GPC: Generative and General Pathology Image Classier

Anh Tien Nguyen and Jin Tae Kwak
School of Electrical Engineering, Korea University, Seoul, Republic of Korea

INTRODUCTION

- In computational pathology, pathological image classification has been extensively studied, in which numerous tasks are closely relevant. For example, cancer grading in different types of organs: prostate, colon, gastric, and breast.
- Current approaches require a separate model per task, which is challenging to transfer a pre-existing model to other related tasks
- We introduce a task-agnostic **Generative and general Pathology image Classifier** ***(GPC)*** that learn from arbitrary datasets of pathology images and perform multiple image classification tasks in a generative image-to-text fashion

METHOD

Problem formulation

- Given M datasets $\{D_1, D_2, ..., D_M\}$, with each dataset is $D_i = \{(x^{i,k}, c^{i,k})\}_{k=1}^{N}$ where $x^{i,k}$ and $c^{i,k}$ denote the k-*th* pathology image and its ground truth in the i-*th* dataset.
- Since $c^{i,k}$ is a text label, it is spited and padded into a sequence of tokens $t^{i,k}$ with a length T, using a tokenizer of a language model $\mathcal{L}$.
- Each pathology image $x^{i,k}$ undergoes a feature extractor $\mathcal{F}$ and a projector $\mathcal{P}$ to produce a feature embedding $f^{i,k}$

$$f^{i,k} = \mathcal{P}(\mathcal{F}(x^{i,k}))$$

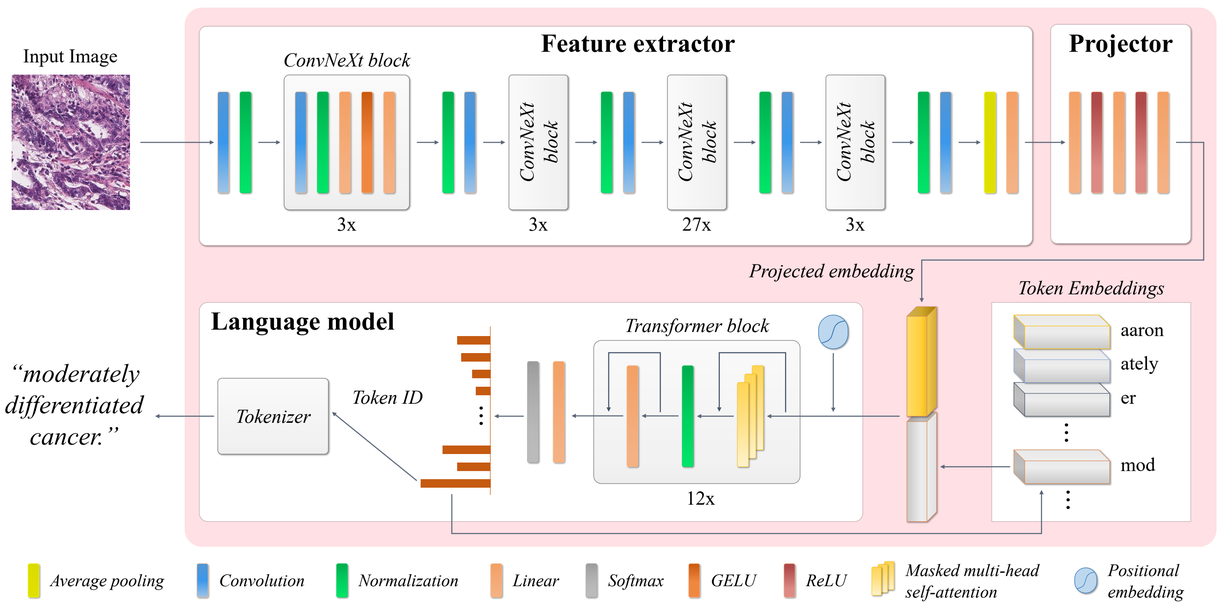
- Afterward, the projected embedding $f^{i,k}$ is used as a condition for the language model $\mathcal{L}$ to generate the m-*th* token $\tilde{t}_m^{i,k}$ autoregressively

$$\tilde{t}_m^{i,k} = \arg\max_{\hat{t}_m^{i,k}} p(\hat{t}_m^{i,k} \mid f^{i,k}, \tilde{t}_1^{i,k}, \tilde{t}_2^{i,k}, \dots, \tilde{t}_{m-1}^{i,k})$$

- As a result, the objective to optimize the parameter θ of GPC is

$$\theta = \underset{\hat{\theta}}{argmax} \sum_{i=1}^{M} \sum_{k=1}^{N} \sum_{m=1}^{T} log \; p_{\hat{\theta}}(\tilde{t}_m^{i,k} \mid f^{i,k}, \tilde{t}_1^{i,k}, \tilde{t}_2^{i,k}, \dots, \tilde{t}_{m-1}^{i,k})$$

