DEEP LEARNING PROJECT

# Project Name:

# IMAGE SUPER RESOLUTION USING CONVOLUTION NEURAL NETWORK

GROUP MEMBERS:

YUSHA ARIF (18K-1289)

MUGHEERA BIN SADIQ (18K-0296)

ABDURRAFFAY (18K-1117)

DATE: 1ST JUNE 2021

**ABSTRACT**

We propose a deep learning method for single image super-resolution (SR). Our method directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolutional network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers. Our deep CNN has a lightweight structure, yet demonstrates state-of-the-art restoration quality, and achieves fast speed for practical on-line usage.

1. **INTRODUCTION**

The problem here is the recovering of high-resolution image from a single low-resolution image. The motivation behind the implementation of this project is to help students who can’t afford smartphones with high resolution camera and the image capturing of class boards is an important resource for a student as they can preserve the work for future use. Due to low-resolution image, they can’t use the image in future as they can’t see the work properly so this model will help these students to map there low-resolution image to high resolution image.

1. **Design**

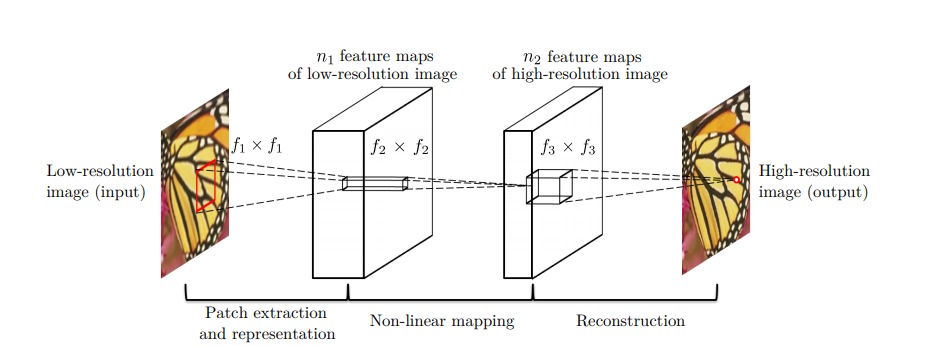
Considering a single low-resolution image, we first upscale it to the desired size using bicubic interpolation, which is the only pre-processing we perform. Let us denote the interpolated image as Y. Our goal is to recover from Y an image F(Y) that is as similar as possible to the ground truth high-resolution image X. For the ease of presentation, we still call Y a “low-resolution” image, although it has the same size as X. We wish to learn a mapping F, which conceptually consists of three operations:

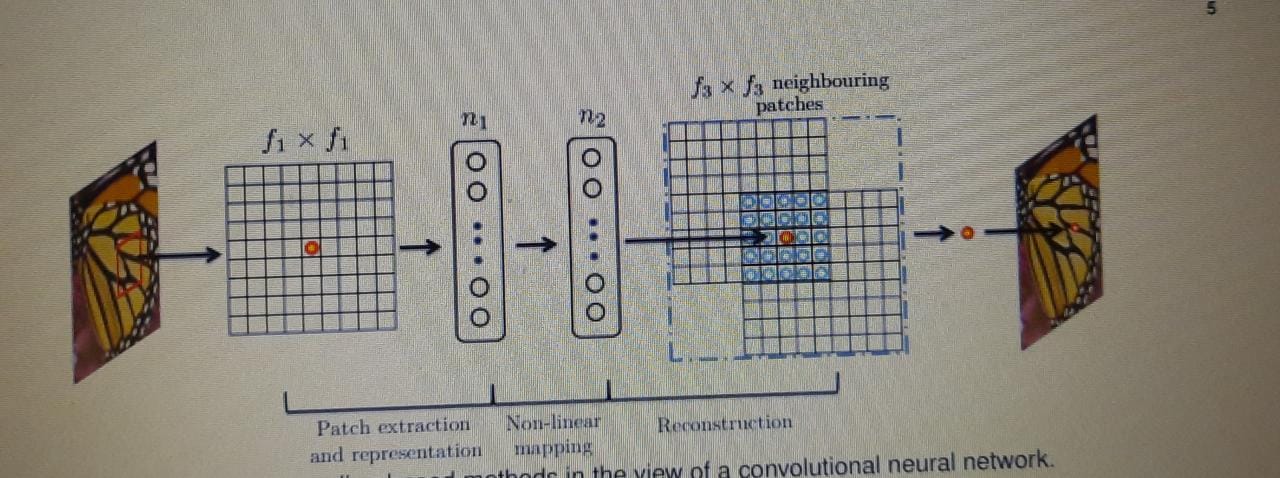
1) Patch extraction and representation: this operation extracts (overlapping) patches from the low-resolution image Y and represents each patch as a high-dimensional vector. These vectors comprise a set of feature maps, of which the number equals to the dimensionality of the vectors.

