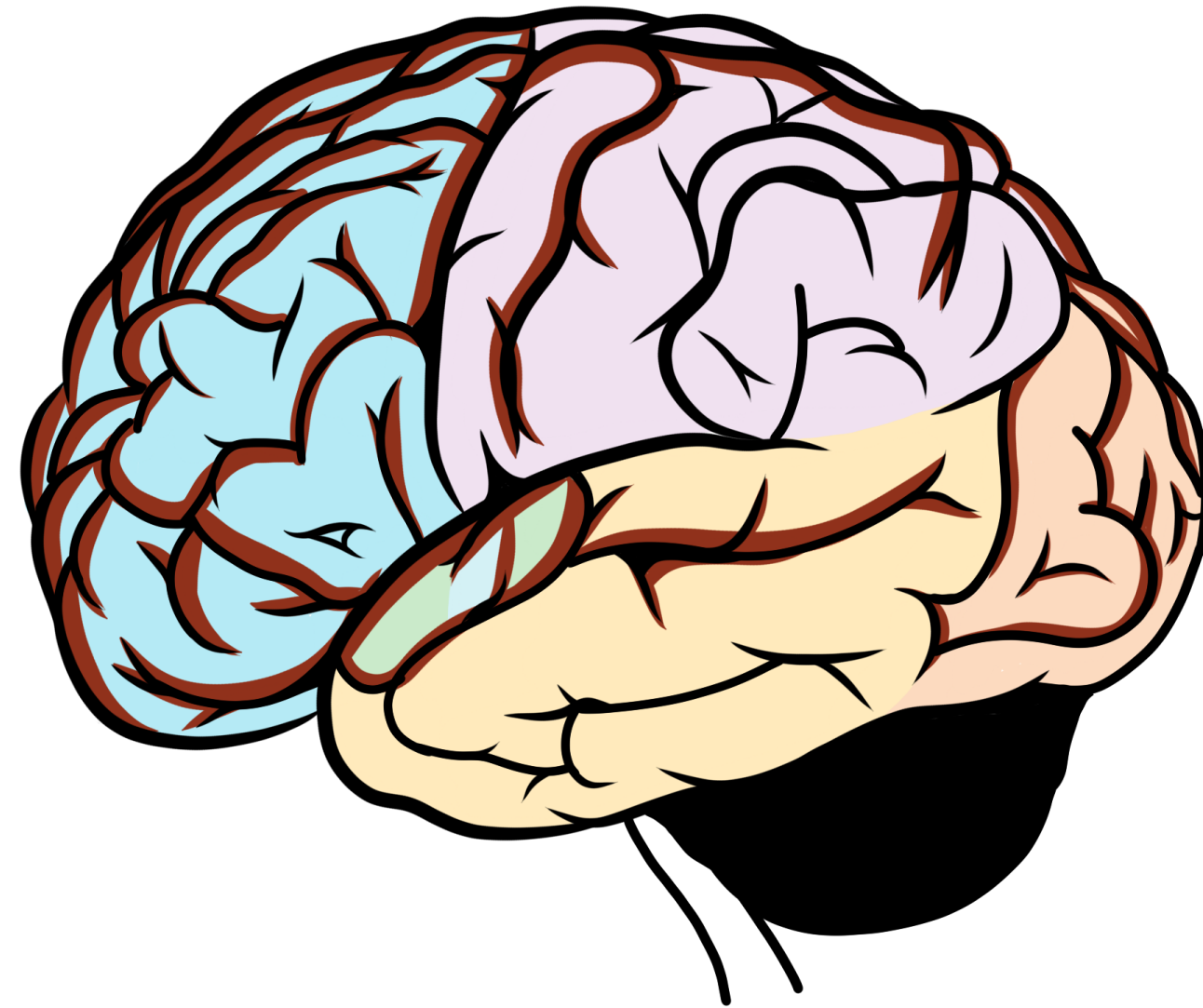


PMLDL. X-Net

Anna Boronina

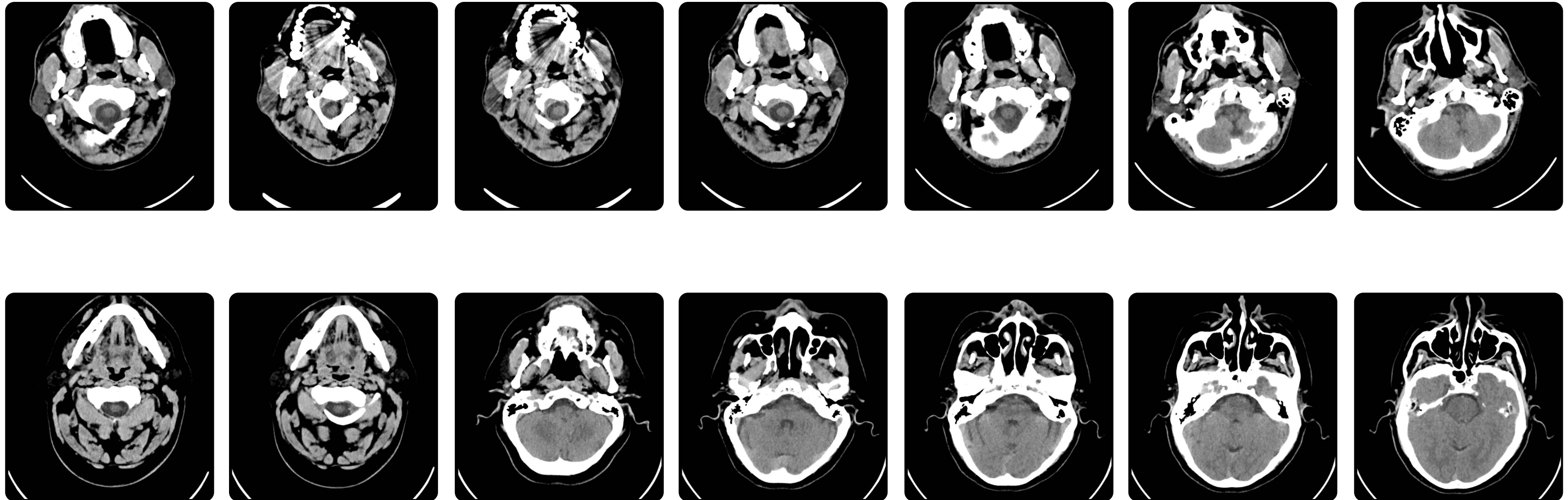
What is a brain stroke (ru: инсульт)?

Brain stroke occurs when the blood supply to part(s) of one's brain is interrupted or reduced.



Problems

- sometimes even doctors cannot spot a brain stroke
- sometimes there is a sequence of MRI scans which should be processed together
- healthy VS stroke:



DL applications

- use Deep Learning model as a complementary tool
- use methods that explain model's decisions
 - GradCam
 - GradCam++
 - LIME
 - ...



