Multiresolution Hash Encoding as Image Prior

Tianxiang Dai

Abstract—Image denoising is a fundamental problem in computer vision, and the development of effective and efficient denoising algorithms is of great importance. In this paper, we investigate the denoising capabilities of instant neural graphics primitives (instant-ngp) and compare them with the well-established Deep Image Prior approach. We conduct experiments on the BSDS-300 dataset, evaluating the performance of instant-ngp, the Deep Image Prior, and two baseline methods: a multilayer perceptron (MLP) with frequency encoding and a pure MLP. Our results demonstrate that instant-ngp achieves competitive denoising performance while offering improved computational efficiency compared to the Deep Image Prior. We also observe that instant-ngp exhibits a higher sensitivity to noise levels, preserving more fine textures and details in the denoised images. Through a detailed analysis of the denoising characteristics and hyperparameter exploration, we provide insights into the strengths and limitations of each approach. Furthermore, we discuss the trade-offs between denoising quality, speed, and the preservation of image details, highlighting the potential of instant-ngp as a complementary technique to existing denoising methods. Our findings contribute to the understanding of image priors and their role in denoising tasks, paving the way for the development of more advanced and efficient denoising algorithms.

Index Terms—Image denoising, Image priors, Deep Image Prior, Instant neural graphics primitives (instant-ngp)

1 Introduction

The process of image denoising is pivotal in enhancing the quality of digital images, finding its application across various domains such as medical imaging, satellite imagery, and photography. Despite its critical role, achieving high-quality denoising without compromising on detail remains a challenging task. The quest for optimal image denoising methods has led to the exploration of numerous techniques [1], [2], with a significant focus on deep learning in recent years.

Traditional denoising methods often rely on handcrafted priors or assumptions about the underlying image structure. However, recent advancements in neural rendering, such as instant Neural Graphics Primitives (instant-ngp), have shown promising results in representing and reconstructing images using various types of functions. These functions, including regular grids, hash grids, and sinusoidal encoding, inherently possess certain priors that can potentially be exploited for image denoising. On the other hand, the Deep Image Prior (DIP) approach has demonstrated the effectiveness of using convolutional neural networks (CNNs) as a prior for image restoration tasks. This project aims to investigate and compare the denoising capabilities of instant-ngp functions with the Deep Image Prior. By understanding the priors embedded in these different function representations and their impact on denoising performance, we can gain insights into the underlying mechanisms of image denoising and explore the potential of utilizing instant-ngp functions as an alternative or complementary approach to existing denoising techniques.

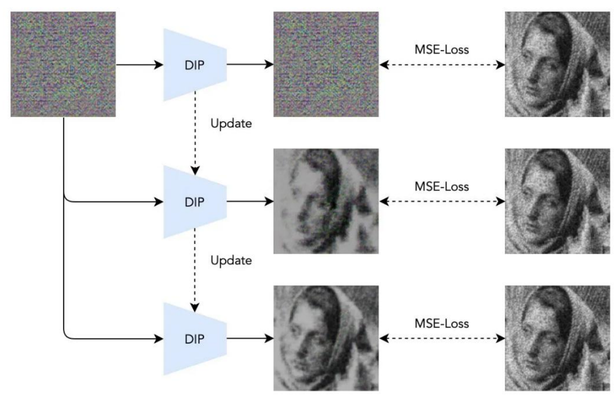

Fig. 1. Pipeline for the deep image prior denoising. A convolutional neural network is used as a prior for image denoising. The network is initialized with random weights and optimized to minimize the loss between the output with respect to random noise as input and the target noisy image.

