CNN Architecture

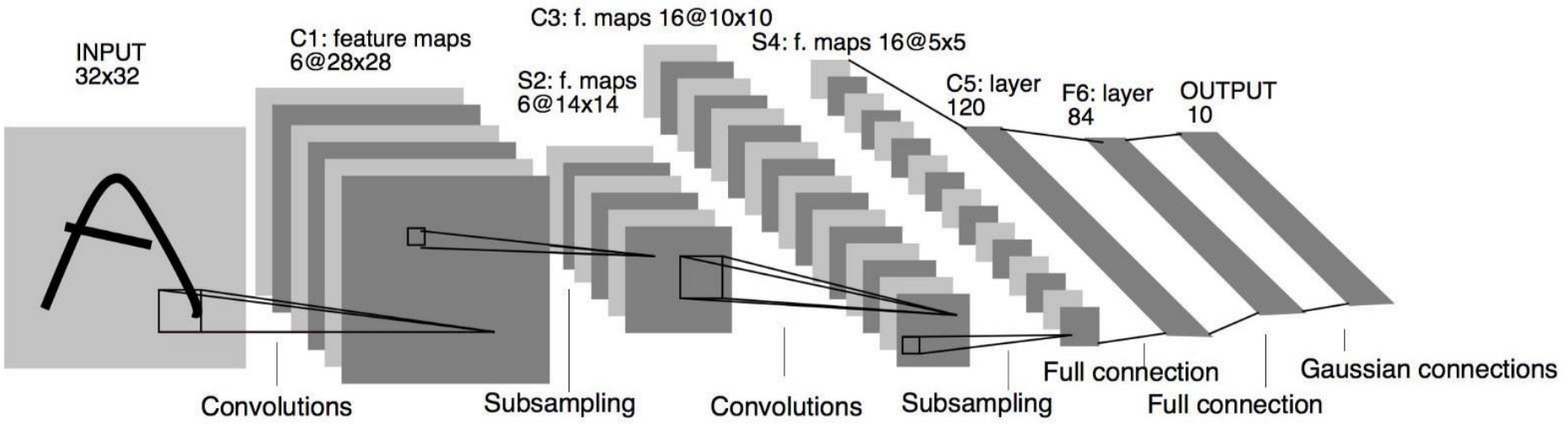
Summary

NCSOFT Vision AI Lab

김준호

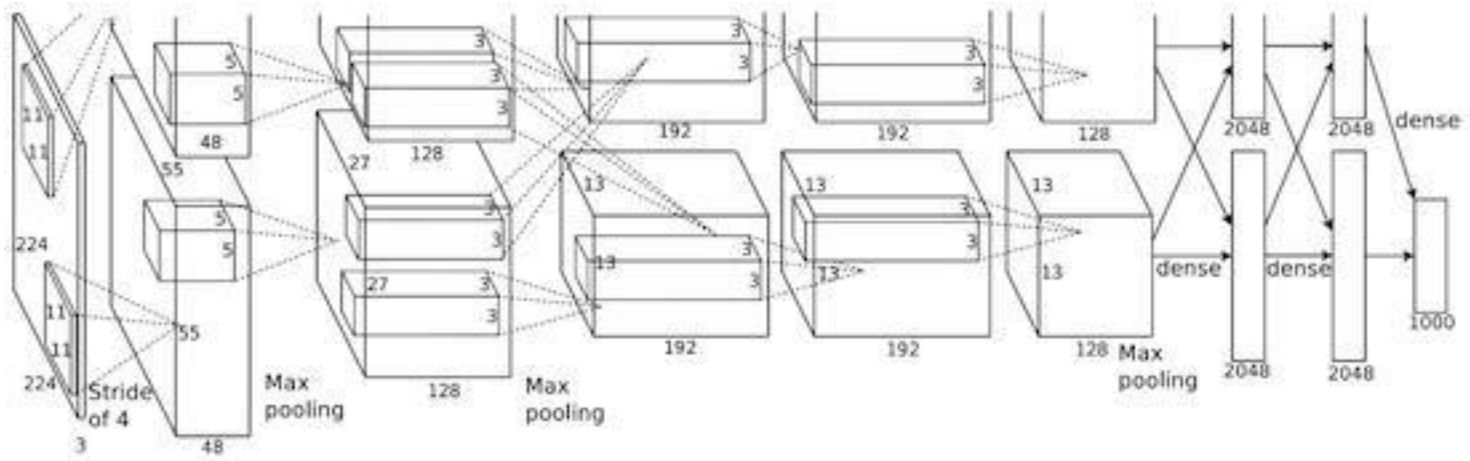
https://github.com/taki0112

LeNet5



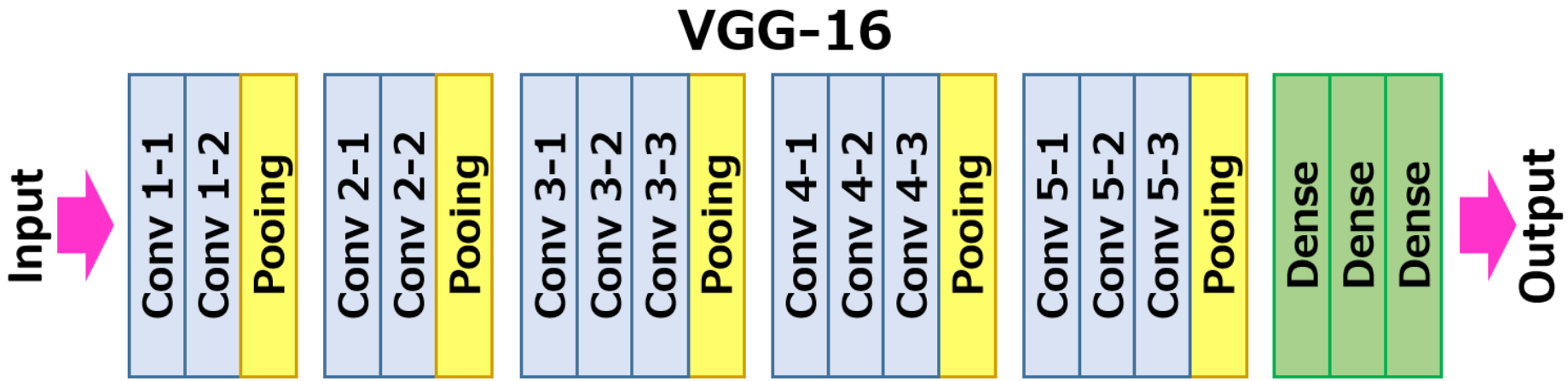
Characteristics	Contribution	Etc
• Conv – Pooling – Activation • Average pooling • Sigmoid activation & tanh activation • 5x5 Convolution filter • 7 layers and less than 1M parameters	• Use of