

Improvement of Cryo-EM Maps by Simultaneous Local and Non-local Deep Learning

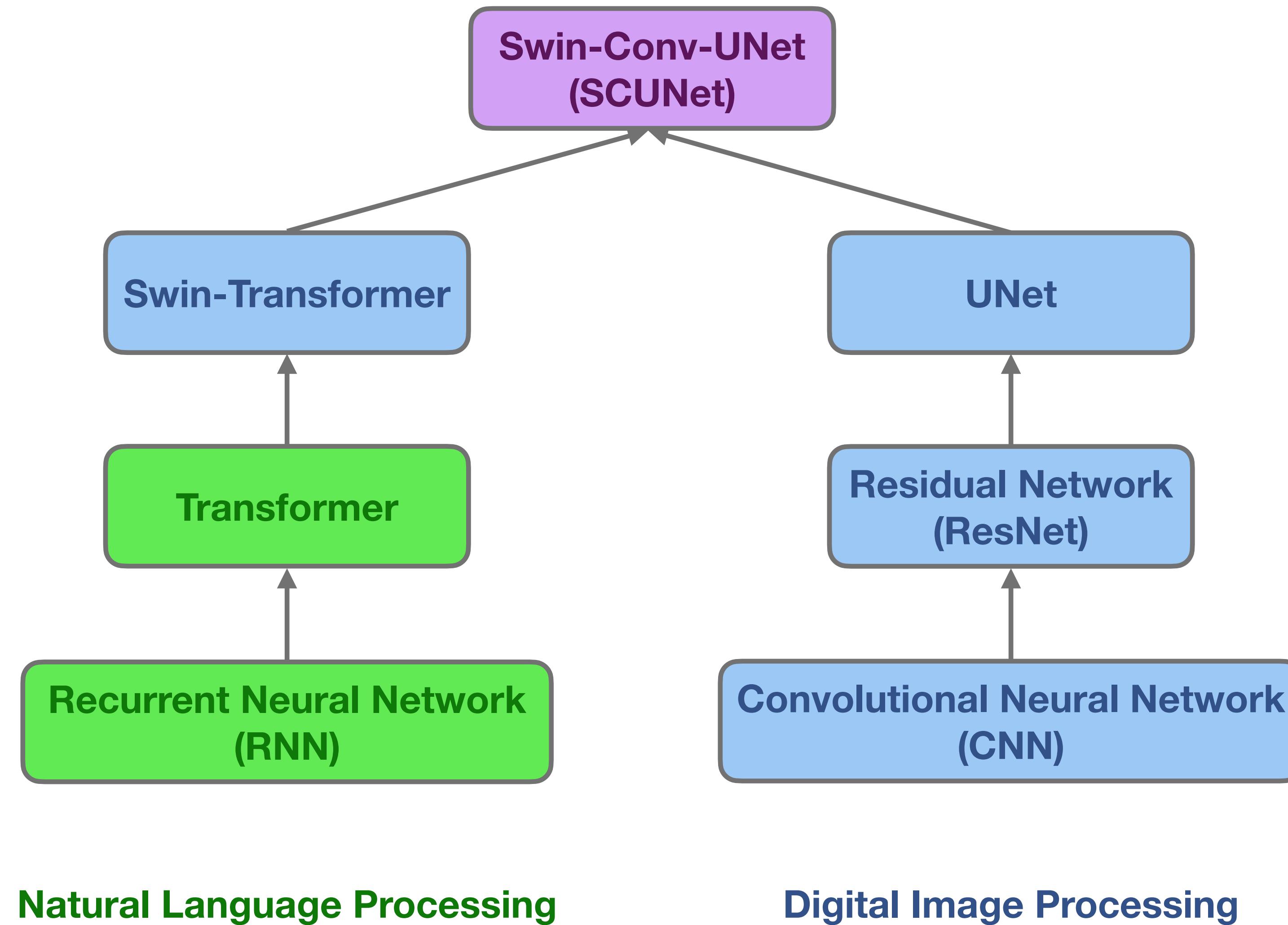
Group Meeting

Yu-Hsiang Lien 連昱翔

References

- 2023, Jiahua He, et. al., Improvement of cryo-EM maps by simultaneous local and non-local deep learning. <https://doi.org/10.1038/s41467-023-39031-1>
- 2022, Kai Zhang, et. al., Practical Blind Denoising via Swin-Conv-UNet and Data Synthesis. <https://doi.org/10.48550/arXiv.2203.13278>

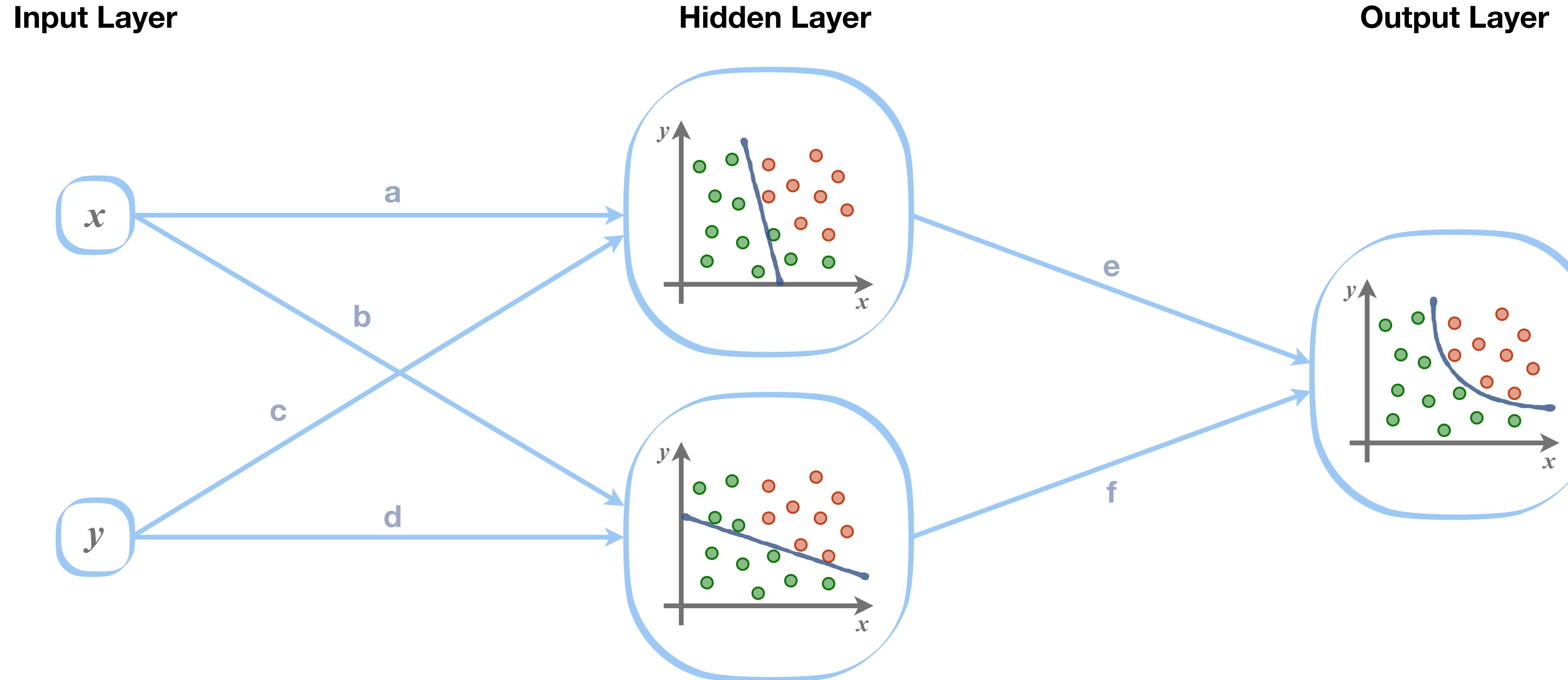
Variants of Deep Learning Model



Outline

- Deep learning model for Digital Image Processing
- Deep learning model for Natural Language Processing
- Swin-Conv-UNet (SCUNet) model
- EMReady — 3D SCUNet model for cryo-EM map improvement
- Test result of improvement of cryo-EM maps from this paper

Recap of Neuron Network



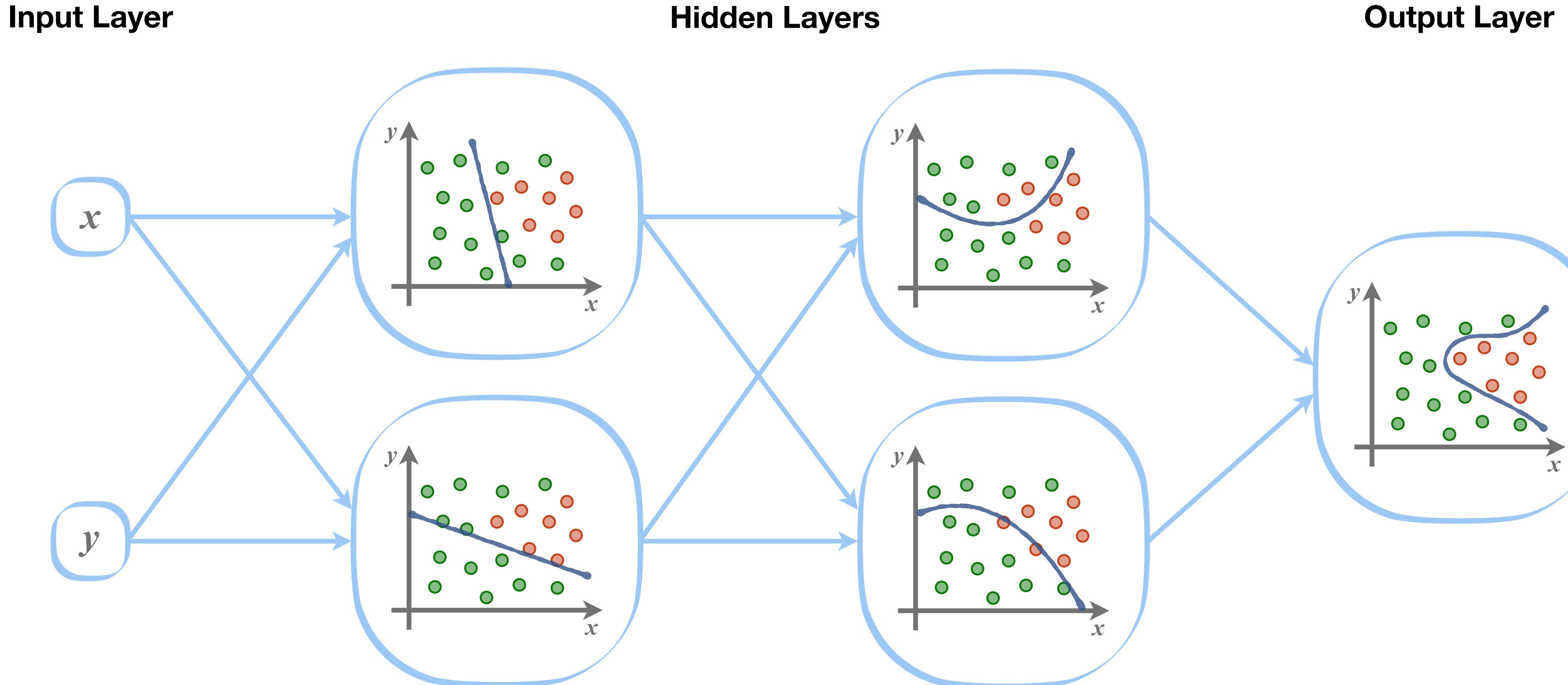
Training process – optimize parameters (a-f) :

Define error/loss function
(e.g. Least square)

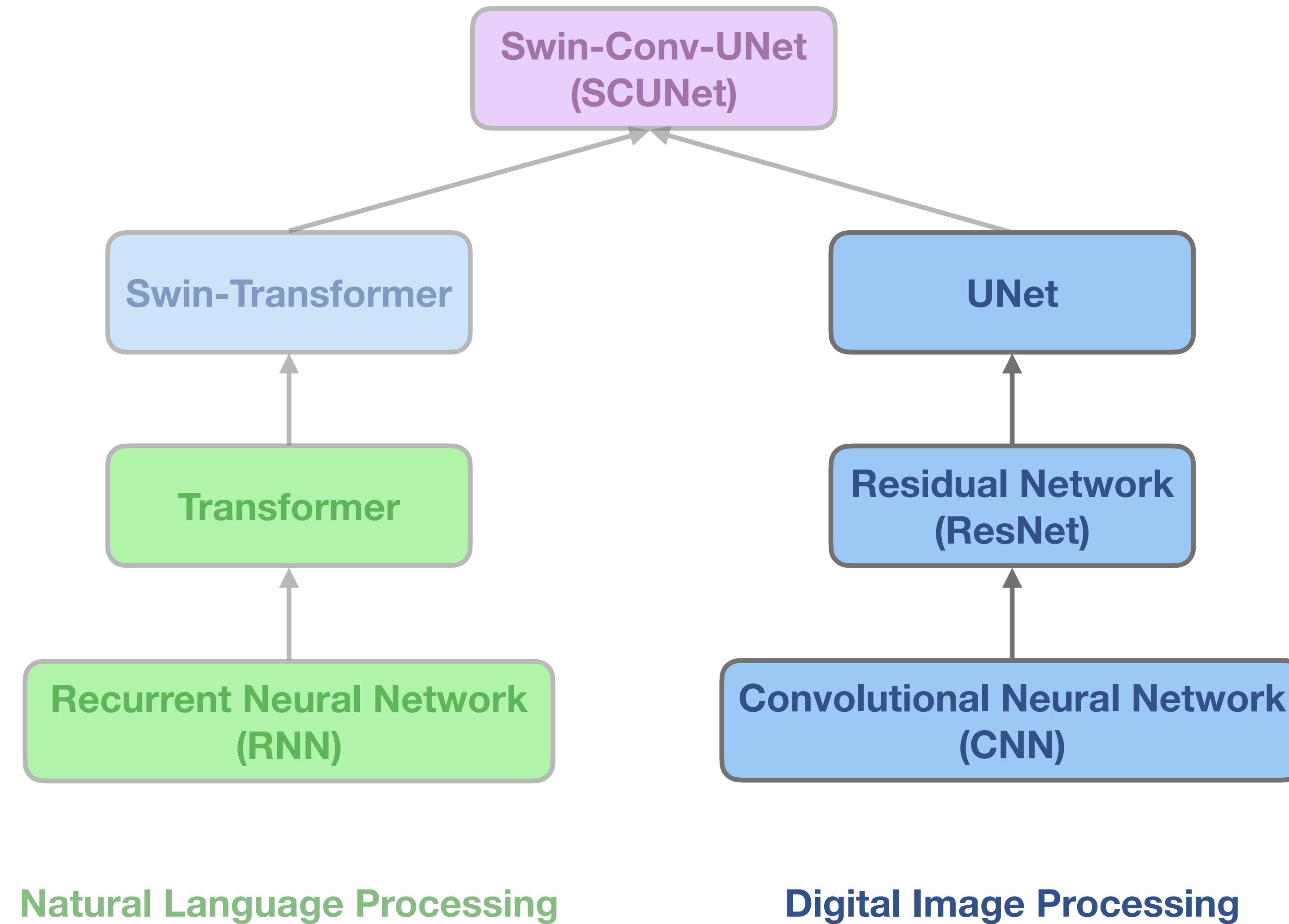


Minimize loss function
(e.g. Gradient descent)

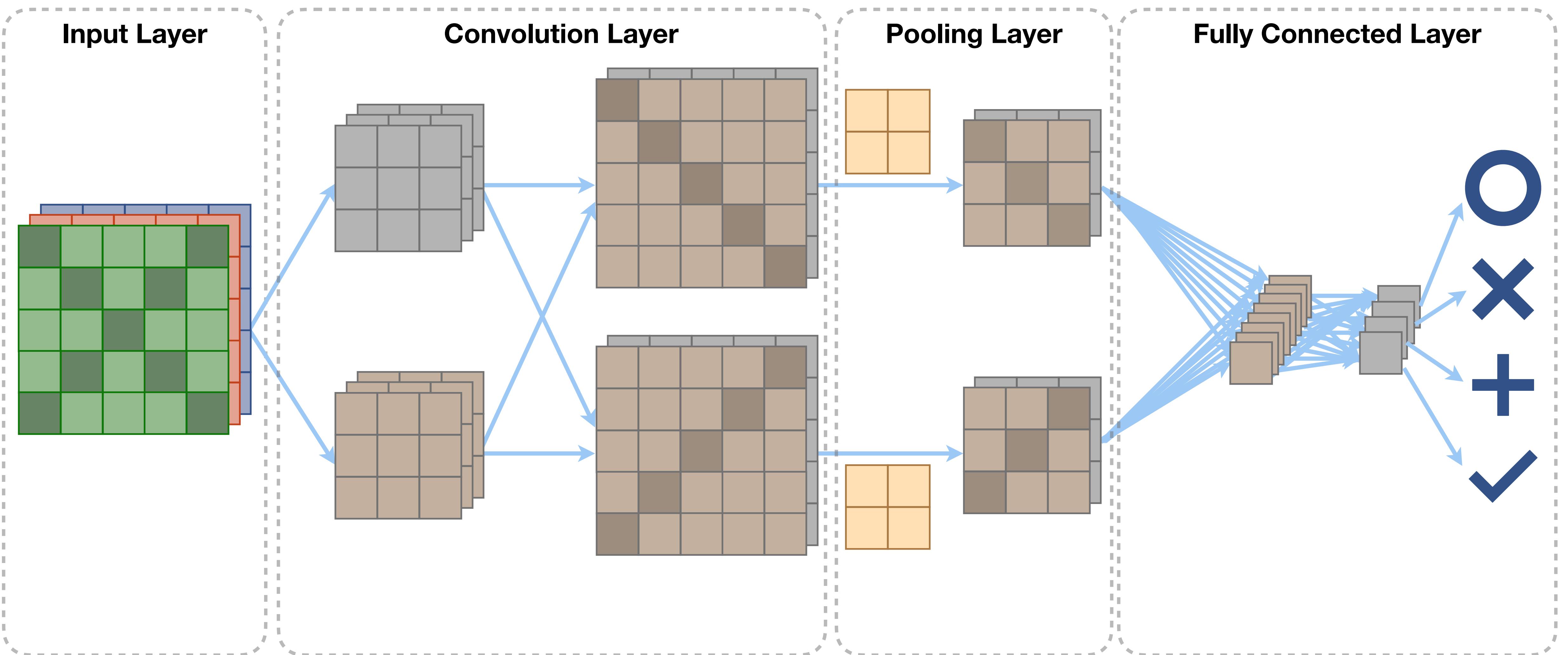
Recap of Deep Neuron Network



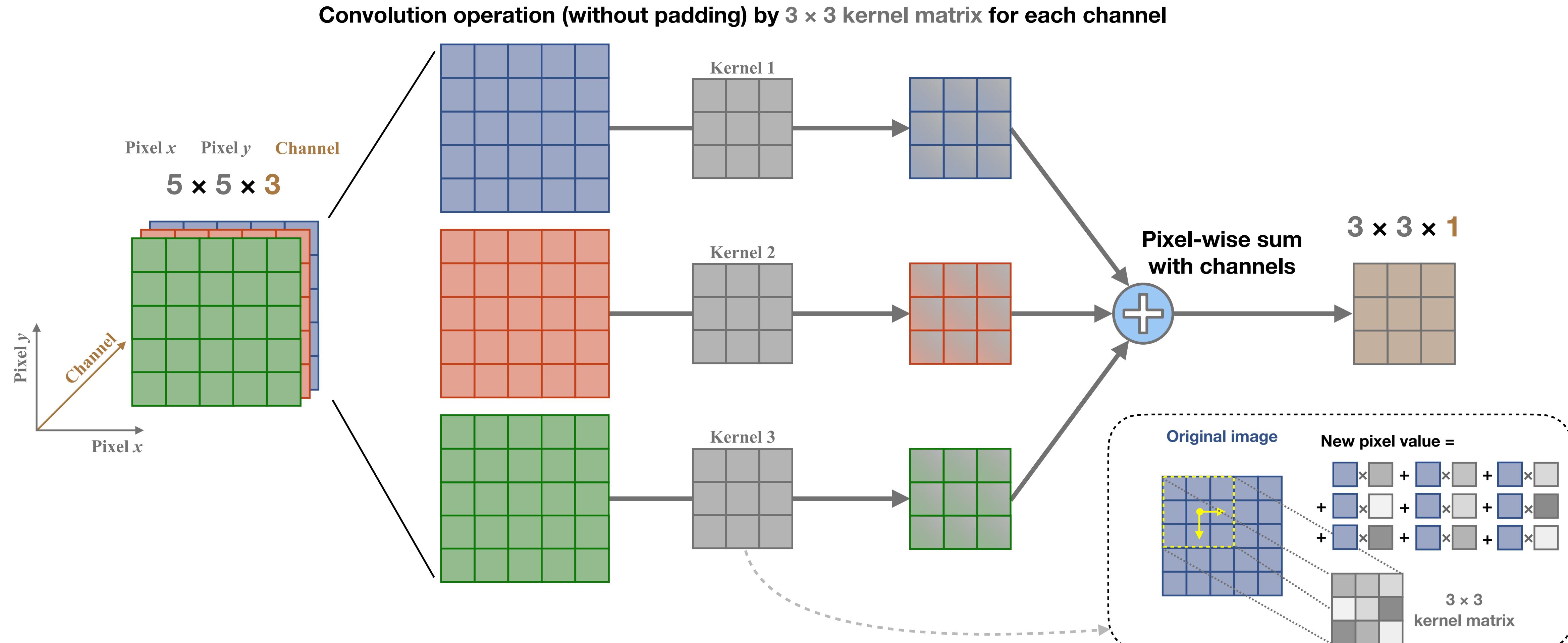
Deep Learning for Image Processing



Convolutional Neural Network (CNN)

