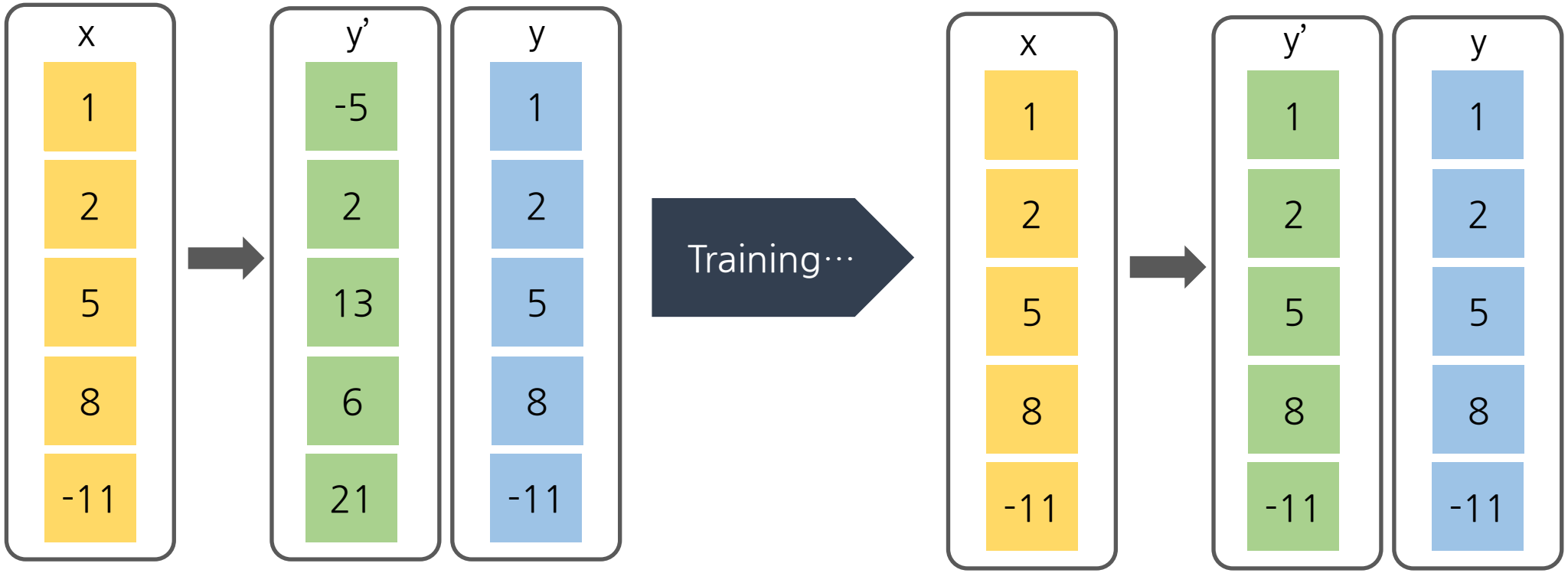
IVCL

Basic Summary of Super Resolution

백전능

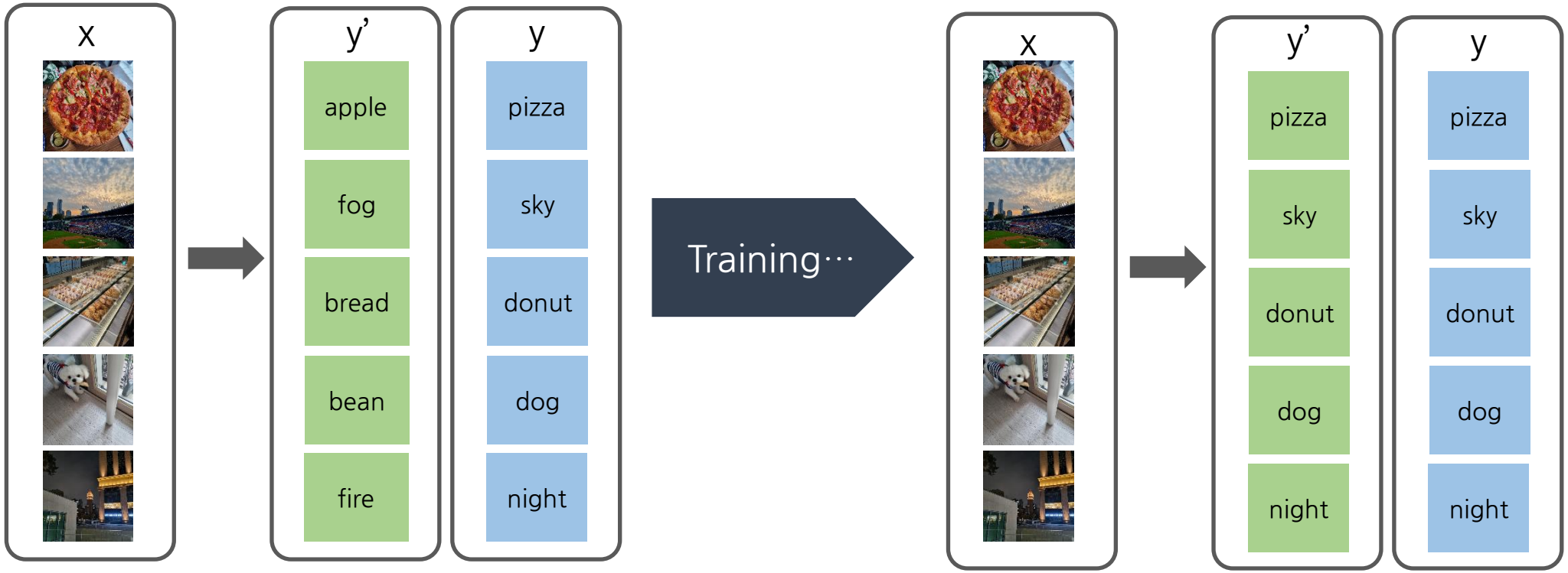
Apply Deep Learning

- Regression



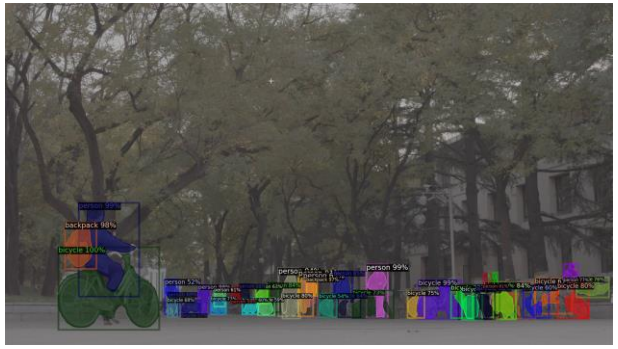
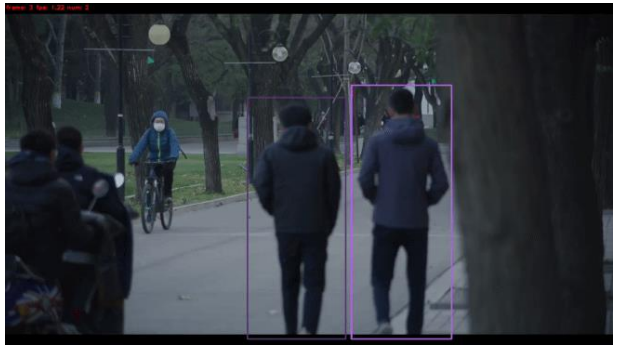
Apply Deep Learning

- Classification



Apply Deep Learning

- Detection, Segmentation, Tracking

