IMAGE SUPER RESOLUTION

by Ivan Maloshtan, Rostyslav Zubko

Objective – enhance the resolution of images using deep learning techniques

DATASETS

DIV2K

Set5

Set14

Preprocessing: downsampling high-resolution images -> Low-resolution

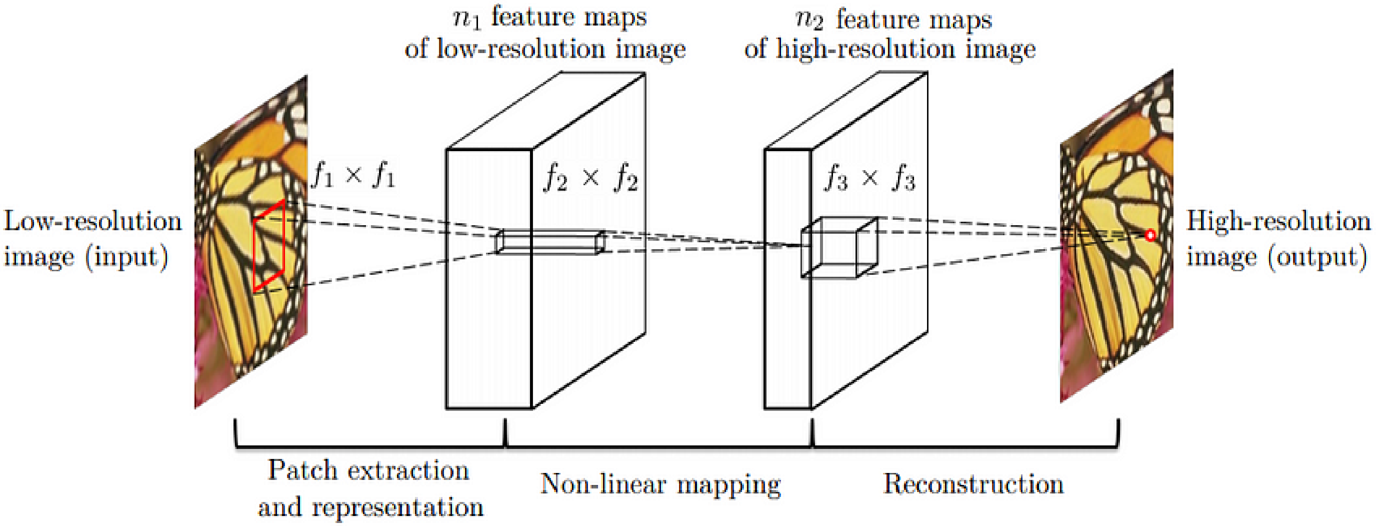
ARCHITECTURES OF LEARNING MODELS:

SRCNN – baseline model;

More advanced:

- EDSR
- SRGAN
- VDSR

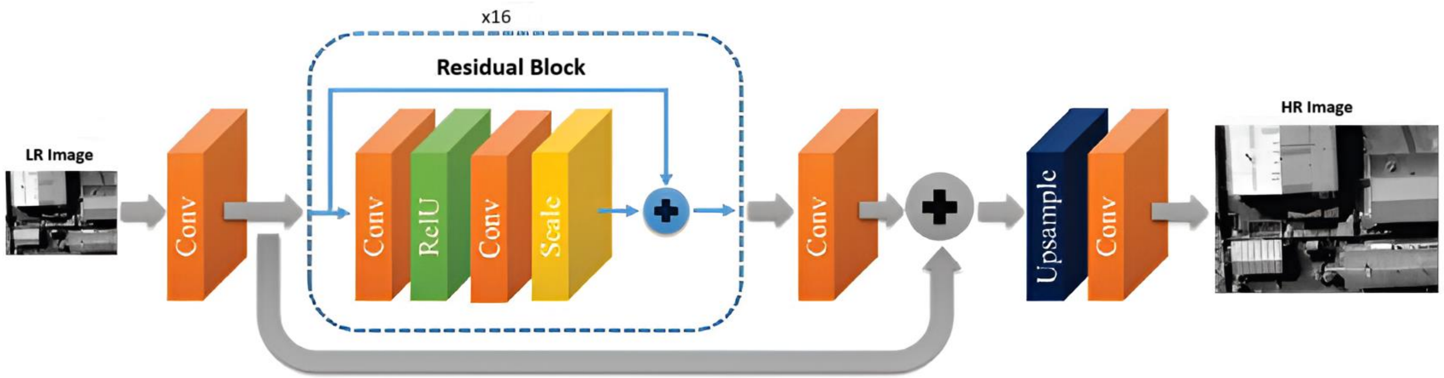
SRCNN



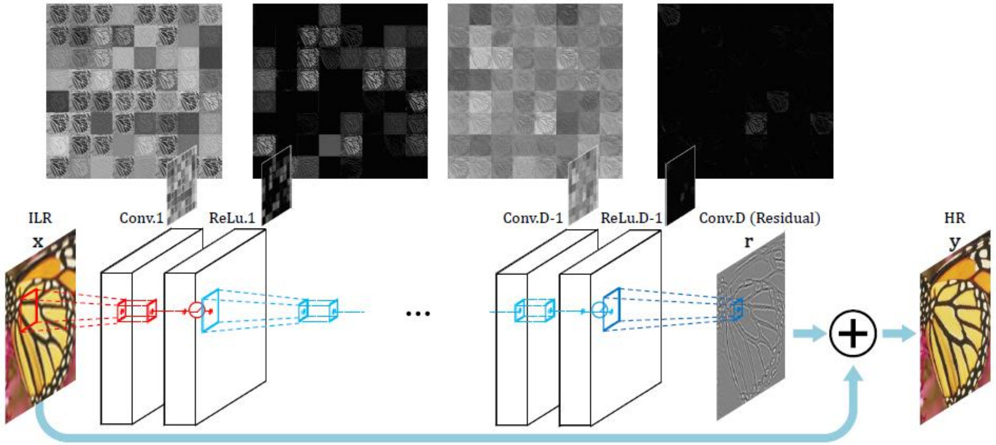
- 3 Convolutional layers
- Filters: *128-64-1*
- Kernel size: *9-3-5*
- Optimization: *Adam optimizer*
- Loss function: *MSE*

EDSR

- Pre-residual block: *1 convolutional layer*
- Residual blocks (x16): ***2 convolutional layers + ReLU activation function***
- **Post-residual block:** ***1 convolutional layer***
- **Optimizer:** ***Adam optimizer***

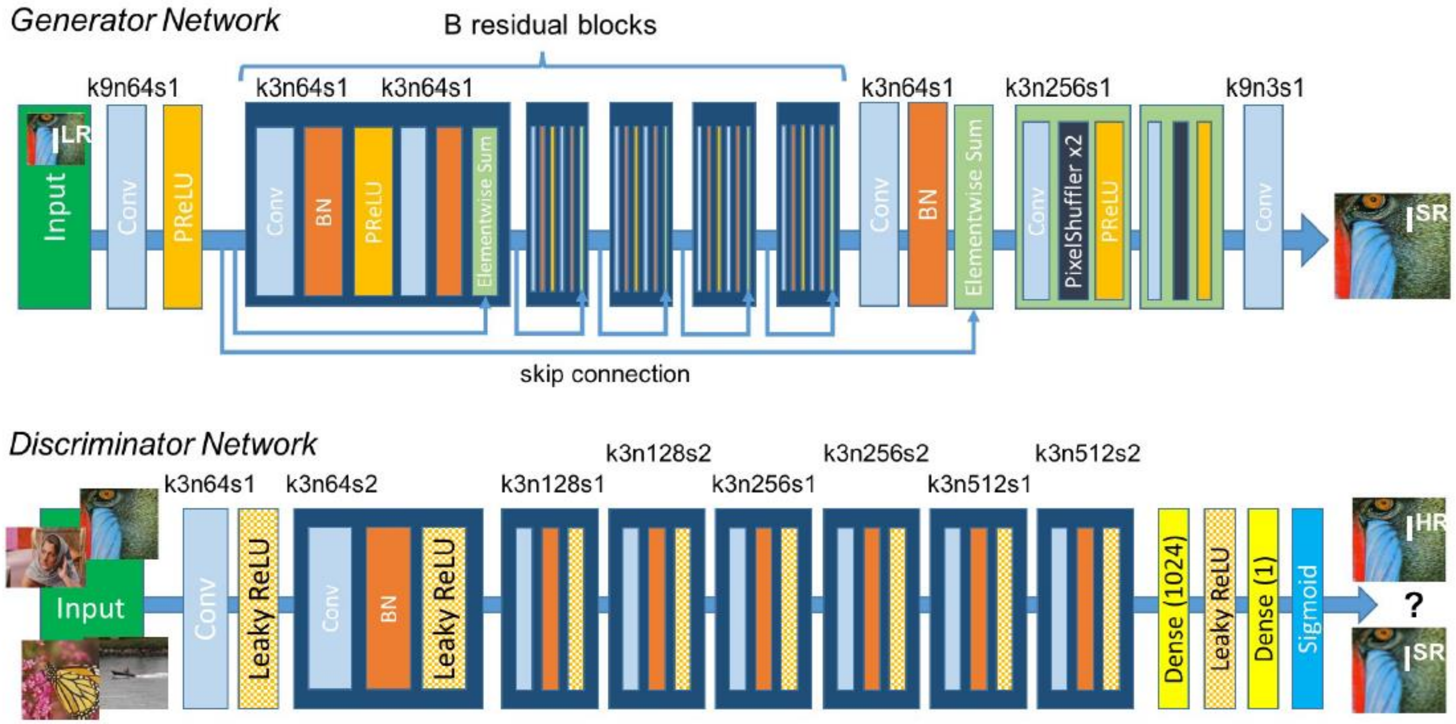


VDSR



1. Input Layer
2. Convolutional Layers
3. Residual Learning
4. Output Layer

SRGAN



<u>Generator</u> - produce fake data (SR image)

<u>Discriminator</u> -
binary classifier; evaluates whether a given data instance is real or fake

<u>Adversarial Training</u>: The generator and discriminator are trained together, refining the image quality through adversarial feedback

***Discriminator* generate the adversarial loss which then backpropagated into the generator architecture.**

SRGAN

Optimization: Adam optimizer

Problem: solutions of MSE optimization problems often lack high-frequency content (details)

Introduces Perceptual Loss - weighted sum of a content loss and adversarial loss component.

MSE-based content loss is replaced with a loss calculated on feature maps of the VGG network (VGG19 classifier); VGG loss is a MSE between the feature representations of a constructed image and reference image.

$$l^{SR} = \underbrace{\underbrace{l^{SR}_{\mathrm{X}}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}}_{\text{perceptual loss (for VGG based content losses)}}$$