

Supplementary Materials: Suppressing Uncertainties in Degradation Estimation for Blind Super-Resolution

Anonymous Authors

1 OVERVIEW

In this supplementary material, we provide additional experimental explanations and details of USR organized as follows:

- In Section 2, we present more implementation details, corresponding to the **Implementation details** of Section 4.1;
- In Section 3, we introduce the datasets that the comparative experiments are based on, corresponding to the **Testing Datasets** of Section 4.1;
- In Section 4, we offer more visual qualitative comparisons, corresponding to the **Qualitative Comparisons** of Section 4.2.

2 IMPLEMENTATION DETAILS

Real-ESRGAN workflow parameters. The classic degradation models cannot simulate some complex degradation issues, particularly unknown noise and complex artifacts, because the synthesized low-resolution images still have a significant gap from the real degraded images. Therefore, Real-ESRGAN [6] extends the classic degradation models to higher-order processes to simulate more realistic degradation. The so-called higher-order degradation model, in layman’s terms, involves arranging and combining classic degradation algorithms. Real-ESRGAN categorizes the degradation algorithms into four types: Blur, Resize, Noise, and JPEG Compression. Below are the specific parameters used in the tests in our work :

```
blur_kernel_size: 13
kernel_list: ['iso', 'aniso',
              'generalized_iso', 'generalized_aniso',
              'plateau_iso', 'plateau_aniso']
kernel_prob: [0.60, 0.40, 0.0, 0.0, 0.0, 0.0]
sinc_prob: 0.1
blur_sigma: [0.2, 0.8]
betag_range: [1.0, 1.5]
betap_range: [1, 1.2]

blur_kernel_size2: 7
kernel_list2: ['iso', 'aniso',
               'generalized_iso', 'generalized_aniso',
               'plateau_iso', 'plateau_aniso']
kernel_prob2: [0.60, 0.4, 0.0, 0.0, 0.0, 0.0]
sinc_prob2: 0.0
blur_sigma2: [0.2, 0.5]
betag_range2: [0.5, 0.8]
betap_range2: [1, 1.2]

final_sinc_prob: 0.2

gt_size: 768
crop_pad_size: 300
use_hflip: False
```

```
use_rot: False
rescale_gt: True
```

```
degradation:
  sf: 4
  # the first degradation process
  resize_prob: [0.2, 0.7, 0.1] # up, down, keep
  resize_range: [0.5, 1.5]
  gaussian_noise_prob: 0.5
  noise_range: [1, 15]
  poisson_scale_range: [0.05, 0.5]
  gray_noise_prob: 0.4
  jpeg_range: [65, 95]

  # the second degradation process
  second_order_prob: 0.0
  second_blur_prob: 0.2
  resize_prob2: [0.3, 0.4, 0.3] # up, down, keep
  resize_range2: [0.8, 1.2]
  gaussian_noise_prob2: 0.5
  noise_range2: [1, 10]
  poisson_scale_range2: [0.05, 0.2]
  gray_noise_prob2: 0.4
  jpeg_range2: [75, 100]
```

Our work is implemented based on [7] and will be made open source upon acceptance of the paper.

3 DATASETS

DIV2K dataset [1] is a popular benchmarking dataset used primarily for tasks related to single image super-resolution (SISR). This dataset was first introduced in the NTIRE challenge at CVPR 2017, focusing on the problem of image upsampling, where the goal is to enhance the resolution of a low-resolution image. DIV2K subdivided into four distinct subsets: a training set (800 images), a validation set (100 images), a test set (100 images), and a challenge set (1,000 images). These images feature multiple resolutions, enabling researchers to experiment and evaluate super-resolution algorithms at different resolutions. The DIV2K dataset was designed to advance the technology of super-resolution image reconstruction. Due to its high quality and diversity, it has become the preferred dataset for many research projects and competitions.

