is 6.09M and 6.45M for the KOALAnet of $s = 2$ and $s = 4$, respectively.

2.2. Comparison to Existing Methods

We provided a comparison on computational complexity in terms of the inference time and GFLOPs with other blind SR methods [1, 2, 3, 4] in Table 1 of the main paper. For all methods, GFLOPs is calculated only for the feed-forward paths. Furthermore, since ZSSR [4], KernelGAN [1] and BlindSR [2] are optimization-based methods, we take the number of iterations into consideration when computing the GFLOPs. The number of iterations needed for the optimization of ZSSR [4] can be different at each run even on the same test image (e.g., “baby” in Set5), while the numbers of iterations of KernelGAN [1] and BlindSR [2] are fixed. Thus for ZSSR, we obtained the average number of iterations from five repetitions each for scale factors 2 and 4, and then computed the GFLOPs using those numbers.

3. Additional Qualitative Results

We provide additional qualitative results on the random anisotropic Gaussian testsets in Fig. 3 and 4 for scale factor

2, and Fig. 5 and Fig. 6 for scale factor 4. The estimated degradation kernels are shown at the bottom right or the bottom left corner for all applicable methods that estimate the degradation kernel. Since IKC [3] estimates the degradation kernels in a lower dimensionality (after PCA), they cannot be visualized along with the other degradation kernels estimated from BlindSR [2], KernelGAN [1] or Ours, from which actual degradation kernels can be generated. The ground truth degradation kernels are also shown at the bottom right or the bottom left corner of images or patches denoted as Ground Truth. In Fig. 3 and Fig. 5, we show full images for comparisons on Set5, Set14 and BSD100, which have relatively low resolutions. In Fig. 4 and Fig. 6, we crop the SR results to better visualize the difference for the readers for comparisons on DIV2K, Urban100 and Manga109, which have higher resolutions of near 2K. For IKC [3], we visualized the results yielding the best PSNR performance among seven iterations (IKC-max). As shown, our KOALAnet produces accurate SR results with sharper edges and realistic textures on various datasets and degradation kernels, even in examples with very high frequency regions.

References

- [1] Sefi Bell-Kligler, Assaf Shocher, and Michal Irani. Blind super-resolution kernel estimation using an internal-gan. In *Advances in Neural Information Processing Systems*, pages 284–293, 2019.
- [2] Victor Cornillère, Abdelaziz Djelouah, Wang Yifan, Olga Sorkine-Hornung, and Christopher Schroers. Blind image super-resolution with spatially variant degradations. *ACM Transactions on Graphics*, 38(6):166.1–166.13, 2019.
- [3] Jinjin Gu, Hannan Lu, Wangmeng Zuo, and Chao Dong. Blind super-resolution with iterative kernel correction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1604–1613, 2019.
- [4] Assaf Shocher, Nadav Cohen, and Michal Irani. “zero-shot” super-resolution using deep internal learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3118–3126, June 2018.

