RadOmni: Advancing Foundation Model for Non-contrast CT with Omni Radiology Knowledge

Fenghe Tang^, Weiwei Cao^, Wenxin Ma, Kai Cao, Huanhuan Liu, Ling Zhang, Shaohua Kevin Zhou*, Jiangpeng Zhang*

達摩院 DAMO ACADEMY

奇跡 Medical Imaging, Robotic, Analytic Computing & Learning

NEURAL INFORMATION PROCESSING SYSTEMS

NCCT is often accompanied By complementary clinical knowledge

NCCT

Refer on CECT

CECT

ROI Confirm

Label ROI

Write Reports

Radiology Report

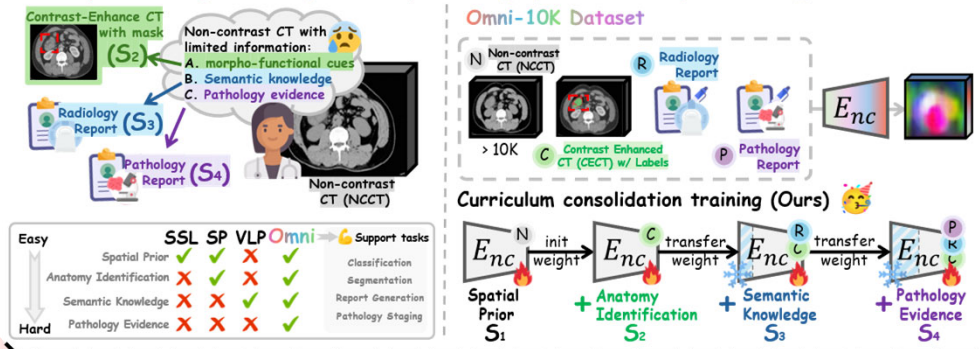
Pathology Confirm

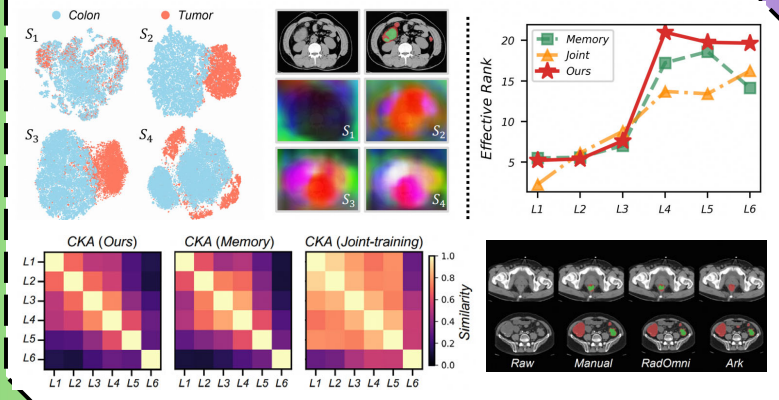
Pathology Report

"I Am RadOmni"

Motivation

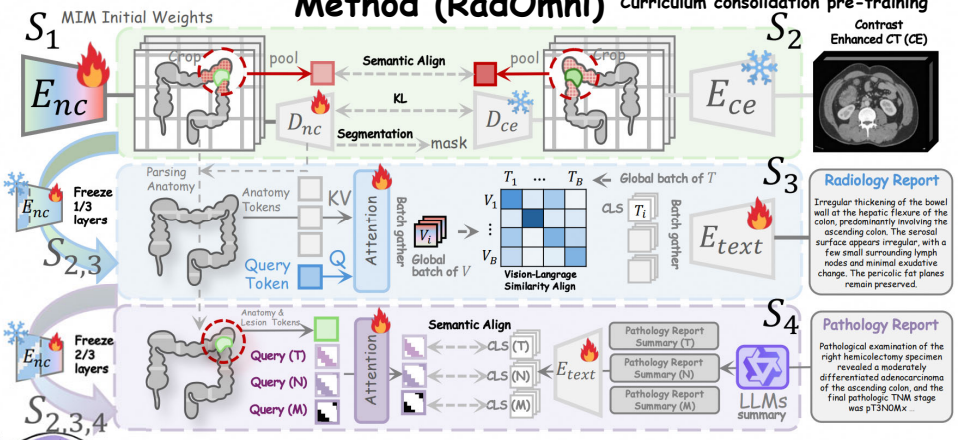
Can pre-training with complementary knowledge push NCCT beyond its diagnostic limits?





Method (RadOmni)

Curriculum consolidation pre-training

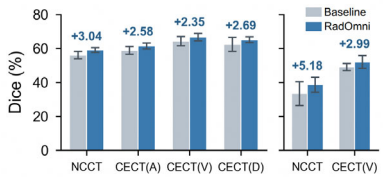


Omni-CT10K

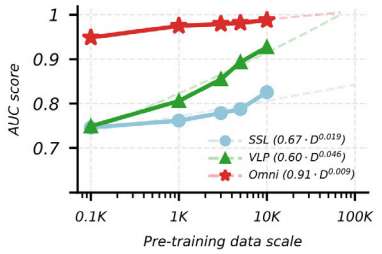
Omni-pretraining Hospital A → Fintune → **Fine-tuning Hospital B** → Generalize → **Zero-shot Hospital C**

