Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) on GS-Blur Dataset

Lin Zhang

1. Background and Motivation

Image deblurring is a fundamental task in computer vision with significant applications in areas such as autonomous driving, surveillance, and remote sensing. Traditional deblurring methods, while effective under controlled conditions, often fail to generalize in real-world scenarios where noise levels and blur patterns vary unpredictably. Many existing deep learning-based deblurring models are trained in low signal-to-noise ratio (SNR) environments, performing well on clean, synthetic datasets but struggling with real-world, high-noise images.

To address this gap, we employ the Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) introduced by Li et al. (2024), a state-of-the-art blind deblurring algorithm, and evaluate its performance on the GS-Blur dataset proposed by Lee et al. (2024). The GS-Blur dataset introduces 3D Gaussian splatting techniques to simulate realistic near-field blur, creating a more dynamic and diverse representation of blur patterns. This study aims to assess whether WEDDM can effectively handle complex, high-noise blur conditions and improve deblurring performance in practical applications.

The full implementation can be accessed through the following GitHub repository: https://github.com/Rising-Stars-by-Sunshine/LinZhang_FinalProject_25Spr_STATS201/tree/main

2. Research Question

The central research question guiding this study is: **How does the Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) perform when trained on the GS-Blur dataset, particularly in restoring near-field images with complex blur patterns?**

Previous studies have demonstrated that WEDDM performs exceptionally well on remote sensing images, particularly in cases of linear motion blur under low-SNR conditions. However, its generalizability to near-field deblurring remains uncertain. This research explores whether the robust noise-handling properties of WEDDM can be transferred to a broader range of blur types, including those encountered in autonomous driving, medical imaging, and consumer photography.

By leveraging the GS-Blur dataset's realistic blur simulation, this study aims to test WEDDM's adaptability in highly dynamic environments, assessing its effectiveness in reducing noise, preserving details, and restoring fine textures compared to other state-of-the-art deblurring models.

3. Application Scenarios

The outcomes of this research have broad applications in fields that require high-quality image restoration under complex and realistic blur conditions. The GS-Blur dataset, derived from computer vision research, provides a more accurate representation of real-world motion blur, making it particularly relevant for industries such as autonomous driving, surveillance, and remote sensing. In autonomous driving, improving deblurring capabilities can enhance vehicle perception in adverse conditions such as fog, rain, or motion-induced blur. In surveillance and security, clearer video frames enable better facial recognition, object tracking, and anomaly detection in low-visibility environments. Additionally, in remote sensing and aerial imaging, reducing blur in satellite and drone-captured images can significantly improve object detection and scene analysis.

By leveraging GS-Blur and evaluating WEDDM's effectiveness in handling near-field complex blur, this research directly contributes to advancements in these industries, ensuring more robust and reliable image processing solutions.

4. Methodology

This study employs a supervised deep learning approach to investigate the performance of deblurring models under different training conditions. The goal is to evaluate how well the Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) performs when trained on different datasets, particularly focusing on their robustness to complex, real-world blur patterns. The comparison is structured around three experimental setups, each designed to assess the impact of dataset realism and model architecture on deblurring effectiveness. GS-Blur dataset that used for training can be found through the link https://drive.google.com/drive/folders/1ZksD7bPl3_ezDLoeHJ2Duwo_LQG1TXB1.

4.1 Experimental Design and Machine Learning Approach

The study follows a comparative experimental design, where different models are trained on varying datasets and their performance is evaluated using PSNR, SSIM, and LPIPS scores. Specifically, three models are trained and tested:

a) WEDDM trained on a simulated dataset (baseline model)

This experiment evaluates WEDDM's performance when trained on synthetic blurry images generated using Gaussian noise and motion blur. The goal is to establish a baseline for WEDDM's effectiveness in low-SNR environments with simplistic blur patterns.

b) MPRNet trained on the GS-Blur dataset

This experiment applies a traditional deblurring model (MPRNet) to the GS-Blur dataset, assessing how well a state-of-the-art CNN-based deblurring model performs on realistic, near-field blur. MPRNet's performance on GS-Blur provides a benchmark for evaluating WEDDM's advantages and limitations.

c) WEDDM trained on the GS-Blur dataset

This setup examines whether WEDDM, originally designed for low-SNR remote sensing images, can generalize to more complex, near-field blur patterns when trained on the GS-Blur dataset. The experiment evaluates whether WEDDM's wavelet-enhanced features improve noise robustness over MPRNet.

