

# Exploring Data Efficiency in Image Restoration: A Gaussian Denoising Case Study

## – Supplementary Materials –

Anonymous Authors

### 1 TRAINING DETAILS

Within the core narrative of our paper, we delineate extensive training outcomes utilizing two eminent CNN architectures: IRCNN [14] and DnCNN [13], alongside a powerful vision transformer framework, Restormer [12]. We meticulously curated four efficacious Mid subsets from the BSD [2], WED [9], Flicker2K [10], DIV2K [1] datasets, respectively. Each model—IRCNN, DnCNN, and Restormer—was trained on these subsets in addition to the original datasets as prescribed in their foundational papers. Collectively, across noise levels of 15, 25, and 50, we have cultivated 45 models, with their Peak Signal-to-Noise Ratio (PSNR) metrics systematically documented in the primary manuscript. This supplementary section is dedicated to elucidating the particulars of our training regimen.

#### 1.1 Ensuring Equitable Comparisons

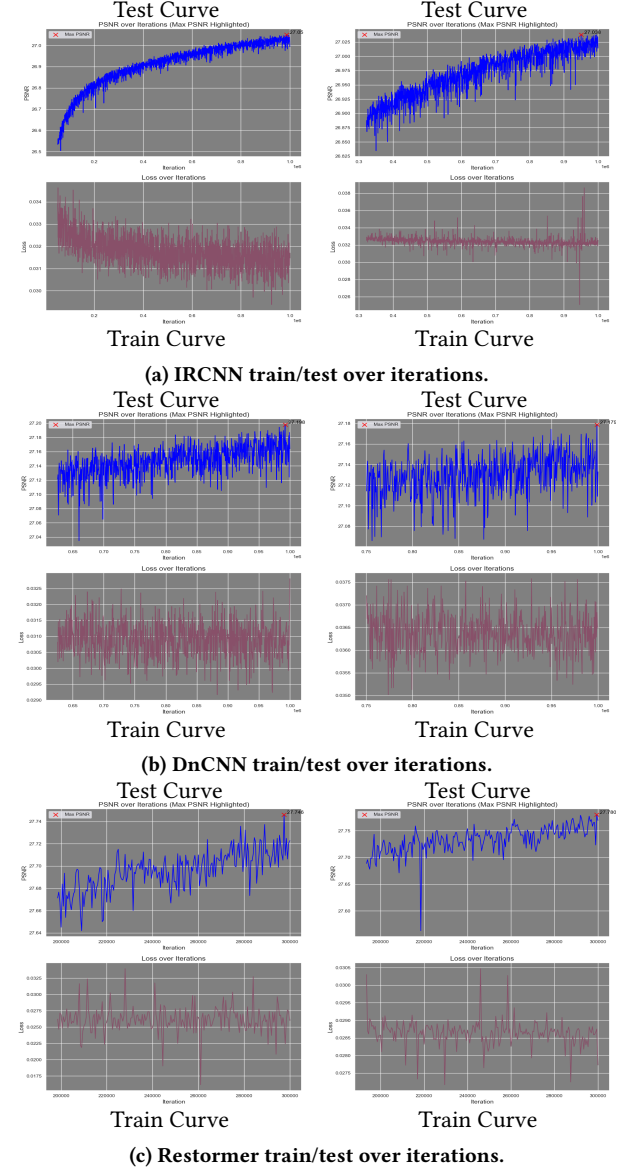
To guarantee a rigorously equitable evaluation, we adhered to a uniform training protocol across all models. Specifically, we employed the Adam optimizer, configuring it with  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , and a steadfast learning rate of  $1e^{-4}$  without any modification. For IRCNN and DnCNN, training batches were composed of 8 samples, with the training extending to 1,000,000 iterations. In contrast, Restormer was trained with batch sizes of 6 over 300,000 iterations. This protocol was consistently applied across both the selected Mid subsets and the comprehensive original datasets.

