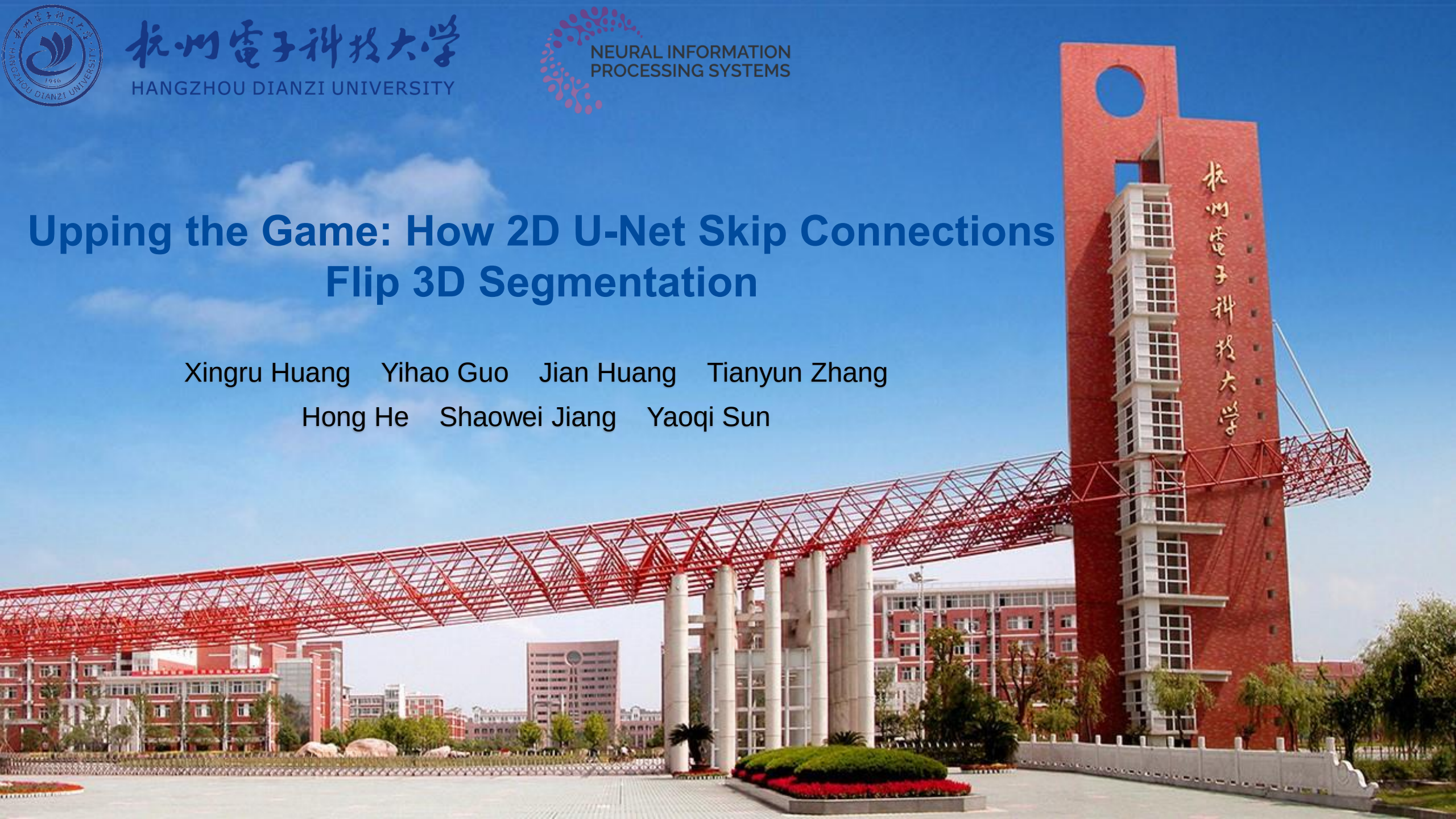
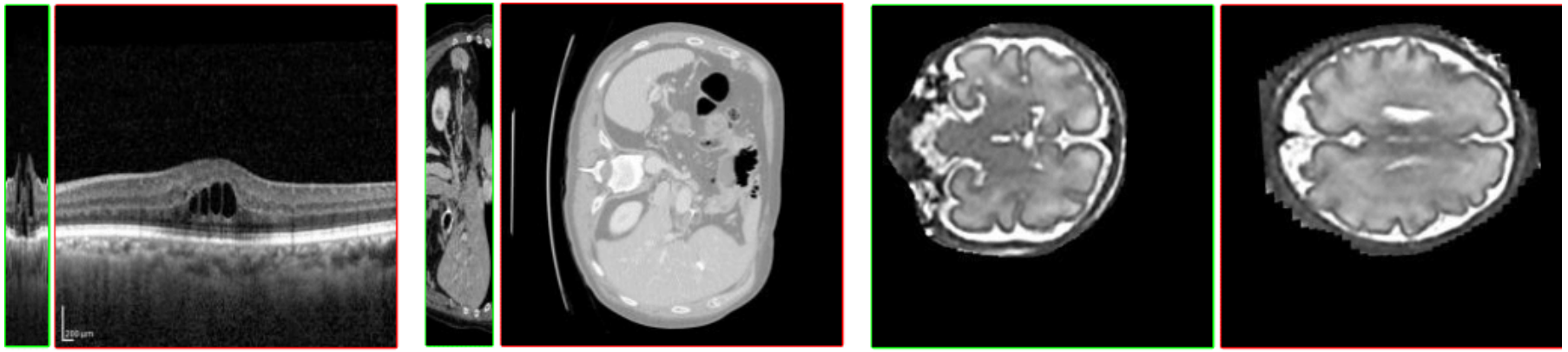