X-Net Paper: Qi *et. al*



[International Conference on Medical Image Computing and Computer-Assisted Intervention](#)

..... MICCAI 2019: [Medical Image Computing and Computer Assisted Intervention – MICCAI 2019](#) pp 247-255 | [Cite as](#)

X-Net: Brain Stroke Lesion Segmentation Based on Depthwise Separable Convolution and Long-Range Dependencies

Authors

[Authors and affiliations](#)

Kehan Qi, Hao Yang, Cheng Li, Zaiyi Liu, Meiyun Wang, Qiegen Liu, Shanshan Wang 

Conference paper

First Online: 10 October 2019

26

Citations

7.8k

Downloads

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 11766)



Previous attempts

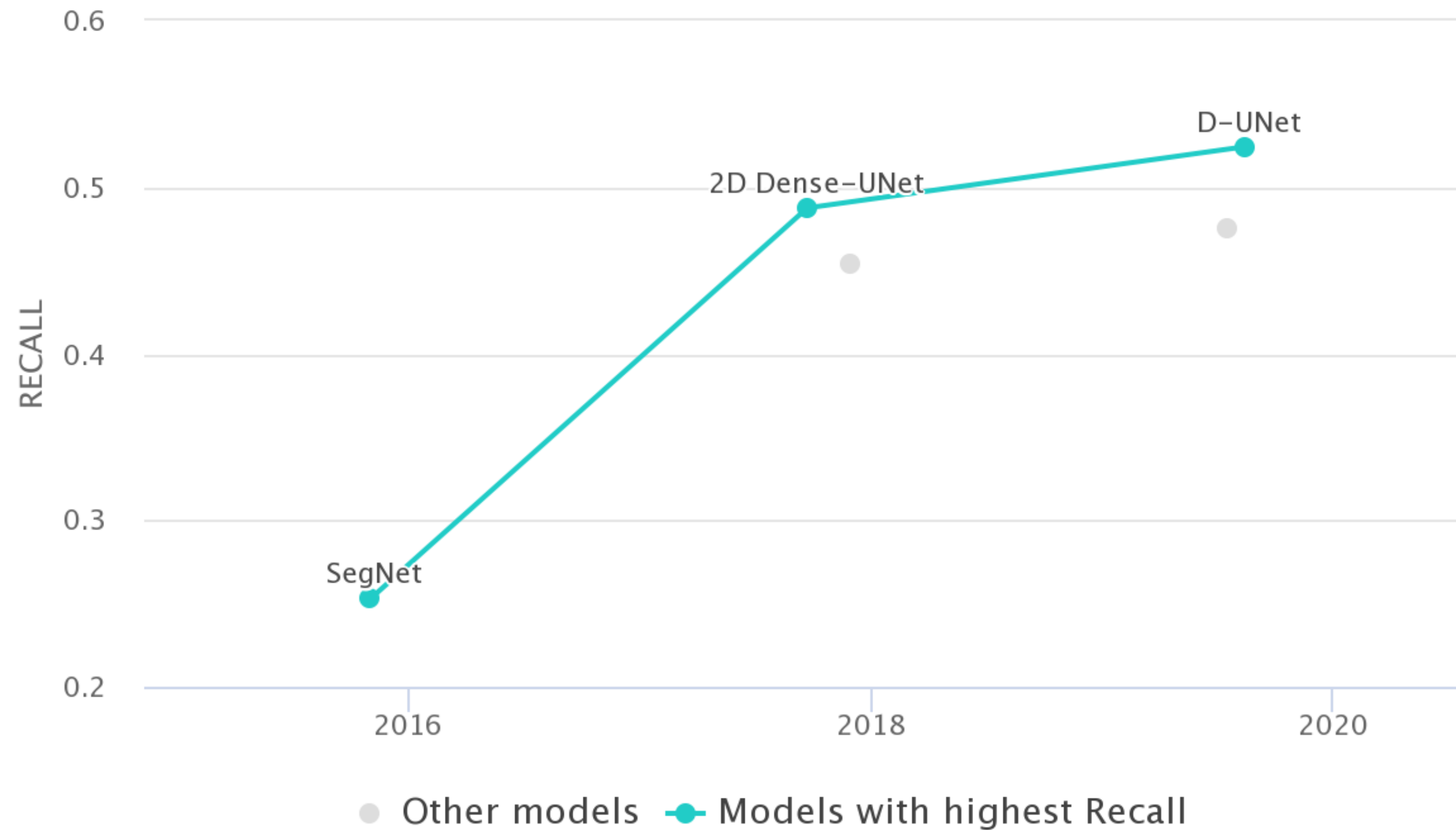
- SegNet - symmetrical convolutional autoencoder
- UNet - originally, non-residual convolutional network with "crop-and-concat"
- 2D Dense Unet - residual dense network

Their problems according to Qi *et. al*

- heavy network parameters
- cannot capture long-range dependencies

Benchmark - ATLAS Dataset

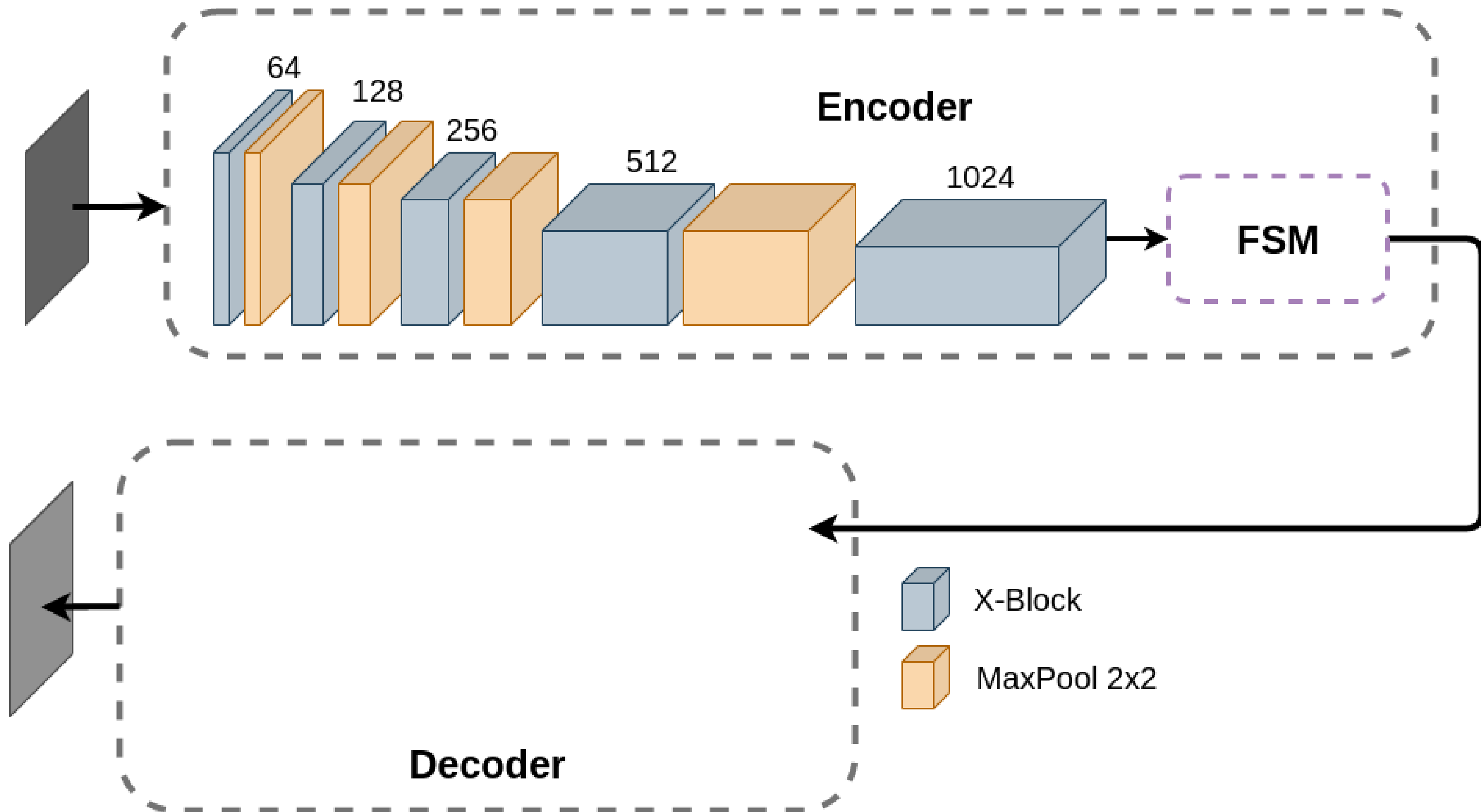
Recall: How many relevant items are retrieved?



