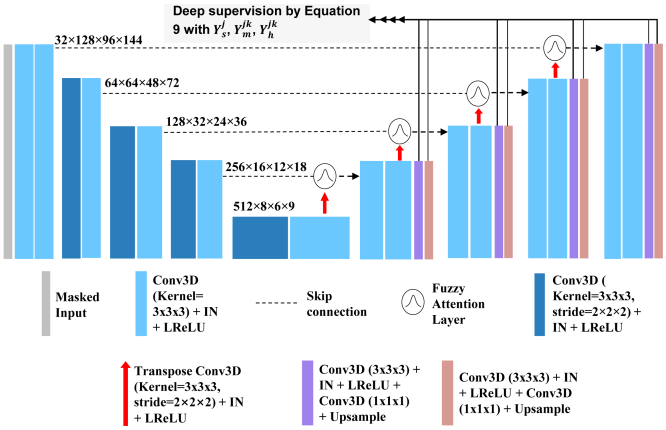


Figure 7. The elaboration of our 3D U-shaped encoder-decoder backbone. IN: Instance Normalization. Fuzzy attention layer is shown in Fig. 8

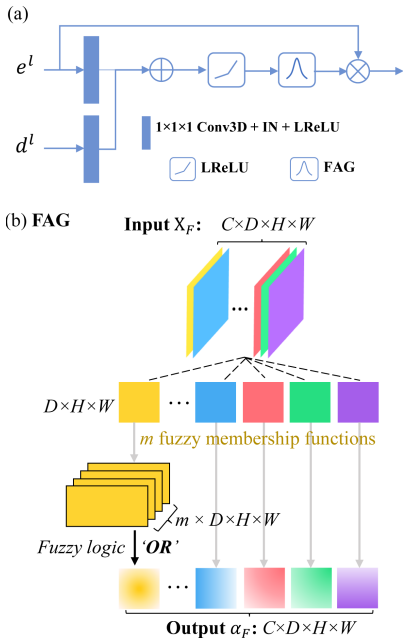
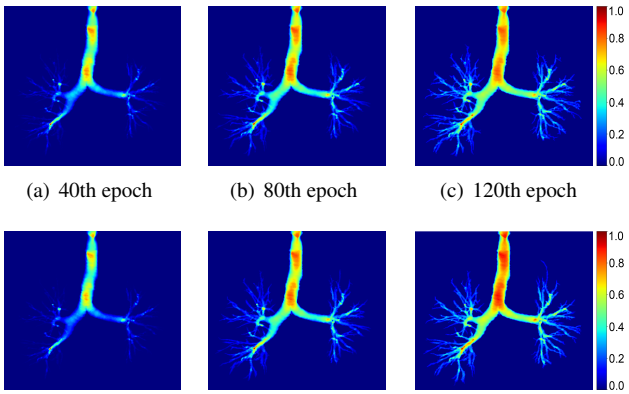


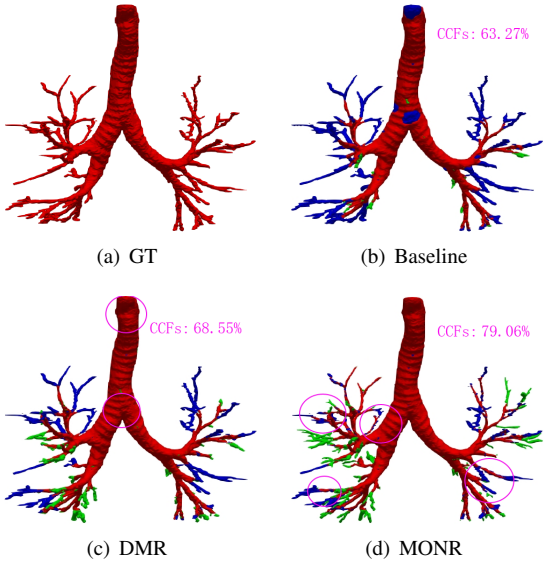
Figure 8. Details of fuzzy attention layer in the 3D U-shaped encoder-decoder backbone in Fig. 7.