杭州電子科技大学
HANGZHOU DIANZI UNIVERSITY
NEURAL INFORMATION PROCESSING SYSTEMS
Upping the Game: How 2D U-Net Skip Connections Flip 3D Segmentation
Xingru Huang Yihao Guo Jian Huang Tianyun Zhang
Hong He Shaowei Jiang Yaoqi Sun

Motivation

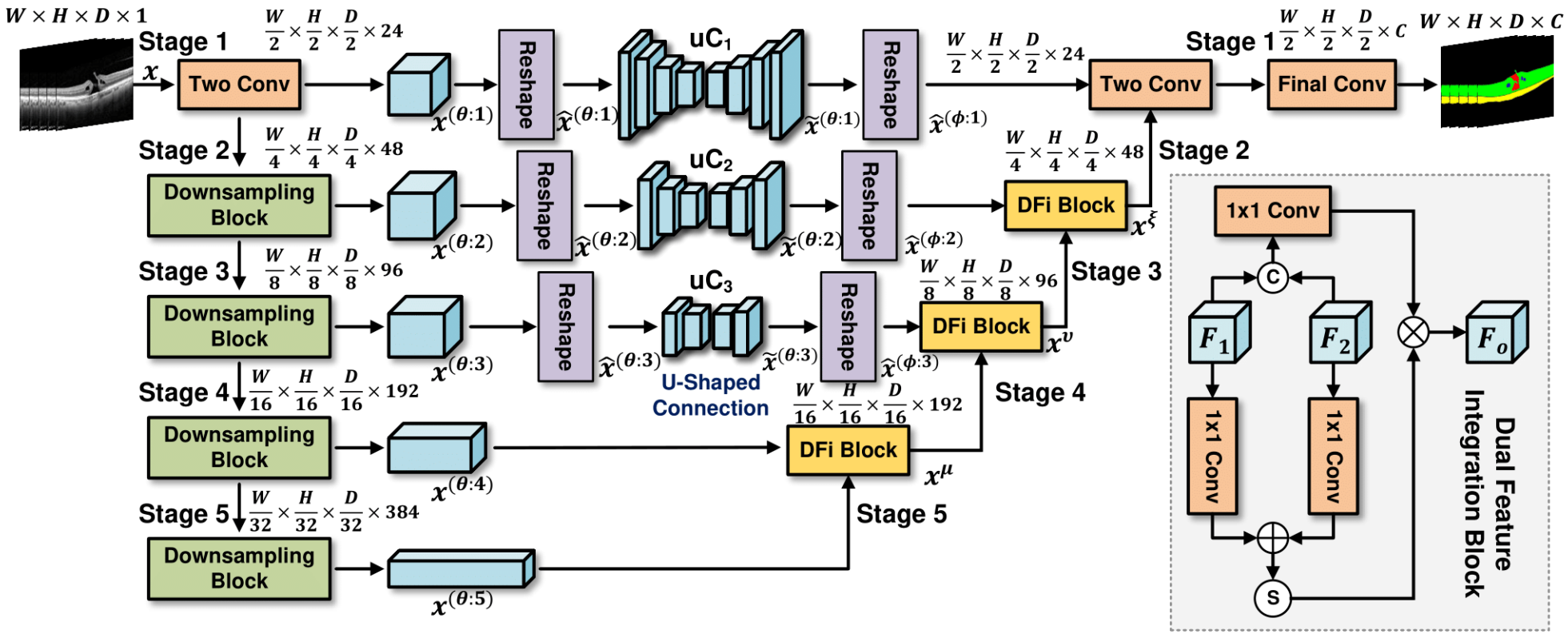
Conventional 3D segmentation techniques predominantly depend on isotropic 3D convolutions for the extraction of volumetric features, which frequently engenders inefficiencies due to the varying information density across the three orthogonal axes in medical imaging modalities such as computed CT and MRI. To address this issue, we introduce the U-shaped Connection (uC), utilizing simplified 2D U-Net in place of standard skip connections to augment the extraction of the axial-slice plane features while concurrently preserving the volumetric context afforded by 3D convolutions.