BSDS [2] is a dataset widely used for image segmentation and edge detection research, developed by researchers at the University of California, Berkeley. It includes natural images of various scenes, such as landscapes, animals, and buildings, reflecting the diversity of the real world. The BSDS dataset is used not only for evaluating edge detection and image segmentation algorithms but also widely in other studies in computer vision and image processing, such as image recognition and super-resolution. Due to its high-quality images and annotations, and standardized evaluation methods, the

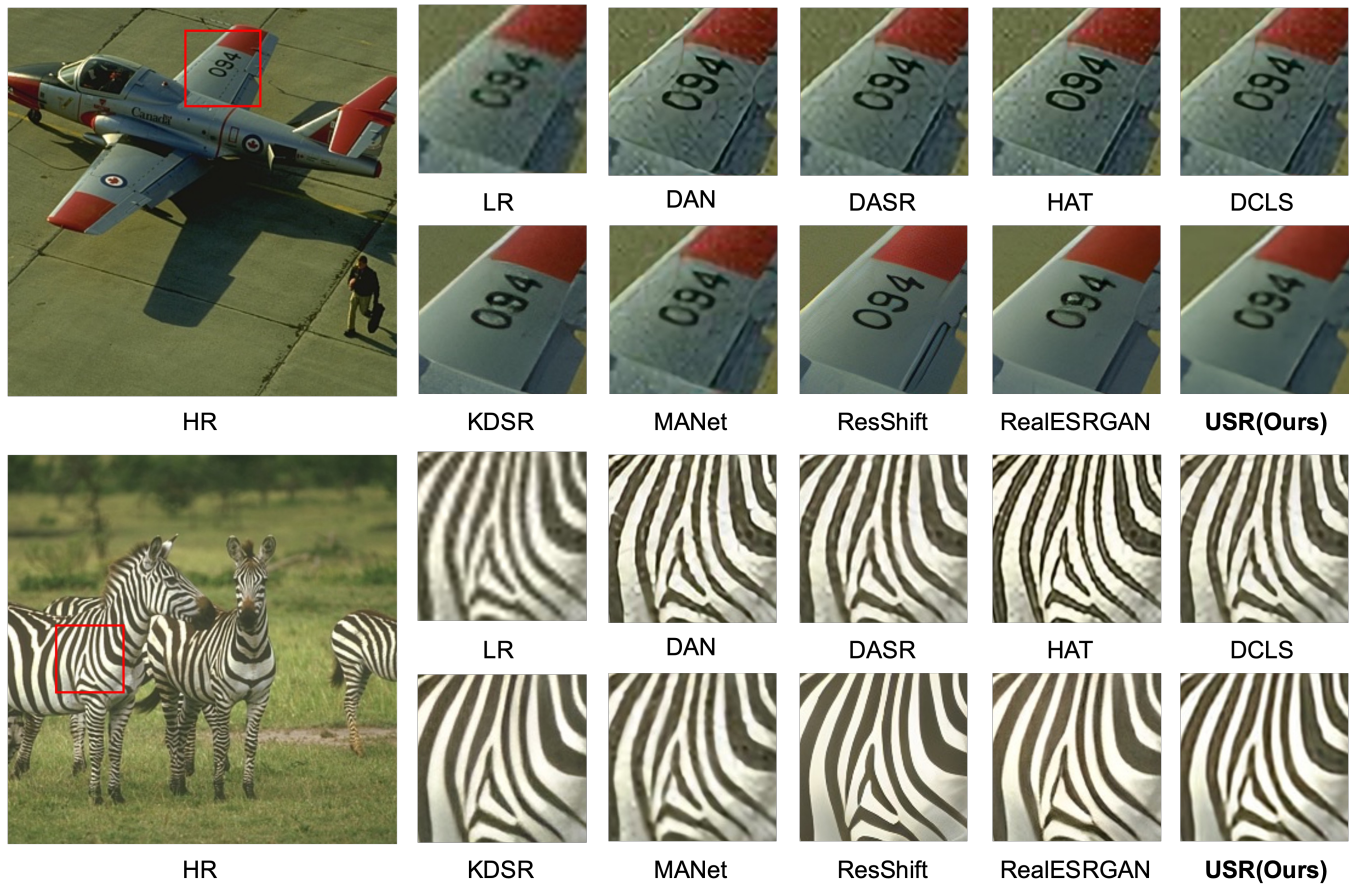


Figure 1: Visual comparisons of several representative methods on examples of the BSDS dataset. The image above is a case of airplane, while the image below depicts a case of zebra.

