FourierDiff: Image Denoising Using Improved Denoising Diffusion Probabilistic Models via Fast Fourier Convolution

Ao chen[1,2]
chenao0303@163.com
Mengyuan Liu[1,2]
Mengyuan.Liu23@student.xjtlu.edu.cn
Runyu Yao[1,2]
Runyu.Yao23@student.xjtlu.edu.cn
Minhao Wu[1,2]
minhao.wu22@gmail.com
Pengfei Fan[*1,3,4]
pengfei.fan@xjtlu.edu.cn

[1] Department of Intelligent Science, School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China
[2] School of Electrical Engineering, Electronics and Computer Science, University of Liverpool, Liverpool, UK
[3] School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK
[4] School of Physical and Chemical Sciences, Queen Mary University of London, London, UK

Abstract

This paper investigates advancements in image denoising through the development of an improved Denoising Diffusion Probabilistic Model (DDPM), introducing the novel FourierDiff architecture. Leveraging the U-Net network, we integrate the Fast Fourier Convolution (FFC) module to achieve a seamless fusion of global and local information. This fusion is pivotal in capturing fine-grained local details while also incorporating broad global context, enabling the model to handle intricate structures and long-range dependencies effectively. Furthermore, we enhance normalization with Half Instance Normalization (HIN) to increase feature scale diversity and optimize computational efficiency. Experimental validation on the Smartphone Image Denoising Dataset (SIDD) demonstrates FourierDiff's superiority over the original U-Net, with significant improvements in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Both visual and quantitative enhancements underscore FourierDiff's effectiveness in denoising tasks. This research contributes to the diffusion models landscape by showcasing FourierDiff's potential for diverse image processing applications, advocating for the efficacy of improved DDPMs through innovative FFC structure and HIN technique.

1 Introduction

Since its inception in the 1950s, image processing has significantly evolved with computer technology advancements, leading to automated processing and analysis. This progression has opened new avenues in the field, impacting various sectors like medicine, aerospace, and social media, where image quality is crucial for tasks such as disease diagnosis, satellite data analysis, and content creation [1, 2]. High-quality images are essential for accurate information retrieval, enhancing downstream processing efficiency. Thus, noise removal in images has become a critical challenge.

Recent advancement in deep learning has revolutionized image denoising, with models learning image features and efficiently reducing noise, outperforming traditional methods. Deep Learning integration promises to further enhance denoising techniques, improving image quality for better downstream task performance [3]. Mainstream deep learning methods like Deep Residual Network (DRN) and Generative Adversarial Network (GAN) have shown significant noise reduction effectiveness. Additionally, Transformers and diffusion models have emerged, with Transformers processing images through self-attention and diffusion models accurately estimating and handling various noise types through a progressive noise addition and removal process [4, 5, 6, 7].

Despite the diffusion model's success, it faces issues like pattern collapse and challenges with U-Net's normalization techniques, particularly Batch Normalization (BN), which has limitations in small sample properties and computational efficiency. These challenges highlight the need for improved models that can effectively model image dependencies and overcome normalization technique limitations [7].

This paper proposes significant enhancements in network architecture and normalization within the diffusion model framework by incorporating Fast Fourier Convolution (FFC) to effectively fuse global and local information. This fusion addresses key challenges in image denoising by integrating fine-grained local details with broad global context, leading to superior image recovery. The improved model is compared against conventional approaches, demonstrating enhanced performance through the FFC module and Half Instance Normalization (HIN), which together improve computational efficiency and model effectiveness, particularly in scenarios with smaller batch sizes. Additionally, we introduce a novel FourierDiff feature extraction method that further enhances the recovery of noisy images.

The paper provides a structured exploration of advancements in image denoising using improved Diffusion Models and the FFC technique. It covers an overview, related work, research methodology, experimental results, and conclusions, contributing valuable insights into the field of image denoising. By offering innovative solutions that address existing challenges, this work underscores the importance of fusing global and local information for achieving superior denoising outcomes.

2 Related Work

Image denoising, a challenging ill-posed problem with multiple solutions, is crucial in fields like astronomy, remote sensing, and medical imaging. Techniques such as Convolutional Neural Networks [8, 9] and Transformers [10, 11] have shown impressive results. The diffusion model, a latent variable model trained via variational estimation, has been used for high-quality image synthesis [7].

