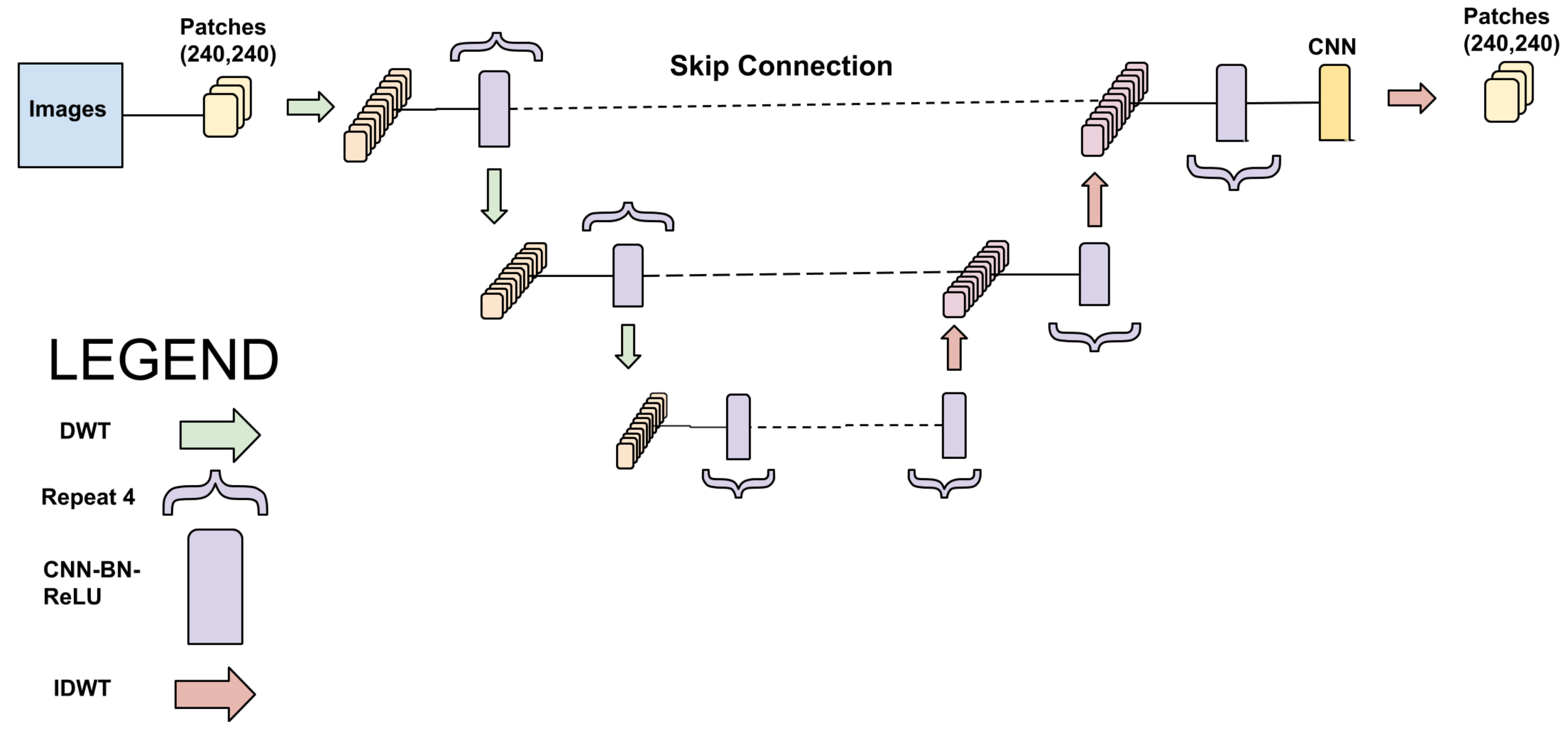
TheMulti-level Wavelett-CNN for Denoising Natural Noise Benchmarks

Arjun Arora
Stanford University

Motivation

- Denoising images is a common application of optimization algorithms. Neural networks of sufficient depth are known to be universal function approximators, which gives them a strong representational power in denoising.[1]
- U-Nets have been shown to do well in denoising tasks, specifically in biological scenarios [2]
- The Multi-level Wavelett-CNN, to reduce parameter size and extract more information, replaces downsampling and upsampling in U-Net with DWT and IDWT [3]
  - In this paper, only independent gaussian noise was used. However, such noise is not representative of all noise images from cameras exhibit.
  - We wanted to test how this model would perform on other types of noise

Model



- The Multi-level Wavelett-CNN used was implemented in Pytorch using the paper described by Liu et al [3] as a guide
- Each Conv-BN-ReLU layer was repeated 4 times per layer, with the first layer having 16 channels per convolution, and the second and third layer having 32 channels per convolution each.
- Skip connections concatenate feature maps across the contracting and expanding networks
- The final Layer is a simple CNN that feeds into a IDWT layer
- The model is trained using MSE loss and evaluated using PSNR

