DEEP CONVOLUTIONAL TRANSFER LEARNING BASED MODEL WITH U-NET ARCHITECTURE AND ATTENTION MECHANISM FOR DENOISING MEDICAL MR IMAGES

Akshaya Thomas
Dawa Namgyal
Devi Krishna K R
Dhanya Babu

CONTENTS

- Introduction
- Literature Survey
- Objectives
- Methodology
  - Network Architecture
    - Data Augmentation
    - Transfer Learning
    - Dilated Convolutions
    - Attention Mechanism
    - VGG Perceptual Loss
- Results
  - Denoising Results for MRI with Gaussian Noise
  - Denoising Results for MRI with Rician Noise
- Conclusion
- References

INTRODUCTION

- MRI images are found to be deeply corrupted by noise distributions like Gaussian and Rician.

- Our aim is to remove noise from brain MRI images using DL in an effective and outstanding way than the existing models.

- We started our project by developing a Convolutional Denoising Autoencoder (CDAE), which proved successful in removing Gaussian noise.

- However, when applied to MR images with Rician noise, CDAE did not yield satisfactory results.

- To address the limitation we advanced our model from Convolutional Denoising Autoencoder (CDAE) to UNet-DCTD-A (Deep Convolutional Transfer Learning Based Denoising Model with U-Net Architecture and Attention Mechanism).

- **UNet-DCTD-A** incorporates
  - UNet based Architecture
  - Transfer learning based encoder
  - Decoder with attention mechanism (scSE blocks)
  - Data augmentation techniques
  - Dilated convolutions

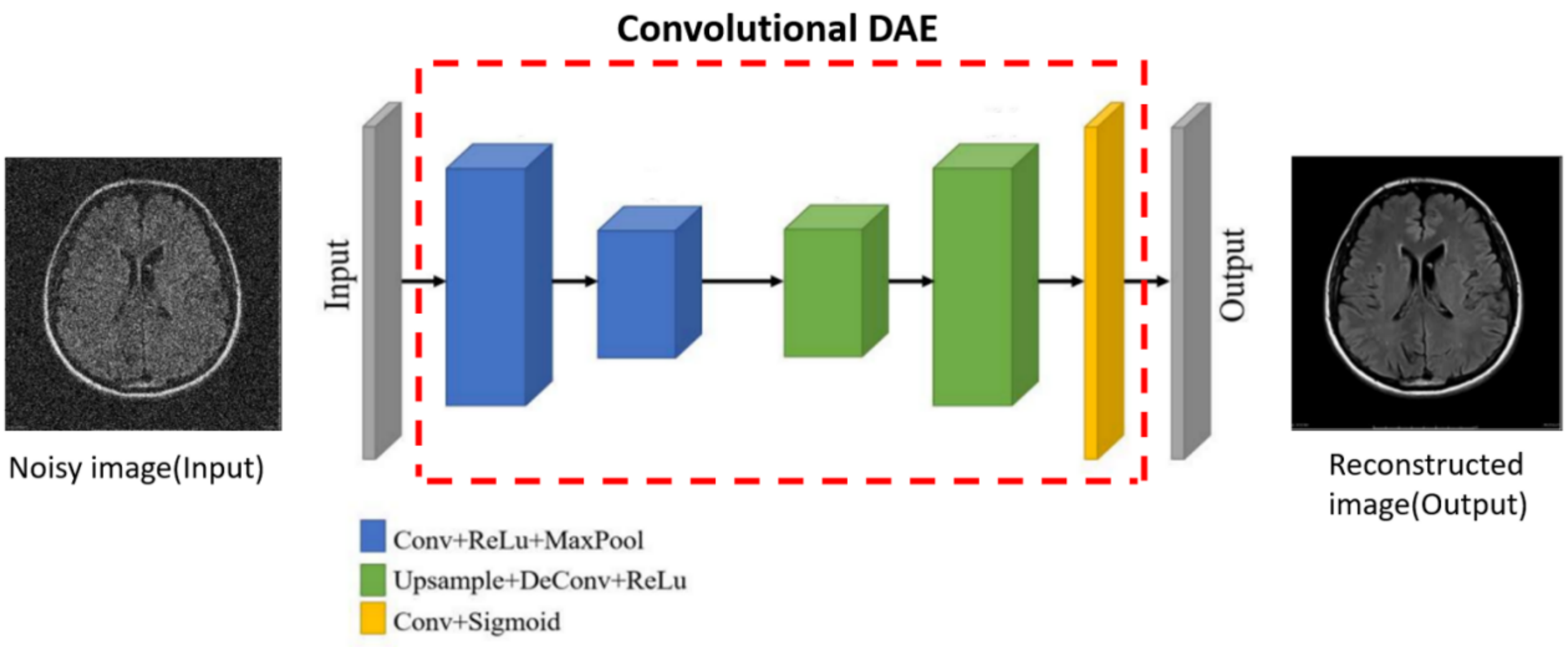
LITERATURE SURVEY

Name of the paper	Authors	Organization & Year of Publication
Improving image quality in low field MRI with deep learning	Hernandez , Fau , Rapacchi, Wojak , Mailleux , Benkreira, Adel	IEEE - 2021
U-net: Convolutional networks for biomedical image segmentation	Ronneberger, Olaf, Philipp Fischer, and Thomas Brox	Springer International Publishing - 2015
Attention-Based Convolutional Denoising Autoencoder for Two-Lead ECG Denoising and Arrhythmia Classification	P. Singh and A. Sharma	IEEE - 2022
Quadratic Autoencoder (Q-AE) for Low-Dose CT Denoising	F. Fan et al	IEEE - 2020
Concurrent spatial and channel 'squeeze & excitation in fully convolutional networks	Roy, Abhijit Guha, Nassir Navab, and Christian Wachinger	MICCAI - 2018