#### 1.2 Training and Evaluation Dynamics

We present the training and validation progression for the IRCNN, DnCNN, and Restormer models with Set12 [13] testset, specifically at a noise level of 50, across the most advantageous Mid subsets (consist with the main result in main manuscript, the optimal Mid subsets were derived from WED for both IRCNN and DnCNN, and from DIV2K for Restormer) and the unabridged original datasets, the results are illustrated in Fig. 1. For additional comparative analyses, refer to Fig. 2.

**Table 1: Comparative summary of subset selection strategies utilizing Top/Bottom/Mid/Uniform/Mix categories, symbolized as T/B/M/U/-. Subsets are selected from BSD dataset and PSNR are tested on Set12 with IRCNN model. The analysis confirms the Mid subset strategy as the optimal choice for data efficiency exploration.**

Size\Type	T-M	B-M	U	T-B	M
200	26.987	27.017	27.014	27.008	27.020



**Figure 1: Comparative analysis of training and validation trajectories across three models (IRCNN/DnCNN/Restormer), the left curves are based on most efficient Mid subset and the right curves are based on full original dataset, which emphasizes the superior efficiency of select Mid subsets (WED for IRCNN and DnCNN; DIV2K for Restormer) against the backdrop of full dataset training.**

**Table 2: Summary of JPEG Compression Artifact Reduction task results using efficient Mid subsets (size 200) from BSD and WED. We train ARCNN, DnCNN-3 respectively on each subset with 1000k iterations and record their best PSNR results on the classic5 testset. We also train them on a mixed large dataset with same iterations for fair comparison.**

Efficient Subset	Ratio	ARCNN [4]				DnCNN-3 [13]			
		$q = 10$	$q = 20$	$q = 30$	$q = 40$	$q = 10$	$q = 20$	$q = 30$	$q = 40$
BSD*	200	29.193	31.429	32.758	33.613	29.662	31.862	33.177	34.047
WED*	(2.30%)	29.262	31.390	32.696	33.570	29.713	31.887	33.198	34.038
Full dataset	8694	29.255	31.400	32.727	33.592	29.715	31.955	33.220	34.084

### 1.3 Additional Methodology for Subset Selection

In our primary manuscript, we undertook a comprehensive evaluation of Top, Mid, and Bottom subsets. Expanding upon this, our supplementary discourse includes alternative subset selection strategies predicated on influence scores. Novel approaches such as top-mid, bottom-mid, and uniform-sampling have been explored. We train IRCNN models on selected subsets from BSD using additional selection methodology, as evidenced in Tab. 1, the Mid subset selection strategy emerges as the most effective, attaining superior PSNR values.

## 2 EXTENSIONS TO OTHER IMAGE RESTORATION TASKS

This section investigates the application of our data-efficient methodology to two pivotal image restoration tasks: JPEG Compression Artifact Reduction and Image Super-Resolution.

### 2.1 JPEG Compression Artifact Reduction

Our approach for JPEG Compression Artifact Reduction involved the use of two classic CNN architectures: ARCNN [4] and DnCNN-3 [13]. Similar to gray image denoising, we initially trained ARCNN on a combined dataset from four sources [1, 2, 9, 10], targeting various quality factors. Subsequently, efficient subsets (200 images) were identified using ARCNN and assessed for their applicability to DnCNN-3, with classic5 [5] serving as the testset. As Tab. 2 illustrates, our methodology achieved comparable PSNR performance with substantially less data (2.30%). In the experiments utilizing ARCNN, our approach surpassed the results achieved with the full-sized dataset. Conversely, in the DnCNN experiments, the observed drop in PSNR was minimal, remaining within a margin of 0.08db. These outcomes collectively attest to the robustness and efficacy of our methodology in the context of JPEG Compression Artifact Reduction. The promising results not only validate our current approach but also underscore the merit in further investigating data efficiency within this domain.