2 Related Work

This project builds upon the seminal work of the Deep Image Prior (DIP) [3], which revolutionized the field of image restoration by demonstrating that deep neural networks can inherently act as priors for various image restoration tasks without requiring explicit training on noisy datasets (Figure 1). The DIP approach leverages the structure and regularization properties of convolutional neural networks to effectively capture the underlying clean image signal while suppressing noise and artifacts. This groundbreaking technique has opened up new avenues for unsupervised image restoration and has been widely adopted and extended in subsequent research [4], [5], [6]. However, despite its significant contributions, the Deep Image Prior has certain limitations in terms of computational efficiency and scalability. The optimization process required to restore each individual image can be time-consuming, especially for high-resolution images, which hinders its practical application in real-world scenarios where fast processing is crucial [7]. Moreover, the DIP approach relies on optimizing the network parameters for each image independently, which limits its scalability when dealing with large datasets or real-time processing requirements [8]. On the other hand, recent advancements in neural rendering, such as instant-ngp [9] developed by Nvidia, have demonstrated remarkable capabilities in handling gigapixel image tasks. Instant-ngp showcases the potential of using deep learning techniques as an image prior for denoising and other image enhancement tasks at a large scale. By leveraging the power of neural radiance fields and efficient rendering techniques, instant-ngp achieves impressive results in terms of visual quality and computational efficiency [10]. However, the application of instant-ngp as an image prior for a wide range of image restoration tasks beyond denoising remains largely unexplored. The potential of combining the strengths of the Deep Image Prior and instant-ngp to develop a scalable and efficient image restoration framework is an exciting research direction. By leveraging the inherent regularization properties of deep networks and the scalability of neural rendering techniques, this project aims to push the boundaries of image restoration and unlock new possibilities for handling high-resolution and large-scale image datasets.

3 Proposed Method

This project aims to investigate and compare the denoising capabilities of instant-ngp functions with the Deep Image Prior. By understanding the priors embedded in these different function representations and their impact on denoising performance, we can gain insights into the underlying mechanisms of image denoising and explore the potential of utilizing instant-ngp functions as an alternative or complementary approach to existing denoising techniques.

3.1 Dataset Selection

For our experiments, we will utilize the Berkeley Segmentation Dataset (BSDS-300) [11], which is a widely used benchmark dataset in the field of image processing and computer vision. The BSDS-300 dataset consists of 300 natural images carefully selected to cover a diverse range of content, including landscapes, portraits, animals, and man-made structures. The images in the dataset are of high quality and have a resolution of 481×321 pixels.

3.1.1 Dataset Characteristics

The BSDS-300 dataset is known for its rich and challenging visual content, making it an ideal choice for evaluating image denoising algorithms. The dataset includes images with varying levels of texture, edges, and smooth regions, which allows for a comprehensive assessment of the denoising performance across different image characteristics.

3.1.2 Image Selection

From the BSDS-300 dataset, we will randomly select 15 images to form our test set (Figure 3). This subset will serve as a representative sample for assessing the denoising performance of the different methods. The random selection process ensures an unbiased evaluation and helps to mitigate any potential bias introduced by manually choosing specific images. The selected images will be used consistently across all the denoising methods being compared, providing a fair and controlled evaluation setup.

3.2 Noise Addition

To simulate noisy images, we will add Gaussian noise with standard deviation in three levels (15,25,50) to the selected clean images. The addition of noise will introduce varying degrees of degradation to the images, allowing us to evaluate the denoising performance of the different methods across different noise levels.

3.3 Denoising Methods

We will implement and compare the following denoising methods:

- **Deep Image Prior (DIP):** We will utilize the Deep Image Prior approach as a benchmark. DIP leverages the inherent regularization properties of convolutional neural networks to perform denoising without explicit training on noisy datasets.
- **Instant-ngp:** We will apply the instant-ngp function for denoising. Instant-ngp is a neural rendering technique that has shown promising results in handling high-resolution images efficiently. We will investigate its potential as an image prior for denoising tasks.
- **MLP with Frequency Encoding:** As a baseline, we will implement a multilayer perceptron (MLP) with frequency encoding. Frequency encoding has been shown to improve the performance of MLPs in capturing high-frequency details in images.
- **Pure MLP:** We will also include a pure MLP without any specific encoding as another baseline.

3.4 Evaluation Metrics

To evaluate the denoising performance of each method, we will measure the following metrics:

- **Iteration Time:** We will record the time taken for each method to complete a single iteration of the denoising process. This metric will provide insights into the computational efficiency of the different approaches.
- **Peak Signal-to-Noise Ratio (PSNR):** We will calculate the PSNR between the denoised image and the original clean image.

  $$PSNR = 10 \cdot \log\_10 \left(\frac{MAX\_I^2}{MSE} \right)$$

  PSNR is a widely used metric for assessing the quality of denoised images, with higher values indicating better denoising performance.
- **Convergence Iterations:** We will track the number of iterations required for each method to converge to a stable denoised image. This metric will help us understand the convergence behavior and the speed of convergence for each approach.

