

ECE 251C- Project Proposal

Experimenting with Multi-level Wavelet-CNN for Image Restoration

Suhrid Subramaniam : A53321749

Punit Agrawal : A53300814

Problem Statement:

Image restoration, fundamentally speaking, is the operation of taking a corrupt or noisy image and estimating the clean/ original image using it. CNNs have played a pivotal role in accelerating the research in this field by helping in achieving state-of-the-art performance in many image restoration applications. However, with an ever increasing demand for higher accuracy and close-to-perfect restoration, the models used for restoration have become computationally intensive owing to the demand for a larger receptive field. Although solutions like dilating filters have been proposed and used in the past for increasing the receptive field, they too suffer from drawbacks like the presence of gridding effect and uncorrelation between neighboring pixels in the resultant pixel map.

The goal of this architecture (MWCNN) is to maximize the receptive field and minimize computational intensity while ensuring that the run-time remains low. The metric used to compare the quality of restoration is PSNR (dB). We also see that PSNR is the highest for MWCNN architecture followed by MemNet, DRRN and more.

In this project, we will implement the [MWCNN architecture](#) and experiment with different inputs and variants of the MWCNN architecture so as to be able to prove the universality of its use or disprove the architecture's usage under certain circumstances. We will use two types of data as input: Daylight input and nightlight input. We will also experiment with the type of noise (Salt and Pepper, Gaussian, Poisson) added to each image and see if changing the input or noise affects the model in a drastic manner. We will also try to tweak the optimization, learning rate, activation and learning rate to see if there is any improvement in the model. Different wavelets can also be experimented with to see if using a particular method changes the results.

Conventional Method:

The conventional methods used for different image restoration purposes are different. For image denoising, the architecture that is usually preferred is the feed-forward denoising convolutional neural networks ([DnCNNs](#)), U-net DnCNN (UDnCNN) or the Dilated convolutions UDnCNN (DUDnCNN) architectures. These architectures make use of residues/ skip connections

to provide better detailing. They also use max pooling and transpose convolution in their contracting and expanding subnetworks.

For Single Image Super Resolution, Super-resolution deep convolutional neural networks (SRCNN) are traditionally used. This architecture uses extracts the feature maps, performs a non-linear mapping from the lower level feature maps to the higher level feature maps and reconstructs the image based on the high resolution feature maps.

JPEG artifacts removal is traditionally done using Artifact removal CNNs (ARCNN) which is basically a Deep Convolutional network with large-stride convolutional and deconvolutional layers.

Advantage of Using Wavelets:

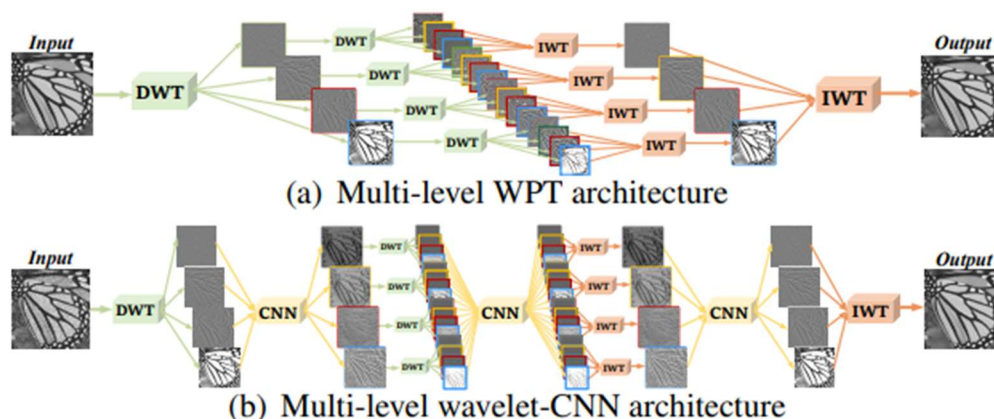
Each of the above proposed methods have their own downsides. Most of these problems eventually result in lower receptive field or larger computational complexity.

