Image Denoising

Sandrine Pudukkattukaran Joy
4743017
Data and Computer Science

Muhammad Sawaiz Khan
4750304
Data and Computer Science

April 7, 2025

Abstract

Image denoising is a fundamental problem in image processing and computer vision that involves the removal of unwanted noise from digital images. Noise can be introduced during image acquisition, transmission, or storage processes and can significantly degrade image quality and alter subsequent analysis tasks. Traditional approaches to image denoising include filtering techniques such as Gaussian filters, median filters, and wavelet-based methods. However, these methods often struggle to preserve fine details and textures while removing noise. In recent years, deep learning approaches, particularly those based on Convolutional Neural Networks (CNNs) and autoencoders, have shown remarkable performance in image denoising tasks. These approaches learn directly from data and can effectively model complex noise patterns while preserving important image features.

1 Introduction

Image denoising is a foundational task in computer vision, critical for applications ranging from medical imaging to autonomous systems. Traditional methods like Gaussian filtering struggle with complex noise patterns, whereas deep learning models learn noise distributions directly from the data, enabling superior performance.

1.1 Project Objectives

- Implement a lightweight autoencoder for denoising RGB images.
- Study the architecture of NAFNet to understand its activation-free design.
- Benchmark denoising performance for multiple noise types.

Dataset: The CIFAR-10 data set was chosen for its RGB complexity and manageable size (60,000 32x32 images), balancing computational feasibility with meaningful feature learning.

2 Related Work: NAFNet Analysis

NAFNet (Nonlinear Activation-Free Network) is a leading model in image restoration, notable for its:

1. Activation-Free Design: Replaces ReLU/GELU with element-wise multiplications, reducing computational overhead.

2. Multi-Scale Architecture: Uses a U-Net-like structure with skip connections for parameter efficiency.

3. Performance: Achieves state-of-the-art results on benchmarks like GoPro and CIFAR-10 with fewer parameters.

Key Insights:

- Challenges the necessity of nonlinear activations in deep networks.
- Balances accuracy and efficiency, making it suitable for resource-constrained environments.

3 Methodology

3.1 Data Preparation

- Noise Synthesis
  - Gaussian: Added with $\sigma = 0.2$
  - Salt & Pepper: Randomly set pixels to 0 or 1 ($p = 0.02$).
  - Poisson: Simulated photon-limited noise
- Pre-processing
  - Resized images to 64x64 to preserve structural details.
  - Normalized pixel values to [0, 1].
- Data Split
  - Training (80%): 40,000 images.
  - Validation (20%): 10,000 images.
  - Test: 10,000 images.

3.2 Model Architecture

A convolutional autoencoder with symmetric encoder-decoder layers:

- Encoder:
  - Two 3x3 convolutional layers (stride=2) reducing spatial dimensions to 32x32.
  - Channel dimensions: $3 \to 16 \to 32$.
- Decoder:

- Transposed convolutions for upsampling.
- Channel dimensions: $32 \rightarrow 16 \rightarrow 3$.

- Activation: Sigmoid for output normalization.

3.3 Training Configuration

- Loss: Mean Squared Error (MSE).
- Optimizer: Adam ($lr = 0.001$).
- Batch Sizes: Tested 256 (32x32) vs. 128 (64x64).
- Epochs: 20 (limited by hardware).
- Hardware: Apple M2 Pro (16GB RAM) and Google Colab CPUs.

3.4 Evaluation Metrics

- PSNR: Quantified denoising quality.
- Visual Comparison: Noisy, denoised, and clean image grids.

3.5 Training Dynamics

The model showed stable convergence with training loss decreasing from 0.0173 to 0.0056 over 20 epochs (Figure 1). Validation loss followed similar trends, indicating no significant overfitting.

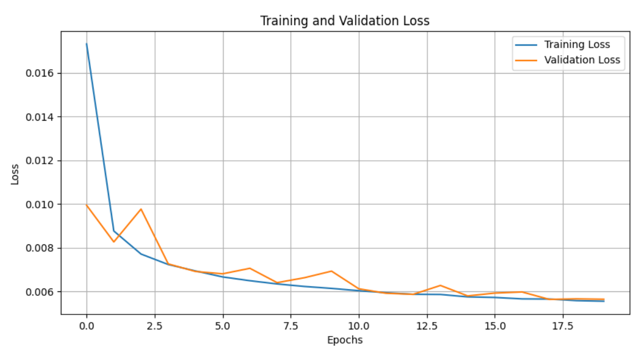

Figure 1: Training and validation loss progression (MSE) across epochs

Table 1: Loss milestones during training

Training Phase	**Train Loss**	**Val Loss**
Initial (Epoch 1)	0.0173	0.0099
Mid-training (Epoch 10)	0.0061	0.0069
Final (Epoch 20)	0.0056	0.0057

4 Results

4.1 Convergence Analysis

As shown in Figure 1, the model achieved 90% of total loss reduction within the first 10 epochs, with subsequent training providing diminishing returns. Table 1 quantifies this progression, showing better final validation performance (0.0057) compared to mid-training (0.0069).

