Optimizing CNN-based Architectures for Image Super-Resolution: Depth, Width, Optimized Upsampling Layers, and GAN-Based Perceptual Quality Enhancements

S.M. Faiaz Mursalin, Igor Tisma

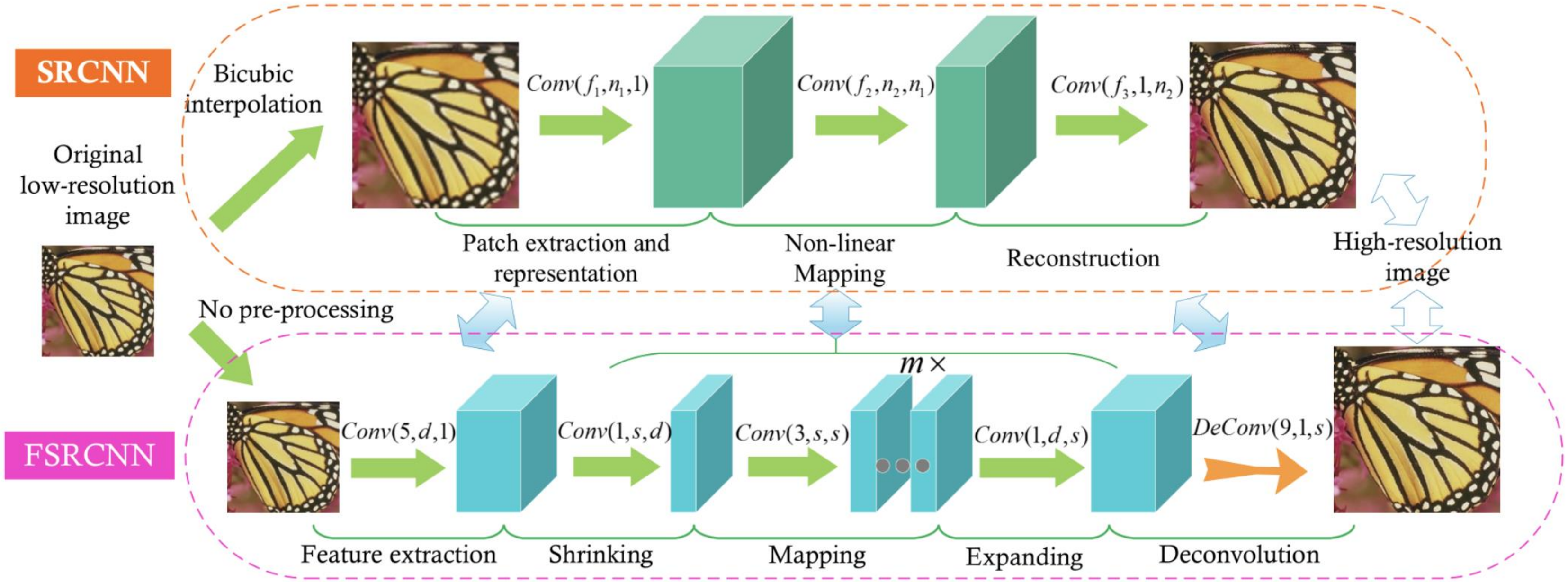
October 30, 2024

Missouri State
UNIVERSITY

Background: Single-Image Super-Resolution

- Enhances the resolution of a single low-resolution (LR) image to produce a high-resolution (HR) output.
- Applications are in Medical imaging, satellite imagery, security, video processing, Restoration of old photographs or low-quality videos
- Challenges:
  - Computational efficiency: balance between model complexity, performance, and processing time
  - Balancing between sharpness and artifact minimization
- Techniques:
  - Interpolation-based: Nearest neighbor, bilinear, bicubic
  - Learning-based: Convolutional Neural Networks (CNNs), GAN
- Popular Models
  - SRCNN: Early CNN-based SISR method, simple and effective
  - ESRGAN: Enhanced SR with GANs for high-quality perceptual improvement
  - EDSR: State-of-the-art with deep residual blocks, designed for high fidelity
- Performance Metrics:
  - PSNR (Peak Signal-to-Noise Ratio): Measures similarity to original HR image
  - SSIM (Structural Similarity Index): Assesses perceived visual quality