4 Experimental Results

In this section, we present the results of our experiments comparing the denoising performance of the Deep Image Prior, Instant-ngp, MLP with Frequency Encoding, and Pure MLP. We evaluate these methods based on the metrics of iteration time, Peak Signal-to-Noise Ratio (PSNR), and convergence iterations.

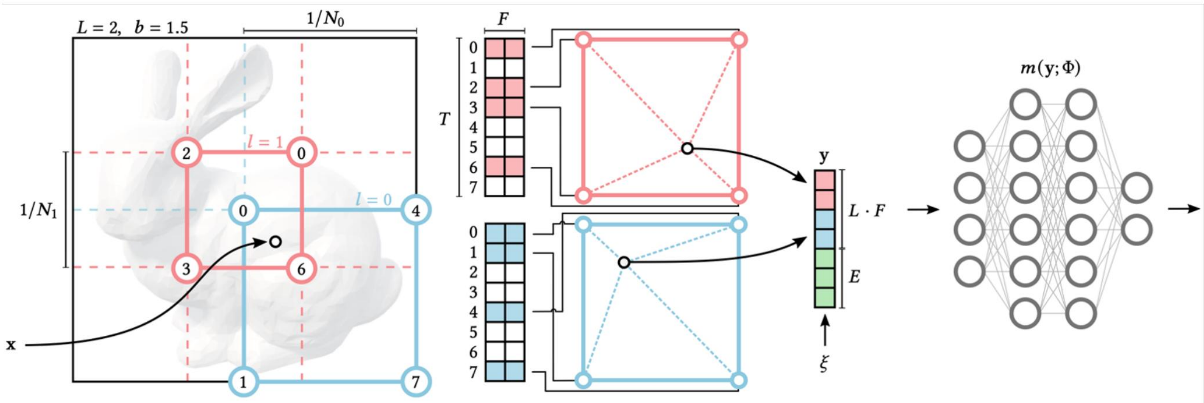

Fig. 2. Model of instant-ngp. In this example, the input image is represented by a mapping function from image coordinates to RGB values. By using a hash grid encoding for the coordinates and then passing the hidden representations of the encodings to a MLP, the model can generate the RGB values for the input image by sampling over the grid.

Fig. 3. Example of images from BSDS-300 dataset. These images all appear in the selected 15 images.

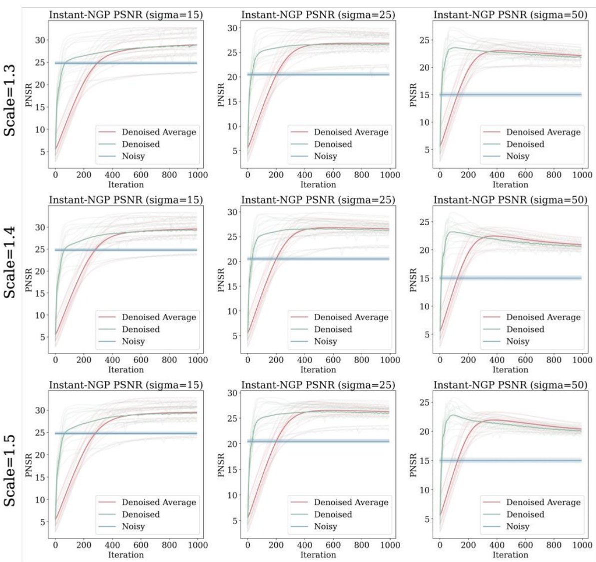

Fig. 4. PSNR plots for Instant-ngp with different per_level_scale values.

