Supplementary Material for
JDEC: JPEG Decoding via Enhanced Continuous Cosine Coefficients

A. Data Processing Inequality in JPEG

In terms of information theory [3], it's feasible to view JPEG compression as resembling a Markov chain as follows:

$$X \rightarrow Y \rightarrow \hat{X}, \tag{1}$$

where, X is a symbol (Images), Y is a encoded file (bit-streams) and $\hat{X}$ is decoded symbol (JPEG Images). Therefore by the data processing inequality [3], the mutual information $I(X;Y) := \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$ between X and $\hat{X}$ cannot exceed the information between X and Y. In JPEG, the equality is satisfied when a conventional decoder does not provide any losses.

$$I(X;Y) \geq I(X;\hat{X}). \tag{2}$$

The equality is satisfied when conditional mutual information $I(X;Y|\hat{X}) = 0$. Theoretically, a JPEG decoder should not provide loss. However, in engineering practice, loss occurs in converting the DCT spectrum back to YCbCr by performing IDCT and rounding to values of 0 and 255. Also, depending on the YCbCr to RGB matrix, precision losses occur too. Consequently, the conventional JPEG provides losses, and Eq. (2) does not satisfy equality.

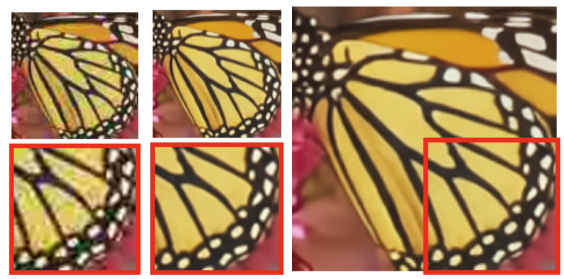
B. Implicit Neural Representation

The proposed JDEC is a function of block coordinates. To demonstrate this, the model should be able to produce results when provided with coordinates that were not observed during training. Therefore, we conducted simple super-resolution experiments using the pre-trained JDEC to represent unobserved coordinates. We maintain all settings identical, with the size of block coordinate δ increased to $rB \times rB$ where r is the upsampling ratio.

q=10, $\downarrow \times 4$				
Concept	SR w/ JAR	SR w/ o JAR	SR w/ JAR	Decoding
Method	RRDB[9]	SwinIR[7]	HST[6]	JDEC (*ours*)
Set5	22.36	22.45	22.49	22.48

Table 1. **Quantitative comparison** for compressed image upsampling on the Set5 [1] dataset (PSNR(dB)). The JAR refers to JPEG Artifact Removal.

Fig. 1 shows that our JDEC is clearly an implicit neural representation by extracting unseen coordinates. It also demonstrates that the resolution of estimated spectra by CCF is a function of continuous frequency. We compare our upsampling results to existing networks that aim to upsample compressed images in Tab. 1.



Input ($q = 10$) JDEC + Bicubic JDEC w/o Bicubic

Figure 1. $\times 2$ **upsampling of compressed image** ($q = 10$) in Set-5 dataset. Note that JDEC upscales the image with the change of additional coordinates δ without additional training.

C. Computational Costs and Performance

With Fig. 2 and Tab. 2, we present an additional comparison of the computational resources including an extended FBCNN model (FBCNN+) and JDEC-CNN+. Our framework overcomes the trade-off between computational complexities and performances.

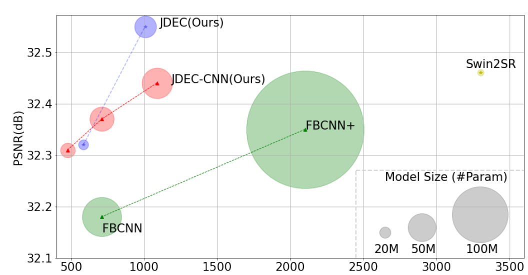


Figure 2. FLOPs and PSNR Comparison with other methods in ICB[8] ($q = 10$).

Method	#Params. (M)	Mem. (GB)	Time (ms)	FLOPs (G)	PSNR|PSNR-B (dB) $q = 10$	PSNR|PSNR-B (dB) $q = 40$
FBCNN [5]	70.1	**0.61**	71.95	709.97	32.18|32.15	36.02|35.95
JDEC-CNN	**26.2**	0.81	**56.59**	**476.33**	**32.31|32.27**	**36.19|36.09**
FBCNN+† [5]	210.3	1.68	218.58	2101.15	32.35|32.31	36.20|36.11
JDEC-CNN+†	**54.6**	**1.62**	**111.65**	**1086.87**	**32.43|32.39**	**36.30|36.19**
Swin2SR [2]	11.5≤	2.79	2203.59	3301.5	32.46| -	36.25| -
JDEC	38.9	**1.76**	**224.79**	**1006.72**	**32.55|32.51**	**36.37|36.28**

Table 2. Computational resources & performance comparison for a 560×560 pixels in ICB [8]. † :Our implementation

