Supplementary Material: EndoFinder: Online Image Retrieval for Explainable Colorectal Polyp Diagnosis
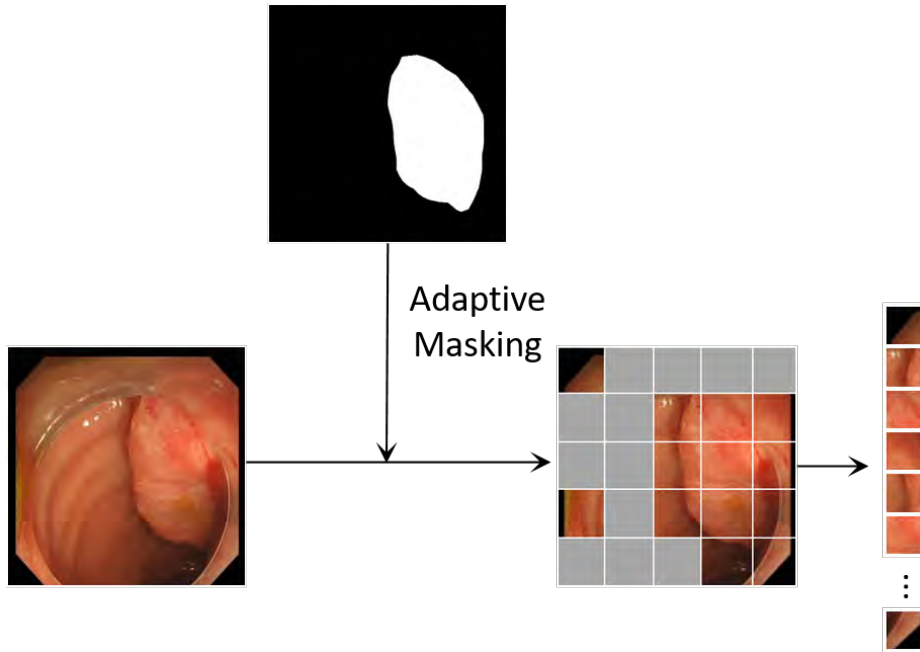

Fig. S1. Adaptive masking is performed based on the mask, and the unmasked parts are fed into the encoder.

Table S1. Implementation Details

mask_ratio	batch_size	Epochs	device	parameters
50%	32	100	RTX 3090	330M
optimizatio	LR	beta1	beta2	weight decay
AdamW	0.000125	0.9	0.95	0.05

Table S2. The loss function weight experiment showed the best re-identification performance at $\lambda = 0.5$ and $\gamma = 5$.

λ	γ	uAP	Acc@1
1	15	0.621	0.643
0.5	15	0.618	0.663
0.5	10	0.624	0.643
0.5	5	**0.695**	**0.693**
0.5	4	0.608	0.603
0.5	3	0.660	0.643
0.5	1	0.638	0.673