# Boosting Diffusion Guidance via Learning Degradation-Aware Models for Blind Super Resolution
## [Supplementary Material]

Shao-Hao Lu[1] Ren Wang[2] Ching-Chun Huang [1] Wei-Chen Chiu[1]
[1]National Yang Ming Chiao Tung University [2]MediaTek Inc.
{shlu2240.cs11, chingchun, walon}@nycu.edu.tw ren.wang@mediatek.com

Here, we provide 1) the details of training data preparation, 2) methodology of the explicit kernel estimator used in our ablation study for both the *explicit* and *combine* experiments, and 3) more qualitative results.

## 1. Training Data Preparation

To train our degradation and restoration models, we need to generate LR-HR paired data. For the DIV2K dataset with unknown degradation, the LR-HR paired images have been provided, and we just randomly crop HR patches of resolution $256 \times 256$ and their corresponding LR patches of resolution $64 \times 64$ during training.

For the CelebA-HQ dataset, which only consists of HR images, we generate LR images by performing convolution with random anti-isotropic Gaussian kernels, where the stride is equal to the degradation scale. Following Wang *et al*. [2], we diversify these kernels by sampling some additional parameters. Specifically, the filter size is sampled from the odd numbers in the range of $[7, 21]$. The weights are determined by two eigenvalues $\lambda_1$ and $\lambda_2$ of a covariance matrix, both of which are sampled from a uniform distribution $\mathcal{U}(0.2, 4)$. We also apply rotation to these kernels with an angle $\theta \sim \mathcal{U}(0, \pi)$.

## 2. Explicit Kernel Estimator

Both of our *explicit* and *combine* approaches involve integrating an explicit kernel estimator into DDNM [3]. Given a degraded image $\mathbf{y}$, the estimator $\mathcal{E}$ predicts the corresponding degradation kernel. Following the architecture proposed by Lian *et al*. [1], we train the estimator using the ground truth kernel as supervision. The training objective aims to minimize the $\mathcal{L}_1$ loss between the estimated kernel and the ground truth kernel $\mathbf{k}$ as

$$\mathcal{L}_k = \|\mathbf{k} - \mathcal{E}(\mathbf{y}; \theta_e)\|_1. \tag{1}$$

After obtaining the predicted kernel, we turn the convolution with this kernel into a matrix multiplication with a matrix $\mathbf{A}$. Furthermore, we can compute its pseudo-inverse $\mathbf{A}^\dagger$ using SVD, and directly plug $\mathbf{A}$ and $\mathbf{A}^\dagger$ into DDNM. However, this explicit kernel estimator assumes the degradation is a convolution with kernel, which might lead to over-simplification. Therefore, our *implicit* method shows better performance than *explicit* and *combine* methods in the ablation study.

Note that we do not have ground-truth kernels for the DIV2K dataset with unknown degradation, which are necessary to train the estimator $\mathcal{E}$. As a result, we adopt the pipeline used for generating training data to sample kernels, and use them to not only generate LR-HR pairs but also train the estimator $\mathcal{E}$.

## 3. More Qualitative Results

We provide more qualitative results on the *ImageNet-Val* and *CelebA-Val* datasets in Figs. 1–3.

## References

[1] Wenyi Lian and Shanglian Peng. Kernel-aware burst blind super-resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 4892–4902, 2023. 1

[2] Longguang Wang, Yingqian Wang, Xiaoyu Dong, Qingyu Xu, Jungang Yang, Wei An, and Yulan Guo. Unsupervised degradation representation learning for blind super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10581–10590, 2021. 1

[3] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*, 2022. 1

Bicubic SR3 DiffBIR SeeSR DDNM MsdiNet Ours Ground Truth

Figure 1. **Qualitative comparison of $4\times$ upsampling on *ImageNet-Val*.**


Figure 2. **Qualitative comparison of 4× upsampling on *CelebA-Val*.**


Figure 3. **Qualitative comparison of $8\times$ upsampling on *CelebA-Val*.**