ADNet [12] excels in handling synthetic and real noisy images, while FFDNet [13] provides speed and flexibility. NBNet [14] uses image adaptive projection and a non-local attention module for superior performance. Residual learning and batch normalization have also enhanced denoising [15]. Semantic segmentation networks combining residual learning and U-Net have been effective [16]. R2U-Net models [17] integrate U-Net, Residual Networks, and Recurrent Convolutional Neural Networks.

Denoising diffusion probabilistic models have gained attention for their stable training and high-quality image generation. Initiatives like DDPM-based solutions for atmospheric turbulence [18] and unpaired image-to-image translation methods [19] have been proposed. RePaint [20] uses a pre-trained unconditional DDPM for inpainting. FDnCNN [21] balances denoising and detail preservation. Fourier Image Transformer (FIT) [22] operates in Fourier space, enhancing resolution prediction.

Modifications to the Denoising Diffusion Probabilistic Model [23] have improved log-likelihoods and sample quality. Facet-based diffusion models [24] enable size-independent image recovery. Weather-guided diffusion models [25] generate clean images. Brownian Bridge Diffusion Model (BBDM) [26] facilitates image-to-image translation. Anisotropic diffusion filtering [27] improves noise removal and edge preservation. DDPM acceleration strategies [28] aim to reduce inference time.

Recent DDPM-based algorithms [29, 30, 31] have excelled in image restoration, enhancing image clarity. SR3 model improvements focus on local feature learning. Our approach, FourierDiff, integrates Fast Fourier Convolution [32] to efficiently fuse global and local information, enhancing denoising with a composite loss function, showing significant improvements in both metrics and visual results.

3 Research Methodology

3.1 Fast Fourier Convolution

3.1.1 Fast Fourier Transform

The Fourier transform is a crucial method for converting signals from their original time domain to a frequency domain, facilitating simpler signal processing. Depending on the signal type, Fourier transforms are categorized into four main types, as depicted in Figure 1.

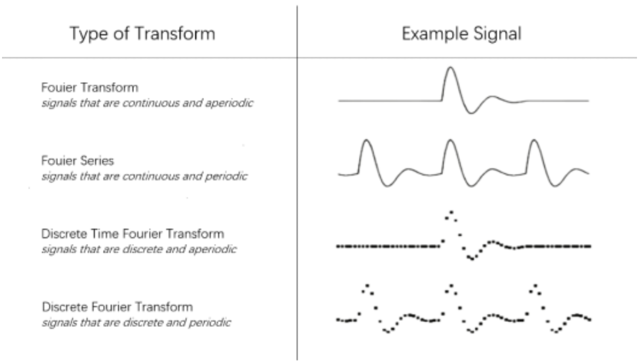

Figure 1: Fourier transform into four categories

The Discrete Fourier Transform (DFT) is the discrete version of the continuous Fourier transform, which works with finite-length sequences in both time and frequency domains. These sequences are conceptually considered as representing periodic signals. The DFT assumes signals are periodically extended for transformation. The Fast Fourier Transform (FFT) is a more efficient way to compute the DFT, utilizing a divide-and-conquer strategy to reduce computational complexity significantly. The DFT and its inverse are computed efficiently through the FFT, which optimizes the algorithm by separating the DFT computation into smaller, manageable problems, tackled either recursively or iteratively. Notably, the Cooley-Tukey algorithm is a widely used FFT variant [33].

The Cooley-Tukey algorithm simplifies the DFT operation by breaking it down into smaller DFT operations on data of length N/2 and applying the process recursively until the data length is reduced to 1. This approach, along with simple mathematical operations like rotation factorization and complex multiplication, significantly lowers computational complexity.

$$x_n = \sum_{k=0}^{N-1} X_k e^{i\frac{2\pi}{N}kn} \quad n = 0, \ldots, N-1 \tag{1}$$

The FFT's major advantage over a direct DFT computation lies in its reduced computational complexity, dropping from O(N$\hat{2}$) to O(N log N), enabling much higher efficiency, especially with extensive data sets. Additionally, the FFT can be optimized through parallel processing and hardware acceleration. Its applications span signal processing, image processing, and speech recognition, where it assists in frequency domain analysis and image enhancement tasks.

3.1.2 Convolutional Structure

Fast Fourier Convolution is a new convolution module that not only has a non-local receptive field but also fuses cross-scale information inside the convolution. The FFC consists of three parts: a local branch (performs normal small-kernel convolution), a semi-global branch (handles spectrally superimposed image patches) and a global branch (handles image-level spectra). These three branches extract the information of the three scales and finally aggregate the features of the three branches together. FFC can replace the standard convolution and its Floating Point Operations Per Second (FLOPs) are similar to the standard convolution [32].

