FA-UNet: An Attention-UNet-based Frequency Domain Image Denoising and Deblurring System

Yiming Jia, Haoyang Ju, Shiyuan Feng

Abstract—FA-Unet introduces a novel approach in digital image processing, focusing on denoising and deblurring through the frequency domain. Diverging from traditional spatial domain methods, this technique utilizes frequency-based separation to effectively isolate and address image distortions. Central to FA-Unet is the application of attention mechanisms, adapted from language processing, to selectively target and improve areas most affected by noise and blur. This targeted restoration approach promises significant improvements in various applications, including medical imaging, satellite imagery, and general photography, offering a new direction in high-quality image enhancement.

Index Terms—Frequency Domain Image Processing, Attention Mechanisms, Image Denoising, Image Deblurring

1 Introduction

In the field of digital image processing, improving the quality of images is a significant challenge. Images often suffer from problems like blur due to movement or camera shake, and noise, which can look like random specks or grain [1]. These issues can make images less clear and less useful, especially in important areas like medical imaging, satellite photos, and everyday photography. Most current methods to fix these problems work in the spatial domain [2], meaning they try to adjust the pixels directly. However, methods for image denoising in the frequency domain have not been fully studied.

Our project, titled ”FA-Unet”, seeks to innovate in this domain by shifting the paradigm from the conventional spatial domain to the frequency domain. The frequency domain offers a unique perspective on image data, representing it in terms of frequency components rather than the spatial distribution of pixel values. This shift is crucial because it allows for the separation of image components based on their frequency characteristics, which is often more effective for identifying and isolating noise and blur elements.

A key feature of our method is using attention mechanisms [3] in the frequency domain. Attention mechanisms, which have been very successful in areas like language processing, help our method focus on the specific parts of an image most affected by noise and blur. This means that our method doesn’t just apply the same changes to the whole image, which can sometimes make the image worse. Instead, it targets only the parts that need fixing.

The potential applications of ”FA Unet” are vast. In medical imaging [4], clearer images can lead to more accurate diagnoses. In satellite imagery [5], enhanced clarity can improve the analysis of geographical and environmental data. In the realm of photography, both professional and amateur photographers can benefit from images that are sharper and free from unwanted noise and blur.

The results achieved by FA-UNet in our experiments are a testament to the method’s effectiveness and the importance of frequency domain analysis. The notable increase in the average Peak Signal-to-Noise Ratio (PSNR) of images, from 13 to 23, empirically demonstrates the substantial improvement in image quality. This significant leap in PSNR not only highlights the ability of FA-UNet to effectively remove noise and blur from images but also underscores the crucial role that frequency domain processing plays in achieving superior image clarity and detail retention.

In conclusion, our project is a new way of fixing images that leverages the untapped potential of the Fourier domain, combined with the power of attention mechanisms. We believe that ”FA Unet” offers a novel solution to a longstanding challenge in the field.

2 Related Work

The proliferation of digital imaging has escalated the development of sophisticated image processing techniques, notably in deblurring and denoising [6] [7]. These advancements are increasingly powered by machine learning technologies.

2.1 Frequency Domain Image Processing

The concept of processing images in the frequency domain has been a cornerstone in the field of digital image enhancement for decades. Unlike spatial domain techniques that manipulate pixel values directly, frequency domain methods involve transforming the image into a frequency representation, typically using Fourier transforms [8]. This approach is particularly advantageous for identifying and isolating periodic patterns, noise, and other high-frequency components that are not as readily apparent in the spatial domain. Early applications of frequency domain techniques primarily focused on basic filtering tasks, such as low-pass and high-pass filtering [9], which were effective for smoothing or enhancing image features, respectively. However, these techniques often overlooked the complexity and variability of noise and blur in real-world images, leading to suboptimal enhancement outcomes.

Recent advancements in frequency domain image processing have seen significant developments, particularly in the realms of noise reduction and image sharpening. Modern approaches now incorporate more sophisticated algorithms that adaptively modify frequency components based on the image content [10]. This has led to more effective handling of diverse types of noise and blur that are commonly encountered in practical scenarios, such as Gaussian noise and motion blur. The advent of machine learning has further revolutionized this field, enabling more intelligent and content-aware frequency domain processing [11]. These advanced methods can now better distinguish between noise and actual image details, significantly enhancing the ability to restore and improve overall image quality. Despite these advancements, the challenge of seamlessly integrating these techniques into real-time processing applications remains, as frequency domain transformations are computationally intensive and often require substantial processing power.

2.2 Attention Mechanisms in Image Processing

Attention mechanisms, originally conceptualized for tasks in natural language processing, have been adeptly adapted for image processing. Their ability to dynamically focus on relevant portions of an image has made them essential in various advanced applications: Vision Transformers (ViTs) [12]: Vision Transformers have been a groundbreaking development in computer vision. For instance, Google's ViT model [13] applies the transformer architecture to image classification tasks, breaking down images into sequences of patches and using self-attention mechanisms to understand the global context of the image. This approach has shown remarkable success in areas like image classification and object detection. Gated Attention Networks for Noise Reduction [14]: In image denoising, gated attention networks have been used to focus on noise patterns within an image selectively. By concentrating on these areas, these networks enhance the effectiveness of the denoising process, leading to clearer and more accurate image restoration.