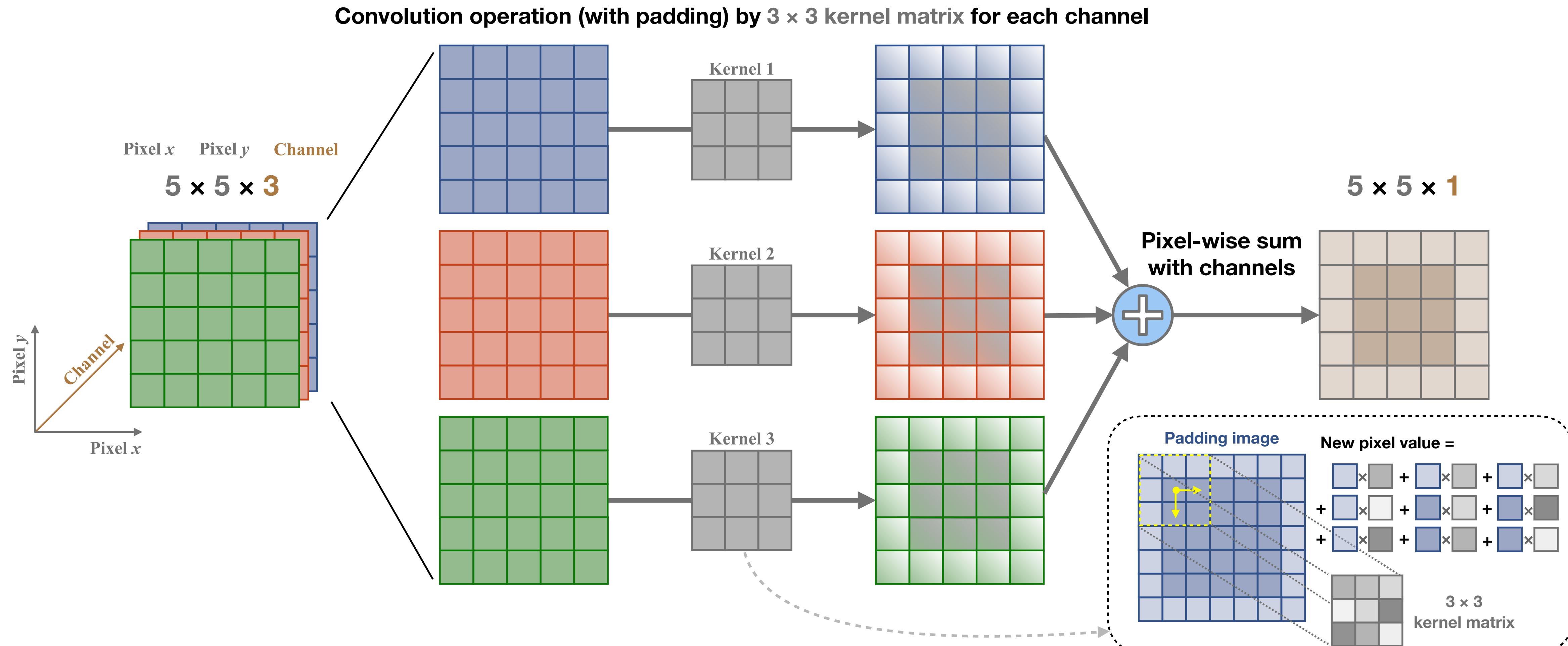


$N \times N$ Convolution Operation w/o Padding



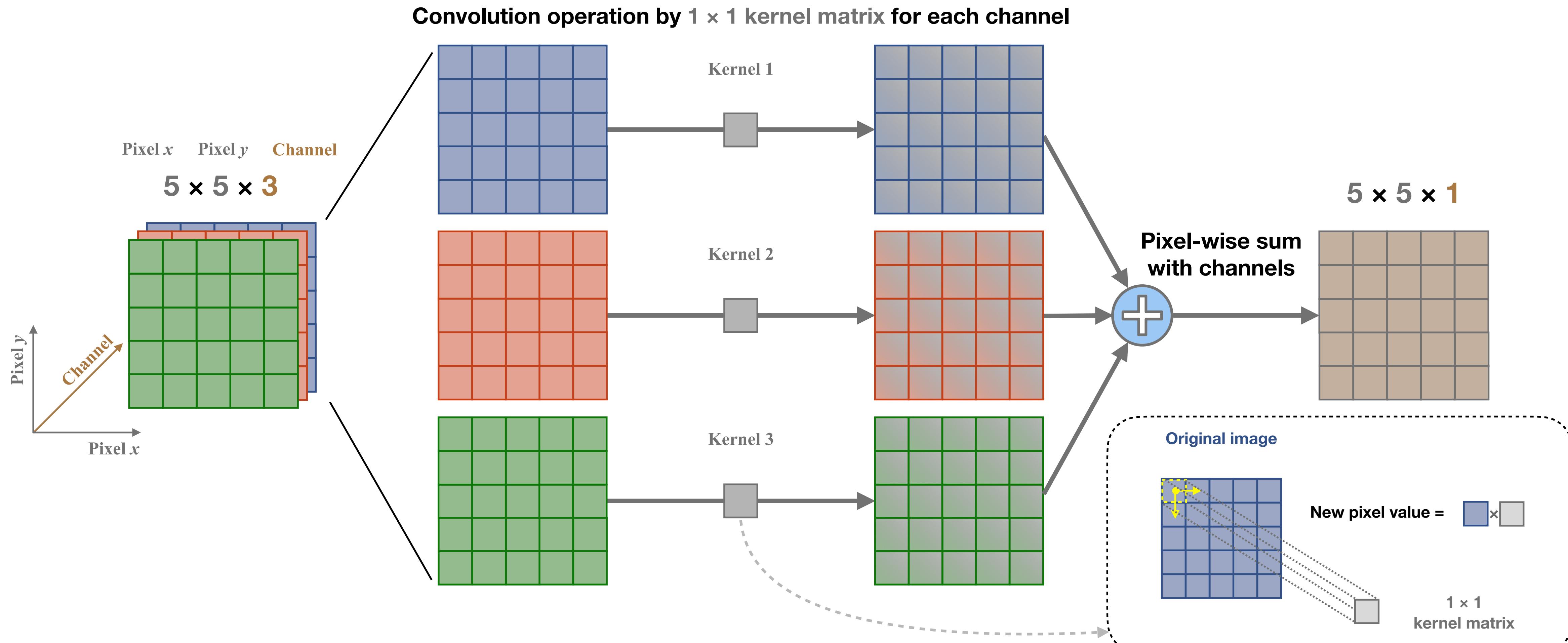
$N \times N$ Convolution operation without padding will merge image information **along channel dimension**, and also merging information along spatial dimension with defined kernel matrix size so that spatial dimensionality will be slightly reduced.

$N \times N$ Convolution Operation w/ Padding



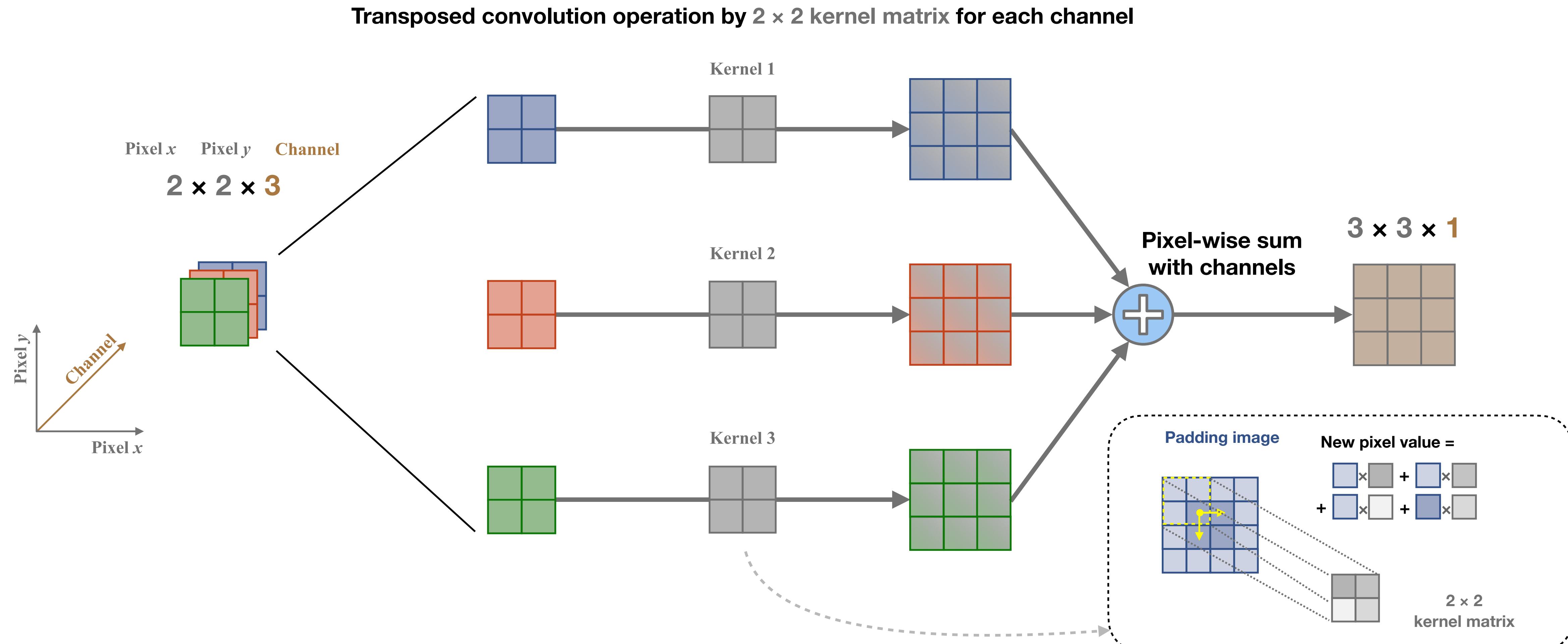
$N \times N$ Convolution operation with padding will merge image information **along channel dimension**, and also merge information along spatial dimension with defined kernel matrix size but hold the same spatial dimensionality (only if operation **stride** set to 1).

Special Case: 1×1 Convolution Operation



1×1 Convolution operation will merge image information **along channel dimension**, but hold the same spatial dimensionality of image.

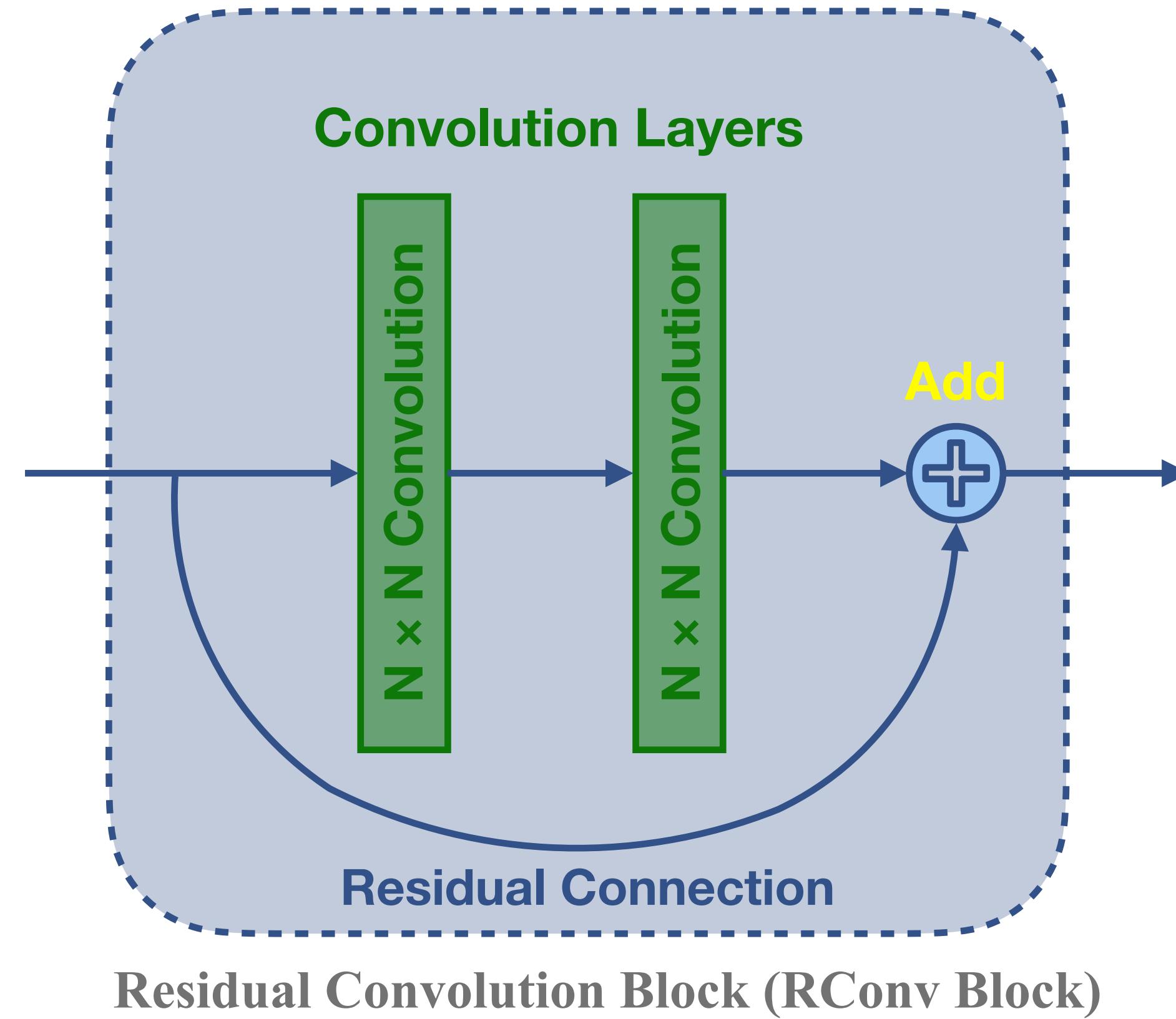
$N \times N$ Transposed Convolution Operation



$N \times N$ transposed convolution operation will merge image information **along channel dimension**, and also extend information **along spatial dimension** with defined kernel matrix size so that spatial dimensionality will be increased.

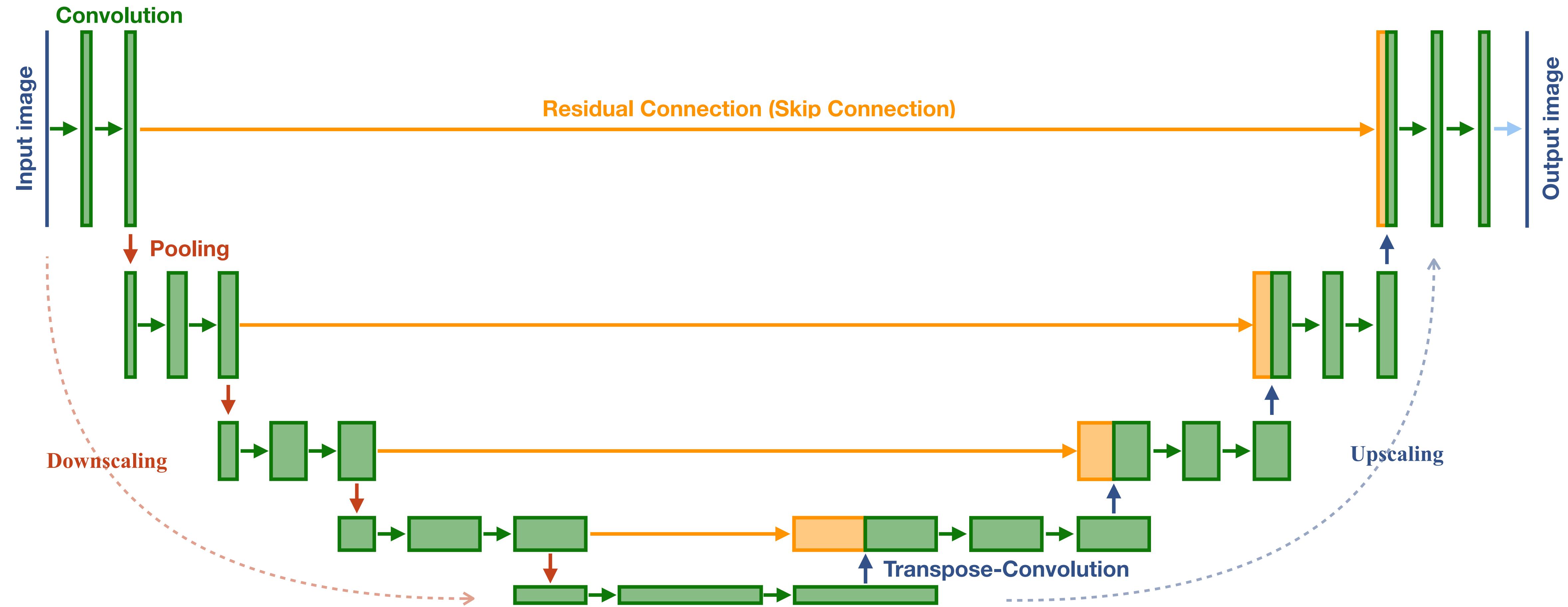
Residual Network (ResNet)

Ref: "Deep Residual Learning for Image Recognition", 2015, <https://doi.org/10.48550/arXiv.1512.03385>



