CV Project Report

Image Super-Resolution

Team 26

D. R. Pailla, K. Gupta, T. Surkar

Introduction

Team

1. Dheeraj Reddy Pailla
2. Kshitij Gupta
3. Trunapushpa Surkar

About the project

The purpose of the project is to...

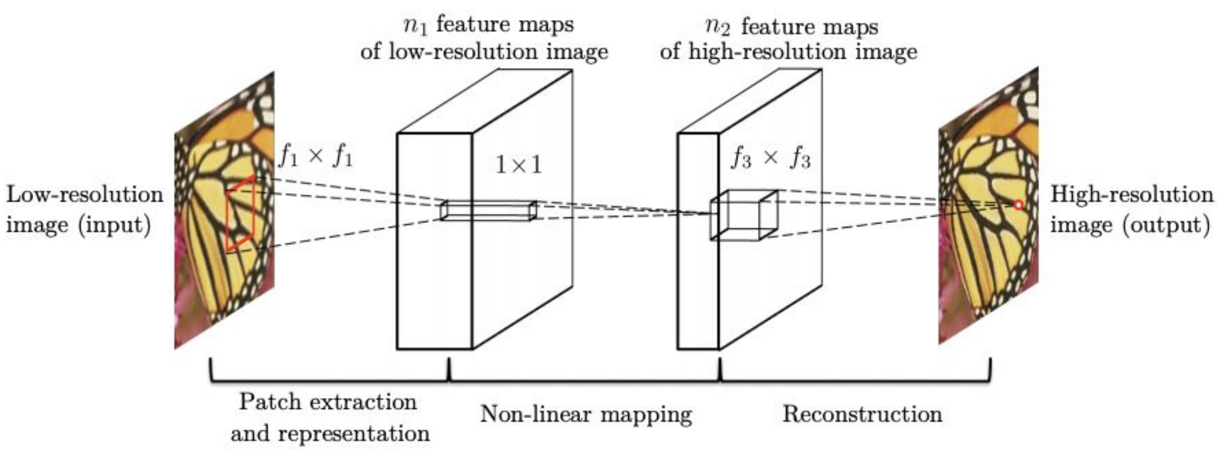
Reference Paper

Learning a Deep Convolutional Network for Image Super-Resolution - Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang [ECCV 2014] available at [this link].

The main themes of the paper are the following:-

1. CNN for image super-resolution
2. Establish a relationship between the deep-learning-based super-resolution method and the traditional sparse-coding-based super-resolution methods. This guides the design of the network structure.

Method



In the above image, f1 = 9, f3 = 5, n1 = 64, n2 = 32.

The system involves three main steps. The first step involves patch extraction and representation. In the above image, f1=9, i.e., 9x9 filters are used with an overlapping stride of 1 over the original input image. We obtain 64 features from each 9x9 match. We use zero padding=4 in this step to maintain the dimensionality of the input image.

The second step involves a non-linear mapping such that the 64 features obtained from the first step are now mapped to a set of 32 features. The final step is the reconstruction phase where the set of 32 features are mapped to the total number of channels in the input image with the usage of 5x5 filters, which is represented as f3xf3 in the above image. We use zero padding=2 in this step to maintain the dimensionality of the input image.

We have also used perceptual loss (output of 7th layer of VGG16 pretrained on ImageNet). The final PSNR values are lesser, but qualitatively the results look better.

We have also tried using gram matrix of the 7th layer output of VGG16 to use as the perceptual loss.

Dataset

2 datasets were used in this project. The T91 database, which consists of 91 images is transformed into a dataset of 22,000 images for this project using overlapping 32x32 windows with stride 14. Furthermore, the 300 images from BSD 300 dataset are also used.

Results

	Input Image	Target Image	Network Output	PSNR value
T91 $f_1 = 9$ $f_3 = 5$ $n_1 = 64$ $n_2 = 32$ MSE Loss Y channel 50 epochs				Avg. PSNR: 28.1654 dB Bicubic Avg. PSNR: 26.4436 dB

T91 $f_1 = 9$ $f_3 = 5$ $n_1 = 64$ $n_2 = 32$ MSE Loss RGB channel 50 epochs				Avg. PSNR: 27.8857 dB Bicubic Avg. PSNR: 26.0454 dB
T91 $f_1 = 9$ $f_3 = 5$ $n_1 = 64$ $n_2 = 32$ Perceptual Loss Y channel 50 epochs				Avg. PSNR: 27.5520 dB Bicubic Avg. PSNR: 26.4436 dB
T91 $f_1 = 9$ $f_3 = 5$ $n_1 = 64$ $n_2 = 32$ Perceptual Loss RGB channel 50 epochs				Avg. PSNR: 24.4652 dB Bicubic Avg. PSNR: 26.0454 dB
T91 $f_1 = 9$ $f_3 = 5$ $n_1 = 64$ $n_2 = 32$ Gram Matrix Perceptual Loss Y channel 50 epochs				Avg. PSNR: 26.7261 dB Bicubic Avg. PSNR: 26.4436 dB

Future Work

A comparative study between super-resolution from a single image and current method implemented.

Codebase

GitHub repo - https://github.com/dheerajpreddy/image-super-resolution