2.3 CNN-Based Denoise and Deblur Models

Advances in denoise models, especially those using deep learning, have dramatically improved noise removal capabilities. CNNs [15], due to its strong expressive abilities, is one of the main techniques widely used for image denoising. For instance, Lan et al. [16] embedded residual block into a CNN to reduce noise for obtaining clean images. Shi et al. [17] integrated hierarchical features to obtain richer information to improve denoising results. Unet [18], as a powerful variant of CNN, has also been applied to image denoising tasks. For example, Fan et al. [19] proposed a way of combining Unet with Transformer layer to enhance the denoising performance. Huibin et al. [20] fused self-attention mechanism into a residual Unet model to guide it suppressing the noise in digital images.

Deblur models have similarly evolved, addressing image blur due to motion or focus issues. Numerous methods [21] [22]train 2D CNNs to sharp images. Tao et al. [23] leverage recurrent layers to extract features across images at multiple scales in a coarse-to-fine manner. To avoid vanishing or exploding gradients during training, DeblurGAN [24] uses global residual layers to directly connect low-level and high-level layers in the area of image deblurring. The attention layer can help deep networks focus on the most important image regions for deblurring. Shen et al. [25] propose an attention-based deep deblurring method consisting of three separate branches to remove blur from the foreground, the background, and globally, respectively.

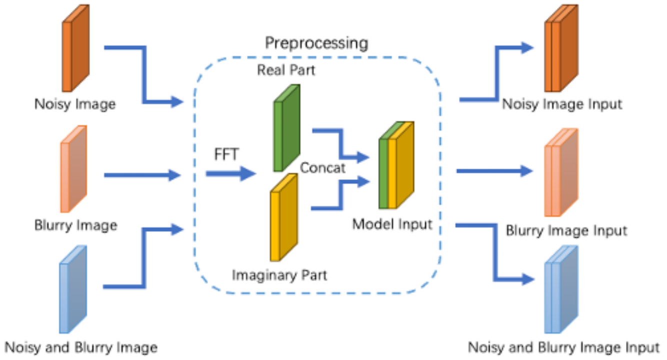

Fig. 1. Data preprocessing pipeline.

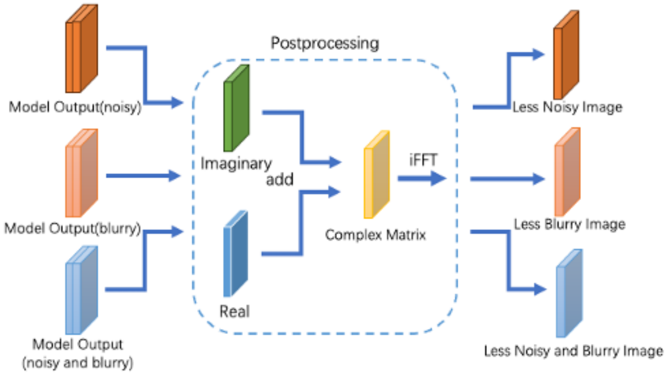

Fig. 2. Data postprocessing pipeline.

These developments reflect the dynamic nature of machine learning in enhancing image quality, showcasing significant progress in attention-based models, denoising, and deblurring techniques.

3 Proposed Method

In our approach, we combine the structures of R2U-Net and Attention U-Net to construct our image reconstruction model and use a three-stage supervised training approach to train it. To adapt our model for frequency domain processing, we have to transform existing datasets from the spatial to the frequency domain.

3.1 Data processing pipeline

Since our model works in the frequency domain, we have to build a data preprocessing pipeline to transform natural image data in the spatial domain into the frequency domain to generate the model input, and a data postprocessing pipeline to do the reverse operation for easier visualization.

In our preprocessing pipeline, as shown in Figure 1, we first took the three-channel images that needed to be reconstructed, applying Fast Fourier Transform (FFT) on them to convert them into the frequency domain tensors, while the dimensionality remains the same. However, a challenge arises due to the tensors in the frequency domain being composed of complex numbers, while neural networks are designed to process real numbers. To address this, we considered two options to fit our data into the neural networks: The first option was to retain only the real part or take the absolute values, but this option preserved only part of the information and resulted in unsatisfying outcomes in experiments. The second option, which we adopted, is to separate the real part and imaginary part of each tensor of an image and concatenate them in the channel dimension. This approach allows us to keep all the information of the tensor. The finial model input is a six-channel tensor.

In our postprocessing pipeline, which is shown in Figure 2, we perform operations that are the inverse of those in the preprocessing pipeline. We take the six-channel tensors outputted by our model and divide them into two sets of three-channel tensors, corresponding to the real and imaginary parts. These parts are then recombined into complex numbers. Finally, we perform Inverse Fast Fourier Transform (iFFT) to get the reconstructed image in the spatial domain.