convolution to extra ct spatial features • Subsample using spatial average of maps	• Slow to train • Hard to train (Neurons dies quickly)

AlexNet



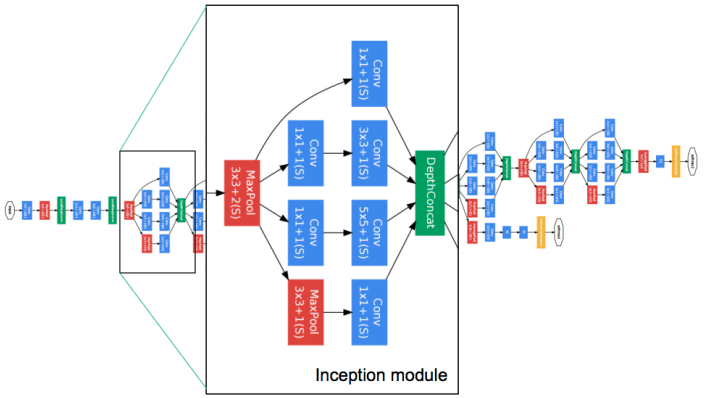
Characteristics	Contribution	Etc
• 11x11, 5x5, and 3x3 convolution • Max pooling • 3 FC • 60M parameters	• Relu activation • Dropout for regularization • Local response normalization • Image augmentation • Multiple GPU	

