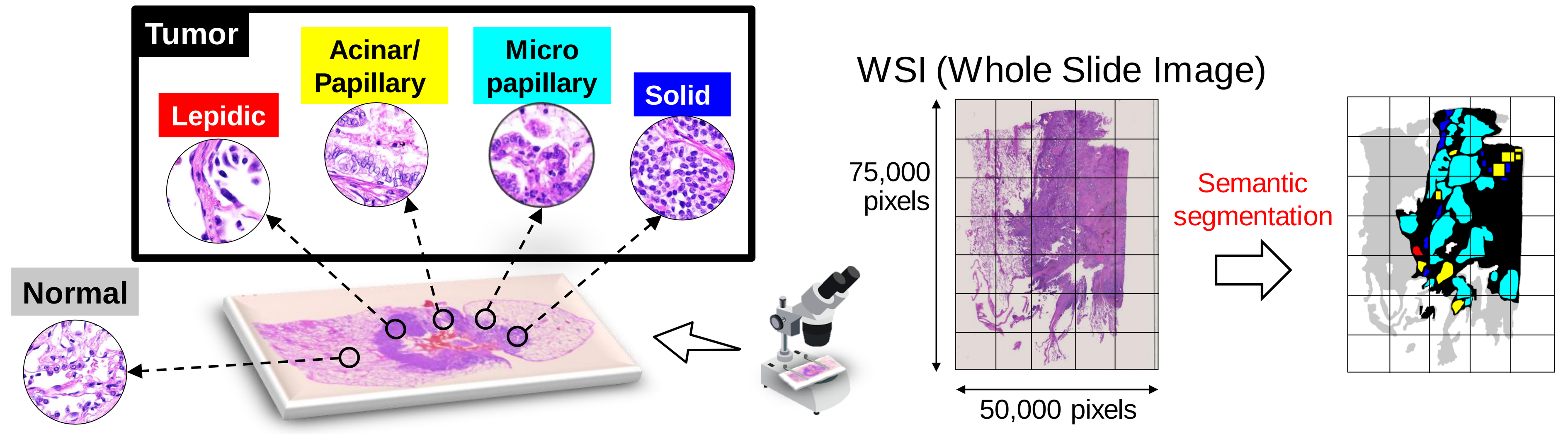
Adaptive Weighting Multi-Field-of-View CNN for Semantic Segmentation in Pathology

Hiroki Tokunaga[1] Yuki Teramoto[2] Akihiko Yoshizawa[2] Ryoma Bise[1,3]
[1]Kyushu University, Fukuoka, Japan [2]Kyoto University Hospital, Kyoto, Japan
[3]Research Center for Medical Bigdata, National Institute of Informatics, Tokyo, Japan

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

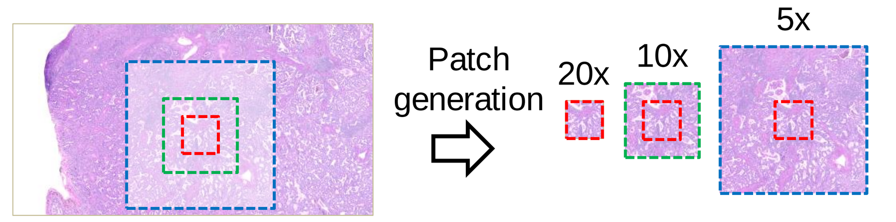
Pathological diagnosis of lung cancer

- Pathological diagnosis of lung cancer is used to determine the treatment policy and predict the prognosis of patients.
- The aim of our study is to automatically segment a lung WSI into cancer subtypes region.