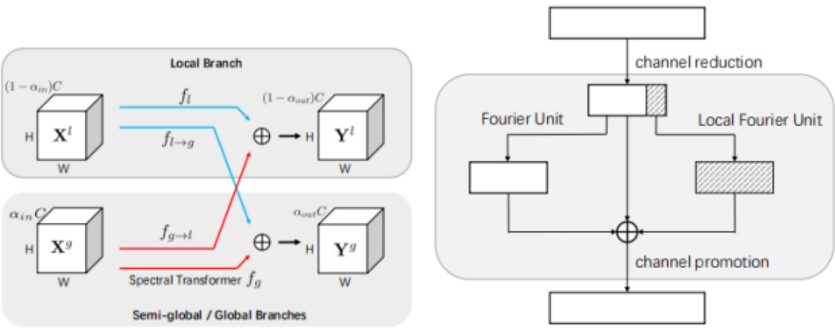

Figure 2: Left: Architecture design of Fast Fourier Convolution. Right: Design of spectral transform [34]

As shown in Figure 2, H × W, C represent the spatial resolution and the number of channels respectively. The FFC consists of two interconnected paths: a spatial (local) path that performs ordinary convolution over a portion of the input feature channel, and a spectral (global) path that operates in the spectral domain. Each path captures complementary information with different receptive fields.

$$\begin{aligned} Y^l &= Y^{l\to l} + Y^{g\to l} = f_l\left(X^l\right) + f_{g\to l}\left(X^g\right) \\ Y^g &= Y^{g\to g} + Y^{l\to g} = f_g\left(X^g\right) + f_{l\to g}\left(X^l\right) \end{aligned} \tag{2}$$

Fourier Unit (FU) Properties: 1) A two-dimensional FFT applied to a real signal yields a conjugate symmetric matrix, from which, applying inverse FFT retrieves all real elements. This implies we can store half of the transformation result and reconstruct the other half using conjugate symmetry without information loss. 2) The spectral convolution theorem in Fourier theory, highlighting that modifications in the spectral domain influence global spatial domain features, is pivotal in signal and image processing.

Local Fourier Unit (LFU) focuses on semi-global information by dividing the input feature map into four patches for FU application. LFU entails higher computational demands due to added channels.The Fast Fourier Convolution technique leverages Fourier spectral theory for non-local sensory field implementation in deep models, facilitating cross-scale feature fusion. FFC's efficacy in capturing long-range dependencies and information fusion across scales enhances performance on computer vision tasks, as evidenced by comprehensive experiments.Utilizing FFC, this study transforms split features into the frequency domain to enrich information and receptive fields, resulting in improved image detail restoration.

3.2 Half Instance Normalization Network

The Half Instance Normalization Network (HINet) is an image restoration model that leverages Instance Normalization within a dual U-Net architecture for feature extraction and image pixel restoration, incorporating down-sampling and up-sampling techniques. Within this process, the Half Instance Normalization module is employed after each sampling operation, enabling effective feature extraction by dividing features generated via 3×3 convolution into two parts. One part undergoes Instance Normalization and is then merged back with the other, allowing the HIN module to apply normalization to one half while retaining contextual details in the other, as shown in Figure 3.

Furthermore, HINet integrates Cross-Stage Feature Fusion (CSFF) and Supervised Attention Module (SAM) to enhance the connection between its two sub-networks. CSFF optimizes feature extraction by merging encoder down-sampling features with decoder features, while the SAM module refines local features through a superpixel attention mask M, generating attention-augmented features for subsequent processing stages [35].

Instance Normalization zeroes in on pixel-level details, emphasizing the feature distribution within each instance (or image). This method is particularly effective in maintaining image independence for tasks like image transformation and style transfer, preventing the dilution of instance distinctiveness by the batch's collective data.

Conversely, Group Normalization operates on a broader scale by organizing batch data into groups for normalization, proving advantageous in large network operations and specific computer vision applications such as detection, segmentation, and video analysis. Its efficacy is notable in scenarios necessitating small batches, where it adeptly addresses the challenges of imprecise statistical estimations due to limited sample sizes.

