Point-Unet: A Context-aware Point-based Neural Network for Volumetric Segmentation

Ngoc-Vuong Ho[1] Tan Nguyen[1] Gia-Han Diep[3] Ngan Le[4] Binh-Son Hua[1,2]

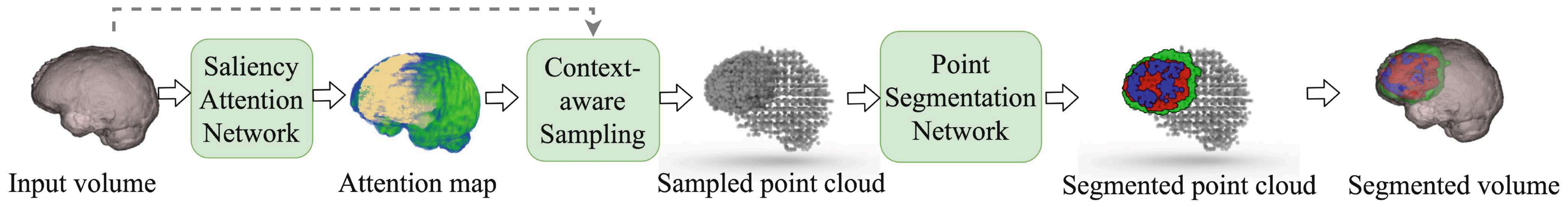
[1]VinAI Research [2]VinUniversity [3]University of Science, VNU - HCM [4]University of Arkansas

MICCAI2021

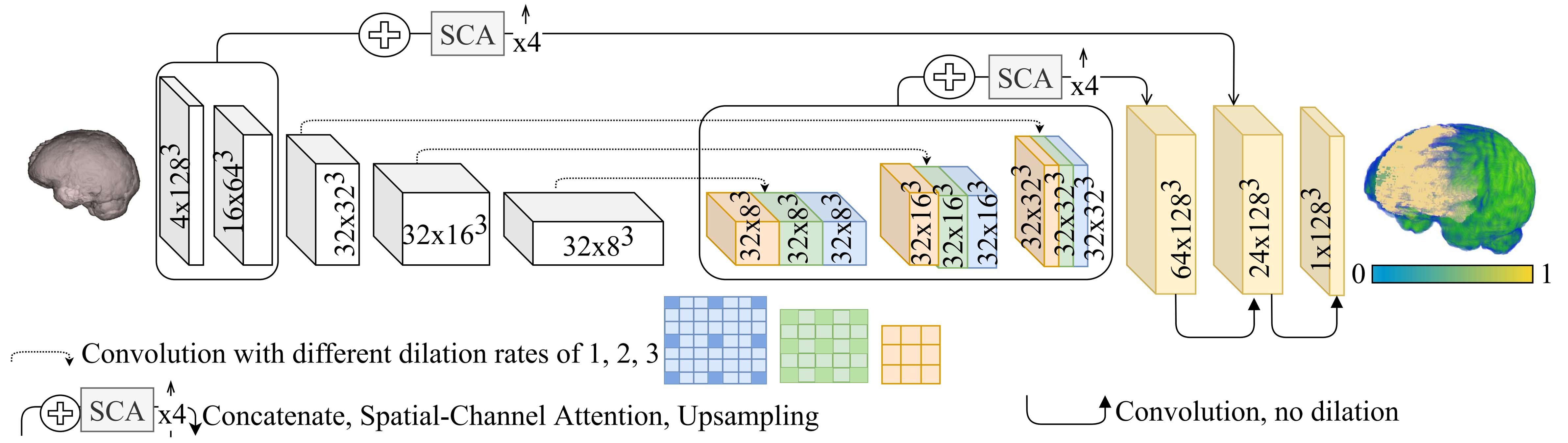
Contributions

- Point-Unet, a new network base on Point Cloud sloved 3D Medical Segmentation
- A saliency proposal network to extract an attentional probability map for an efficient point sampling
- The state-of-the-art accuracy and efficiency are achieved on 3D medical image semantic segmentation