2) Non-linear mapping: this operation nonlinearly maps each high-dimensional vector onto another high-dimensional vector. Each mapped vector is conceptually the representation of a high-resolution patch. These vectors comprise another set of feature maps. The first layer extracts an n1-dimensional feature for each patch. In the second operation, we map each of these n1-dimensional vectors into an n2-dimensional one. This is equivalent to applying n2 filters which have a trivial spatial support 1 × 1. This interpretation is only valid for 1×1 filters. But it is easy to generalize to larger filters like 3 × 3 or 5 × 5. In that case, the non-linear mapping is not on a patch of the input image; instead, it is on a 3 × 3 or 5 × 5 “patch” of the feature map

3) Reconstruction: this operation aggregates the above high-resolution patch-wise representations to generate the final high-resolution image. This image is expected to be similar to the ground truth X.

The project resides on 3 convolutional network layers. The loss function which we used for training is mean square error and with the help of this loss function we will be able to learn weights of each layer and train our model more accurately so that the predictions in future are more accurate and the loss is also low. The architectural design of the model is as follows:



****

1. **Methodology**

The dataset we used for training was the STL10 dataset. This dataset has the following features:

* 10 classes: airplane, bird, car, cat, deer, dog, horse, monkey, ship, truck.
* Images are 96x96 pixels, color.
* 500 training images (10 pre-defined folds), 800 test images per class.

For testing, we used the Set14 dataset which is a dataset consisting of 14 images commonly used for testing performance of Image Super-Resolution models

The work was divided equally amongst the group members. Yusha was responsible for defining the architecture of the SRCNN model and creating the training module. Abduraffay worked on the evaluation module whereas Mugheera was responsible for extracting the key information from the research papers and setting up the tensorboard for displaying the training results.

1. **Results**

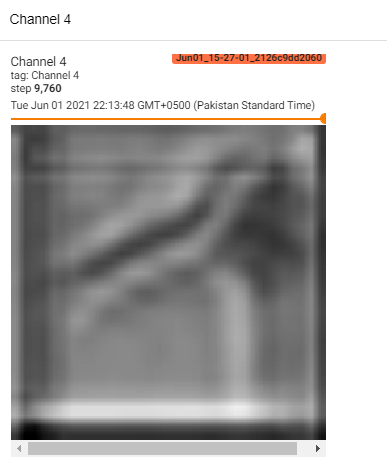
The results of our model are as follows:

**Outputs of Various Channels During Training:**



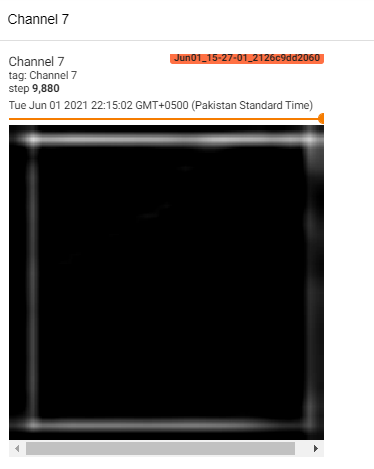


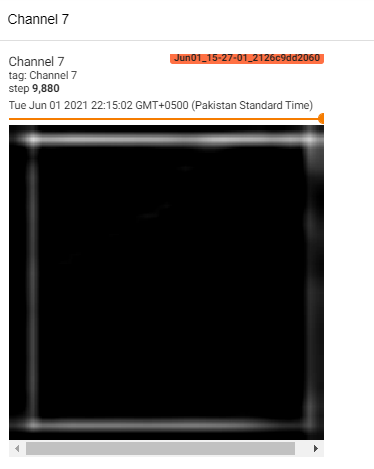


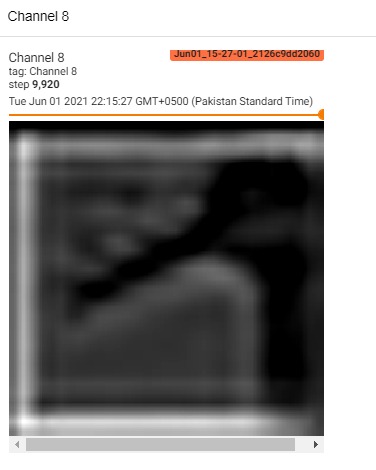


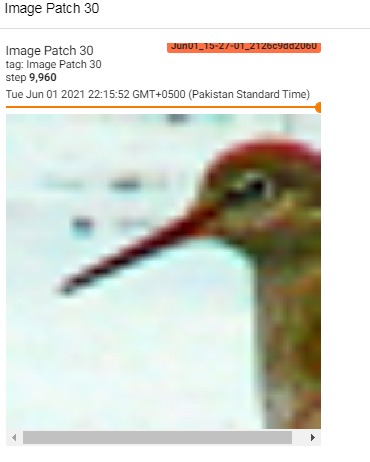


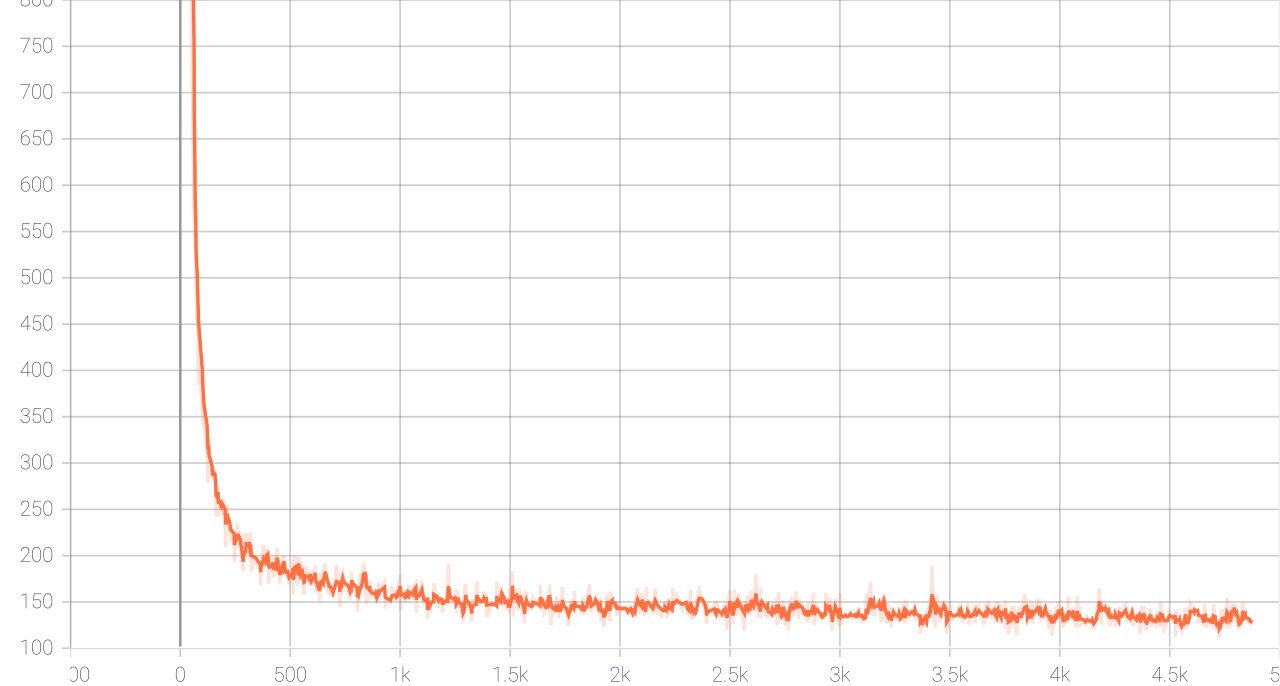




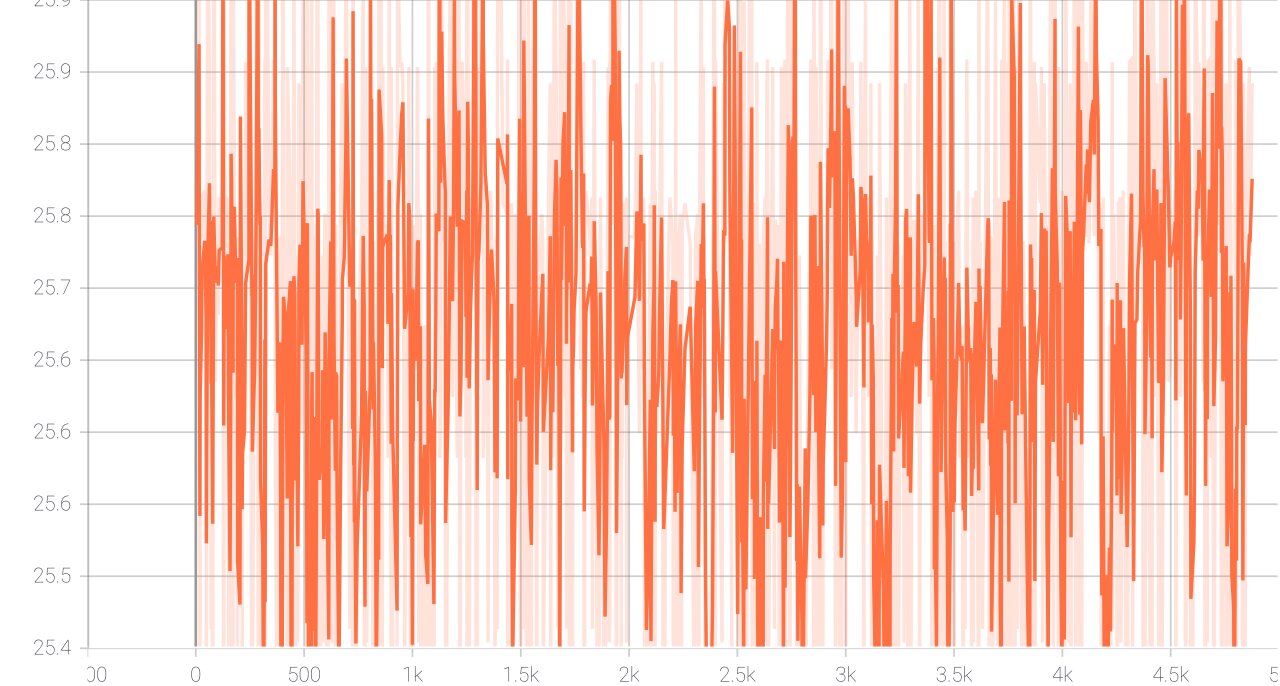




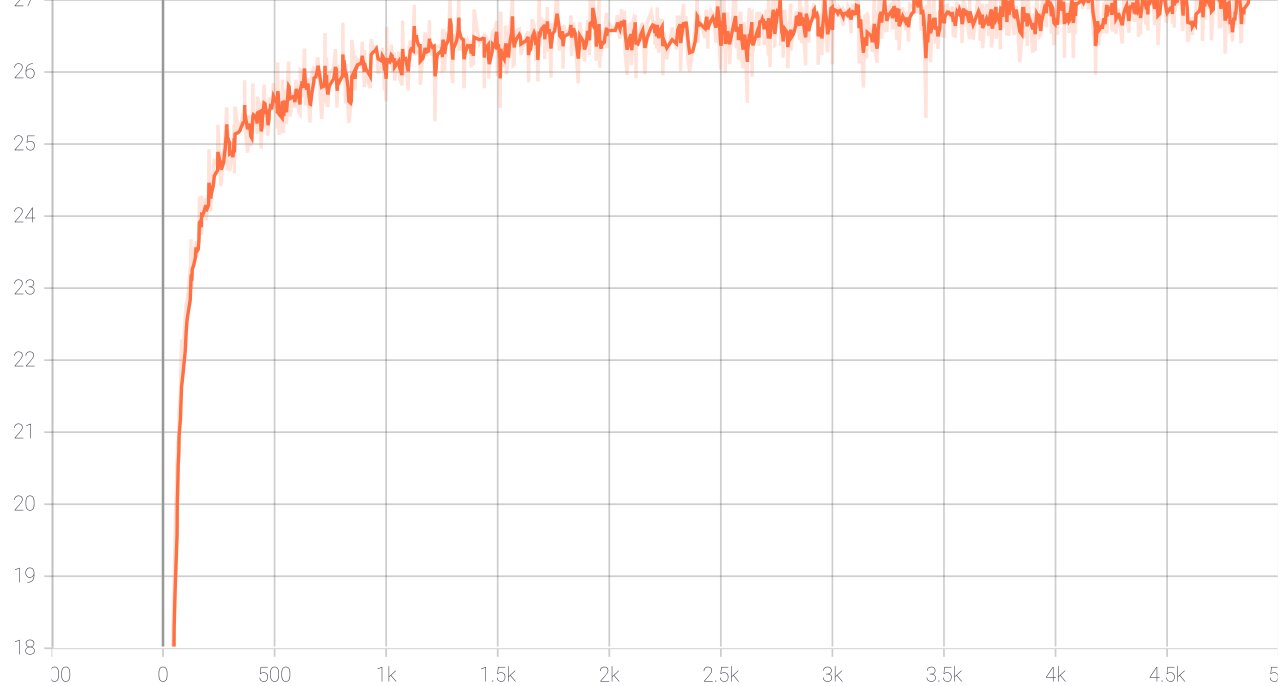




**MSE Loss**



**PSNR of BiCubic Interpolation (For Comparison)**



**PSNR of Reconstruction**

**TEST RESULTS:**

** **

** **

1. **Conclusion and future work**

The accuracy of the model will improve more when it will be trained on much larger dataset and much greater values of epoch size. In addition to these measures the performance of this model can also improve by exploring more filters and different training strategies. The model can be extended to use in real life applications like installing this model in the surveillance cameras used in traffic signals to help the traffic police in monitoring the vehicle speeds and recognizing the vehicles which are over-speeding and due to this over-speeding, the images captured by the cameras are not good enough to recognize so these images can be mapped to high resolution images.

1. **Reference**

* Chao Dong, Chen Change Loy, Member, IEEE, Kaiming He, Member, IEEE, and Xiaoou Tang, Fellow, IEEE *“Image Super-Resolution Using Deep Convolutional Networks”* (2015)
* Chang, H., Yeung, D.Y., Xiong, Y*.: Super-resolution through neighbor embedding*. In: IEEE Conference on Computer Vision and Pattern Recognition (2004)
* Dai, D., Timofte, R., Van Gool, L.: *Jointly optimized regressors for image super-resolution. In: Eurographics*. vol. 7, p. 8 (2015)
* Dong, C., Loy, C.C., He, K., Tang, X.: *Learning a deep convolutional network for image super-resolution. In: European Conference on Computer Vision*, pp. 184–199 (2014)