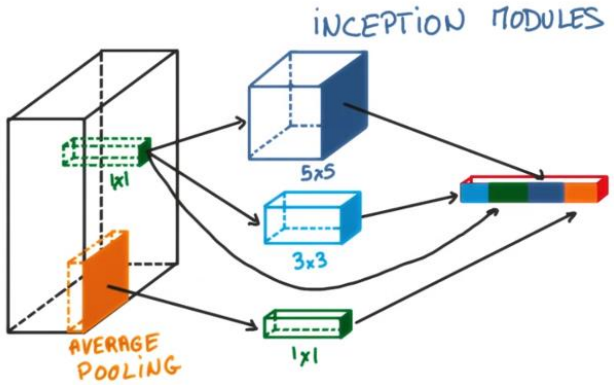
VggNet

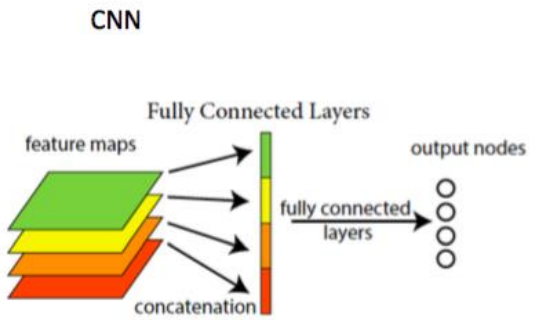


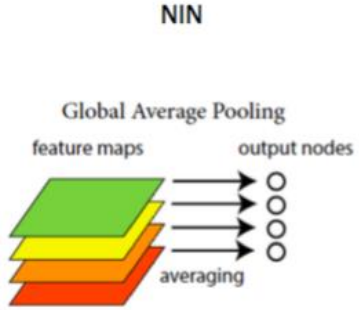
Characteristics	Contribution	Etc
• 3x3 conv • Max pooling • 130M parameters	• Simple architecture • Easy to use • Practical	• Many parameters

GoogLeNet





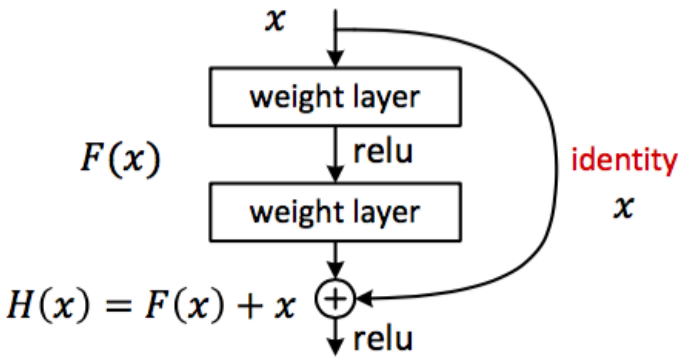


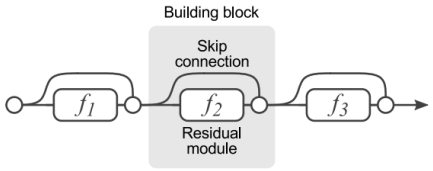


Characteristics	Contribution	Etc
• 1x1 convolution • Global average pooling • 6.8 M parameters	• Depth reduction using 1x1	• Hard to use • Inception v2, v3, v4 ~