BSDS dataset has become an important benchmark in the field of image processing.

Urban100 [3] dataset is specifically used for image super-resolution research, focusing particularly on images in urban environments. It aims to provide a collection of urban scene images with complex structures and rich details for evaluating and improving image super-resolution algorithms. The dataset includes 100 high-resolution images of urban scenes, featuring buildings, streets, vehicles, windows, billboards, and other urban characteristics, with high structural complexity and rich textural details. Its focus on urban environments and high-quality images make Urban100 one of the frequently used datasets in super-resolution research, especially in assessing algorithms' effectiveness in handling complex structures and details.

T91 [10] dataset plays a significant role in the development and evaluation of super-resolution algorithms in image super-resolution research. Composed of 91 diverse images, it is used for training and testing super-resolution algorithms. Despite the limited number of images, the dataset includes various types of images, such as natural landscapes. This diversity makes the dataset suitable for evaluating

the performance of super-resolution algorithms on different types of images.

DPED [4] is a dataset designed specifically for research on enhancing the quality of mobile photography. It aims to assist in improving and evaluating image enhancement algorithms, particularly those designed to enhance the quality of photos taken with mobile phone cameras. DPED includes images from different smart-phone cameras, often paired with images of the same scenes taken with high-quality reference cameras. By providing real-world image samples and high-quality reference images, it fosters technological advancements in this area.

DRealSR [8] is designed to enhance the development of image super-resolution algorithms under realistic conditions. Unlike traditional datasets that generate low-resolution images through synthetic methods like bicubic downscaling, DRealSR offers images captured in natural settings, providing a true test of real-world imaging challenges. This dataset includes pairs of low and high-resolution images, captured using different camera settings to mimic optical and sensor-based imperfections typically absent in synthetic datasets. DRealSR covers a wide range of scenes and subjects, making it ideal for training robust super-resolution models capable