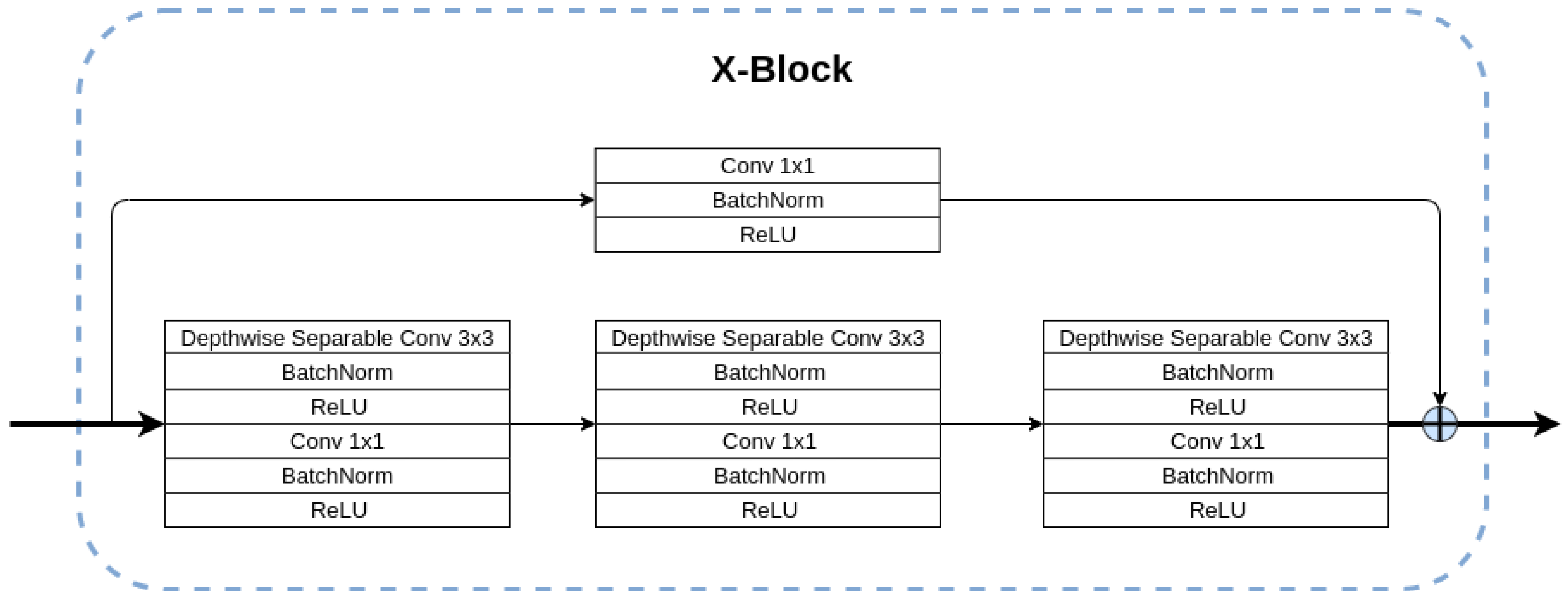
Solution proposed by Qi *et. al*

- reduce number of parameters
 - separable convolution
 - it is efficient
 - gives better segmentation results
 - basic upsampling
- Feature Similarity Module (FSM) to capture long-range dependencies
- skip-and-concat to solve vanishing gradient problem and overfitting

Architecture. Encoder



Encoder. XBlock



Encoder. X-Block. Separable Convolution

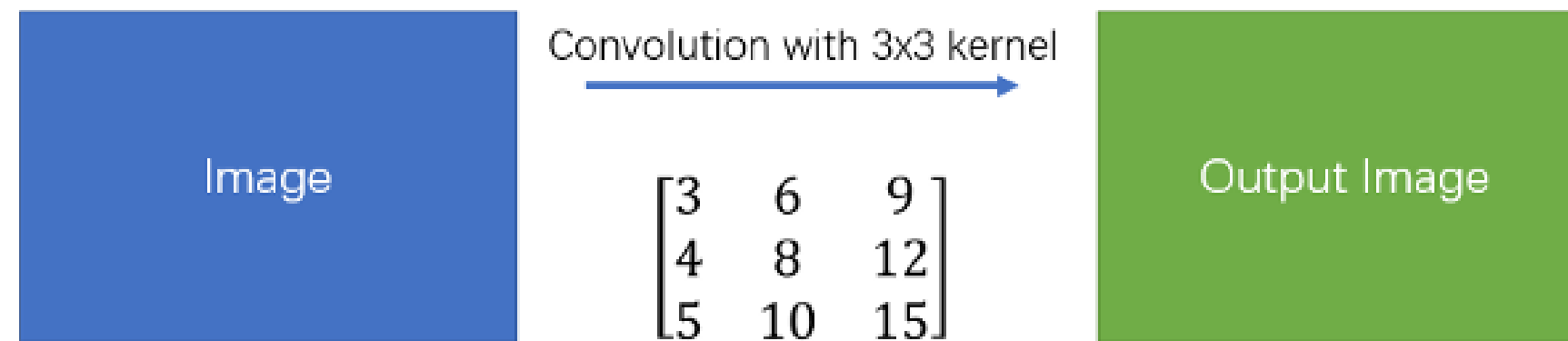
Separating a 3x3 kernel spatially

$$\begin{bmatrix} 3 & 6 & 9 \\ 4 & 8 & 12 \\ 5 & 10 & 15 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