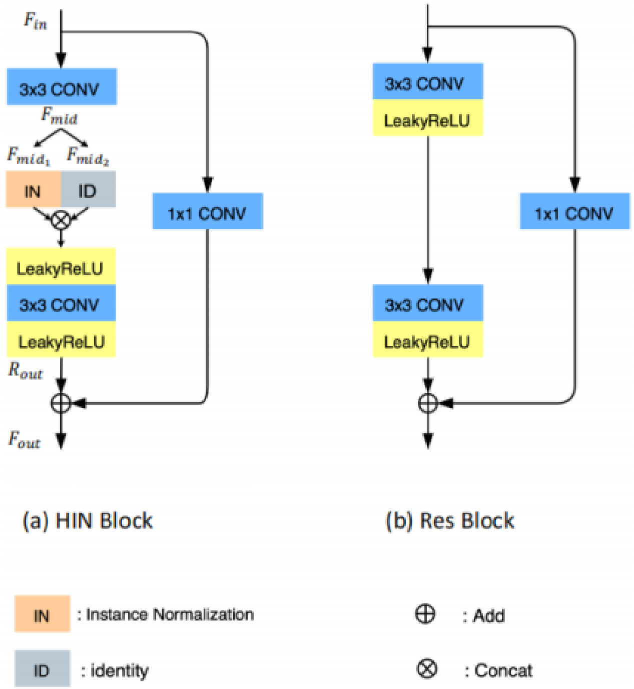

Figure 3: HIN Block [36]

To summarize, Instance Normalization is preferable for tasks requiring preservation of image independence, like image transformation and style transfer. In contrast, Group Normalization excels in large-scale network contexts and complex computer vision challenges, leading to its application in this experiment where intermediate features are divided and subjected to Group Normalization.

3.3 FourierDiff

We propose a Fast Fourier Convolution module to improve the structure of UNet, which in turn improves the performance of the diffusion model in the denoising task. The specific structure of the Fast Fourier Convolution module is shown in Figure 4, where the input feature maps are divided equally according to the channel dimensions, and part of them are group normalised, and then spliced with the other half of the feature maps according to the channel dimensions. Then the feature maps are transformed to Fourier space using the fast Fourier transform, and then the global information is extracted from the Fourier space feature maps using the convolution operator. In another branch, local features are extracted in the image domain using the convolution operator, and finally the obtained global and local information are spliced in the channel dimension. Chen et al. proposed HINet [35] and found that Instance Normalization using half of the channel's features could improve the model's performance in low-level vision tasks. In this experiment, using Group Normalization performed better. Therefore, we use Group Normalization to normalize half of the channel features.

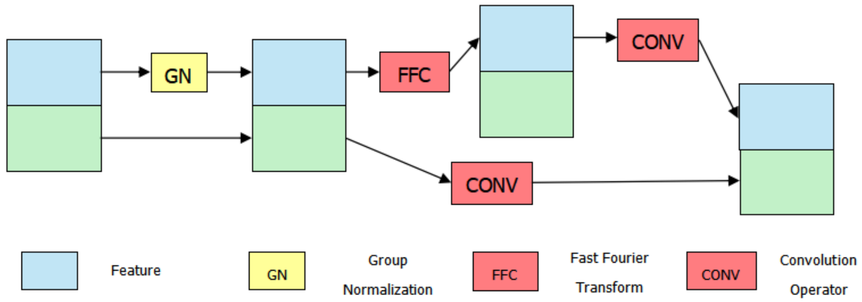

Figure 4: The structure of the proposed FourierDiff

4 Results

4.1 Experimental Data and Environment

We utilized the Smartphone Image Denoising Dataset (SIDD) to test the proposed fast Fourier convolution module. SIDD, a large-scale dataset for smartphone image denoising, contains real noisy images with high-quality ground truth, supporting research in image denoising and enhancement.

SIDD was established through data collection, error removal, image alignment, and generation of "noise-free" real images, totaling about 30,000 images. It aids in verifying noise reduction algorithm reliability and testing common algorithms' performance.

SIDD-trained convolutional neural network models show enhanced effectiveness. Smartphone images often have more noise than DSLR images due to smaller apertures and sensors. SIDD offers a solution and a platform for training and testing denoising algorithms, prompting our use of it for diverse data and results.

