The Role of 1×1 Convolution

PRIMARY FUNCTIONS

Dimensionality control · Channel mixing · Adding non-linearity

Channel Reduction

Projects high-dimensional features to lower dimensions

Channel Expansion

Increases feature map channels without spatial processing

Cross-Channel Mixing

Learns relationships between different channels

Example: Dimensionality Reduction

256

Input Channels

→

1×1 Conv

64

Output Channels

✓

Efficient

Much cheaper than 3×3 or 5×5

Computational efficiency vs larger kernels

Network-in-Network: Adds depth without overhead

Universal Use: Inception · ResNet · MobileNet · Transformers (feedforward)