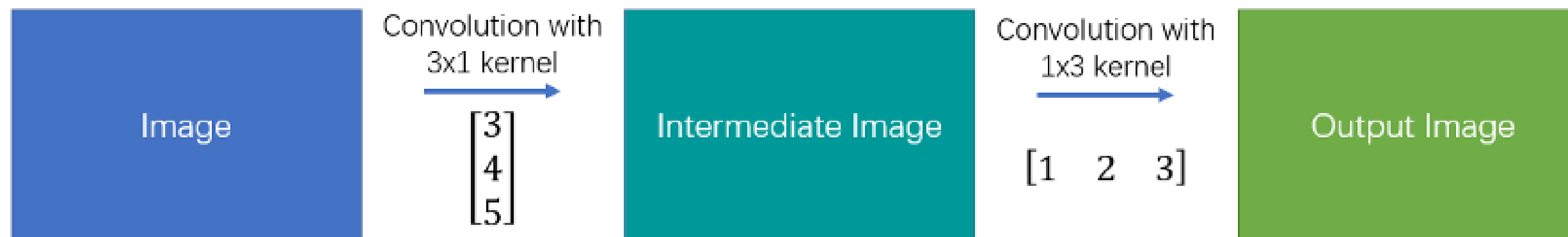
Encoder. X-Block. Separable Convolution

Separable Convolution can enhance efficiency and decrease overfitting without significantly reducing effectiveness

