Image Restoration from Noisy Images

Gokul M K , Keerthi Vasan M , Vaidehi Bagaria
ED21B026 ED21B034 ED21B070

September 23, 2025

1 Introduction

In this work, we aim to address the problem of image denoising by leveraging the state-of-the-art RIDNet architecture. RIDNet is a CNN-based approach for real image denoising, known for effectively reducing noise while preserving structural details in images. It is the first model to incorporate feature attention in a denoising context. We used RIDNet to the denoising dataset provided and evaluated its performance across multiple object classes. We also tried to use GAN, Vision Transformer, Diffusion Models and found that it is not easy to train as they require powerful GPUs. This made us stick to a CNN-based denoising model, and through various experiments, we found out that RIDNet performed the best.

2 Architecture Details

2.1 RIDNet Overview

RIDNet (Real Image Denoising Network) is an architecture designed for efficient denoising tasks, incorporating residual learning and feature attention. Our implementation follows the original RIDNet structure as proposed in [1]. The key components include residual blocks, EAM blocks, global pooling, element-wise addition and multiplication, which enable the network to focus on areas with prominent noise.

2.2 Network Details

The model is composed of 3 main modules, **feature extraction**, **feature learning residual** and **reconstruction**.

- **Feature Extraction:** This module is composed of only one convolutional layer to extract initial features
- **Feature Learning Residual:** This module is composed of ”Enhancement Attention Module” (EAM) that are cascaded together. EAM provides wide receptive field through kernel dilation. 4 EAMS blocks are used in the model.
- **Reconstruction:** This module is composed of a single convolutional layer that reconstructs the features from the previous layer.

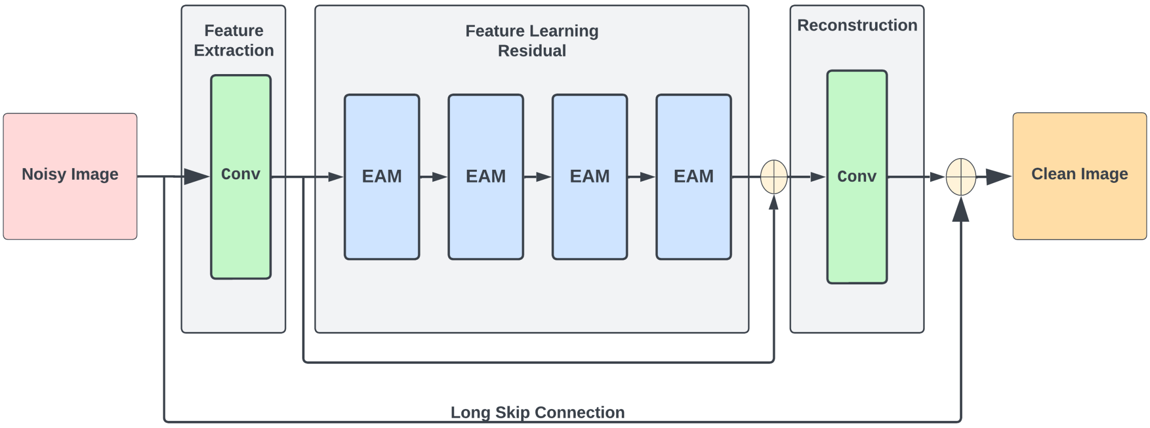

Figure 1: RIDnet Architecture

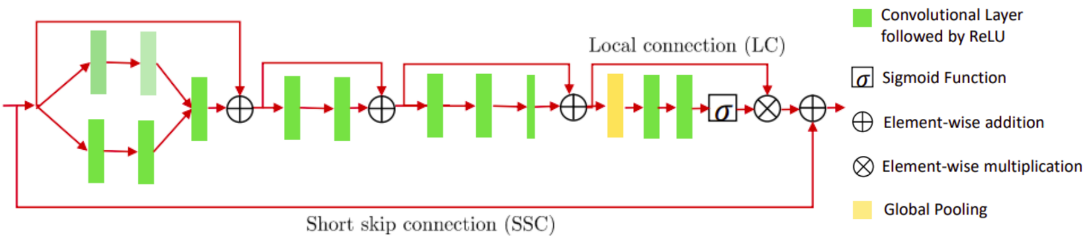

Figure 2: EAM Block Architecture

2.3 Implementation

Our model architecture follows closely the original RIDNet architecture to ensure fidelity to the original design and benchmark its performance on the provided dataset. However, we conducted extensive experiments, testing various optimizers, loss functions, learning rates and activation functions. Based on these trials, we used the following functions for our final model.