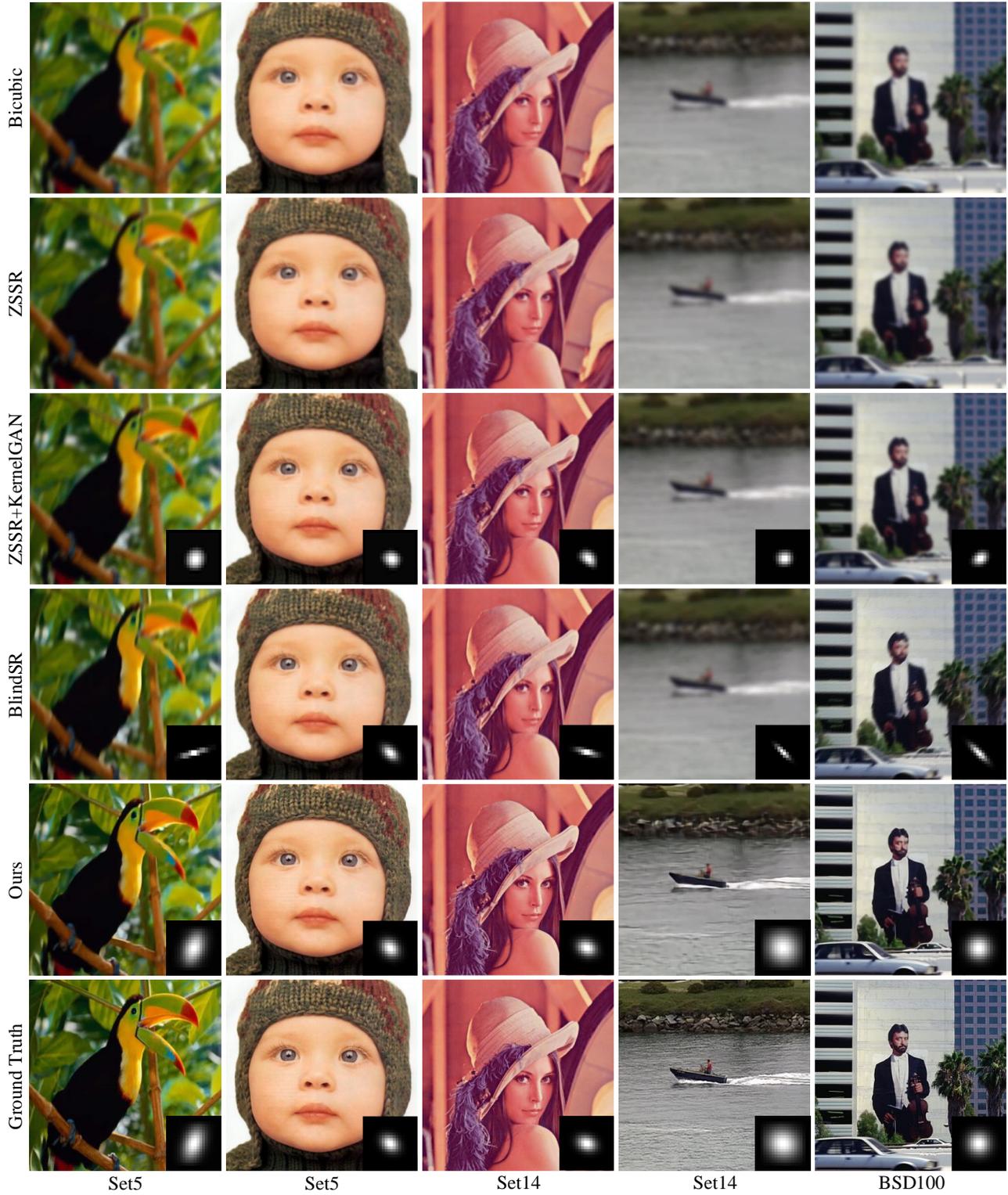


Figure 3: Additional qualitative comparison with ZSSR [4], ZSSR+KernelGAN [1] and BlindSR [2] for scale factor 2 on Set5, Set14 and BSD100 datasets. The estimated (or ground truth) degradation kernels are placed on the bottom right corner for all applicable methods that estimate the degradation kernel. Our KOALAnet is able to predict accurate degradation kernels and generate sharp SR results on various datasets and degradation kernels.