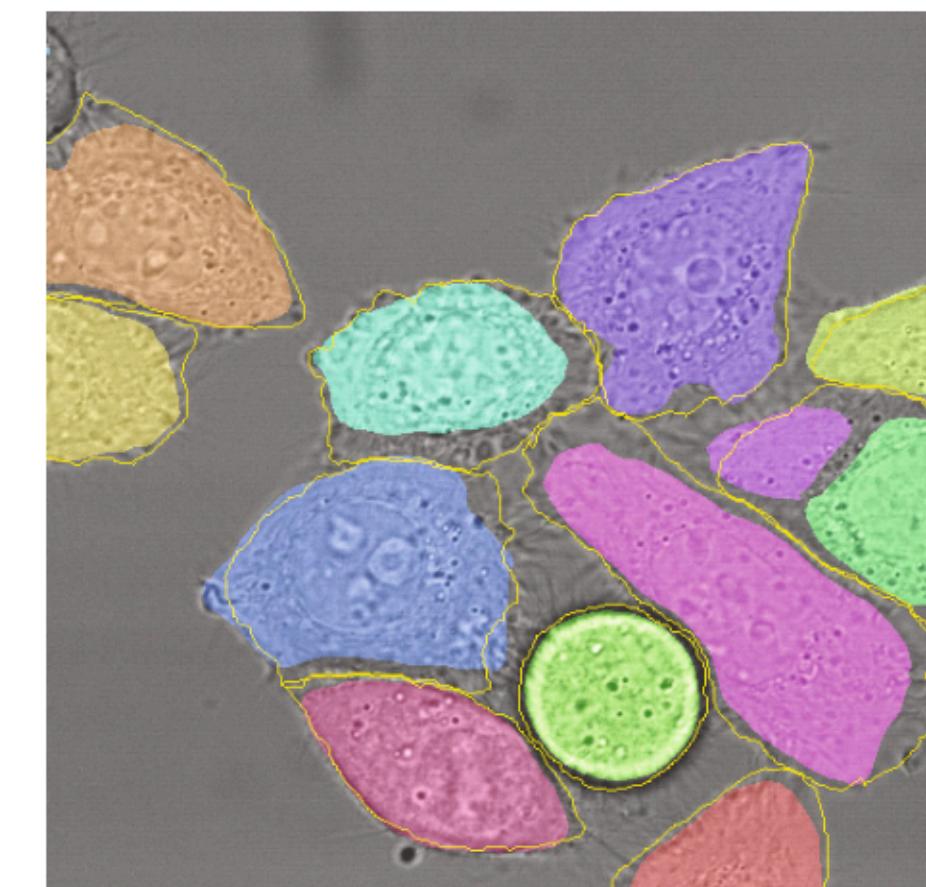
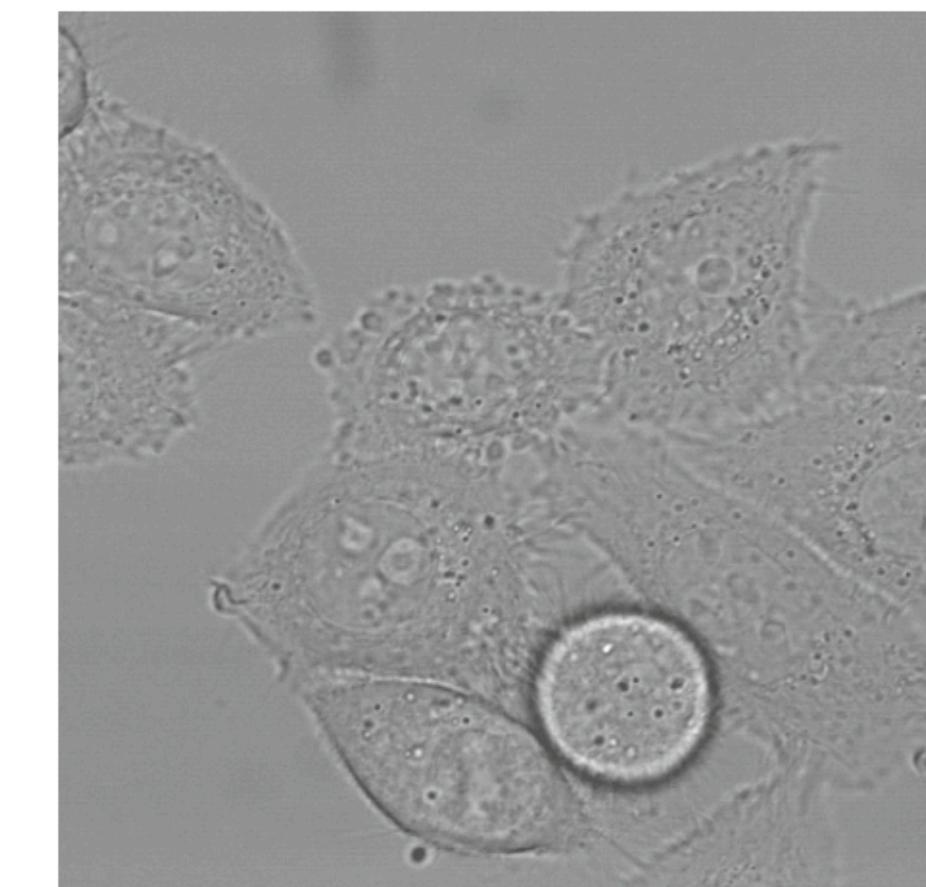
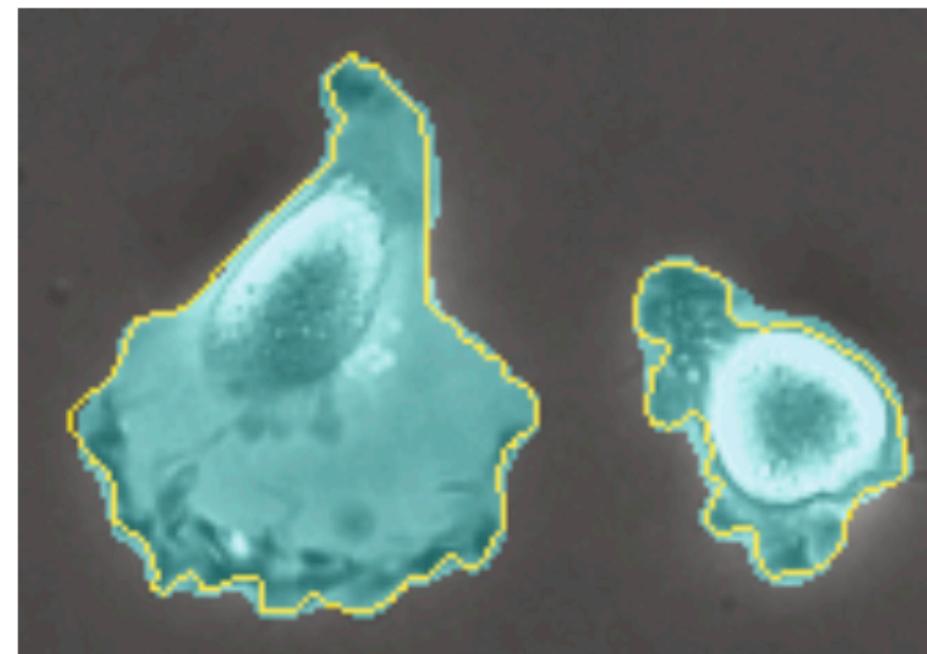
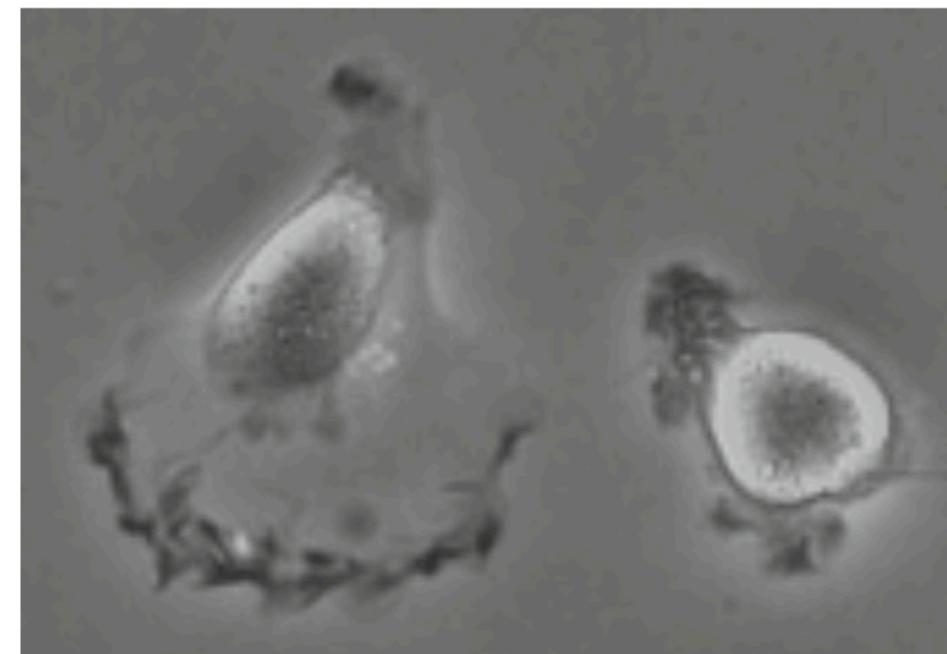
By introducing residual convolution block, the training of network in **deeper CNN model** become more easy.

UNet Architecture



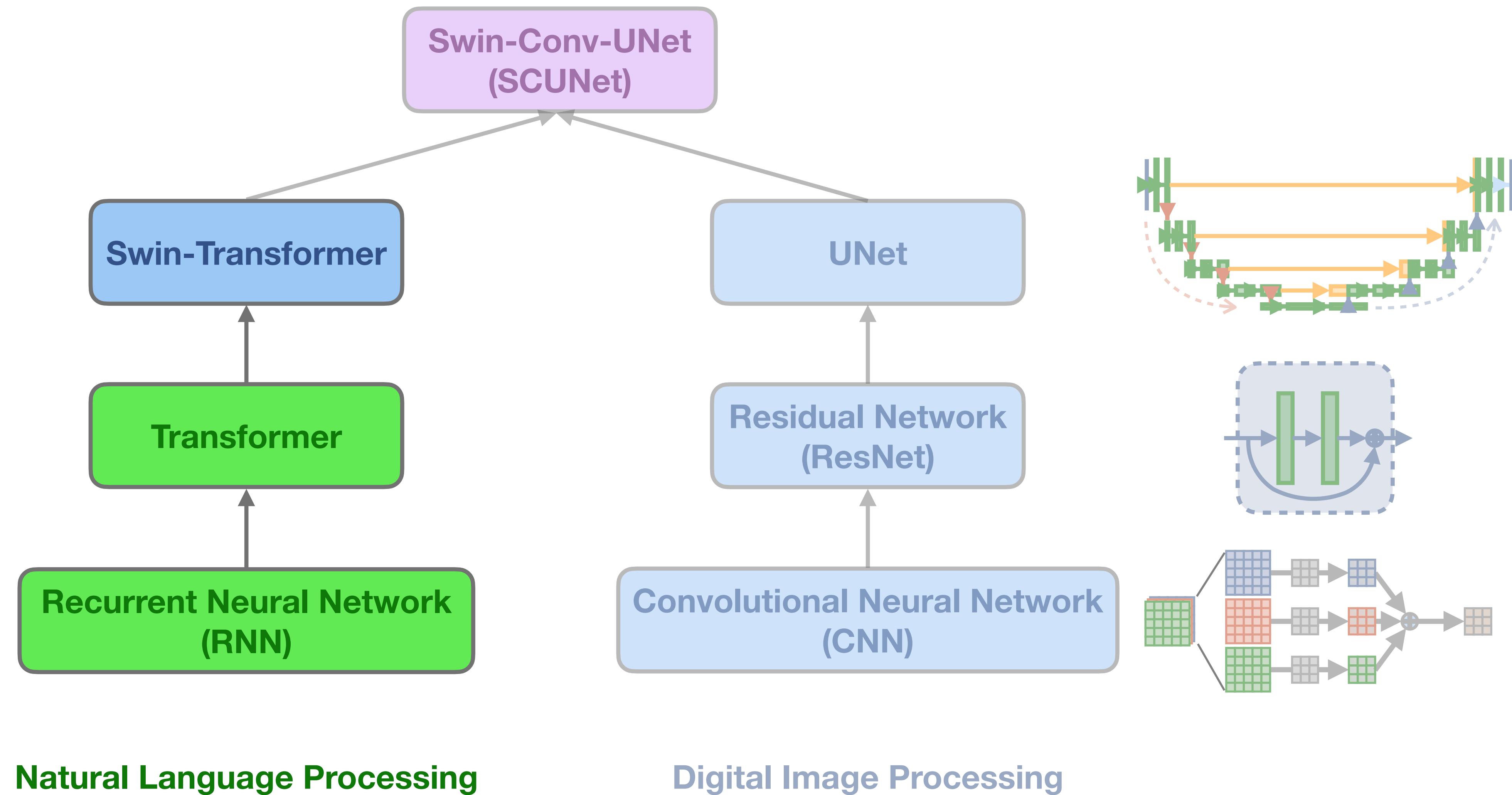
By adopting transposed convolution for upscaling and residual connection from original features information, UNet architecture achieve to output target image.

Image-to-Image Model — UNet

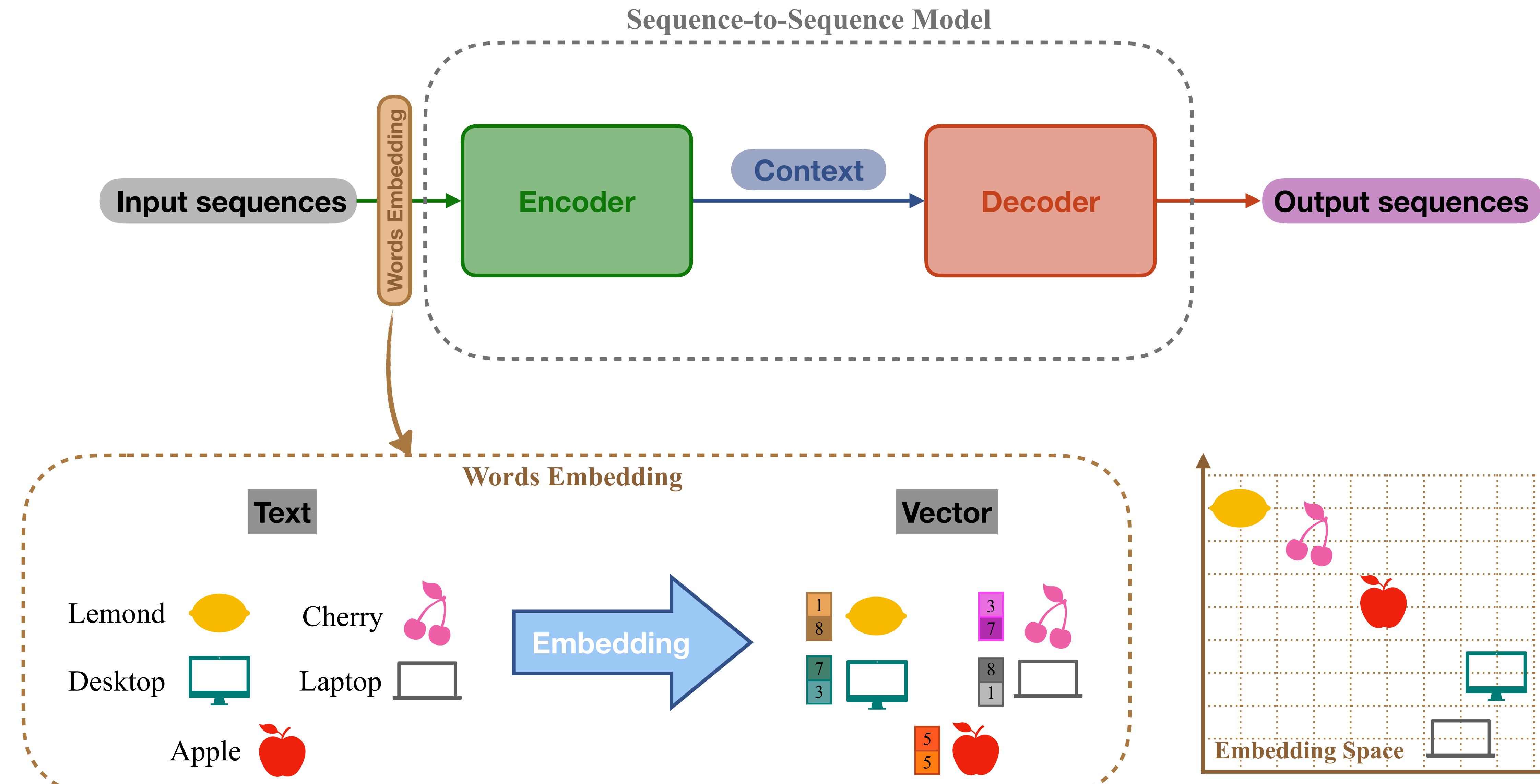


- The original paper which proposed UNet model is apply for biomedical image segmentation task.

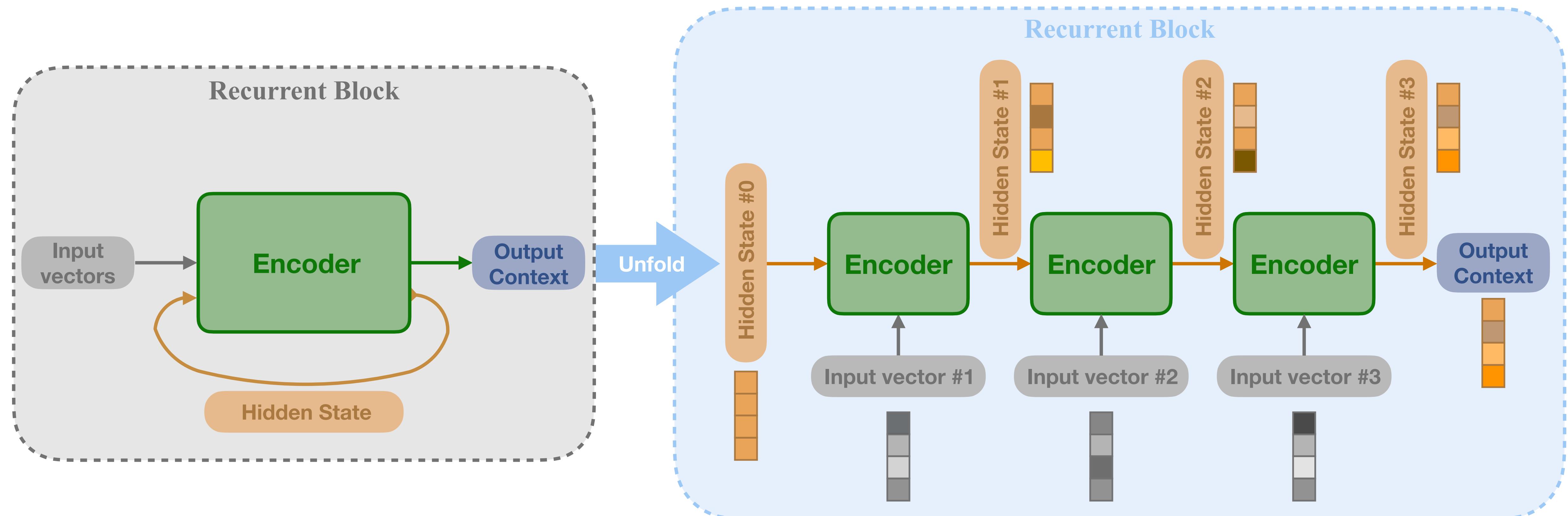
Deep Learning for Natural Language Processing



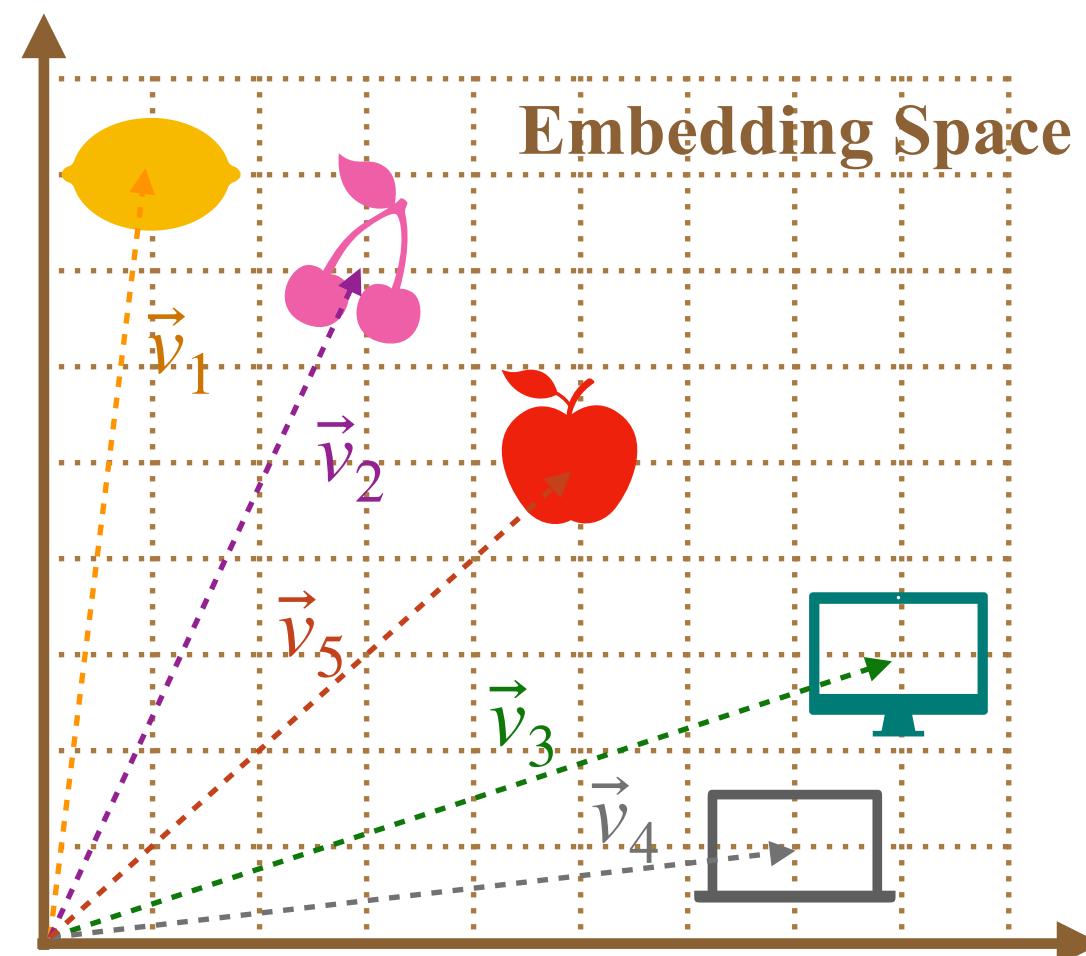
Sequence-to-Sequence Model



Recurrent Neural Network (RNN)