- **Optimizer:** ADAM with $\alpha = 1e^{-5}$ and Cosine Annealing Schedule with T=50
- **Loss Function:** Mean-Squared-Error loss (L1 loss by the Authors)
- **Activation:** Sigmoid Linear Unit (SiLU)

For better denoising, while preserving the defects, while training, we divide each image into 16 sub-images and then train the model on these 16 sub-images with corresponding clean sub-images. In our experiments we found that dividing the images into 16 parts proved to have better denosing capability while preserving the defects. For the validation set we pass the entire image.

2.4 Defect Mask Segmentation

For obtaining the defect mask from noisy image, we trained a U-Net model with ground truth images and their respective mask defect label. U-Net is a semantic segmentation technique. The model architecture is fairly simple: an encoder and a decoder with skip connections. We utilised the segmentation_models library to implement U-Net with weights pretrained on imagenet.

3 Experimental Results

3.1 Sample Outputs

Figure 34 shows sample output images generated by RIDNet, illustrating the model's denoising effectiveness on different objects within the dataset.

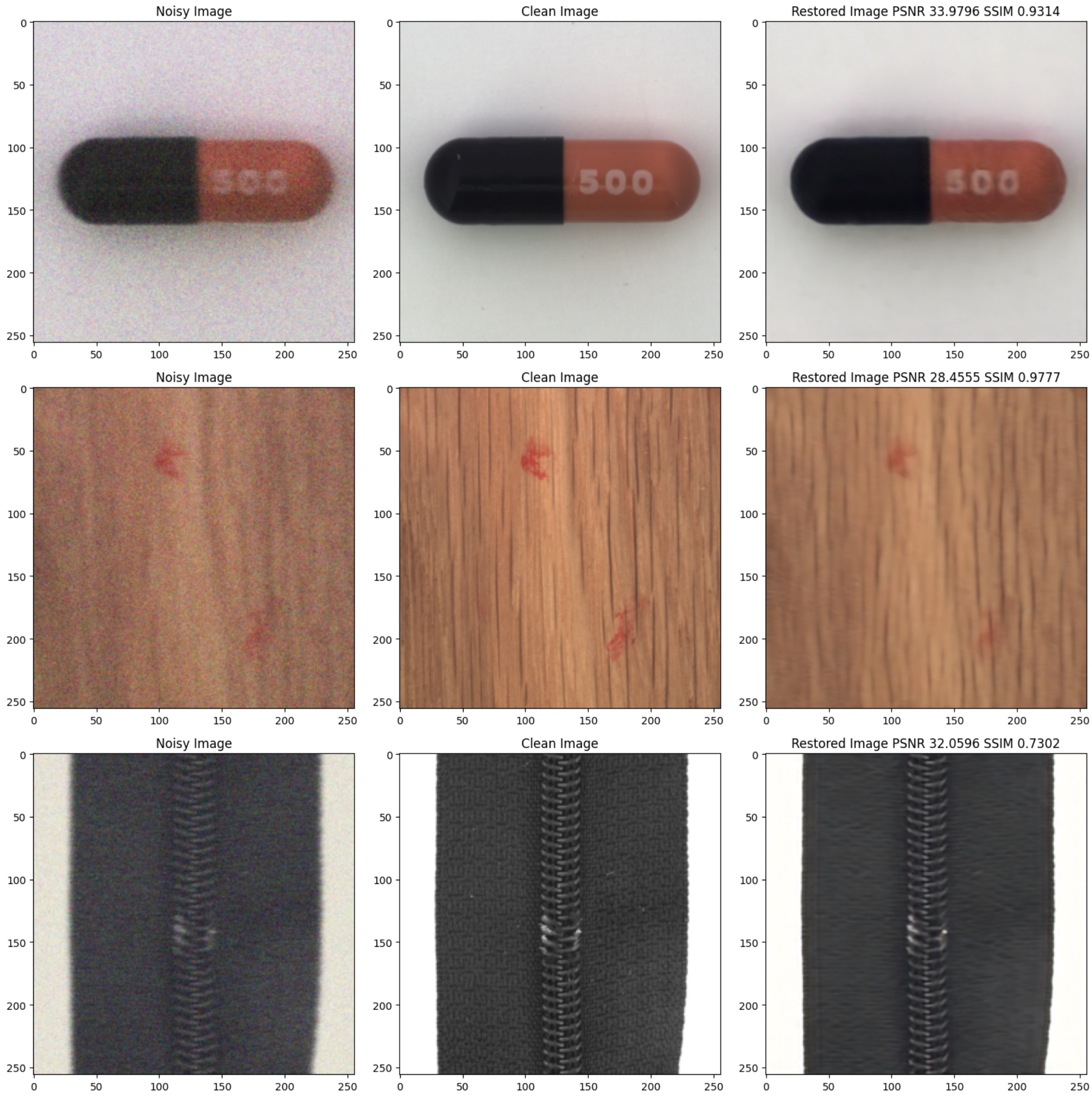

Figure 3: Sample outputs for denoised images from the validation dataset

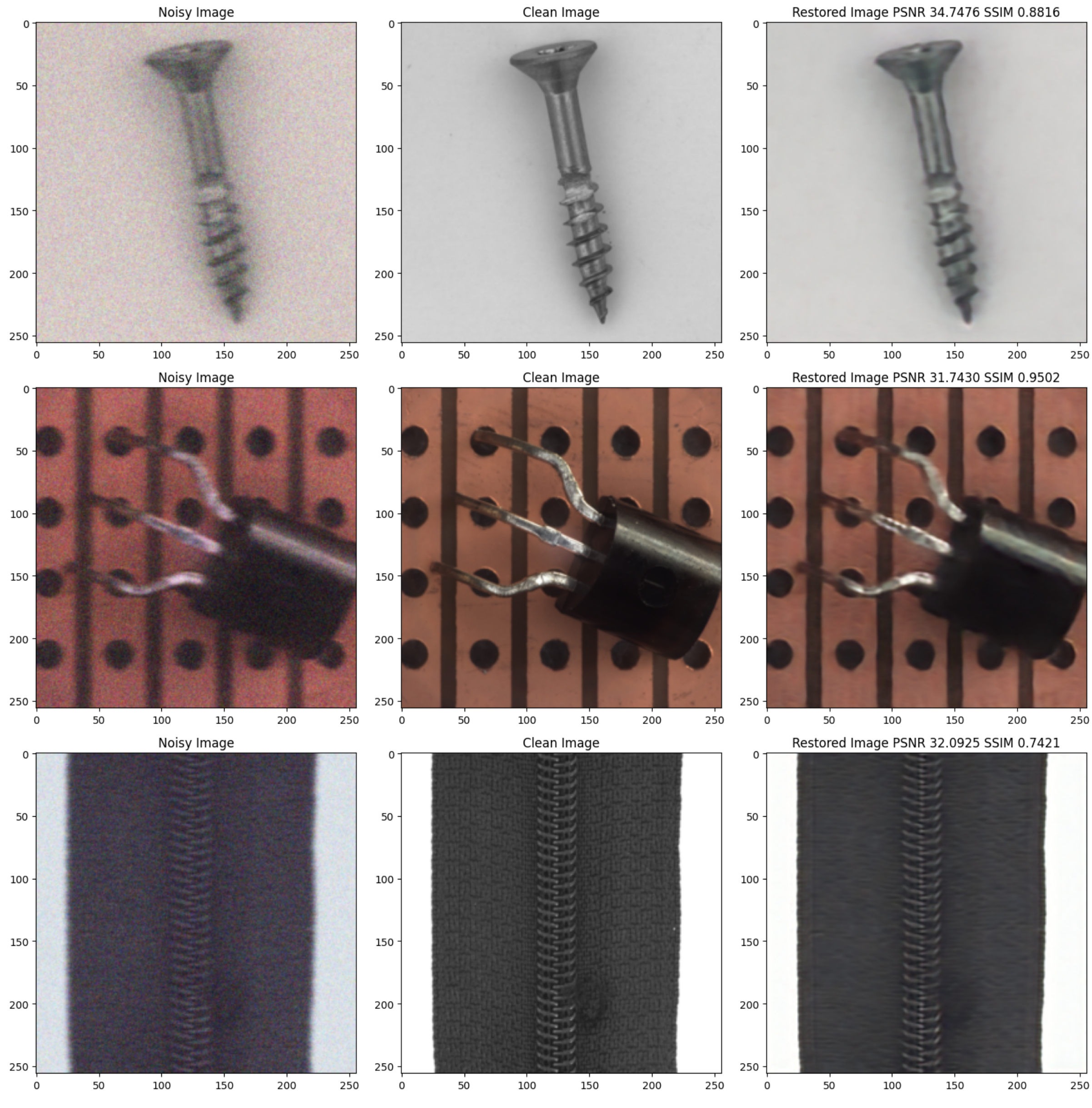

Figure 4: Sample outputs for denoised images from the validation dataset

3.2 PSNR and SSIM Evaluation

We assessed the model's denoising performance using the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics. Figures 5 and 6 show the PSNR and SSIM values, respectively, for