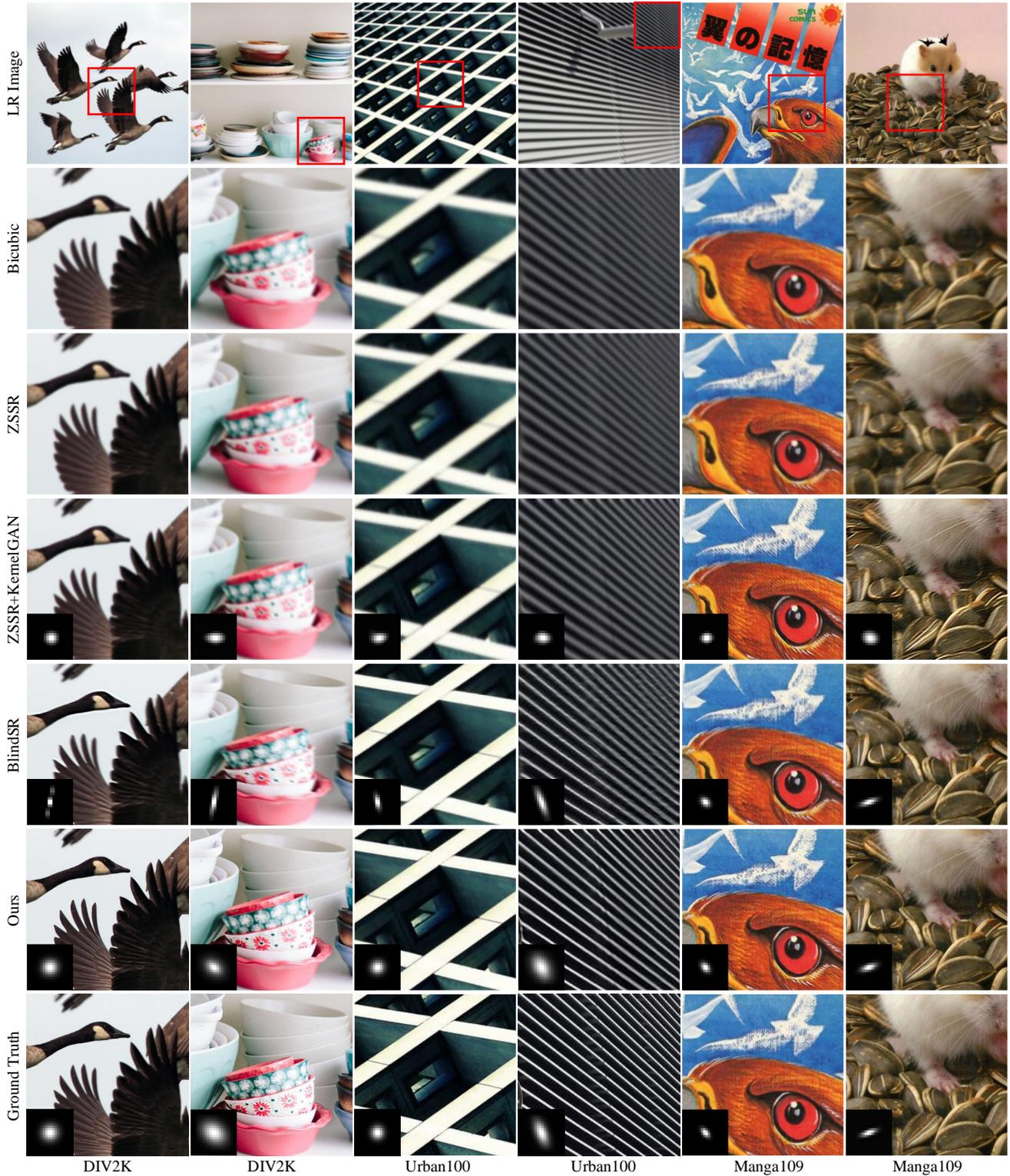


Figure 4: Additional qualitative comparison with ZSSR [4], ZSSR+KernelGAN [1] and BlindSR [2] for scale factor 2 on DIV2K, Urban100 and Manga109 datasets. The estimated (or ground truth) degradation kernels are placed on the bottom left corner for all applicable methods that estimate kernel information. Our KOALANet is able to predict accurate degradation kernels and generate sharp SR results on various datasets and degradation kernels.

