Image Enhancement using Deep Learning

By:
1. Srinidhi C - 19BPS1005
2. Mohanapriya Singaravelu - 19BPS1046
3. Harini Jaishanker - 19BPS1058

What is Single Image Super Resolution?

Image Super Resolution refers to the task of enhancing the resolution of an image from low-resolution (LR) to high (HR). This is essentially done by upsampling the LR image to HR image.

Deep learning models are utilized to perform this upsampling. Loss functions are used to measure reconstruction error and guide the model optimization.

The different type of Loss Functions used

Pixelwise L1 loss – Absolute difference between pixels of ground truth HR image and the generated one.

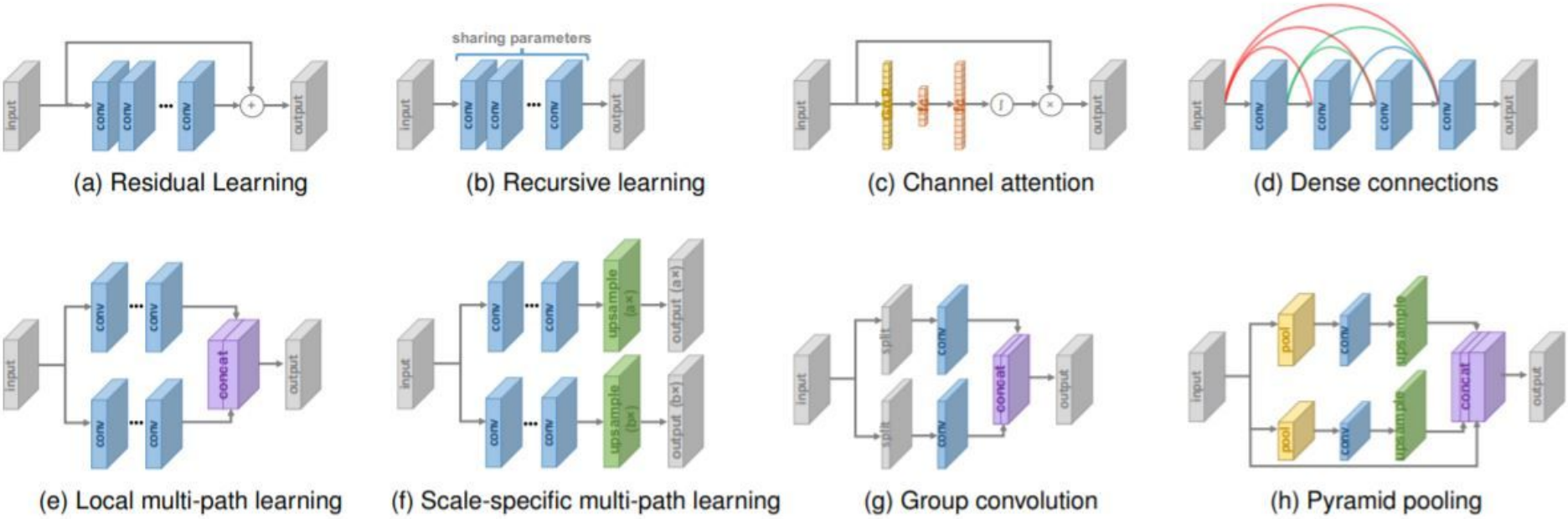
Pixelwise L2 loss – Mean squared difference between pixels of ground truth HR image and the generated one.

Content loss – the content loss is indicated as the Euclidean distance between high-level representations of the output image and the target image. High-level features are obtained by passing through pre-trained CNNs like VGG and ResNet.

Adversarial loss – Based on GAN where we treat the SR model as a generator, and define an extra discriminator to judge whether the input image is generated or not.

PSNR – Peak Signal-to-Noise Ratio (PSNR) is a commonly used objective metric to measure the reconstruction quality of a lossy transformation. PSNR is inversely proportional to the logarithm of the Mean Squared Error (MSE) between the ground truth image and the generated image.