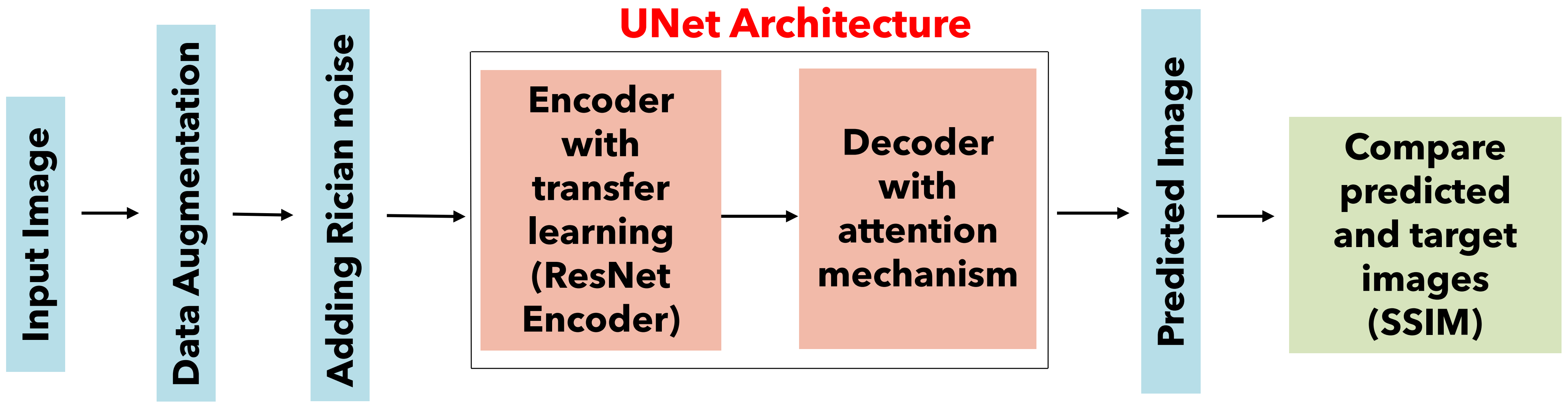
OBJECTIVES

- Develop a robust deep learning-based model to effectively remove noises like Gaussian and Rician from brain MR images.
- Comparative analysis with existing models to verify the performance of the developed model in denoising brain MR images.
- Investigate and integrate various advanced deep learning approaches into the developed denoising model to enhance the SSIM metric.

METHODOLOGY



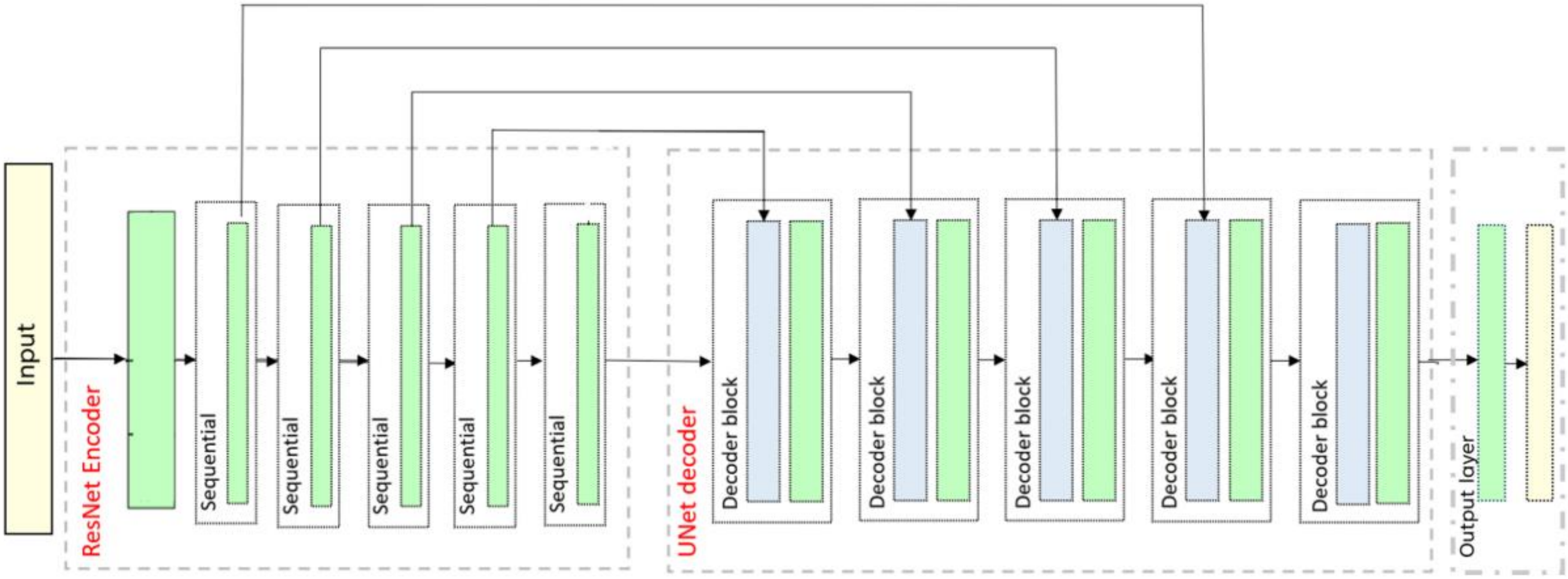
Block Diagram of Convolutional DAE for MRI image denoising
(Base Model)



Block Diagram of UNet-DCTD-A for MRI image denoising

(Basic Model integrated with various deep learning techniques)