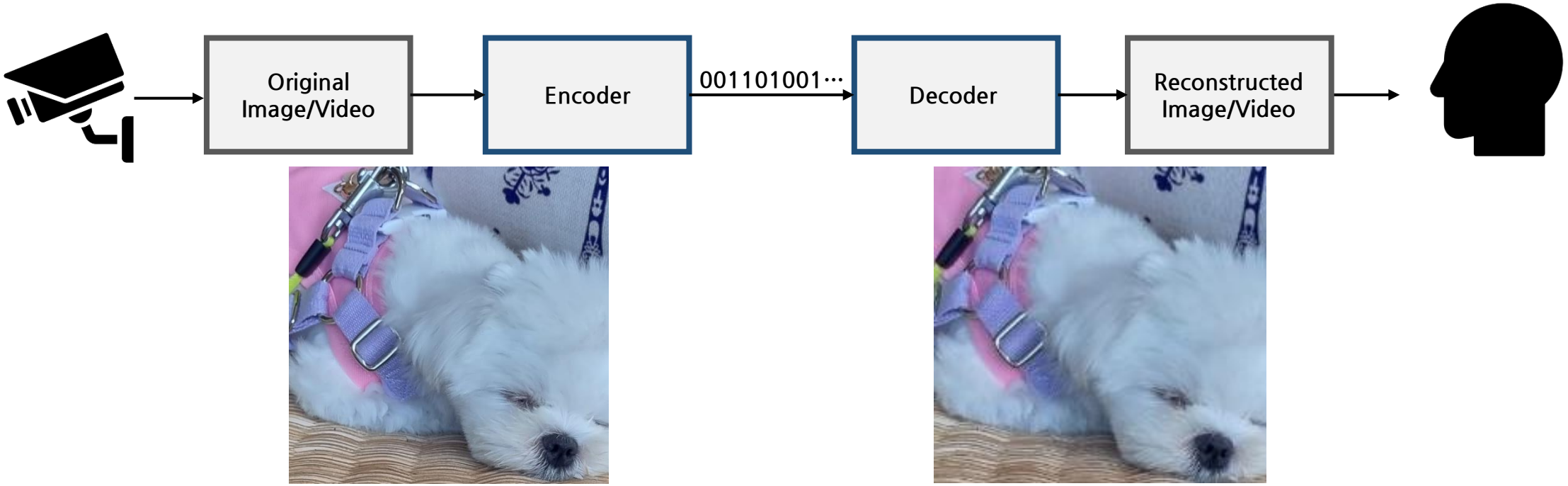
- Image generation

Input : floating city street

Super Resolution

- SR(super resolution) : Upsampling with deep learning

Necessity of SR



- Noise occur during image, video transmission.
- Some image, video may be of low quality in the original
- SR is required to improve the resolution of image, video

