Automatic Segmentation of GTV of NPC for Radiotherapy Planning via a Fine-tuned UniSeg Model (Task 02)

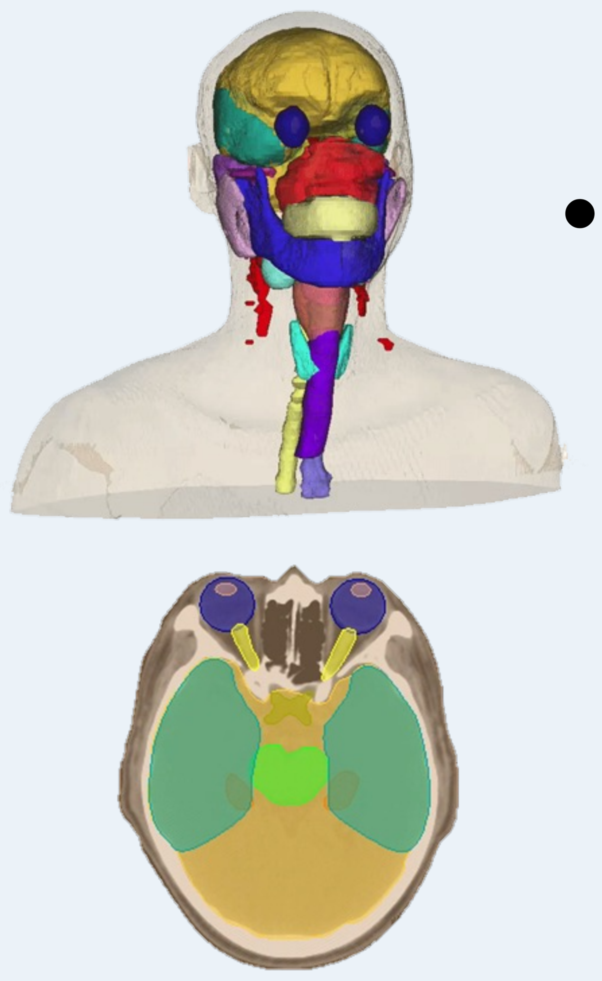
Yiwen Ye[1], Ziyang Chen[1], and Yong Xia[1,2]

[1]National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering

[2]Ningbo Institute of Northwestern Polytechnical University, Ningbo 315048, China