each object in the dataset, arranged alphabetically. These metrics provide insights into the quality and similarity of the denoised outputs compared to the original images.

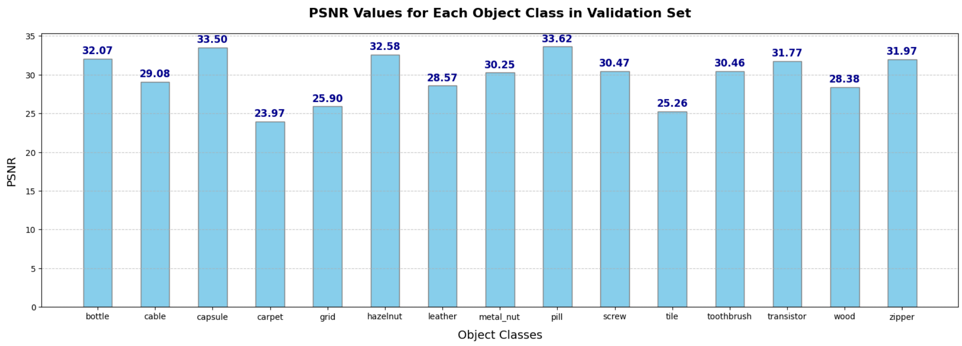

Figure 5: Average PSNR value for each object class

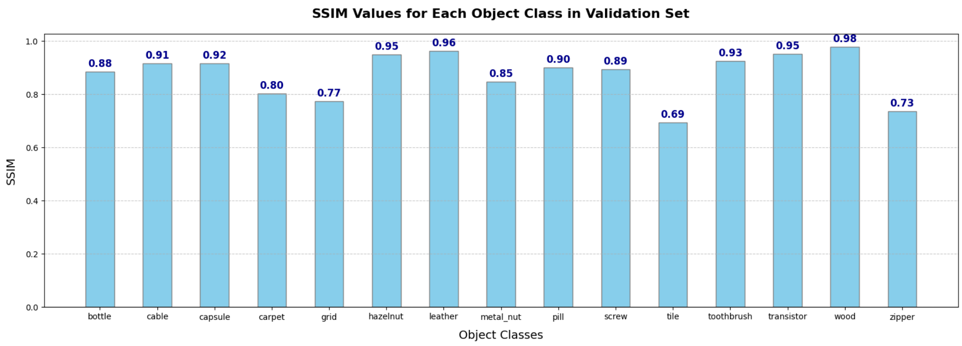

Figure 6: Average SSIM value for each object class

The PSNR chart in Figure 5 illustrates the model's effectiveness in noise reduction across various objects, with higher PSNR values indicating better preservation of image fidelity. The SSIM chart in Figure 6 complements this by measuring structural similarity, where higher values reflect closer alignment with the original image structure.We got an average PSNR value of **30.2348** and SSIM value of **0.8756** on the given validation dataset.