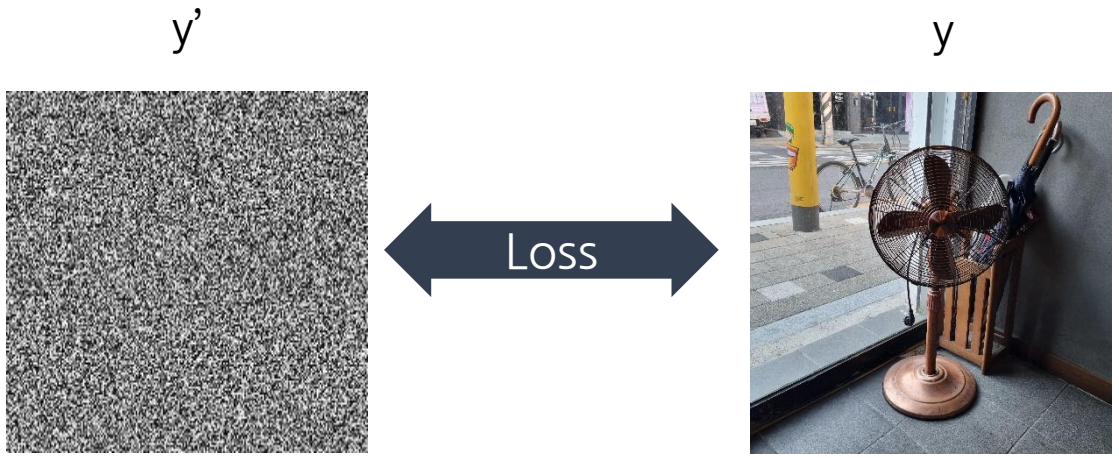
SR Loss Function

- L1 (MAE)

$$\mathcal{L}_{\text{pixel\_l1}}(\hat{I}, I) = \frac{1}{hwc}\sum_{i,j,k}|\hat{I}_{i,j,k} - I_{i,j,k}|$$

- L2 (MSE)

$$\mathcal{L}_{\text{pixel\_l2}}(\hat{I}, I) = \frac{1}{hwc}\sum_{i,j,k}(\hat{I}_{i,j,k} - I_{i,j,k})^2$$



SR Assessment

- PSNR (Peak Signal-to-Noise Ratio)

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{\frac{1}{N}\sum_{i=1}^{N}(I(i) - \hat{I}(i))^2}\right)$$

- SSIM (Structural Similarity)

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

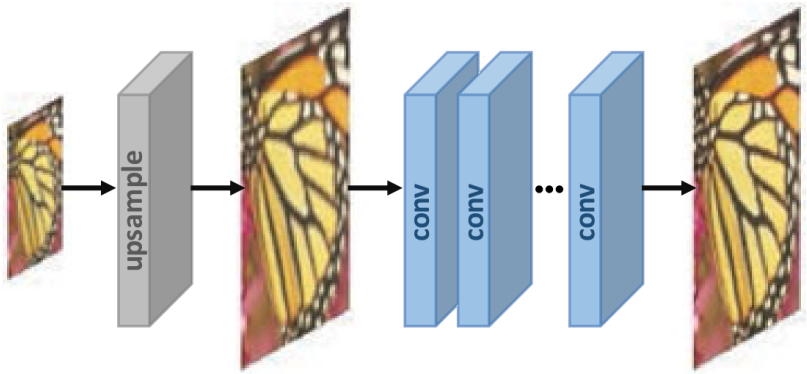
y’