### 2.2 Image Super-Resolution

In the realm of image super-resolution, we conducted experiments using Modified SRResNet (MSRResNet) [6, 11] and EDSR [8, 11], considering scale factors of 2 and 4. Following [7], we employ DIV2K [1] as original training dataset (full size) and Set5 [3] as the evaluation testset. Experiments are conducted in a similar way with gray image denoising. MSRResNet, initially trained on 900

DIV2K images, was subsequently applied to identify efficient subsets (comprising 200 images). These subsets were evaluated for their generalization to EDSR. Notably, we modified the model depth from the standard in BasicSR [11], reducing the number of residual blocks to accelerate training. The comprehensive results are tabulated in Tab. 3. Notably, for scale x2, our methodology achieved superior PSNR performance using only 22.2% of the data. For scale x4, the PSNR reduction was marginal (0.04db and 0.001db), underscoring the effectiveness of our approach in image super-resolution. These findings highlight the potential for further research into data efficiency in this domain.

**Table 3: Summary of Image Super-Resolution task results using efficient Mid subsets (size 200) from DIV2K. We train MSRResNet, EDSR respectively on each subset with 500k iterations and record their best PSNR results on the Set5 testset. We also train them on original DIV2K (full size) with same iterations for fair comparison.**

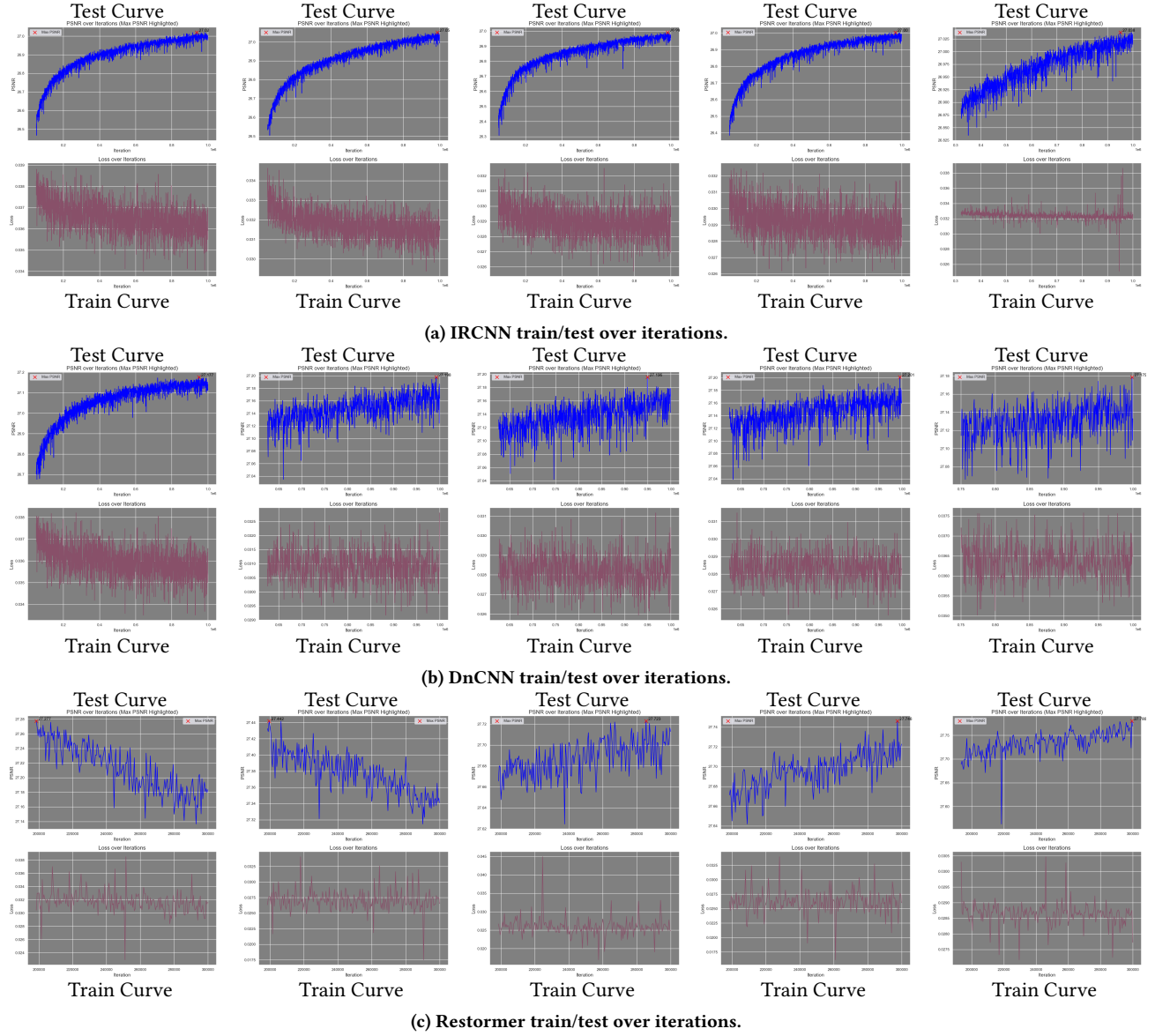
DIV2K	Scale: x2	Scale: x4
MSRResNet [11]	35.650	30.019
EDSR [8]	35.709	30.053
Efficient (Ratio)	Scale: x2	Scale: x4
MSRResNet [11] (22.2%)	35.666	29.979
EDSR [8] (22.2%)	35.720	30.052