Effect of field of view

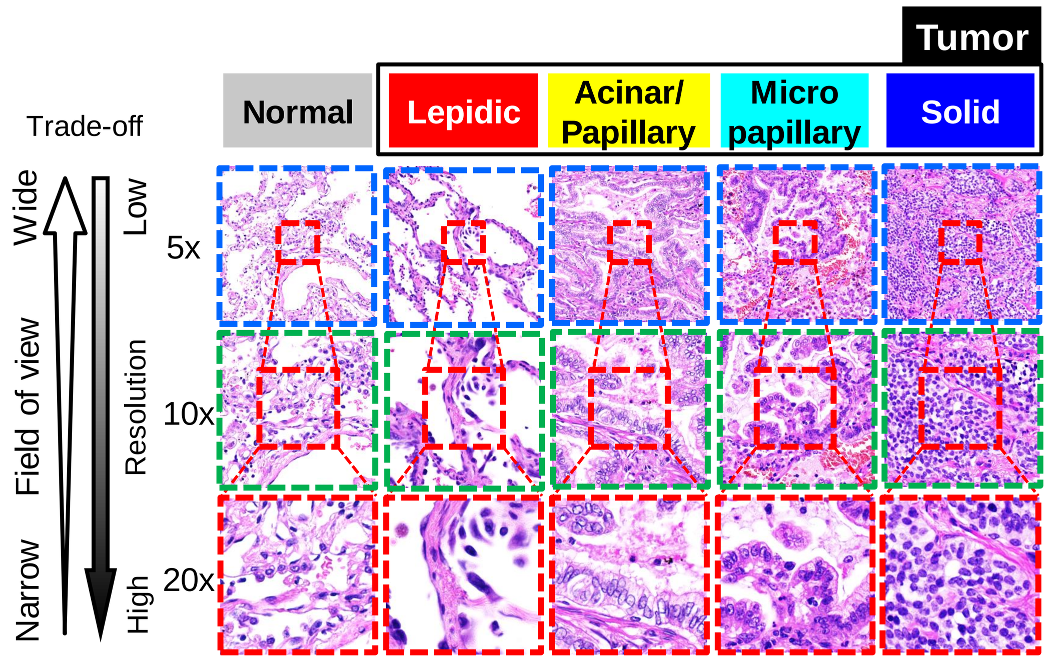
- Patch image segmentation:
  - The size of WSI is too large.
    ⇨ Patch-based approach



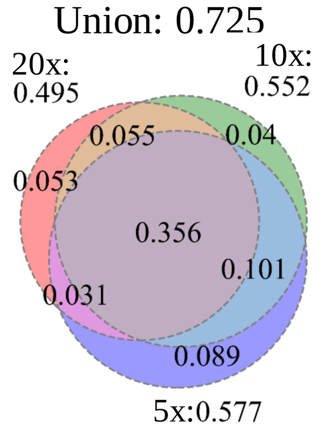
A Part of a pathological image

Patches with different FOV

- Trade-off between the field of view and the spatial resolution:
  - We Investigated how the contextual information is related to the discriminative features for segmenting cancer subtypes.
  - The Venn diagrams show that the contextual information from different magnification images is effective.



The similarity of the predictions of the CNNs which were trained particular magnification images



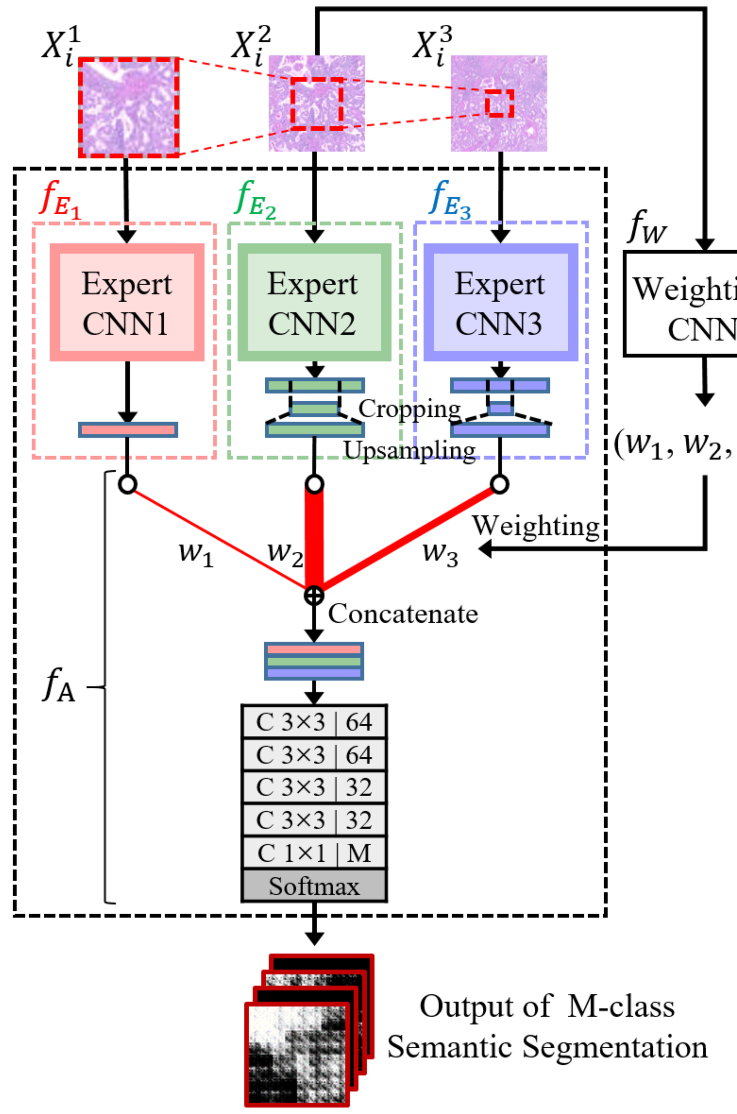
A circle: a set of pixels predicted correctly