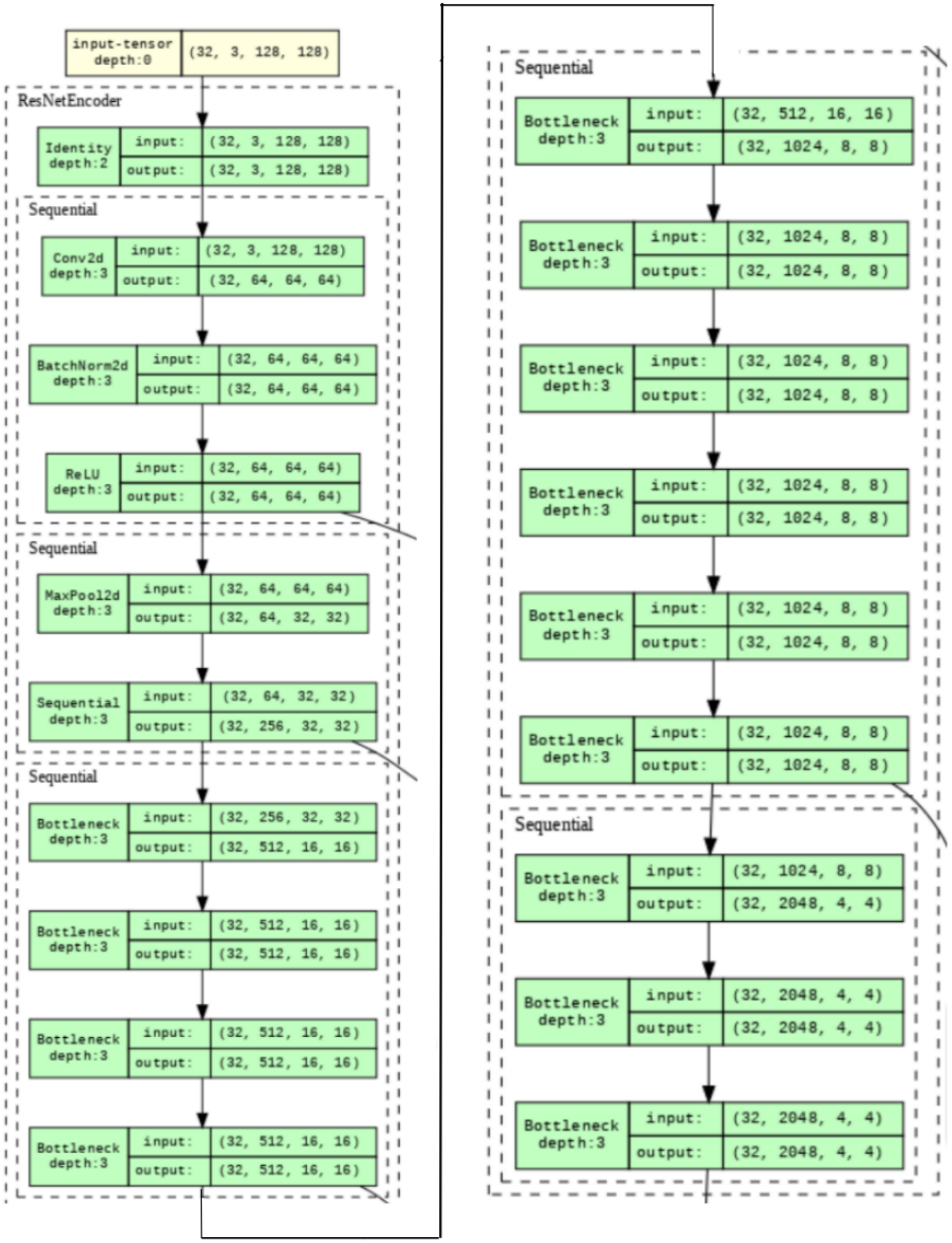
NETWORK ARCHITECTURE



Detailed view of UNet-DCTD-A

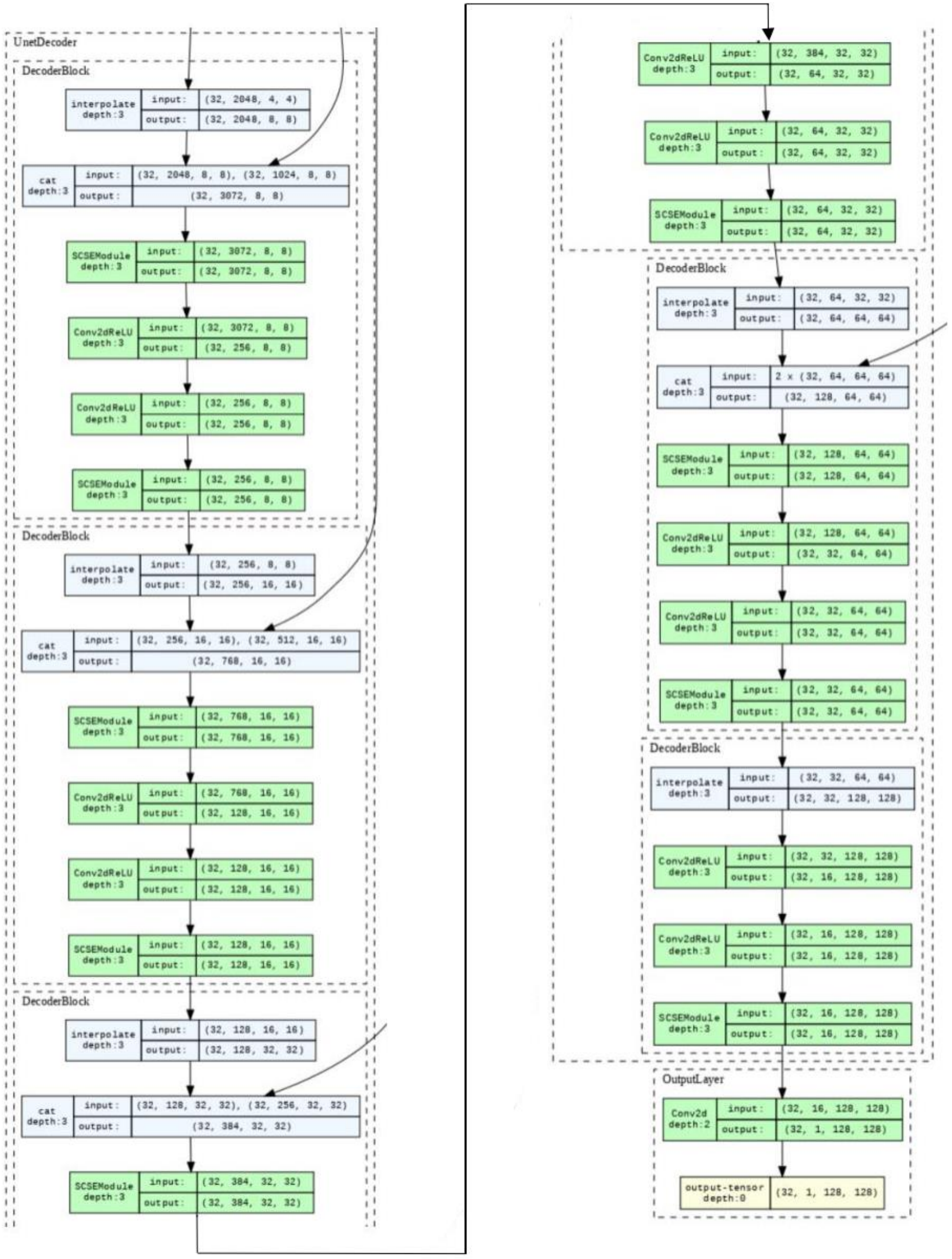
ENCODER

- Based on ResNet architecture
- Consists of series of convolutional layers with batch normalization and ReLU activation
- 5 stages
- Incorporates Transfer learning
- A pretrained ResNet18 model is used



DECODER

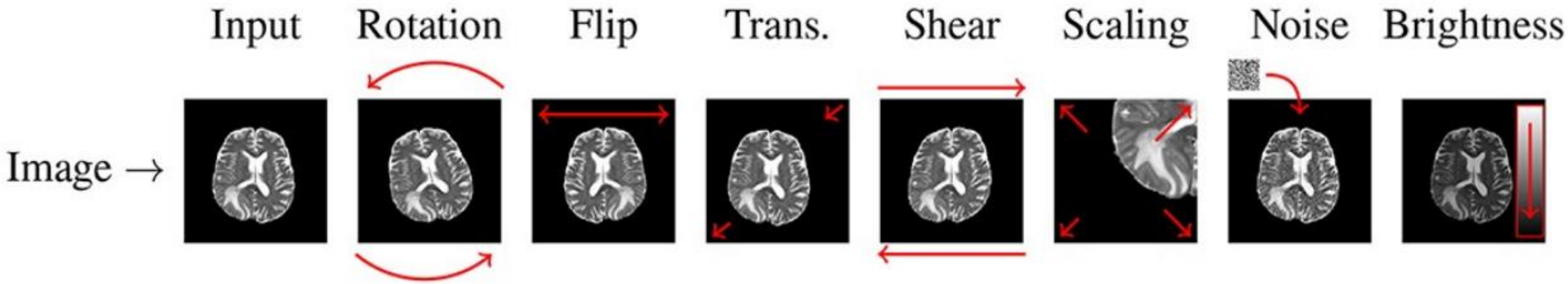
- Based on UNet architecture
- Includes attention modules (SCSE Module) for capturing spatial and channel-wise attention.
- Attention modules contain combinations of global average pooling, convolutions, ReLU activation & sigmoid activation to compute attention maps.
- Final output of the decoder is the denoised Image



DATA AUGMENTATION

- Strategy for increasing the amount and variety of a training dataset by modifying and transforming the current data in numerous ways.