4.1 Selection of Parameters

n our experiments, we aimed to ensure a fair comparison between the Deep Image Prior and Instant-ngp by using the recommended settings and parameters from their respective original papers. For the Deep Image Prior, we adhered to the model architecture and hyperparameters specified in the original paper, which have been shown to yield good denoising results. For Instant-ngp, we started with the model settings used in the gigapixel image task, as described in the original paper. However, to optimize the performance of Instant-ngp for the denoising task, we explored various hyperparameters and their impact on denoising quality and speed. Specifically, we focused on the following hyperparameters:

- **n_levels**: The number of levels in the hash map hierarchy.
- **n_features_per_level**: The number of features per level in the hash map.
- **log2_hashmap_size**: The logarithm base 2 of the hash map size.
- **base_resolution**: The base resolution of the hash map.
- **per_level_scale**: The scale factor applied at each level of the hash map hierarchy.

During our exploration, we observed that the **n_levels** parameter did not significantly impact the denoising quality. Therefore, to improve computational efficiency, we reduced **n_levels** from 16 to 12. Similarly, we found that the hash map size, controlled by **log2_hashmap_size**, had a minimal effect on the denoising quality. Consequently, we reduced **log2_hashmap_size** from 15 to 12, further optimizing the memory usage and speed of Instant-ngp. We also investigated the impact of the **per_level_scale** parameter on the denoising performance. We tested three different values: 1.3, 1.4, and 1.5. Figure 4 shows the training plots comparing the PSNR values achieved by Instant-ngp with these different scale values.

From the training plots, we observed that a **per_level_scale** of 1.4 yielded slightly better denoising results compared to 1.3 and 1.5. However, the difference in PSNR values was marginal, indicating that the choice of **per_level_scale** had a relatively minor impact on the overall denoising quality. Based on these findings, we proceeded with the following optimized hyperparameters for Instant-ngp in our experiments:

- **n_levels**: 12
- **n_features_per_level**: 2
- **log2_hashmap_size**: 12
- **base_resolution**: 8
- **per_level_scale**: 1.4

4.2 Running Speed Comparison

Table 1 shows the running speed comparison of the different denoising methods in terms of milliseconds per iteration (ms/it).

From the running speed comparison, we observe that the Deep Image Prior has the highest iteration time of 49.12 ms/it, indicating that it is computationally more expensive compared to the other methods. Instant-ngp shows a significant improvement in running speed, with an iteration time of 21.23 ms/it, which is less than half of the Deep Image Prior. The MLP-based methods, both with and without frequency encoding, exhibit the fastest running speeds, with iteration times of 7.52 ms/it and 7.47 ms/it, respectively. These results suggest that the MLP-based methods are computationally more efficient than the Deep Image Prior, given their smaller network size.

4.3 PSNR and Convergence Iterations

Table 2 presents the PSNR and convergence iterations achieved by each denoising method on the selected 15 images from the BSDS-300 dataset.

From the PSNR results, we observe that the Deep Image Prior achieves the highest PSNR values across all noise levels, indicating its superior denoising performance. At a noise level of 15 (sigma), the Deep Image Prior reaches a PSNR of 30.79 dB after 3000 iterations. As the noise level increases to 25 and 50, the PSNR values for the Deep Image Prior decrease to 29.30 dB and 25.52 dB, respectively, still maintaining its lead over the other methods. Instant-ngp shows the second-best PSNR performance, with values of 29.59 dB, 26.82 dB, and 22.48 dB at noise levels of 15, 25, and 50, respectively. However, it is worth noting that Instant-ngp converges faster than the Deep Image Prior, requiring fewer iterations to reach its final PSNR values. For example, at a noise level of 50, Instant-ngp converges after 380 iterations, while the Deep Image Prior takes 1400 iterations. The MLP with Frequency Encoding demonstrates consistent PSNR performance across different noise levels, with values ranging from 24.93 dB to 23.30 dB. Although its PSNR values are lower than the Deep Image Prior and Instant-ngp, it still outperforms the pure MLP by a significant margin. The pure MLP exhibits the lowest PSNR values among all the methods, with a consistent PSNR around 11.5 dB regardless of the noise level. This suggests that the pure MLP struggles to effectively denoise the images and highlights the importance of incorporating appropriate priors or encodings to improve denoising performance.

These results demonstrate the trade-off between denoising performance and computational efficiency. The Deep Image Prior achieves the highest PSNR values but requires more iterations and computational resources. Instant-ngp provides a good balance between PSNR performance and convergence speed, being 30x faster to train until convergence, making it an attractive option for practical applications.