3.2 Neural Networks Structure

Our model incorporates recurrent convolutional layers within a U-Net framework, enabling the network to revisit the input data iteratively. This feature is crucial for complex tasks such as image denoising and deblurring, where distinguishing between useful signal and disruptive noise can be challenging. The recurrent layers enhance the model's ability to capture and emphasize important features while suppressing irrelevant ones over successive iterations.

Moreover, the attention modules depicted in the schematic are strategically placed to focus the model's capacity on areas most affected by noise and blur. These modules act selectively, improving the clarity of the output by adjusting the processing power applied to different regions of the image. The attention-driven focus is especially beneficial for maintaining the integrity of edges and textures—a common concern in the denoising and deblurring process.

Together, these components form a powerful neural network, as illustrated, that is finely tuned for the intricate demands of image enhancement, achieving a balance between depth of feature extraction through recurrence and precision through attention.

3.3 Training Method

In FA-Unet, we propose an image denoising + deblurring model training pipeline for Unet-based models and it consists of two pretraining stages and one E2E training stage as shown in Figure 4. Through experiments, this structure outperforms various other structure candidates, for example, based on transfer learning on a single UNet model (missing stage 3 in Figure 4) or structures similar to language model encoder pretraining which is usually used in NLP field (missing stage 2 in Figure 4).

Model Pretraining Stage In model pretraining stages, two Unet-based models are separately initialized. One of them will be trained on noisy-only images to perform image denoising task (Stage 1) and the other will be trained on blurry-only images to train its image deblurring ability (Stage 2). For image denoising task, we choose to use MSE as loss function as it is based on L2 norm. For image deblurring task, we utilize MAE loss which is based on L1 norm (Appendix A). Stage1 and Stage2 in Figure 4 can be performed in parallel to increase model pretraining efficiency.

Model E2E Training Stage In model E2E training stage, above two pretrained models are concatenated in ”first denoising, second deblurring” sequence and the whole system is trained end-to-end on both noisy and blurry image inputs. This stage's final output will be the FA-Unet system's image output.

4 Experimental Results

4.1 Datasets

The **GoPro Dataset**, specifically designed for image deblurring tasks, comprises a collection of 3,214 images of 1,280×720 resolution. This dataset is divided into 2,103 training images and 1,111 test images. It features realistic pairs of blurred and corresponding sharp images, captured using a high-speed camera. This makes the dataset a valuable resource for research and development in image and video deblurring algorithms.

For more details, please refer to the dataset page on *Papers With Code*: https://paperswithcode.com/dataset/gopro.

The **Intel Image Classification** dataset, hosted on *Kaggle*, includes approximately 25,000 images categorized into six classes: *buildings*, *forest*, *glacier*, *mountain*, *sea*, and *street*. These images are provided in high resolution and a standardized resolution of 150x150 pixels. The dataset is divided into three subsets: 14,000 images for training, 3,000 for testing, and 7,000 for prediction tasks. This dataset is particularly useful for machine learning and computer vision applications, focusing on natural scene classification.

For more detailed information, you can visit the dataset page on *Kaggle*: https://www.kaggle.com/datasets/puneet6060/intel-image-classification.

For the Intel dataset, since the original images are free from noise and blur, we first convolved them with a Gaussian kernel to blur the images, and then added Gaussian noise to create a suitable training set. In the case of the GoPro dataset, the blur is inherent and naturally present in the dataset, so we only needed to manually add Gaussian noise.

4.2 Experiment across different system structures on GoPro

We conducted a comprehensive series of experiments on different model structures and training methods in the frequency doamin on the GoPro dataset, including the following four experimental settings:

a) **Only U-Net**: Train one basic U-Net on a dataset of noisy and blurry images, and measure the performance of U-Net on general image reconstruction in the frequency domain.

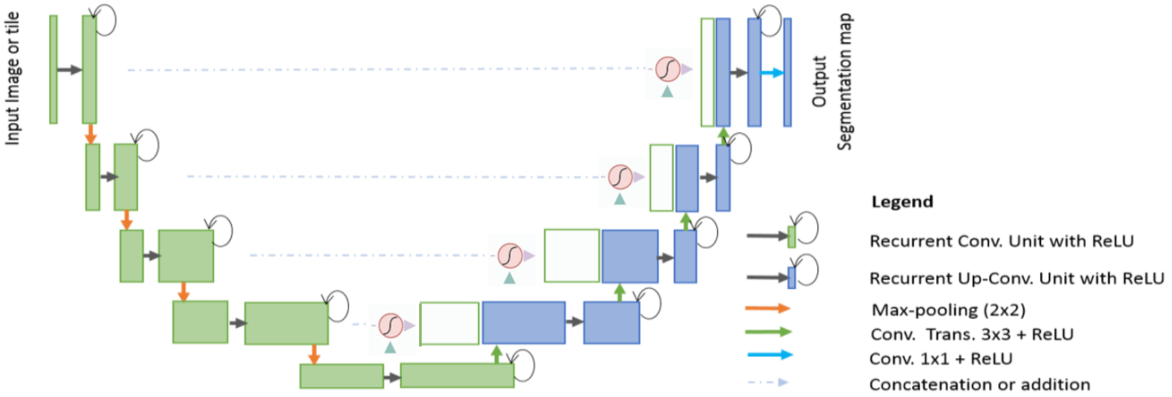

Fig. 3. Attention R2U-Net structure.