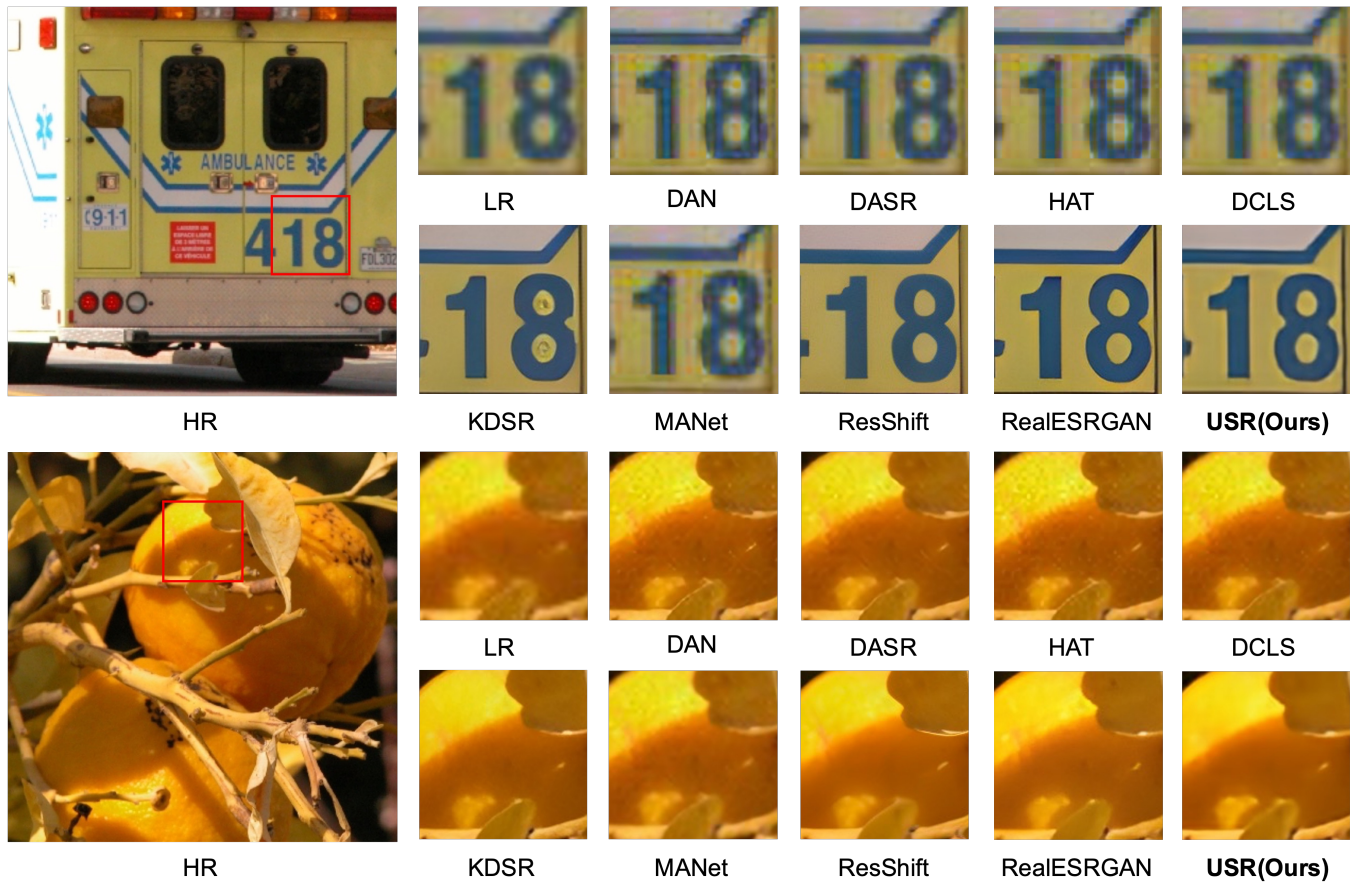


Figure 2: Visual comparisons of several representative methods on examples of the T91 dataset. The image above is a case of ambulance, while the image below depicts a case of lemon tree.

of performing across diverse real-world scenarios. Its realistic approach helps bridge the gap between academic research and practical applications in fields such as digital photography, surveillance, and medical imaging.

4 QUALITATIVE COMPARISONS

Comparison on BSDS dataset. In Figure 1, we particularly focused on two visual cases: airplane and zebra images. The quality of detail restoration in magnified regions is crucial for evaluating super-resolution methods.

Firstly, examining the magnified area of the airplane image, the superiority of USR in detail restoration is evident. The numbers “094” on the airplane are exceptionally clear when restored using the USR method, exhibiting well-defined edges and high contrast. Compared to traditional super-resolution techniques, USR performs better in maintaining the original shape and clarity of the edges, which often suffer from excessive smoothing in other methods. Moreover, the airplane’s color restoration is both natural and accurate, demonstrating USR’s capability in color fidelity. This is

significant because many algorithms tend to introduce color distortions while enhancing image resolution, resulting in restored images that look unnatural or overly processed.

In the zebra image, USR’s exceptional performance is evident once more. The stripes of a zebra are a challenging feature because their detail and high-contrast edges demand precise restoration of sharp lines while avoiding artificial artifacts. Here, USR significantly outperforms other methods by generating clear, sharp edges of stripes while maintaining the natural contrast and spatial relationships between them. It avoids over-smoothing or blurring the stripes, which in some algorithms could strip the image of its natural texture. This capability to restore zebra stripes highlights USR’s strength in handling images with complex patterns and high-frequency details.

Comparison on T91 dataset. In Figure 2, we observe the effects of super-resolution reconstruction methods in two distinct scenes: the rear of an ambulance and a close-up of a lemon tree.

