Inverse Problems in Imaging and Implicit Representations

Chinnaswamy Purra
chinnaswamy.purra@student.uni-siegen.de
Matr: 1710342

Anil Babu Yadav Jenige
anil.jenige@student.uni-siegen.de
Matr: 1719064

Universität Siegen
July 31, 2024

Abstract

Image deblurring is essential in recovering sharp images from those affected by factors such as camera motion, defocus, and atmospheric conditions. The study delves into both Explicit and Implicit representations of deblurring, comparing their computational efficiency, robustness, and effectiveness. Explicit representation rely on predefined mathematical models and assumptions about the blur kernel. In contrast, implicit representation, including SIREN, Gaussian Fourier Feature Networks, and WIRE, leverage data-driven approaches and neural networks to capture intricate patterns and details. Our findings indicate that the WIRE method, utilizing wavelet-based transformations, excels in handling complex blur patterns, outperforming other methods in terms of Peak Signal-to-Noise Ratio (PSNR). The study provides a comprehensive analysis of the strengths and limitations of each technique, offering valuable insights for future advancements in image restoration. The complete code for this study is available at `https://github.com/chinnaswamy-P/Inverse-Problem-in-Imaging-RAML`

1 Introduction

This project focuses on Image deblurring, a fundamental problem in the field of image processing which aims to recover sharp images from blurred ones caused by factors like camera motion or defocus and atmospheric conditions. This process is difficult due to the loss of information during blurring, making it a best example of an inverse problem in image processing. Inverse problems involve recovering an original signal from observed data that has undergone a known transformation, often complicated by noise.

Image deblurring has significant implications across various domains, including medical imaging, astronomy, and digital photography. The ability to recover clear images from blurred data can enhance diagnostic accuracy in medical scans, improve the resolution of astronomical observations, and rescue otherwise unusable photographs. Moreover, as imaging technology advances, the demand for sophisticated deblurring techniques continues to grow, driving innovation in both hardware and software solutions. During training, the Peak Signal-to-Noise Ratio (PSNR) value (in dB) can be used as a metric to evaluate and compare the performance of different deblurring methods, with higher PSNR indicating better reconstruction quality.

This study evaluates the robustness, computational efficiency, and effectiveness of Explicit and Implicit deblurring techniques. By evaluating both approaches, we aim to provide insights into their respective strengths and limitations, guiding future advancements in image restoration.

2 Related Work

Recent advancements in solving inverse problems in imaging have leveraged deep learning and implicit neural representations. Ongie et al. (2020) [2] provide a comprehensive overview of deep learning techniques for inverse problems, categorizing methods based on their use of forward models and learning paradigms. Sitzmann et al. (2020) [4] introduced SIREN, which uses sinusoidal activation functions to capture fine details and high-frequency variations in data, making it effective for tasks like image deblurring.

Tancik et al. (2020) [5] proposed Gaussian Fourier Feature Networks, which enhance the network's ability to represent high-frequency details by using Gaussian Fourier feature mappings. Lindell et al. (2021) [1] developed WIRE, a novel approach using wavelet-based transformations to efficiently represent signals across multiple scales, demonstrating superior performance in image deblurring tasks. Additionally, the Sinusoidal Representation Network (SIREN) [3] has been recognized for its ability to represent complex natural signals and their derivatives. This section succinctly captures the essence of each referenced work and their contributions to the field of inverse problems in imaging, fitting well within the constraints of a project report.

3 Methodology

Image deblurring task is an inverse problem (1) in which the objective is to recover the original data from transformed observations, often complicated by noise.

$$\mathbf{f} = \mathbf{A}(\mathbf{u}) + \mathbf{n} \tag{1}$$

where, $\mathbf{f}$ is the observed data or the blurred image, $\mathbf{A}$ is the transformation applied to the original image $\mathbf{u}$, where $\mathbf{A}$ represents the Gaussian blur, $\mathbf{u}$ is the original image that we aim to recover and $\mathbf{n}$ is the noise introduced during the observation or measurement process.

This can be formulated as an optimization problem (2) where we minimize the difference between the blurred image and the estimated image, often incorporating regularization terms to handle noise and ensure smoothness.

$$\min_{\mathbf{u}} \|\mathbf{A}(\mathbf{u}) - \mathbf{f}\|^2 + \alpha \|\nabla \mathbf{u}\| \tag{2}$$

where, $\mathbf{u}$ is the original image that we aim to recover, $\mathbf{A}$ is the Gaussian blur kernel, representing the blurring operator, $\mathbf{f}$ is the observed blurred image, $\|\mathbf{A}(\mathbf{u}) - \mathbf{f}\|^2$ is the data fidelity term, which ensures that the deblurred image, when transformed, matches the observed blurred image, $\|\nabla \mathbf{u}\|$ is the regularization term, where $\nabla \mathbf{u}$ represents the gradient of $\mathbf{u}$, imposing smoothness or other desired properties on the recovered image, α is the regularization parameter that balances the data fidelity term and the regularization term.

3.1 Explicit Representation

Explicit Representation in image deblurring rely on predefined mathematical models and blur kernel assumptions. These methods invert the blurring process by using algorithms based on known blur and noise properties.