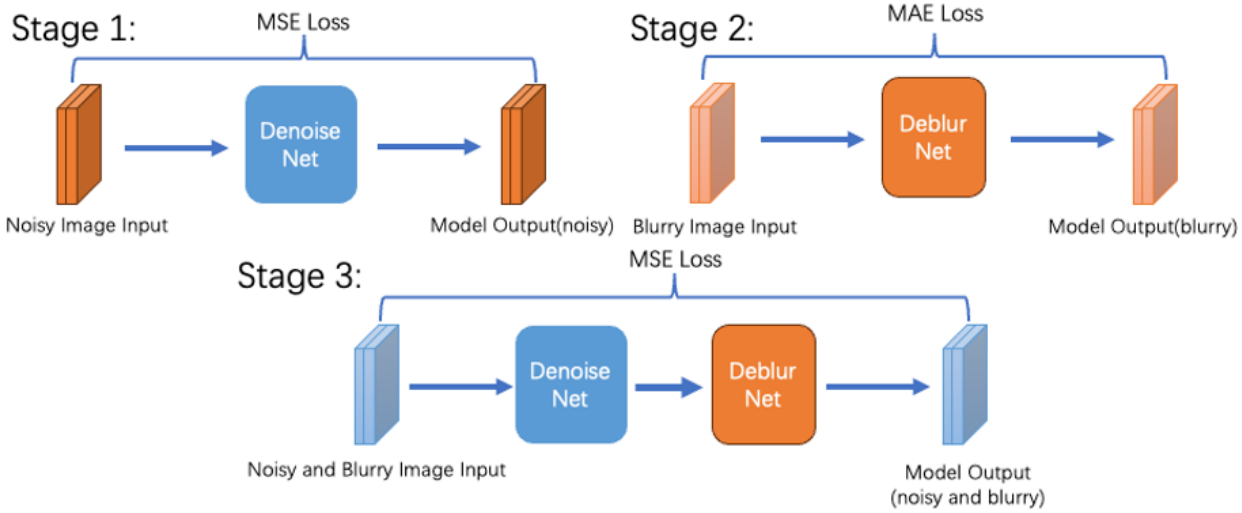

Fig. 4. 3-stages training pipeline.

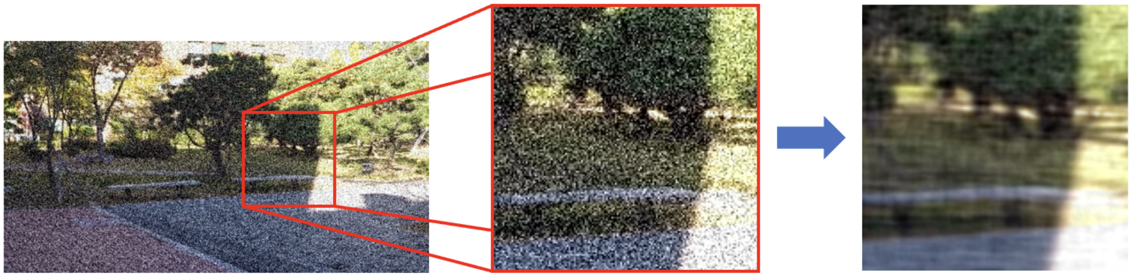

Fig. 5. FA-UNet can denoise and deblur images in detail.

b) **Only Attention R2U-Net**: Train one Attention R2U-Net on a dataset of noisy and blurry image, and measure the performance of Attention R2U-Net on general image reconstruction in the frequency domain.

c) **FA-UNet without pertaining Deblur Net**: First pretrain the Denoise Net on noisy-only images, then train FA-UNet end-to-end on both noisy and blurry images. In other words, we implemented Stage 1 and Stage 3 in Figure 4 based on the hypothesis that the blur patterns in the blur-only images differ from those in the images after denoising, thereby making the effect of pretraining the Deblur Net uncertain.

d) **FA-UNet**: Train the FA-UNet following the three-stage process described in section 3.3.

The performance of image reconstruction was quantified using the Peak Signal-to-Noise Ratio (PSNR) metric. The results of different experimental settings are shown in Table 1.

