

Supplementary Tables and Figures

	Split	TCGA-CRC		TCGA-GASTRIC	
		ImageNet	MoCo-CRC	ImageNet	MoCo-Gastric
MeanPool	CV	0.84 (0.05)	0.87 (0.05) +0.03	0.76 (0.04)	0.82 (0.05) +0.06
MeanPool	CV centers	0.78 (0.10)	0.85 (0.07) +0.07	0.72 (0.12)	0.85 (0.07) +0.13
DeepMIL	CV	0.82 (0.05)	0.88 (0.05) +0.06	0.74 (0.01)	0.85 (0.05) +0.11
DeepMIL	CV centers	0.79 (0.059)	0.84 (0.11) +0.05	0.73 (0.15)	0.85 (0.05) +0.12
Chowder	CV	0.81 (0.05)	0.88 (0.04) +0.07	0.73 (0.07)	0.84 (0.06) +0.11
Chowder	CV centers	0.75 (0.15)	0.83 (0.12) +0.08	0.72 (0.08)	0.86 (0.05) +0.14

Table S1: Cross-validation performances (AUC) on TCGA-CRC and TCGA-Gastric. We report mean and standard deviation on 3×5 folds. We split the data into 5 fold either randomly (CV) or by ensuring that all samples from a center are in the same set (CV centers).

Feature extractor	Train dataset	MeanPool	Chowder	DeepMIL
ImageNet	CRC	$C = 1$	$R = 100$ 100 epochs	$N = 64$ 30 epochs
ImageNet	Gastric	$C = 0.0$	$R = 25$ 30 epochs	$N = 64$ 20 epochs
MoCo-CRC	CRC	$C = 0.5$	$R = 10$ 10 epochs	$N = 32$ 10 epochs
MoCo-Gastric	Gastric	$C = 1.0$	$R = 100$ 30 epochs	$N = 128$ 10 epochs
MoCo-CRC-Gastric	CRC	$C = 0.5$	$R = 100$ 30 epochs	$N = 64$ 10 epochs

Table S2: Hyperparameters used for the different models. All hyperparameters were tuned on the training sets of the TCGA-CRC-Kather or TCGA-Gastric-Kather cohorts. C refers to L2 penalization coefficient, R to the number of extreme tiles used in Chowder and N to the size of the attention layer in DeepMIL.

Methods	MoCo-CRC-Gastric
MeanPool	0.86 (0.04)
Chowder	0.88 (0.05)
DeepMIL	0.88 (0.06)

Table S3: Cross-validation performances (AUC) on TCGA-CRC, for the models trained using the features of MoCo-CRC-Gastric. The result reported are the mean and the standard deviation from 5-fold cross-validation repeated 3 times.

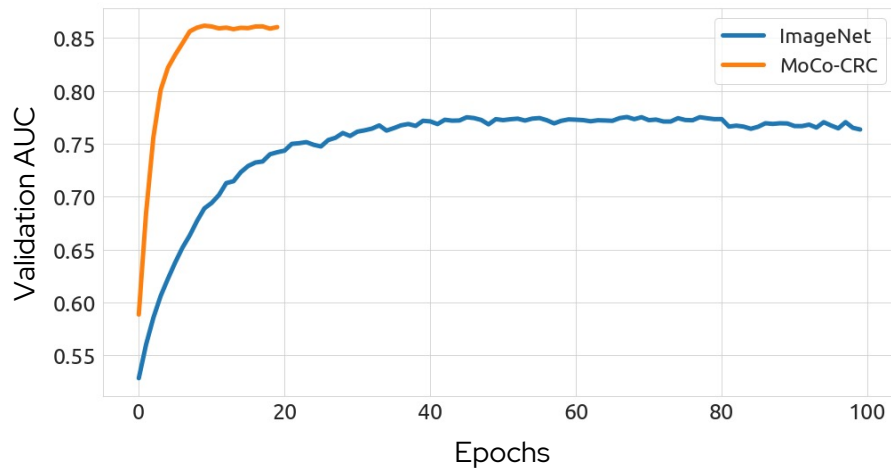


Figure S1: Validation Curves of Chowder with ImageNet and MoCo-CRC features. The plot represents the average validation AUCs per epoch of all runs of the cross-validation for Chowder trained with ImageNet and MoCo-CRC features, depicted in Table 3.