ResNet

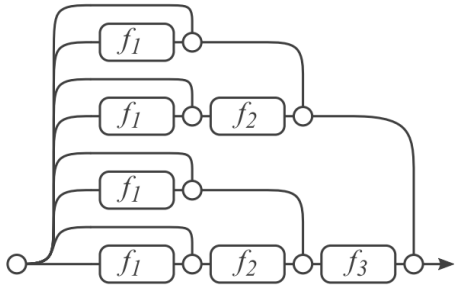
• Residual net





(a) Conventional 3-block residual network

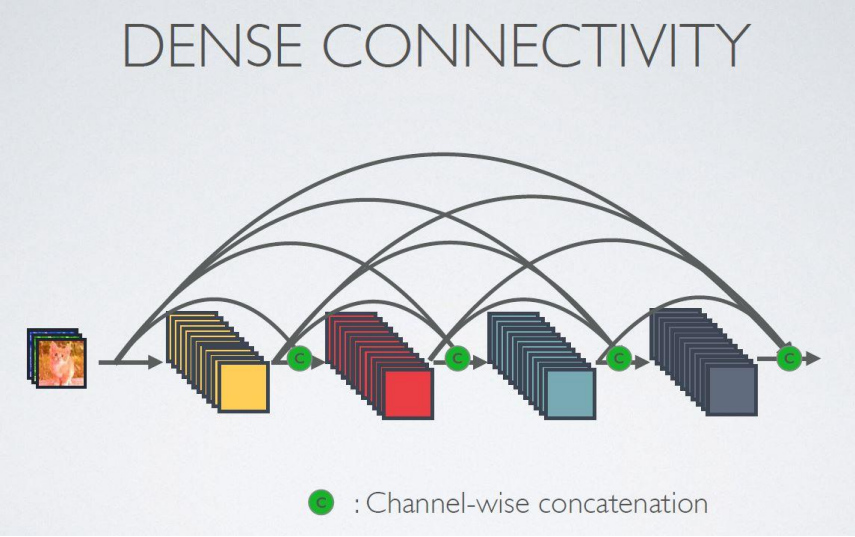
=

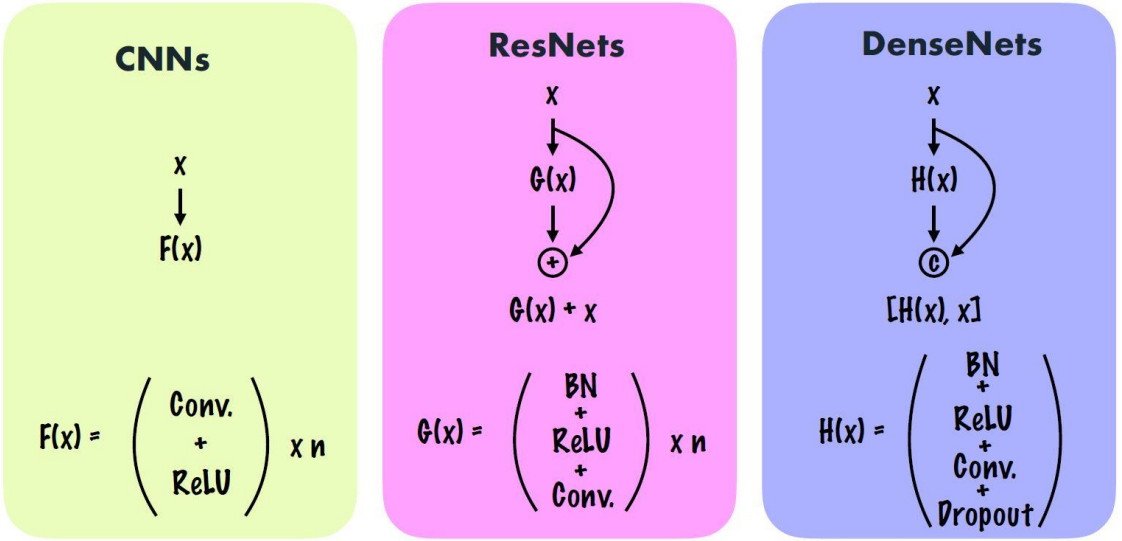


(b) Unraveled view of (a)

Characteristics	Contribution	Etc
• Going deep	• Skip connection	• Pre-activation resnet
• 1M parameters	• Easy to use	• ResNeXt
	• Practical	