TABLE 1
Speed comparison of different denoising methods. The time taken for each method to complete a single iteration is recorded in milliseconds.

	Deep Image Prior	Instant-NGP	Frequency & MLP	MLP
ms/ it	49.12	21.23	7.52	7.47

TABLE 2
PSNR and convergence comparison of different denoising methods. The PSNR values are averaged over the selected 15 images from the BSDS-300 dataset.

noise (sigma)	Deep Image Prior	Instant-NGP	Frequency & MLP	MLP
15	30.79 (it 3000)	29.59 (it 1000)	24.93 (it 1000)	11.55 (it 1000)
25	29.30 (it 3000)	26.82 (it 580)	24.62 (it 1000)	11.53 (it 1000)
50	25.52 (it 1400)	22.48 (it 380)	23.30 (it 1000)	11.54 (it 1000)

4.4 Detailed Comparison

To gain further insights into the denoising characteristics of the Deep Image Prior and Instant-ngp, we visually examine and compare their denoised results. Figure 4.4 shows a side-by-side comparison of the denoised images produced by the Deep Image Prior and Instant-ngp at different noise levels.

From the visual comparison, we observe that Instant-ngp tends to preserve more fine textures and details in the denoised images compared to the Deep Image Prior. The Deep Image Prior, on the other hand, produces smoother and more visually pleasing results, especially at higher noise levels.

At a noise level of 15 (sigma), both the Deep Image Prior and Instant-ngp effectively remove the noise while retaining the overall image structure and details. However, upon closer inspection, we notice that Instant-ngp preserves slightly more intricate textures, such as the individual leaves and branches in the image.

As the noise level increases to 25 and 50, the differences between the two methods become more apparent. The Deep Image Prior continues to produce smooth and coherent results, successfully suppressing the noise while maintaining the main image content. In contrast, Instant-ngp struggles to completely eliminate the noise, resulting in some residual noise artifacts in the denoised images. However, Instant-ngp still manages to retain more of the original textures and fine details compared to the Deep Image Prior.

The preservation of textures and details by Instant-ngp can be attributed to its ability to capture high-frequency information through the use of neural graphics primitives, which is evident when we compare it with the baseline models with frequency encoding and pure MLP (Figure 4.4). However, this sensitivity to high-frequency content also makes Instant-ngp more susceptible to higher levels of noise. As the noise level increases, the high-frequency noise components can be mistaken for genuine image details, leading to the persistence of noise artifacts in the denoised results.

On the other hand, the Deep Image Prior's smoothness prior helps to suppress noise more effectively, especially at higher noise levels. By prioritizing the overall image