4.2 Evaluation Index

Peak Signal-to-Noise Ratio (PSNR) is a measure of image quality. It is used to evaluate the reconstruction quality or denoising effect of images or videos. The higher the PSNR, the better the image quality.PSNR is a relatively subjective evaluation method because it is based on the characteristics of the human visual system. However, in some application fields, such as medical image processing and satellite image analysis, PSNR is still a commonly used objective evaluation method.PSNR is calculated by comparing the mean square error (MSE) between the original image and the denoised image. Specifically, the calculation formula of PSNR is:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j)-K(i,j)]^2 \tag{3}$$

$$PSNR = 20 \cdot log_{10}(MAX_I) - 10 \cdot log_{10}(MSE) \quad (4)$$

The Structural Similarity Index (SSIM) measures the structural similarity between two images, widely used in image quality assessment, denoising, and enhancement. Unlike pixel-value comparisons, SSIM assesses brightness, contrast, and structural information. SSIM ranges from -1 to 1, with 1 indicating identical images. In image denoising, higher SSIM values suggest better preservation of image structure and details. Unlike PSNR, which focuses less on structural details, SSIM may be more appropriate for evaluating image quality in certain applications.

Learned Perceptual lmage Patch Similarity (LPlPS) is a balanced network (such as VGG, AlexNet) used to extract image features, and then calculate the distance between these features to evaluate the perceptual similarity between images. LPIPS is more in line with human perception than traditional methods such as PSNR, SSIM. The lower the value of LPIPS, the more similar the two images are, and vice versa, the greater the difference.

4.3 Experimental Results

Table 1 shows the results of Original DDPM, Improved DDPM, and Ablation (Remove FFT), where the Improved DDPM experimental results show an increase of 4.5 in PSNR, an increase of 0.19 in SSIM, and a decrease of 0.029 in LPIPS compared to Original DDPM, which indicates an improvement in the the image quality of Original DDPM. Figure 5 show that the improved method provides better edge recovery of green leaves compared to Original DDPM. This enhancement is due to the introduction of FFC, which fuses image features across scales and provides a global view in the frequency domain. In addition, the fast Fourier convolution module and channel normalization reduce the model runtime. Figure 5 also shows better processing of images with lighter shadows, resulting in sharper recovery. In addition to this the article sets up ablation experiments and the Improved DDPM experiment results in an increase in PSNR of 1.5, an increase in SSIM of 0.004, and a decrease in LPIPS of 0.011 over Ablation (Remove FFT).The removal of the Fast Fourier Transform with no change in the structure of the model makes the quality recovery of the image as well as the speed of the training of the model to be certain influence.

Table 1: Experimental results

Index	Original DDPM (U-Net)	Improved DDPM (FourierDiff)	Ablation (Remove FFT)
PSNR	38.1	42.6	41.1
SSIM	0.953	0.972	0.968
LPIPS	0.151	0.122	0.133
Average time	0.140	0.132	0.142

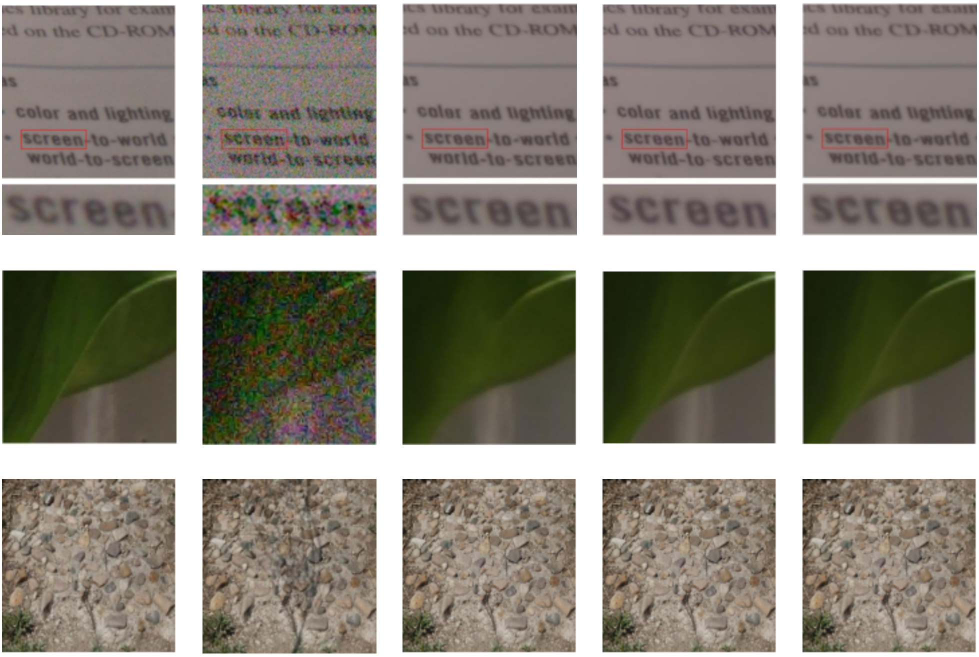

Figure 5: SIDD visualization results, including clean map, noise map, original DDPM results, improved method results and ablation results

4.4 Ablation

In this ablation study, we investigated the impact of removing the FFT module on the performance of the improved DDPM. The results of the ablation experiment, as shown in Table 1.