Attention Mechanism



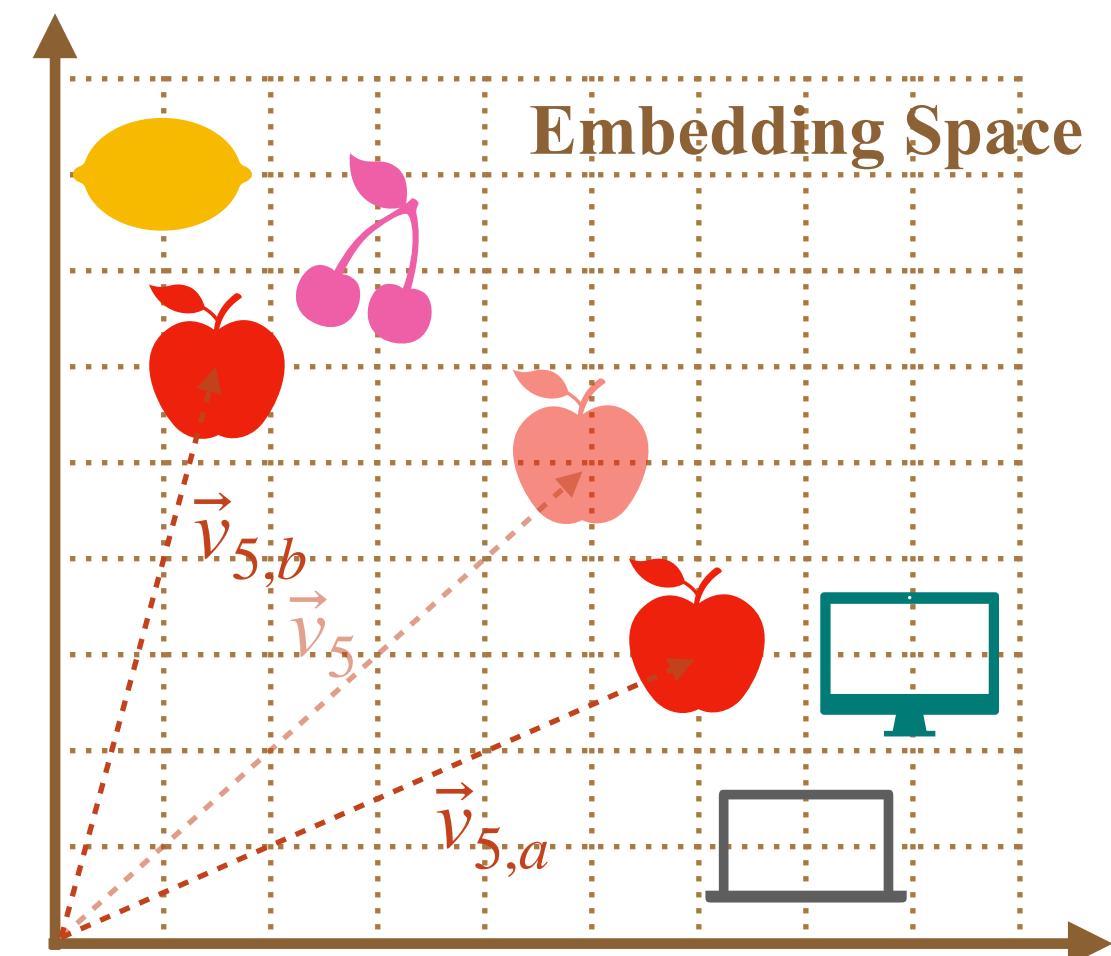
An **apple** laptop
 \vec{v}_5

An **apple** and a **lemon**
 \vec{v}_5

Calculate similarity of words – Scaled dot product

$$\text{Similarity}(i, j) = \frac{1}{\sqrt{L}} \mathbf{v}_i^T \mathbf{v}_j$$

Weighting word's vector by similarity

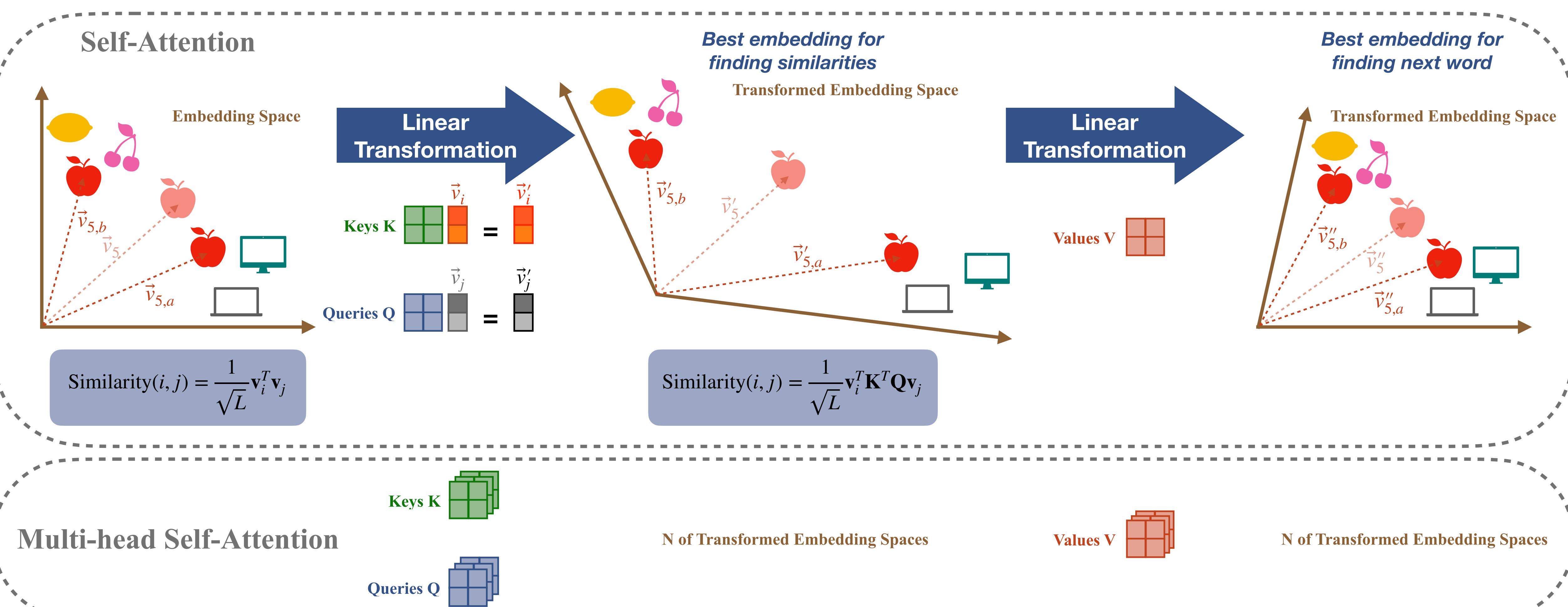


An **apple** laptop
 $\vec{v}_5 \rightarrow \vec{v}_{5,a}$

An **apple** and a **lemon**
 $\vec{v}_5 \rightarrow \vec{v}_{5,b}$

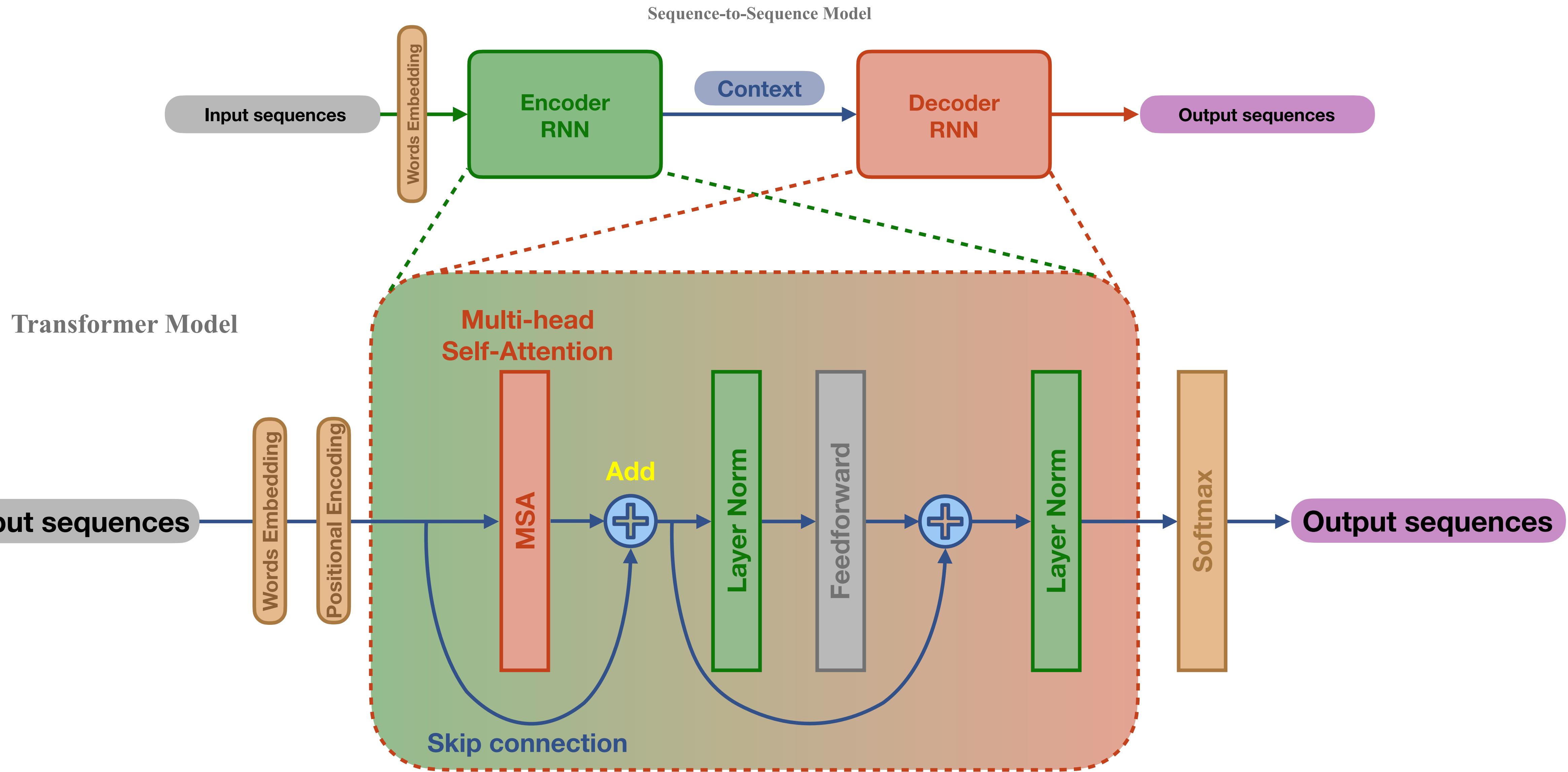
Attention step is trying to modify each **word's coordinate** in embedding space base on **similarity** between words.

Self-Attention & Multi-head Self-Attention

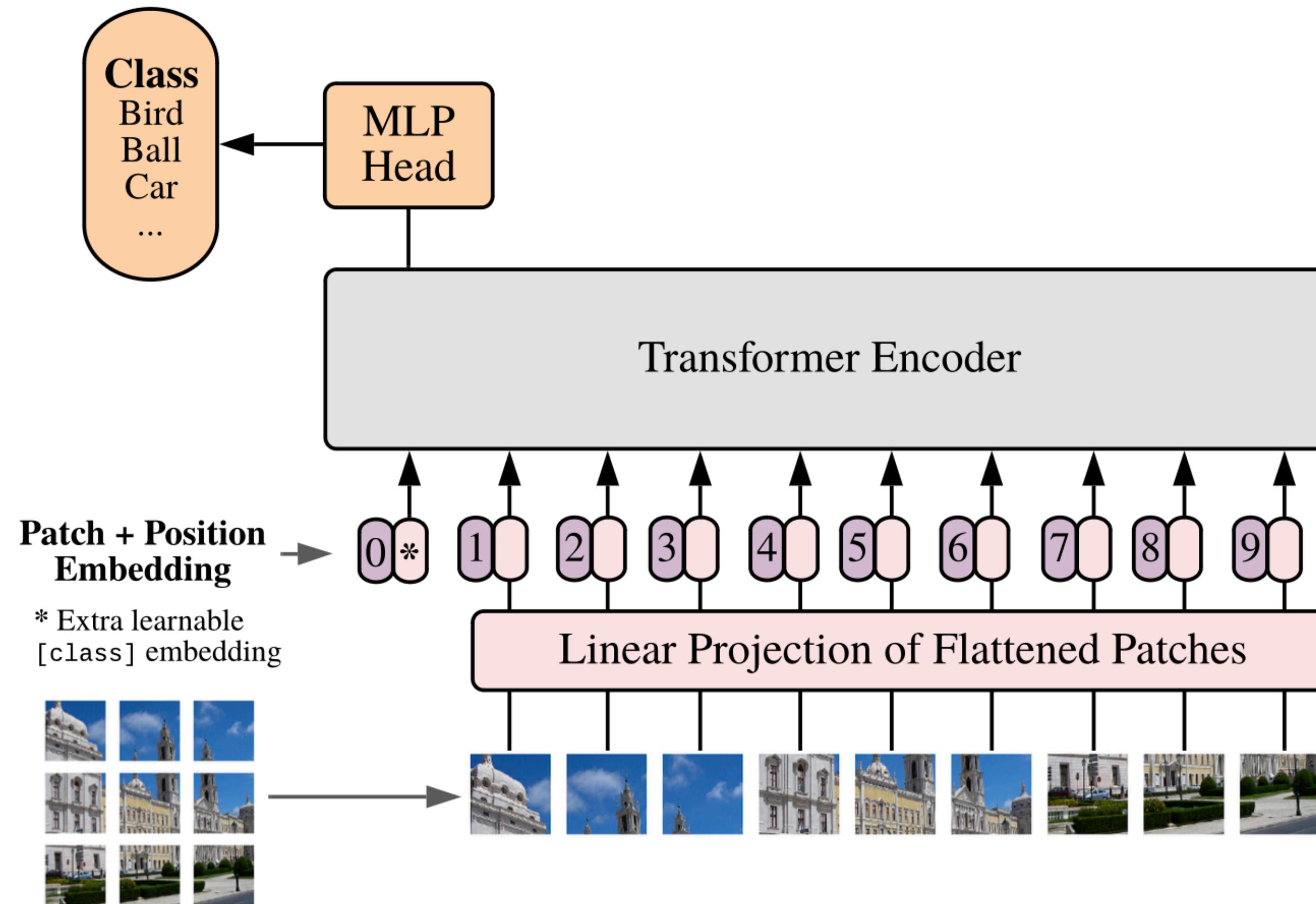


Self-attention is trying to find two **new embedding spaces** for improving attention step, one embedding for finding similarities by **keys** and **queries** matrices, another embedding for finding next word by **values** matrix.

Transformer

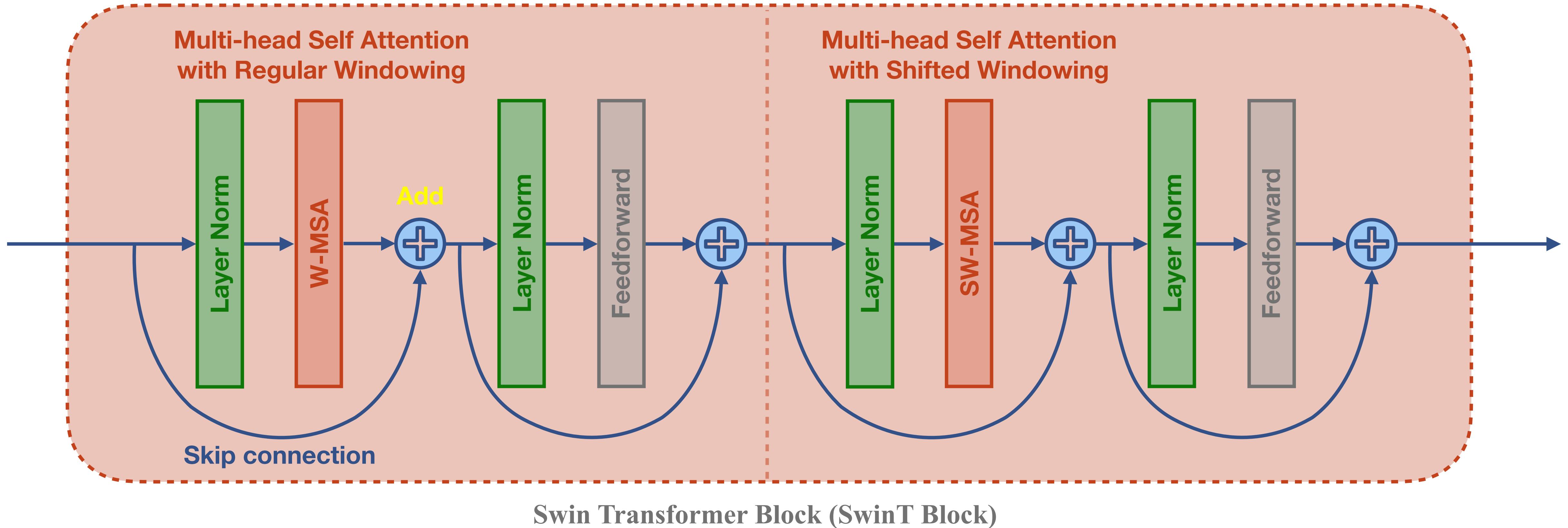


Vision Transformer



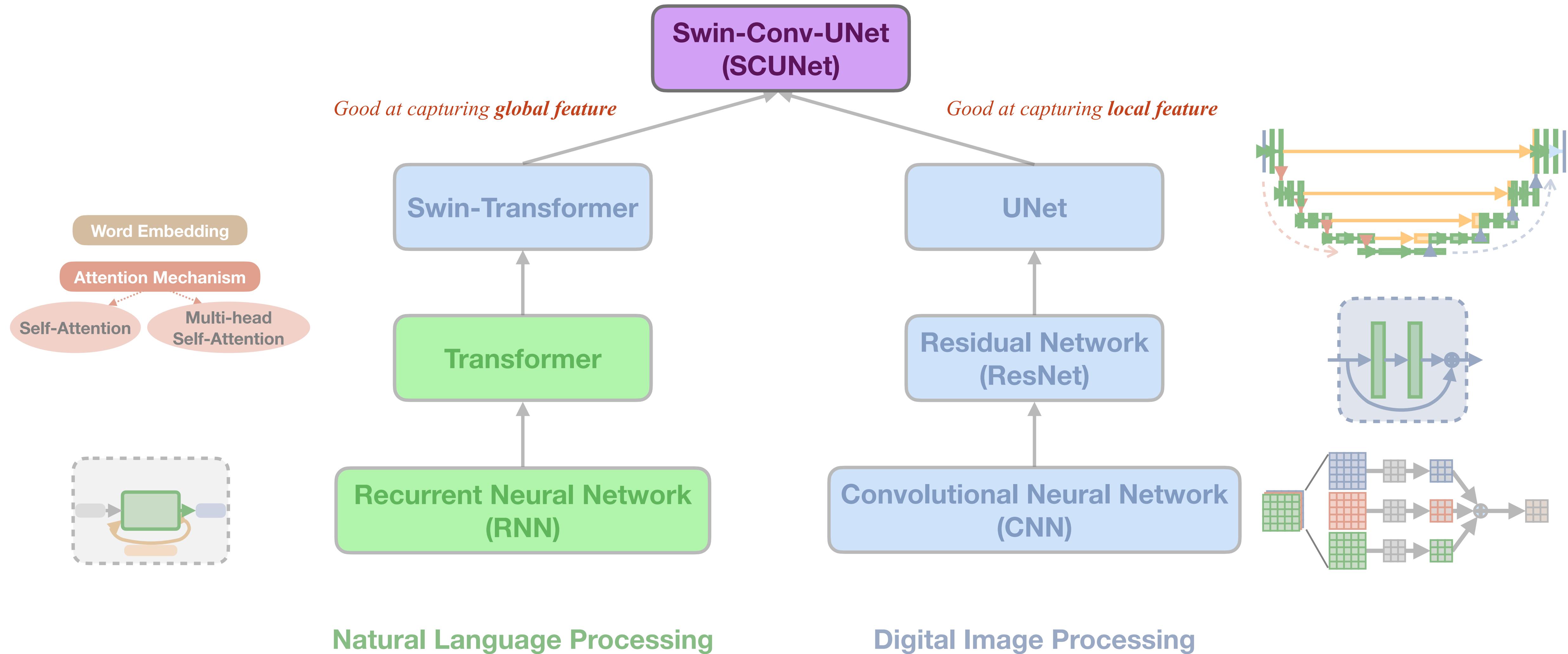
Vision Transformer using transformer's architecture for image input, image cut into several **patches** and treat as sequences for attention step.