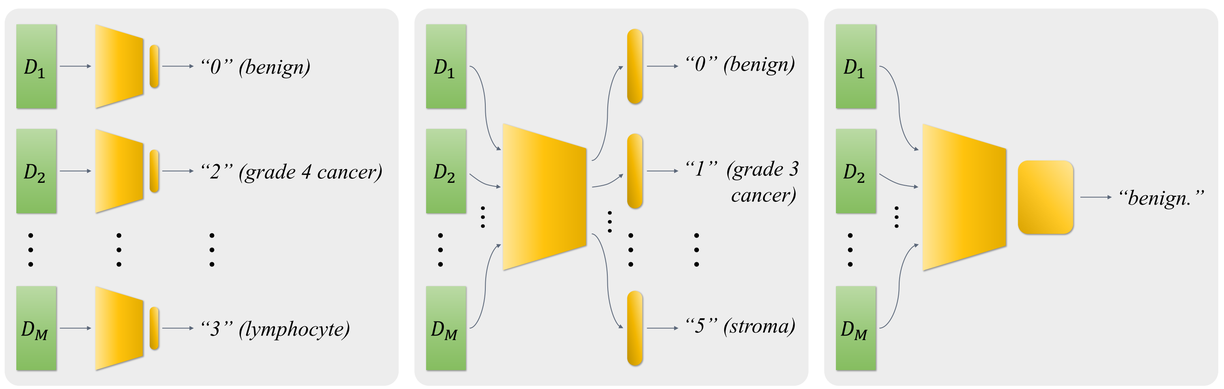
Network architecture



- Feature extractor $\mathcal{F}$ is built upon ConvNeXt to extract high-level representations from input images.
- Projector $\mathcal{P}$ maps the output features from the space of $\mathcal{F}$ to that of a language model $\mathcal{L}$, constructed from a three-layer perceptron.
- Language model $\mathcal{L}$, based on OPT, generates the correct pathological text labels for the projected embeddings obtained from $\mathcal{P}$.

EXPERIMENT

Experimental design



Task-specific model (E_{TS}) | Task-agnostic model (E_{TA}) | Task-agnostic generative model (E_{TAG})

Datasets

Task	Dataset	Mag.	Patch size	# patches	Text labels
Colorectal cancer grading	Colon-1	20x	512 x 512	9,857	*benign, well differentiated cancer, moderately differentiated cancer, poorly differentiated cancer*
	Colon-2	20x	512 x 512	110,170	
Prostate cancer grading	Prostate-1	40x	750 x 750	22,022	*benign, grade 3 cancer, grade 4 cancer, grade 5 cancer*
	Prostate-2	40x	690 x 690	17,066	
Gastric cancer grading	Gastric	40x	512 x 512	265,066	*benign, tubular well differentiated cancer, tubular moderately differentiated cancer, tubular poorly differentiated cancer*
Colorectal tissue typing	K19	20x	224 x 224	100,000	*adipose, background, debris, normal, lymphocyte, stroma, epithelium, muscle, mucus*