Methodology

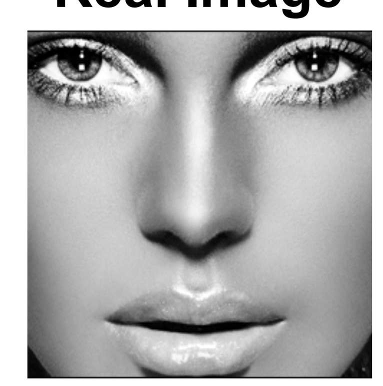
- In this project, we explored three different types of noise:
  - Independent gaussian noise
  - A mixture of independent gaussian noise and image dependent noise
  - Impulse noise
- For the first two noise types, we used a std dev of 6% for the gaussian noise
- The images we used were pulled from a combination of Berkeley Segmentation Dataset, DIV2K, and the Waterloo Exploration Database
  - All images were segmented into ~24 240 x 240 patches and converted into grayscale
- Due to Memory/ GPU constraints the training set was reduced to ~ 22,000 (compared to the ~137K of Liu et al [3]) 240 x 240 pixel images. Val set and train set sizes were 2400 images each

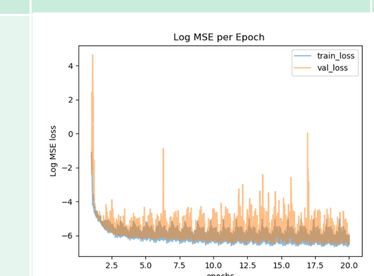
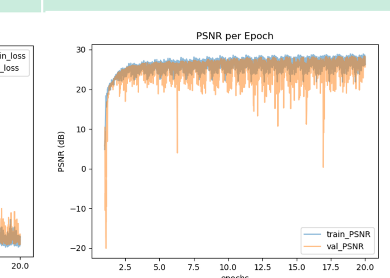
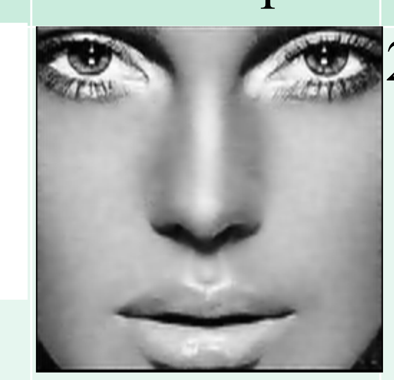
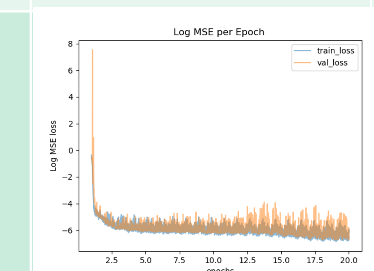
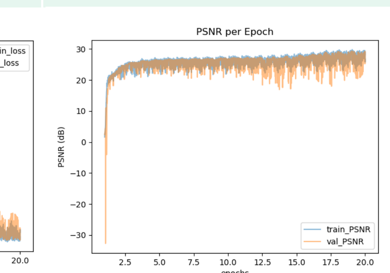
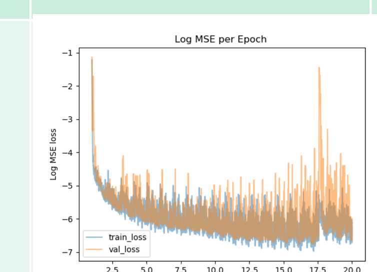
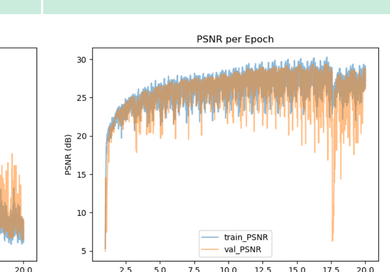
Experimental Results

Analysis

- Overall, the dependent + gaussian noise model did the best on the test set, having ~1.0 dB better PSNR than pure gaussian noise
  - This may be due to the fact that the noise in the mixed noise model has noise that is correlated to the magnitudes of the pixel values in images.
- The impulse noise model produced the worst results of all
  - Most likely, this is related to impulse noise essentially removing information at certain pixel values, requiring the model to do a pixel wise "inpainting" task rather than just denoising

Real Image



	Training Loss	Training PSNR	Ex. Output	Test PSNR
Gaussian Noise				28.05 dB
Dependent + Gaussian Noise				29.16 dB
Impulse Noise				26.04 dB

References

[1] Kur Hornik, "Multilayer Feedforward Networks are Universal Approximators",Neural Networks, 1989
[2] Heinrich, Stille, and Thorsten, "Residual U-Net Convolutional Neural Network Architecture for Low-Dose CT Denoising", Current Directions in Biomedical Engineering, 2018
[3] Liu et al , "Multi-level Wavelet-CNN for Image Restoration", arXiv, 2018