1 Introduction

➢Task & Challenges

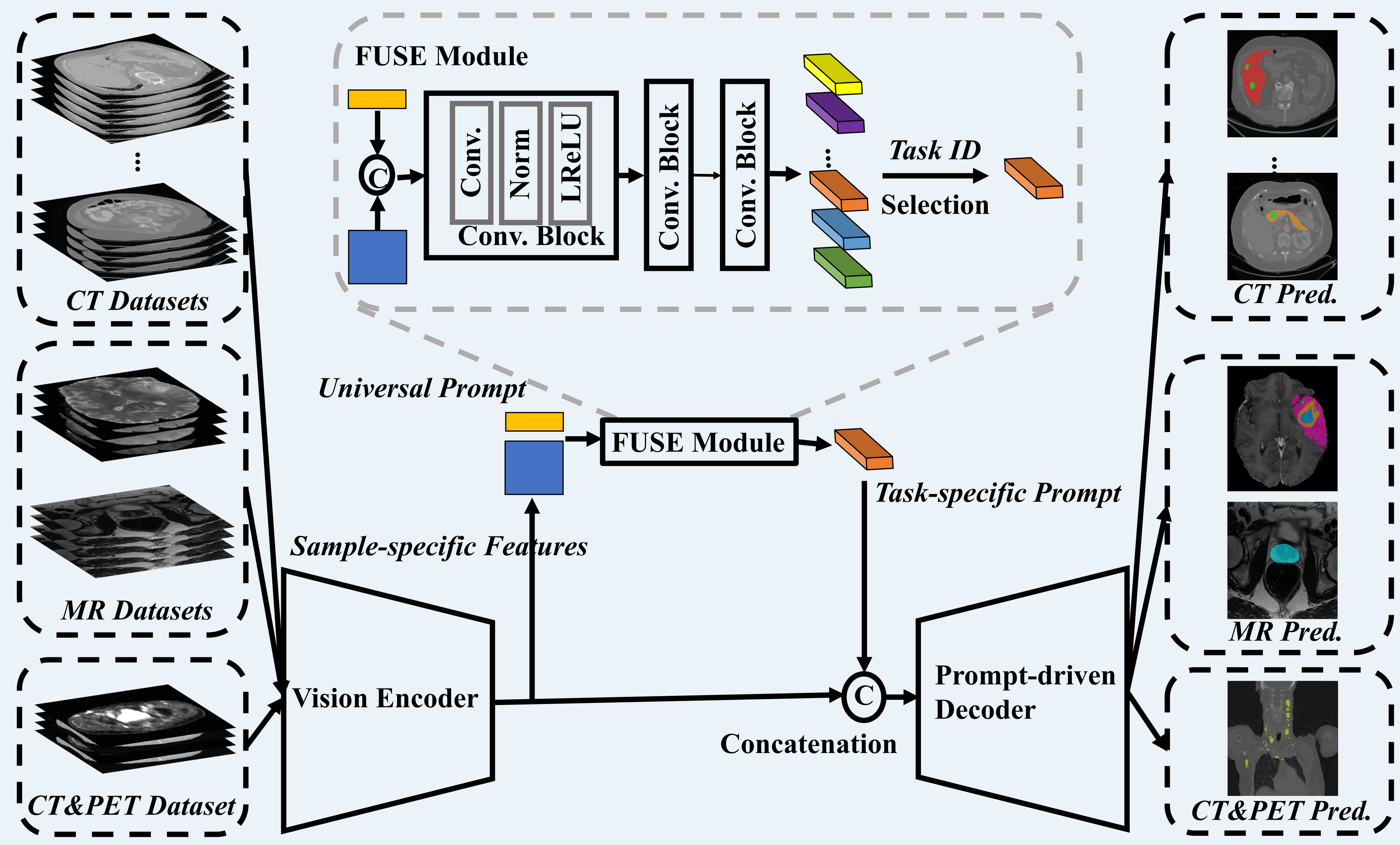


- Task:
  a) Segment 2 Gross Target Volumes from CT images with high accuracy and robustness.
- Challenges:
  a) **Heterogeneity:** Tumors may vary greatly in size, shape, and density.
  b) **Proximity to Critical Organs:** The GTV may be close to organs-at-risk, necessitating extremely precise segmentation to avoid collateral damage to healthy tissue.
  c) **Class Imbalance:** The GTV usually occupies a much smaller portion of the total image volume, leading to class imbalance problems.

➢Motivation

Despite nnUNet has laudable performance, even outpacing seasoned experts in some clinical scenarios. nnUNet trained from scratch still leaves room for enhancement, particularly through transfer learning. To address this gap, our previous work developed a supervised pre-trained nnUNet model called UniSeg[1].

2 A Brief on UniSeg



Technical pipeline of our UniSeg, including a vision encoder, FUSE module, and a prompt-driven decoder. The sample-specific features produced by the encoder are concatenated with a learnable universal prompt as the input of the FUSE module. Then the FUSE module produces the task-specific prompt, which enables the model to be `aware' of the ongoing task.

3 Powerful Transfer Performance of UniSeg

Details of eleven upstream datasets and two downstream datasets.