- Aims to increase the model's performance and capacity to generalize.



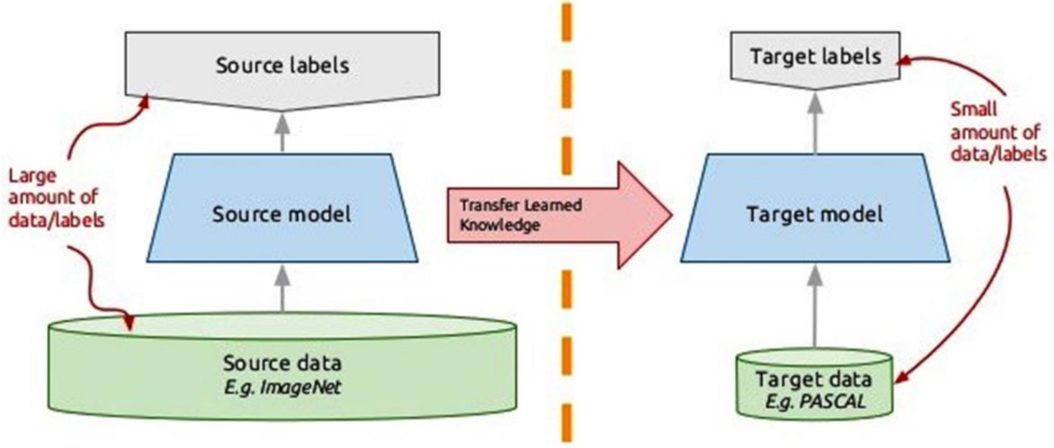
Data Augmentation Example

TRANSFER LEARNING

- Used to leverage pre-trained models on large datasets and adapt them to solve different tasks or work with smaller datasets.

- In our model we have used **ResNet-18** as a pretrained model for transfer learning.

- ResNet-18 is a 18-layer convolutional neural network (16 convolutional layers, one MaxPool layer, and one average pool layer).

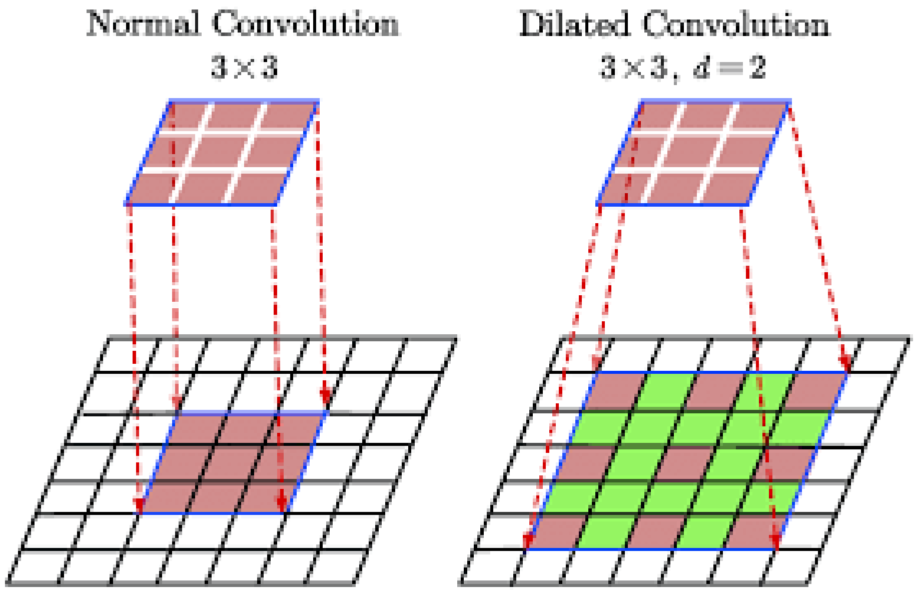


DILATED CONVOLUTIONS

- Introduce gaps or holes between the filter weights, allowing the filter to have a larger effective receptive field.

- Advantage: Allow for exponentially increasing the receptive field without increasing the number of parameters.

- Used when a large receptive field is required, but computational resources are limited.



ATTENTION MECHANISM

- It will enhance the representation power of the network by selectively attending to the most relevant information.

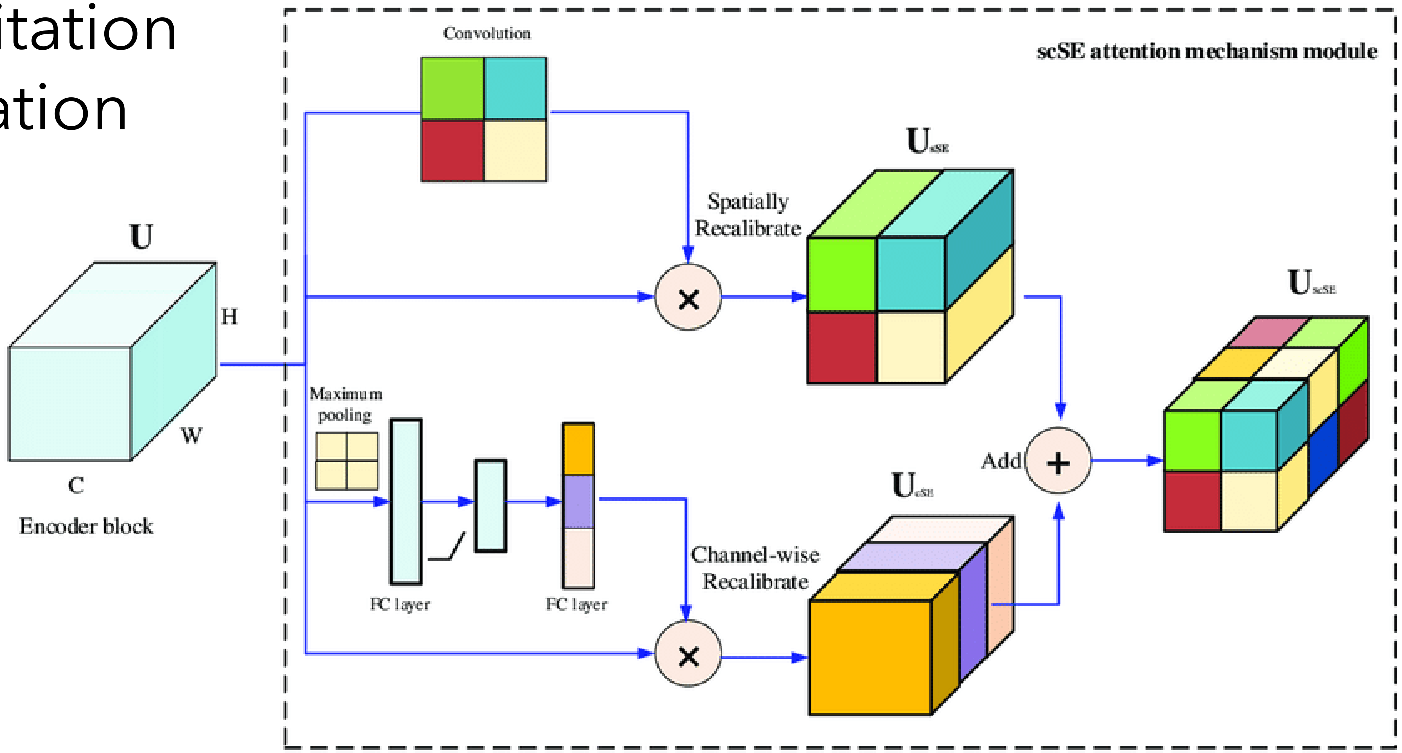
- It can help the network to focus on discriminative features, suppress noise or irrelevant information, and improve its ability to understand complex patterns.

