

Supplementary Materials

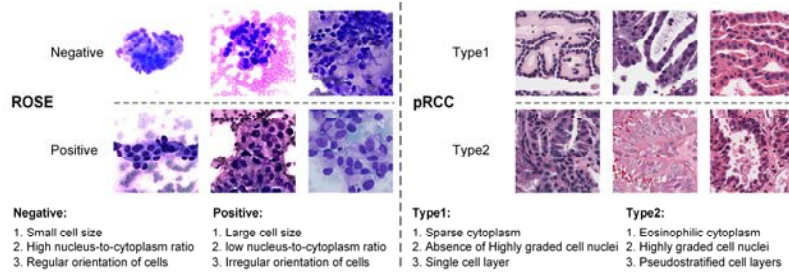


Fig. 1 Examples of two pathological datasets with their classification criteria.

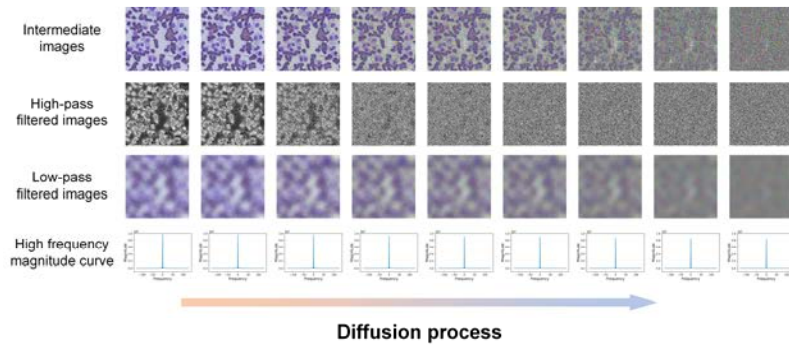


Fig. 2 Visualization of the diffusion process with FFT.

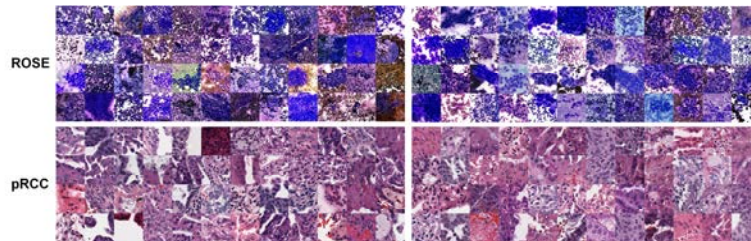


Fig. 3 Generated samples from IDDPM.

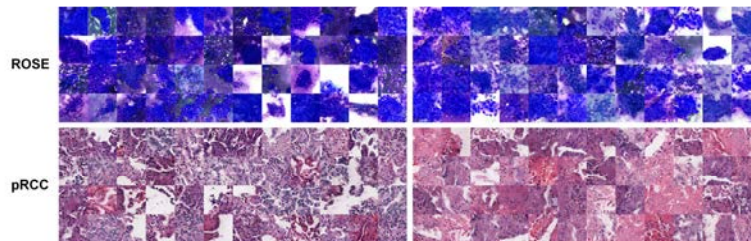


Fig. 4 Generated samples from LoFGAN.

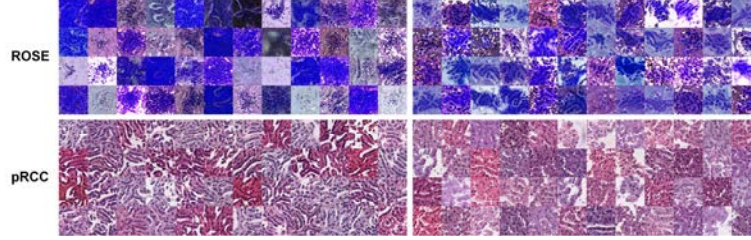


Fig. 5 Generated samples from MixDL.

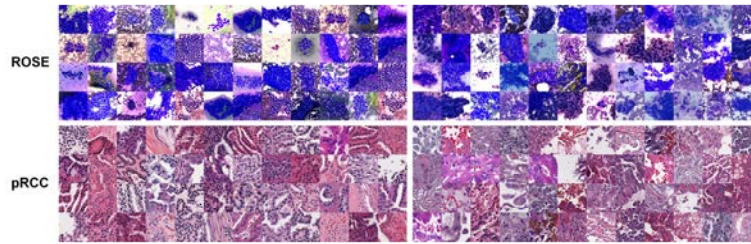


Fig. 6 Generated samples from ADD (ours).

Table 1. Comparisons in the downstream classifications using generated training set. Bold marks the highest of all results, while underlined marks the highest in generated models

Model	ROSE		pRCC	
	Accuracy (%)	F1-score (%)	Accuracy (%)	F1-score (%)
Baseline	78.13	79.09	83.50	85.36
ProGAN	64.64	67.31	71.78	72.45
IDDP	76.51	76.60	82.25	83.24
LoFGAN	74.05	73.13	81.06	81.14
MixDL	61.25	60.14	68.12	70.93
ADD (Ours)	<u>77.92</u>	<u>78.15</u>	84.04	85.49

Table 2. Comparisons in the downstream classifications using aug-trainset. Bold marks the highest of all results, while the arrow indicates the augmentation of generated models.

Model	ROSE		pRCC	
	Accuracy (%)	F1-score (%)	Accuracy (%)	F1-score (%)
Baseline	78.13	79.09	83.50	85.36
ProGAN	74.70	73.02	75.02	75.25
IDDP	79.32↑	80.81↑	83.91↑	84.89
LoFGAN	75.49	74.26	83.97↑	85.55↑
MixDL	69.66	71.83	76.28	77.26
ADD (Ours)	80.24↑	80.97↑	84.76↑	86.43↑