DLP Project Report
Image Super Resolution

Prepared By:
Muhammad Rehan Khan 21K-3172
Syed Muhammad Shayan Hussain 21K-3386

Instructor: Miss Sumaiyah Zahid

Objective

The primary goal of this project is to enhance low-resolution images by reconstructing high-quality versions using deep learning techniques. By implementing and fine-tuning various super-resolution models, this project aims to restore fine details and reduce noise, making the solutions applicable to domains such as medical imaging, satellite data analysis, security surveillance, and digital photo restoration.

Problem Statement

Real-world low-resolution images often suffer from information loss due to downscaling, compression artifacts, and noise. Traditional interpolation methods (like bicubic) fail to recover high-frequency details. The challenge is to develop and compare learning-based methods that can infer fine textures, maintain structural integrity, and generalize well to unknown degradations.

Methodology

Dataset and Preprocessing

- Dataset: ILSVRC2013 Validation Data from ImageNet.
- Preprocessing: Images downscaled by a factor of 3× using bicubic interpolation.

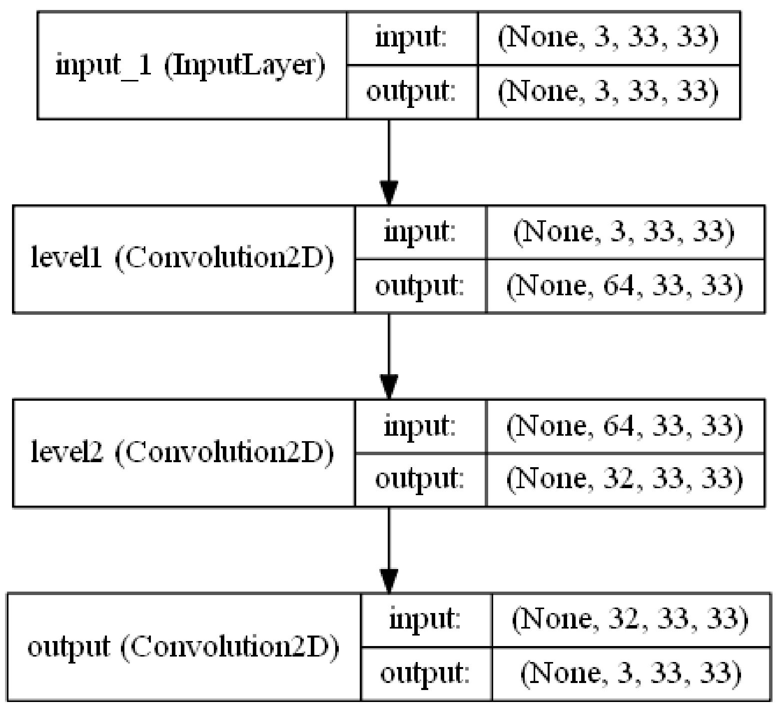
All models received LR-HR pairs for supervised learning (except Real-ESRGAN used only for inference).

Hyper Parameters

- Epochs: 20
- Batch Size: 32
- Learning Rate: 1e-4

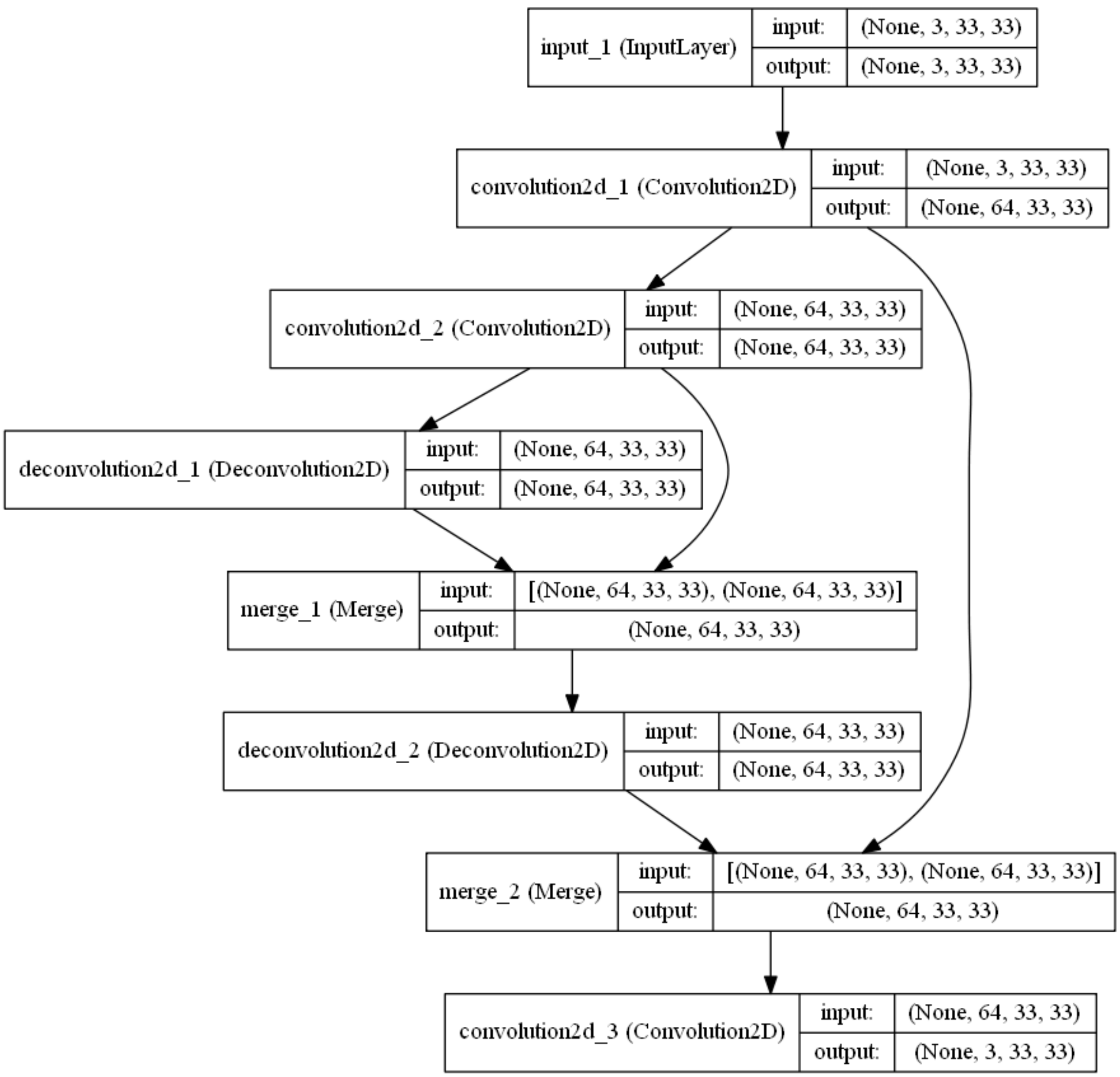
CNN-based Super Resolution (SRCNN 9-5-5)