A value: the ratio of pixels correctly predicted

Union: the set of pixels that either expert predicted a correct class

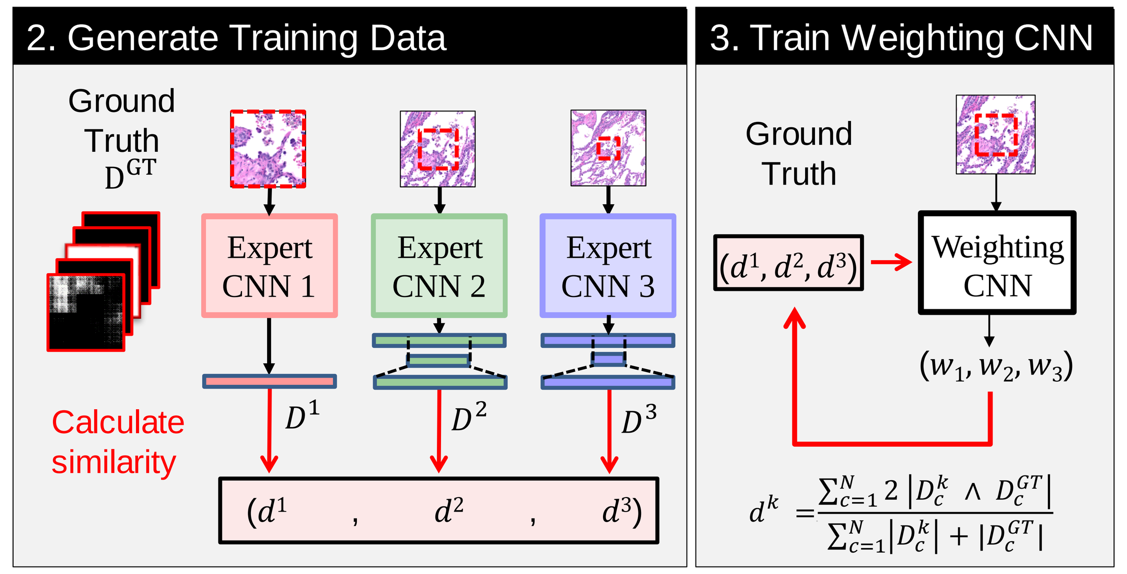
Adaptive Weighting Multi-Field-of-View CNN

- Our proposed model aggregates expert CNNs by adaptively weighting each expert depending on the input image.
  - Expert CNN: is specialized in segmenting particular magnification images.
  - Weighting CNN: adaptively estimates the weights of the expert CNNs.
  - Aggregating CNN: concatenates the outputs of expert CNNs with the estimated weights. This outputs the final segmentation result.

