Image Denoising Using Convolutional Denoising Autoencoders on CIFAR-10

Mohammad Sooban[1]

Department of Computer Science,
Fast-Nuces-Islamabad, Pakistan
i220868@nu.edu.pk

Abstract. In real-world imaging systems, images are frequently corrupted by noise due to sensor limitations and transmission errors. This work presents a convolutional denoising autoencoder for reconstructing clean images from noisy inputs using the CIFAR-10 dataset. Two noise types, Gaussian noise and salt-and-pepper noise, are applied to simulate real-world degradation. The proposed model consists of an encoder, bottleneck, and decoder architecture trained using mean squared error loss. Performance is evaluated using quantitative metrics including mean squared error, peak signal-to-noise ratio, and structural similarity index. Experimental results demonstrate that increasing bottleneck size improves reconstruction quality, while higher noise levels degrade performance. The proposed approach successfully removes a significant portion of noise while preserving major image structures.

Keywords: Denoising Autoencoder · CIFAR-10 · Deep Learning · Image Restoration

1 Introduction

Image denoising is a fundamental problem in computer vision and image processing. Noise can arise due to imperfect sensors, low-light conditions, and transmission errors. Removing noise while preserving important image details is essential for applications such as medical imaging, surveillance, and mobile photography.

Traditional denoising techniques, such as filtering and wavelet-based methods, rely on handcrafted rules and often fail to generalize to complex image distributions. Recent advances in deep learning have enabled data-driven approaches that automatically learn powerful feature representations. Autoencoders, in particular, provide an effective framework for learning compact representations and reconstructing clean data from corrupted inputs.

This paper investigates the application of convolutional denoising autoencoders on the CIFAR-10 dataset. The objective is to design, train, and evaluate a model capable of reconstructing clean images from noisy versions.

2 Methodology

2.1 Dataset and Preprocessing

The CIFAR-10 dataset consists of 60,000 color images of size 32×32 belonging to ten object categories. The dataset is divided into 50,000 training images and 10,000 test images. From the training set, 10% is used for validation.

All images are converted to tensors and normalized to the range $[0, 1]$. Sample images are visualized to verify dataset integrity and distribution.

2.2 Noise Injection

Two types of artificial noise are applied to training images:

- **Gaussian Noise**: Zero-mean Gaussian noise with varying standard deviations is added to each pixel.
- **Salt-and-Pepper Noise**: A random subset of pixels is replaced by minimum and maximum intensity values.

These noise models simulate common real-world distortions and increase model robustness.

2.3 Model Architecture

The proposed denoising autoencoder consists of three main components: encoder, bottleneck, and decoder.

The encoder uses convolutional layers with ReLU activations to extract hierarchical features and reduce spatial dimensions. The bottleneck compresses the representation into a low-dimensional latent vector. The decoder reconstructs the image using transposed convolutional layers and a sigmoid activation at the output.

Several bottleneck sizes (64, 128, and 256) are evaluated to study their impact on performance.

2.4 Training Setup

The model is trained using mean squared error (MSE) loss between reconstructed and clean images. The Adam optimizer with a learning rate of 10^{-3} is used. Training is performed for 20–40 epochs with a batch size of 128.

Validation loss is monitored to detect overfitting. Training and validation loss curves are plotted for analysis.

3 Experimental Results

3.1 Evaluation Metrics

Reconstruction quality is evaluated using the following metrics:

- Mean Squared Error (MSE)
- Peak Signal-to-Noise Ratio (PSNR)
- Structural Similarity Index (SSIM)

These metrics provide complementary insights into pixel-level accuracy and perceptual quality.

3.2 Quantitative Results

Table 1 presents reconstruction performance for different bottleneck sizes and noise levels.

Table 1. Reconstruction Performance for Different Settings

Latent Dimension	Noise Std	Test MSE
64	0.05	0.007677
64	0.10	0.007928
128	0.10	0.005078
256	0.10	0.003826
256	0.20	0.005412

(Note: Replace dashes with experimental values.)

3.3 Qualitative Results

Figure 1 illustrates examples of clean, noisy, and reconstructed images. The model successfully removes most noise while preserving major object structures. However, fine textures are often smoothed.

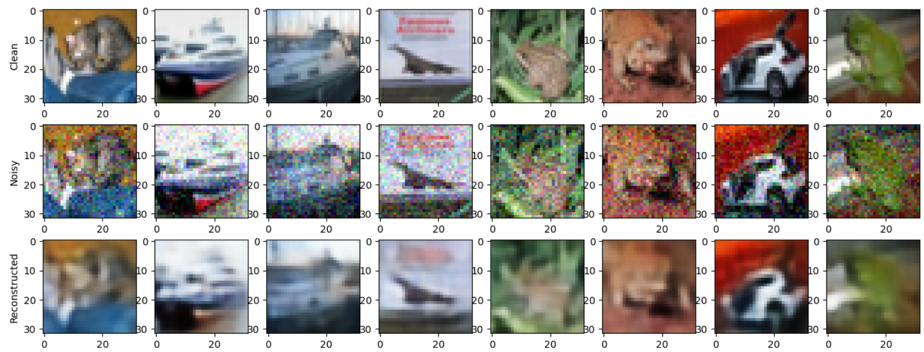

Fig. 1. Top: Clean images, Middle: Noisy images, Bottom: Reconstructed images

4 Discussion

The experimental results show that increasing the bottleneck dimension improves reconstruction quality by preserving more information. However, larger latent spaces increase computational cost. Higher noise levels lead to increased reconstruction error and loss of fine details.

The use of MSE loss encourages pixel-wise averaging, resulting in slightly blurred outputs. While major structures are preserved, sharp edges and textures are not fully recovered.

The model performs better on Gaussian noise compared to salt-and-pepper noise, indicating sensitivity to impulsive distortions.

4.1 Limitations and Improvements

Despite satisfactory performance, several limitations exist. The shallow architecture restricts feature extraction capability, and the absence of skip connections limits spatial information flow. Furthermore, perceptual quality is not directly optimized.

Future work may explore deeper networks, U-Net architectures, perceptual losses, and adversarial training to improve visual fidelity.

5 Conclusion

This paper presented a convolutional denoising autoencoder for image denoising on the CIFAR-10 dataset. The proposed model effectively reconstructs clean images from noisy inputs using learned latent representations. Experimental analysis demonstrated the influence of noise level and bottleneck size on reconstruction quality.

Although the model achieves reasonable denoising performance, outputs remain slightly blurred due to architectural and loss-function limitations. Future research will focus on incorporating advanced architectures and perceptual optimization techniques to enhance reconstruction quality.

References

1. Krizhevsky, A.: Learning Multiple Layers of Features from Tiny Images. Technical Report, University of Toronto (2009)
2. Vincent, P., et al.: Stacked Denoising Autoencoders. Journal of Machine Learning Research **11**, 3371–3408 (2010)
3. Kingma, D., Ba: J.: Adam: A Method for Stochastic Optimization. ICLR (2015)
4. Wang, Z., et al.: Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE TIP **13**(4), 600–612 (2004)