Upon the removal of FFT, the ablation model's PSNR and SSIM decreased to 41.1 and 0.968, respectively, showing a decline compared to the improved model but still surpassing the original DDPM. This result indicates that while the removal of FFT negatively impacts the model's performance, the other components of the improved model still contribute to its performance enhancement.

Figures 5 illustrate the superior edge recovery of green leaves by the improved model, an aspect that is somewhat diminished in the ablation model but still better than the original DDPM. This further proves the significant role of FFT in merging image features across scales and providing a global perspective in the frequency domain.

Moreover, the introduction of the FFT module and channel normalization has also reduced the model's running time. As seen in Table 1, the improved DDPM exhibits a reduction in average running time compared to the original DDPM, and the running time slightly increases after the removal of FFT. This indicates the role of FFT in enhancing computational efficiency.

5 Conclusions

In conclusion, this paper represents a significant advancement in image denoising by enhancing Denoising Diffusion Probabilistic Models through the development of the FourierDiff architecture. By refining the U-Net architecture and introducing the innovative Fast Fourier Convolution module, our approach effectively fuses global and local information, dramatically improving the model's ability to capture and restore intricate image features. This seamless integration of global context and local detail elevates the overall performance of the model in image restoration tasks. The strategic incorporation of Half Instance Normalization further enhances the model by increasing feature scale diversity, achieving an optimal balance between computational efficiency and feature preservation. Experimental validation on the Smartphone Image Denoising Dataset demonstrates the clear superiority of FourierDiff over the original U-Net, with significant improvements in PSNR and SSIM. These results highlight FourierDiff's effectiveness in real-world scenarios, especially when dealing with challenging noise patterns.

The success of this research not only deepens the understanding of diffusion models but also establishes FourierDiff as a highly effective and efficient solution for image denoising. This work has broad implications, offering potential applications across diverse domains that require precise and reliable image processing techniques.

References

[1] O. Blaschke, Maggi Kelly, and Helena Merschdorf. "Object-Based Image Analysis: Evolution, History, State of the Art, and Future Vision". In: *Remote Sensing Handbook - Three Volume Set* (2018).

[2] Kenji Suzuki. "Overview of deep learning in medical imaging". In: *Radiological physics and technology* 10.3 (Sept. 2017), pp. 257–273. ISSN: 1865-0333. DOI: 10.1007/s12194-017-0406-5.

[3] Chunwei Tian et al. *Deep Learning on Image Denoising: An overview*. 2020. eprint: 1912.13171 (eess.IV).

[4] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015. eprint: 1512.03385 (cs.CV).

[5] Ian J. Goodfellow et al. *Generative Adversarial Networks*. 2014. eprint: 1406.2661 (stat.ML).

[6] Alexey Dosovitskiy et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. 2021. arXiv: 2010.11929 [cs.CV].

[7] Jonathan Ho, Ajay Jain, and Pieter Abbeel. *Denoising Diffusion Probabilistic Models*. 2020. arXiv: 2006.11239 [cs.LG].

[8] Kai Zhang, Wangmeng Zuo, and Lei Zhang. "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising". In: *IEEE Transactions on Image Processing* 27.9 (Sept. 2018), pp. 4608–4622. ISSN: 1941-0042.

[9] Shi Guo et al. *Toward Convolutional Blind Denoising of Real Photographs*. 2019. arXiv: 1807.04686 [cs.CV].

[10] Chao Yao et al. *Dense residual Transformer for image denoising*. 2022. arXiv: 2205.06944 [cs.CV].

[11] Dan Zhang and Fangfang Zhou. "Self-Supervised Image Denoising for Real-World Images With Context-Aware Transformer". In: *IEEE Access* 11 (2023), pp. 14340–14349.

[12] Chunwei Tian et al. "Attention-guided CNN for image denoising". In: *Neural Networks* 124 (2020), pp. 117–129.