Assment

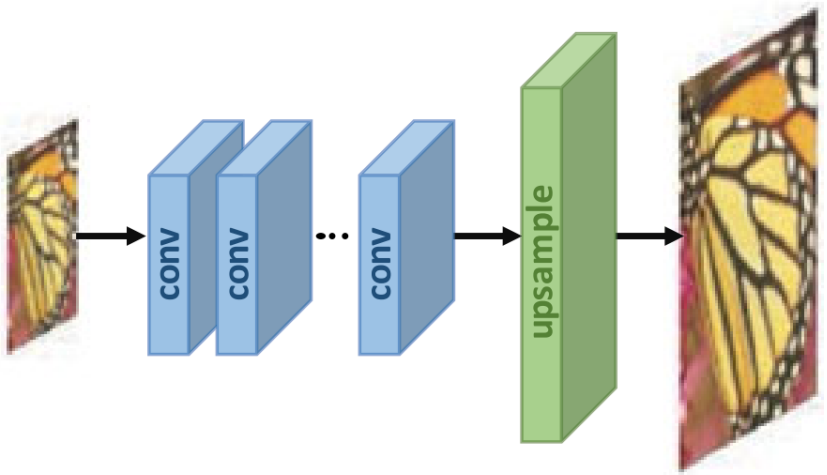
y

SR Frameworks

- Pre-upsampling
  - Traditional upsampling method
  - cost of time and space

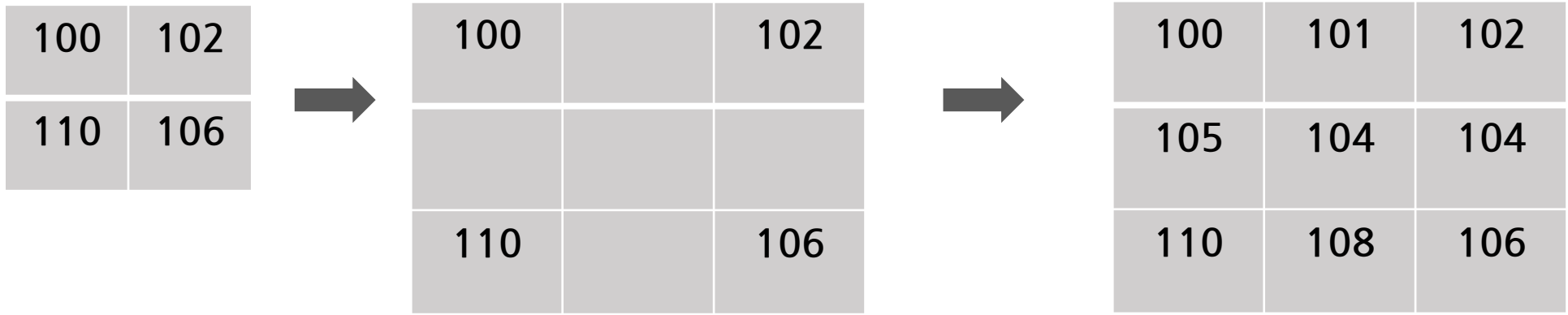


- Post-upsampling
  - learnable upsampling layer
  - end-to-end, high efficiency



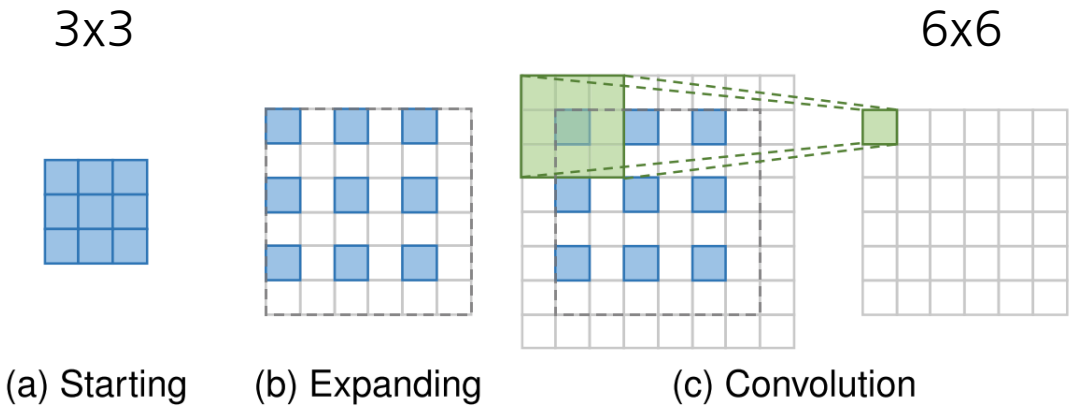
Upsampling Method

- Interpolation
  - Traditional upsampling method
  - Nearest Neighbor, Bilinear, Bicubic, etc…

