VQ-Seg: Vector-Quantized Token Perturbation for Semi-Supervised Medical Image Segmentation

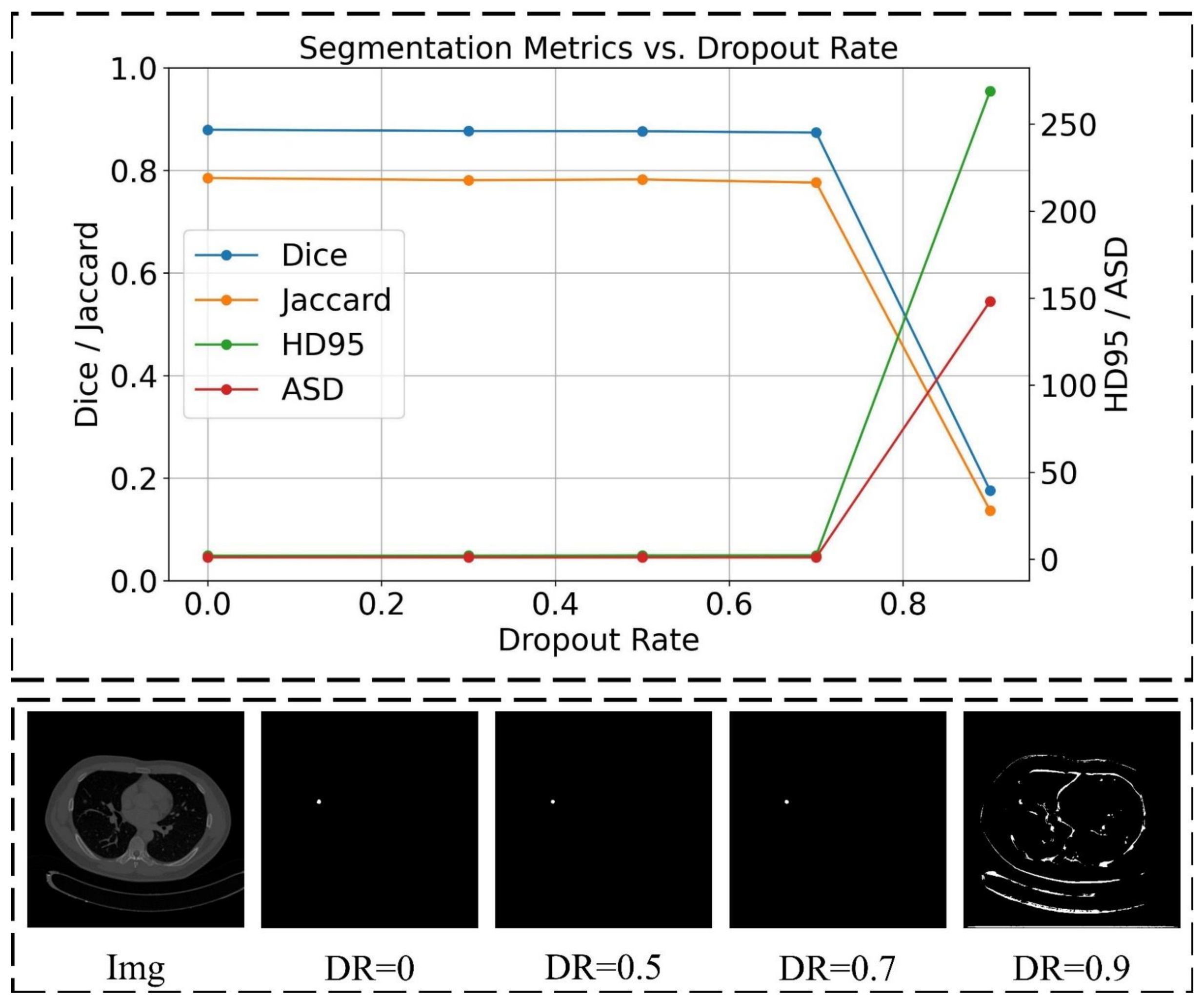
Sicheng Yang*[1], Zhaohu Xing*[1], Lei Zhu[1,2]✉

[1]The Hong Kong University of Science and Technology (Guangzhou)

[2]The Hong Kong University of Science and Technology

Email: syang671@connect.hkust-gz.edu.cn

Problems



Effect of DR on seg performance.