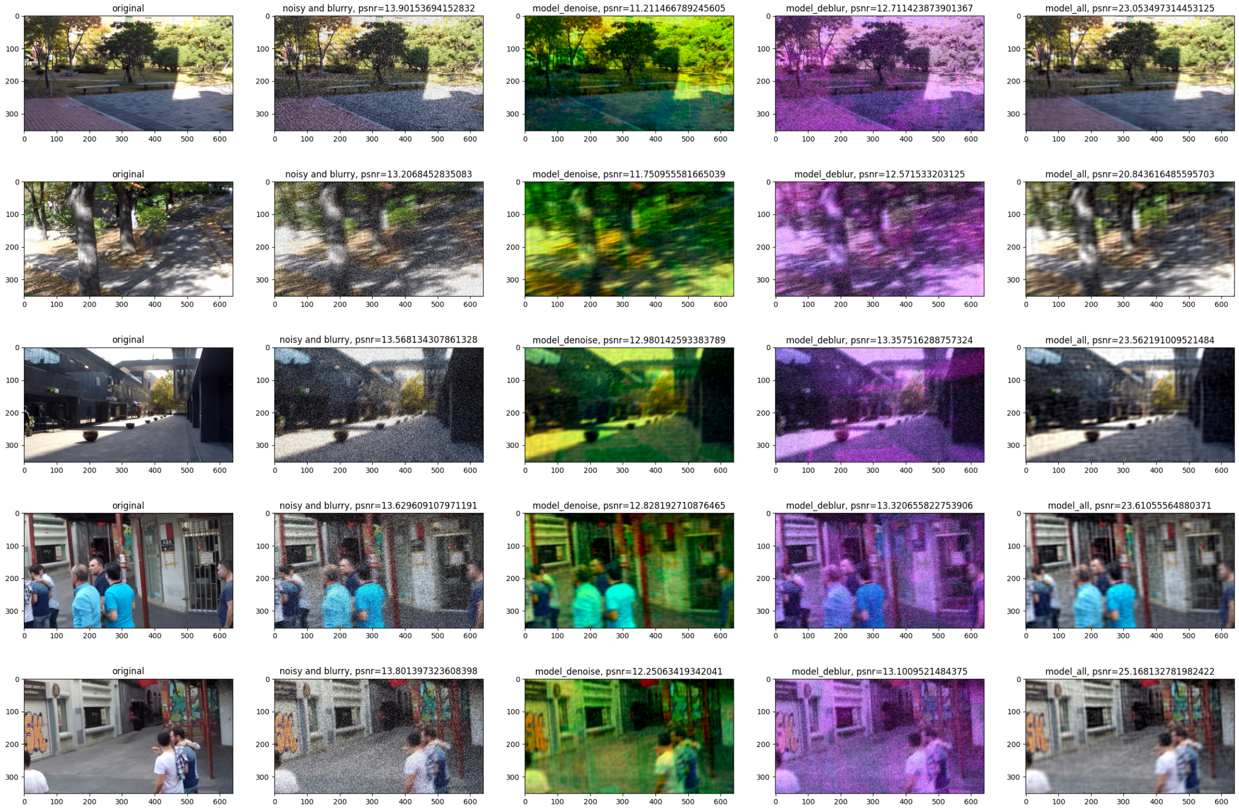
original
noisy and blurry, psnr=13.90153694152832
model_denoise, psnr=11.211466789245605
model_deblur, psnr=12.711423873901367
model_all, psnr=23.053497314453125
original
noisy and blurry, psnr=13.2068452835083
model_denoise, psnr=11.750955581665039
model_deblur, psnr=12.571533203125
model_all, psnr=20.843616485595703
original
noisy and blurry, psnr=13.568134307861328
model_denoise, psnr=12.980142593383789
model_deblur, psnr=13.357516288757324
model_all, psnr=23.562191009521484
original
noisy and blurry, psnr=13.629609107971191
model_denoise, psnr=12.828192710876465
model_deblur, psnr=13.320655822753906
model_all, psnr=23.61055564880371
original
noisy and blurry, psnr=13.801397323608398
model_denoise, psnr=12.25063419342041
model_deblur, psnr=13.1009521484375
model_all, psnr=25.168132781982422
0
100
200
300
400
500
600

Fig. 6. Comparison between FA-UNet, Denoise Net and Deblur Net. From left to right, each column includes the original high-quality images, noisy and blurry images, images after being processed by the Denoise Net, images after being processes by the Denoise Net, and images after being processed by FA-UNet.

Table 1: PSNRs of different experimental settings.

method	PSNR
Unprocessed images	13.351
Only U-Net	22.246
Only Attention R2U-Net	22.637
FA-UNet without pertaining Deblur Net	22.872
FA-UNet	**22.937**

The PSNR of the original images serves as a baseline and represents the quality of the images before processing. All the methods substantially increase the PSNR compared to the baseline, indicating the effectiveness of image reconstruction in the frequency domain. The Attention R2U-Net yields a PSNR higher than the U-Net, which suggests that the attention mechanisms and R2U-Net structure provide improvement in image reconstruction quality over the basic U-Net. The PSNRs of FA-UNet both with and without pretraining the Deblur Net are higher than those from using a single net. This indicates that separating the task of denoising and deblurring task is more effective. Moreover, pretraining the Deblur Net yields a slightly higher PSNR. Another thing worth mentioning is that the images reconstructed using FA-UNet exhibit a slight blurriness. This may suggest that the deblurring ability of UNet-based in the frequency domain is not good enough.

4.3 Experiment in different datasets: GoPro and Intel Image Classification

Next, we applied two different image datasets with distinct characteristics to our FA-Unet to further evaluate its performance. GoPro dataset holds higher resolution images and most of them contain the street scenes with complex environmental detail. On the other hand, Intel Image Cla'ssification dataset includes smaller size images with 150*150 resolution and their scene contain less objects.

Through experiments, for GoPro dataset, our system achieves average validation PSNR of 22.937. For Intel Image Classification dataset, average validation PSNR of 23.092 is achieved. Under different image datasets with various characteristics, our system maintains stable performance which suggests it can be applied to image denoising and deblurring tasks in diverse scenarios.