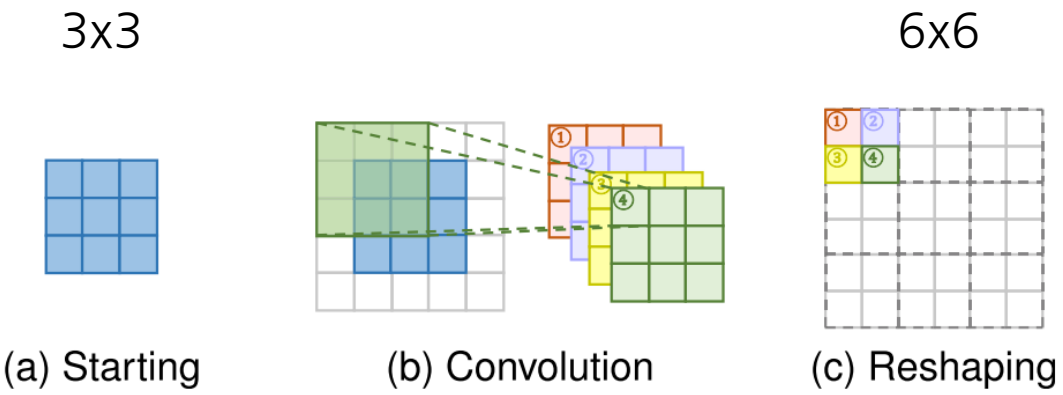


Upsampling Method

- Transposed Convolution Layer
  - Learning-Based Upsampling

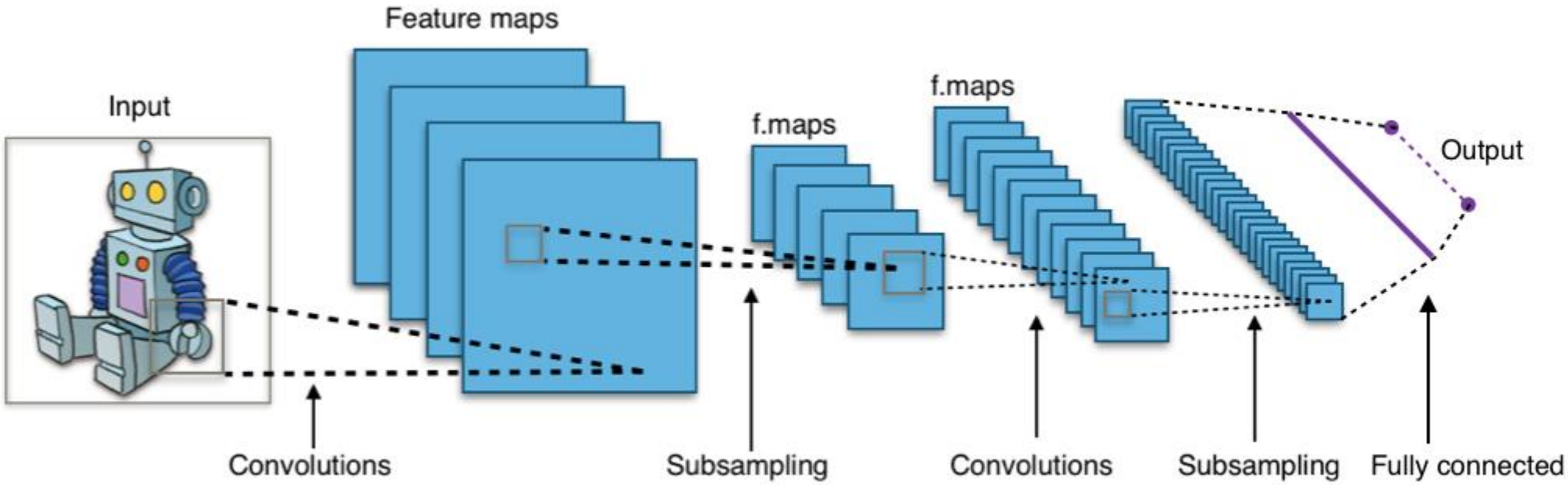


- Pixel Shuffle Layer
  - Learning-Based Upsampling



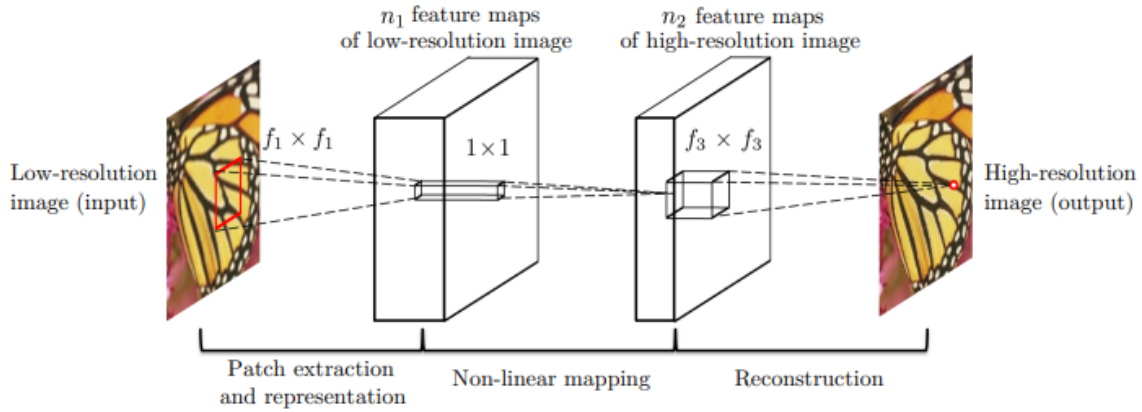
SR Network : SRCNN

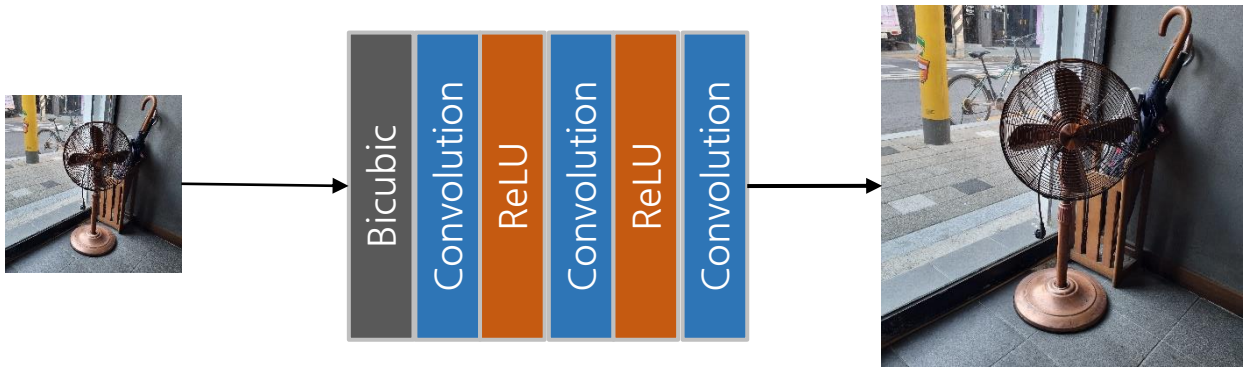
- CNN(Convolutional Neural Network)