Swin Transformer



Swin Transformer improve the calculation efficiency from Vision Transformer by defining **window** in order to limit the attention calculation in certain range.

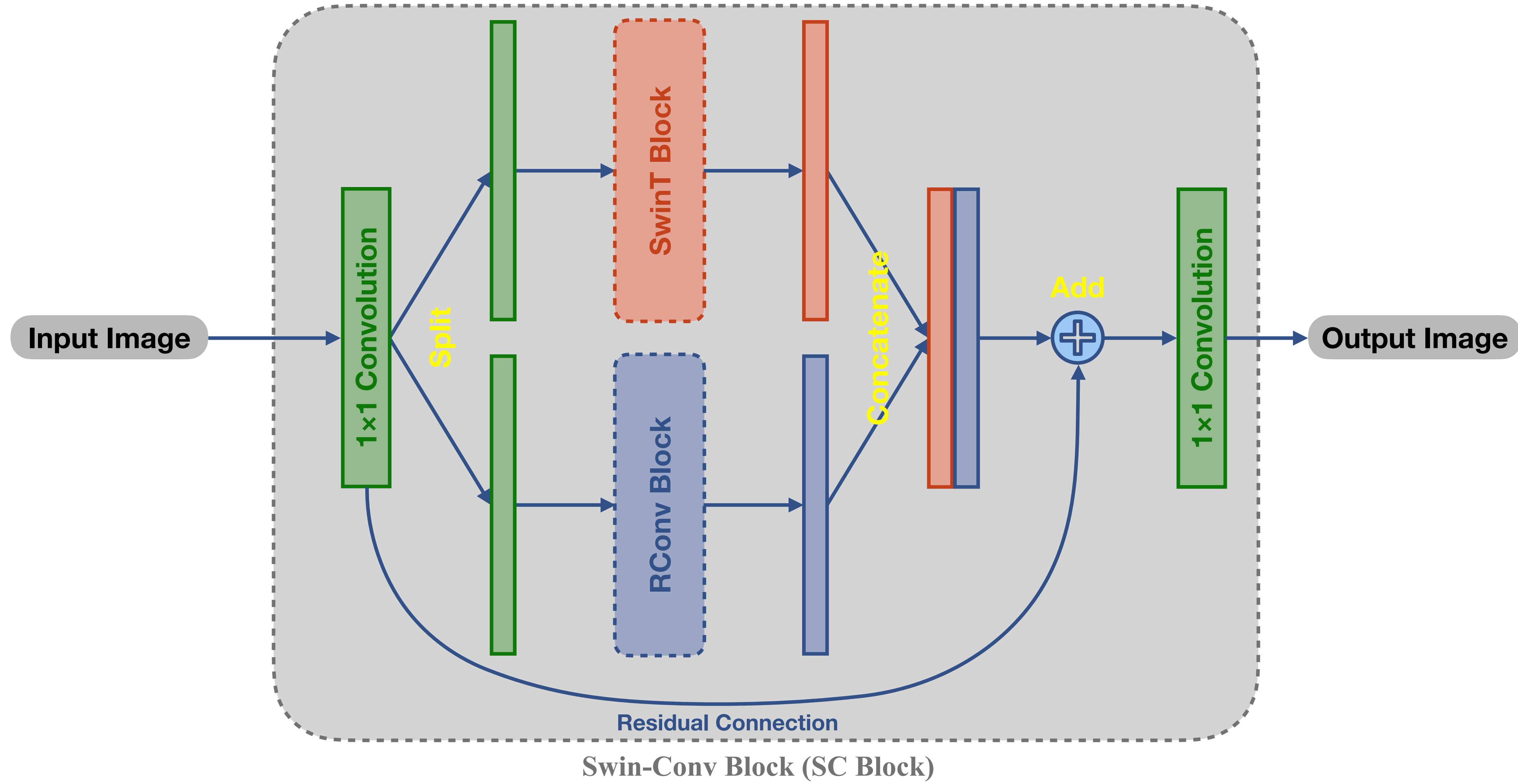
Deep Learning Model — SCUNet



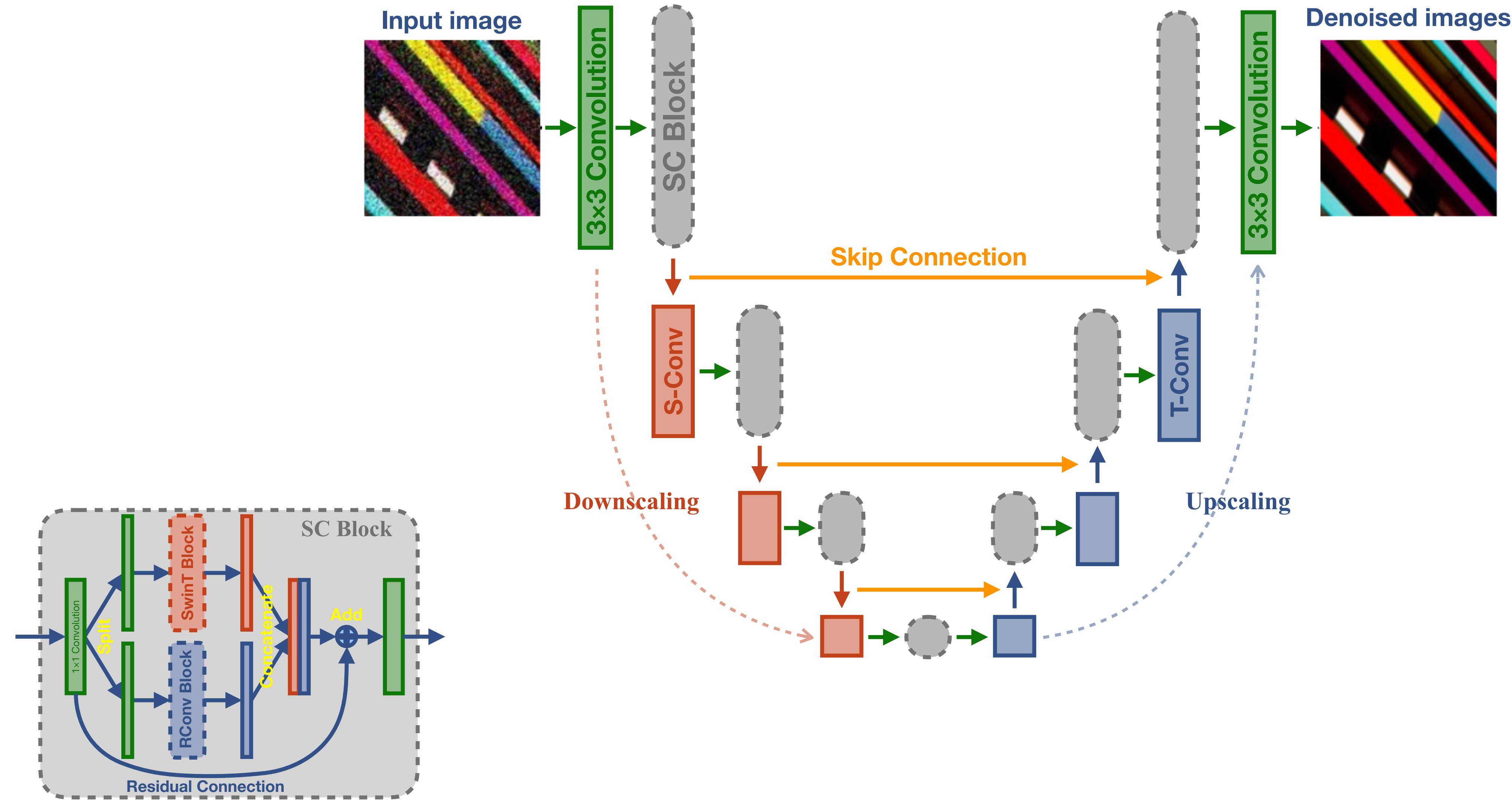
Natural Language Processing

Digital Image Processing

Swin-Conv Block (SC Block)



Swin-Conv-UNet (SCUNet) Architecture



Example of SCUNet on Image Denoising Task

Noisy



SCUNet
Process





Article

<https://doi.org/10.1038/s41467-023-39031-1>

Improvement of cryo-EM maps by simultaneous local and non-local deep learning

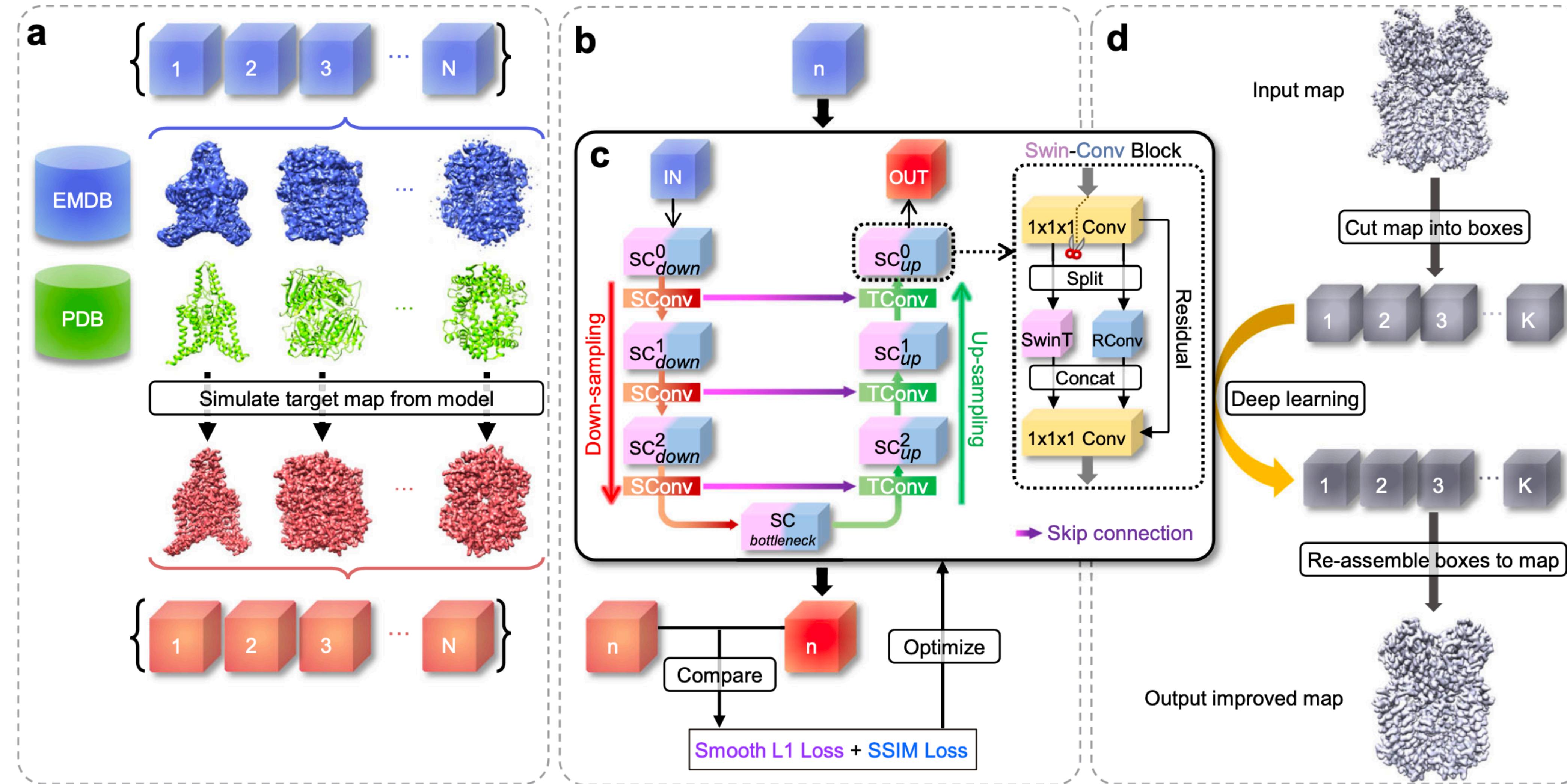
3D Swin-Conv-UNet-based Deep Learning Framework (EMReady)

Received: 3 August 2022

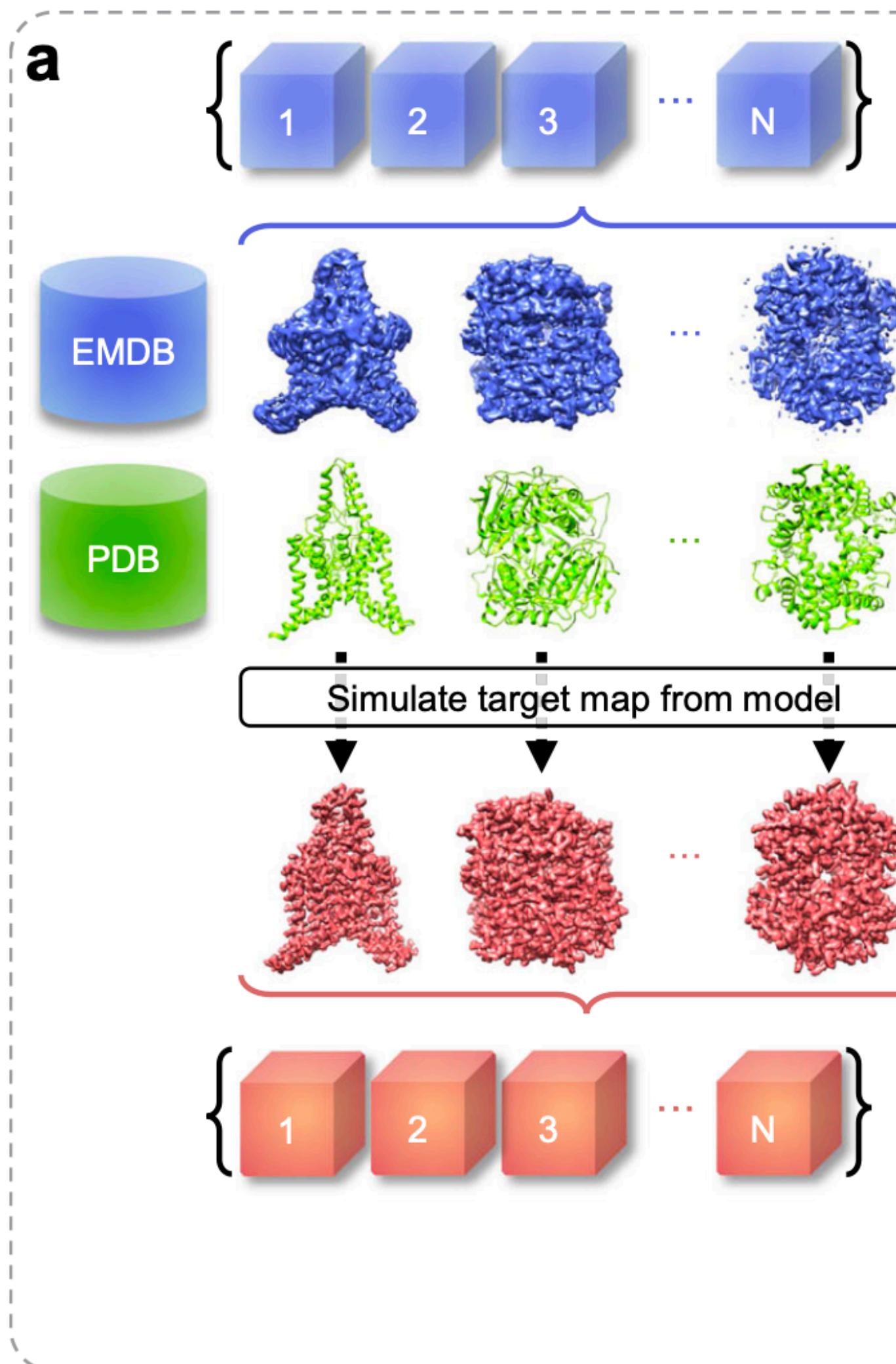
Jiahua He ¹, Tao Li ¹ & Sheng-You Huang ¹

Accepted: 25 May 2023

Overview of EMReady



Training Data Set of EMReady



Cryo-EM Density Map from Electron Microscopy Data Bank (EMDB)

The grid size of the maps is unified to 1Å by a cubic interpolation

Collect 436 pairs of cryo-EM density maps and associated PDB structures with resolutions ranging from 3Å to 6Å

(256 maps used for **training set**, 70 maps used for **validation set**, 110 maps used for **testing set**)

Associated Atomic Structure from Protein Data Bank (PDB)