Dataset	Upstream											Downstream	
	CT								MR		CT&PET	CT	MR
	Liver	Kidney	HepaV	Pancreas	Colon	Lung	Spleen	VerSe20	Prostate	BraTS21	AutoPET	BTCV	VS
Organ	✓	✓	✓	✓	×	×	✓	×	✓	×	×	✓	×
Tumor	✓	✓	✓	✓	✓	✓	×	×	×	✓	✓	×	✓
Vertebrae	×	×	×	×	×	×	×	✓	×	×	×	×	×
Train	104	168	242	224	100	50	32	171	91	1000	400	21	193
Test	27	42	61	57	26	13	9	43	25	251	101	9	49

Results of SSL and SL models. Blue numbers are used to indicate the performance gain using per-trained weights. We repeat all experiments three times and report mean values.

Dataset		MG	GVSL	SMIT	UniMiSS	DeSD	DoDNet	UniSeg
BTCV	Sp	86.8	90.6	90.7	95.0	96.1	96.4	96.2
	RKi	85.5	92.3	92.1	92.9	94.6	94.5	94.4
	LKi	83.0	91.2	91.9	91.5	93.2	89.7	91.6
	Gb	63.5	63.7	63.0	67.1	64.4	68.3	68.4
	Es	70.5	72.5	74.8	73.6	75.2	76.9	77.9
	Li	92.4	95.6	95.7	96.4	96.6	96.8	96.7
	St	78.3	80.1	75.9	82.4	88.7	86.5	87.8
	Ao	88.5	87.5	88.6	88.9	90.0	89.8	90.1
	IVC	85.3	84.4	86.4	83.9	87.5	87.7	87.6
	PSV	70.7	71.7	72.8	73.2	75.1	76.1	76.7
	Pa	71.4	72.7	74.3	76.2	79.9	81.9	83.3
	RAG	68.7	68.1	71.3	67.1	70.4	73.2	73.4
	LAG	58.2	63.6	69.5	67.0	70.5	75.2	75.1
	Mean	77.1 (+2.7)	79.5 (+1.9)	80.6 (+1.3)	81.2 (+3.0)	83.3 (+0.8)	84.1 (+0.9)	84.6 (+1.4)
VS	Tumor	79.3 (+7.2)	91.0 (+2.2)	92.2 (+2.3)	91.4 (+2.0)	92.2 (+1.5)	91.8 (+1.1)	92.9 (+2.1)

4 Implementation Details

➢Transfer Learning:
Transfer the pre-trained encoder-decoder to the SegRap Task.

➢Dataset:
Totally 120 no-contrast and contrast-enhanced CT with voxel-level annotations of 45 organs

➢Training and Inference Framework
nnUNet framework

➢Preprocessing
Partition each training sample into more compact patches, each with a size of 96 × 256 × 256,
Perform nnUNet's preprocessing operation without resampling.

➢Key settings:
Patch size: 64 × 192 × 192; Batch size: 2; 5-fold cross-validation.

5 Results on the Validation Set

Results of our method and the other top 5 methods on the validation leaderboard for Task 2. The best result for each metric is highlighted in blue.

Rank	DSC_Mean	NSD_Mean	DSC_Std	NSD_Std
1 (Ours)	79.19	0.7441	0.0780	0.1103
2	79.01	0.7397	0.0895	0.1198
3	79.02	0.7393	0.0824	0.1119
4	78.87	0.7307	0.0732	0.1079
5	78.48	0.7312	0.0874	0.1103

Performance of several solutions on Task 2. All results are obtained using 5-fold cross-validation.The best result on each metric is highlighted in blue.

Method	DSC_Mean	NSD_Mean	DSC_Std	NSD_Std
ResUNet w/ 1000 epoch	77.99	0.7234	0.0805	0.1110
UniSeg w/ 1500 epoch	79.13	0.7424	0.0808	0.1142
UniSeg w/ 1000 epoch (Ours)	79.19	0.7441	0.0780	0.1103

[1] Ye Y, Xie Y, Zhang J, et al. UniSeg: A Prompt-driven Universal Segmentation Model as well as A Strong Representation Learner[J]. arXiv preprint arXiv:2304.03493, 2023.