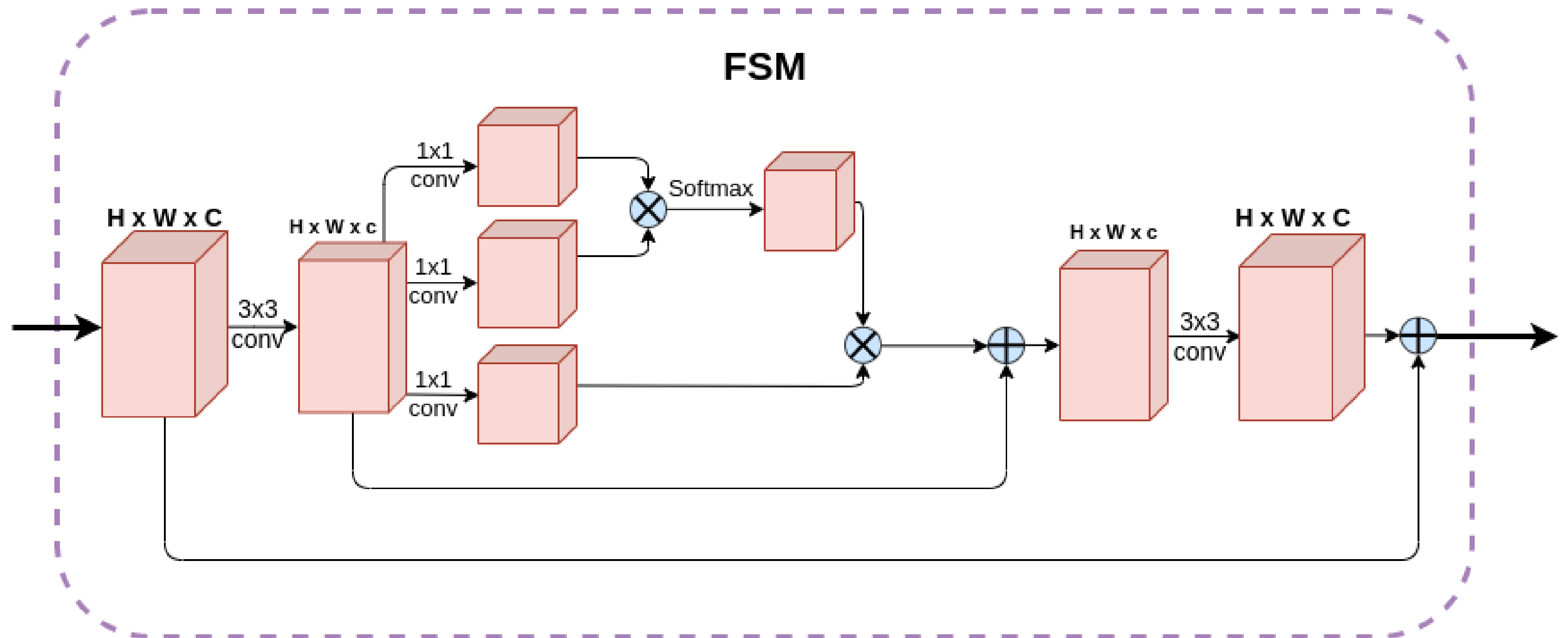
Simple Convolution



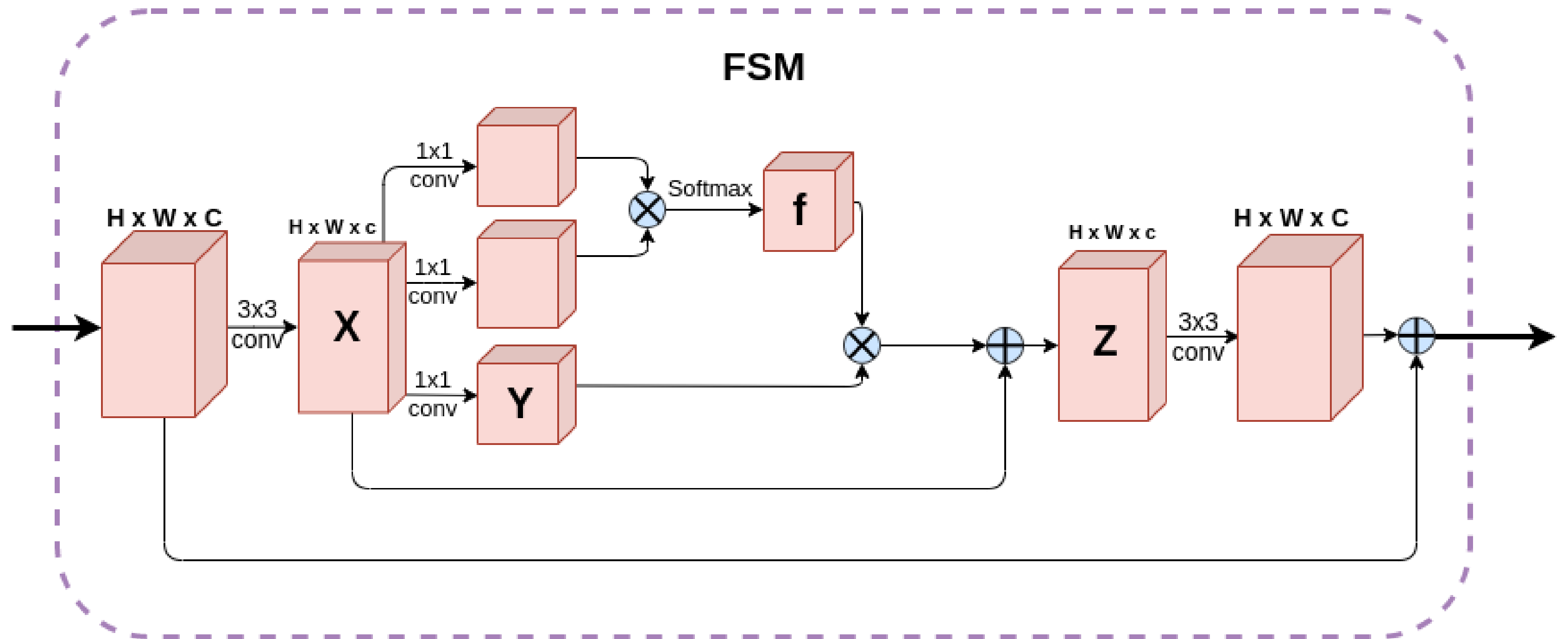
Spatial Separable Convolution



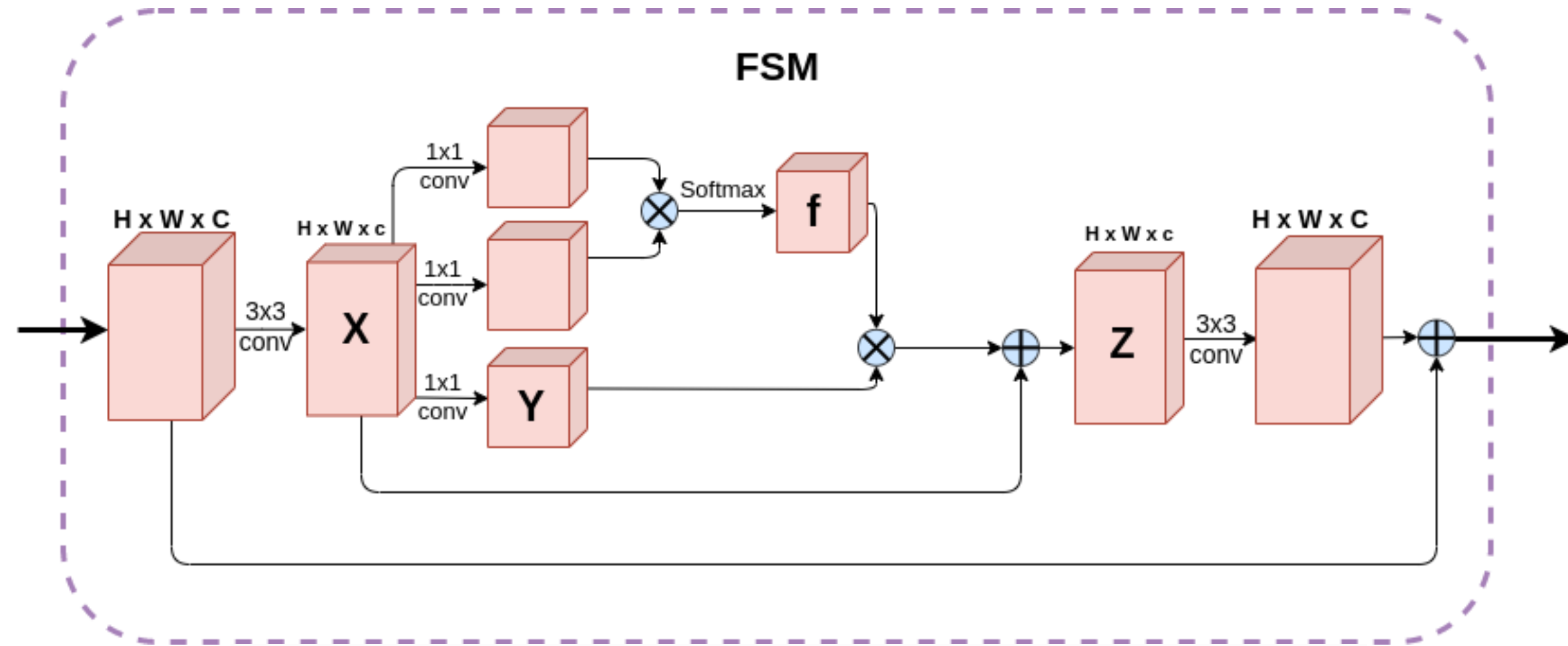
Encoder. FSM



Encoder. FSM



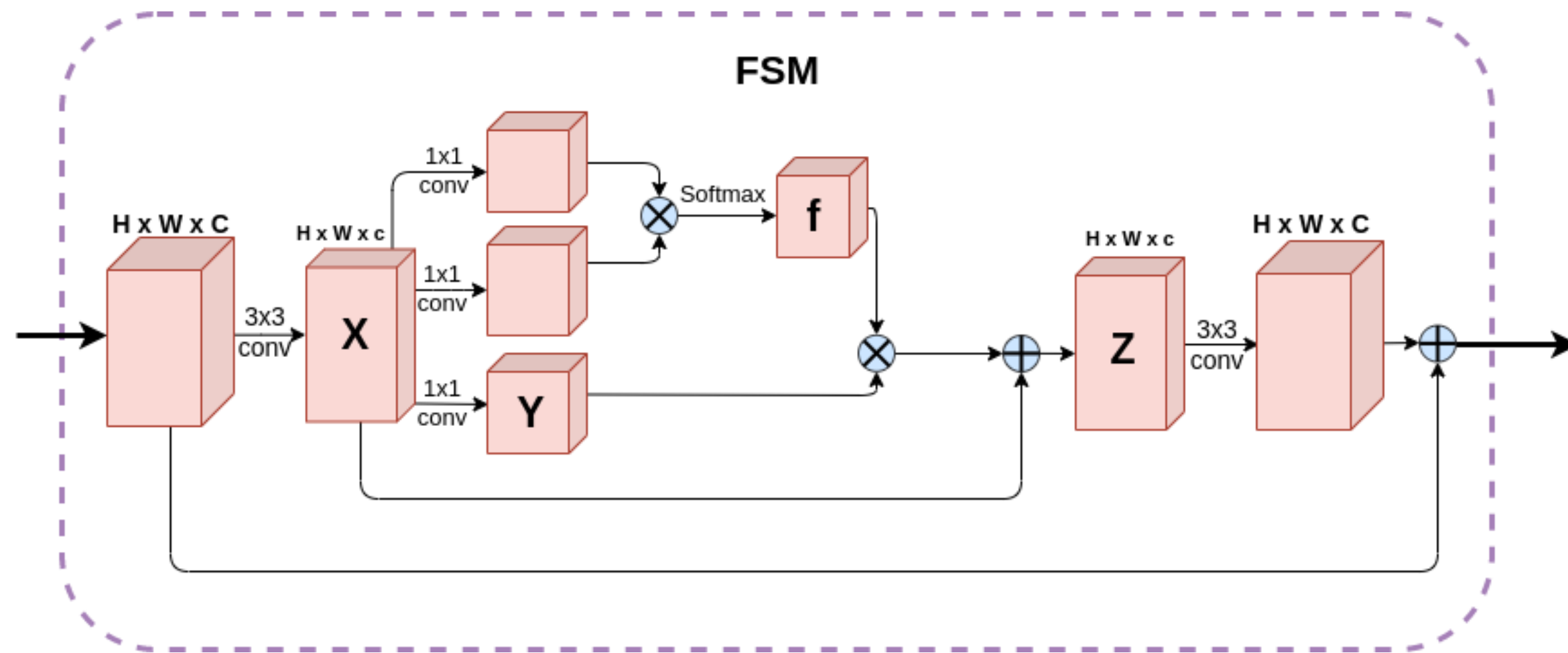
Encoder. FSM



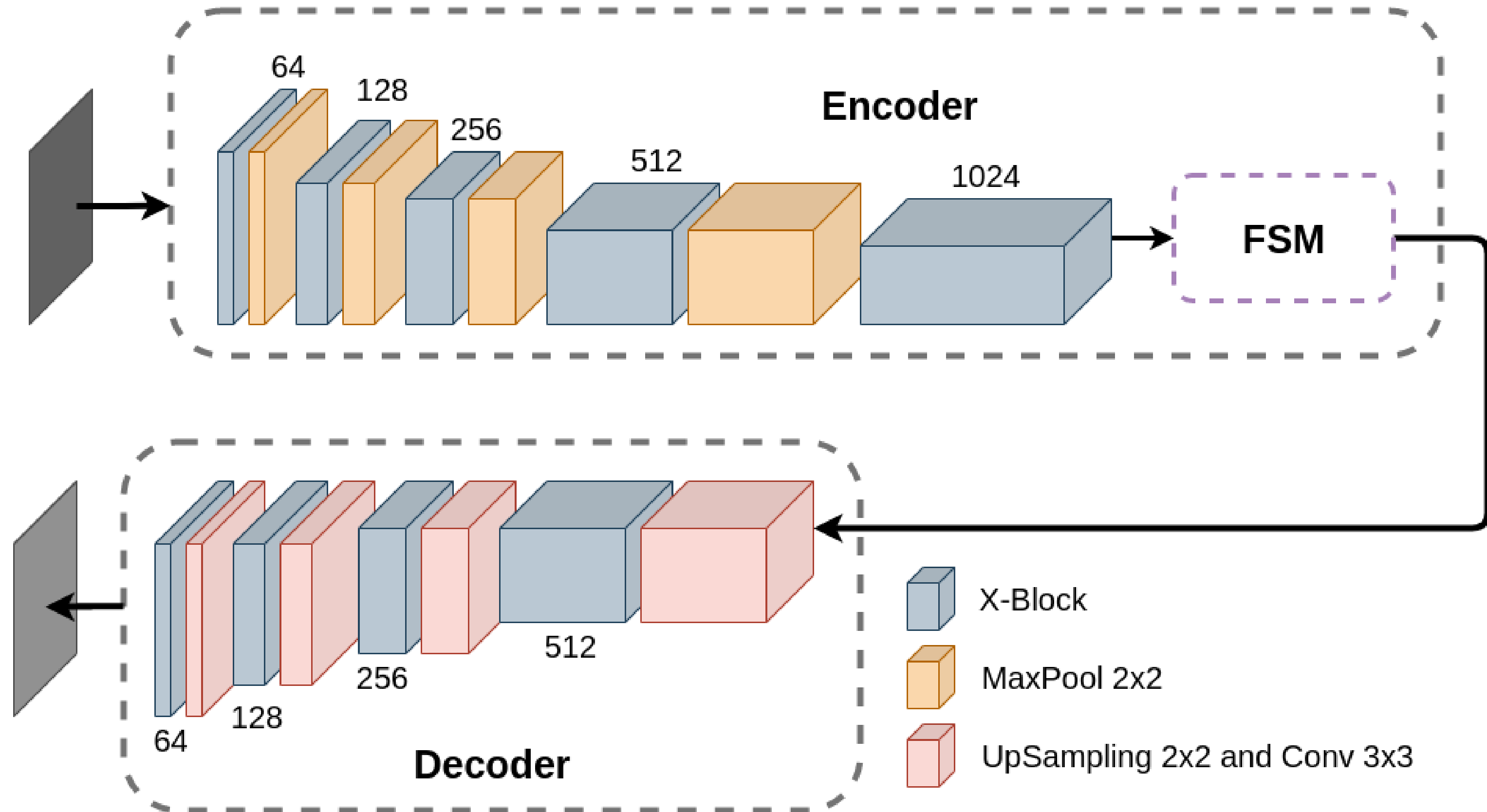
$$Z_i = \sum_{j=1}^N f(x_i, x_j) Y_j + X_i$$