Anisotropy in Medical Imaging



Comparison of information density between time-axial (Green) and slice planes (Red). Panels (a), (b), and (c) show these differences for volumetric data of the abdomen, retina, and brain tissues, respectively.

Network Structure:



Overview of the proposed uC 3DU-Net architecture. We propose 2D U-Net skip connection (uC) to replace conventional skip connections in 3D segmentation networks and intergate it in 3D U-Net.

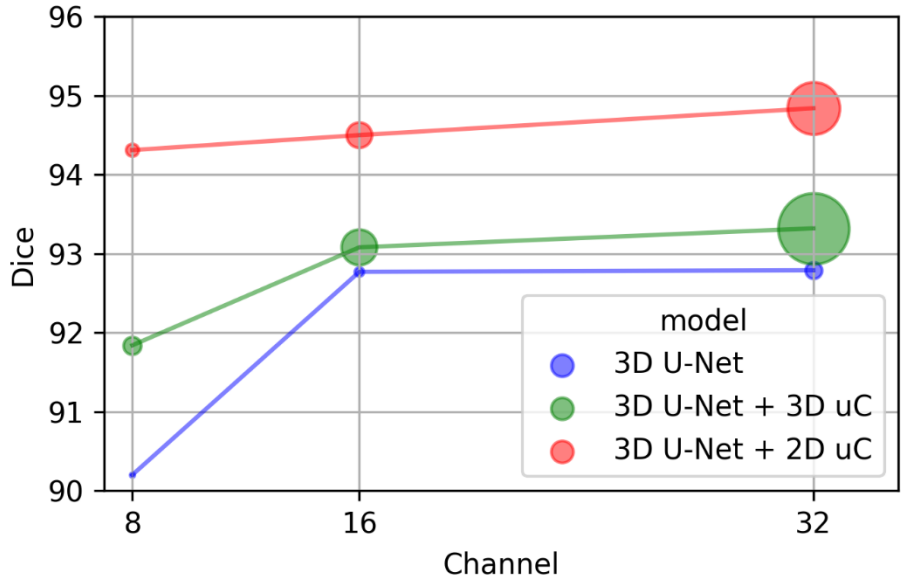
Image Reconstruction Fidelity Comparison

	3D U-Net(32)	+3D uC	+2D uC
PSNR	28.98db	34.93db	**36.29db**

PSNR results of the reconstruction experiments on the OIMHS dataset, including three methods: 3DU-Net, 3DU-Net + 3DuC, and 3DU-Net + 2DuC. 3DU-Net + 2DuC achieve best reconstruction performance.