Simulated Target Density Map from Atomic Structure

$$\rho(\mathbf{x}) = \sum_i^M \theta Z_i e^{-k|\mathbf{x}-\mathbf{r}_i|^2}$$

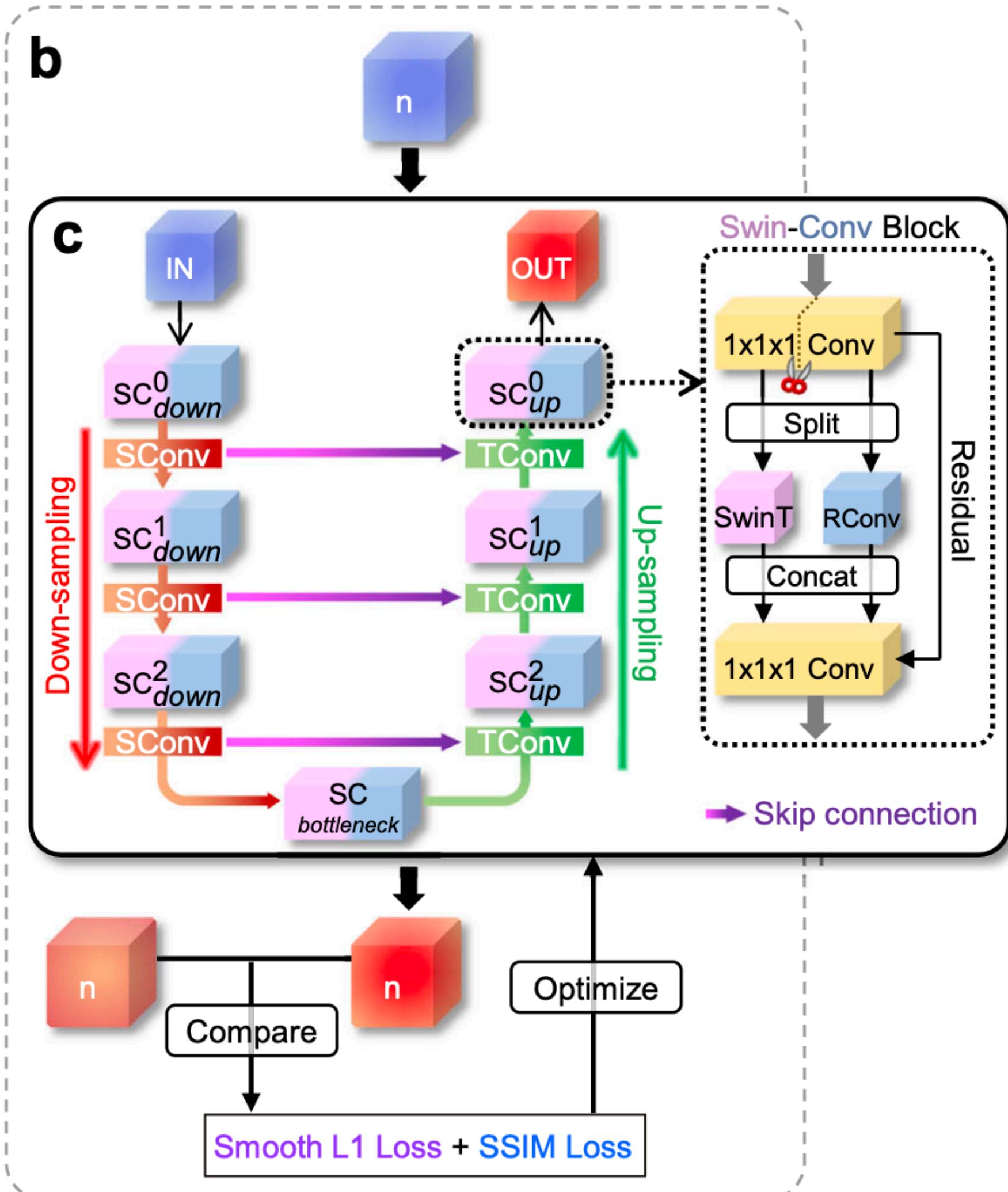
$$k = \left(\frac{\pi}{1.2 + 0.6R} \right)^2$$

$$\theta = \left(\frac{k}{\pi} \right)^{3/2}$$

Cut into several pairs of boxes

Size of box = 48×48×48 voxels with 1Å grid size

Training Procedure of EMReady



- The loss function for optimization = **Smooth L1 Loss** + **SSIM Loss**

(Local)

(Global)

- Smooth L1 Loss:**

$$\text{SmoothL1Loss}(X, Y) = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{l_{i,j,k}}{N^3}$$

$l_{i,j,k}$: Smooth L1 distance

N : Slice size of box

- Structural Similarity (SSIM) Loss:**

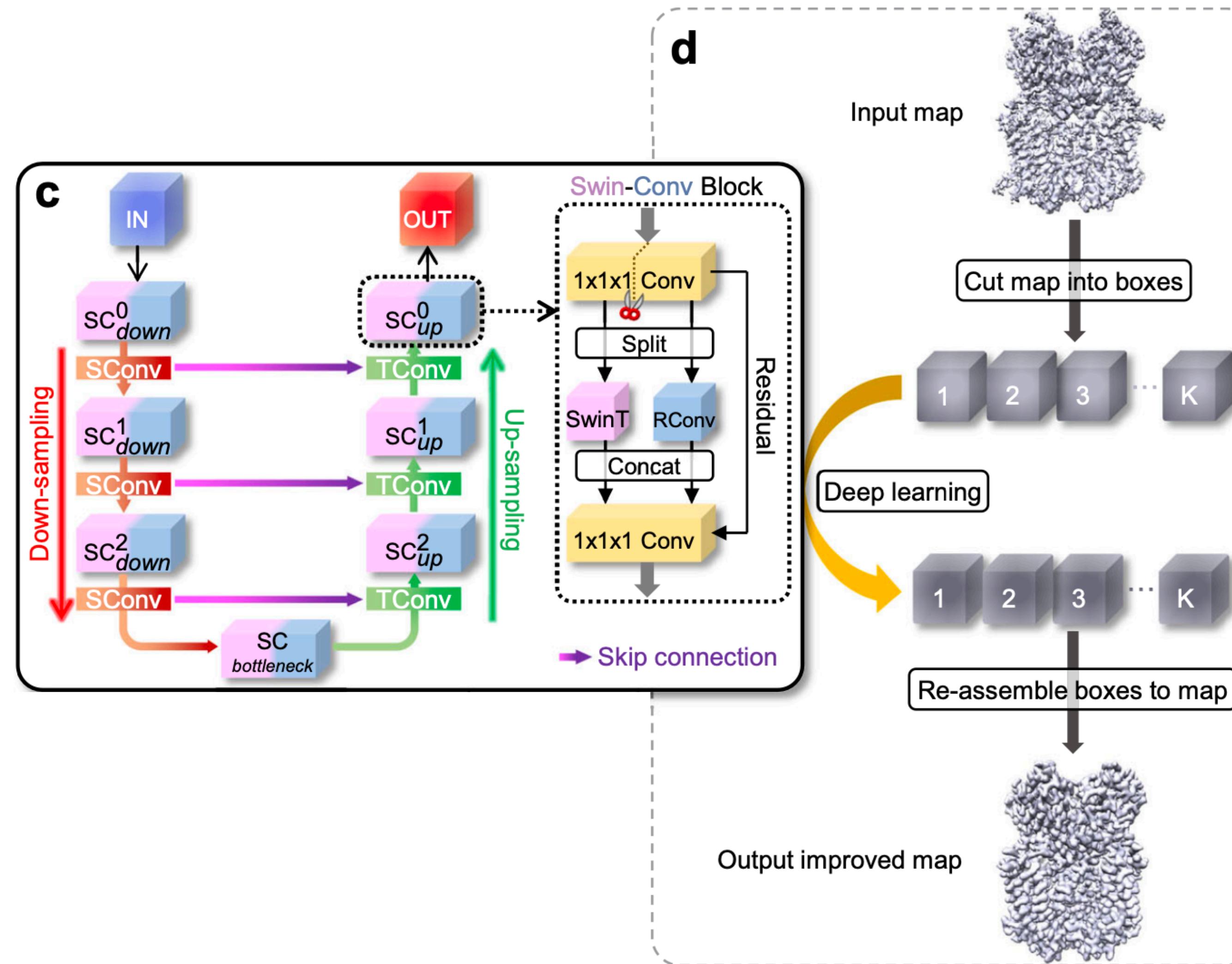
$$\text{SSIMLoss}(X, Y) = 1 - c(X, Y) \cdot s(X, Y) = 1 - \frac{2\sigma_{XY} + \epsilon}{\sigma_X^2 + \sigma_Y^2 + \epsilon}$$

$c(X, Y)$: Contrast similarity

$s(X, Y)$: Structure similarity

ϵ : Small constant to prevent dividing by zero

The Map Processing Workflow of EMReady



- For the map evaluation:
 - A given cryo-EM density map is cut into **K** of small boxes.
 - Each box is processed by **EMReady**.
 - The output processed boxes are re-assembled into one final density map.

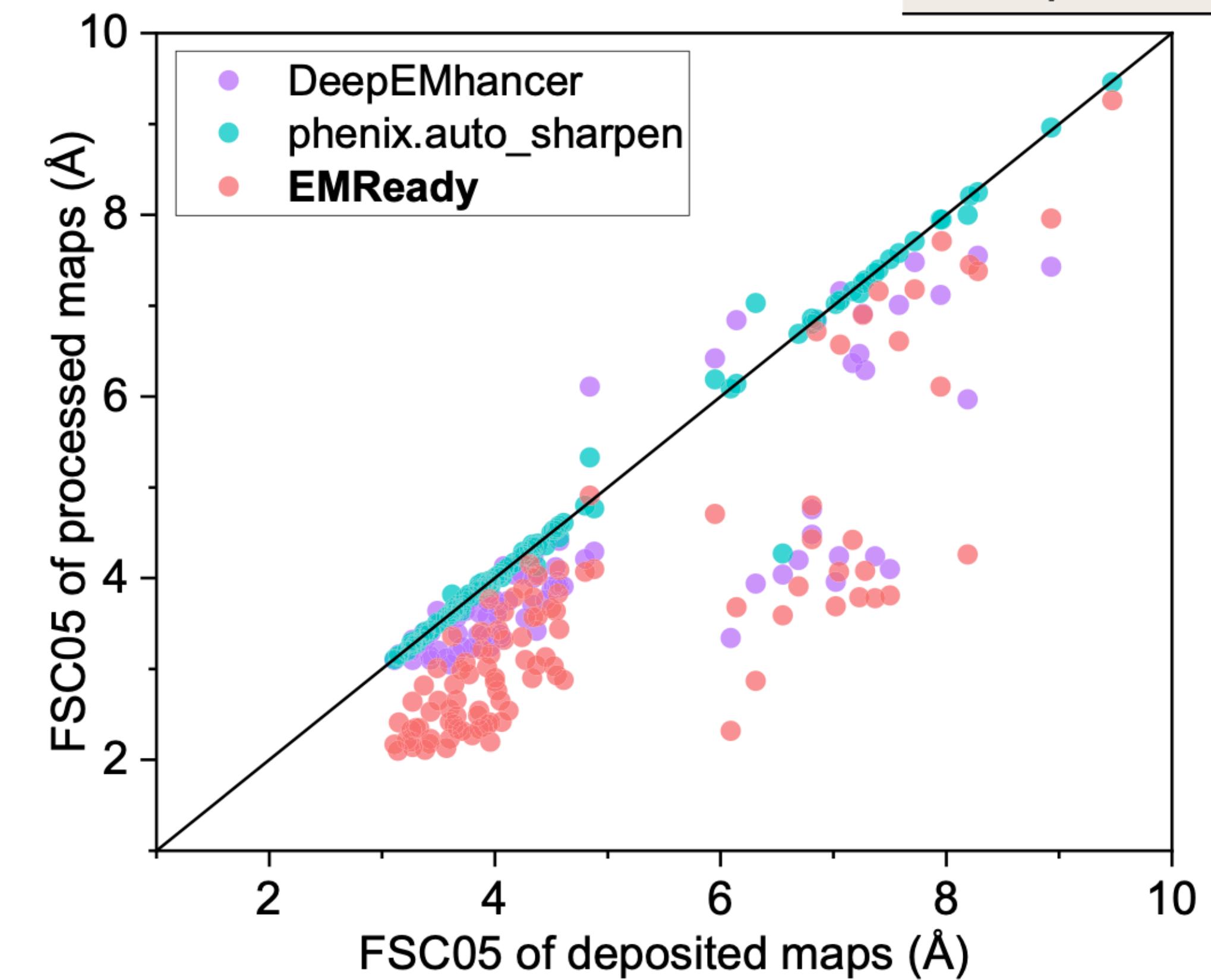
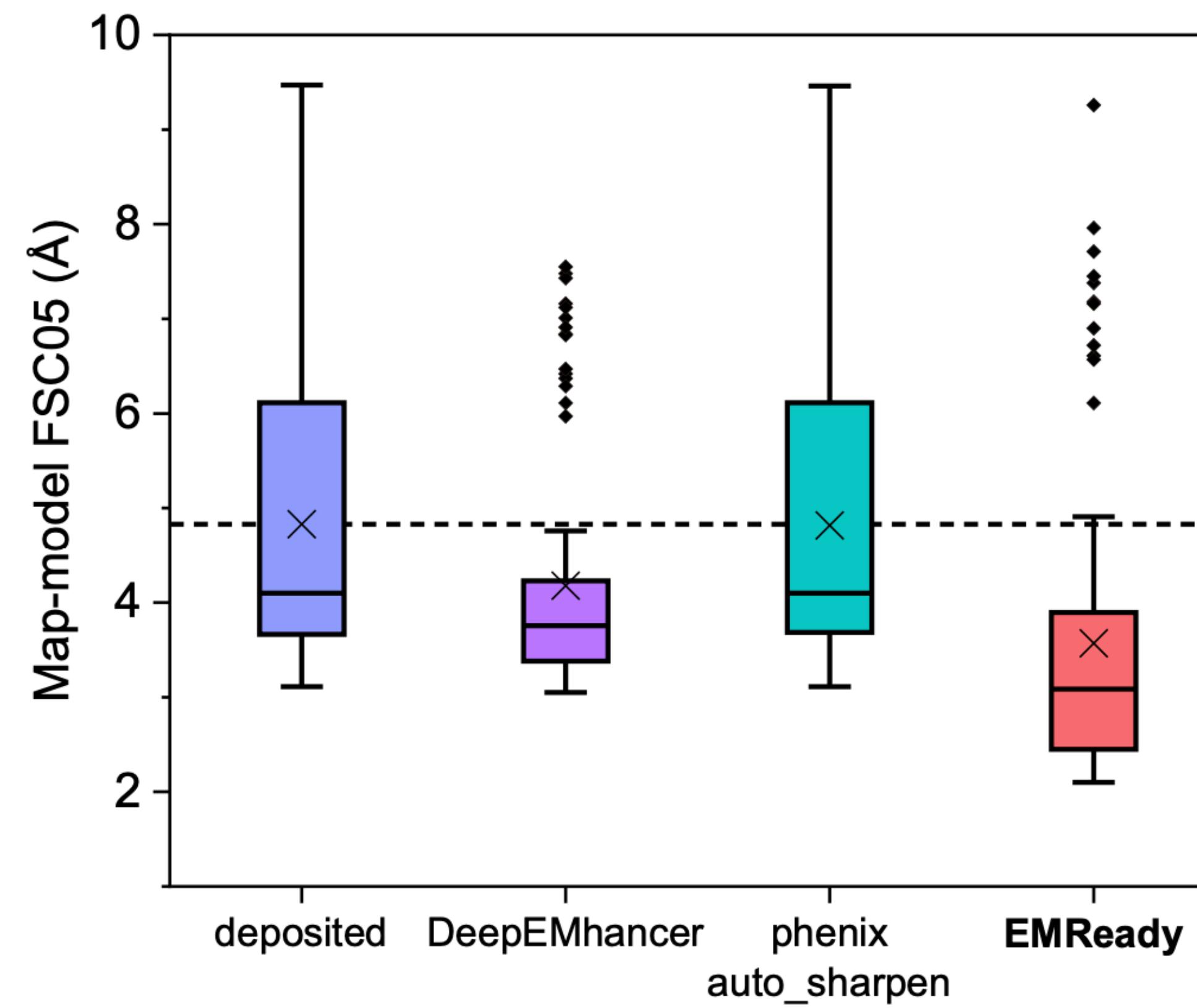
List of Tests from this Paper

- Evaluations on **primary maps**.
The primary map indicates the final reconstruction result (usually post-processed) deposited in the EMDB.
- Evaluations on **half-maps**.
- Improvement in **map interpretability**.
- Evaluation against higher-resolution structures and maps.
- Validation of density modifications by **EMReady**.
- Ablation experiments.

Evaluations on Primary Maps

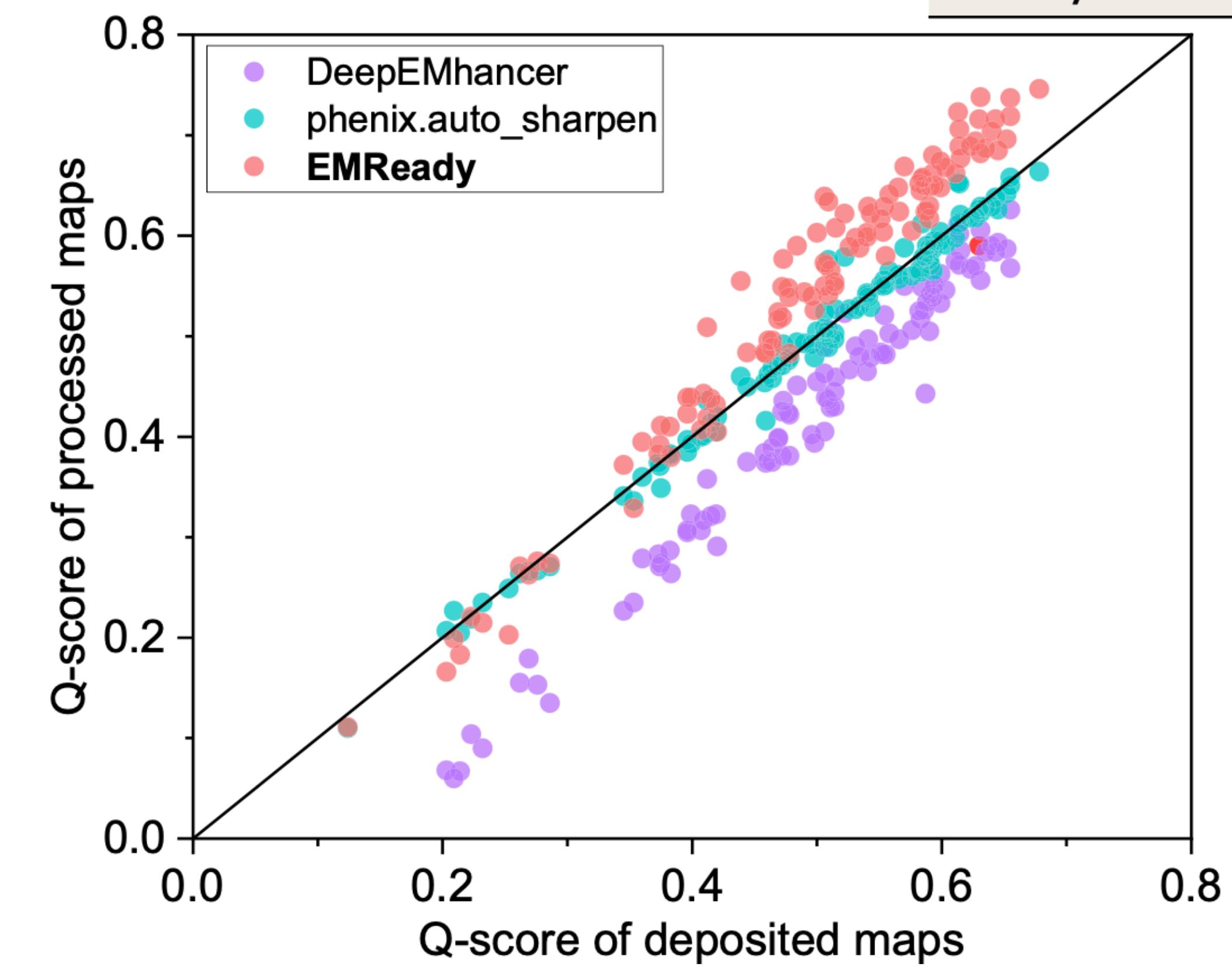
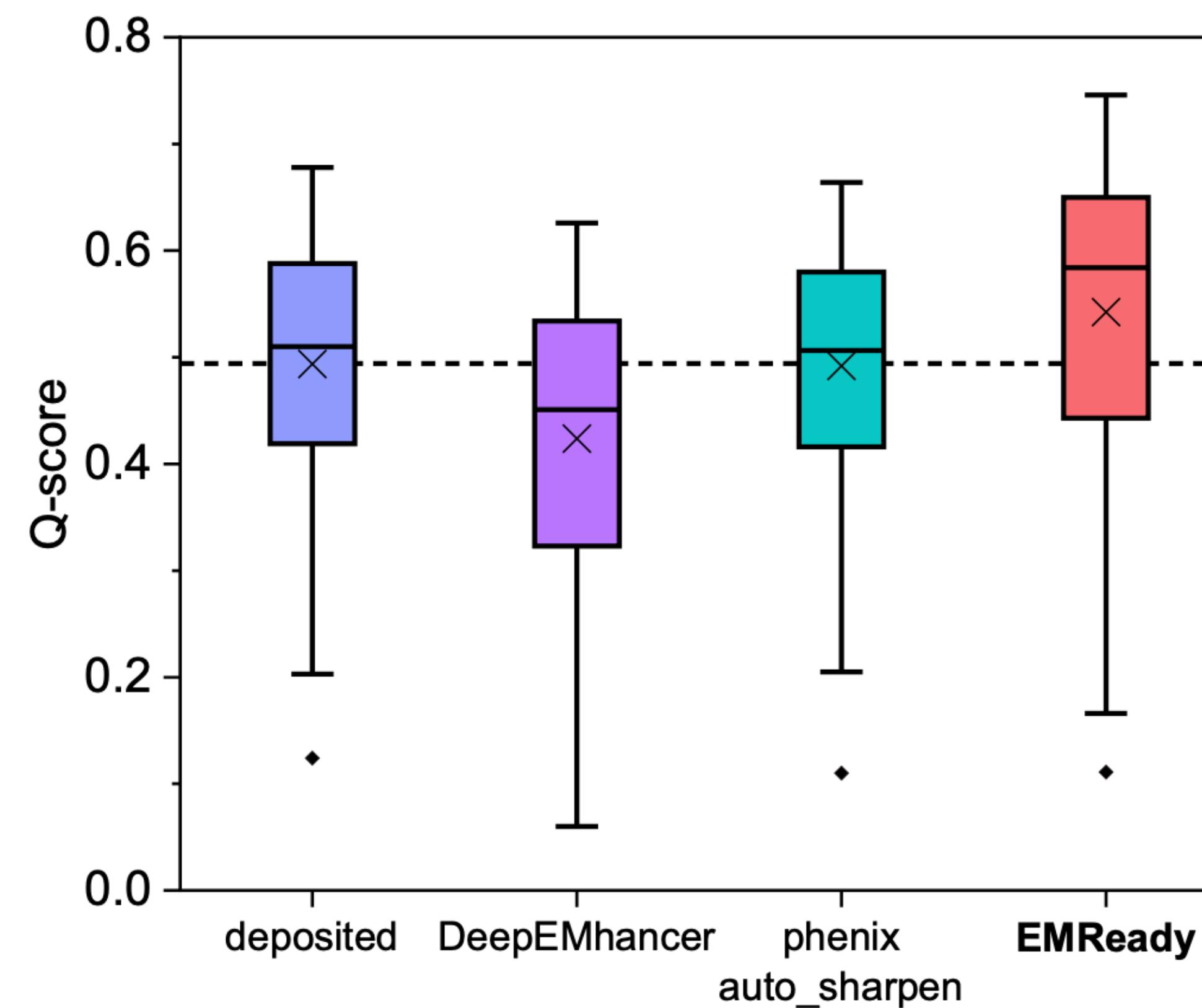
Method 1: the map-model Fourier shell correlation (FSC) (calculated by *phenix.mtriage*).