SR Network : SRCNN

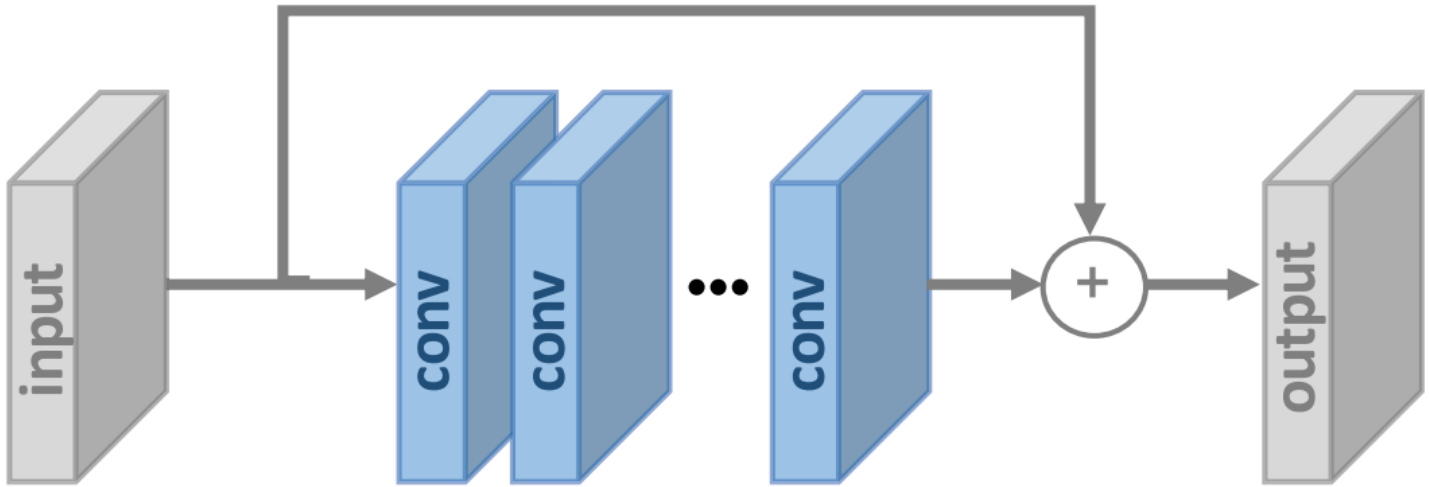
- Light weight structure
- Fast learning speed





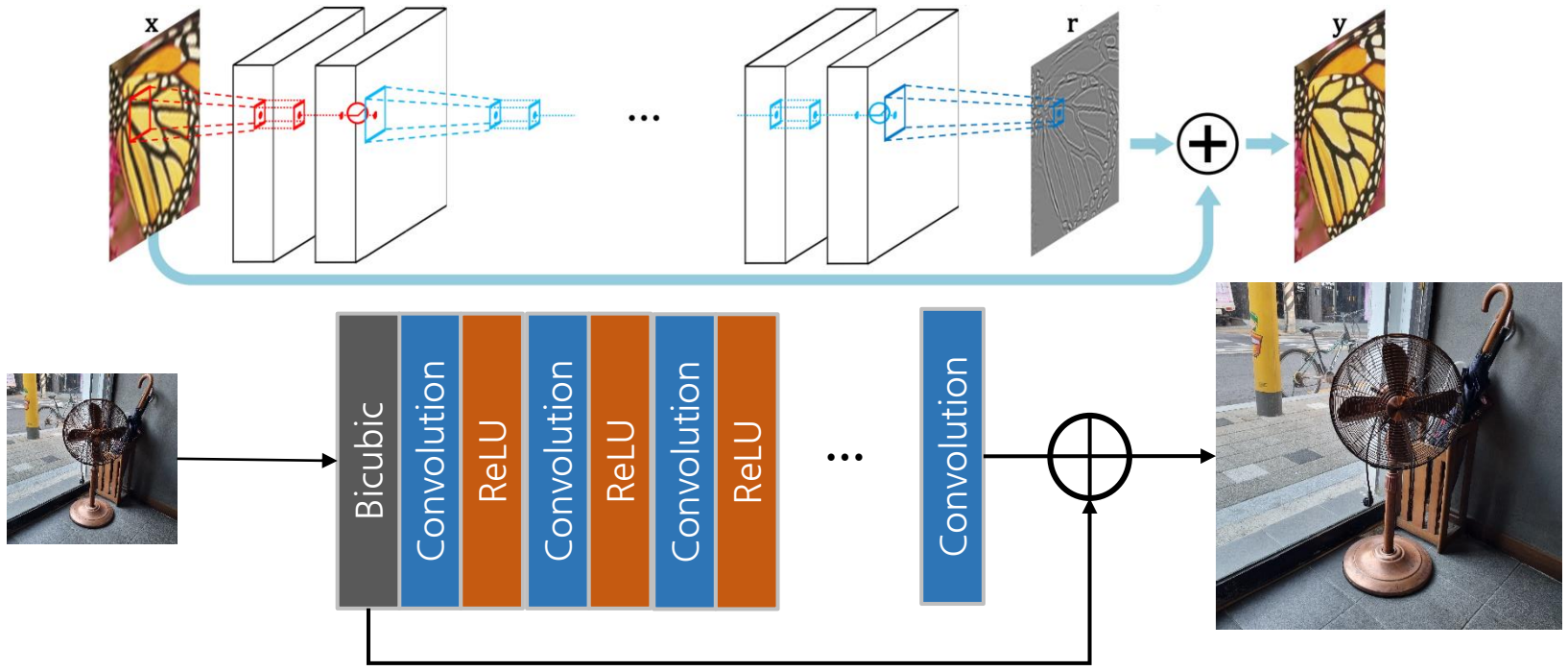
SR Network : VDSR

- Residual Learning
  -Reduce the impact of vanishing gradients
  -Deeper neural network design