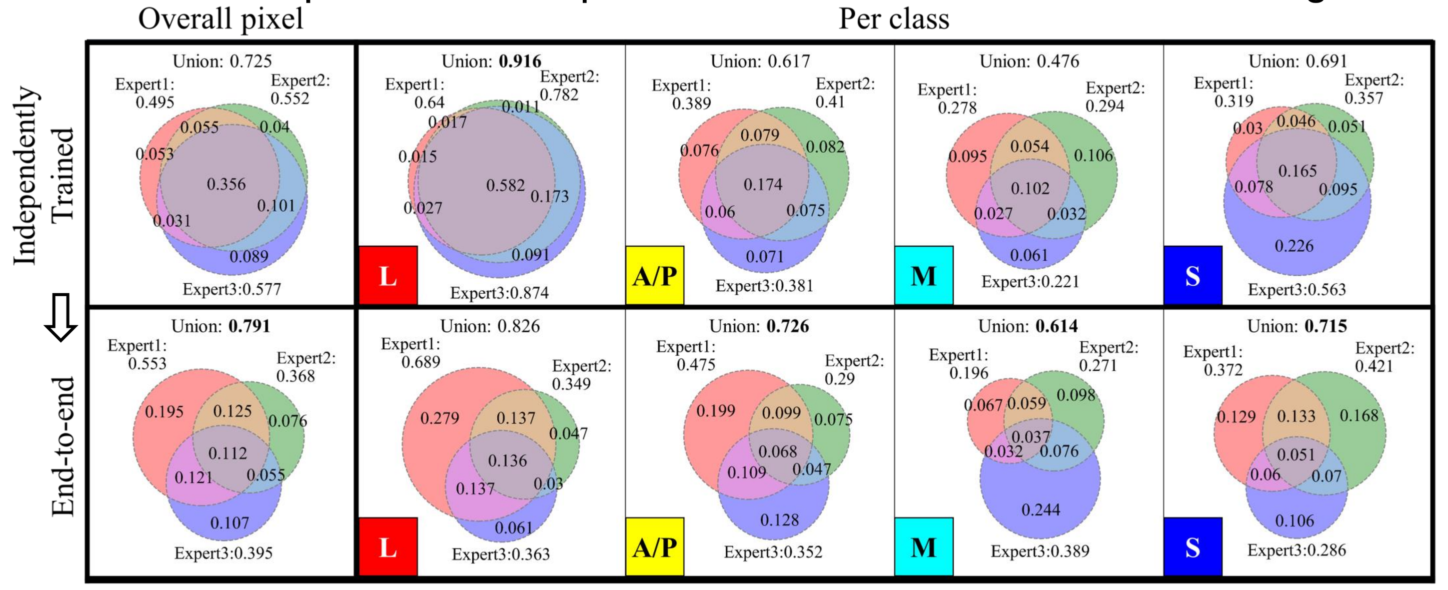


Training algorithm

1. Pre-train Expert CNNs $\{f_{E_k}\}_{k=1}^3$
2. Generate training data for Weighting CNN
3. Train Weighting CNN f_W
4. Train the integrated network $(\{f_{E_k}\}_{k=1}^3, f_A)$
5. Repeat 2-4 until convergence



