

## Supplementary

TABLE 1. Performance of each CNN model measured by AUC using 5-folds cross-validation.

Comparison Method	Training Method	Small Input Size	Middle Input Size	Large Input Size	Fusion
2D slice image transverse plane	scratch	0.715±0.036	0.718±0.047	0.664±0.051	0.737±0.042
	fine-tune	0.743±0.048	0.722±0.034	0.674±0.055	0.769±0.043
2D slice image multi-view plane	scratch	0.723±0.026	0.702±0.043	0.610±0.033	0.753±0.037
	fine-tune	0.798±0.030	0.820±0.021	0.824±0.016	0.846±0.021
3D volume image	scratch	0.776±0.030	0.782±0.041	0.763±0.017	0.817±0.023