Different type of Network architectures



(a) Residual Learning (b) Recursive learning (c) Channel attention (d) Dense connections

(e) Local multi-path learning (f) Scale-specific multi-path learning (g) Group convolution (h) Pyramid pooling

Enhanced Super Resolution GAN

- SRResNet-based architecture with residual-in-residual blocks;
- Mixture of context, perceptual, and adversarial losses. Context and perceptual losses are used for proper image upscaling, while adversarial loss pushes neural network to the natural image manifold using a discriminator network that is trained to differentiate between the super-resolved images and original photo-realistic images.
- Uses Sub-pixel upsampling method
- Improvement over SRGAN: network architecture, adversarial loss and perceptual loss.

Enhanced Deep Super Resolution Network

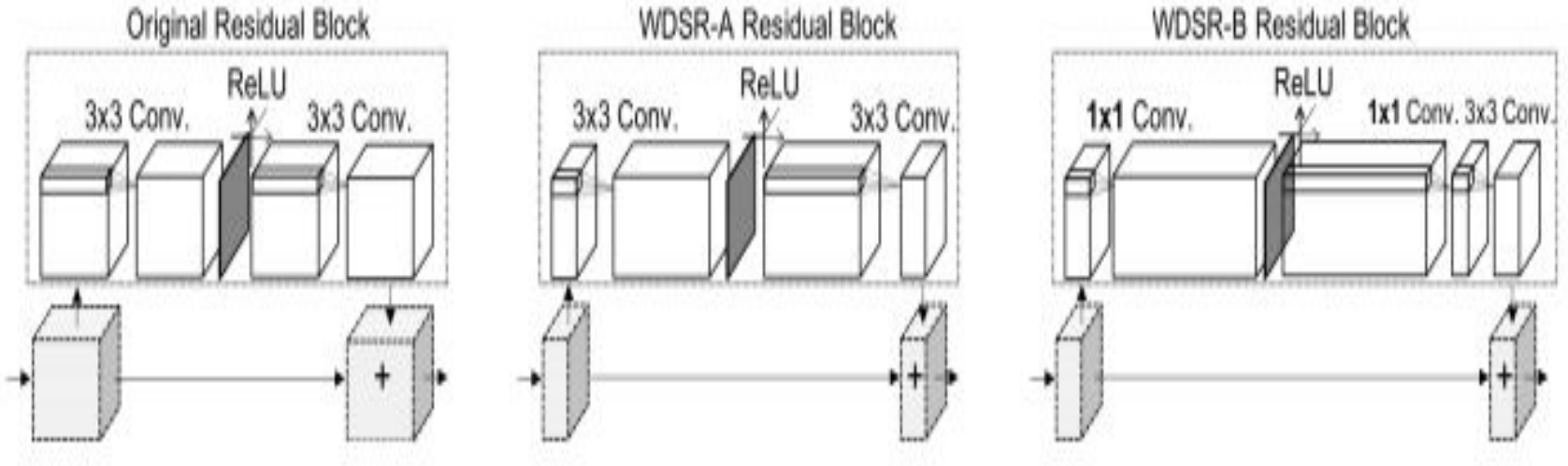
The EDSR architecture is based on the SRResNet architecture, consisting of multiple residual blocks. It’s a more efficient network as the Batch Normalization layers are removed.

The structure is similar to SRResNet, but our model does not have ReLU activation layers outside the residual blocks

Wide Activation Deep Super Resolution

In the WDSR models the number of convolutional filters is reduced to 32 and the skip connection and residual blocks are distinctly different to the SR-Resnet and EDSR models. In the WDSR-A model , residual blocks are structurally identical to the EDSR residual block with the only difference being that the number of filters in the first convolutional layer in block is expanded by a factor of 4 (totalling 128).

The residual block in WDSR-B contains a convolutional layer with a filter expansion factor of 6 (totalling 192) with reLU activation and a kernel size of 1. This is followed by another convolutional layer of kernel size 1, with 154 layers (80% of 192). The block ends with a convolutional layer with kernel size 3 and residual block addition.


Original Residual Block
3x3 Conv.
ReLU
3x3 Conv.
WDSR-A Residual Block
3x3 Conv.
ReLU
3x3 Conv.
WDSR-B Residual Block
1x1 Conv.
ReLU
1x1 Conv.
3x3 Conv.

