Image Super Resolution using Deep Learning

Shish Pal (IT-243/16)
Saurabh Verma (IT-23/16)
Information Technology
NIT Srinagar

Project Supervisor
Dr. Shabir Ahmad Sofi
Information Technology
NIT Srinagar

Motivation

- Our work is motivated by the advancement in deep learning algorithms for various computer vision problems e.g. face recognition, self-driving car, captcha recognition.
- In some Radar and Sonar imaging applications (e.g. magnetic resonance imaging (MRI), high-resolution computed tomography) Super-resolution can be very useful.
- Converting SDTV content to HDTV
- Daily Surveillance Camera
- Improving the Biometric Authentication System
- Signature Approval (cheque), Earth Remote Sensing, Astronomical Observation

Problem Statement

- Image Super-Resolution is the process of reconstructing a high-resolution image from a low-resolution image.
- An image may have a “lower resolution” due to a smaller spatial resolution (i.e. size) or due to a result of degradation (such as blurring).
- In Computer Vision low-resolution images used for Machine learning and deep learning based solution is the down-sampled image with some blurring and noise added to them.
- *LR = degradation(HR).* Usually degradation function is not known before hand. Directly estimating the inverse degradation function is an ill-posed problem.

Approaches for Image super resolution

- Machine Learning Traditional Models: poor results
- Interpolation based method: results in ringing and jagging effect
- Manifold based approaches: discover mapping using geometry i.e computationally complex & loss of high frequency area.
- Noise resilient image super resolution: lot of info. Loss during denoising.
- Single channel image super resolution: only works with grey scale
- Our Approach: Using Deep Learning based model to convert low resolution image to high resolution image.

Our Approach

- We designed a deep convolutional neural network that learns the end-to-end mapping of low-resolution to high-resolution images.
- We make use of a unified architecture to integrate Image preprocessing, multidimensional mapping and image reconstruction phase for novel end-to-end neural network training for image super resolution.
- An end to end trainable neural network is the one where all parameters of the model are simultaneously trained. There's some data processing system or learning systems that require multiple stages of processing
- what end to end deep learning does is, it can take all those multiple stages and replace it with just a single neural network.