Encoder. FSM

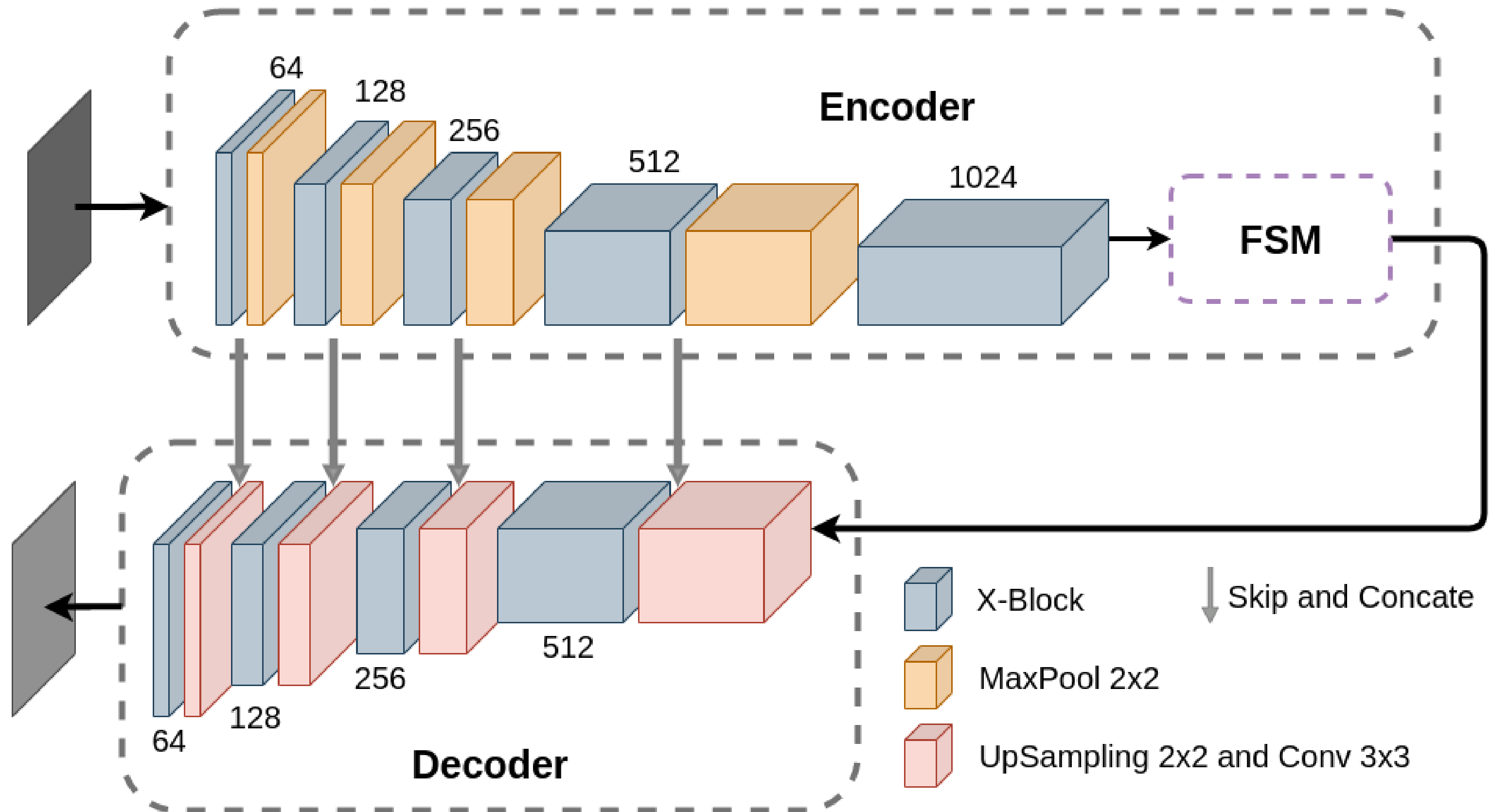
1. FSM is a non-local operation for capturing long-range dependencies by learning a relationship between any two positions of a feature map
2. FSM has a reduced number of training parameters (thanks to separable convolution)
3. Can be plugged into any model