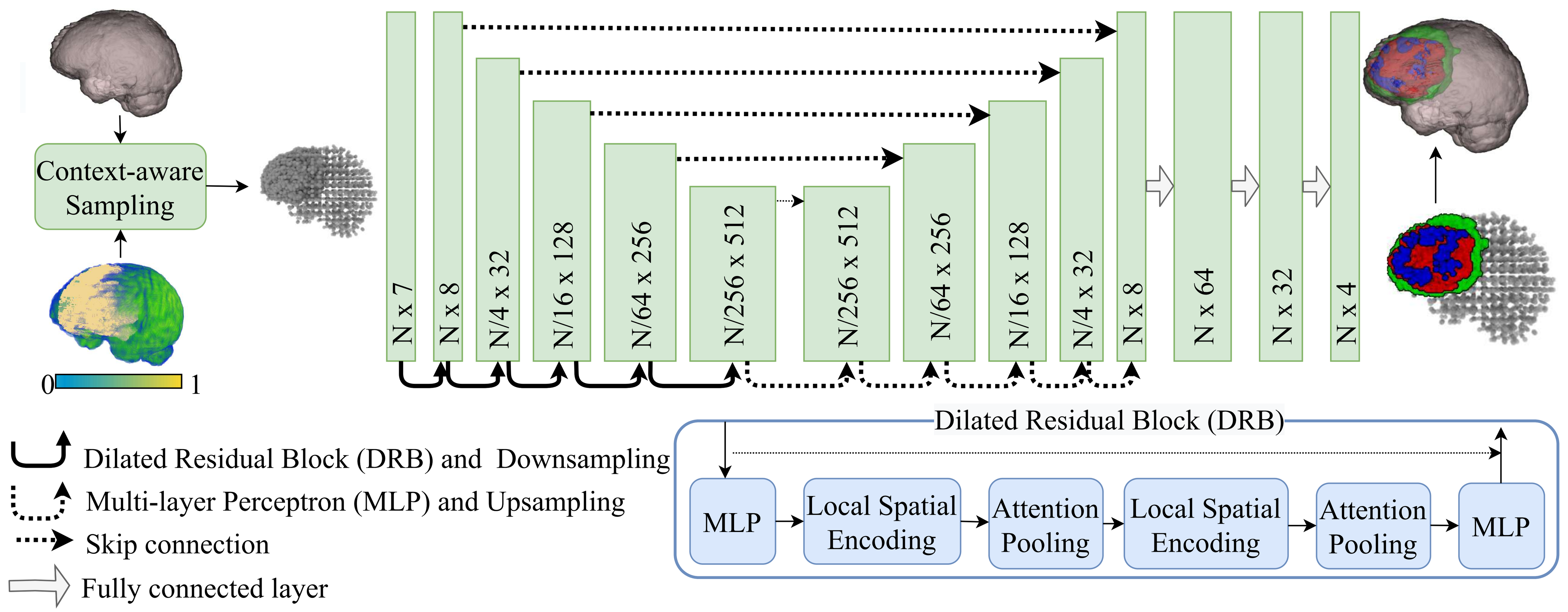
Point-Unet



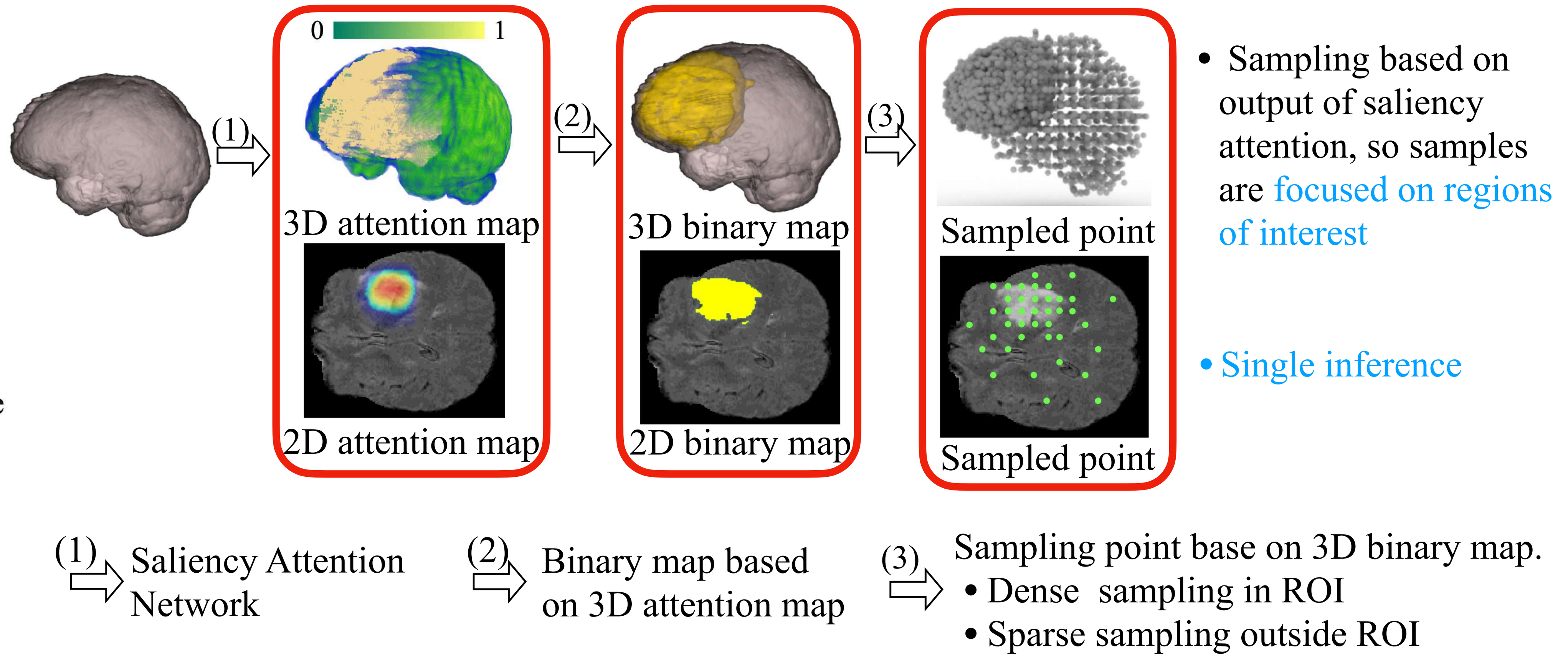
Saliency Attention Network



Point Segmentation Network



Context-aware Sampling



- Sampling based on output of saliency attention, so samples are focused on regions of interest
- Single inference

Quantitative Results

BraTS20	Offline validation set			Online validation set			
Methods	Dice score ↑		HD95 ↓	Methods	Dice score ↑		HD95 ↓
	ET/WT/TC	AVG	AVG		ET/WT/TC	AVG	AVG
3DUnet [15]	66.92/82.86/72.98	74.25	30.19	3DUNet [35]	67.66/87.35/79.30	78.10	21.16
nnNet [12]	73.64/80.99/81.60	78.74	*14.33*	nnNet [12][a]	68.69/81.34/78.06	78.03	24.30
aeUnet [25]	*71.31*/*84.72*/*79.02*	*78.35*	15.43	nnUNet [10][c]	*77.67*/**90.60**/**84.26**	**84.18**	15.30
				aeNet [25][a]	64.00/83.16/74.66	73.95	33.91
				Cascade [21][4]	78.81/89.92/*82.06*	*83.60*	*12.00*
				KiUNet [33]	73.21/87.60/73.92	78.24	**8.38**
RandLA [8]	67.40/87.74/76.85	77.33	**7.03**	RandLA [8]	66.31/88.01/77.03	77.17	16.65
Ours	**76.43/89.67/82.97**	**83.02**	8.26	**Ours**	**78.98**/*89.71*/82.75	83.81	11.73

[a] Reproduce the results on the network trained with 100 epochs.
[b] We choose the model with as similar batch size as ours.
[c] We choose single model with 190 epoches, stage 1. The best model at Brats2020.

