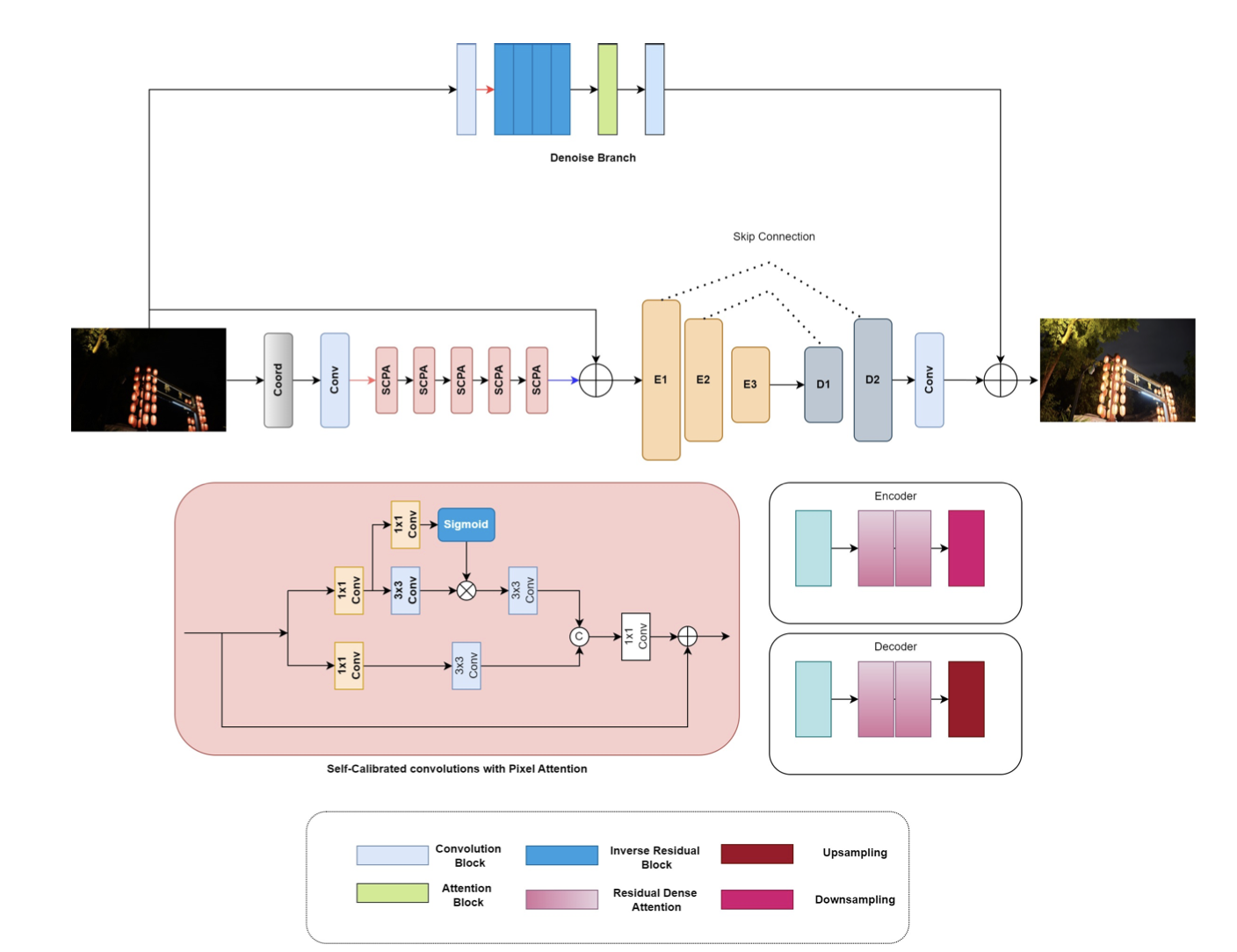
**REPORT FOR LOW-LIGHT IMAGE DENOISING**

AKSHIT MANOCHA

Enrollment no. 23112011

**Overview:** The model that has been used here is ImageNet that was used in NTIRE-2024 Low Light Image Enhancement competition



### **Architecture Overview:**

1. **CoordConv Layer**: The input image is processed through a CoordConv layer to add spatial information channels.
2. **SCPA Layers**: The enhanced image representation is downsampled and passed through five consecutive SCPA layers with Pixel Attention Blocks to capture intricate spatial patterns and features. The output is then upsampled and combined with the original input image.
3. **Modified U-Net with RDCA Blocks**: The combined representation goes through a Modified U-Net architecture. The U-Net has three encoder blocks and two decoder blocks. Each encoder block has two RDCA blocks followed by downsampling. Each decoder block has two RDCA blocks followed by upsampling. This setup allows for feature extraction and refinement at multiple scales.
4. **Denoising Block**: Simultaneously, the input image is processed by a denoising block. The denoising block includes four inverse convolutional layers and an attention mechanism to suppress noise and maintain image details.This block outputs a three-channel denoised image, enhancing the overall image quality.

**Training Process**:

The training process used Adam optimizer. The learning rate started at 0.001, meaning the model made relatively larger updates in the beginning. Over the course of 100 epochs, the learning rate gradually decreased to 0.00001

**Training Objective**: The objective of training the model is to minimize a loss function, which measures how well the model's predictions match the target outputs. The loss function used in this training was a combination of three different types of loss:

1. **L1 Loss**: This measures the absolute differences between the predicted and target images. It's like checking how much each pixel value is off by and summing these differences. Minimizing the L1 loss helps ensure that the predicted images are as close as possible to the actual high-quality images.
2. **LSSIM Loss**: SSIM (Structural Similarity Index) assesses the visual similarity between two images. Unlike L1 loss, which focuses on individual pixel values, SSIM considers changes in structural information, such as luminance, contrast, and texture. By incorporating SSIM loss, the model is encouraged to produce images that not only have accurate pixel values but also look structurally similar to the target images. In the final loss calculation, this component is weighted by 0.1, indicating it has a lesser but still significant impact compared to L1 and LGrad losses.
3. **LGrad Loss**: This loss term focuses on the gradients of the images, which are related to the edges and textures within the images. By minimizing the gradient loss, the model learns to produce sharper images with more detailed textures and edges.

The combined loss function can be represented as:

L = 0.1 ×L SSIM + L1 + LGradL

This means that the total loss is a mix of all three components, with SSIM loss having a smaller contribution. The model aims to reduce this total loss during training, resulting in enhanced low-light images that are visually similar, have accurate pixel values, and retain sharp edges and textures.

In summary, the training process for the proposed network involves a carefully designed optimizer and loss function, enabling the model to effectively learn how to improve the quality of low-light images by focusing on different aspects of image quality.

### **Summary:**

This architecture combines spatial information enhancement, intricate feature capture, multi-scale feature refinement, and effective noise suppression to improve image denoising and enhancement.

The output Enhanced Image got an average PSNR value of around 24 dB on Training Set

The proposed network, designed to enhance low-light images, was trained using NVIDIA Tesla P100 GPU on Kaggle Notebook

Research Paper used for reference : <https://arxiv.org/pdf/2404.14248>