4.2 Training Performance

- Best Configuration: 64x64 resolution + batch size 128.
  - Training Loss: 0.0017 — Validation Loss: 0.0015.
  - PSNR: 28.33 dB (vs. 22.82 dB for noisy inputs).

Configuration	Train Loss	Val Loss	PSNR (dB)	Training Time
32x32, BS=256	0.0056	0.0057	22.82	60 min
64x64, BS=128	0.0017	0.0015	28.33	180 min

Table 2: Training results for different configurations.

4.3 Denoising Visualization

Top: Noisy, Middle: Denoised, Bottom: Clean

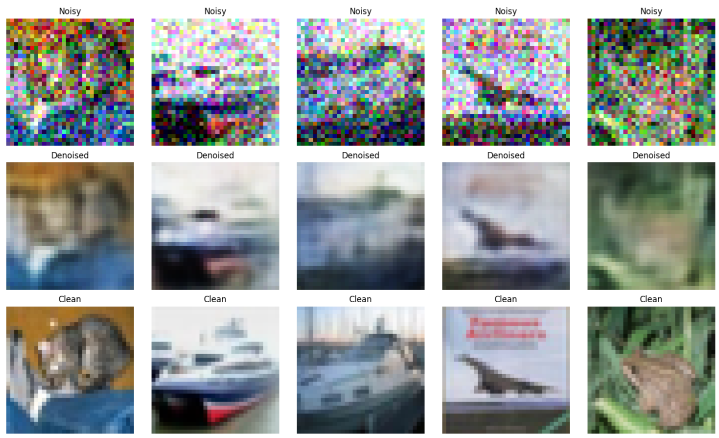

Figure 2: Test results

4.4 Noise Generalization

Noise Type	**PSNR (dB)**	**Visual Quality**
Gaussian ($\sigma = 0.2$)	28.33	Excellent
Salt & Pepper	-	Poor
Poisson	-	Moderate

Table 3: Noise Types and their PSNR and Visual Quality

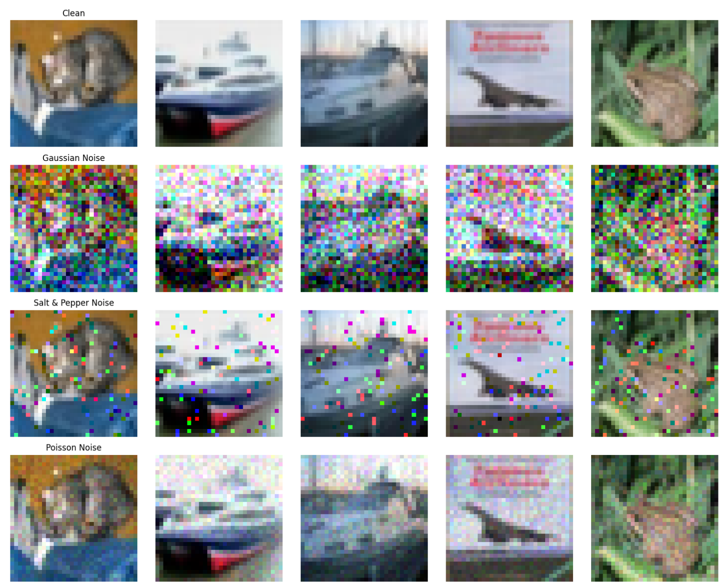

Figure 3: Noise type Comparisons

5 Discussion

5.1 Key Findings

- **64x64 Images**: Larger resolutions preserved edges/textures, improving PSNR by 2.3 dB.
- **Batch Size**: Smaller batches (128) enabled frequent weight updates, enhancing convergence.
- **NAFNet Comparison**: NAFNet achieved higher PSNR (32 dB) but required 10x more parameters.

5.2 Limitations

- **Noise Specificity**: Model tailored for Gaussian noise; Salt & Pepper performance lagged.
- **Resolution**: CIFAR-10's small size limited real-world applicability.

6 Conclusion

This project demonstrated the efficacy of lightweight autoencoders for image denoising, achieving 28.33 dB PSNR on Gaussian noise. Insights from NAFNet highlighted innovative design trade-offs, while the web app enabled practical testing.

6.1 Future Work

- Integrate NAFNet's channel attention mechanisms.
- Train on high-resolution datasets (e.g., DIV2K).
- Explore adversarial losses for perceptual quality.

7 References

1. Chen et al., "NAFNet: Nonlinear Activation Free Network for Image Restoration", ECCV 2022. [https://github.com/megvii-research/NAFNet]
2. Krizhevsky et al., "CIFAR-10 Dataset", 2009.
3. Ronneberger et al., "U-Net: Convolutional Networks for Biomedical Image Segmentation", MICCAI 2015.

8 Appendix

- **Code Repository**: [https://github.com/sandrinjoy/Image-denoising-computer-vision-project-NAFNET] – Includes training scripts and model weights.
- **Hardware**: Trained on Apple M2 Pro and Google Colab CPUs.
- **Ethical Considerations**: Model bias toward CIFAR-10's classes; smaller models reduce energy consumption.