➢Dropout-based feature perturbation in semi-supervised segmentation is sensitive and unstable to tune, motivating the need for a more controllable perturbation mechanism.

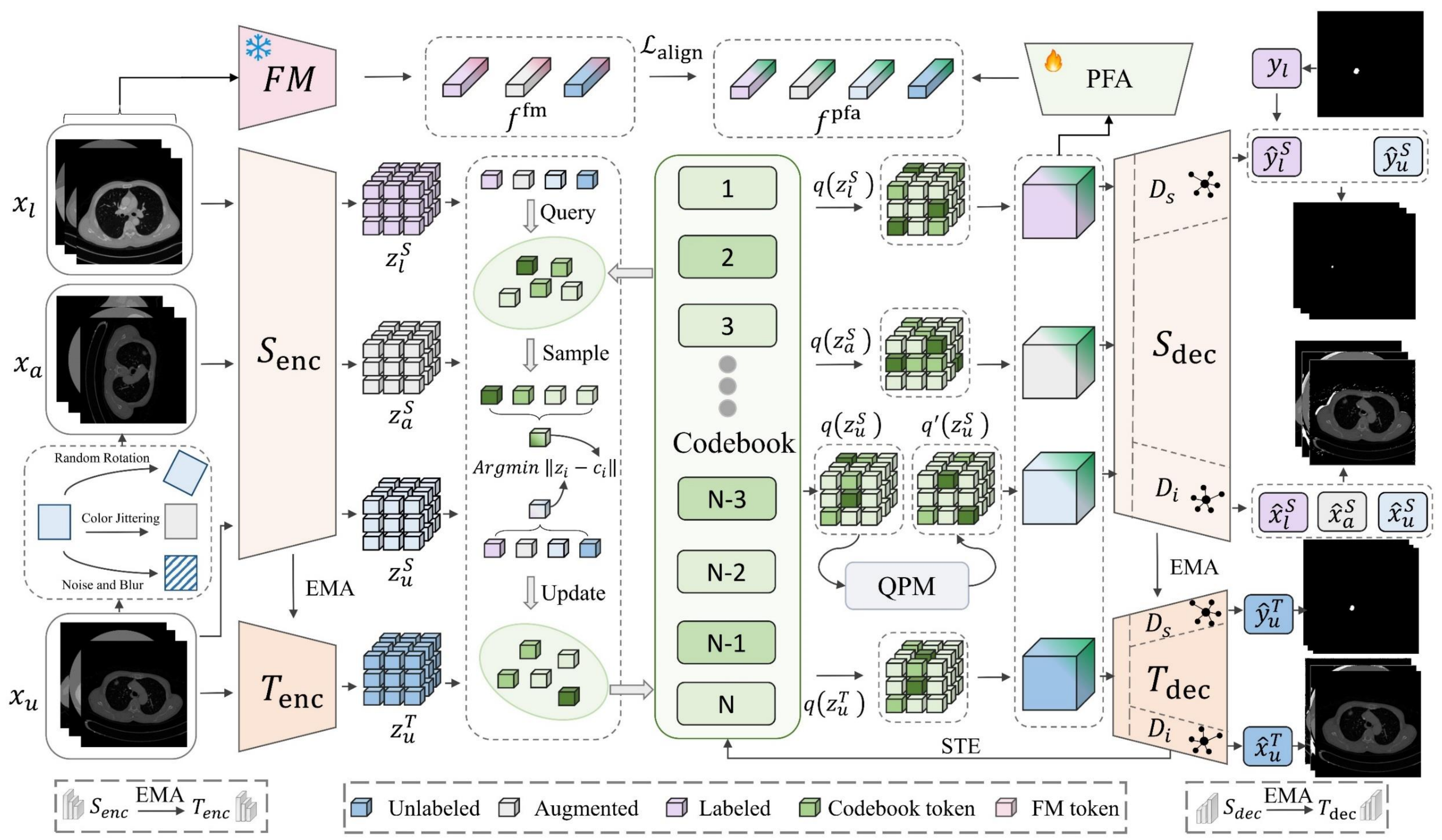
Contribution

➢We propose VQ-Seg with controllable VQ-space perturbation, a dual-branch architecture, and foundation-model-guided alignment.