Impact of Convolution Dimensionality on uC performance and Parameters.

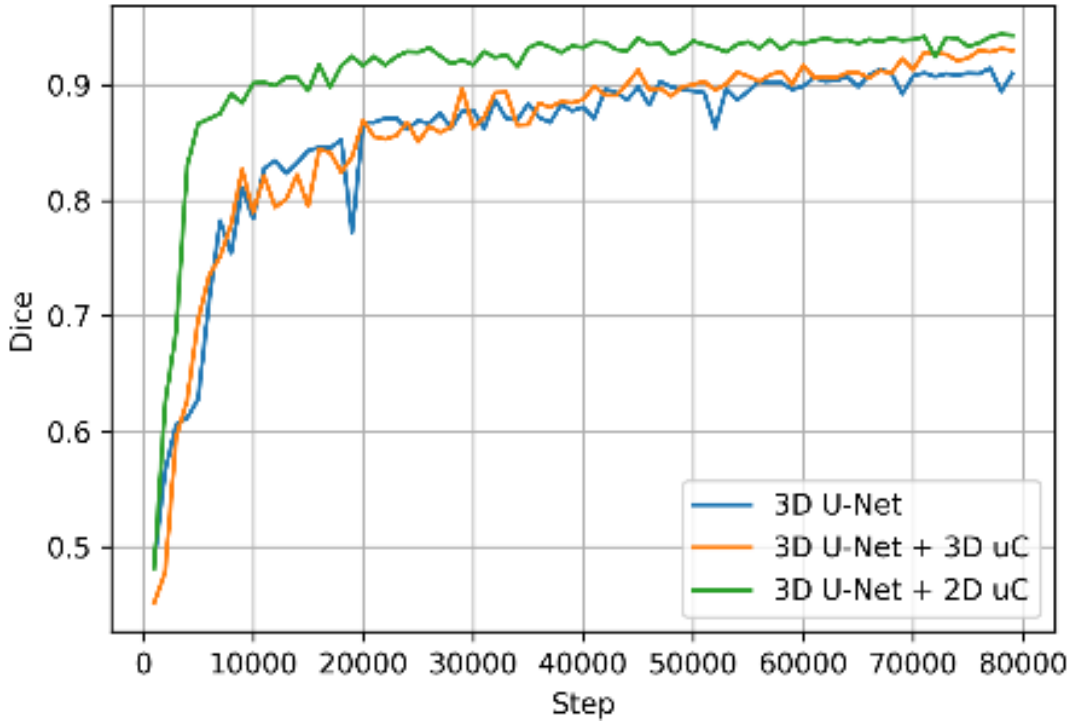
Methods	#Params	FLOPs	mIoU	Dice	ASSD	HD95	AdjRand
3D U-Net (8)	0.37M	16.50G	83.17±4.51	90.20±3.20	2.88±2.16	16.92±21.89	89.51±3.23
+3D uC	5.56M	46.80G	85.58±3.60	91.84±2.35	2.25±1.35	8.64±14.07	91.19±2.40
+2D uC	2.85M	42.54G	**89.59±2.61**	**94.31±1.62**	**0.52±0.48**	**2.57±4.74**	**93.84±1.65**
3D U-Net (16)	1.47M	65.09G	87.08±3.43	92.77±2.24	0.92±0.92	4.57±8.92	92.20±2.29
+3D uC	22.21M	186.27G	87.50±3.34	93.08±2.11	4.61±1.25	43.45±19.56	92.41±2.15
+2D uC	11.38M	169.20G	**89.98±3.43**	**94.50±2.23**	**0.39±0.42**	**2.56±4.96**	**94.06±2.26**
3D U-Net (32)	4.81M	135.90G	87.08±3.32	92.79±2.09	1.31±1.35	7.39±14.69	92.22±2.15
+3D uC	88.67M	596.06G	87.93±3.33	93.32±2.09	3.09±1.31	14.13±16.99	92.73±2.12
+2D uC	45.33M	527.75G	**90.52±2.77**	**94.84±1.68**	**0.35±0.33**	**2.20±4.07**	**94.43±1.76**