Architecture

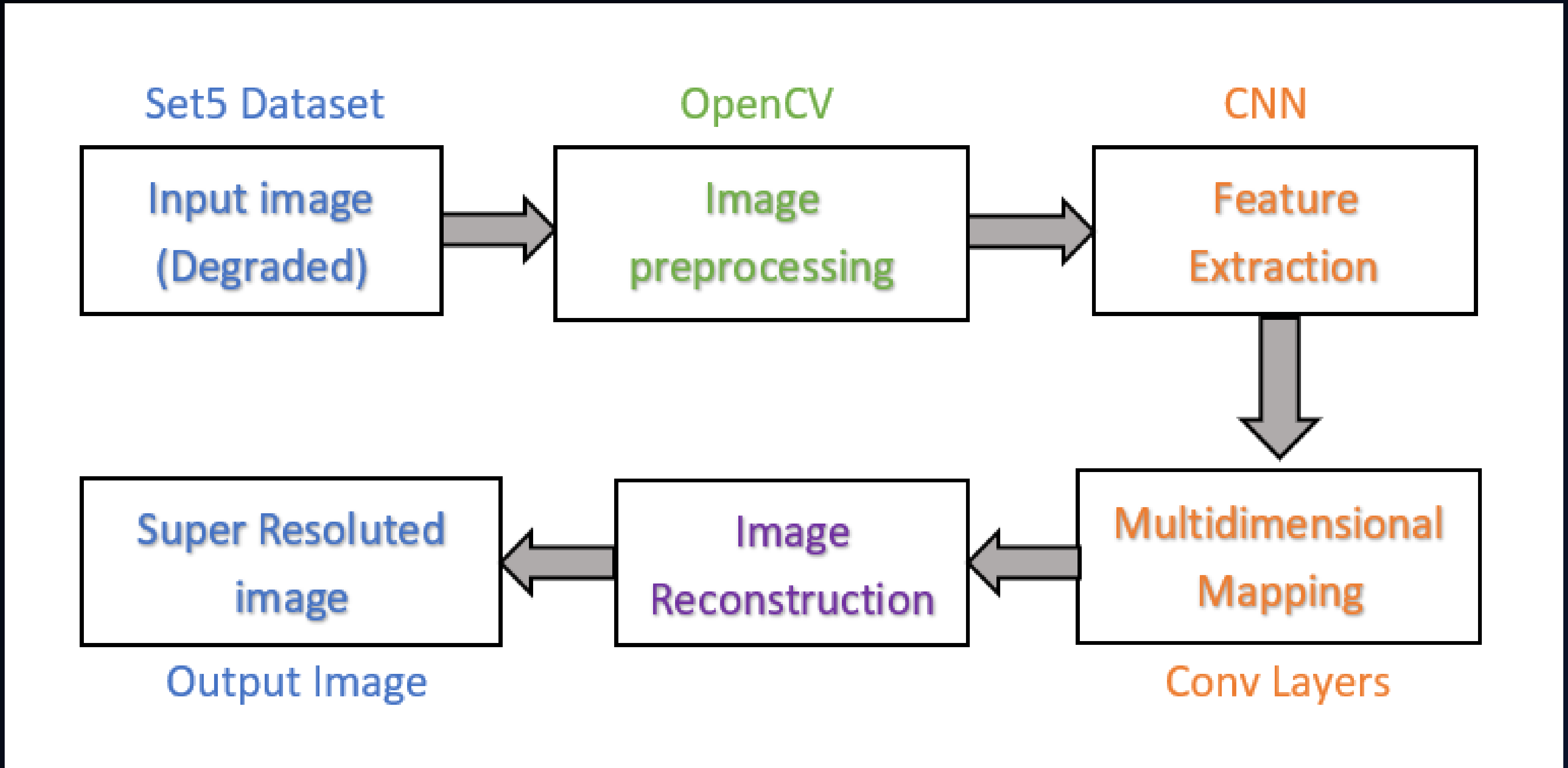


Fig: Neural Network Architecture for Image super Resolution

Input Image

- Our network takes degraded image and converts it to a high resolution image.
- Data Preparation: original high resolution data to low resolution data(degraded).
- First high resolution reduced by a factor(f) of its actual size than up-scaled(zoom) to the original size by Inter Linear interpolation. This gives a degraded image for input.
- To ensure our images are effectively degraded, we calculated MSE,PSNR, and SSIM of degraded image with reference image.