Architecture. Decoder



Final Architecture. Skip-and-concate



A bit more information

Optimizer: Adam, lr=0.001

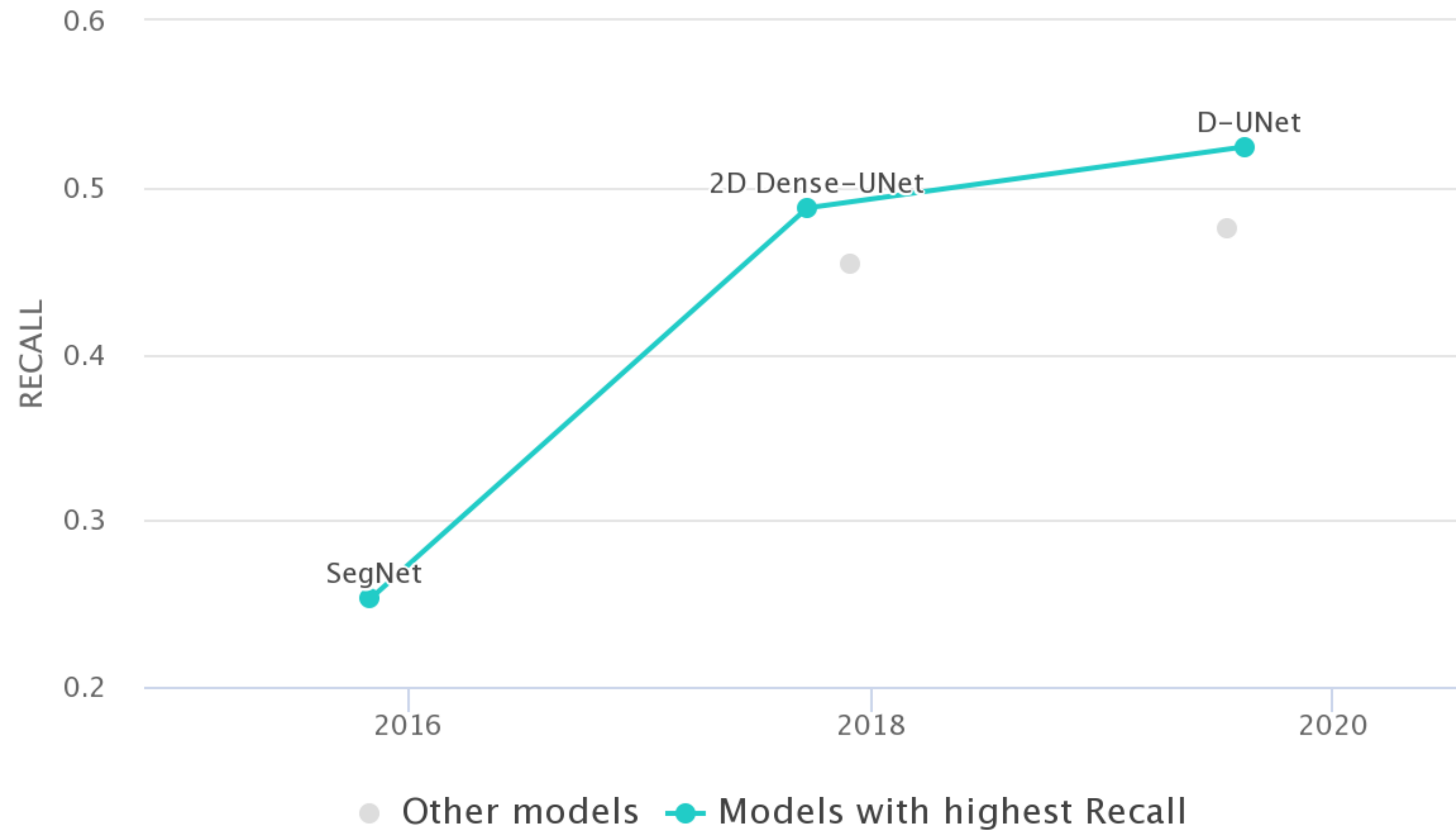
Loss Function: Dice + Cross Entropy losses