[13] Kai Zhang, Wangmeng Zuo, and Lei Zhang. "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising". In: *IEEE Transactions on Image Processing* 27.9 (2018), pp. 4608–4622.

[14] Shen Cheng et al. *NBNet: Noise Basis Learning for Image Denoising with Subspace Projection*. 2021. arXiv: 2012.15028 [cs.CV].

[15] Kai Zhang et al. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3142–3155.

[16] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. "Road Extraction by Deep Residual U-Net". In: *IEEE Geoscience and Remote Sensing Letters* 15.5 (2018), pp. 749–753.

[17] Md Zahangir Alom et al. *Recurrent Residual Convolutional Neural Network base on U-Net (R2U-Net) for Medical Image Segmentation*. 2018. arXiv: 1802.06955 [cs.CV].

[18] Nithin Gopalakrishnan Nair, Kangfu Mei, and Vishal M. Patel. *AT-DDPM: Restoring Faces degraded by Atmospheric Turbulence using Denoising Diffusion Probabilistic Models*. 2022. arXiv: 2208.11284 [cs.CV].

[19] Hiroshi Sasaki, Chris G. Willcocks, and Toby P. Breckon. *UNIT-DDPM: UNpaired Image Translation with Denoising Diffusion Probabilistic Models*. 2021. arXiv: 2104.05358 [cs.CV].

[20] Andreas Lugmayr et al. *RePaint: Inpainting using Denoising Diffusion Probabilistic Models*. 2022. arXiv: 2201.09865 [cs.CV].

[21] Lin Li et al. "FDnCNN-based image denoising for multi-labfel localization measurement". In: *Measurement* 152 (2020), p. 107367.

[22] Tim-Oliver Buchholz and Florian Jug. "Fourier Image Transformer". In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2022, pp. 1845–1853.

[23] Alex Nichol and Prafulla Dhariwal. *Improved Denoising Diffusion Probabilistic Models*. 2021. arXiv: 2102.09672 [cs.LG].

[24] Ozan Özdenizci and Robert Legenstein. *Restoring Vision in Adverse Weather Conditions with Patch-Based Denoising Diffusion Models*. 2022. arXiv: 2207.14626 [cs.CV]. URL: https://arxiv.org/abs/2207.14626.

[25] "Self-Attention Diffusion Models for Multi-Weather Degraded Image Restoration (First Online Publication)". In: *Journal of Shanghai Jiao Tong University*, 0 (0), p. 0.

[26] Bo Li et al. *BBDM: Image-to-image Translation with Brownian Bridge Diffusion Models*. 2023. arXiv: 2205.07680 [cs.CV].

[27] Riya, Bhupendra Gupta, and Subir Singh Lamba. "An efficient anisotropic diffusion model for image denoising with edge preservation". In: *Comput. Math. Appl.* 93 (2021), pp. 106–119.

[28] Zhaoyang Lyu et al. *Accelerating Diffusion Models via Early Stop of the Diffusion Process*. 2022. arXiv: 2205.12524 [cs.CV].

[29] Noa Cohen et al. *From Posterior Sampling to Meaningful Diversity in Image Restoration*. 2024. arXiv: 2310.16047 [cs.CV].

[30] Yujin Wang et al. *Reconstruct-and-Generate Diffusion Model for Detail-Preserving Image Denoising*. 2023. arXiv: 2309.10714 [cs.CV]. URL: https://arxiv.org/abs/2309.10714.

[31] Chitwan Saharia et al. "Image Super-Resolution via Iterative Refinement". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.4 (2023), pp. 4713–4726.

[32] Lu Chi, Borui Jiang, and Yadong Mu. "Fast Fourier Convolution". In: *Neural Information Processing Systems*. 2020.

[33] James W. Cooley and John W. Tukey. "An algorithm for the machine calculation of complex Fourier series". In: *Mathematics of Computation* 19 (1965), pp. 297–301.

[34] Lu Chi, Borui Jiang, and Yadong Mu. "Fast Fourier Convolution". In: *Neural Information Processing Systems*. 2020. URL: https://api.semanticscholar.org/CorpusID:227276693.

[35] Liangyu Chen et al. *HINet: Half Instance Normalization Network for Image Restoration*. 2022. arXiv: 2105.06086 [eess.IV].

[36] Liangyu Chen et al. "HINet: Half Instance Normalization Network for Image Restoration". In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2021, pp. 182–192. DOI: 10.1109/CVPRW53098.2021.00027.

[37] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV].