Degraded Image From HR Image

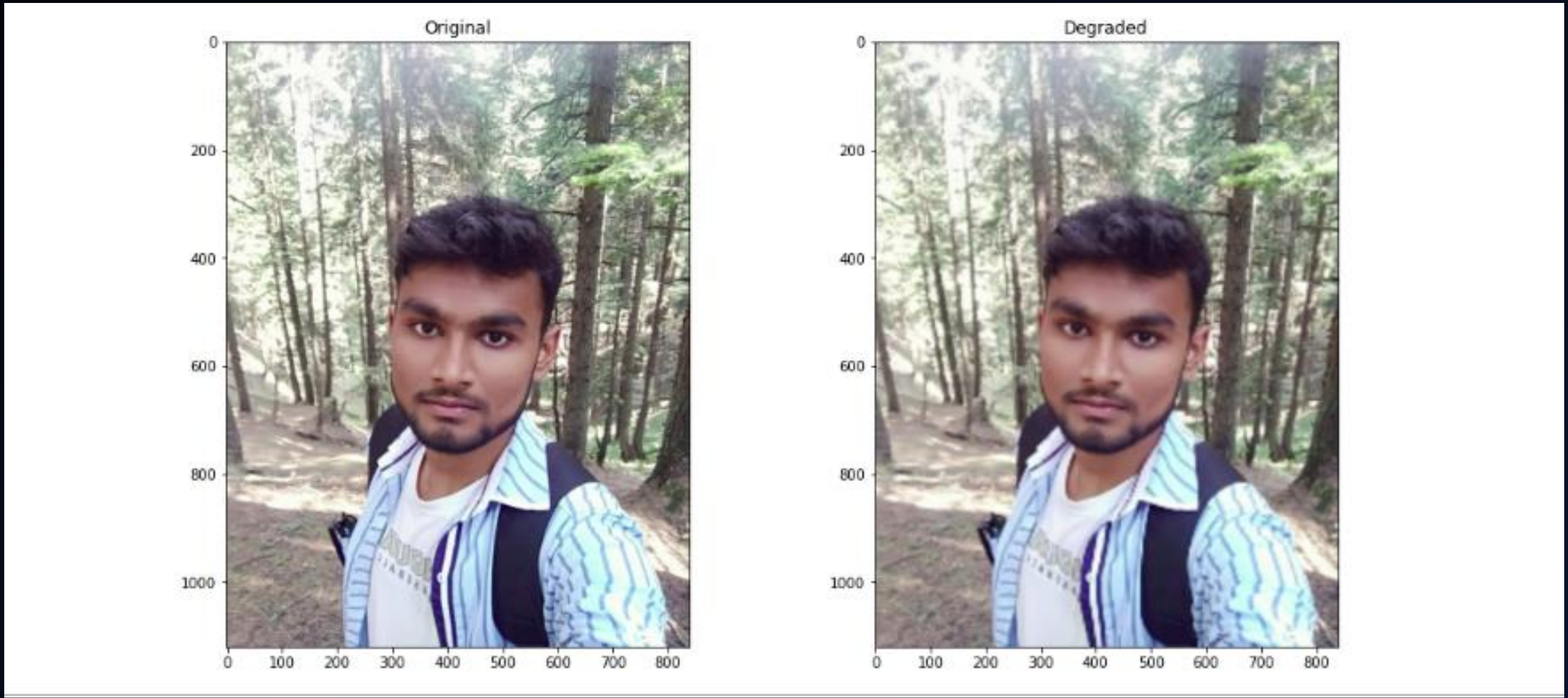


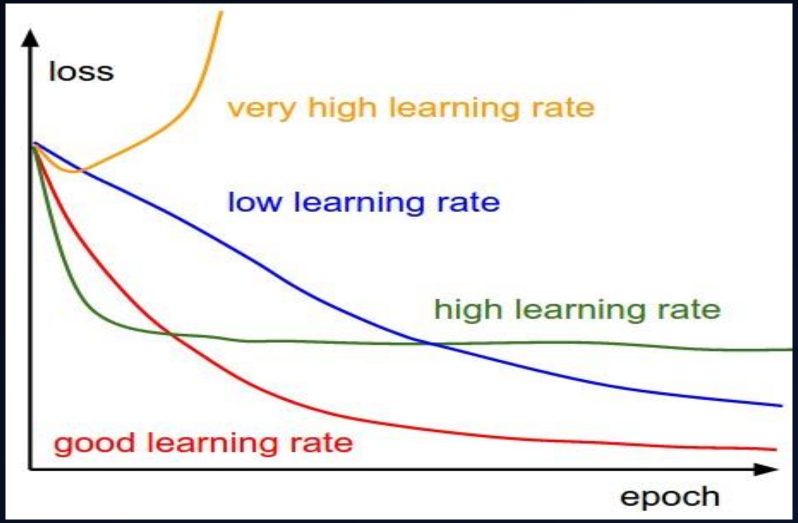
Fig: Blurred output from original Image

Image Preprocessing

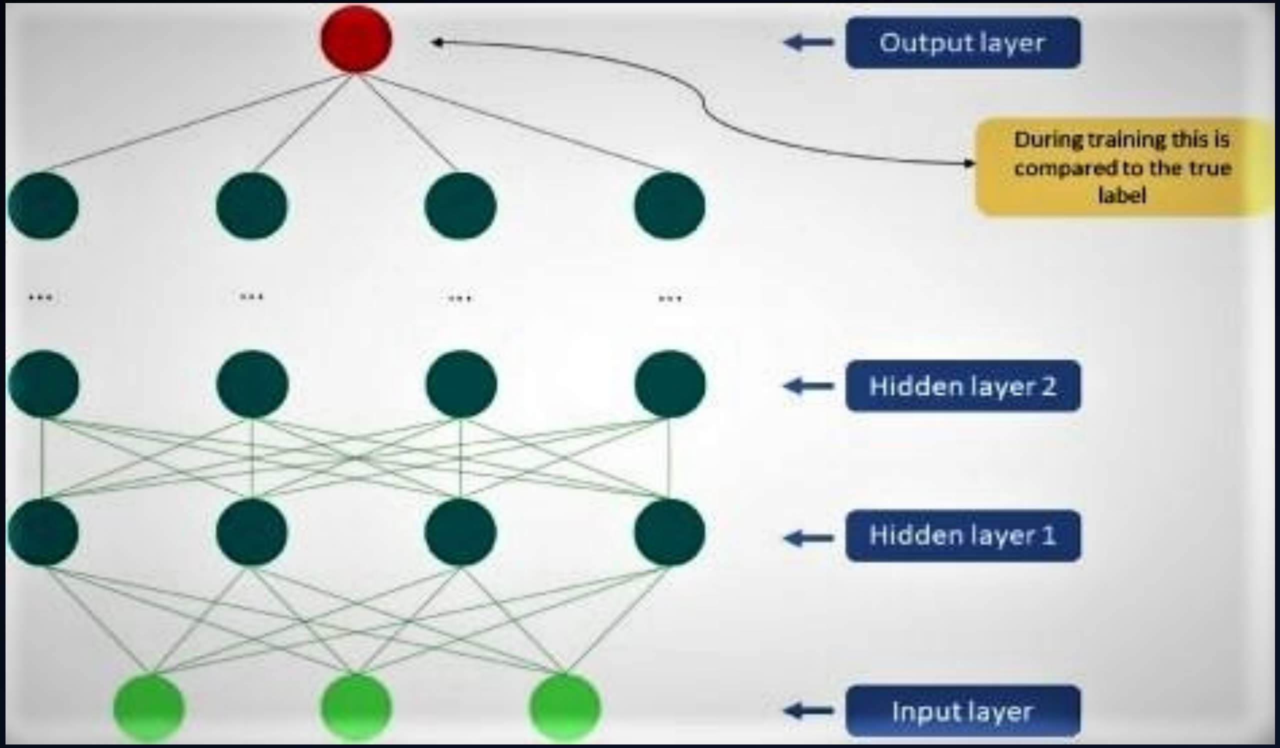
- Input Image is given as input to the image preprocessing step.
- OpenCV works with and generates BGR (Blue, Green, Red) color format images. In our model we are working with YCrCb color space. So, we converted our input images to YCrCb color space.
- The Y'CrCb is quite useful considering the phosphor emission characteristics of newer CRT screens. And is hence used in televisions, HDTV, etc.
- The Y' channel (luma) is basically the grayscale version of the original image. Human eye is more receptive to black and white images, so this channel isn't compressed much. This preserves a lot of details. The Cr and Cb channels contain the colour information. They can be highly compressed.

Feature Extraction

- Used to extract useful features from input image.
- Used CNN based customized model similar to VGG network.
- A batch normalization layer is added after this layer to take care of internal covariate shift problem by normalizing the inputs of each layer.
- Used Adam Optimizer with learning rate of 0.003.