By comparing these three models, the study aims to determine whether dataset realism or model architecture has a greater impact on deblurring performance in near-field images.

4.2 Data Preprocessing and Input Features

The datasets used in this study require careful preprocessing to ensure fair comparisons between models. The preprocessing pipeline includes:

a) Resizing all images to 96×96 pixels for computational efficiency.

b) Data Augmentation:

i. For the simulated dataset, synthetic blur is generated using Gaussian noise and motion blur kernels.

ii. The GS-Blur dataset inherently contains complex 3D Gaussian splatting-based blur, eliminating the need for additional augmentation.

The primary input features for all models include RGB pixel values from blurry images, while the ground truth consists of their corresponding sharp images.

4.3 Models and Algorithms

a) Wavelet-Enhanced Deblurring Diffusion Model (WEDDM)

A diffusion-based blind deblurring algorithm that integrates wavelet transformations for multi-scale noise reduction. This is designed to handle low-SNR remote sensing images, now adapted for near-field blur in the GS-Blur dataset. The model is trained using Mean Squared Error (MSE) loss with an Adam optimizer (learning rate = 1e-4).

b) Multi-Scale Progressive Restoration Network (MPRNet)

A CNN-based deblurring model that progressively restores images in a multi-stage architecture. It uses spatial attention mechanisms to enhance important image features. The model is trained on the GS-Blur dataset as a baseline for non-diffusion-based deblurring performance.

4.4 Strategies for Interpretability and Explainability

To understand how different models restore blurred images, we apply the following interpretability techniques:

a) Saliency Maps: Used to visualize which parts of an image are most influential in the model's decision-making. This could be found in the appendix.

b) Feature Comparisons: By comparing WEDDM's wavelet-based reconstructions to MPRNet's CNN-based approach, we assess how each model processes blur and noise differently.

4.5 Visualization and Evaluation Metrics

All models are compared using three primary performance metrics:

a) **Peak Signal-to-Noise Ratio (PSNR)**: Measures image reconstruction quality (higher is better).

- **i.** **Scale:** 0 to $+\infty$ (theoretically)
- **ii.** **Common Range:** 20-50 dB (higher values indicate better quality, typically >30 dB for high-quality images).
- **iii.** **Interpretation:** A higher PSNR value indicates that the restored image is closer to the original clean image in terms of pixel-wise similarity, meaning less distortion and better reconstruction.

b) **Structural Similarity Index (SSIM)**: Assesses perceptual similarity between images (higher is better).

   i. **Scale:** -1 to 1

   ii. **Common Range:** 0 to 1 (negative values are rare in practice).

   iii. **Interpretation:** SSIM considers luminance, contrast, and structural information. A value closer to 1 indicates that the restored image retains more structural details of the original, with SSIM > 0.8 typically signifying high-quality restoration.

c) **Learned Perceptual Image Patch Similarity (LPIPS)**: Evaluates perceptual distance between images (lower is better).

   i. **Scale:** 0 to 1 (though it can sometimes slightly exceed 1 in rare cases).

   ii. **Interpretation:** LPIPS is based on deep neural network embeddings that model human perception. A lower LPIPS score indicates that the restored image is perceptually closer to the original, with values <0.1 generally indicating high perceptual quality.

Additionally, qualitative visualizations are generated to directly compare the deblurred outputs from each model. These visualizations will highlight:

a) Differences in detail preservation between simulated and GS-Blur-trained models.

b) WEDDM's ability to handle high-noise images compared to MPRNet.

c) The impact of dataset realism on deblurring performance.

This methodology ensures a comprehensive evaluation of both dataset effects and model design choices, providing insights into the optimal training strategy for robust deblurring in real-world conditions.