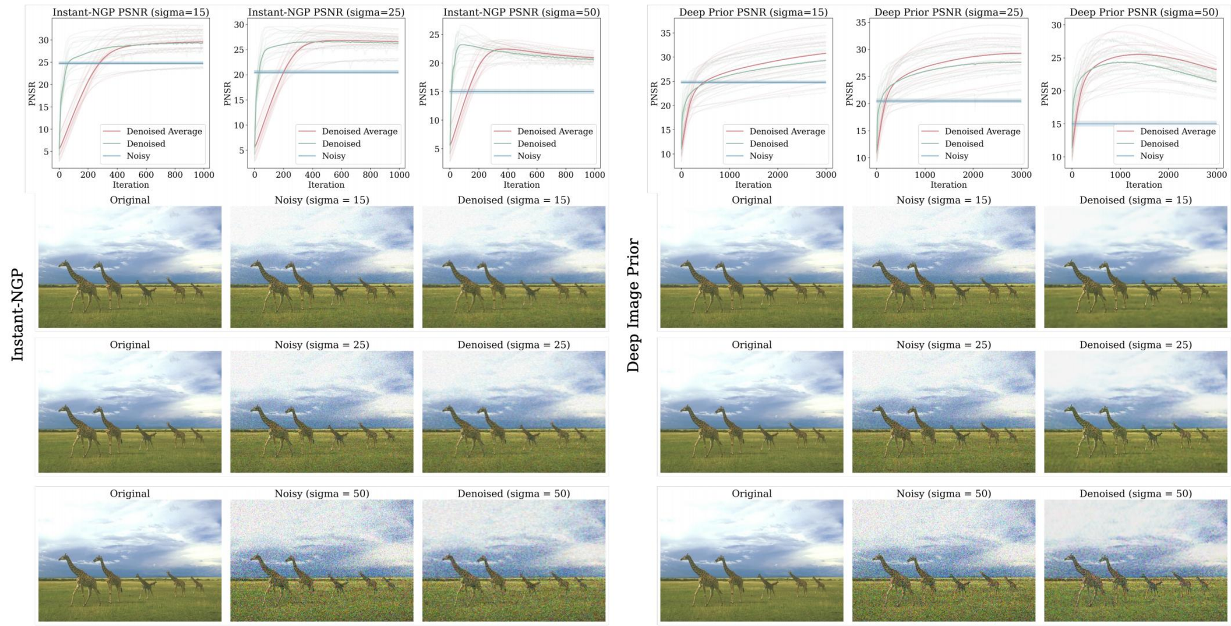

Fig. 5. PNSR plot and denoising results of the instant-ngp and deep image prior on a sample image from the BSDS-300 dataset. The images are shown at three different noise levels (15, 25, 50). The PSNR values are plotted against the number of iterations for each method, with the sloid line being the averaged value.

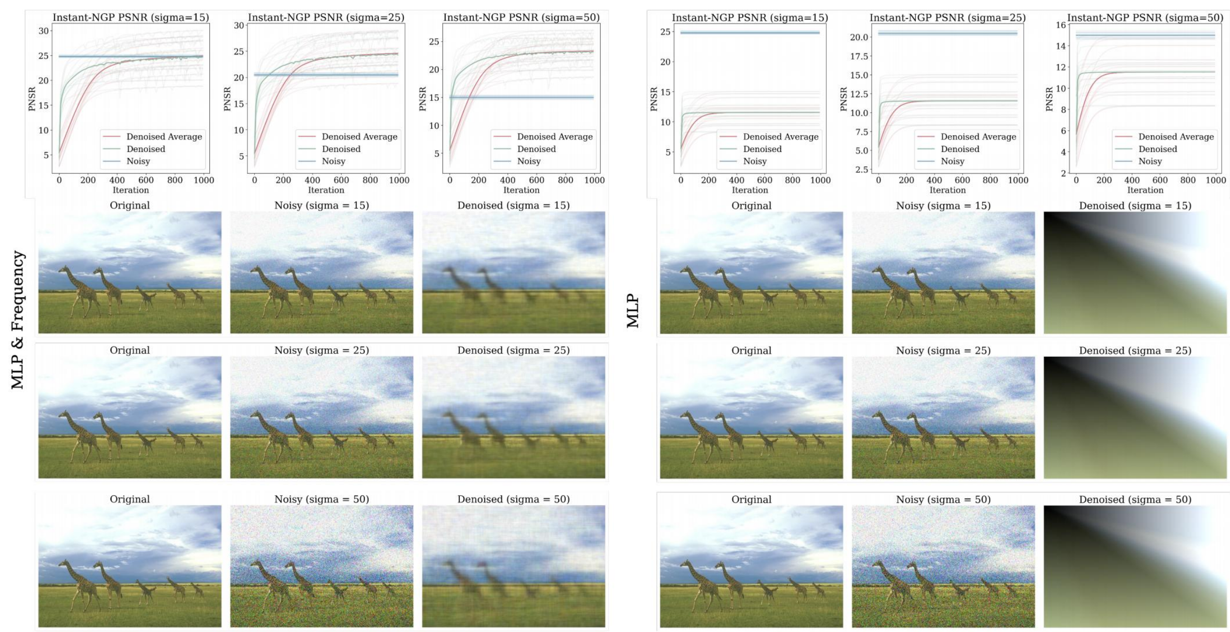

Fig. 6. PNSR plot and denoising results of the MLP with frequency encoding and pure MLP as a baseline.

structure and low-frequency content, the Deep Image Prior is able to generate visually pleasing and noise-free results, albeit at the cost of some loss of fine textures and details.