Using multi-level wavelet packet transform (WPT), we try to provide meaningful feature maps to the CNNs by ensuring good frequency and spatial information. The operations of max pooling and transpose convolution are replaced by Discrete Wavelet Transform (DWT) and Inverse Discrete Wavelet Transform which helps in preventing information loss. Haar Wavelet Transform is used at every stage.

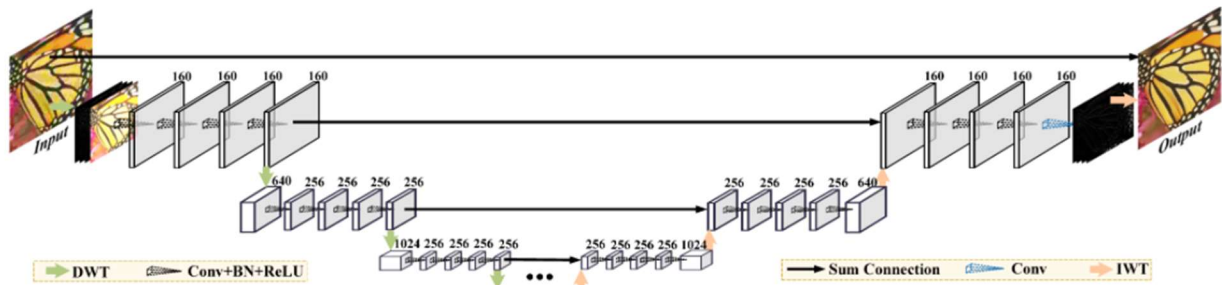
We also see that some parts of the network like skip connections in UDNnCNNs can be omitted as they are already embedded in MWCNN architecture.

We can also use MWCNN as an alternative to dilating filters which cause information loss because of sparse sampling of locations with the checkerboard pattern.

As each DWT produced sub-band is treated dependent to the other sub-bands, we get a much larger receptive field even in comparison to other Wavelet based networks like Deep convolutional framelets. The network architecture is shown below:



Each CNN block consists of 4-layer FCN without pooling, and takes all the sub-band images as inputs. A more detailed version of the above network is shown below:



Project Plan:

Firstly, we intend on implementing the DnCNN, UDnCNN and DUDnCNN architectures and getting the PSNR as well as timing results for them after image denoising.

We then plan on implementing the MWCNN architecture, retrain it for image denoising, super resolution and JPEG artifact removal and check the PSNR and run-time for the same.

Our implementation of MWCNN is broken down into three phases:

PHASE 1: Reimplementing the paper

- 1) Read up on related paper and collect datasets.
- 2) Clean and preprocess datasets.
- 3) Implement WPT architecture.
- 4) Add CNN blocks to the WPT architecture to get the MWCNN architecture.

PHASE 2: EXPERIMENTING WITH DIFFERENT INPUTS AND NOISES:

- 1) Retrain the model using nightlight datasets
- 2) Retrain the model for salt and pepper noise, gaussian noise and Poisson noise.

PHASE 3 (Optional): EXPERIMENTING WITH NETWORK

- 1) Try to see if the network accuracy improves by using different activation functions (leaky relu), different optimizations (adagrad, adam), learning rate, different types of wavelets (haar, Daubechies).

Datasets:

- 1) DIV2K dataset – (<https://data.vision.ee.ethz.ch/cvl/DIV2K/>)

References:

- 1) Pengju Liu and Hongzhi Zhang and Wei Lian and Wangmeng Zuo (2019). Multi-level Wavelet Convolutional Neural NetworksCoRR, abs/1907.03128.
- 2) Kai Zhang and Wangmeng Zuo and Yunjin Chen and Deyu Meng and Lei Zhang (2016). Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image DenoisingCoRR, abs/1608.03981.