Results of 3D U-Net, 3D U-Net + 3D uC, and 3D U-Net + 2D uC across different channel depth numbers on the OIMHS dataset. The best values for each metric are highlighted in bold.

Performance comparison of 3D U-Net, 3D U-Net + 3D uC, and 3D U-Net + 2D uC across varying channel depths, with circle size representing parameter count. The horizontal axis represents channel depth numbers.

Impact of Convolution Dimensionality on uC performance and Parameters.



Validation curve showing Dice scores on the OIMHS dataset, with the channel depth for all three models(3D U-Net, 3D U-Net + 3D uC, and 3D U-Net + 2D uC) set to 32. The horizontal axis represents the number of training iterations.

Comparative Experiments Results:

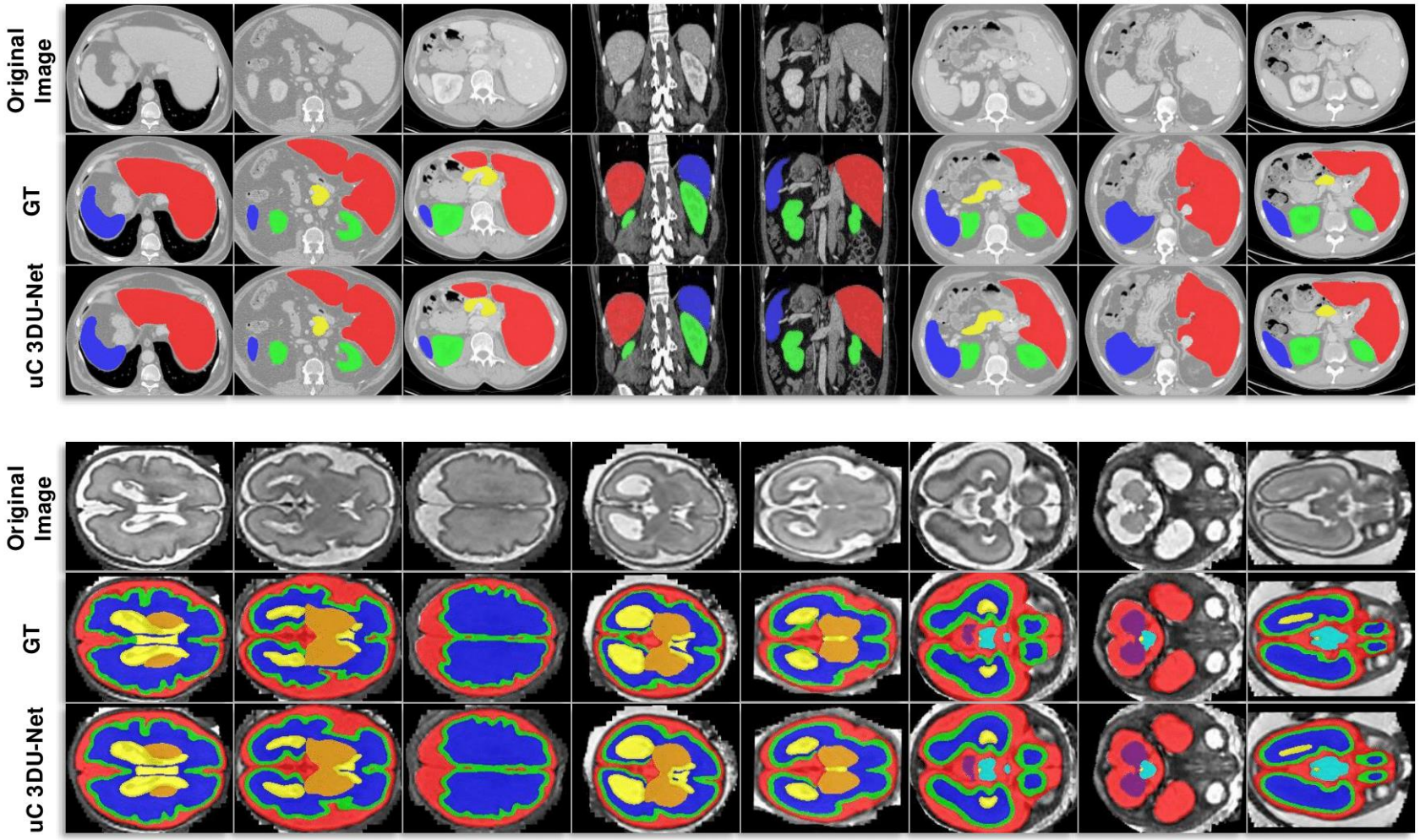
Methods	#Params	FLOPs	FLARE2021					FeTA2021							
			Mean	Spleen	Kidney	Liver	Pancreas	Mean	ECF	GM	WM	Vent	Cereb	DGM	BS
3D U-Net [57]	4.81M	135.9G	89.2	91.1	96.2	90.5	78.9	85.7	86.7	76.2	92.5	86.1	91.0	84.5	82.7
