Super-Resolution using Convolutional Neural Network

Varun Varma Thozhiyoor
Ashik Raveendran

November 17, 2024

Table of Contents

Problem Statement

Image super-resolution has gained a lot of interest recently because of its applications in a variety of disciplines. Super-resolution tries to improve image quality and resolution, resulting in greater clarity and detail. Traditional methods often fall short in producing high-quality outcomes. Convolutional Neural Networks (CNNs) are a potential alternative that can learn complex information and produce clearer images. The project will explore CNN-based super-resolution approaches, comparing them to traditional methods and examining factors that influence performance across various image types.

Introduction

- **Super Resolution (SR)** refers to the process of enhancing the resolution of an image, typically from a low-resolution (LR) image to a high-resolution (HR) version.
- **Traditional SR Methods:**:
  - Relied on interpolation techniques such as bilinear, bicubic, etc.
  - Limited in handling complex image details and textures.
- **CNN-based Super Resolution:**
  - Leverages deep learning models to learn mapping from low-resolution to high-resolution images.
  - CNNs capture complex patterns, textures, and fine details more effectively than traditional methods.
- **Key Advantages:**
  - Produces high-quality results
  - Capable of reconstructing finer details from low-resolution inputs

Bicubic Interpolation

- **Bicubic interpolation** is one of the traditional methods for super resolution, which involves the sixteen nearest neighbors of a point. The intensity value assigned to point (x, y) is obtained using the equation

$$v(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j$$

  The sixteen coefficients are determined from the sixteen equations with sixteen unknowns that can be written using the sixteen nearest neighbors of point (x, y) .

- Bicubic interpolation causes blurring, leading to a loss of sharpness in fine details and edges. Deep learning-based super-resolution can mitigate these issues.

Model Architecture

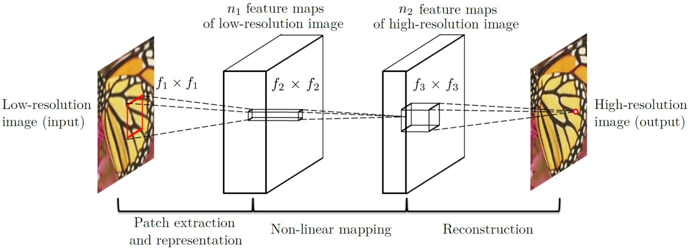

Figure: SRCNN model architecture[1]

Model: "sequential_4"

Layer (type)	Output Shape	Param #
conv2d_12 (Conv2D)	(None, 25, 25, 64)	5,248
activation_8 (Activation)	(None, 25, 25, 64)	0
conv2d_13 (Conv2D)	(None, 25, 25, 32)	2,080
activation_9 (Activation)	(None, 25, 25, 32)	0
conv2d_14 (Conv2D)	(None, 21, 21, 1)	801

Figure: Model info

VGG16 Perceptual Loss

- In our project we are using **VGG16 Perceptual loss[2]** as evaluation metric for generated images. It compares the perceptual similarity between images rather than pixel-by-pixel accuracy.
- **Per-pixel loss functions?**
  - Comparing two images based on their individual pixel values. So, if two images, that are perceptually the same, but different from each other based on even one pixel, then based on per-pixel loss functions they will be very different from each other. (Eg: PSNR, SSIM, etc)
- **Perceptual loss functions?**
  - Comparing two images based on high-level representations from pretrained Convolutional Neural Networks.

Results

(a) Original high-resolution image,

(b) Bicubic upscaling of the low-resolution image.

(c) SRCNN enhanced high-resolution output

Figure: Results

DFT of Bicubic and SRCNN Images

Figure: Bicubic image

Figure: SRCNN image

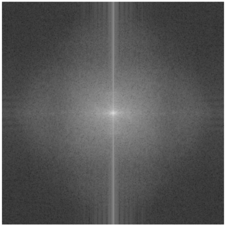

Figure: DFT of Bicubic image

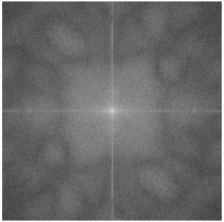

Figure: DFT of SRCNN image

Results

- **PSNR and SSIM Metrics**
  - For the Bicubic interpolated image:
    - PSNR: 35.76 dB
    - SSIM: 0.9515
  - For the SRCNN-enhanced image:
    - PSNR: 25.48 dB
    - SSIM: 0.8171

The predicted image shows lower PSNR and SSIM values due to slight intensity variations. To address this, we are using **VGG16 perceptual loss**.

- **VGG16 perceptual loss**
  - For Bicubic : 14.275507926940918
  - For SRCNN : 1.213773250579834

Results

- **Good Performance on Smoothened Images:** Here model 2 was trained using a dataset containing blurry images (Gaussian blur).

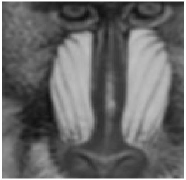

Figure: Blurred

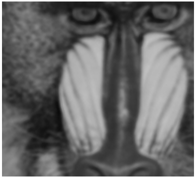

Figure: Model 1

Figure: Model 2

- **Not Robust to Noise:** Here model was trained using a dataset containing noisy images (Gaussian noise).

Figure: Noisy

Figure: SRCNN

References

[1] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2015.
https://doi.org/10.1109/tpami.2015.2439281

[2] Johnson, J., Alahi, A., Fei-Fei, L. (2016c). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. arXiv (Cornell University).
https://doi.org/10.48550/arxiv.1603.08155