TransXGBoost

Objective

To build a deep learning model that can convert a low-light image to a high-quality, well-lit image.

Introduction

In the current technological era, visual content has become an essential aspect of our lives, impacting many fields like entertainment, social media, education, and business but capturing images of high quality in low-light settings presents a significant challenge, often resulting in images with poor brightness, noise, and blurred details. Therefore, we have tried to use an AI model to combat this problem.

Overview

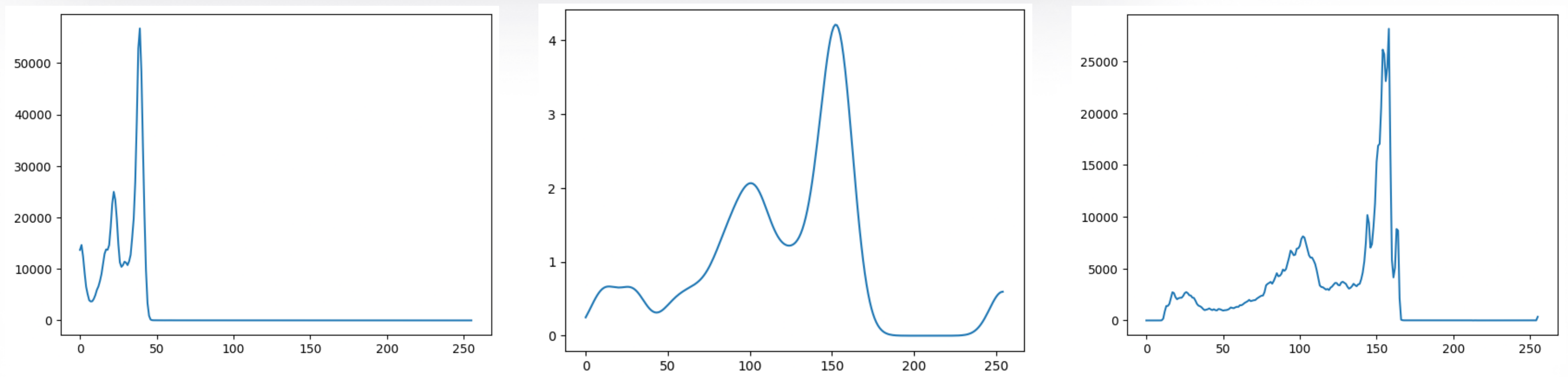
In our project we used the combined power of machine learning and deep learning to achieve the task. First we used XGBoost to predict the histogram of target image as a preprocessing step for the transformer model .

Dataset

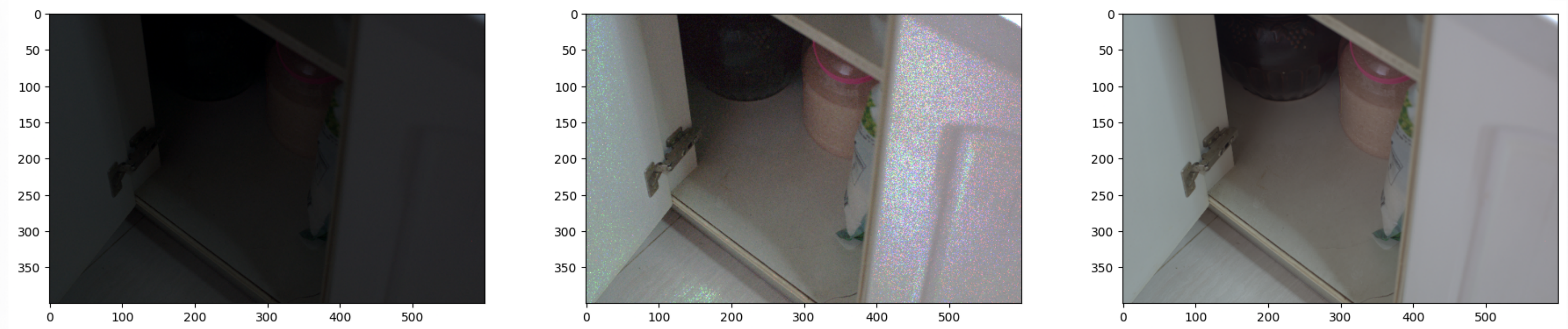
We are using the LOL dataset to train our model. It consists of 500 images having resolution 400 by 600 , out of which 485 are images are used for training and 15 are used for evaluating the model.