SegResNet [58]	1.18M	15.6G	90.2	96.3	93.4	96.5	74.5	86.2	86.8	77.0	92.7	86.5	91.1	86.7	82.5
RAP-Net [59]	38.2M	101.2G	91.3	94.6	96.7	94.0	79.9	86.5	88.0	77.1	92.7	86.2	90.7	87.9	83.2
nn-UNet [35]	31.2M	743.3G	92.6	97.1	96.6	97.6	79.2	87.0	88.3	77.5	93.0	86.8	**92.0**	88.0	84.0
TransBTS [32]	31.6M	110.4G	90.2	96.4	95.9	97.4	71.1	86.8	**88.5**	77.8	93.2	86.1	91.3	87.6	83.7
UNETR [29]	92.8M	82.6G	88.6	92.7	94.7	96.0	71.0	86.0	86.1	76.2	92.7	86.2	90.8	86.8	83.4
nnFormer [60]	149.3M	240.2G	90.6	97.3	96.0	97.5	71.7	86.3	88.0	77.0	93.0	85.7	90.3	87.6	82.8
SwinUNETR [34]	62.2M	328.4G	92.9	97.9	96.5	98.0	78.8	86.7	87.3	77.2	92.9	86.9	91.4	87.5	84.2
3D UX-Net [56]	53.0M	639.4G	93.4	**98.1**	**96.9**	98.2	80.1	87.4	88.2	78.0	**93.4**	87.2	91.7	**88.6**	84.5
uC 3DU-Net	21.7M	286.4G	**94.1±1.61**	98.0±1.02	96.7±1.45	**98.3±0.76**	**83.2±5.37**	**87.8±1.99**	87.8±2.57	**80.7±2.29**	92.6±1.89	**89.1±3.52**	91.0±2.47	86.3±5.01	**87.5±2.32**