3.3 Defect Mask Segmentation

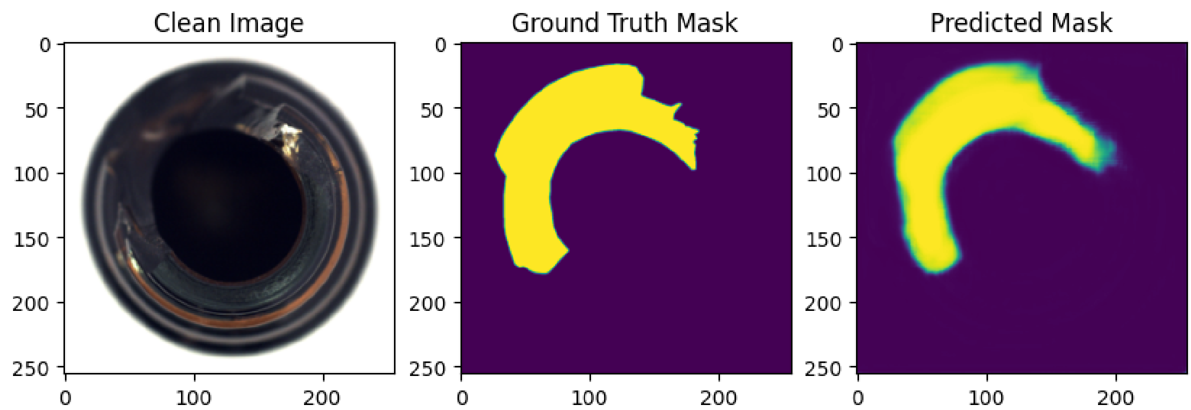

Figure 7: Defect Mask Prediction for Clean Image

The Dice Coefficient was used to evaluate U-Net's performance in mask segmentation. For mask predictions from clean images, the Dice Coefficient reached approximately 90.9%, while for predictions from denoised images, it was around 59.2%..

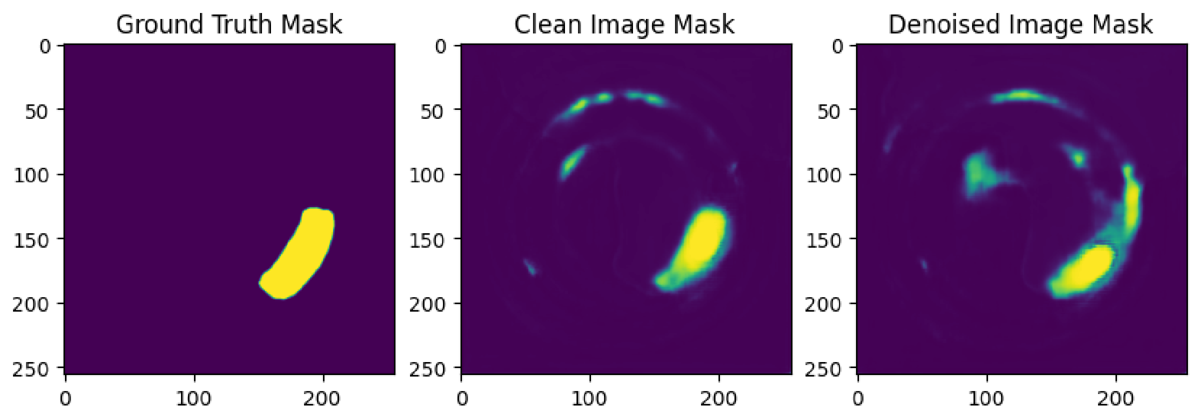

Figure 8: Defect Mask Prediction for Denoised Image

4 GitHub Repository

The codebase and model weights are publicly available on GitHub at the following link:

- `https://github.com/gokulmk-12/KLA-Denoising-DLI`

The repository includes:

- Code files for model training and evaluation.
- Pre-trained model weights.
- A list of dependencies with version specifications.
- Instructions to replicate our experiments.

References

[1] Saeed Anwar, Nick Barnes, "Real Image Denoising with Feature Attention", 2020.