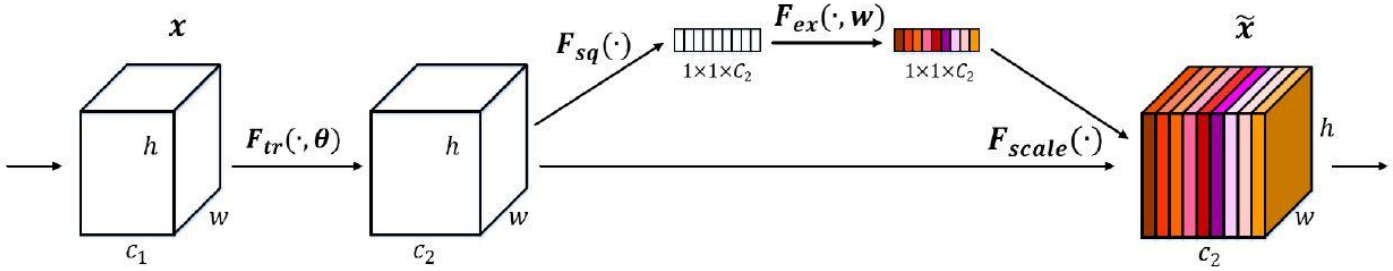
DenseNet



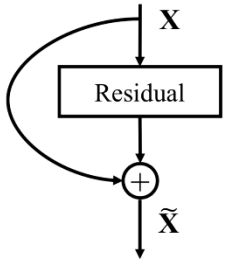


Characteristics	Contribution	Etc
• BN – Relu – conv • 1M parameters	• Concatenation • Easy to use • Practical	• More memory

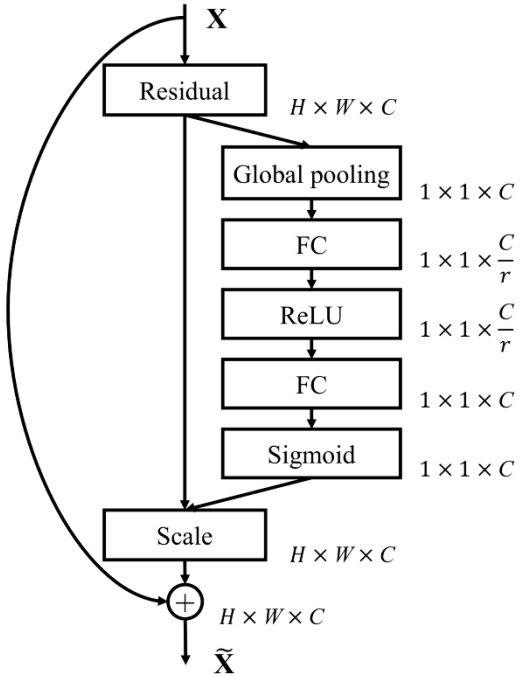
SENet



Squeeze	Excitation	Scale
• Shrinking feature maps $\in \mathbb{R}^{w\times h\times c_2}$ through spatial dimensions $(w \times h)$ • Global distribution of channel-wise responses	• Learning $W \in \mathbb{R}^{c_2\times c_2}$ to explicitly model channel-association • Gating mechanism to produce channel-wise weights	• Reweighting the feature maps $\in \mathbb{R}^{w\times h\times c_2}$