## 3 LIMITATIONS AND FUTURE WORKS

While our method demonstrates effectiveness in exploring data efficiency for restoration tasks, it is inherently bounded by the existing datasets' quality ceiling. This limitation highlights the necessity for approaches that can transcend the constraints of current datasets. Future research will pivot towards a generative strategy, leveraging the characteristics of efficient subsets identified through our methodology. This generative direction is anticipated to broaden the scope and improve the efficacy of data utilization in the context of image restoration.

## REFERENCES

- [1] Agustsson et al. 2017. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- [2] Pablo Arbeláez, Michael Maire, Charless Fowlkes, and Jitendra Malik. 2011. Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions*



**Figure 2: Comparative analysis of the training and validation trajectories for three models—IRCNN, DnCNN, and Restormer. The results are presented from left to right for the BSD Mid subset, WED Mid subset, Flickr2K Mid subset, DIV2K Mid subset, and the original full-size dataset.**

on *Pattern Analysis and Machine Intelligence* 33, 5 (2011), 898–916. <https://doi.org/10.1109/TPAMI.2010.161>

[3] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. 2012. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. (2012).

[4] Chao Dong, Yubin Deng, Chen Change Loy, and Xiaoou Tang. 2015. Compression artifacts reduction by a deep convolutional network. In *Proceedings of the IEEE international conference on computer vision*. 576–584.

[5] Foi et al. 2007. Pointwise Shape-Adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images. *IEEE Transactions on Image Processing* 16, 5 (2007), 1395–1411. <https://doi.org/10.1109/TIP.2007.891788>

[6] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4681–4690.

[7] Jingyun Liang, Jiezhong Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. 2021. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*. 1833–1844.

[8] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. 2017. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 136–144.

[9] Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. 2017. Waterloo Exploration Database: New Challenges for Image Quality Assessment Models. *IEEE Transactions on Image Processing* 26, 2 (2017), 1004–1016. <https://doi.org/10.1109/TIP.2016.2631888>

[10] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, and Lei Zhang. 2017. NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results. In *Proceedings of the IEEE Conference on Computer Vision and Pattern*

- Recognition (CVPR) Workshops.
- [11] 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/XPiXeLGroup/BasicSR>.
- [12] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shah-baz Khan, and Ming-Hsuan Yang. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5728–5739.
- [13] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing* 26, 7 (2017), 3142–3155.
- [14] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. 2017. Learning Deep CNN Denoiser Prior for Image Restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.