- The design and implementation of attention mechanisms in our model is based on the attention modules called **Spatial & Channel Squeeze-and-Excitation (SE) blocks.**

-

- Convolutional neural networks (CNNs) employ the scSE (Spatial-Channel Squeeze & Excitation) module as a strategy to enhance the representation and modelling skills of the network.

- Steps are:
  - Squeeze operation
  - Channel-wise excitation
  - Spatial-wise excitation
  - Combination

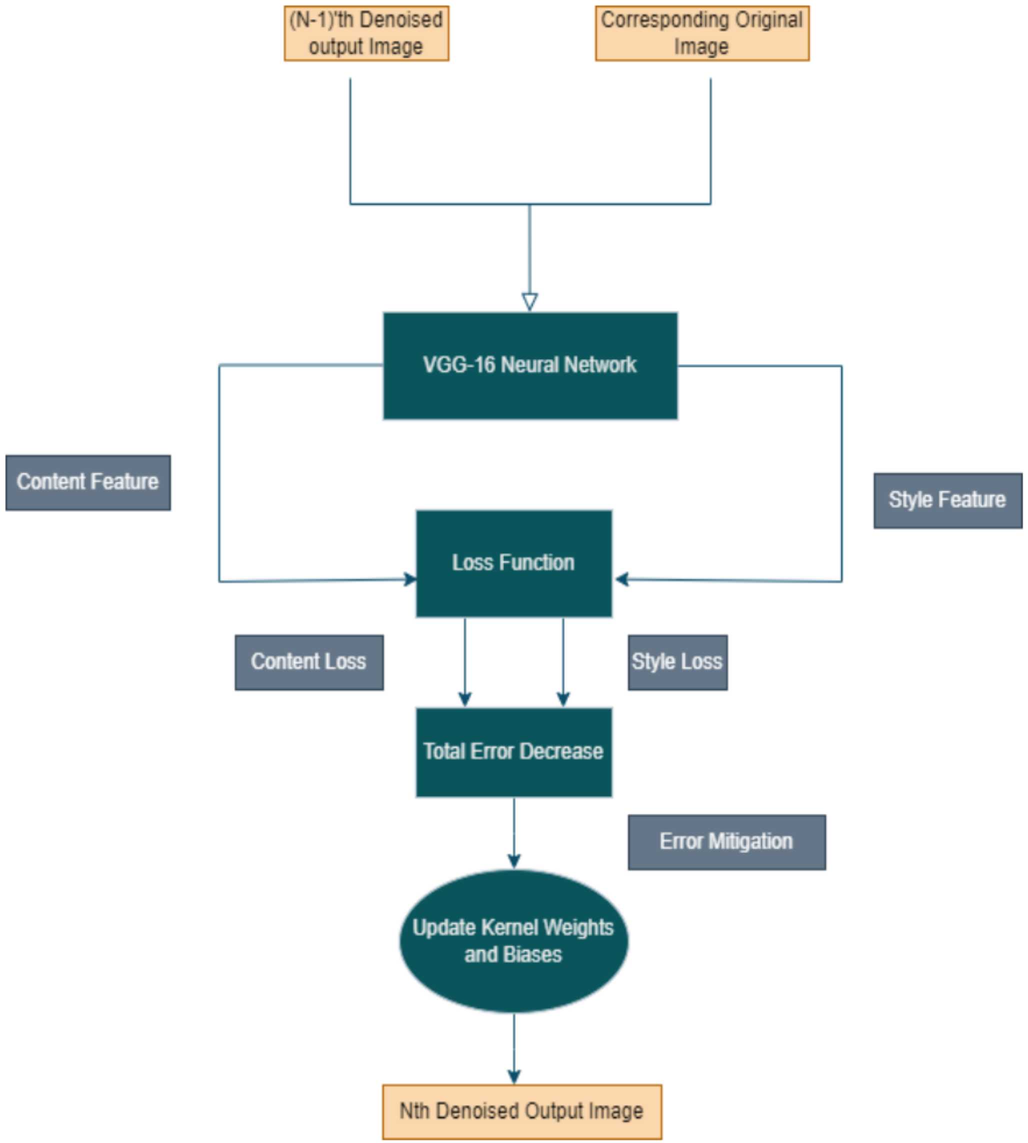


VGG PERCEPTUAL LOSS

- VGG perceptual loss measures the difference between two images in terms of their perceptual similarity, rather than pixel-wise differences.
- It utilizes a pre-trained VGG network (VGG16), to extract high-level feature representations from the input images.
- These feature representations capture the content and style information of the images at different layers of the network.
- Training the denoising autoencoder with this loss preserves structural and textural information, resulting in denoised brain MRI scans resembling the clean originals.

- To compute the VGG perceptual loss, the input images and the generated images (e.g., stylized images or high-resolution images) are passed through the VGG network, and the feature maps at specified layers are extracted.

- The loss is then calculated as the mean squared error (MSE) between the feature maps of the input and generated images.