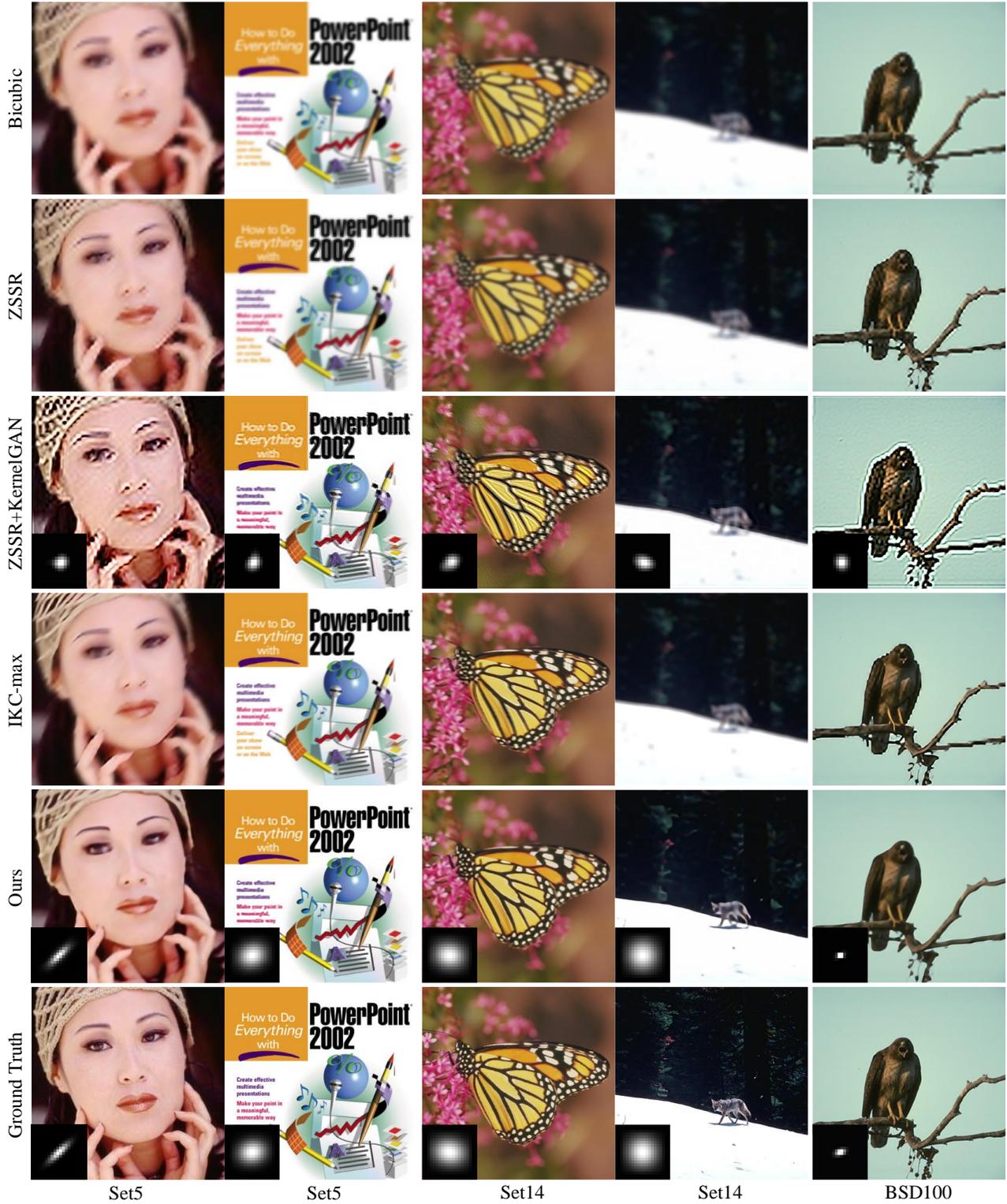


Figure 5: Additional qualitative comparison with ZSSR [4], ZSSR+KernelGAN [1] and IKC [3] for scale factor 4 on Set5, Set14 and BSD100 datasets. The estimated (or ground truth) degradation kernels are placed on the bottom left corner for all applicable methods that estimate the degradation kernel. We show the results yielding the best PSNR among seven iterations for IKC. Our KOALAnet is able to predict accurate degradation kernels and generate sharp SR results on various datasets and degradation kernels.