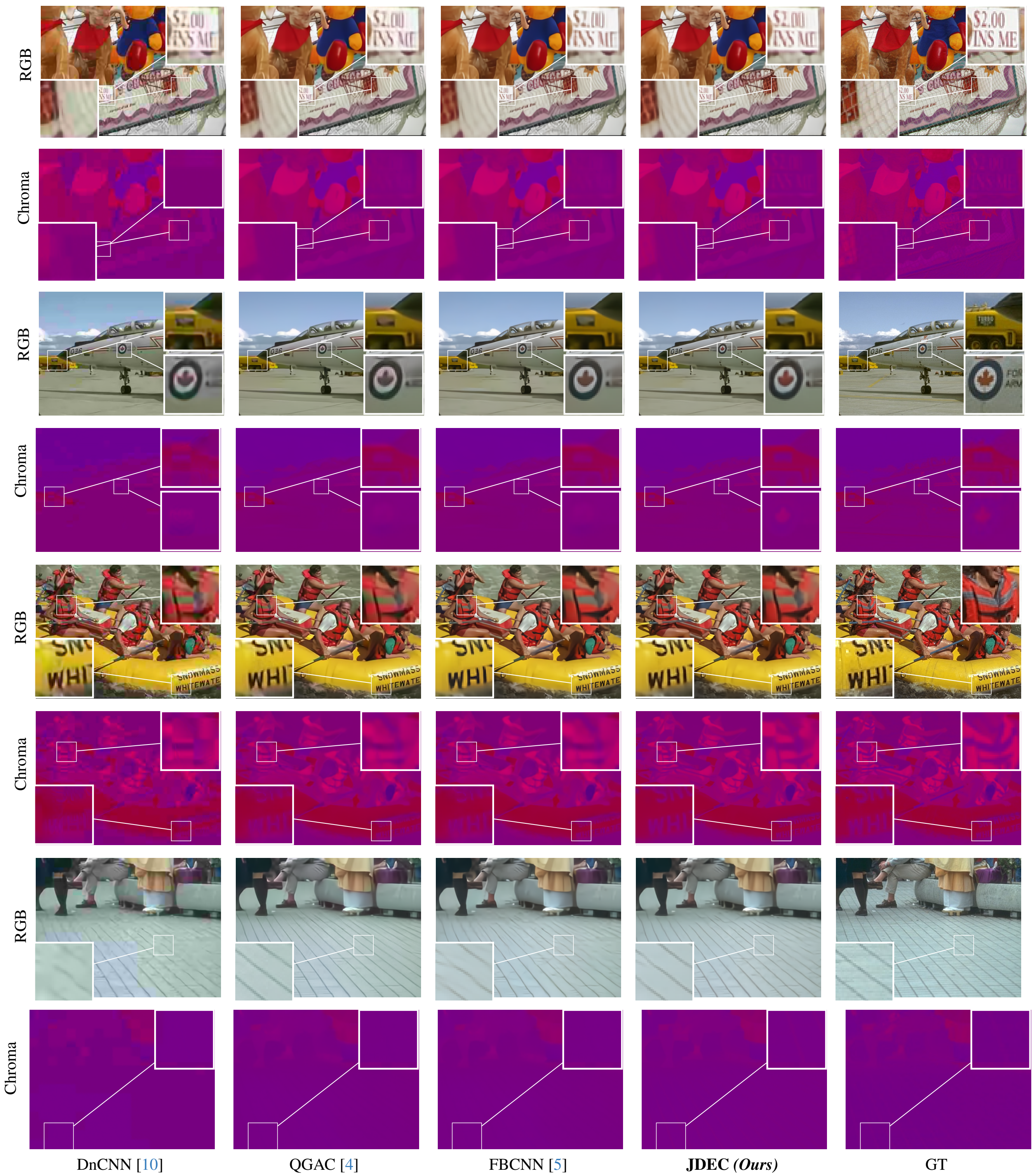


Figure 3. Additional qualitative comparison in color JPEG artifact removal ($q = 10$).

D. Additional Qualitative Results

We present additional qualitative results for comparison. Additionally, we divide RGB and chroma images to demonstrate the robustness of our JDEC to color distortion.

References

[1] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. 2012. 1

[2] Marcos V Conde, Ui-Jin Choi, Maxime Burchi, and Radu Timofte. Swin2sr: Swinv2 transformer for compressed image super-resolution and restoration. In *European Conference on Computer Vision*, pages 669–687. Springer, 2022. 1

[3] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999. 1

[4] Max Ehrlich, Larry Davis, Ser-Nam Lim, and Abhinav Shrivastava. Quantization guided jpeg artifact correction. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16*, pages 293–309. Springer, 2020. 2

[5] Jiaxi Jiang, Kai Zhang, and Radu Timofte. Towards flexible blind jpeg artifacts removal. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4997–5006, 2021. 1, 2

[6] Bingchen Li, Xin Li, Yiting Lu, Sen Liu, Ruoyu Feng, and Zhibo Chen. Hst: Hierarchical swin transformer for compressed image super-resolution. In *European Conference on Computer Vision*, pages 651–668. Springer, 2022. 1

[7] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. SwinIR: Image Restoration Using Swin Transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 1833–1844, 2021. 1

[8] Rawzor. Image compression benchmark. . *url: http://imagecompression.info/*. 1

[9] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018. 1

[10] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7):3142–3155, 2017. 2