RESULTS

1. Denoising Results for MRI with Gaussian Noise

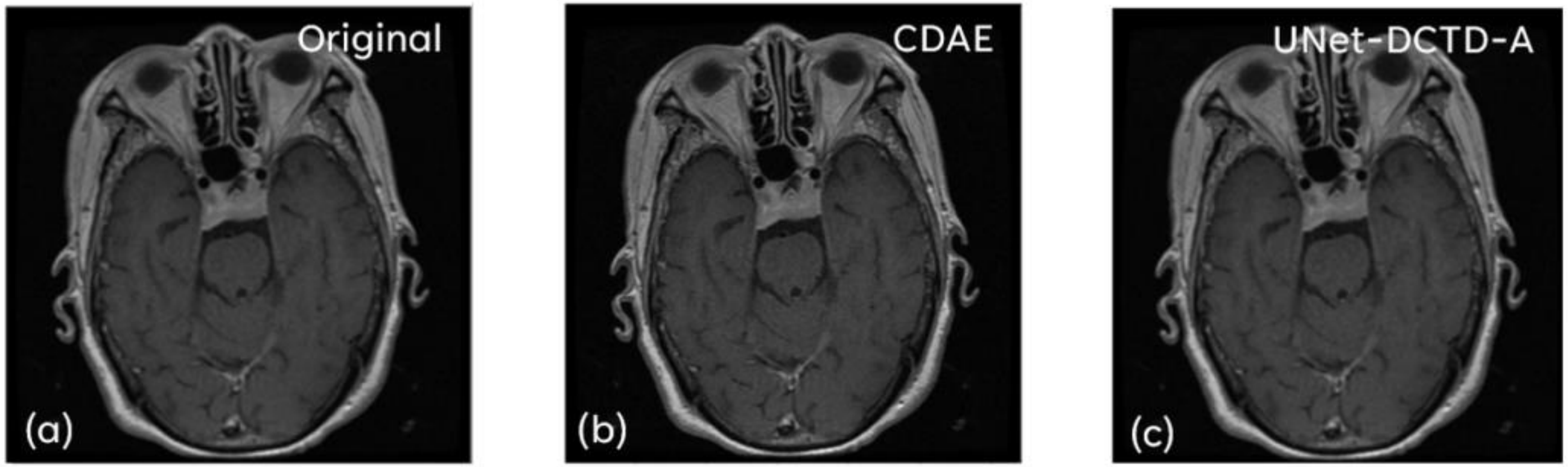


Figure 1: a) Original Image b) Reconstructed Image using CDAE c)Reconstructed Image using UNet-DCTD-A

Epochs : 200 SSIM (b) : 0.92 SSIM (c) : 0.96

1.1 Comparison with Traditional Methods

Table 1: Comparison of denoising performance of different methods with our model for gaussian noise with noise factor=0.1, mean=0 and standard deviation =1

Filter	**SSIM**	**PSNR**
Gaussian	0.5681	24.6228
Bilateral	0.6805	21.9754
Total Variation(TV)	0.6824	22.3618
Wavelet	0.7574	24.9854
Shift invariant	0.7982	26.1192
Non Local Means(NLM)	0.8214	26.3957
Block Match 3D(BM3D)	0.8343	29.1254
CDAE	0.9235	33.7564
UNet-DCTD-A	0.9613	38.4684

2. Denoising Results for MRI with Rician Noise

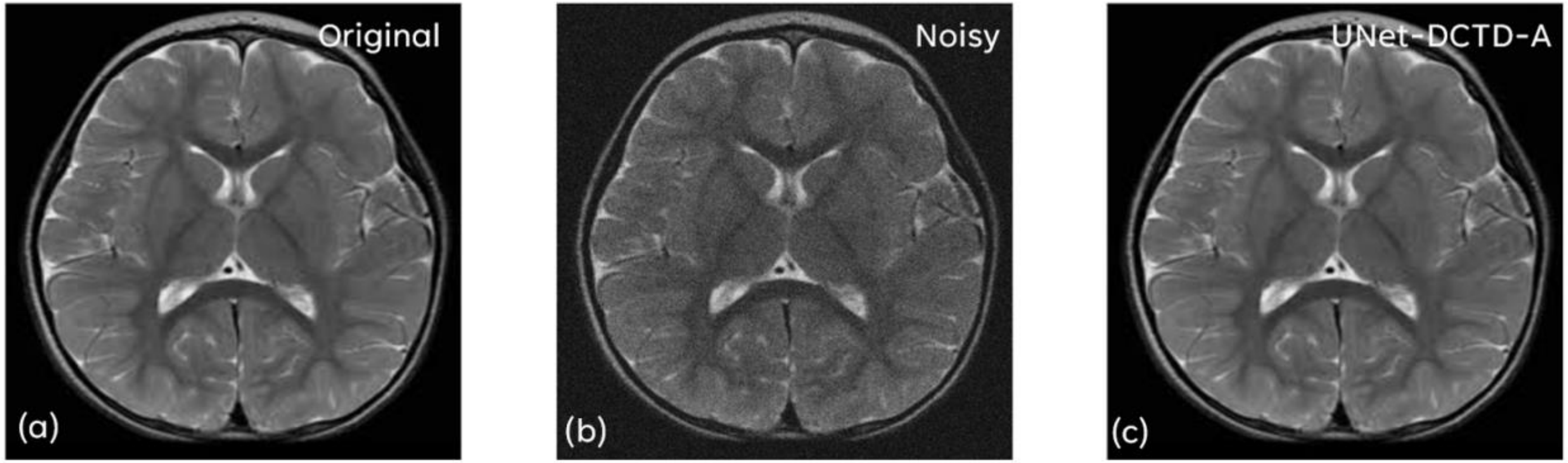


Figure 2: a) Original Image b) Rician Noise Added Image c) Reconstructed Image using UNet-DCTD-A