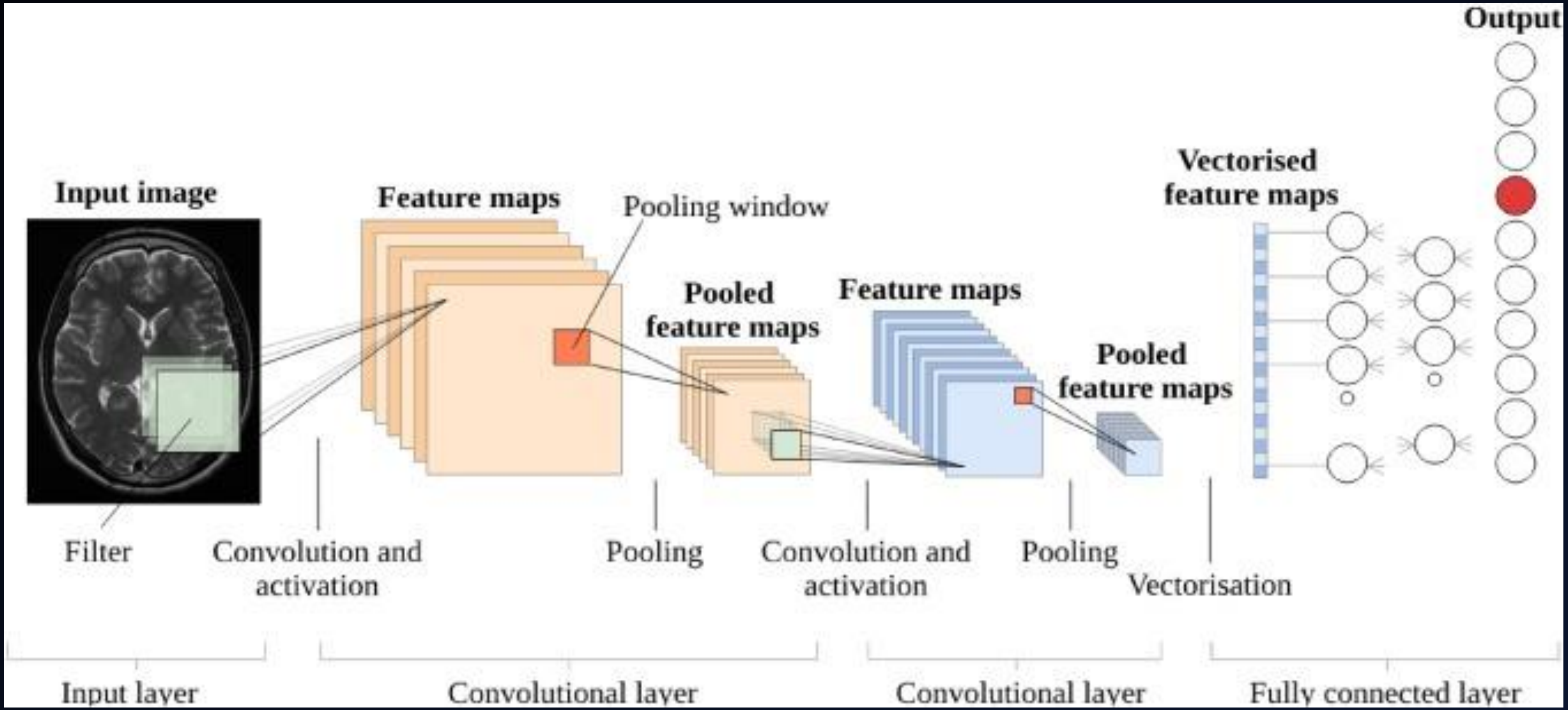
Convolution Neural Network



Source: sciencedirect

Figure: Simple CNN

CNN Architecture



Source: sciencedirect

Multidimensional Mapping

- This layer maps low resolution features to high resolution features.
- Extracted feature from CNN are given as input to this layer.
- We mapped m dimensional vector to n dimensional vector.
- Adding more layer results in slower convergence and better results.
- More deeper network tends to overfitting resulting in degraded results.
- For our model we adopted three layer network.

Image Reconstruction

- Converts high resolution features into a high resolution image.
- In Traditional Algorithm, Features were averaged to produce final image, but we can’t average them if they are in different domain.
- We used convolution layer which projects different domain features into same domain and then convolves them to produce final output image.
- In the end, we again converted output image to BGR color space for comparison.

Output

- Output of our model is a high resolution image of degraded input image.

Original Degraded ISR

Reconstructed image for Human perceptual evaluation

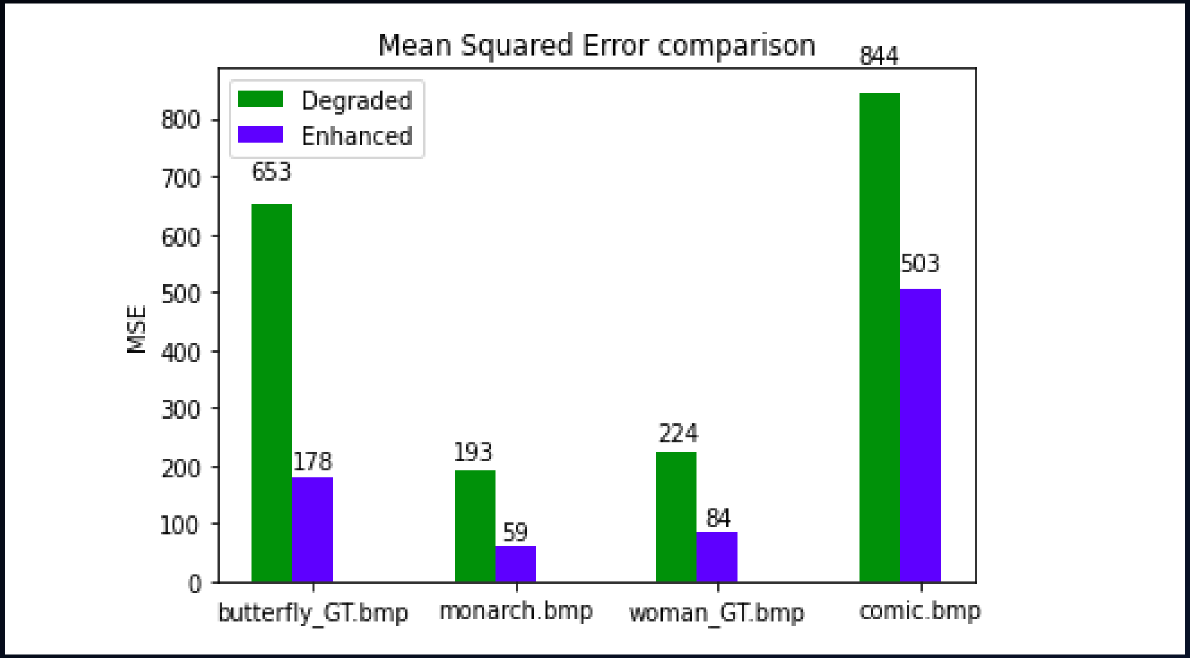
Performance Evaluation:

- A number of Image Quality Assessment (IQA) techniques are used for the same. These metrics can be broadly classified into two categories — Subjective metrics and Objective metrics.
- Subjective metrics are based on the human's (observer) perceptual evaluation where a number or viewers rate their opinion based on their perceptions of image quality.
- whereas objective metrics are based on quantitative evaluation that try to assess the image quality.
- Subjective metrics are often more "perceptually accurate", however these metrics are **inconvenient, expensive or time-consuming** to compute.

Quantitative evaluation metrics:

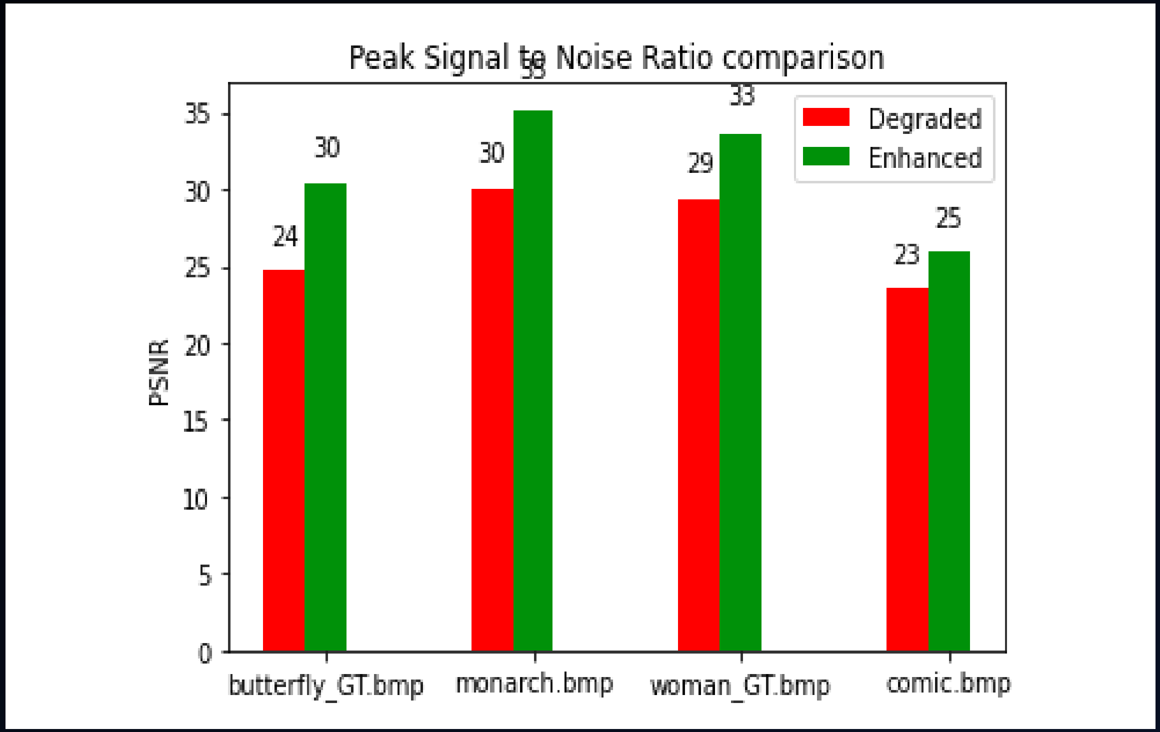
- These are evaluated in **Full Reference method** ie. Comparing images w.r.t to original one.