From Figure 7, we can virtually observe that our system performs similarly on both large-complex images and small-simple images. Again, in both cases, the resulting photos are a little bit blurry. However, during experiment, we find out that if the input photos only contain Gaussian blur, our system can achieve better deblurring performance, as shown in Figure 8.

4.4 Experiment between spatial approach and frequency approach

Finally, we conducted a comparative study between the traditional spatial domain deblurring approach, exemplified

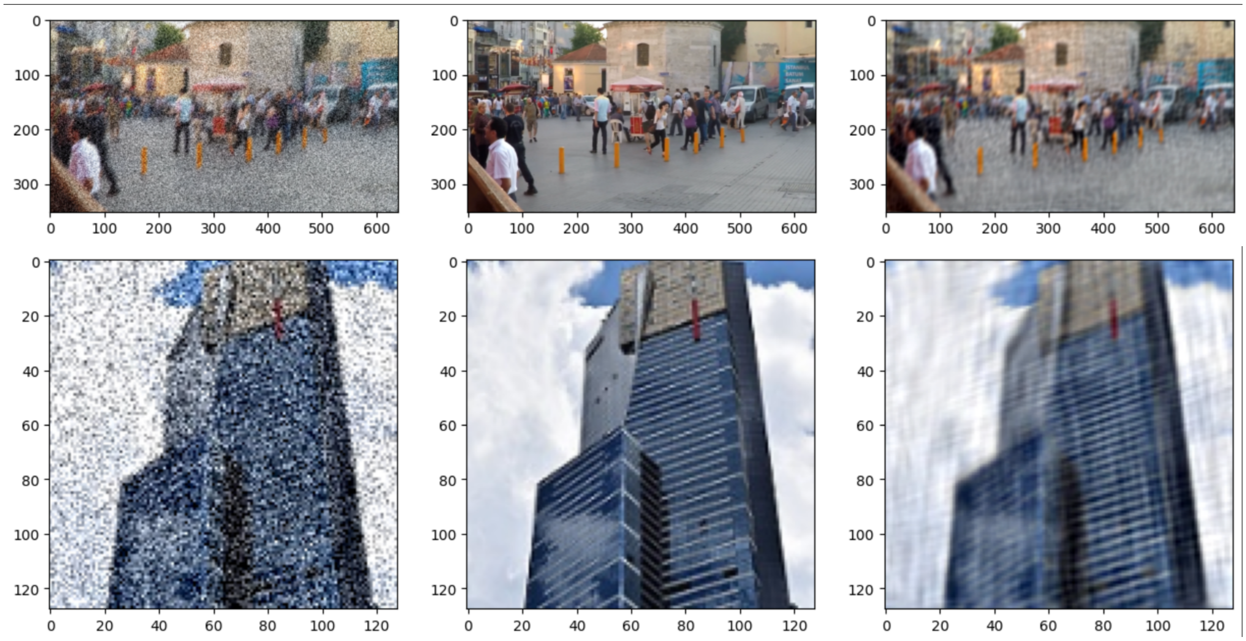

Fig. 7. Example Image Denoising and Deblurring Result on GoPro and Intel Image Classification datasets.

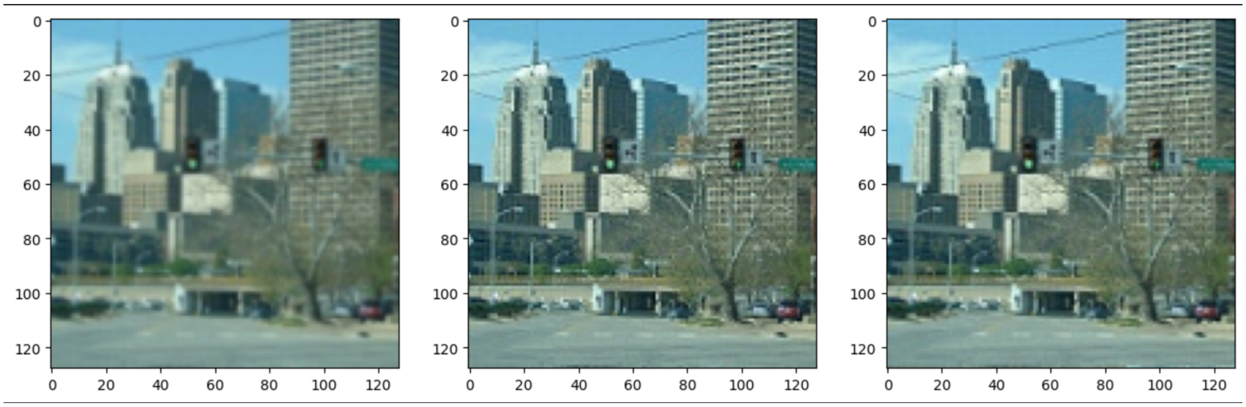

Fig. 8. Image Deburring Result on Photo only Contains Gaussian Blur

by UNet, and our proposed FA-UNet method which operates in the frequency domain. Our result of this experiment is in Fig. 8. Despite FA-UNet exhibiting a slightly lower Peak Signal-to-Noise Ratio (PSNR) in quantitative evaluations, qualitative assessments paint a different picture. Visual inspection reveals that images processed by FA-UNet retain a richer set of details, especially textural information that is often lost in conventional deblurring. This not only highlights the subjective improvement in image quality as perceived by the human eye but also underscores the significance of incorporating frequency domain information—a testament to the efficacy of FA-UNet in preserving essential image characteristics.