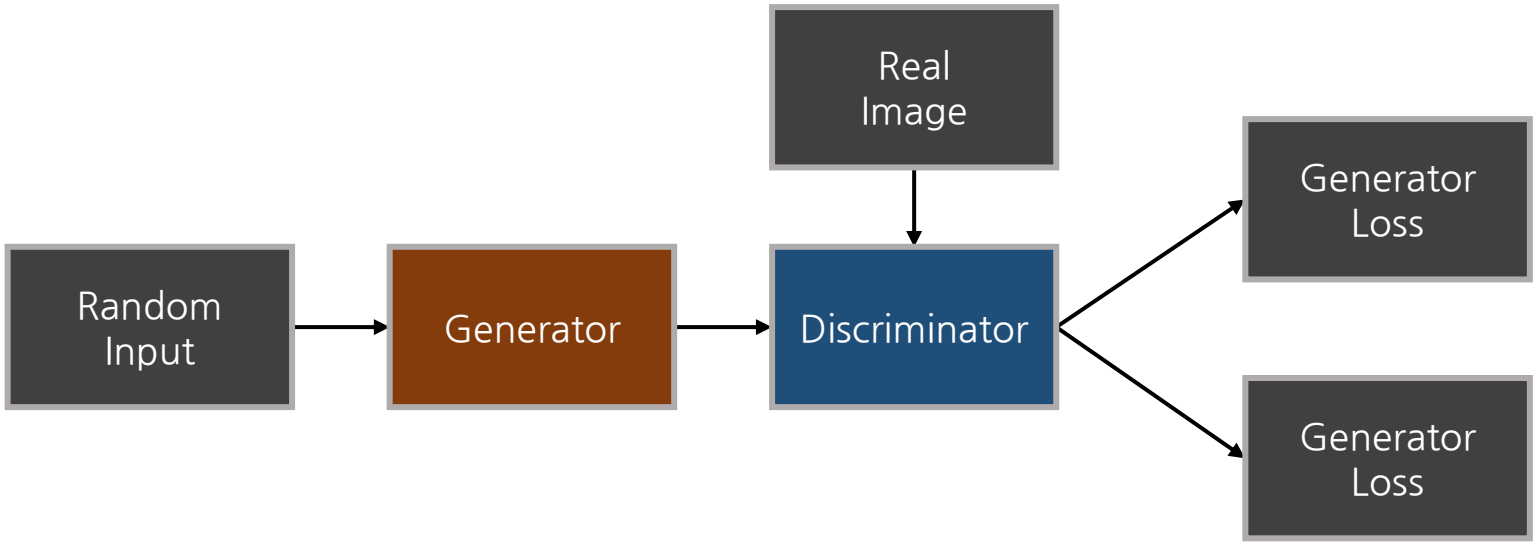
SR Network : VDSR

- Multi scale factor
- Deeper neural networks due to residual learning
- Fast convergence due to high learning rates