Related Work: FSRCNN - Fast Super-Resolution Convolutional Neural Network

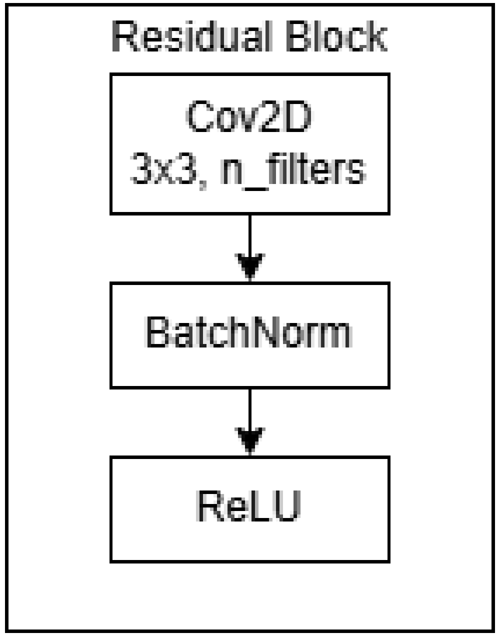


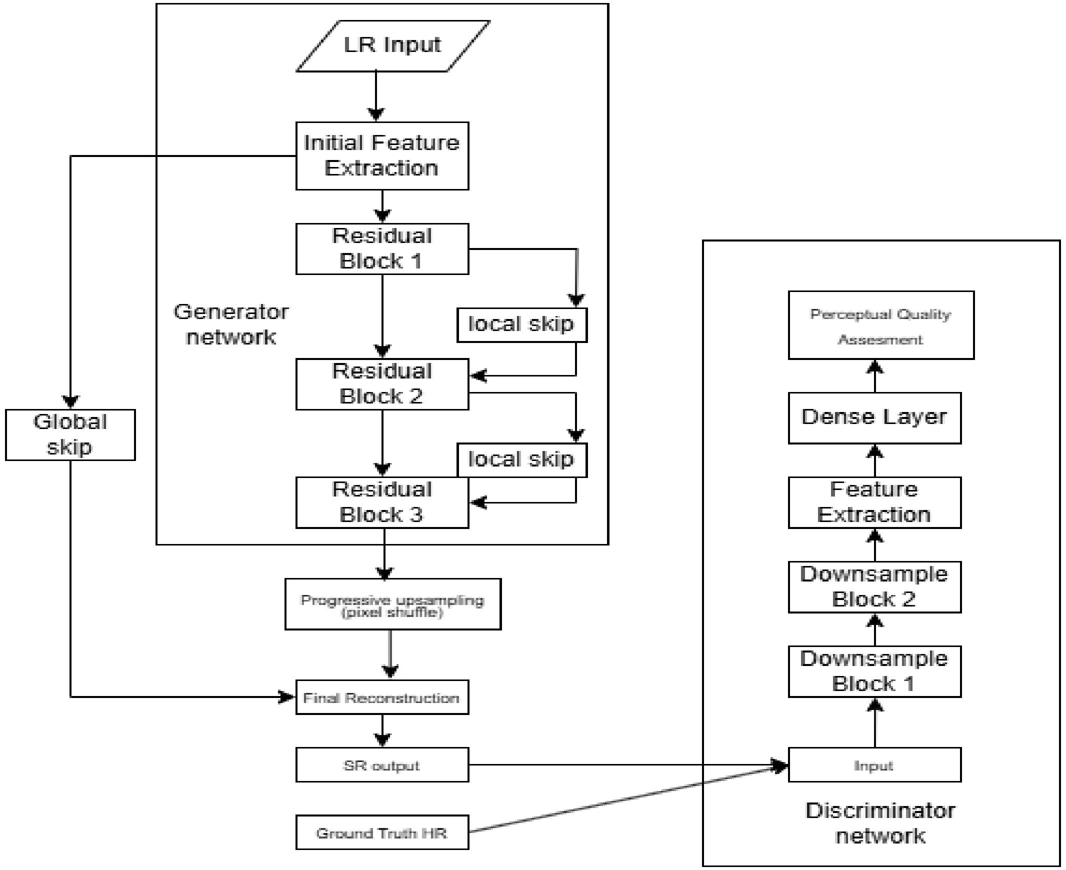
- Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391-407). Springer, Cham.

Proposed Improvements

- Depth (Adding More Layers)
  - Purpose: Increases the network's capacity to capture complex, hierarchical features.
  - Benefit: Enhances the model's ability to reconstruct high-frequency details in the image.
- Width (Increasing Channel Capacity)
  - Purpose: Expands the number of filters per layer to capture richer image representations.
  - Benefit: Provides more descriptive features, aiding in improved texture and color recovery.
- Optimized Upsampling Layers
  - Purpose: Refines the upsampling process using techniques like transpose convolution or sub-pixel convolution.
  - Benefit: Produces sharper, high-resolution outputs while reducing artifacts commonly seen in standard upsampling.
- Enhanced Perceptual Quality using GAN:
  - Fine-tuning with GANs has been effective in models like ESRGAN
  - Applying GAN with CNN based model
    - Implementing GANs for super-resolution often leads to improved structural similarity, finer texture details, and reduced artifacts compared to conventional upsampling methods.

Proposed Model





Evaluation Metrics and Baseline Comparison

- Evaluation Metrics
  - Peak Signal-to-Noise Ratio (PSNR): Measures the fidelity of reconstructed images, with higher PSNR values indicating better quality.
  - Structural Similarity Index (SSIM): Assesses perceived image quality by comparing structural information, with a focus on luminance, contrast, and structure similarity.
- Datasets
  - Set5 and Set14: Commonly used benchmark datasets for evaluating super-resolution models.
  - Maybe: BSD100 and Urban100 for additional evaluation on diverse image sets.
- Baseline Comparisons
  - Compare performance of optimized CNN-based models against traditional upsampling methods and baseline architecture:
    - Bilinear Upsampling
    - Bicubic Upsampling
    - Compare with baseline SRCNN, FSRCNN and EDSR
  - Expected PSNR targets:
    - Factor 2x: PSNR ≥ 34 on Set5
    - Factor 3x: PSNR ≥ 31 on Set5
    - Factor 4x: PSNR ≥ 29 on Set5

Thank you

QUESTIONS?

Missouri State
UNIVERSITY