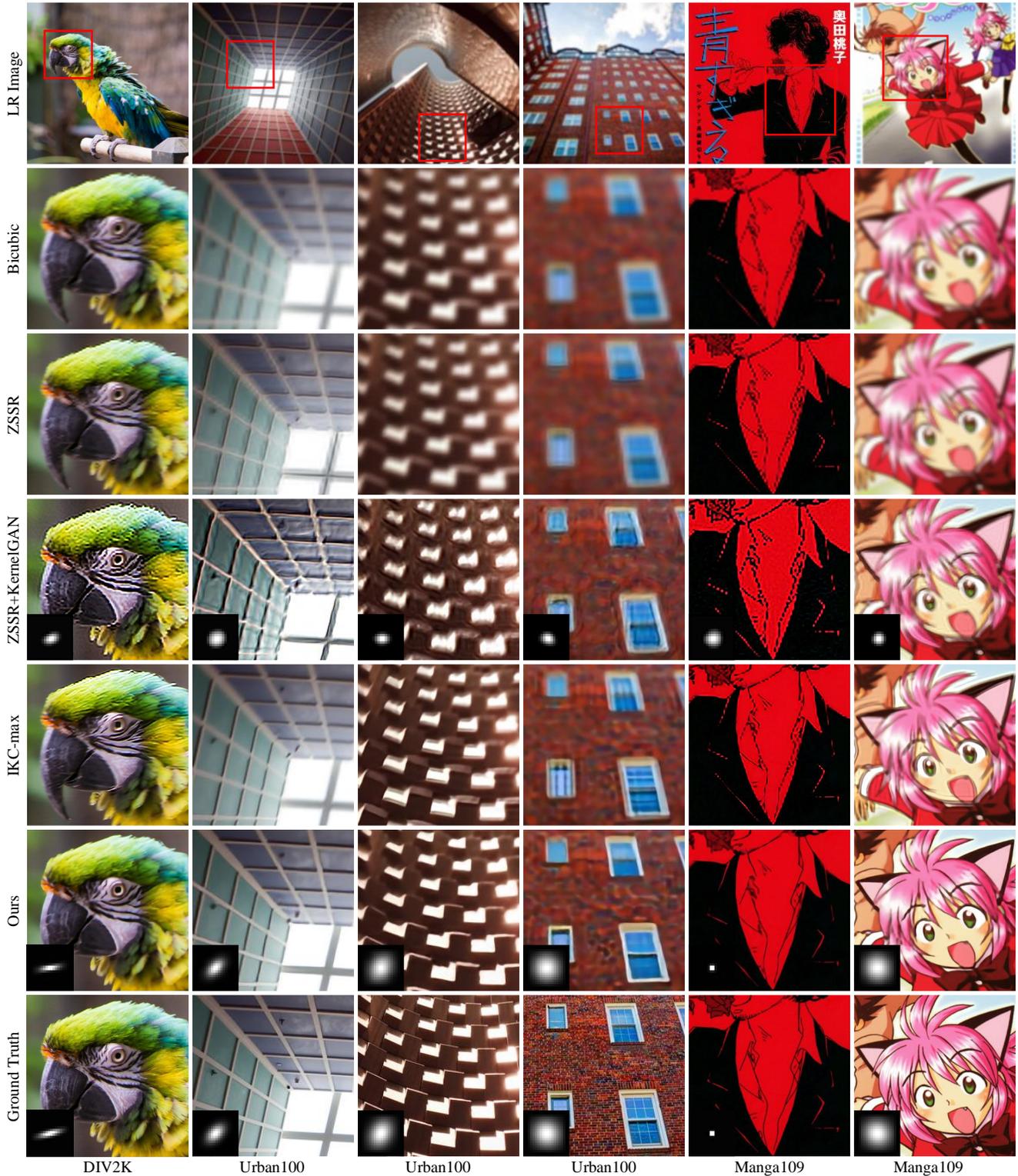


Figure 6: Additional qualitative comparison with ZSSR [4], ZSSR+KernelGAN [1] and IKC [3] for scale factor 4 on DIV2K, Urban100 and Manga109 datasets. The estimated (or ground truth) degradation kernels are placed on the bottom left corner for all applicable methods that estimate the degradation kernel. We show the results yielding the best PSNR among seven iterations for IKC. Our KOALAnet is able to predict accurate degradation kernels and generate sharp SR results on various datasets and degradation kernels.