Method	#Params	FLOPs	mIoU	Dice	VOE	HD95	AdjRand
3D U-Net [57]	4.81M	135.9G	86.02	92.05	13.98	6.77	91.34
Swin UNETR [34]	62.2M	328.4G	86.73	92.53	13.27	5.09	91.85
3D UX-Net [56]	53.0M	639.4G	87.43	92.90	12.57	4.41	92.27
SASAN [61]	22.96M	282.92G	88.44	93.53	11.56	3.14	92.96
nnFormer [60]	149.3M	240.2G	72.16±7.91	81.60±7.41	27.84±25.07	23.49±7.91	80.36±7.73
TransBTS [32]	31.6M	110.4G	74.80±7.31	83.08±6.55	25.20±23.82	31.43±7.31	82.05±6.85
UNETR [29]	92.8M	82.6G	80.52±6.68	88.11±5.41	19.48±30.31	30.07±6.68	87.21±5.56
uC 3DU-Net	21.7M	286.43G	**89.48±3.56**	**94.13±2.56**	**10.52±5.61**	**2.98±3.56**	**93.62±2.66**

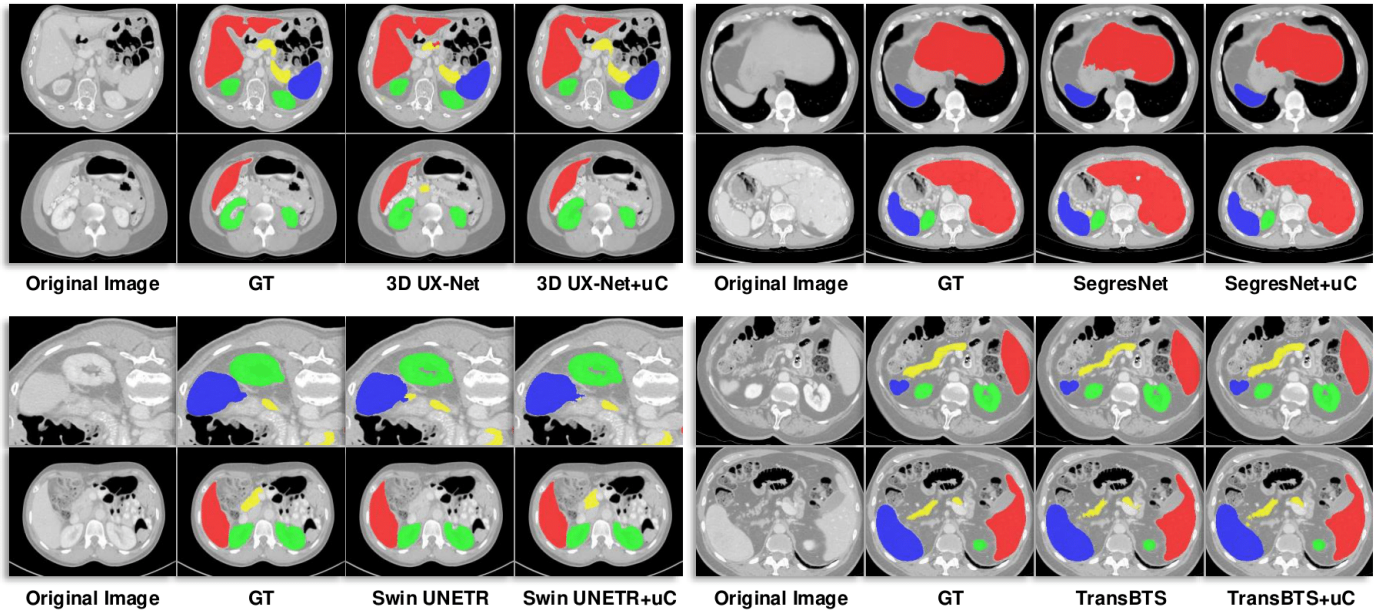
Method	#Params	FLOPs	mIoU	Dice	ASSD	HD95	AdjRand
3D U-Net [57]	4.81M	135.9G	86.69±4.30	92.28±2.93	2.31±2.06	12.68±18.54	92.14±2.98
TransBTS [32]	31.6M	110.4G	70.06±9.11	79.64±8.12	4.88±1.59	35.75±10.00	79.35±8.19
UNETR [29]	92.8M	82.6G	84.17±5.12	90.42±3.75	2.36±1.68	13.41±16.77	90.27±3.81
nnFormer [60]	149.3M	240.2G	80.69±7.76	87.56±6.36	2.01±1.80	9.75±9.95	87.38±6.41
Swin UNETR [34]	62.2M	328.4G	86.76±4.94	92.36±3.20	2.67±1.81	14.87±18.78	92.22±3.26
3D UX-Net [56]	53.0M	639.4G	86.56±4.76	92.21±3.18	2.60±1.98	15.46±19.75	92.07±3.24
uC 3DU-Net	21.7M	286.4G	**88.29±4.04**	**93.35±2.60**	**1.48±1.09**	**8.53±10.74**	**93.22±2.66**