Epochs : 200 SSIM : 0.94 PSNR: 37dB

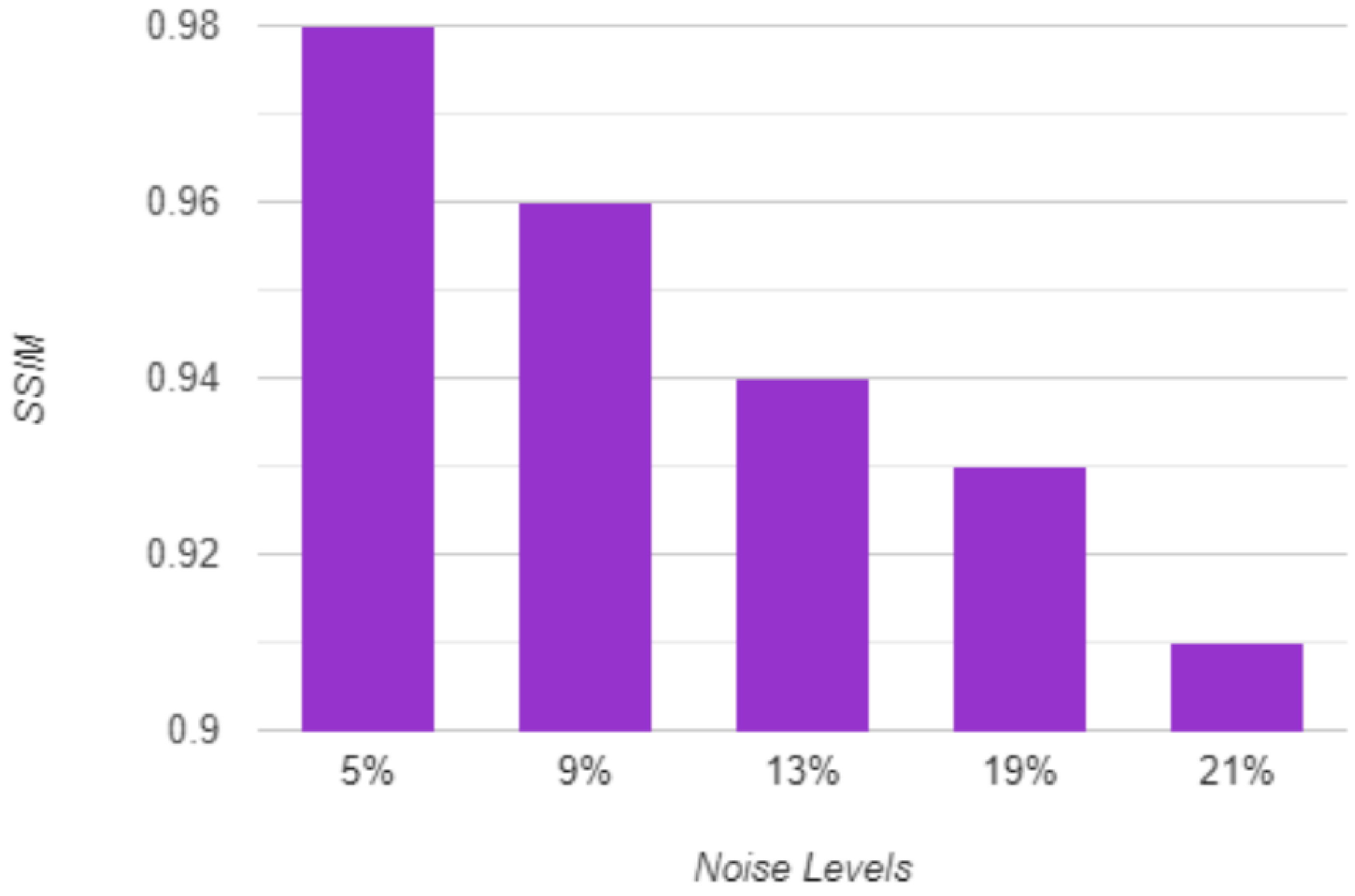


Figure 3 : Performace of the model on various noise levels

2.1 Comparison with Traditional Methods

Table 2: Comparison of denoising performance of different methods with our model for rician noise with sigma=0.11

Filter	SSIM	PSNR
Gaussian	0.3562	16.8386
Bilateral	0.4709	17.7383
Total Variation(TV)	0.4895	17.0059
Wavelet	0.4105	16.8780
Shift invariant	0.44084	16.9021
Non Local Means(NLM)	0.4824	16.9431
Block Match 3D(BM3D)	0.4835	16.9234
CDAE	0.8012	21.8254
UNet-DCTD-A	0.9415	36.7546