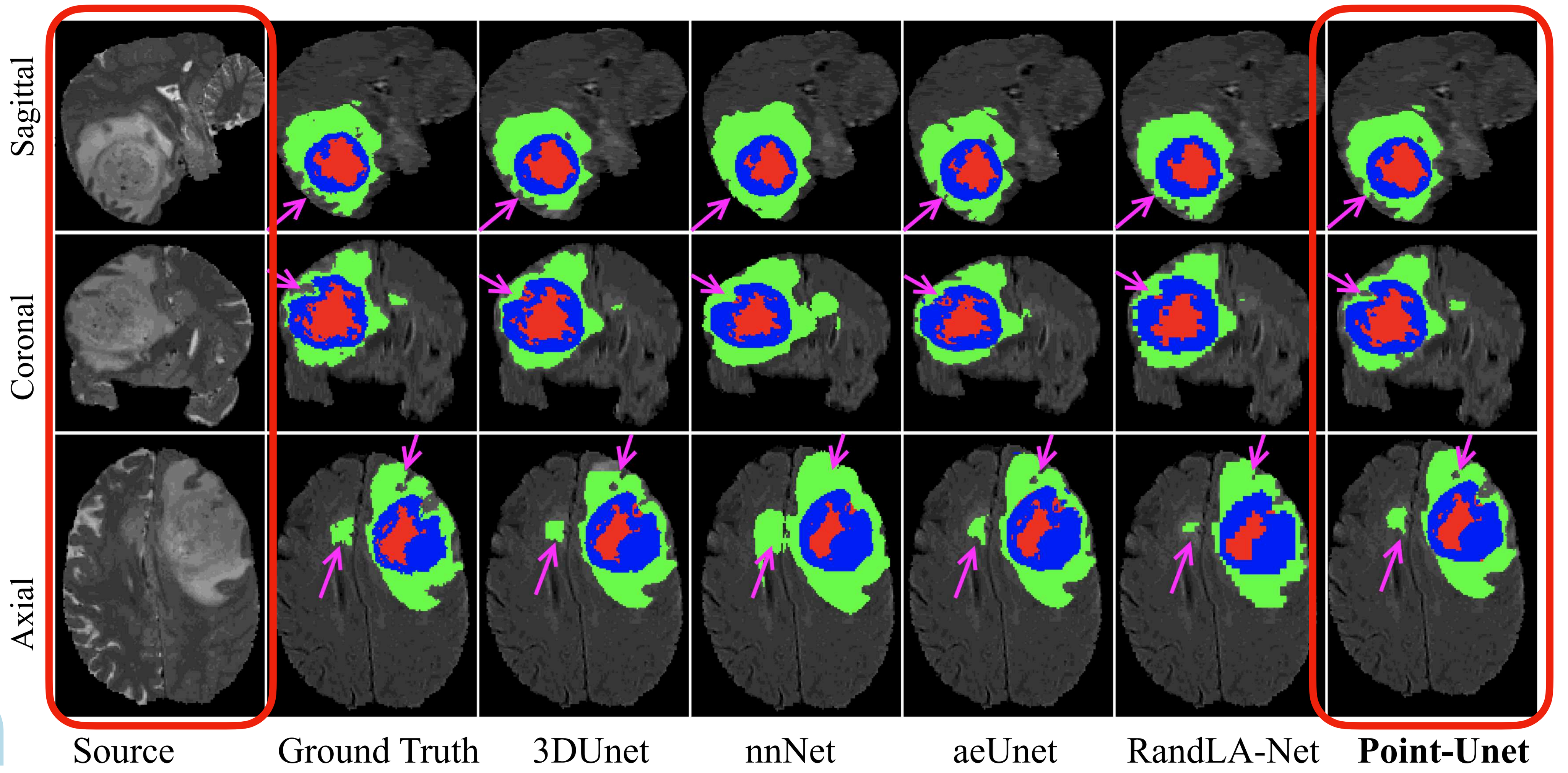
Comparison on BraTS20. The **best**, second best and *third best* are highlighted

Method	Average ↑	Method	Average ↑
Oktay et al. [27]	83.10 ± 3.80	Yu et al. [40]	84.50 ± 4.97
Zhu et al. [43]	84.59 ± 4.86	**Ours**	**85.68** ±5.96