Method	FSC-0.5 (Å)
deposited	4.83
DeepEMhancer	4.18
phenix.auto_sharpen	4.82
EMReady	3.57



Evaluations on Primary Maps

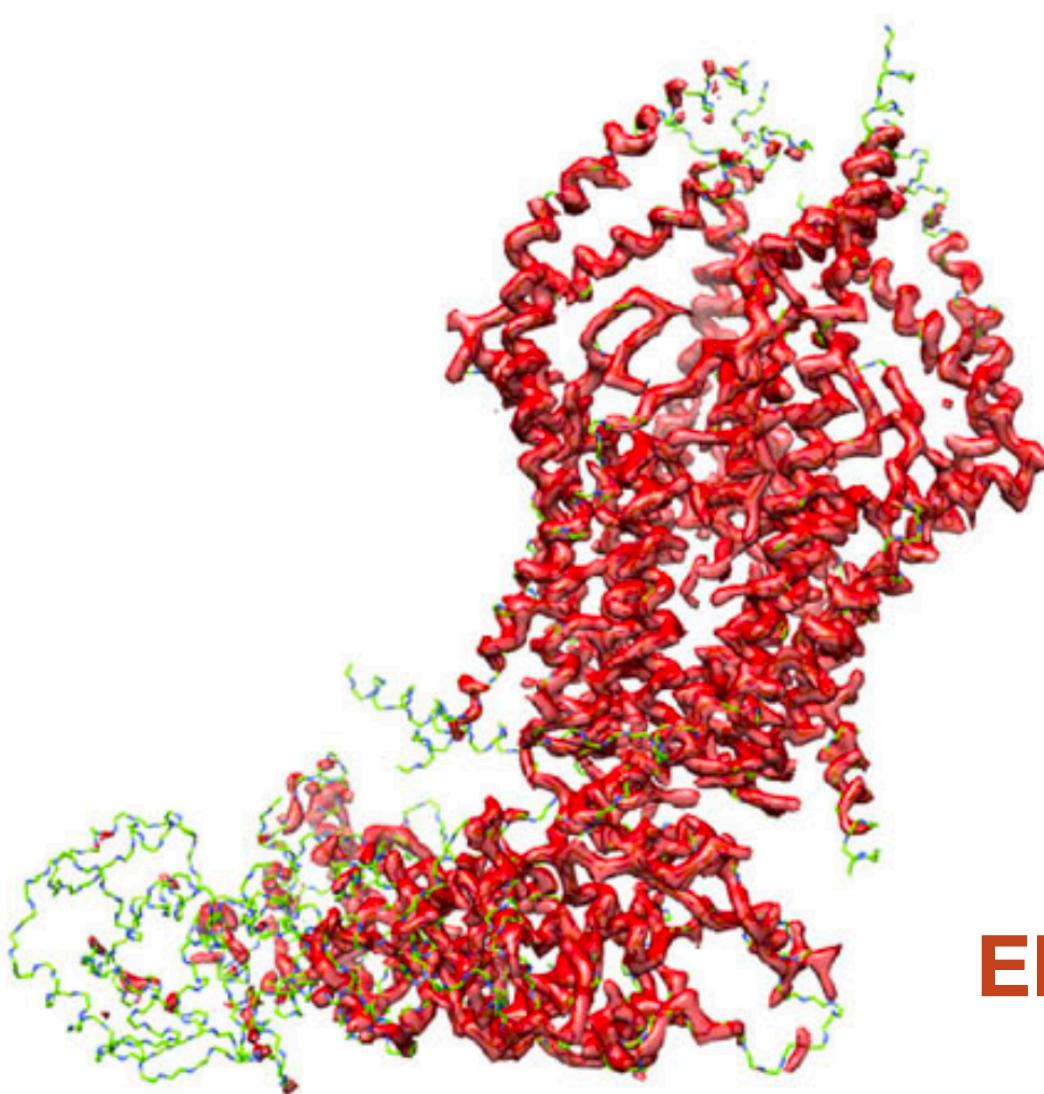
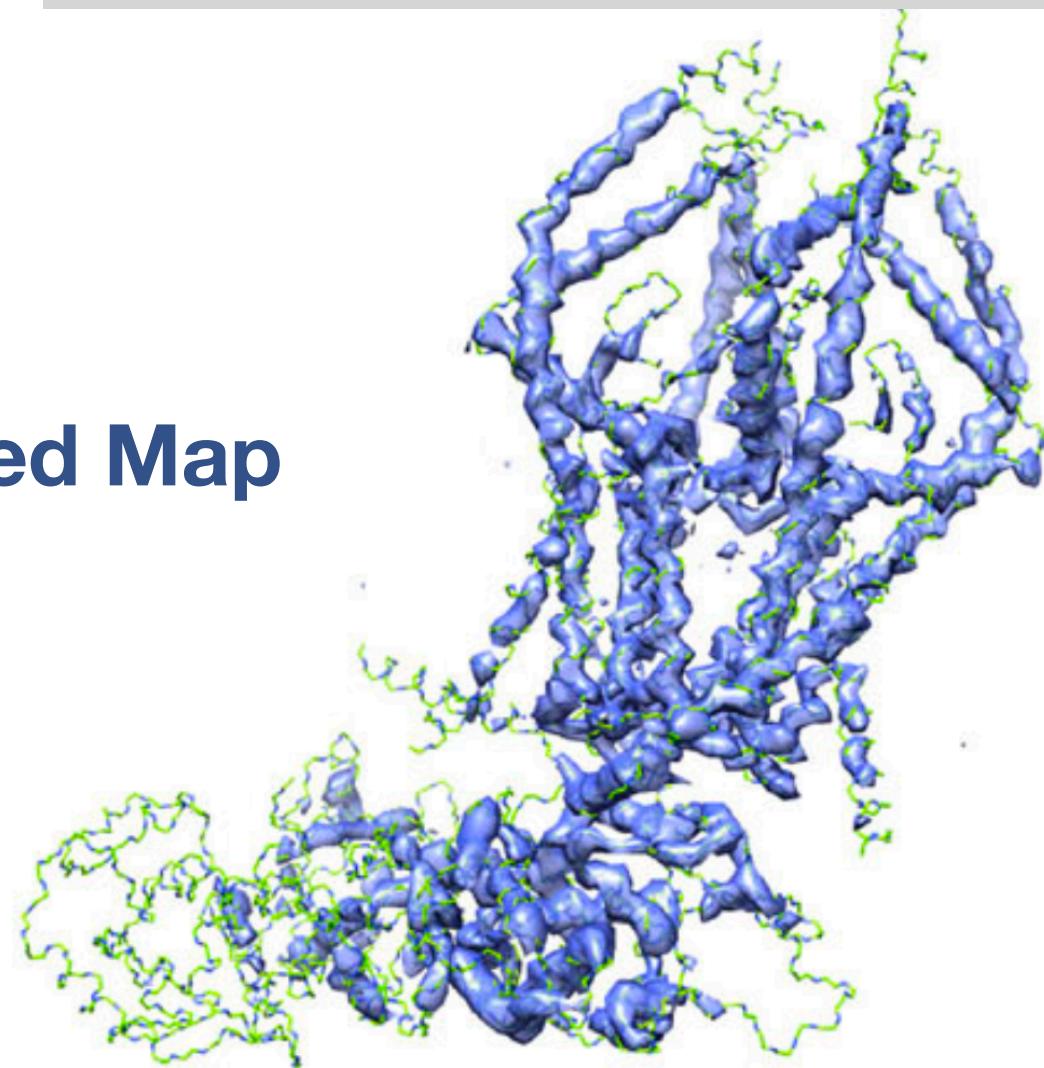
Method 2: the averaged Q-Score of protein atoms.



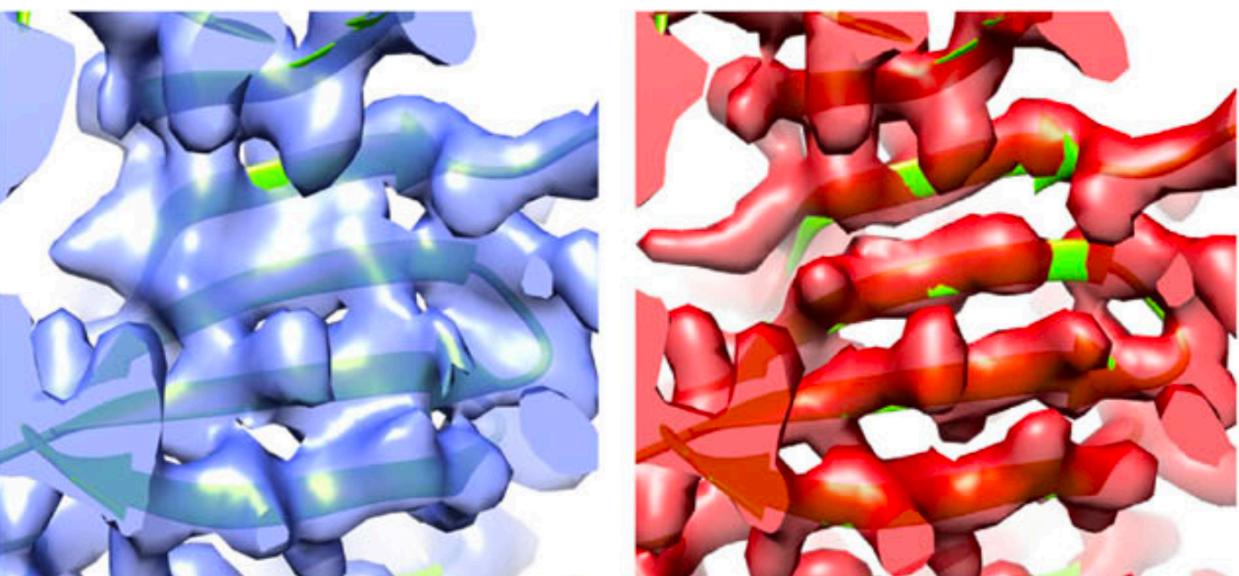
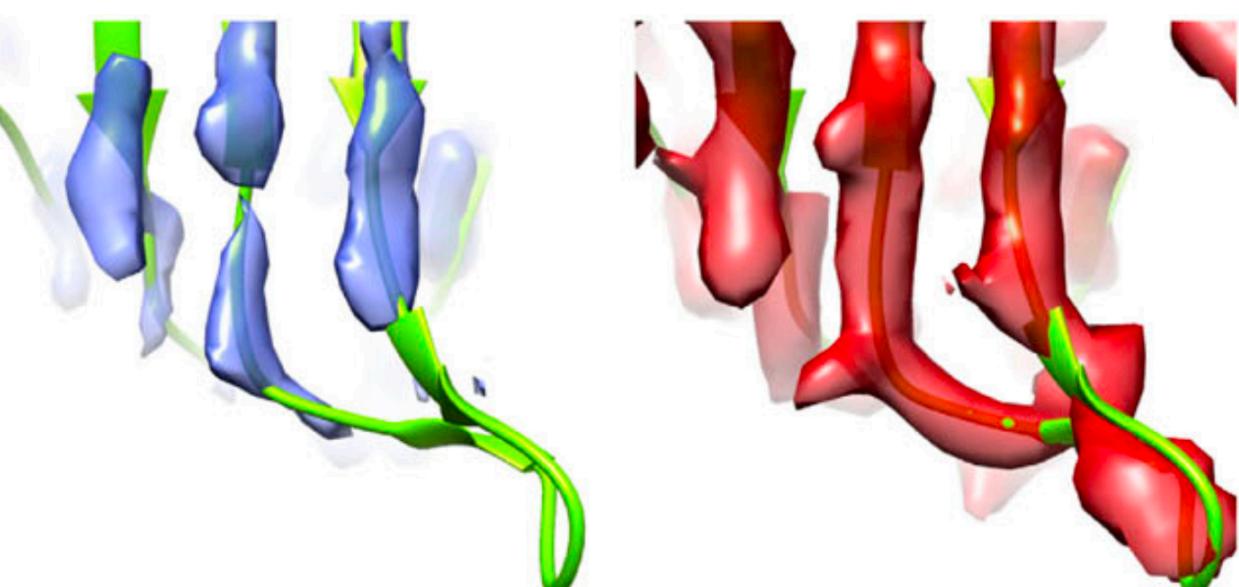
Examples of Improved Maps by EMReady

Improvement in α -helix structure

Deposited Map

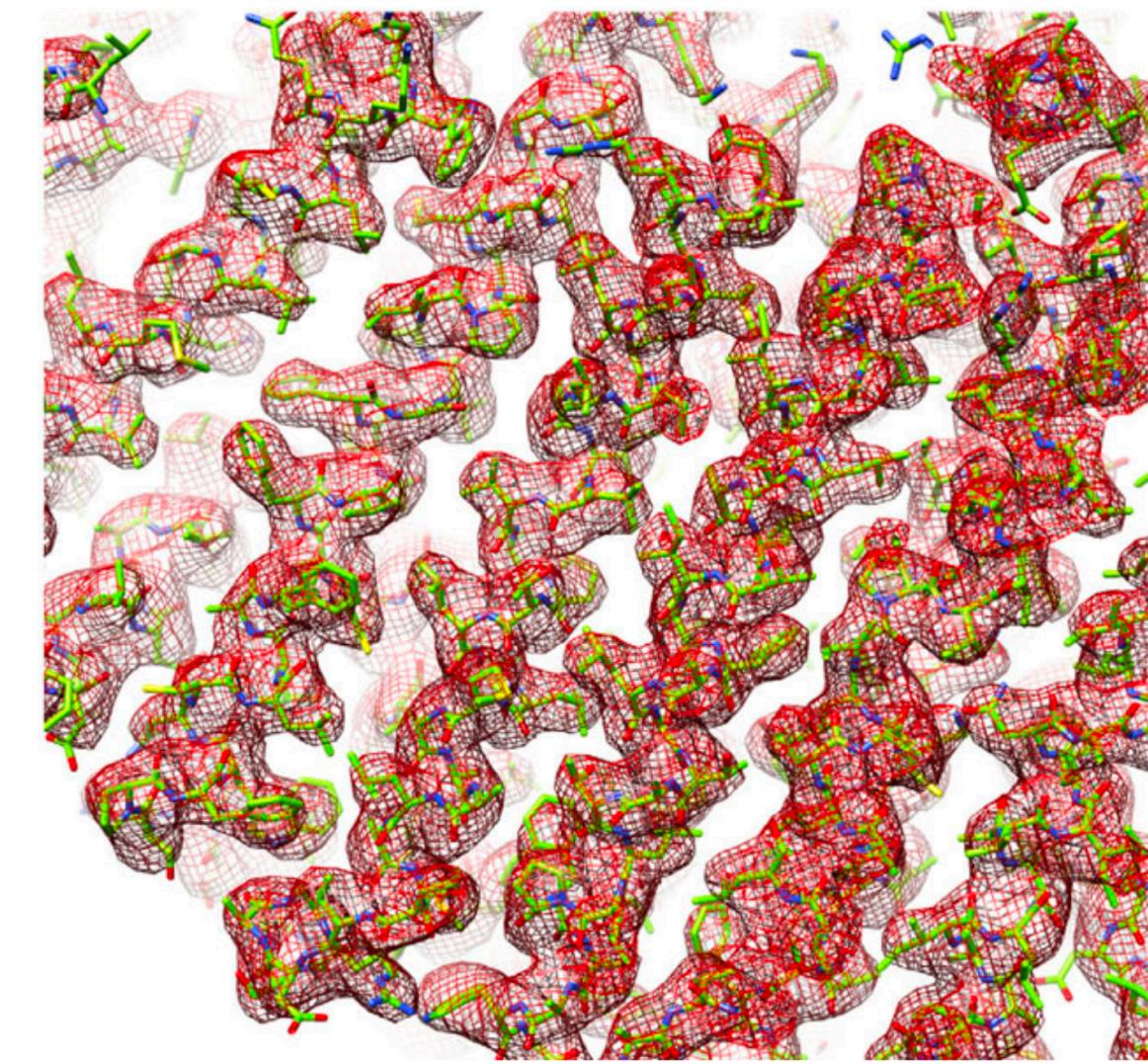
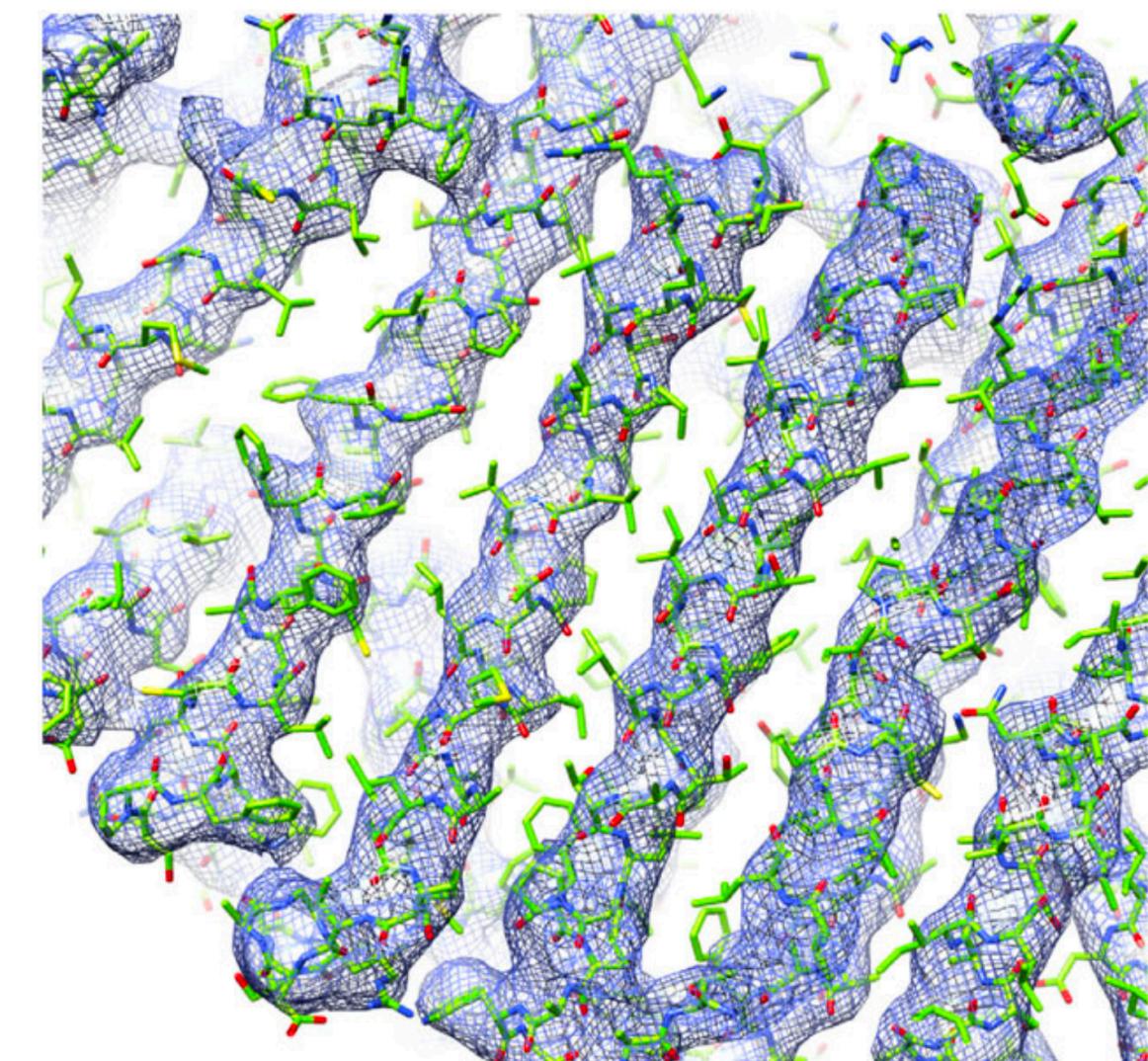