➢We construct a large-scale dataset containing 828 annotated CT scans.

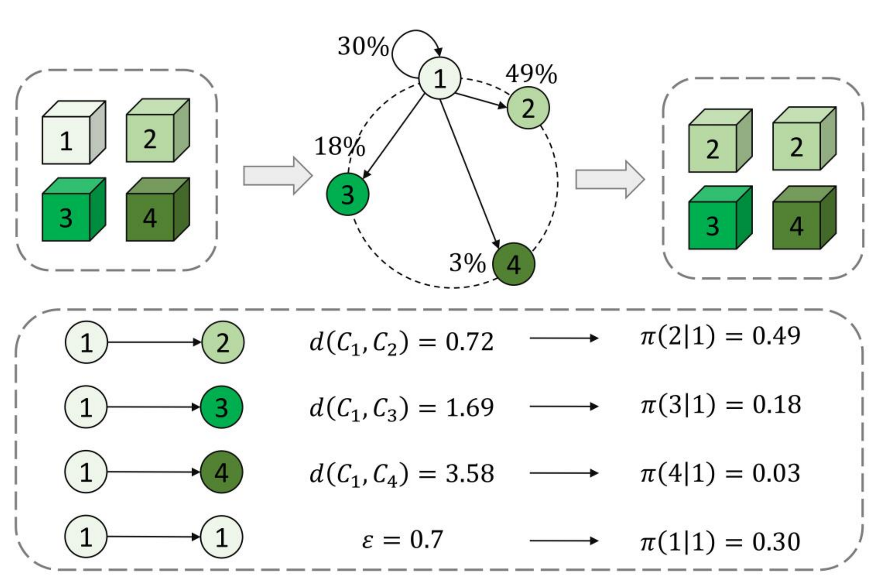
➢Extensive evaluations on multiple benchmarks demonstrate clear state-of-the-art performance.

Methodology

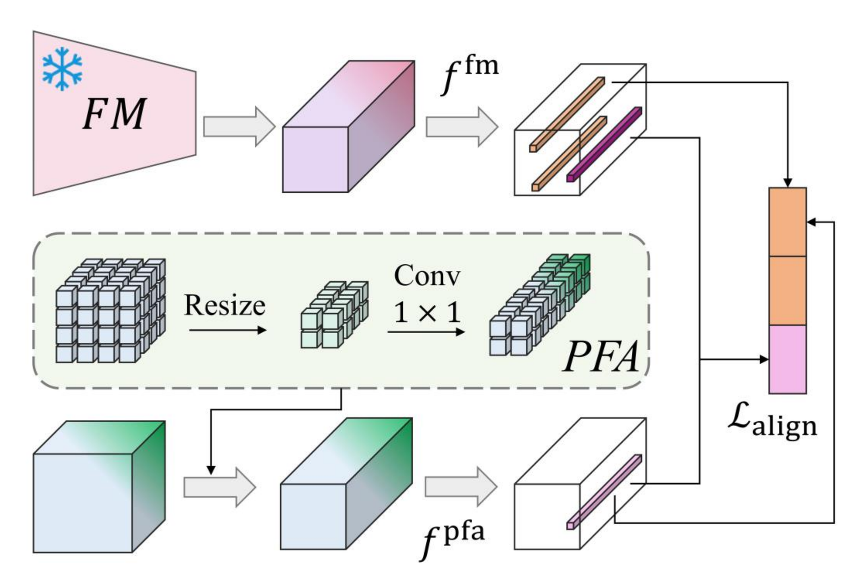


Overview of the VQ-Seg framework.