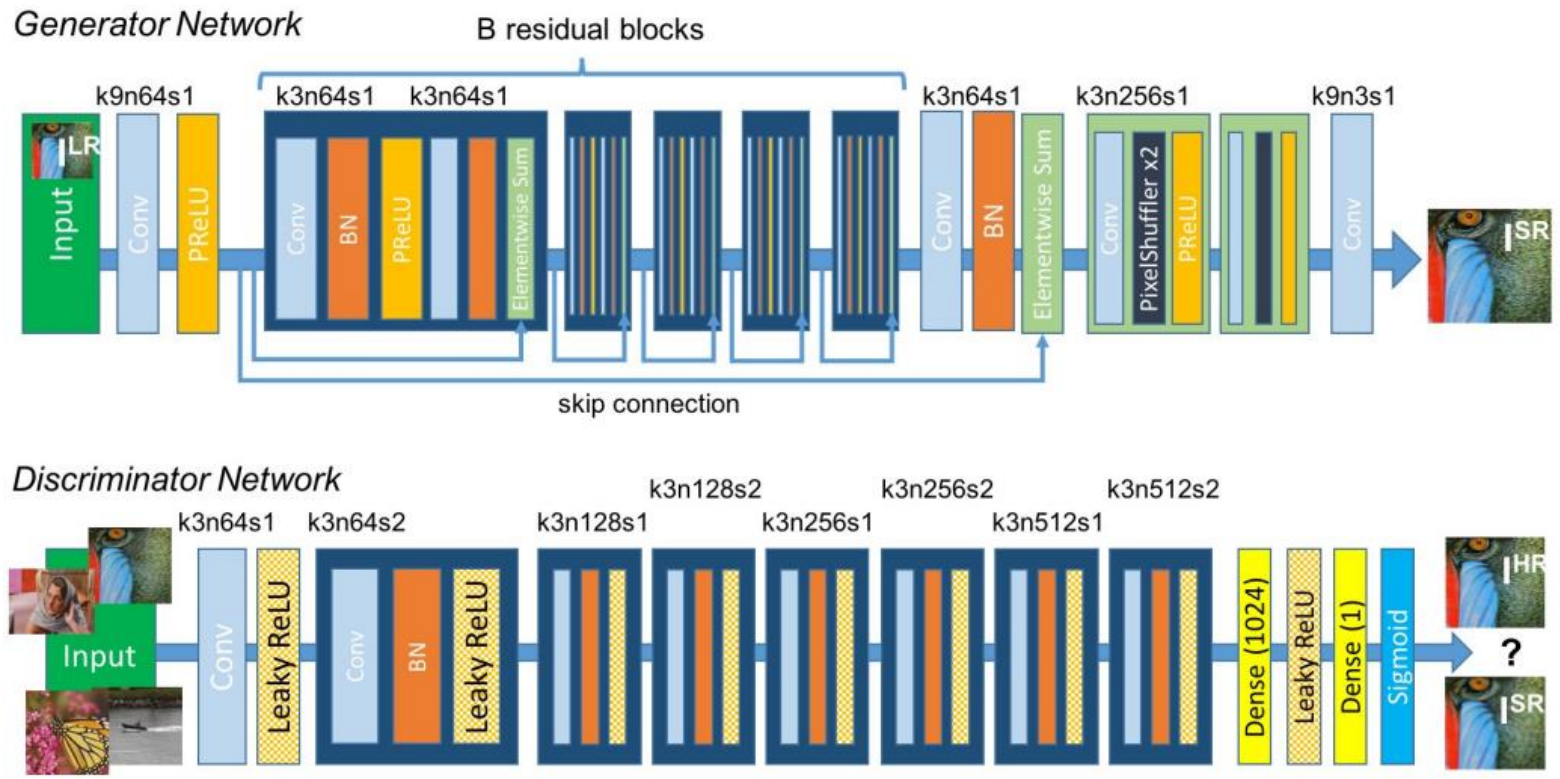
SR Network : SRResNet, SRGAN

- GAN (Generative Adversarial Network)



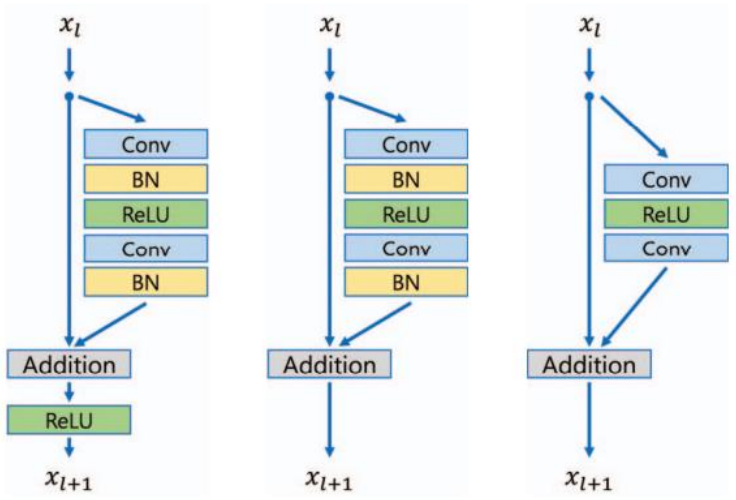
SR Network : SRResNet, SRGAN

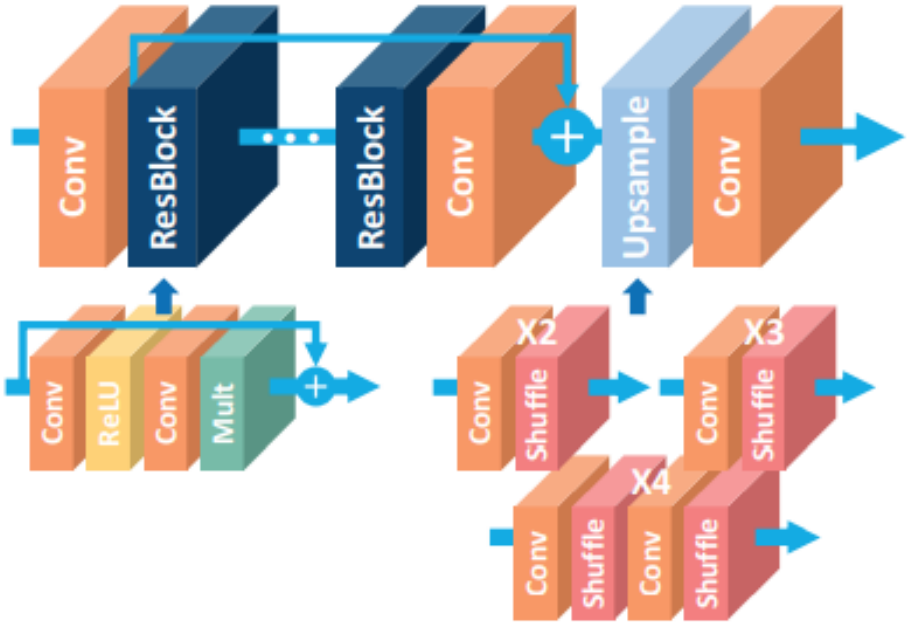
- For HVS(Human Visual System), use perceptual loss instead of MSE



SR Network : EDSR

- Residual block modification





SR Performance Comparison

Models	PSNR/SSIM(×4)	Train data	Parameters	Mult&Adds
SRCNN_EX [48]	30.49/0.8628	ImageNet subset	57K	52.5G
ESPCN [49]	30.90/-	ImageNet subset	20K	1.43G
VDSR [61]	31.35/0.8838	G200+Yang91	665K	612.6G
DRCN [63]	31.53/0.8838	Yang91	1.77M(recursive)	17974.3G
DRRN [70]	31.68/0.8888	G200+Yang91	297K(recursive)	6796.9G
LapSRN [84]	31.54/0.8855	G200+Yang91	812K	29.9G
SRResNet [68]	32.05/0.9019	ImageNet subset	1.5M	127.8G
MemNet [76]	31.74/0.8893	G200+Yang91	677K(recursive)	2265.0G
RDN [78]	32.61/0.9003	DIV2K	22.6M	1300.7G
EDSR [71]	32.62/0.8984	DIV2K	43M	2890.0G
MDSR [71]	32.60/0.8982	DIV2K	8M	407.5G
DBPN [90]	32.47/0.898	DIV2K+Flickr+ ImageNet subset	10M	5715.4G

Thank you for listening.