Improvement in β -sheet structure



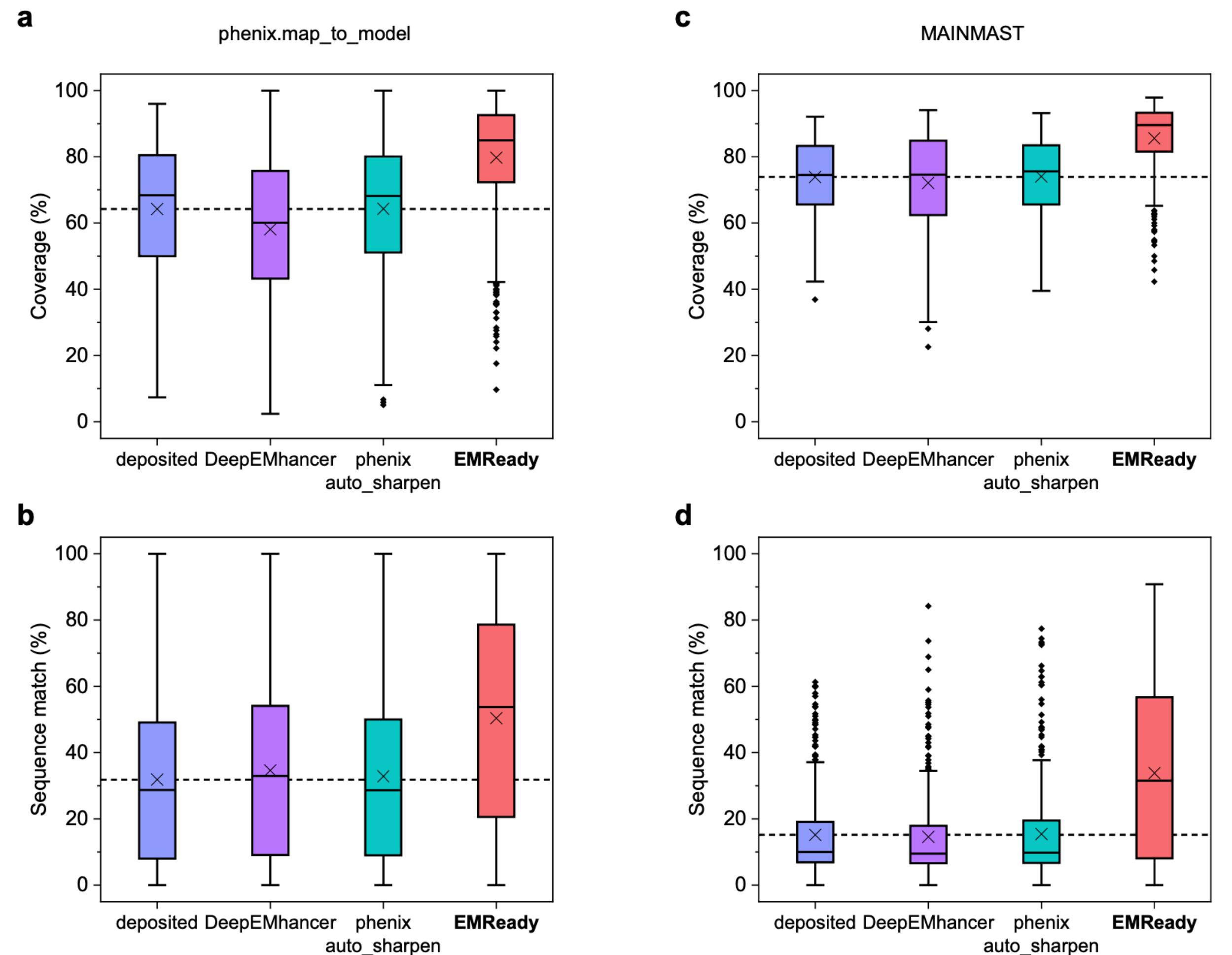
EMReady-processed Map

Improvement in side-chain structure

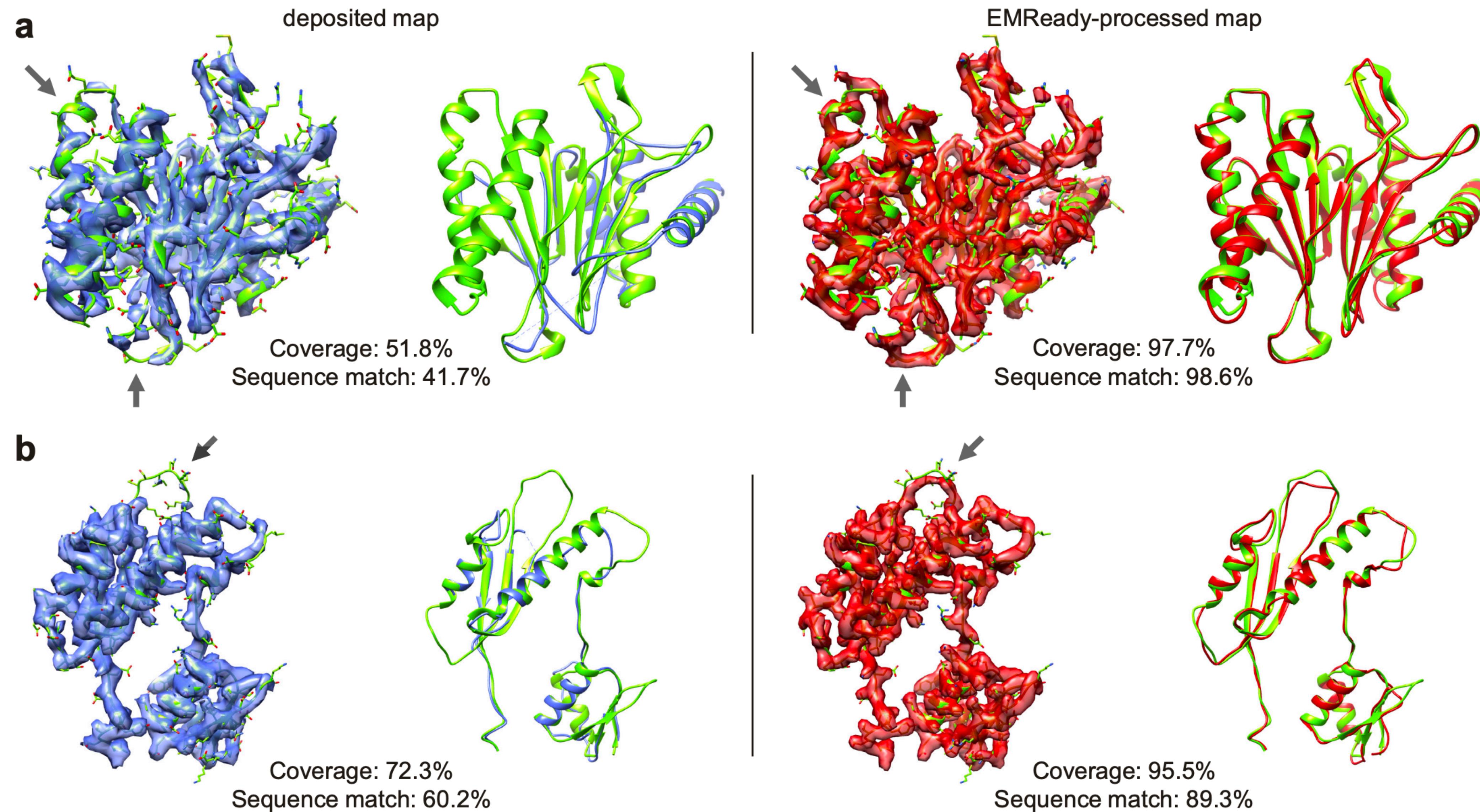


Improvement in Map Interpretability

- Comparison: **built atomic model** \leftrightarrow **deposited atomic model**. (Tool: *phenix.chain_comparison*)

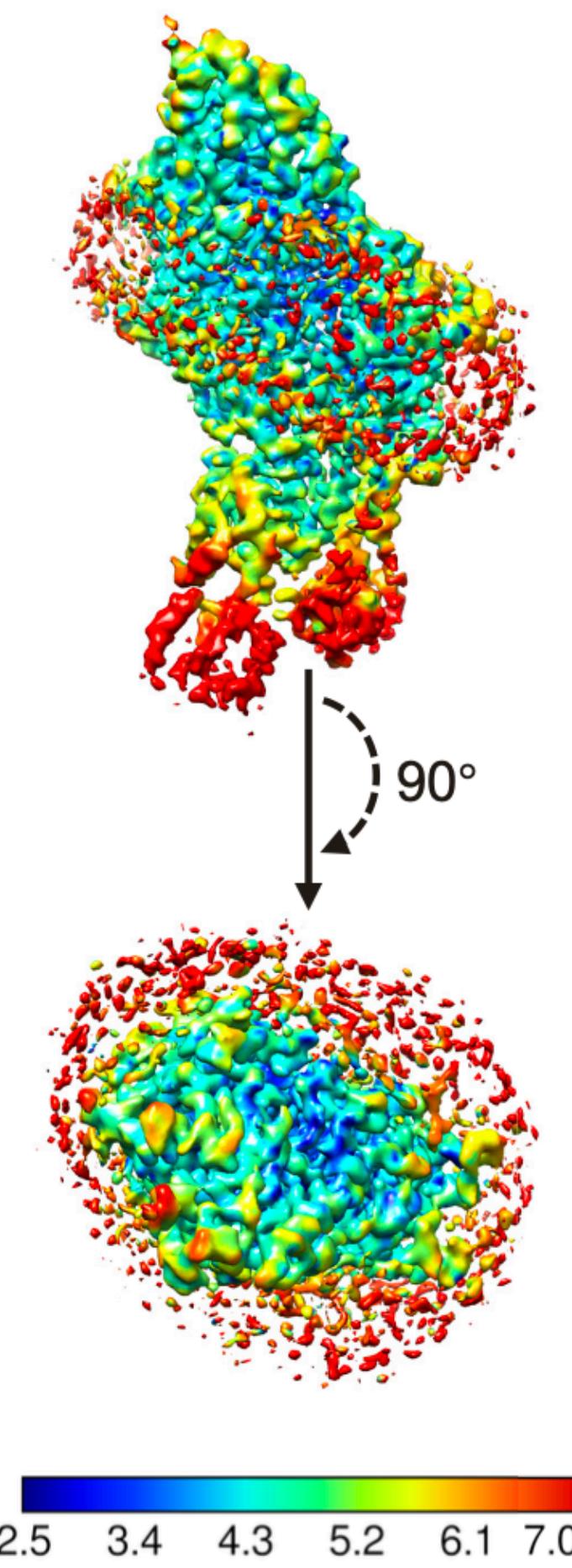


Improvement in Map Interpretability

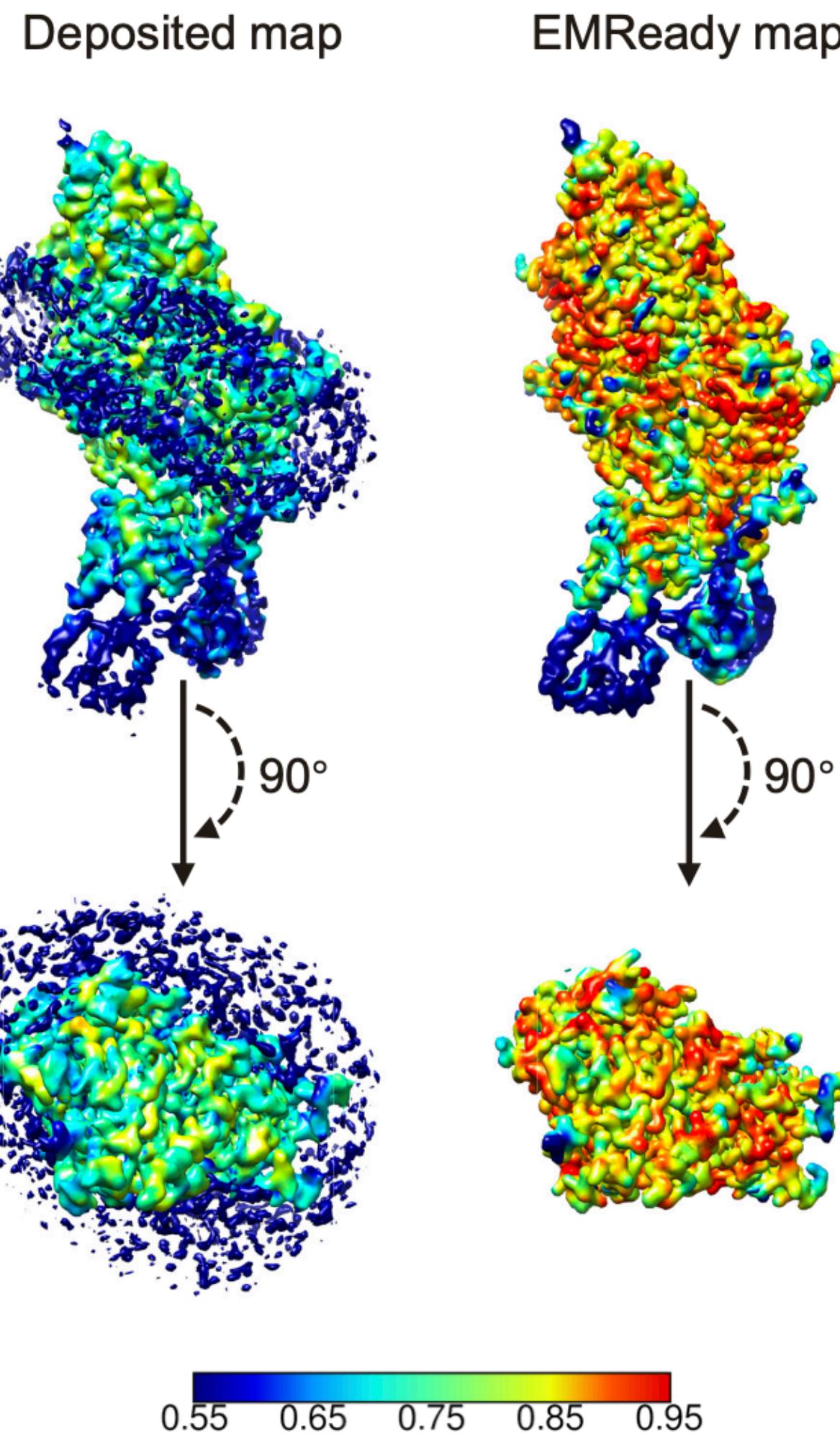


Validation of Density Modifications by EMReady

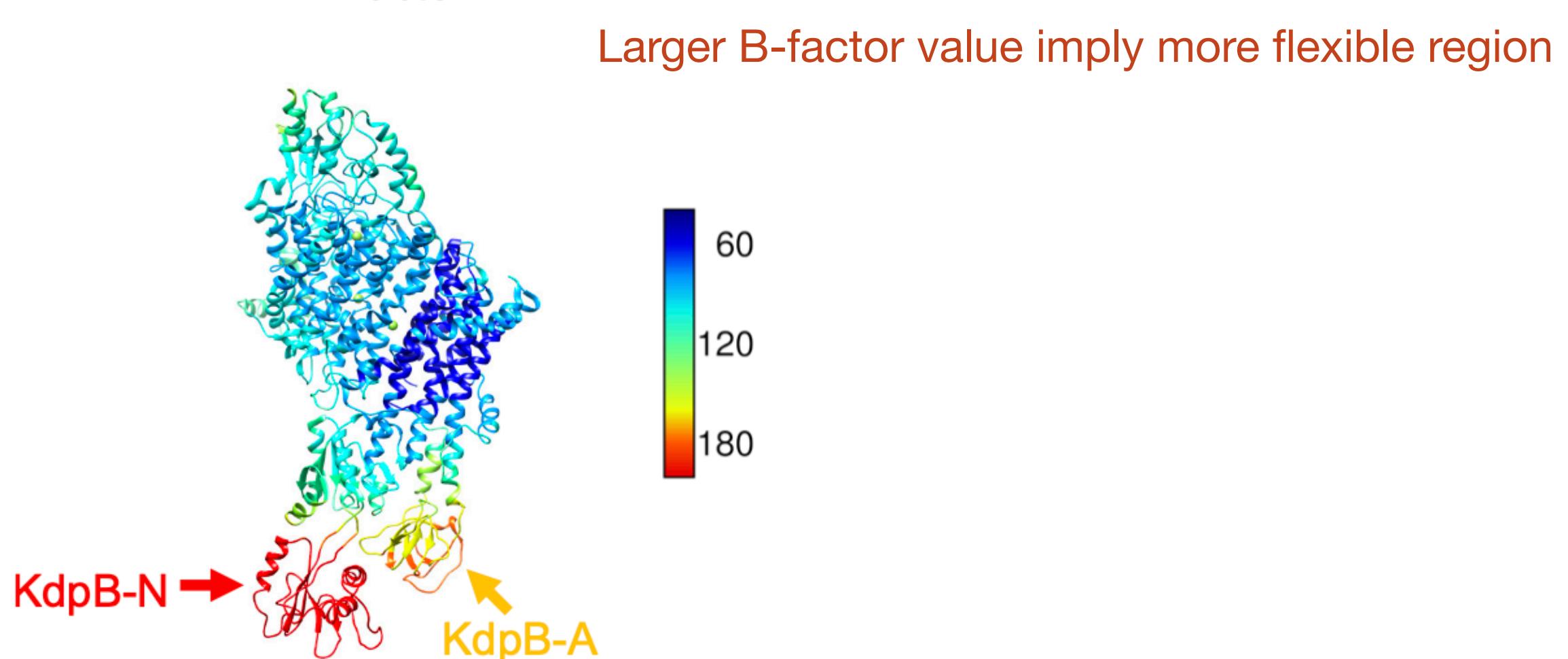
Local resolution



Local correlation



B-factor



- EMReady tends to modify the density towards the PDB structure for the higher resolution (more reliable) regions, but does not try to overfit the density towards the lower resolution (more flexible) parts.
- EMReady succeeds in suppressing the noises of lipid solvents around the protein without removing the protein density signals or adding artifacts from noises.