5. **Results and Discussion**

Our comparative study evaluated three experimental settings:

a) Simulated Dataset → WEDDM (10 Epochs)

b) GS-Blur Dataset → MPRNet (10 Epochs)

c) GS-Blur Dataset → WEDDM (10 Epochs)

d) GS-Blur Dataset → WEDDM (50 Epochs, this is appended as WEDDM trained on GS-Blur Dataset for 10 epochs has not reached the desired result)

The models were assessed using three primary evaluation metrics: PSNR, SSIM, and LPIPS. The results are summarized in Table 1 below.

Table 1. Comparison of PSNR, SSIM, and LPIPS Scores Across Models and Datasets

Model	Dataset	Training Epochs	PSNR (↑)	SSIM (↑)	LPIPS (↓)
WEDDM	Simulated	10	**31.0597**	**0.9563**	**0.0105**
MPRNet	GS-Blur	10	8.8505	0.2413	0.6165
WEDDM	GS-Blur	10	20.3270	0.7277	0.2872
WEDDM	GS-Blur	50	21.1060	0.7588	0.2575

Key Observations:

a) **WEDDM trained on a traditional Simulated Dataset** outperforms all GS-Blur models in terms of PSNR, SSIM, and LPIPS, showing that the model learns more effectively when trained on traditional synthetic datasets.

b) **WEDDM trained on GS-Blur Dataset (50 Epochs)** shows only marginal improvement over 10 Epochs, suggesting that extended training alone does not significantly enhance performance. Interestingly, LPIPS decreased (higher perceptual distance) despite PSNR and SSIM improving slightly.

c) **MPRNet trained on GS-Blur** performs the worst across all three metrics, indicating that WEDDM has a clear advantage over MPRNet in handling high-noise, complex blur scenarios. Moreover, the training time for WEDDM is dramatically decreased compared to training MPRNet (10 epochs training for WEDDM needs only about 5 minutes, but 10 epochs training for MPRNet acquires more than 4 hours).

Figures 1 to 3 give visualization for comparisons.

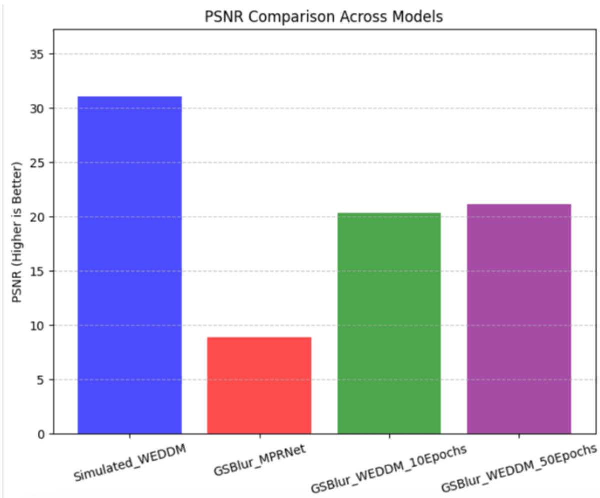

Figure 1. Models' PSNR Scores (Higher is Better)

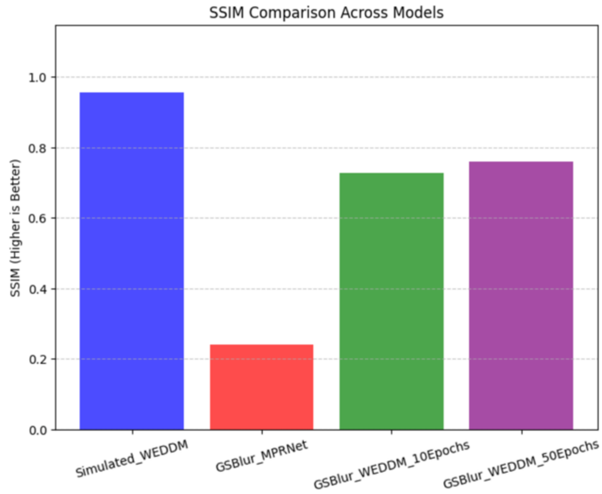

Figure 2. Models' SSIM Scores (Higher is Better)