Histogram Mapping

Upon doing a thorough analysis of the data, we observed a significant correlation between the low-light image and high-light image pixels. Specifically, we discovered that the pixel values of the low-light and high-light images are closely linked in terms of percentile i.e. the same pixel positions correspond to the same small percentile range in both images. Using this idea, we employ a quantile regression model to perform histogram mapping. We trained an XGBoost regressor model, which utilizes low-light image's histogram as input and generates the pixel values of the high light image at every 5th percentile. We then use these pixel values and a KDE model to reconstruct the histogram of the high-light image pixels. Finally, by applying a histogram transformation on the low-light image to the generated histogram, we are able to reconstruct a high light version of the low light image.



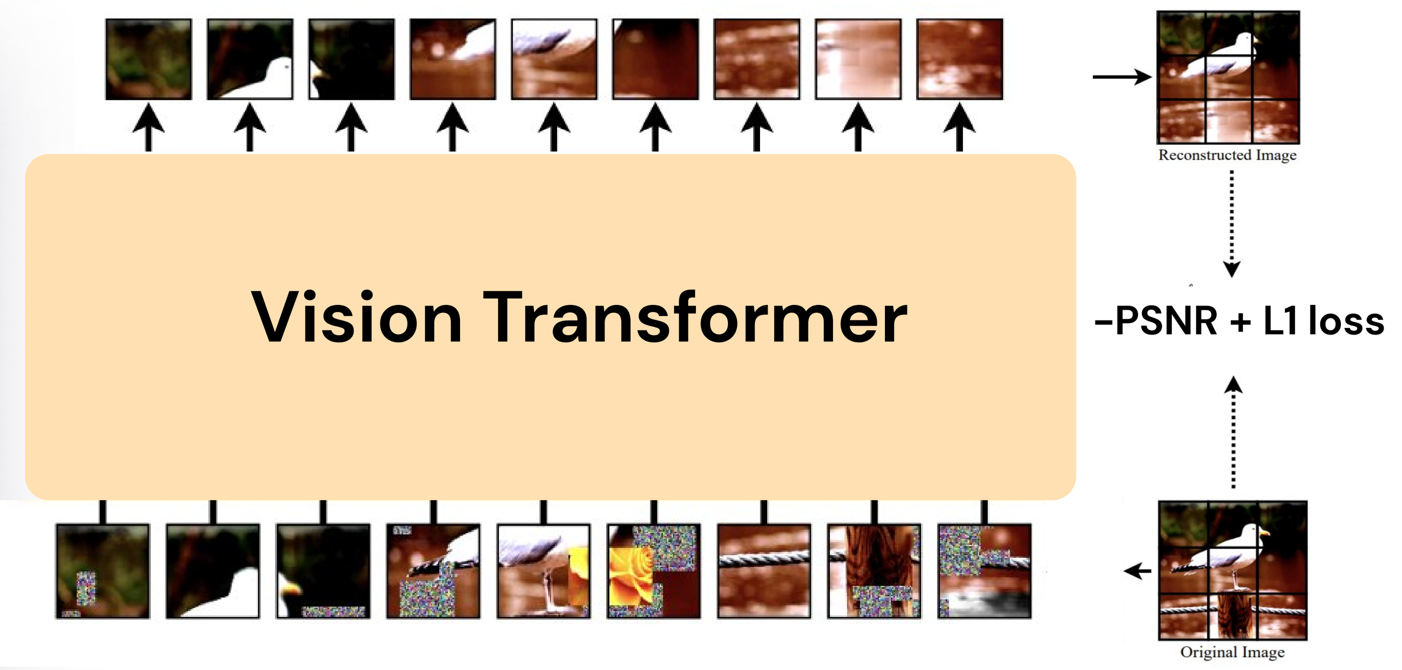
Histograms of [low, predicted, high]