Omni-CT10K	ExtBench A (4K)	ExtBench B (0.5K)
6K normal & 4K CRC	2K normal & 2K CRC	0.2K normal & 0.3K CRC
10K paired reports	4K paired reports	0.5 K paired reports
~6K paired CECT	~1.3K paired pathology reports	~0.3K paired pathology reports
~4K paired pathology report		

Method (Backbone)	Cls. (AUC %)	Seg. (DSC %) Organ	Seg. (DSC %) Tumor	Pathology (AUC %) T	N	M	Avg	Report Generation (%) BLEU-4	ROUGE-L	CIDEr	SPICE	GREEN	MSD Colon
Train from scratch													
Swin [40]	90.55	92.90	45.75	62.23	58.05	39.25	53.18	0.01	21.62	2.21	18.58	32.34	35.13*
nnUNet [39]	85.44	94.41	53.82	55.94	51.40	30.89	29.41	0.61	23.62	2.19	21.64	20.59	41.69*
Self-supervised pretraining													
VoCov2 [29] (Swin)†	97.51	93.18	48.43	60.63	57.23	38.70	52.19	0.01	20.28	1.85	7.59	19.20	42.57*
SparK [25, 26, 11] (nnUNet)	95.80	94.51	53.95	57.30	52.53	41.99	50.61	0.01	26.83	1.83	28.97	14.48	44.47
Supervised pretraining													
SuPreM [18] (Swin)†	98.43	92.78	46.23	61.77	59.25	48.70	56.57	0.66	25.66	2.49	23.29	18.12	38.78*
SuPreM [18] (nnUNet)	97.39	94.85	58.30	63.91	57.26	53.42	58.20	0.01	27.29	1.97	24.43	24.34	48.42*
UniModel [17] (UNet)†	94.07	91.53	39.16	55.93	39.73	59.20	51.62	1.02	25.91	2.73	25.15	35.27	45.43
Vision-Language pretraining													
CLIP [41, 33] (nnUNet)	88.08	94.32	51.82	62.40	54.23	39.11	51.91	3.38	29.03	8.76	28.84	28.34	37.93
fVLM [6] (nnUNet)	97.28	94.18	54.47	63.74	52.05	11.44	42.41	1.97	28.41	4.78	27.41	22.31	41.76
Merlin [7] (ResNet)†	94.76	94.40	57.89	52.36	24.73	58.48	45.19	0.01	28.33	1.71	28.03	17.37	43.21
Multi-modal supervised pretraining													
VoComni [29] (nnUNet)†	98.50	93.87	48.97	59.84	53.15	41.37	51.45	0.01	28.47	1.90	27.59	15.06	43.85
Triad [34] (nnUNet)	98.06	94.54	54.44	64.63	58.17	35.62	52.81	3.30	28.54	11.15	29.58	32.70	45.61
Argus [35] (nnUNet)	98.14	94.85	56.24	66.24	56.52	48.15	56.97	5.36	32.65	16.47	31.52	34.57	49.43
Ark [36] (nnUNet)	98.23	94.98	61.62	66.46	60.43	**74.59**	67.16	0.13	19.17	0.43	21.81	22.05	50.22
Omni-supervised pretraining													
RadOmni (nnUNet)	**98.54**	**95.01**	**63.67**	**67.44**	**63.35**	73.56	**68.12**	**5.97**	**34.34**	**18.10**	**32.18**	**37.02**	**51.02**



Method	Cls. (AUC %)	Pathology (AUC %) T	N	M	Avg
Train from scratch					
Swin [40]	78.97	69.39	54.03	50.39	57.94
nnUNet [39]	61.09	64.27	48.74	47.90	53.64
Self-supervised learning					
VoCov2 [29] (Swin)†	92.03	70.51	53.79	49.96	58.09
SparK [25, 26, 11] (nnUNet)	69.33	64.64	54.20	44.68	54.51
Supervised pretraining					
SuPreM [18] (Swin)†	90.51	66.75	58.33	48.62	57.90
SuPreM [18] (nnUNet)	81.80	67.22	64.04	47.89	59.72
UniModel [17] (UNet)†	68.37	66.32	54.51	51.49	57.44
Vision-language pretraining					
CLIP [41, 33] (nnUNet)	62.08	67.73	56.89	45.22	56.61
fVLM [6] (nnUNet)	88.12	64.67	50.47	44.76	53.30
Merlin [7] (ResNet)†	74.06	69.60	48.96	51.67	56.74
Multi-modal supervised pretraining					
VoComni [29] (nnUNet)†	93.97	67.37	51.94	44.25	54.52
Triad [34] (nnUNet)	92.01	72.97	58.94	43.87	58.59
Argus [35] (nnUNet)	96.81	71.00	60.32	**55.21**	62.18
Ark [36] (nnUNet)	97.58	70.51	63.45	52.01	61.99
Omni-supervised pretraining					
RadOmni (nnUNet)	**98.53**	**74.64**	**64.61**	53.54	**64.27**



Method	MSD (%) Liver	Vessel	Pancreas	Lung	Avg
Swin	67.85	64.80	47.88	70.28	62.70
nnUNet	67.43	68.92	46.08	64.73	61.79
SuPreM (Swin)	68.20	64.71	50.02	72.55	63.87
VoCo (Swin)	71.22	67.25	51.34	**75.74**	66.39
VoCo (nnUNet)	70.09	68.99	50.43	68.99	64.63
RadOmni (nnUNet)	**71.36**	**70.92**	**64.48**	74.38	**70.29**