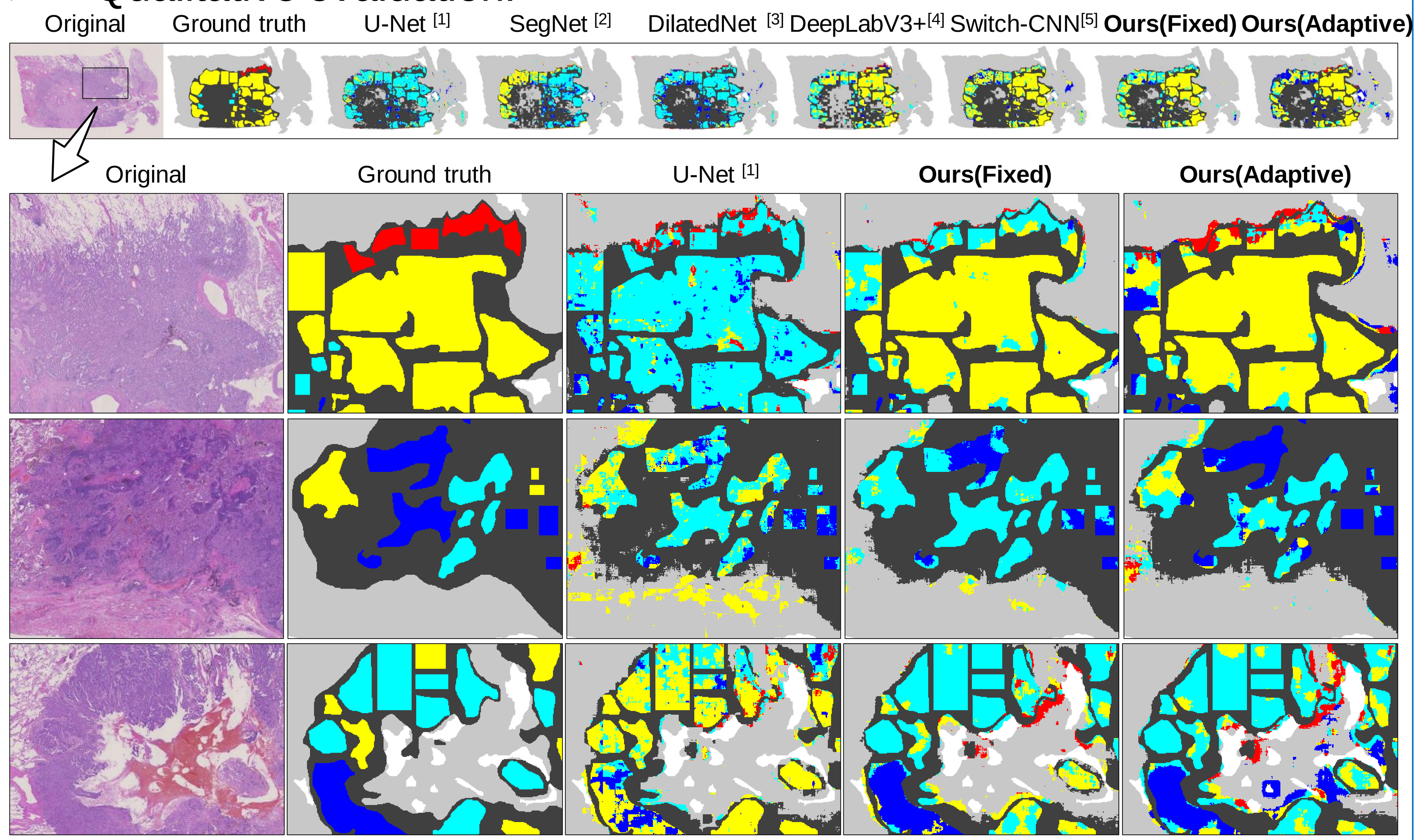
The change in the correct answer rate for pre-trained Experts and after the end-to-end learning



After the end-to-end learning, although the circle of each expert became small on average, the union of the prediction results became large.

Experiments

- Two-version proposed models:
  - Ours (Adaptive): Adaptive Weighting Multi-Field-of-View CNN
  - Ours (Fixed): Fixed Weighting Multi-Field-of-View CNN that uses the fixed weight 1.0 for aggregation
- Qualitative evaluation:



- Quantitative evaluation:

Four-class subtypes segmentation

Network	Magnification	OP	PC	mIoU
U-net	20x	0.446	0.446	0.300
U-net	10x	0.484	0.481	0.331
U-net	5x	0.524	0.537	0.379
SegNet	20x	0.477	0.477	0.320
SegNet	10x	0.547	0.544	0.398
SegNet	5x	0.492	0.525	0.326
Dilated-net	20x	0.433	0.422	0.274
Dilated-net	10x	0.445	0.456	0.314
Dilated-net	5x	0.515	0.528	0.378
DeepLabv3+	20x	0.585	0.580	0.438
DeepLabv3+	10x	0.625	0.624	0.474
DeepLabv3+	5x	0.588	0.583	0.433
Hard-Switch-CNN	(20x,10x,5x)	0.486	0.484	0.347
Ours (Fixed)	(20x,10x,5x)	0.641	0.642	0.505
Ours (Adaptive)	(20x,10x,5x)	**0.672**	**0.676**	**0.536**

Four-class segmentation accuracy of Ours(Adaptive) using other experts

Expert network	Magnification	mIoU	improvement
U-net	(20x,10x,5x)	0.536	0.157
SegNet	(20x,10x,5x)	0.459	0.061
Dilated-net	(20x,10x,5x)	**0.537**	0.159
DeepLabv3+	(20x,10x,5x)	0.510	0.036

- Conclusion:
  - Our model achieved the best performance and the fixed version was second best on four-class subtypes segmentation task.

References:
[1] O. Ronneberger+, U-net: Convolutional networks for biomedical image segmentation, MICCAI 2015.
[2] V. Badrinarayanan+, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, TPAMI 2017
[3] F. Yu+, Multi-scale context aggregation by dilated convolutions, ICLR 2016.
[4] L. Chen+, Rethinking atrous convolution for semantic image segmentation, arXiv 2017.
[5] D. B. Sam+, Switching convolutional neural network for crowd counting, CVPR 2017.