3.2 Implicit Representation

Implicit Representation does not rely on predefined models of the blur kernel. Instead, they learn to map blurred images to their sharper counterparts using data. These methods capture intricate patterns and details through data-driven learning, often leveraging coordinate-based neural networks.

3.2.1 Sinusoidal Representation Network (SIREN)

SIREN is a neural network architecture (3) that uses sine as a periodic activation function. It is defined as:

$$\mathbf{\Phi}(x) = \mathbf{W}_n(\phi_{n-1} \circ \phi_{n-2} \circ ... \circ \phi_0) \tag{3}$$

where, $\phi_i(x) = \sin(\mathbf{w}_i x + \mathbf{b}_i)$ is the activation function for layer i. SIREN is particularly effective for representing complex natural signals and their derivatives, making it suitable for solving challenging boundary value problems and allows it to capture fine details and high-frequency variations in data, making effective for tasks requiring precise reconstructions like image deblurring.

3.2.2 Gaussian Fourier Feature Network

This network uses a Gaussian Fourier feature (4) mapping as an input encoding before passing the data through a standard MLP. The input encoding is defined as:

$$\gamma(x) = [a_1 \cos(2\pi \mathbf{b}_1^T x), a_1 \sin(2\pi \mathbf{b}_1^T x), ..., a_m \cos(2\pi \mathbf{b}_m^T x), a_m \sin(2\pi \mathbf{b}_m^T x)] \quad (4)$$

where, $\mathbf{b}_i$ are sampled from a Gaussian distribution. The Fourier feature transformation enables the network to represent high-frequency components of the image effectively.

3.2.3 Wavelet Implicit Neural Representation (WIRE)

WIRE is a new implicit neural representations (INR) based on a continuous complex Gabor wavelet activation function (5). This uses wavelet decomposition as a basis for implicit neural representations (7). It decomposes the signal into different frequency bands using wavelet transforms and learns to reconstruct the signal from these decompositions. This approach allows for efficient representation of signals across multiple scales.

the complex Gabor wavelet activation function is defined as:

$$\psi(x) = e^{-|x|^2/(2\sigma^2)} e^{i\omega^T x}, \quad (5)$$

where, The function $\psi(x)$ captures both spatial and frequency information by combining a Gaussian envelope and a complex exponential. Here, σ controls the width of the Gaussian, and ω is the frequency vector. WIRE network layer is defined as:

$$\mathbf{\Phi} l(x) = \sum k = 1^K \mathbf{w}_k \psi(\mathbf{a}_k (x + \mathbf{b}_k)), \quad (6)$$

This layer transforms the input signal $\mathbf{x}$ using a weighted sum of wavelet functions, where $\mathbf{w}_k$ are weights, and $\mathbf{a}_k$ and $\mathbf{b}_k$ are parameters that scale and shift the input. And full WIRE network is

$$\mathbf{f}(x) = \mathbf{W}_L(\mathbf{\Phi}_{L-1}(\cdots \mathbf{\Phi}_1(x))) \quad (7)$$

The full network is composed of $\mathbf{L}$ layers, with each layer applying a wavelet-based transformation. The final output $\mathbf{f}(x)$ is the result of these successive transformations, allowing the network to learn a compact representation of the input.

4 Experiments

During training, we optimize the loss function that minimizes the difference between the observed blurred image and the re-blurred estimated image, incorporating regularization terms to handle noise and ensure smoothness (2). In the process of training an implicit neural representation, pairs of coordinates from the blurred image are fed into the network during training to map the pixel intensities, which optimizes the network to minimize the reconstruction error against the original sharp image.

4.1 Result & Analysis

4.1.1 Comparative Analysis of Image Deblurring Techniques: Explicit, Gaussian Fourier, and WIRE Methods

WIRE using wavelet-based transformations in its architecture (7), the network can execute image analysis and processing operations at multiple scales. WIRE is able to extract finer details and complex patterns from the data more efficiently than other techniques because of its multi-scale analysis capability (1).

Hence, the WIRE technique yields deblurred images that are more reliable and accurate, and it performs well in situations where the blur patterns are intricate or irregular.

(a) Original clear image and It's Blurred version

(b) Resulting deblurred images

Figure 1: Comparative Analysis of Image Deblurring Techniques

Gaussian Fourier Feature network's capacity to represent high-frequency details is improved by using Gaussian Fourier transforms on the input coordinates. However, its performance can degrade when the blur patterns deviate significantly from Gaussian assumptions. The network struggles to adapt to diverse and intricate blur patterns, resulting in less effective deblurring compared to more adaptive methods like WIRE (1b).

Explicit representation can be computationally intensive and sensitive to the choice of regularization parameters and may have trouble with extremely intricate or mysterious blur patterns (1b). While these methods can be effective for certain blur types, they often struggle with highly complex or unknown blur patterns. The explicit approach requires precise modeling of the blur kernel, which is not always feasible in real-world scenarios.