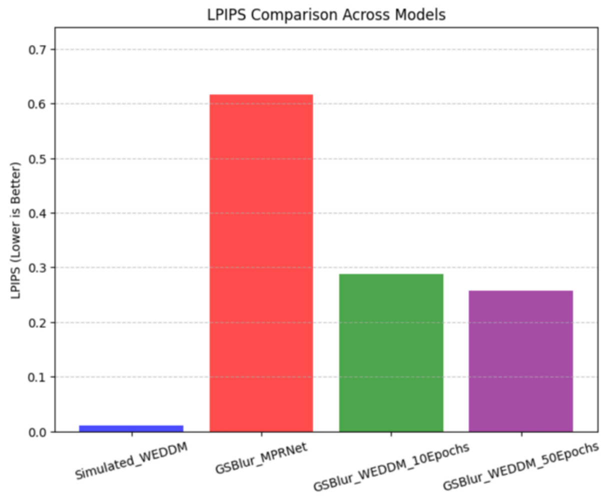

Figure 3. Models' LPIPS Scores (Lower is Better)

However, why is WEDDM trained on the GS-Blur dataset performing worse than expected? Potential reasons are listed below:

a) Training Data Characteristics: The GS-Blur dataset simulates more realistic, near-field blur conditions, which are inherently more complex than the simpler motion blur in traditional synthetic datasets. WEDDM's architecture may need even much longer training cycles or fine-tuning hyperparameters to adapt effectively.

b) Noise-Robustness Issue: The simulated dataset provides ideal training conditions (low noise, clear blur patterns), whereas GS-Blur introduces significant real-world variability. This shift could make it harder for WEDDM to generalize well to more realistic blur conditions.

c) Potential Overfitting or Underfitting: Since GS-Blur contains more complex blur types, WEDDM might require additional data augmentation, alternative loss functions, or adaptive learning rate strategies to better capture these features.

6. Intellectual Merit and Practical Impacts

Advancement of Existing Literature

This research contributes to the growing body of knowledge in image restoration by introducing the application of the Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) to real-world blur scenarios, specifically those represented by the GS-Blur dataset. While much of the existing literature on image deblurring has primarily focused on simpler, synthetic blur patterns (e.g., motion blur or Gaussian blur), this study extends those methods by incorporating complex, near-field blur simulations that mimic real-world environmental conditions. By utilizing the GS-Blur dataset, which introduces more diverse and dynamic blur types, this work enriches the literature on image deblurring, particularly in high-noise, complex conditions where previous models, such as MPRNet, have struggled. Furthermore, this research sheds light on the transferability of techniques designed for remote sensing images (WEDDM) to more challenging, practical scenarios like those found in autonomous driving and medical imaging. By comparing different model architectures and datasets, this work offers valuable insights into the trade-offs between dataset realism and model design choices, providing a roadmap for future deblurring research.

Inspiring Future Research Directions

This research uncovers several avenues for future exploration in the field of image deblurring. First, it highlights the need for additional techniques to improve the generalizability of models like WEDDM to a broader range of blur types, particularly those encountered in dynamic, real-world settings. The current results suggest that further advancements in model architecture or longer training times could yield better performance on complex datasets like GS-Blur. One potential direction is the incorporation of hybrid models that combine the strengths of diffusion-based approaches, such as WEDDM, with CNN-based methods like MPRNet, creating a more versatile system capable of handling both noise reduction and blur restoration more effectively. Additionally, future research could focus on adaptive training strategies, such as progressive learning or curriculum learning, to help models better adapt to evolving complexities in real-world blur conditions. Given the growing importance of AI in real-time applications, there is also an opportunity to explore model optimization techniques that reduce training time while maintaining high-quality results. Finally, since the research focused solely on image deblurring, expanding the scope to address other restoration tasks, such as super-resolution or inpainting, could provide further insights into the broader potential of wavelet-based techniques in image processing.

7. Practical Impact

Societal Benefits

The research can improve image restoration in critical sectors such as autonomous driving, surveillance, and medical imaging. By enhancing the clarity of images in challenging conditions, it could improve safety on the road, support better security systems, and enable more accurate medical diagnoses, ultimately benefiting society.