The dataset used to train - DIV2K dataset

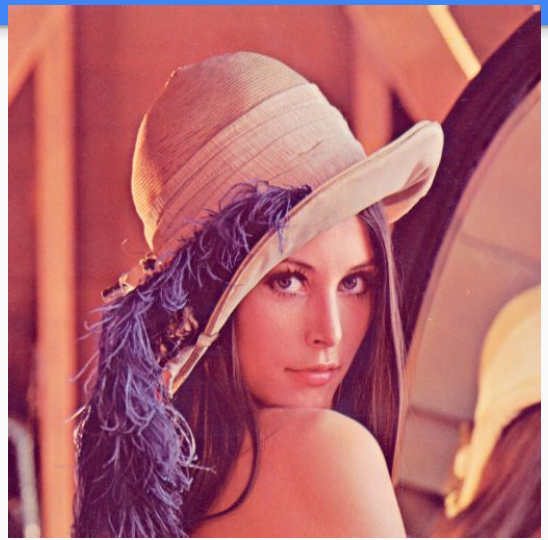
The DIV2K dataset is divided into:

train data: starting from 800 high definition high resolution images we obtain corresponding low resolution images and provide both high and low resolution images for 2, 3, and 4 downscaling factors

validation data: 100 high definition high resolution images are used for generating low resolution corresponding images, the low res are provided from the beginning of the challenge and are meant for the participants to get online feedback from the validation server; the high resolution images will be released when the final phase of the challenge starts.

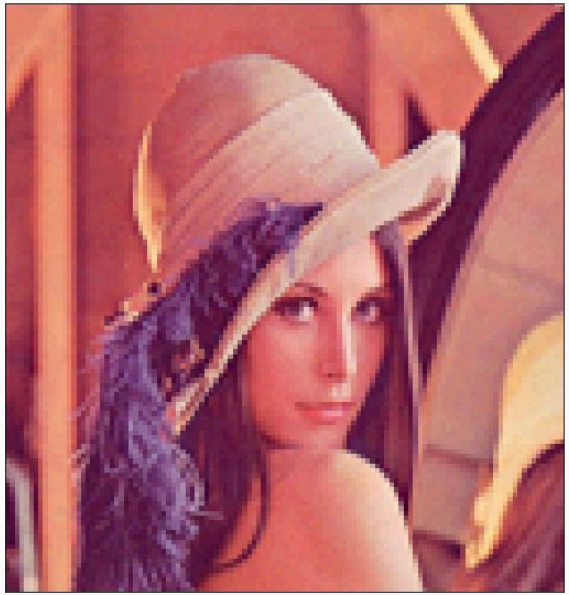
test data: 100 diverse images are used to generate low resolution corresponding images; the participants will receive the low resolution images when the final evaluation phase starts and the results will be announced after the challenge is over and the winners are decided.

Input Image



EDSR

LR

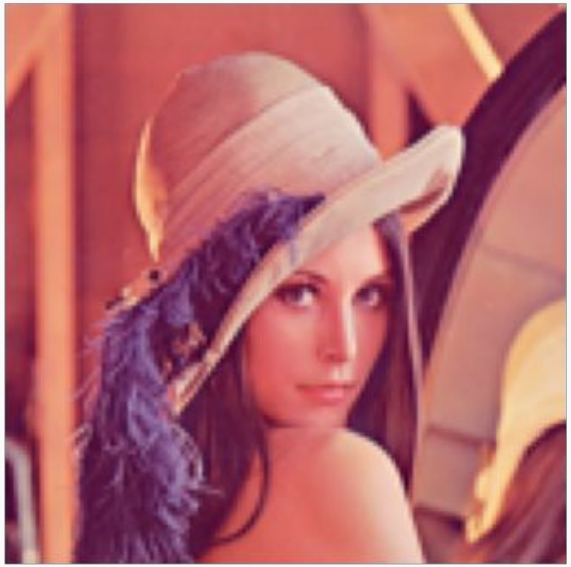


SR (x4)

ESRGAN

Original

x4 Bicubic



Super Resolution