(a) 40th epoch (b) 80th epoch (c) 120th epoch

(d) 40th epoch (e) 80th epoch (f) 120th epoch

Figure 9. Feature response maps during the training with different mask-ratio strategies. First row: fixed static mask-ratio training. Second row: training with the proposed DMR module.



(a) GT (b) Baseline

(c) DMR (d) MONR

Figure 10. Predicted results by sequentially accumulating each module (Red color: true positive. Green: false positive. Blue: false negative). Each segmentation result is evaluated by the comprehensive metric CCFs. Some false positives may come from the imperfect annotation in Fig. 11.

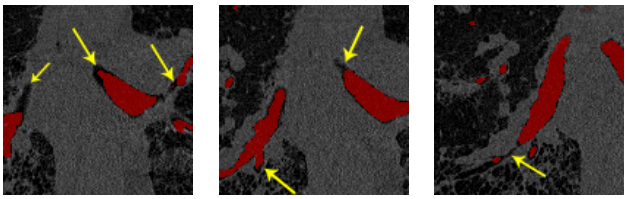


Figure 11. The minority of imperfect manual annotations (Red regions represent ground truth masks).

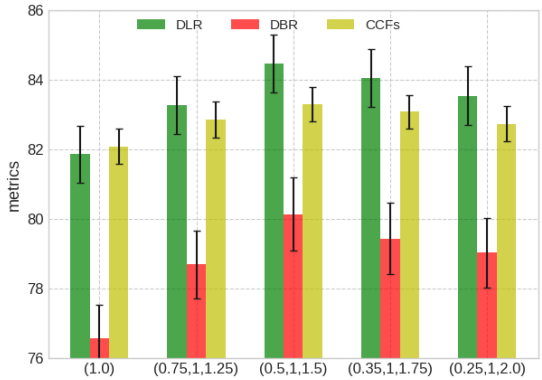


Figure 12. Ablation study of MONR module on Lung fibrosis dataset with different orders k, standard deviations are scaled by 0.12 for a better view.

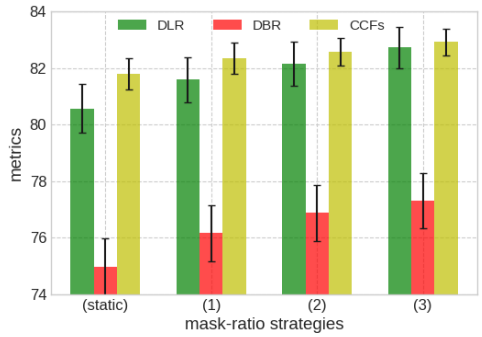


Figure 13. Ablation study of DMR module on Lung fibrosis dataset with different dynamic mask-ratio strategies in Fig. 2(a), standard deviations are scaled by 0.12 for a better view.

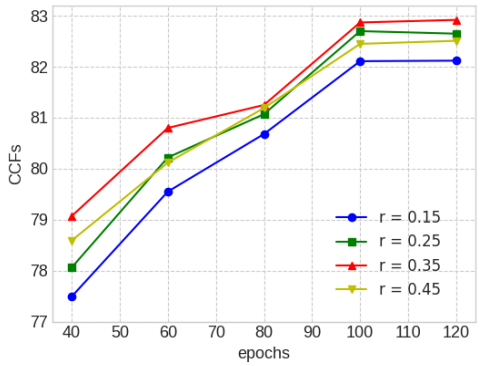


Figure 14. The ablation studies of DMR module on Lung fibrosis dataset with different maximal mask-ratios r.

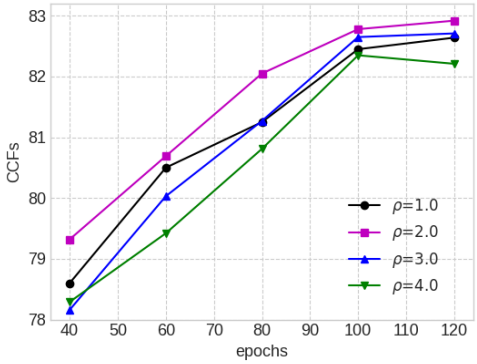


Figure 15. The ablation studies of DMR module on Lung fibrosis dataset with different powers ρ.