Images of [low, predicted, high] using Histogram Mapping

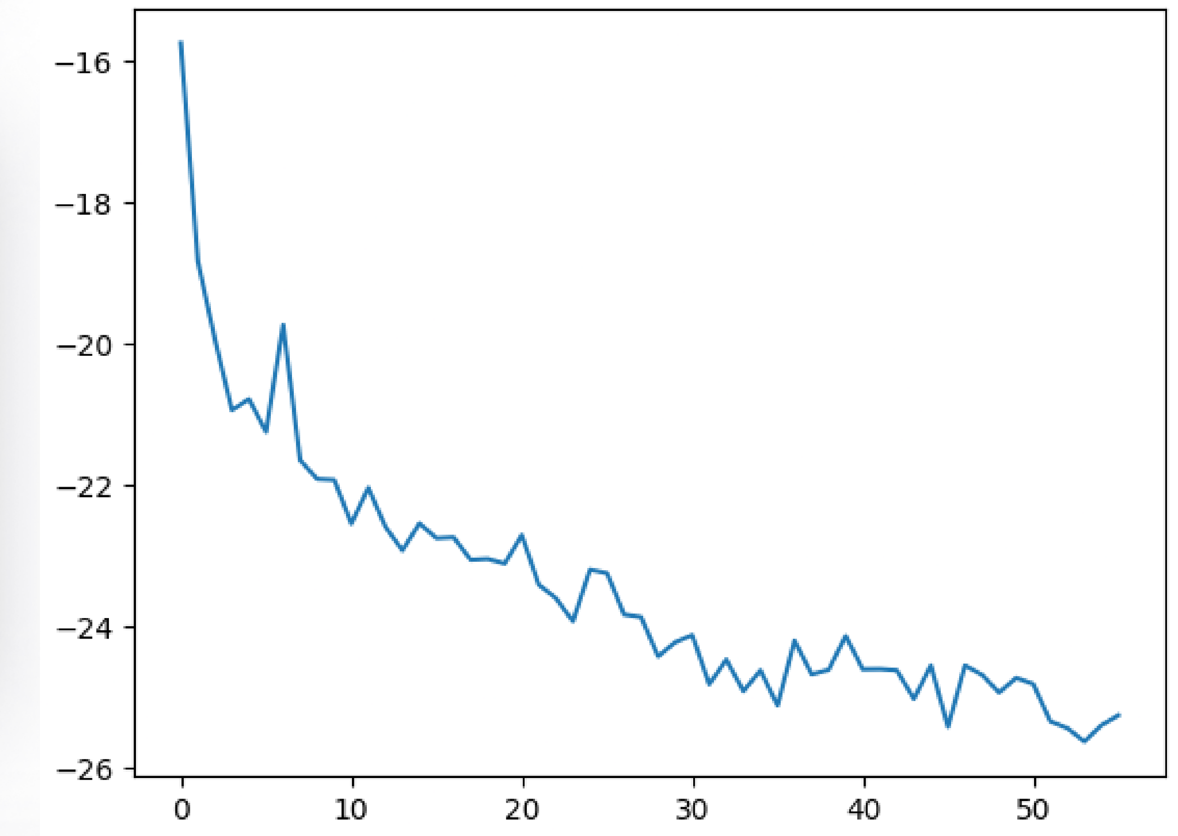
Vision Transformer

The output image after histogram mapping still had amplified noise and colour distortions due to low light settings so we applied a vision transformer which takes in the input images as patches of 10x10 and gives out new patches which after concatenating generates new image.



Training

- For loss function our objective was to minimize -PSNR + L1 Loss .
- We trained for about 150 epochs and achieved a training loss of -28.06
- Fig beside : Training loss v/s epoch for first 50 epochs.



Results

- Mean PSNR on test set obtained after histogram mapping : 19.6
- Mean PSNR on test set obtained after applying vision transformer to histogram mapping : 20.52

Conclusion

Through our experiments we have learnt that preprocessing and histogram mapping is a crucial step before applying models before applying deep learning models as it allows models to learn other features.

IIT ROORKEE