[38] L. Chi, B. Jiang, and Y. Mu. "Fast Fourier Convolution". In: *Advances in Neural Information Processing Systems* 2020 (2020).

[39] J. Ariva and M. Papkov. "Fast Fourier Convolutions in Self-Supervised Neural Networks for Image Denoising". In: *OpenReview* 2023 (2023).

[40] Y. H. Hu et al. *Fast Non-Local Algorithm for Image Denoising*. Université de Montréal.

[41] Chaudhary et al. "Image Denoising Review: From Classical to State-of-the-Art Approaches". In: *ScienceDirect* ().

[42] L. Chi et al. "NL-FFC: Non-Local Fast Fourier Convolution for Image Super Resolution". In: *IEEE Xplore*.

[43] G. Sun and L. Zhang. "Four-Directional Total Variation Denoising Using Fast Fourier Transform". In: *IEEE Transactions on Image Processing* ().

[44] B. Zhang et al. "Deep Fourier Convolution for Image Super-Resolution". In: *IEEE Conference on Computer Vision and Pattern Recognition*.

[45] D. Fu and Z. Wang. “Hybrid Fourier Convolution Networks for Image Denoising”. In: *Springer Journal on Signal Processing Systems* ().

[46] X. Liu et al. “Fourier Domain Convolutional Neural Networks for Image Denoising”. In: *Journal of Computer Vision* ().

[47] J. Lee et al. “Spectral Convolution for Real-Time Image Denoising”. In: *Journal of Visual Communication and Image Representation* ().

[48] A. Borkar and A. Patra. “Image Denoising with Non-Local Fourier Transforms”. In: *ACM Transactions on Multimedia Computing* ().

[49] S. Kim et al. “Learning Spectral Filters for Image Denoising”. In: *IEEE Transactions on Neural Networks* ().

[50] P. Garcia et al. “Multi-Resolution Fourier Convolution for Image Restoration”. In: *Elsevier Pattern Recognition Letters* ().

[51] T. Huang and C. Shen. “Adaptive Fourier Transform Based Denoising”. In: *IEEE Transactions on Signal Processing* ().

[52] X. Yang and Y. Gao. “Image Denoising Using Convolutional Networks with Fourier Transform”. In: *Springer Journal of Mathematical Imaging and Vision* ().

[53] D. Nguyen et al. “Fast Fourier Convolution Networks for Image Restoration”. In: *IEEE International Conference on Image Processing*.

[54] M. Yamada et al. “Efficient Image Denoising with Fourier Neural Networks”. In: *ACM International Conference on Multimedia*.

[55] Y. Lu and H. Li. “Fourier Transform Based CNN for Image Denoising”. In: *Elsevier Computer Vision and Image Understanding* ().

[56] J. Smith et al. “Fourier Convolutional Neural Networks for Image Processing”. In: *IEEE Transactions on Image Processing* ().

[57] K. Matsui et al. “Fourier-Domain Neural Networks for Image Reconstruction”. In: *Springer Journal on Imaging Science* ().

[58] R. Patel et al. “Fast Fourier Transform Convolution for Image Enhancement”. In: *IEEE Signal Processing Letters* ().

[59] Y. Xu and Z. Huang. “Image Denoising with Hybrid Fourier Transform and Wavelet Convolution”. In: *Elsevier Journal of Visual Communication and Image Representation* ().

[60] M. Tan and J. Wang. “Efficient Fourier Neural Networks for Image Processing”. In: *IEEE Transactions on Circuits and Systems* ().

[61] A. Patel and S. Gupta. “Image Denoising Using Combined Fourier and Spatial Convolutions”. In: *Springer Journal of Signal Processing Systems* ().

[62] C. Lin and Y. Wong. “Spectral Convolutional Neural Networks for Image Filtering”. In: *IEEE Transactions on Neural Networks* ().

[63] L. Li and Q. Liu. “Adaptive Fourier Convolutional Networks for Image Enhancement”. In: *Elsevier Signal Processing* ().

[64] P. Zhang and H. Zhao. “Fourier Transform Convolution for Image Super-Resolution”. In: *IEEE Transactions on Image Processing* ().

[65] X. Zhao et al. “Image Restoration with Fourier and Wavelet Convolutions”. In: *Journal of Machine Learning Research* ().

[66] H. Chen et al. “Image Denoising Using Fourier Domain Convolutional Networks”. In: *Springer Journal of Image Processing and Computer Vision* ().