In the example of the ambulance, the focus is on the vehicle’s tail number “418”. The USR (Ours) method excels in restoring the clarity and readability of the digits, showing less blurring and distortion. In contrast, results from other methods display softer edges and a

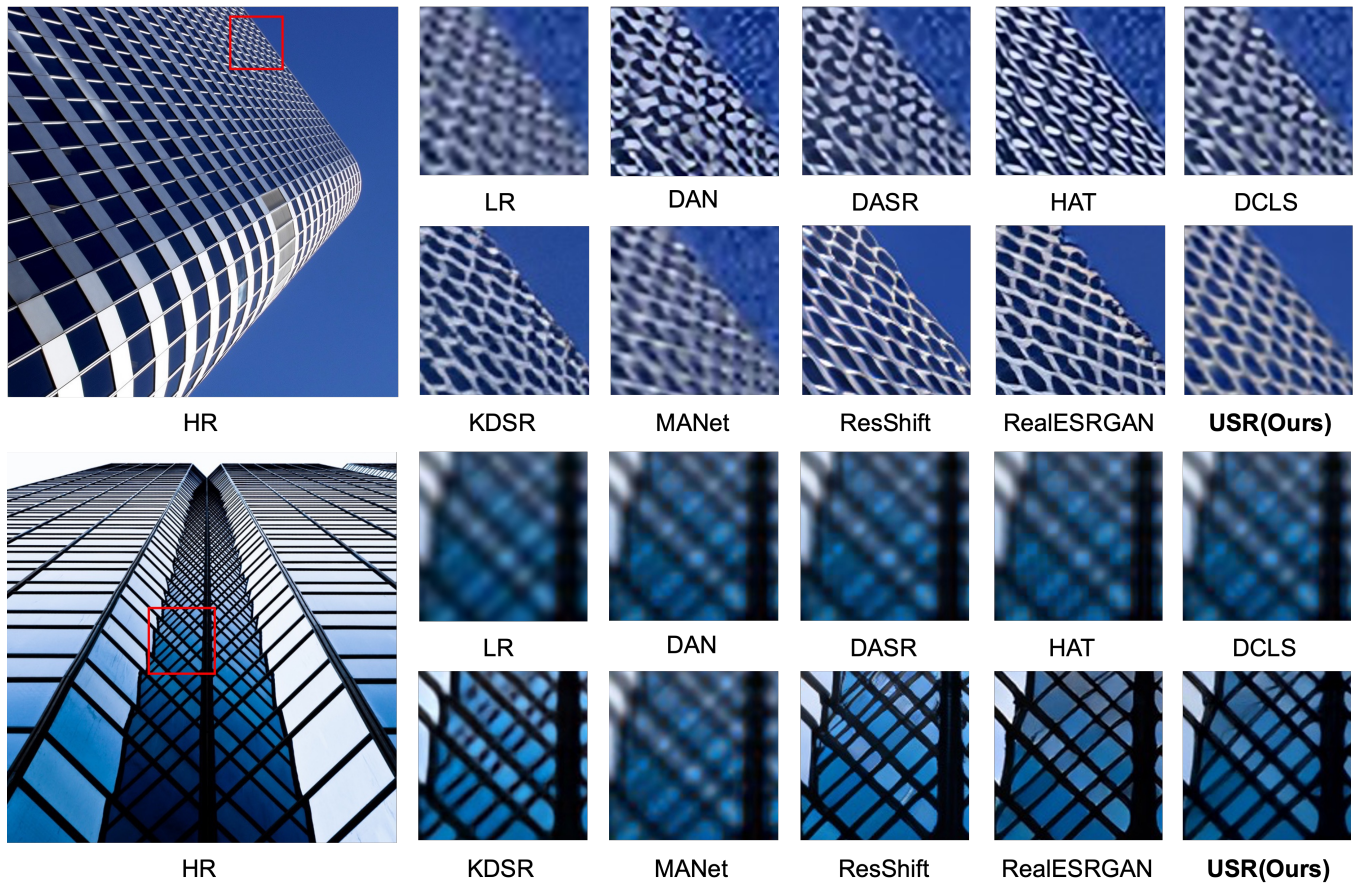


Figure 3: Visual comparisons of several representative methods on examples of the Urban100 dataset.

reduction in detail clarity, particularly noticeable on the "18". The USR method preserves the geometric shapes and sharpness of the edges, which is crucial for information recognition in real-world scenarios.