(a) Quantized Perturbation Module (QPM) introduces controllable perturbations for consistency learning.
(b) Dual-branch architecture jointly optimizes image reconstruction and segmentation using the Post-VQ features.
(c) Post-VQ Feature Adapter (PFA) aligns the quantized features with semantic embeddings from a foundation model.



The QPM module.

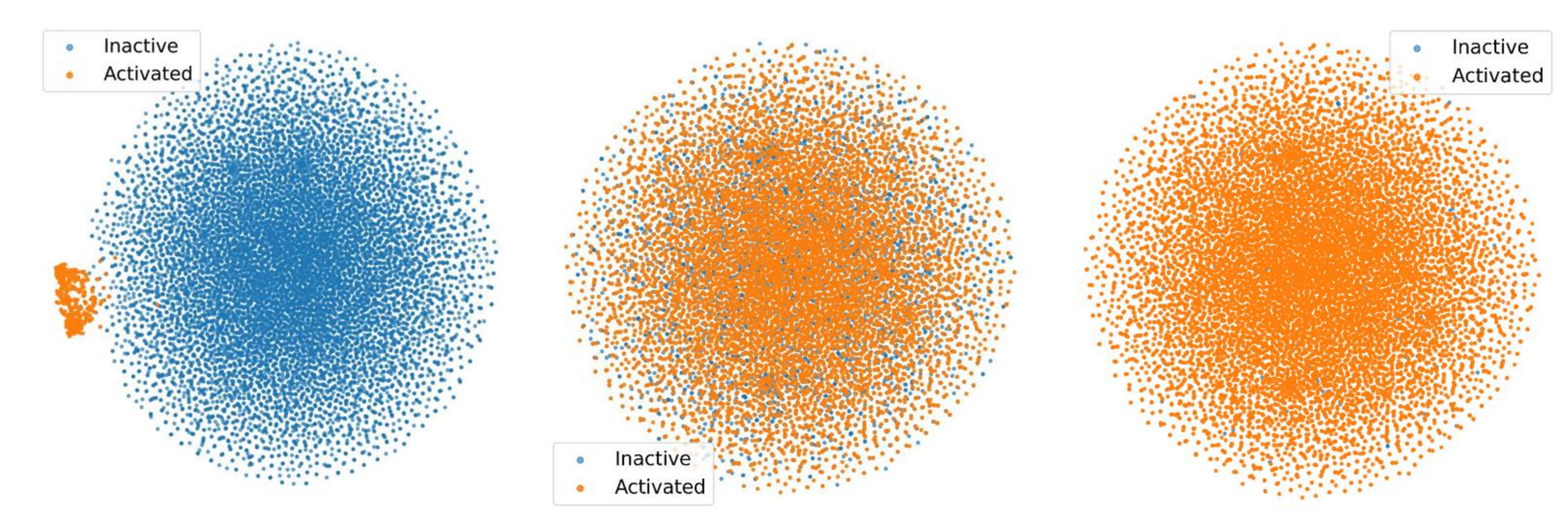


The PFA module.