- Based on the architecture from Dong et al. [SRCNN, arXiv:1501.00092
- A three-layer CNN with 9×9, 5×5, and 5×5 convolutional filters
- Trained from scratch for 20 epochs on a dataset of over 20,000 images (downscaled by 3x
- Batch size: 32, Optimizer: Adam, Loss: MSE
- Goal: Minimize pixel-wise differences to achieve high PSNR.



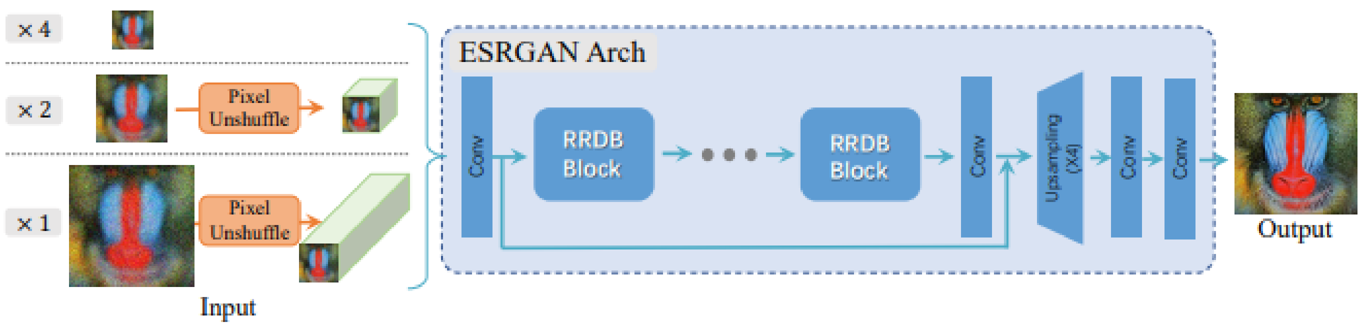
Denoising Super Resolution CNN (DSRCNN)

- Extended the base SRCNN by incorporating a denoising autoencoder module.
- Trained under similar conditions as SRCNN to allow fair performance comparison.
- This model focuses on reducing noise artifacts while enhancing resolution.



GAN-based Super Resolution (Real-ESRGAN)

- Based on Real-ESRGAN [Wang et al., arXiv:2107.10833].
- Utilized a pre-trained model with no fine-tuning due to already excellent perceptual results.
- Used for qualitative comparison to assess visual fidelity beyond PSNR.



Results

PSNR is used for evaluation, following the original papers. It's a standard metric for super-resolution and allows consistent comparison across models. PSNR (Peak Signal-to-Noise Ratio) measures how close the reconstructed super-resolved (SR) image is to the original high-resolution (HR) image. The testing set was 10 percent of the training set; 2000+ images were used for testing.

SRCNN

- Average PSNR: 31.42 dB

DSRCNN

- Average PSNR: 31.58 dB

GAN

- Average PSNR: ~30 dB (Note: PSNR should not be used for GANs because it penalizes perceptual enhancements—like added texture or detail—that improve visual quality but differ from the ground truth at the pixel level.)

Reference PSNR Scores from Papers

- SRCNN (9-5-5) – PSNR: 32.39 dB (scale ×3)【paper: 1501.00092v3】

References

1. Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image Super-Resolution Using Deep Convolutional Networks. *IEEE TPAMI*. arXiv:1501.00092
2. Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. arXiv:2107.10833v2
3. Russakovsky, O., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *IJCV*. ILSVRC 2013 Dataset
4. Image Super Resolution Github
5. ESR-GAN Github