1. MSE:



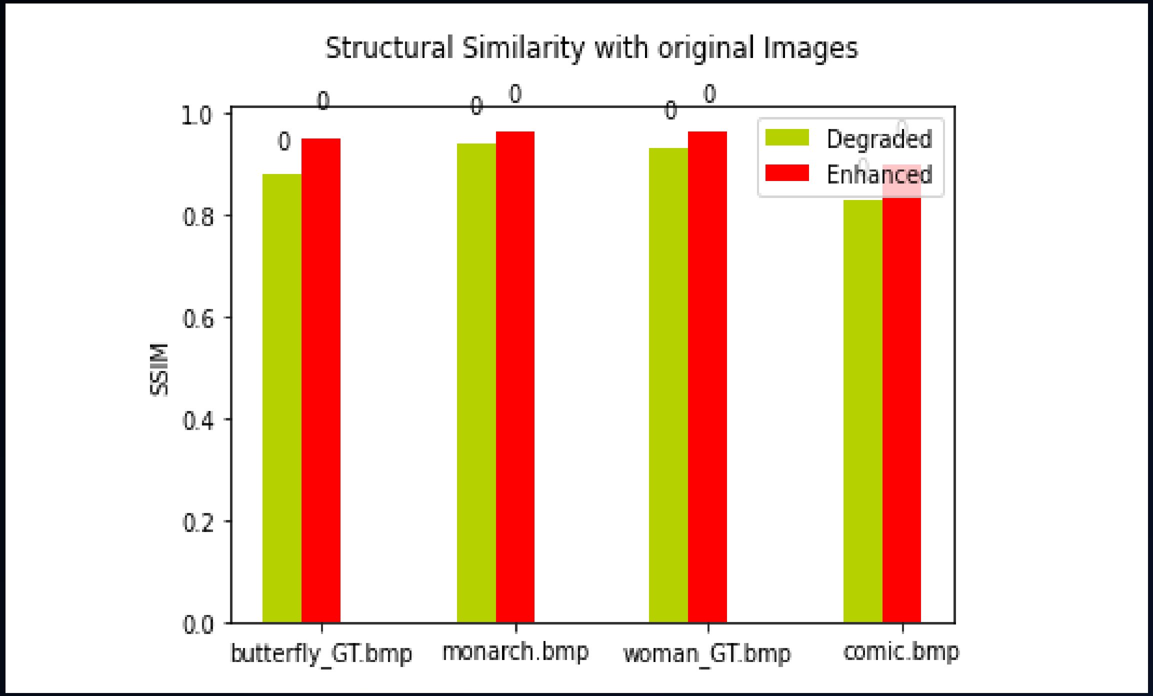
- comparative plot of mean root squared error of degraded input image vs Reconstructed output image w.r.t Original ground truth image.

2. PSNR:



- comparative plot of peak signal to noise ratio of degraded input image vs Reconstructed output image w.r.t Original ground truth image

3. SSIM:



- Although structure of degraded image was not much changed after degradation for 2x- up-sampling but still we can see an improvement in the SSIM value of reconstructed image.

Conclusion

- We used cutting edge end-to-end trainable convolutional neural network-based model, Rather than either training or tuning different components of neural network separately for image super resolution.
- Our model integrates Preprocessing, Feature extraction, Multidimensional mapping and Reconstruction phase into a unified architecture.
- In future we are planning to add more layers to our model which may help in improving accuracy further. To train complex model high computation GPU enabled machine is required.

REFERENCES

- Chao Dong, Chen Change Loy, Kaiming He and Xiaoou Tang, "Learning a deep convolutional network for image super-resolution", *Computer Vision-ECCV 2014*, 2014.
- Ezequiel López-Rubio, "Superresolution from a single noisy image by the median filter transform", *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, 2016.
- H. Weng, J. Lin, G. Lin and Y. Niu, "Multi-scale Fusion for Arbitrary-scale Image Super-resolution," 2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), Xiamen, China, 2019.
- Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: IEEE International Conference on Computer Vision. pp. 349–356 (2009)
- Yang, J., Wang, Z., Lin, Z., Cohen, S., Huang, T.: Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing 21(8), 3467–3478 (2012)
- Hong Chang, Dit-Yan Yeung, and Yimin Xiong, "Super-resolution through neighbor embedding", Computer Vision and Pattern Recognition 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 1, 2004.
- Yang Jianchao, Wright John, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation", Image Processing IEEE Transactions on, vol. 19, no. 11, 2010.

Thank You!

Questions?