Additionally, the need for proper tuning of regularization parameters can make the process cumbersome and less adaptive to varying conditions (Table1). Consequently, the Explicit method may not always achieve the same level of deblurring quality as data-driven approaches like WIRE or neural networks with Fourier features.

4.1.2 Effect of Hyperparameters on Deblurring Performance

In our experiments, we examined the effects of hyperparameters on the performance of image deblurring methods. The key hyperparameters investigated include the regularization parameter (alpha) and the learning rate (lr). The following observations in Table 1 highlight how variations in these parameters influence the deblurring results.

High Alpha: A higher value of alpha, in Table 2, places more emphasis on the regularization term in the loss function. This helps in enforcing smoothness and reducing noise in the deblurred image. However, an excessively high alpha can lead to underfitting, where the model fails to capture important details of the image, resulting in overly smooth outputs that lack sharpness and clarity.

Table 1: Effect of alpha & lr on Deblurring Performance (PSNR in dB) at no. of iterations=5000 (Exp - Explicit, GF - Gaussian Fourier Feature Network)

Table 2: Effect of alpha at lr=0.005

alpha	Exp	GF	WIRE
0.0005	16.52	20.08	24.56
0.001	19.21	21.96	24.71
0.002	21.30	23.89	25.07
0.003	22.19	24.57	25.06
0.004	22.51	24.36	25.11
0.005	22.33	24.13	25.04
0.01	22.23	24.23	24.89
0.02	22.01	23.57	24.17
0.05	21.23	22.31	22.64

Table 3: Effect of lr at alpha=0.004

lr	Exp	GF	WIRE
0.001	22.83	24.22	25.01
0.005	22.51	24.36	25.11
0.006	22.50	24.01	24.90
0.007	22.32	24.20	24.70
0.0001	21.19	23.91	22.74
0.0005	22.45	24.01	24.66
0.0006	22.30	23.80	24.01
0.0007	21.31	22.73	23.57
0.00005	20.19	23.50	21.03

Low Alpha: A lower value of alpha, in Table 2, places more emphasis on the data fidelity term, which ensures that the deblurred image closely matches the observed blurred image. This can lead to better preservation of details, but at the risk of overfitting, especially if the data is noisy. Overfitting can result in retaining noise along with the details, which may degrade the overall image quality.

High Learning Rate: A higher learning rate, in Table 3, leads to faster convergence during the training process. This means that the model's parameters are updated more quickly, potentially reaching a solution faster. However, a high learning rate can cause instability and overshooting, where the model might skip over the optimal solution and fail to converge properly.

Low Learning Rate: A lower learning rate, in Table 3, results in slower convergence, meaning the model's parameters are updated more gradually. This can provide more stable updates and reduce the risk of overshooting, increasing the likelihood of finding a precise optimal solution. However, the trade-off is that training takes longer, which may be impractical for very large datasets or complex models.

Our experiments revealed optimal results at alpha = 0.004 and learning rate (lr) = 0.005 with the models trained for 5000 iterations (1b). Using these optimal hyperparameters, we observed the following PSNR values for different models (1):

i. Explicit model: PSNR = 22.51 dB

ii. Gaussian Fourier Feature Network: PSNR = 24.36 dB

iii. WIRE (Wavelet Implicit Neural Representation): PSNR = 25.11 dB

These results indicate that the WIRE model outperformed both the Explicit model and the Gaussian Fourier Feature Network in terms of PSNR, suggesting its superior capability in image deblurring tasks under the given conditions.

5 Conclusion

In this study, we compared three methods for image deblurring: the WIRE method, Gaussian Fourier method, and Explicit method (1). Our findings indicate that the Wire outperformed the other two approaches, demonstrating superior adaptability and effectiveness in restoring sharp images from blurred versions (1). In contrast, the Gaussian Fourier method, while computationally efficient, struggled with non-Gaussian blurs, and the Explicit, despite its rigorous mathematical framework, was sensitive to parameter choices and computationally intensive.

Future work could explore a broader range of blur types, including motion blur and atmospheric distortions, to test generalizability. Enhancing neural network architectures, such as using generative adversarial networks, could further improve implicit methods like the Wire method. Applying these techniques to other inverse imaging problems and real-world scenarios, such as medical imaging and surveillance, could demonstrate their practical utility and impact

References

[1] David B. Lindell, Vincent Sitzmann, and Gordon Wetzstein. Wire: Wavelet implicit neural representations. *arXiv preprint arXiv:2301.05187*, 2021. Retrieved from `https://arxiv.org/abs/2301.05187`.

[2] Greg Ongie, Ajil Jalal, Christopher Metzler, Richard Baraniuk, Alexandros G. Dimakis, and Rebecca Willett. Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory*, 2020. Retrieved from `https://ieeexplore.ieee.org/abstract/document/9084378`.

[3] Devin Schumacher. Sinusoidal representation network. Retrieved from `https://serp.ai/sinusoidal-representation-network/`.

[4] Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *arXiv preprint arXiv:2006.09661*, 2020. Retrieved from `https://arxiv.org/abs/2006.09661`.

[5] Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Neha Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *arXiv preprint arXiv:2006.10739*, 2020. Retrieved from `https://arxiv.org/abs/2006.10739`.