Batch Size: 8



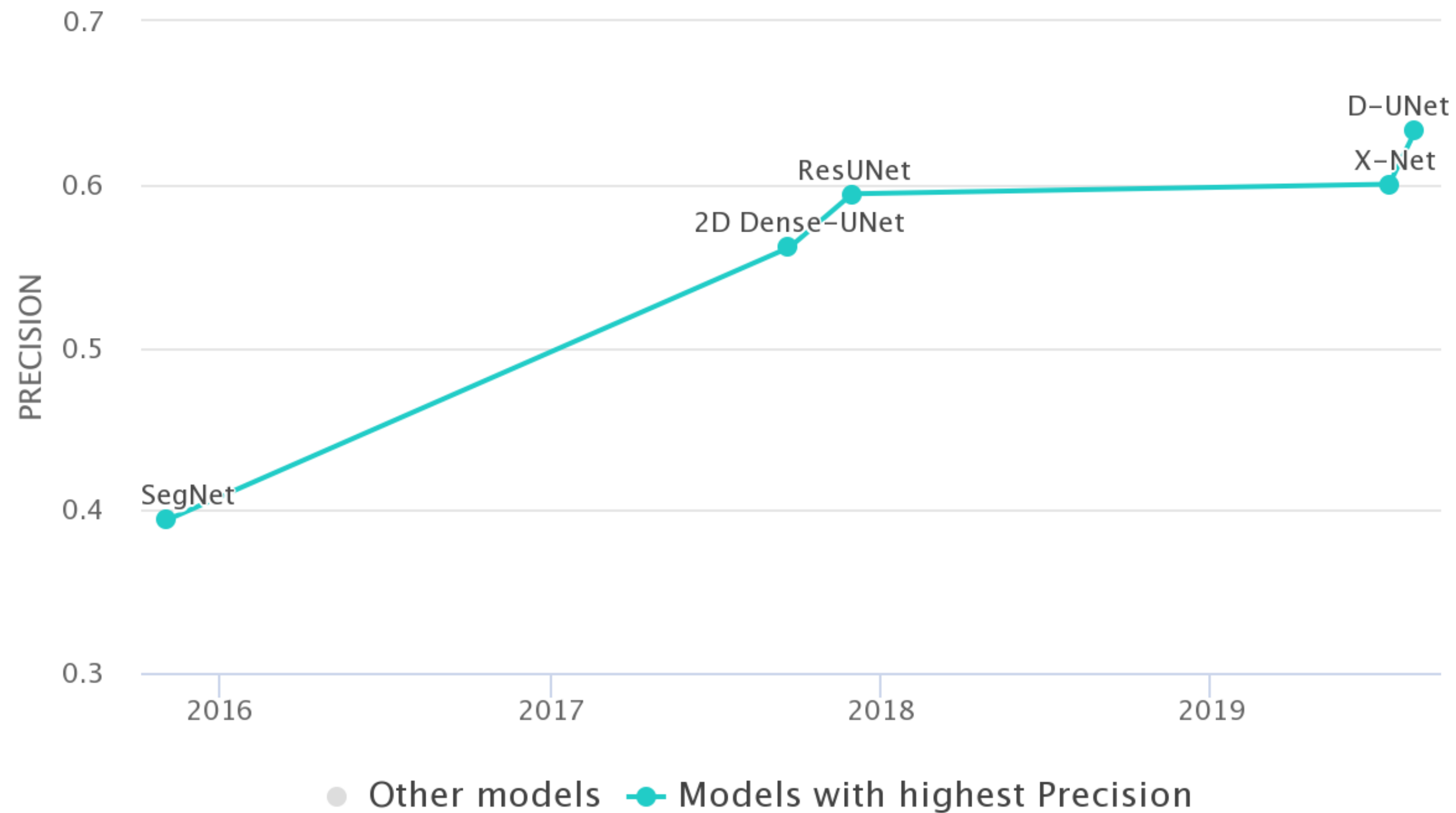
Benchmark - ATLAS dataset

Recall: How many relevant items are retrieved?



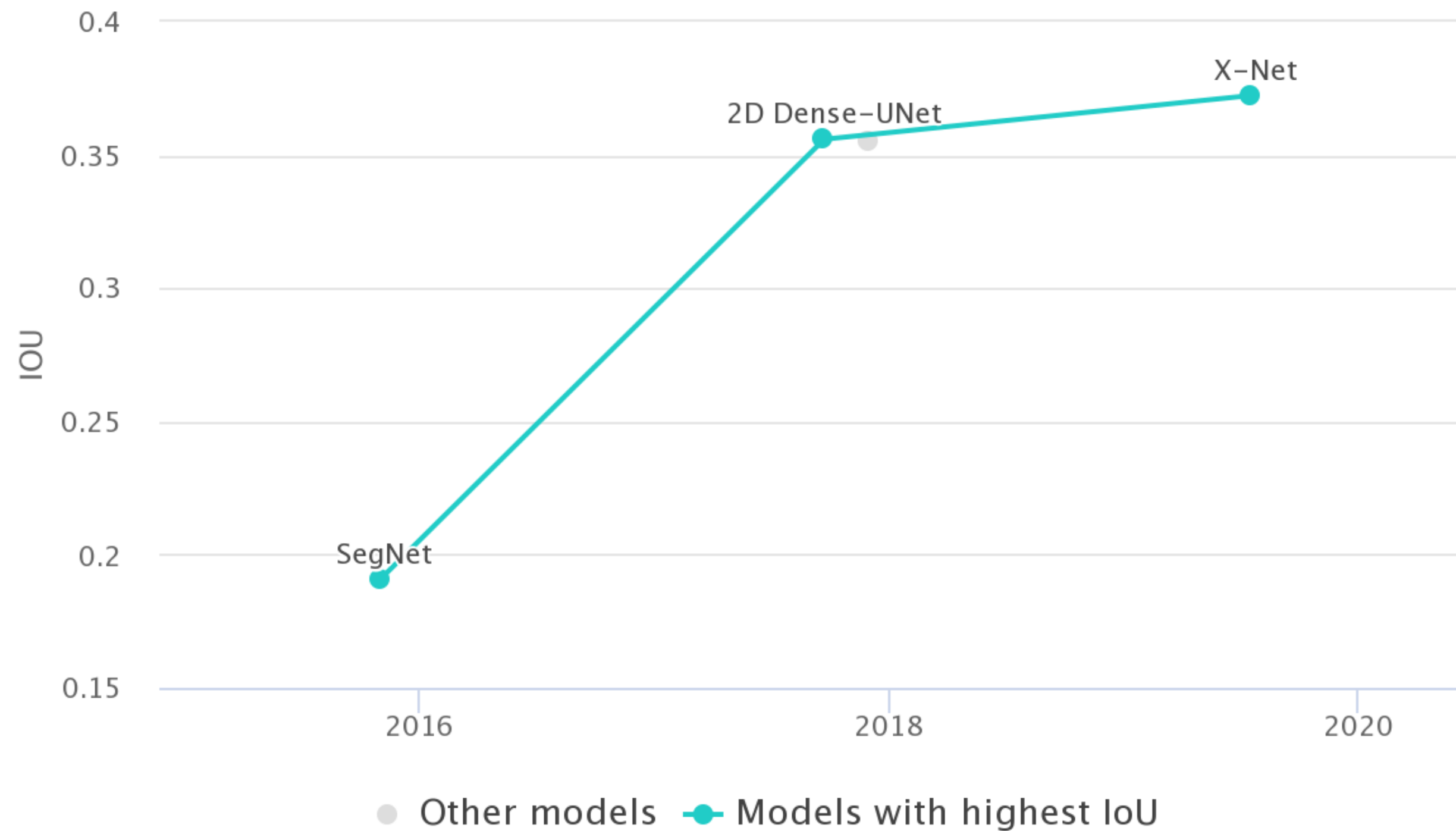
Benchmark - ATLAS dataset

Precision: How many retrieved items are relevant?



Benchmark - ATLAS dataset

IoU: What is the % of overlap between ground-truth segment and predicted segment?



The authors' results

Method	IoU	# parameters
2D Dense-UNet	0.35	50.0M
U-Net	0.34	34.5M
SegNet	0.35	29.5M
X-Net	0.37	15.1M