5 Conclusion

In conclusion, our project presents a novel approach for image denoising and deblurring, focusing on frequency domain processing. We have successfully implemented and integrated frequency-domain image preprocessing and post-processing methods to fully transfer image information into frequency domain and back. A key feature of our approach is the use of a UNet architecture enhanced with attention mechanisms, specifically tailored for handling tensors in the frequency domain. This method not only preserves the integrity of the image but also significantly improves the denoising and deblurring process.

However, it is important to note that this project represents just the beginning of exploring the vast potential of frequency domain image processing. Future improvements could include leveraging the symmetric properties of the Fast Fourier Transform (FFT) results to reduce tensor sizes, enhancing computational efficiency, and further refining image quality. Additionally, designing neural networks specifically tailored for the frequency domain could also significantly boost performance by optimizing processing techniques for this unique context. In essence, our work lays the foundation for a promising new solution in image processing, one with substantial room for advancement and optimization.

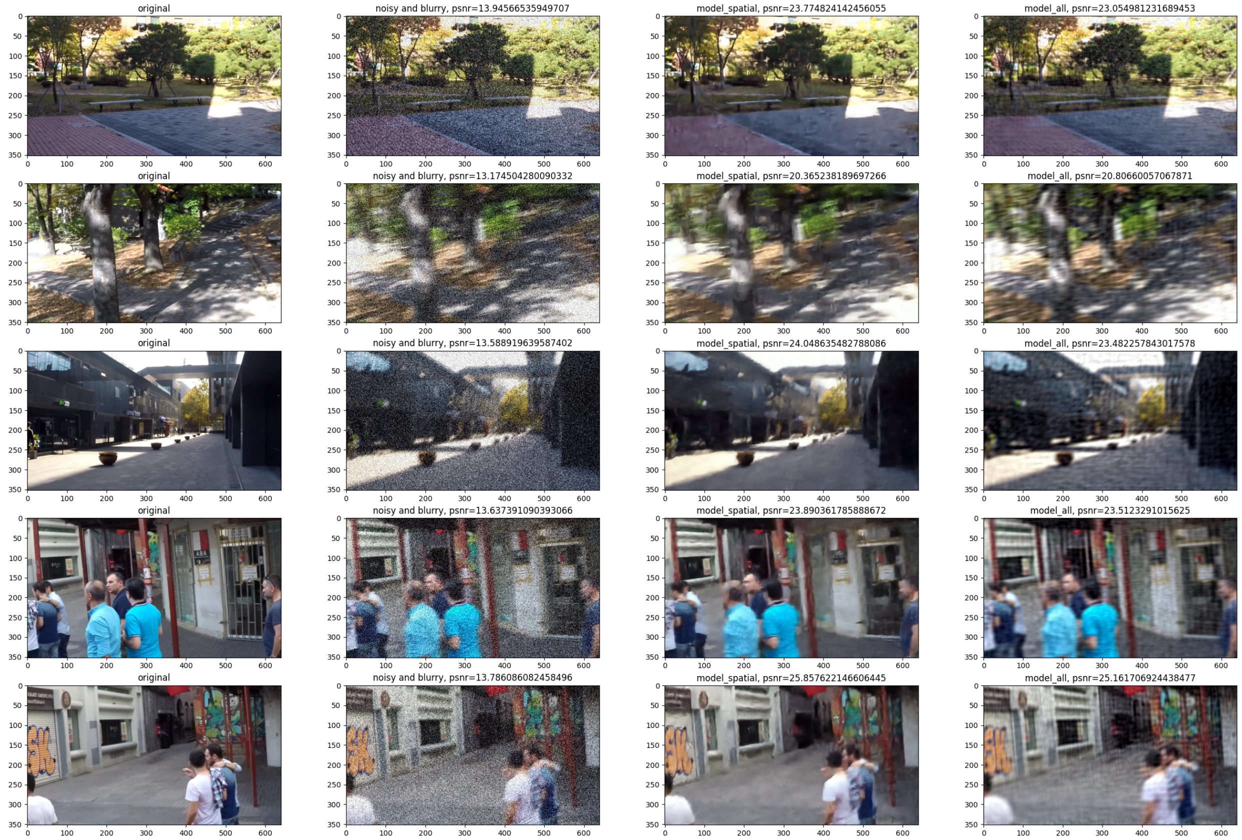
original
noisy and blurry, psnr=13.94566535949707
model_spatial, psnr=23.774824142456055
model_all, psnr=23.054981231689453
original
noisy and blurry, psnr=13.174504280090332
model_spatial, psnr=20.365238189697266
model_all, psnr=20.80660057067871
original
noisy and blurry, psnr=13.588919639587402
model_spatial, psnr=24.048635482788086
model_all, psnr=23.482257843017578
original
noisy and blurry, psnr=13.637391090393066
model_spatial, psnr=23.890361785888672
model_all, psnr=23.5123291015625
original
noisy and blurry, psnr=13.786086082458496
model_spatial, psnr=25.857622146606445
model_all, psnr=25.161706924438477

Fig. 9. Comparison between UNet in spatial domain and FA-UNet.