ResNet Module



SE-ResNet Module

Characteristics

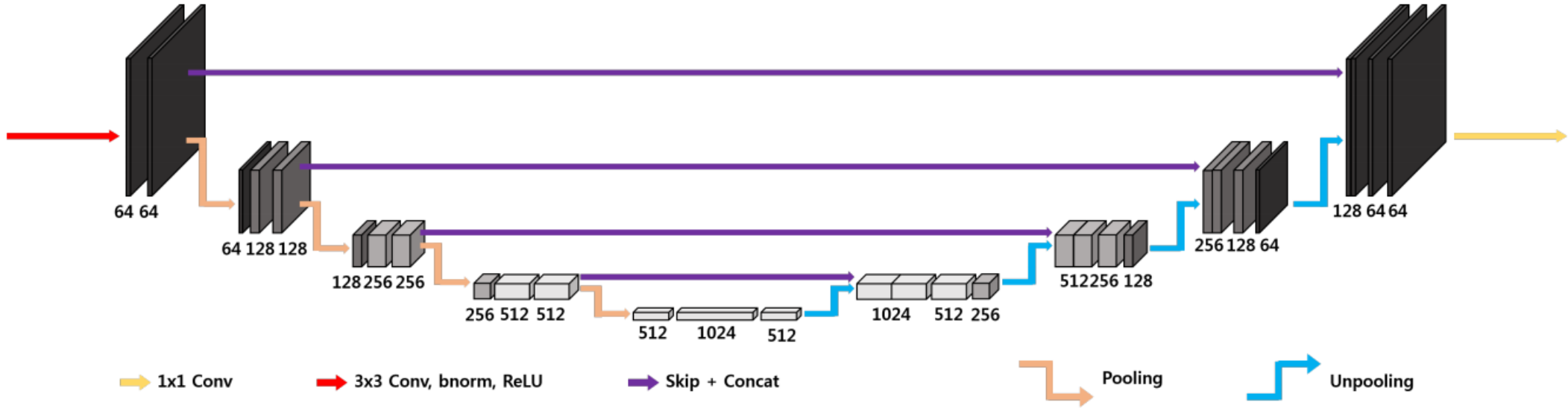
- Attention

Contribution

- Easy to use
- Practical

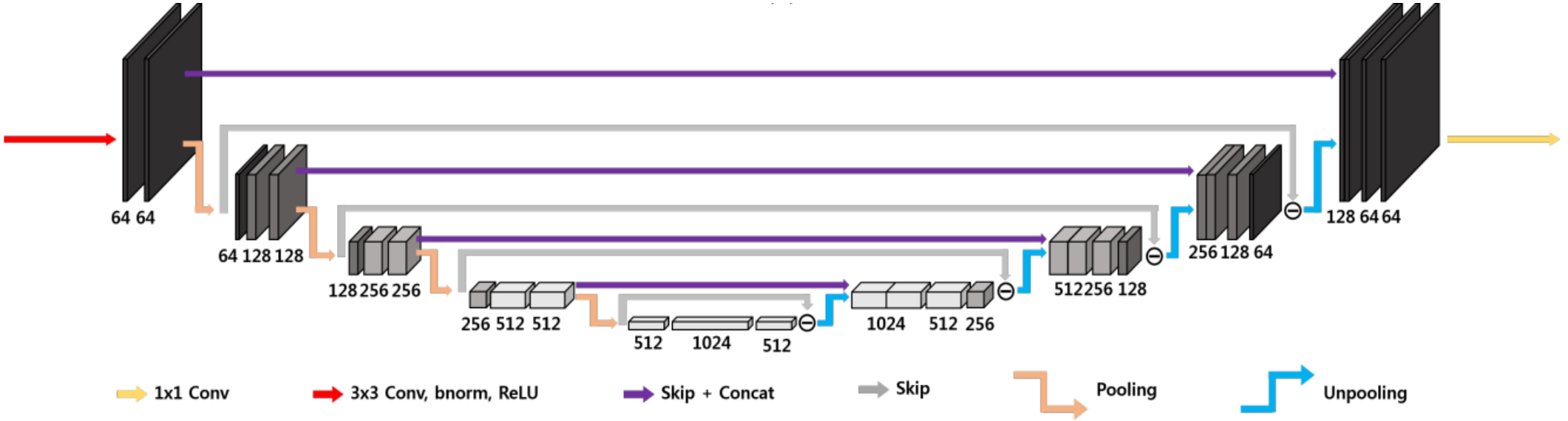
Etc

U-Net



Characteristics	Contribution	Etc
• Intuitive	• Easy to use • Practical	• Segmentation, Generation

Framing U-Net



Characteristics	Contribution	Etc
• Mathematical	• Easy to use • Practical	• Segmentation, Generation

END

https://github.com/taki0112

김준호