Key applications include:

a) **Autonomous Driving**: Enhanced image clarity helps vehicles navigate in poor conditions like rain or fog.

b) **Surveillance**: Better video quality improves object detection and facial recognition in low-visibility environments.

c) **Medical Imaging**: Sharper images lead to more accurate diagnoses and treatment plans.

d) **Remote Sensing**: Enhanced clarity improves environmental monitoring and disaster management.

AI Governance and Ethical Considerations

This project aligns with AI governance principles by promoting transparency and fairness. It encourages inclusivity in AI development, ensuring that technologies are accessible and effective across diverse contexts. Additionally, it supports SDGs by advancing AI for safer, more sustainable transportation and better environmental monitoring. Responsible innovation, with ethical oversight, ensures that AI development benefits society without compromising privacy or fairness.

References

Lee, Dongwoo, Joonkyu Park, and Kyoung Mu Lee. 2024. "GS-Blur: A 3D Scene-Based Dataset for Realistic Image Deblurring." ArXiv.org. 2024. https://arxiv.org/abs/2410.23658.

Li, Zhiyuan, Jie Li, Yueting Zhang, Jiayi Guo, and Yirong Wu. 2024. "A Noise-Robust Blind Deblurring Algorithm with Wavelet-Enhanced Diffusion Model for Optical Remote Sensing Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17: 16236–54. https://doi.org/10.1109/jstars.2024.3422175.

Appendix: Model Training and Saliency Map Analysis

WEDDM Training Process

The Wavelet-Enhanced Deblurring Diffusion Model (WEDDM) was trained on the GS-Blur dataset to evaluate its performance in restoring near-field images with complex blur patterns. The training setup followed a supervised learning approach, where the model learned from pairs of blurred and clean images using Mean Squared Error (MSE) loss.

Key components of the training process include:

a) Data Preprocessing: Images were resized to 96×96 pixels and normalized.

b) Model Architecture: WEDDM consists of a denoising module followed by a diffusion module, using convolutional layers and ReLU activation functions.

c) Optimization Strategy: The model was trained using Adam optimizer with a learning rate of 1e-4 for 10 epochs.

After training, the model was evaluated using PSNR, SSIM, and LPIPS metrics, which provided a quantitative assessment of its deblurring performance.

Saliency Map Analysis

To interpret WEDDM's behavior, we generated saliency maps that highlight the most important regions influencing the model's predictions. The saliency maps were computed using gradient-based backpropagation, where the most relevant pixels for deblurring were identified. See figure 4 as reference.

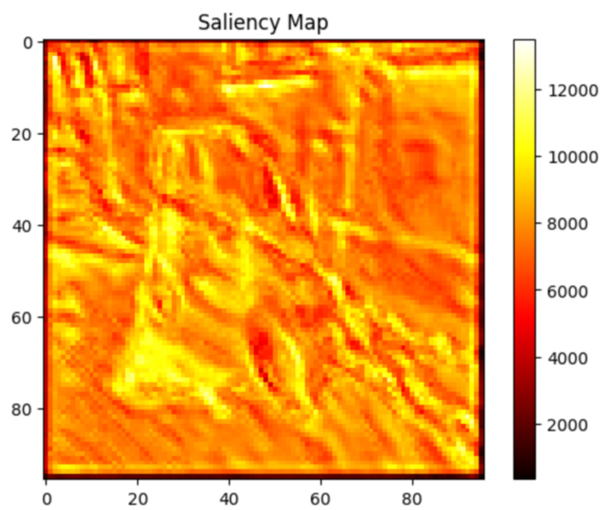

Figure 4. Saliency Map for WEDDM Trained on GS-Blur Dataset

a) Objective: Understand which parts of the image the model focuses on when performing deblurring.

b) Method: Compute gradients with respect to the input image and visualize the most impactful regions.

c) Visualization: The resulting heatmaps reveal whether WEDDM correctly attends to blurred edges and important textures rather than background noise.

This interpretability analysis helps us assess whether WEDDM effectively distinguishes between noise and meaningful structures and how it generalizes across different blur conditions.