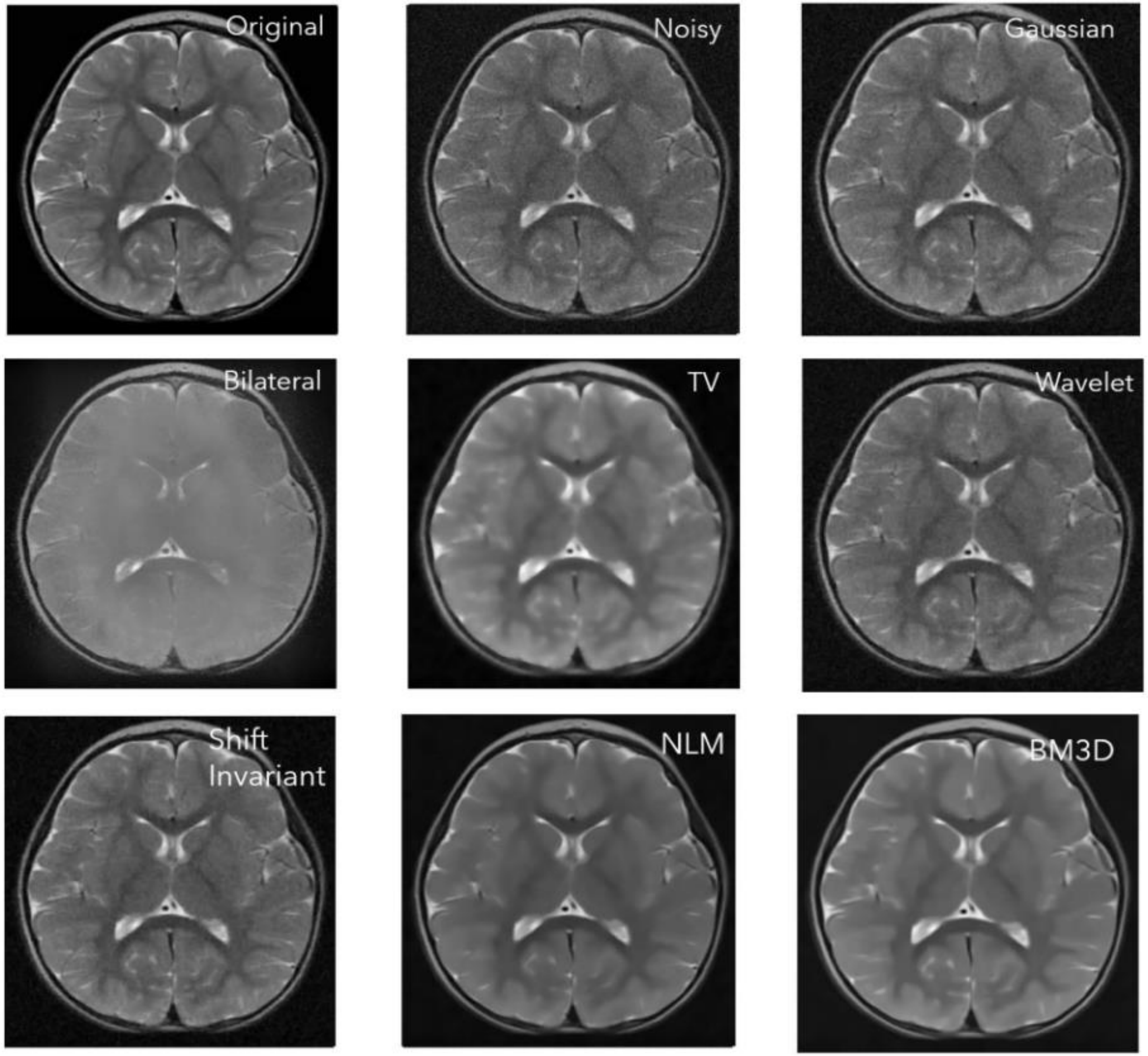
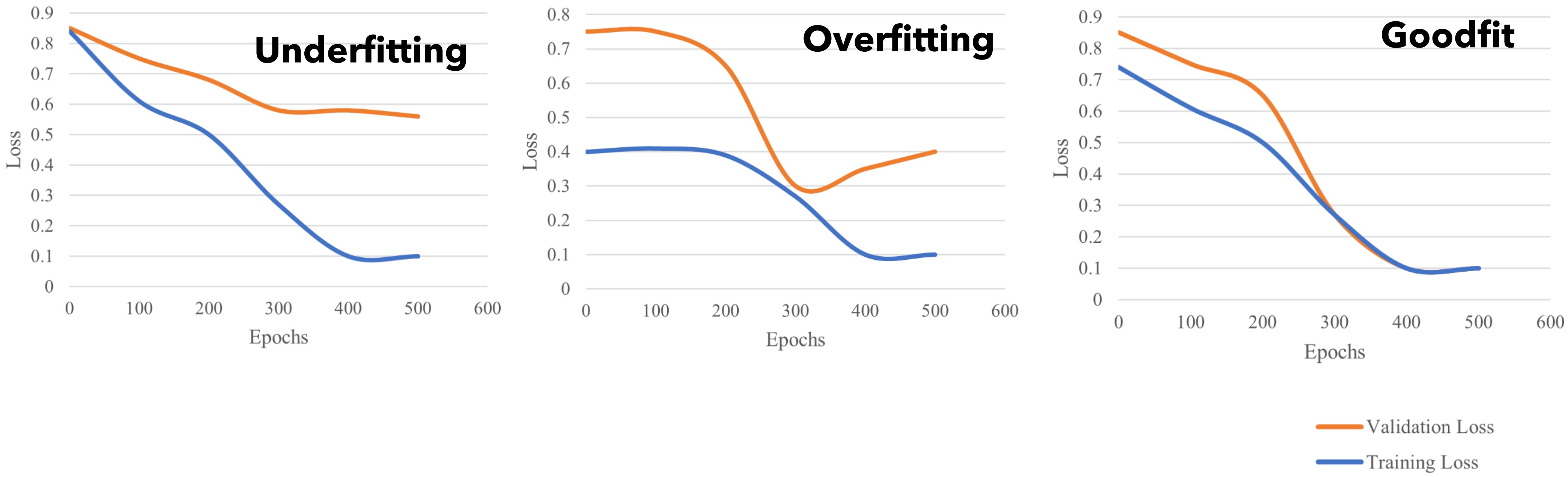
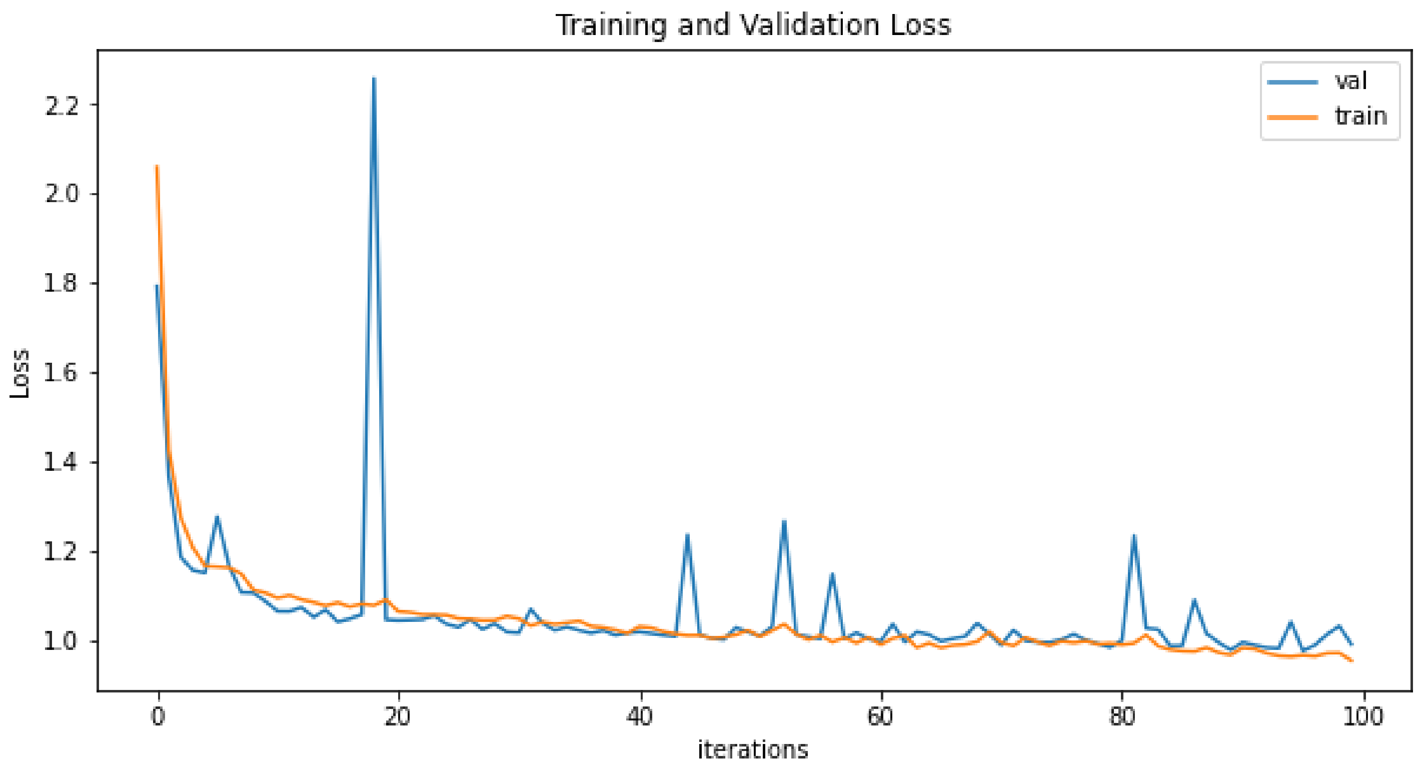


Fig 4 : Visual inspections of brain MRI with rician noise denoised by traditional methods

TRAINING & VALIDATION

- Training loss is used to assess how a deep learning model fits the training data while validation loss is used to assess the performance of a deep learning model on the validation set.





Loss function V/S No: of epochs

- The graph shows that as training proceeds over time, the loss function is significantly decreasing.

- This demonstrates that our model is fitting the data well.

CONCLUSION

- In 200 epochs, we obtained an SSIM score of 0.96 for MRI with Gaussian noise and 0.94 for MRI with Rician Noise (13% noise).
- The previous graph shows that data is fitting the model well.
- So increasing the epoch count could potentially enhance the SSIM performance. But practical considerations such as time and resource constraints limited our training duration to 200 epochs.
- The obtained SSIM values represent a remarkable achievement, indicating the exceptional denoising capabilities of our proposed model. Thus we can say that our model has achieved state of the art performance.

REFERENCES

1. A. G. Hernandez et al, “Improving Image Quality In Low-Field MRI With Deep Learning” , 2021 IEEE International Conference on Image Processing (ICIP)

2. Gurrola-Ramos, Javier, Oscar Dalmau, and Teresa E. Alarc ́on. “A residual dense u-net neural network for image denoising”, IEEE Access 9 (2021): 31742-31754

3. Jifara W, Jiang F, Rho S, Cheng M, Liu S. “Medical image denoising using convolutional neural network: a residual learning approach”, The Journal of Supercomputing. 2019 Feb 6;75:704-18.

4. Roy AG, Navab N, Wachinger C. “Concurrent spatial and channel squeeze excitation in fully convolutional networks”, In Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I 2018 (pp. 421-429). Springer International Publishing.

5. L. Gondara, “Medical Image Denoising Using Convolutional Denoising Autoencoders”, 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 2016

THANKYOU