For the lemon tree example, the magnified view highlights the details on the lemon surface and the contrast between the branches. USR's method also demonstrates its superiority in restoring these details, especially in maintaining the texture of the lemon peel and the varied color gradients. Compared to other methods, USR's results are closer to the high-resolution original, with the lemon's tiny pits and shadows clearly reconstructed without the issues of excessive smoothing or loss of detail.

Comparison on Urban100 dataset. In Figure 3, we observe a comparison of super-resolution reconstruction effects for two urban buildings. Initially, the high-resolution image provides exceptionally sharp and detailed architectural features, including the clear edges of windows and detailed textures of glass reflections. Conversely, the low-resolution image exhibits significant loss of detail, with the building's lines and patterns becoming blurred.

The magnified area highlights the performance differences among various methods in super-resolution reconstruction. In the comparison of buildings, the USR (Ours) method demonstrates particularly

effective restoration of details and patterns, with the lines between windows clearly visible and the patterns closely resembling those of the high-resolution original image. Moreover, the USR method successfully restores the geometric symmetry of the buildings, a challenge for other methods which sometimes introduce blurring or ripple-like artifacts at these edge areas.

In the magnified view of the windows, the USR method maintains the natural texture of glass reflections, while other algorithms tend to overly smooth these areas, resulting in a loss of reflection and texture authenticity. Furthermore, in the images processed by USR, the lines of the buildings are straight and precise, and the division lines of the windows are more accurate, with no issues of staggering or misalignment.

When compared with other methods such as KDSR [9], MANet [5], ResShift [11], and RealESRGAN [6], it is evident that they face varying degrees of challenges when dealing with such complex structures. Some methods may perform well in restoring details but fall short in maintaining straight lines and natural textures. In contrast, the USR method not only delivers high-quality details but also exhibits significant advantages in overall geometric fidelity and visual impact of the images.

REFERENCES

- [1] Eirikur Agustsson and Radu Timofte. 2017. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- [2] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. 2011. Contour Detection and Hierarchical Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 5 (May 2011), 898–916. <https://doi.org/10.1109/TPAMI.2010.161>
- [3] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. 2015. Single Image Super-Resolution From Transformed Self-Exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5197–5206.
- [4] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. 2017. DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks. In *Proceedings of the IEEE International Conference on Computer Vision*. 3277–3285.
- [5] Jingyun Liang, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. 2021. Mutual Affine Network for Spatially Variant Kernel Estimation in Blind Image Super-Resolution. In *IEEE International Conference on Computer Vision*.
- [6] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. 2021. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF international conference on computer vision*. 1905–1914.
- [7] Xintao Wang, Liangbin Xie, Ke Yu, Kelvin C.K. Chan, Chen Change Loy, and Chao Dong. 2022. BasicSR: Open Source Image and Video Restoration Toolbox. <https://github.com/XPiPixelGroup/BasicSR>.
- [8] Pengxu Wei, Ziwei Xie, Hannan Lu, Zongyuan Zhan, Qixiang Ye, Wangmeng Zuo, and Liang Lin. 2020. Component divide-and-conquer for real-world image super-resolution. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII* 16. Springer, 101–117.
- [9] Bin Xia, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Radu Timofte, and Luc Van Gool. 2022. Knowledge distillation based degradation estimation for blind super-resolution. *arXiv preprint arXiv:2211.16928* (2022).
- [10] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma. 2010. Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing* (Nov 2010), 2861–2873. <https://doi.org/10.1109/tip.2010.2050625>
- [11] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. 2024. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in Neural Information Processing Systems* 36 (2024).