

# Single Image Super Resolution: An In-Depth Analysis

Ammar bin Ayaz (aba450), Umar Farooq (uf247)

## Introduction

Image super-resolution (SR), which refers to the process of recovering high-resolution (HR) images from low-resolution (LR) images, is an important class of image processing techniques in computer vision. In general, this problem is very challenging and inherently ill-posed since there are always multiple HR images for a single LR image but with the rapid development of deep learning techniques, Super-resolution models based on Deep learning have been extensively explored and they often achieve state-of-the-art performance on different benchmarks of Super-Resolution.

## Objectives

- A deeper understanding of the topic of Super-Resolution and architectural analysis of some of the state-of-the-art techniques of Super-Resolution such as **SRGAN**, **SRResNet**, **SRCNN**, **ESRGAN**.
- Comparing the performance of above-mentioned techniques on benchmarks like **Set5**, **Set14** and **BDS100**.

## Data sets

- **Training:**
  - T-91 image dataset – A dataset of 91 images
  - DIV-2k dataset - The image quality is of 2K resolution and is composed of 800 images for training while 100 images each for testing and validation

- **Evaluation:**

Benchmark datasets to evaluate the performance of different models.

- **Set5** - Five test images of a baby, bird, butterfly, head, and a woman.
- **Set14** - Consists of more categories as compared to **set5** i.e 14 categories.
- **BDS100** - The dataset has 100 images ranging from natural images to object-specific such as plants, people, food etc.

# Models

- **SRCNN**: A 3-layer 9-1-5 deep convolution neural network with filter width  $n_1 = 64$  and  $n_2 = 32$
- **SRGAN/SRResNet**: SRGAN combines the concept of ResNet and GAN to train the model, in SRResNet we use a no GAN network consisting only of the residual blocks and use MSE as loss.
- **ESRGAN**: Two modifications made in the above mentioned SRGAN model.
  - 1) Removing all the batch normalization layers
  - 2) Replacing the basic block with Residual in Residual dense block (RRDB)

**Input**: Bicubic interpolated output of the low-resolution image (scale 2, 4)

**Output**: A high resolution up-scaled image

**Evaluation metric**: Peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM)

**Implementation**:

- **Experiment 1:**

Limited training of the models (mentioned in the references) with input datasets and evaluating on the test sets.

- **Experiment 2:**

Using pre-trained weights of the same models to evaluate the test datasets for better inference of the results. This approach is used to see the actual performance of the models mentioned in the project, as we can't train the whole network proposed in the papers because of computational limitations.

## Anticipated results:

We expect **ESRGAN** and **SRGAN/SRResNet** to perform better than the simpler **SRCNN** model, but the prior models are complex with deep structures resulting in extensive training.

# References

## Paper:

- A Deep Journey into Super-resolution: A Survey - <https://arxiv.org/abs/1904.07523>
- Photo-Realistic Image Super-Resolution Using GAN - <https://arxiv.org/abs/1609.04802>
- Image SuperResolution Using Convolutional Networks-<https://arxiv.org/abs/1501.00092>
- ESRGAN : Enhanced Super-Resolution Generative Adversarial Networks - <https://arxiv.org/abs/1809.00219>

## Code:

- SCRNN Implementation - <https://github.com/yjn870/SRCNN-pytorch>
- ESRGAN : <https://github.com/xinntao/ESRGAN>
- SRResNet : <https://github.com/PacktPublishing/Generative-Adversarial-Networks-Projects>