Acknowledgments

We would like to thank Professor Lindell for the informative lectures that helped shape our project idea, and our teaching assistant Anagh Malik for his helpful guidance and practical suggestions that contributed to the development of our project.

Appendix A

L2 Norm for Denoising: The L2 norm is adept at handling Gaussian noise, commonly found in noisy images. By focusing on the squared differences, it effectively reduces overall noise while maintaining image integrity. This choice is ideal for denoising, as it smooths out random variations without significantly altering the core image structure.

L1 Norm for Deblurring: On the other hand, L1 norm excels in preserving edges and finer details, which are crucial in deblurring tasks. It's less sensitive to outliers, a common challenge in deblurring, enabling the model to focus on the essential elements like edges and textures without being misled by extreme pixel values.

References

[1] L. Fan, F. Zhang, H. Fan, and C. Zhang, "Brief review of image denoising techniques," *Visual Computing for Industry, Biomedicine, and Art*, vol. 2, pp. 1–12, 2019.

[2] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, "Deep learning on image denoising: An overview," *Neural Networks*, vol. 131, pp. 251–275, 2020.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[4] S. V. M. Sagheer and S. N. George, "A review on medical image denoising algorithms," *Biomedical signal processing and control*, vol. 61, p. 102036, 2020.

[5] J. Song, J.-H. Jeong, D.-S. Park, H.-H. Kim, D.-C. Seo, and J. C. Ye, "Unsupervised denoising for satellite imagery using wavelet directional cyclegan," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 8, pp. 6823–6839, 2020.

[6] P. Kaur, G. Singh, and P. Kaur, "A review of denoising medical images using machine learning approaches," *Current medical imaging*, vol. 14, no. 5, pp. 675–685, 2018.

[7] H. Dahlberg, D. Adler, and J. Newlin, "Machine-learning denoising in feature film production," in *ACM SIGGRAPH 2019 Talks*, 2019, pp. 1–2.

[8] E. O. Brigham, *The fast Fourier transform and its applications*. Prentice-Hall, Inc., 1988.

[9] I. W. Selesnick, H. L. Graber, D. S. Pfeil, and R. L. Barbour, "Simultaneous low-pass filtering and total variation denoising," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1109–1124, 2014.

[10] K.-H. Yap, L. Guan, S. W. Perry, and H. San Wong, *Adaptive image processing: a computational intelligence perspective*. Crc Press, 2018.

[11] H.-C. Shih, "A survey of content-aware video analysis for sports," *IEEE Transactions on circuits and systems for video technology*, vol. 28, no. 5, pp. 1212–1231, 2017.

[12] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM computing surveys (CSUR)*, vol. 54, no. 10s, pp. 1–41, 2022.

[13] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[14] S. Wang, B. Huang, T. H. Wong, J. Huang, and H. Deng, "Clga net: Cross layer gated attention network for image dehazing." *Computers, Materials & Continua*, vol. 74, no. 3, 2023.

[15] A. E. Ilesanmi and T. O. Ilesanmi, “Methods for image denoising using convolutional neural network: a review,” *Complex & Intelligent Systems*, vol. 7, no. 5, pp. 2179–2198, 2021.

[16] R. Lan, H. Zou, C. Pang, Y. Zhong, Z. Liu, and X. Luo, “Image denoising via deep residual convolutional neural networks,” *Signal, Image and Video Processing*, vol. 15, pp. 1–8, 2021.

[17] W. Shi, F. Jiang, S. Zhang, R. Wang, D. Zhao, and H. Zhou, “Hierarchical residual learning for image denoising,” *Signal Processing: Image Communication*, vol. 76, pp. 243–251, 2019.

[18] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, “H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes,” *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.

[19] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, “Swin-unet: Unet-like pure transformer for medical image segmentation,” in *European conference on computer vision*. Springer, 2022, pp. 205–218.

[20] H. Zhang, Q. Lian, J. Zhao, Y. Wang, Y. Yang, and S. Feng, “Ratunet: residual u-net based on attention mechanism for image denoising,” *PeerJ Computer Science*, vol. 8, p. e970, 2022.

[21] A. Chakrabarti, “A neural approach to blind motion deblurring,” in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*. Springer, 2016, pp. 221–235.

[22] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, “Learning to deblur,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 7, pp. 1439–1451, 2015.

[23] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia, “Scale-recurrent network for deep image deblurring,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8174–8182.

[24] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, “Deblurgan: Blind motion deblurring using conditional adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8183–8192.

[25] Z. Shen, W. Wang, X. Lu, J. Shen, H. Ling, T. Xu, and L. Shao, “Human-aware motion deblurring,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 5572–5581.