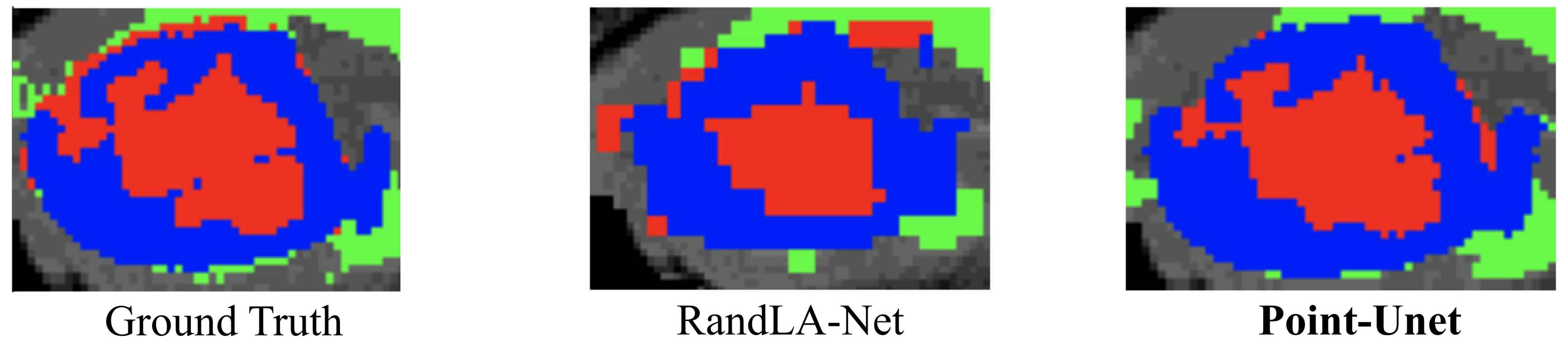
Dice score comparison on Pancreas dataset

Code: https://github.com/VinAIResearch/Point-Unet

Quantitative Results

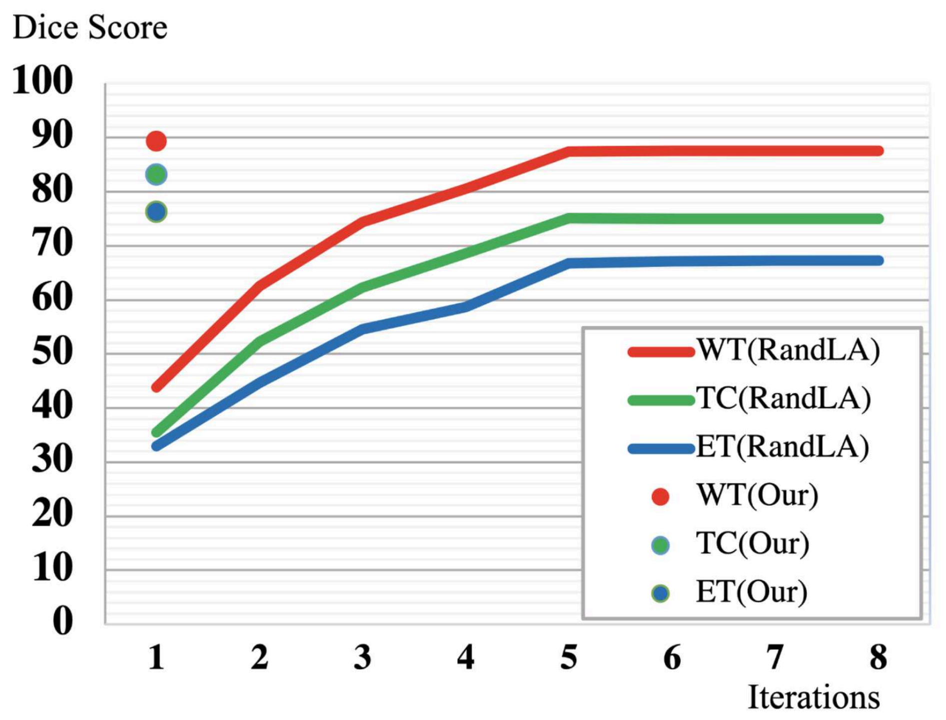


Efficiency of Context-aware Sampling



Point-Unet has good boundaries. RandLA-Net has several zigzag artifact at boundaries

Performance Analysis



Networks	Patch size	Memory ↓	Inference ↓
3DUnet baseline [15]	128 × 128 × 128	8.75 GB	7.80 s
	160 × 192 × 128	16.70 GB	0.28 s
	240 × 240 × 144	32.00 GB	0.23 s
nnNet [12]	128 × 128 × 128	7.20 GB	55.30 s
	160 × 192 × 128	11.10 GB	26.50 s
	240 × 240 × 144	21.70 GB	2.30 s
aeUnet [25]	128 × 128 × 128	17.21 GB	110.40 s
	160 × 192 × 128	31.42 GB	78.80 s
	240 × 240 × 144	>48.00 GB	7.10 s
RandLA[8]	240 × 240 × 155	15.98 GB	8.00 s
(Ours)	240 × 240 × 155	17.22 GB	1.24 s

Point-Unet requires just a single iteration for inference, having high accuracy and relatively low memory footprint.