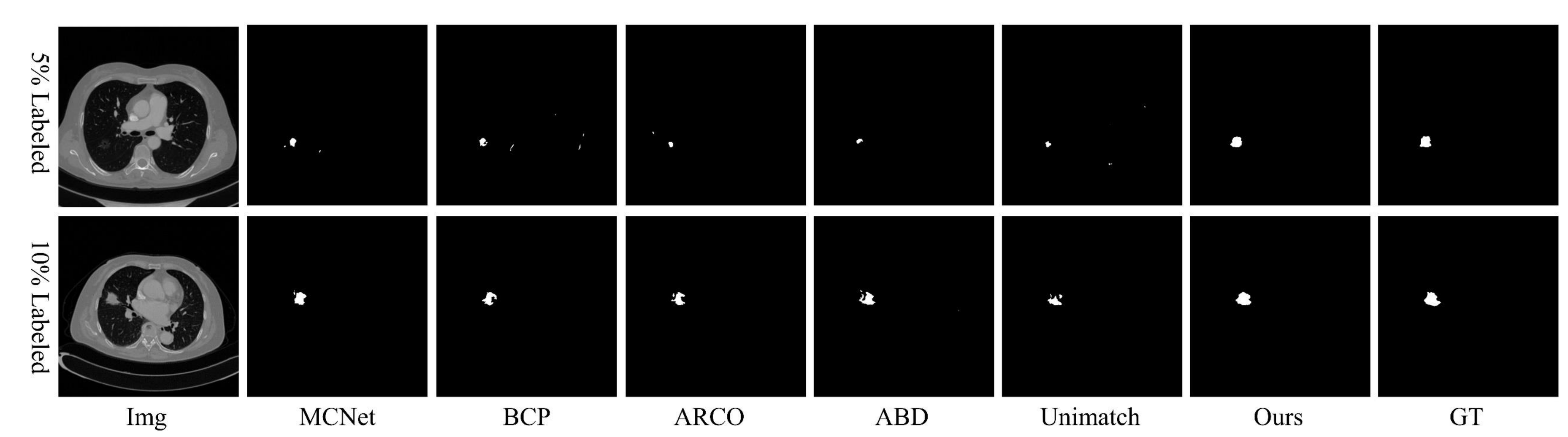
Experiments

Table 1: Quantitative comparison on the LC dataset with two labeled ratio settings (5%, 10%) using four metrics: Dice and Jaccard (↑), HD95 and ASD (↓). Best results are in **bold**, second best are underlined.

Method	5% Labeled				10% Labeled			
	Dice↑	Jaccard↑	HD95↓	ASD↓	Dice↑	Jaccard↑	HD95↓	ASD↓
UNet-F [47]	0.8345	0.7386	6.9634	2.2913	0.8345	0.7386	6.9634	2.2913
UNet-S [47]	0.4343	0.3118	26.0498	12.6188	0.6490	0.5175	21.4063	7.3382
nnUNet-F [54]	0.8259	0.7236	4.2533	1.4216	0.8259	0.7236	4.2533	1.4216
nnUNet-S [54]	0.4590	0.3438	13.2746	8.8636	0.6538	0.5194	25.2100	8.9332
UA-MT [16]	0.6029	0.4647	48.6681	24.6020	0.7222	0.5989	11.6724	5.4939
MCNet [48]	0.6378	0.4970	15.2759	4.9231	0.7555	0.6414	16.1903	9.9647
SSNet [49]	0.6328	0.4886	25.1005	9.3180	0.7480	0.6278	14.9581	7.3399
BCP [50]	0.6243	0.4854	26.9303	10.4789	0.7252	0.5994	18.9768	6.5105
ARCO [51]	0.6162	0.4778	36.2256	14.6243	0.7246	0.5945	14.4803	4.3660
ABD [52]	0.6414	0.5024	12.5608	5.9661	0.7468	0.6244	12.6570	6.7437
Unimatch [53]	0.6493	0.5071	17.8700	5.4526	0.7511	0.6333	17.0178	5.7388
Ours	**0.6643**	**0.5257**	**12.2525**	**4.2276**	**0.7852**	**0.6731**	**11.6179**	**4.2094**



T-SNE visualization of codebook evolution.



Visual results on the LC dataset.