These observations highlight the trade-off between texture preservation and noise suppression in image denoising. Instant-ngp's focus on preserving textures and details makes it more suitable for applications where maintaining the original image characteristics is crucial, such as in digital art or texture synthesis. However, its sensitivity to higher noise levels may limit its performance in scenarios with severe noise corruption. The Deep Image Prior, with its emphasis on smoothness and noise suppression, is more robust to higher noise levels and produces visually appealing results, making it a preferred choice for general image denoising tasks.

5 CONCLUSION

In this study, we investigated the denoising capabilities of instant-ngp functions and compared them with the Deep Image Prior and other baseline methods. Our experiments on the BSDS-300 dataset provided valuable insights into the performance, computational efficiency, and characteristics of these denoising approaches.

The Deep Image Prior demonstrated the highest PSNR values across all noise levels, showcasing its effectiveness in removing noise while preserving the overall image structure. However, it required more iterations and computational resources to achieve these results. Instant-ngp emerged as a promising alternative, offering a good balance between denoising performance and convergence speed. Although its PSNR values were slightly lower than the Deep Image Prior, Instant-ngp converged faster and showed a notable ability to preserve fine textures and details in the denoised images.

The MLP-based methods, particularly the MLP with Frequency Encoding, exhibited computational efficiency but had lower PSNR values compared to the Deep Image Prior and Instant-ngp. The pure MLP struggled to effectively denoise the images, highlighting the importance of incorporating appropriate priors or encodings to improve denoising performance.

Our detailed analysis revealed that Instant-ngp's focus on preserving textures and details made it more sensitive to higher levels of noise, resulting in the persistence of some noise artifacts in the denoised results. In contrast, the Deep Image Prior's smoothness prior enabled it to suppress noise more effectively, especially at higher noise levels, at the cost of some loss of fine details.

The insights gained from this study contribute to the understanding of image priors and their role in denoising tasks. The Deep Image Prior remains a powerful and effective approach for general image denoising, particularly when visual quality and noise suppression are the primary concerns. Instant-ngp, with its ability to preserve textures and details, offers a complementary approach that can be beneficial in applications where maintaining the original image characteristics is crucial.

Furthermore, the computational efficiency of the MLP-based methods, especially the MLP with Frequency Encoding, highlights the potential for developing fast and lightweight denoising algorithms. While their denoising performance may not match that of the Deep Image Prior or Instant-ngp, they can serve as valuable baselines or be used in scenarios where computational resources are limited.

Future research directions could explore the integration of instant-ngp functions with other denoising techniques to leverage their complementary strengths. Additionally, investigating the application of instant-ngp functions to a wider range of image restoration tasks, such as super-resolution or inpainting, could uncover new possibilities and further extend their utility.

REFERENCES

[1] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, 2005, pp. 60–65 vol. 2.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[3] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, p. 1867–1888, Mar. 2020. [Online]. Available: http://dx.doi.org/10.1007/s11263-020-01303-4

[4] G. Mataev, M. Elad, and P. Milanfar, "Deepred: Deep image prior powered by red," *ArXiv*, vol. abs/1903.10176, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:85498757

[5] Y. Gandelsman, A. Shocher, and M. Irani, "Double-dip: Unsupervised image decomposition via coupled deep-image-priors," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, ser. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, Jan. 2019, pp. 11 018–11 027, project funded by the European Research Council (ERC) under the Horizon 2020 research & innovation programme (grant No. 788535); IEEE/CVF Conference on Computer Vision and Pattern Recognition ; Conference date: 16-06-2019 Through 20-06-2019.

[6] D. Ren, K. Zhang, Q. Wang, Q. Hu, and W. Zuo, "Neural blind deconvolution using deep priors," 06 2020, pp. 3338–3347.

[7] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, pp. 1–1, 06 2021.

[8] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," 10 2019, pp. 3844–3851.

[9] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *CoRR*, vol. abs/2201.05989, 2022. [Online]. Available: https://arxiv.org/abs/2201.05989

[10] A. Yu, S. Fridovich-Keil, M. Tancik, Q. Chen, B. Recht, and A. Kanazawa, "Plenoxels: Radiance fields without neural networks," 12 2021.

[11] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, July 2001, pp. 416–423.