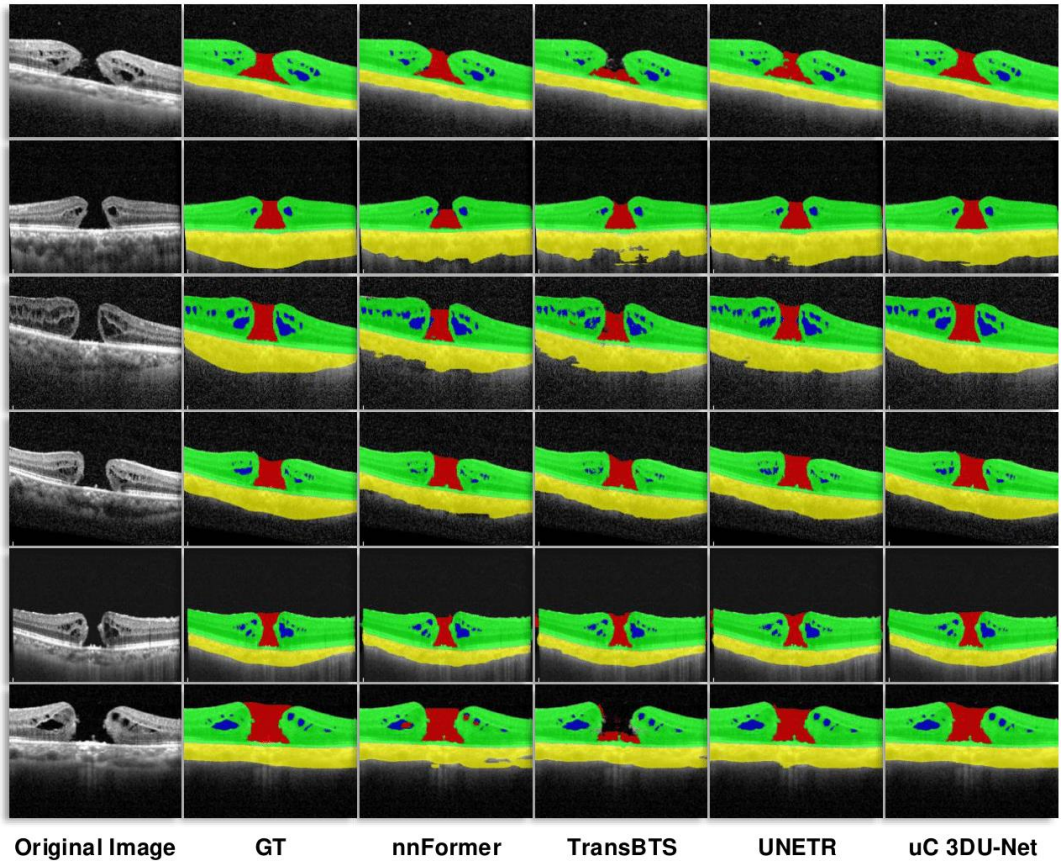
Comparative experimental results of the proposed uC3DU-Net and previous methods on the FLARE2021, FeTA 2021, OIMHS, and AbdomenCT-1K datasets. The best values for each metric are bolded. The proposed approach achieves SOTA performance.

Visualization Results



Visualization Results





Thank you!

Code: [IMOP-lab/U-Shaped-Connection]

Contact information: j.huang@hdu.edu.cn