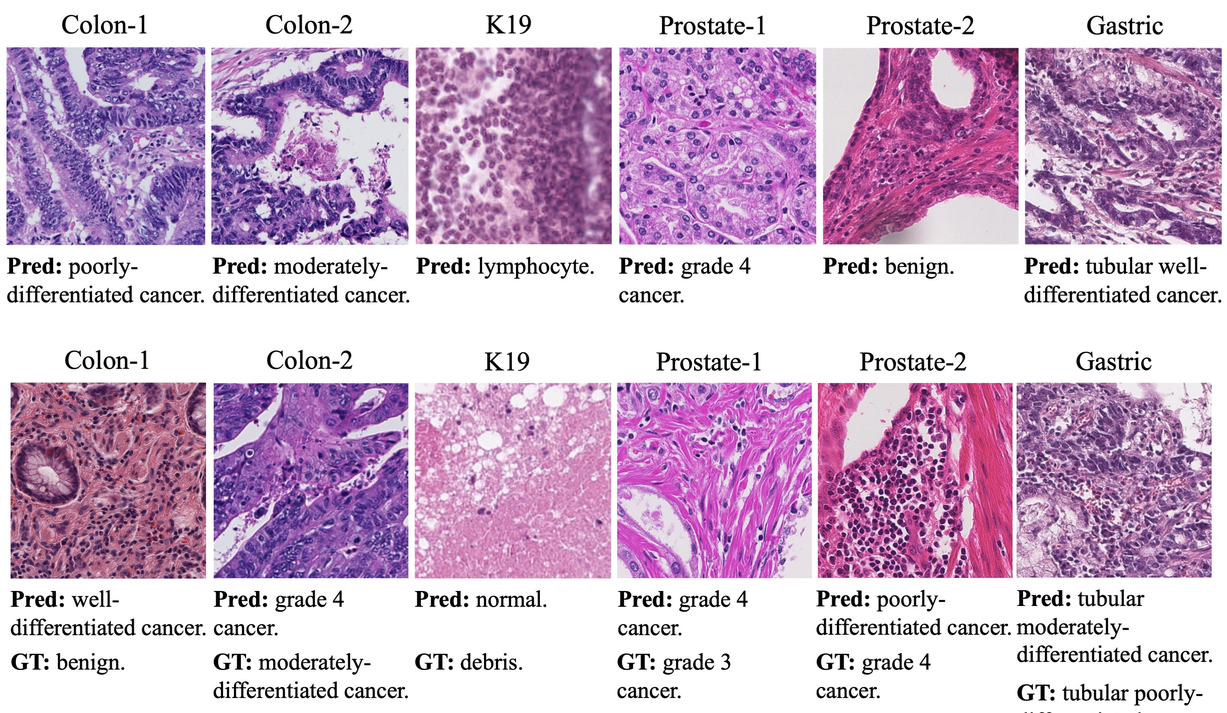
RESULT

- Results of colorectal cancer grading and tissue typing

Model	Type	Colon-1				Colon-2				K19			
		Acc	Acc_g	*F1*	κ_w	*Acc*	Acc_g	*F1*	κ_w	*Acc*	*Pre*	*Re*	*F1*
ConvNext	E_{TS}	87.7	82.8	0.832	0.940	78.1	71.9	0.731	**0.908**	**99.6**	**0.996**	**0.996**	0.994
EfficientN		85.9	80.9	0.819	0.914	76.9	68.4	0.708	0.701	98.0	0.973	0.968	0.985
Resnet50		86.8	82.9	0.838	0.806	79.5	68.2	**0.733**	0.688	98.7	0.988	0.988	0.987
MaxViT		87.9	**84.0**	0.838	0.805	76.3	72.8	0.723	0.895	98.3	0.988	0.991	0.988
SwinV2		88.0	82.7	0.829	0.839	77.9	73.7	0.729	0.885	99.4	0.996	0.993	0.991
ViT		87.5	82.0	0.838	0.838	**79.8**	72.8	0.728	0.899	98.2	0.989	**0.996**	0.988
ConvNext	E_{TA}	85.9	80.4	0.823	0.933	74.4	66.5	0.698	0.868	98.8	0.986	0.991	0.988
EfficientN		83.2	79.2	0.793	0.882	72.5	63.4	0.670	0.722	98.5	0.982	0.982	0.974
Resnet50		84.1	81.0	0.807	0.824	70.1	61.9	0.622	0.671	97.7	0.984	0.983	0.986
MaxViT		86.8	82.6	0.809	0.813	71.3	68.8	0.720	0.888	98.3	0.985	0.991	0.974
SwinV2		86.5	81.2	0.822	0.933	70.4	69.0	0.671	0.842	98.4	0.985	0.980	**0.996**
ViT		86.0	79.7	0.812	0.831	72.1	67.1	0.701	0.833	98.1	0.985	0.989	0.988
GIT	E_{TAG}	85.3	79.7	0.811	0.924	67.9	58.6	0.596	0.839	98.9	0.989	0.988	0.990
CLIP,OPT		82.5	75.6	0.795	0.914	72.7	67.4	0.653	0.791	99.0	0.989	0.992	0.985
GPC		**88.4**	83.8	**0.848**	**0.944**	79.0	**74.0**	0.722	0.898	99.4	0.995	0.995	**0.996**

- Results of prostate and gastric cancer grading

Model	Type	Prostate-1				Prostate-2				Gastric			
		Acc	Acc_g	*F1*	κ_w	*Acc*	Acc_g	*F1*	κ_w	*Acc*	Acc_g	*F1*	κ_w
ConvNext	E_{TS}	70.6	70.1	0.630	0.597	77.8	78.2	0.639	0.696	83.8	68.1	0.760	0.925
EfficientN		69.7	66.4	0.582	0.504	74.3	**77.3**	0.599	0.633	81.3	68.1	0.712	0.890
Resnet50		70.9	67.5	0.643	0.512	**77.3**	78.7	0.608	0.619	82.2	66.9	0.707	0.901
MaxViT		71.6	70.2	**0.652**	**0.649**	75.9	76.7	0.605	0.678	83.2	68.5	0.758	0.926
SwinV2		**71.9**	72.0	0.637	0.639	73.9	75.1	0.623	0.669	83.9	68.5	0.771	**0.935**
ViT		**71.9**	**72.2**	0.641	0.643	75.4	75.9	0.608	0.690	**84.4**	69.2	**0.774**	0.930
ConvNext	E_{TA}	68.5	69.7	0.576	0.578	73.3	76.3	0.562	0.616	83.0	67.2	0.757	0.930
EfficientN		65.2	62.1	0.522	0.511	71.9	73.1	0.512	0.589	80.5	63.5	0.701	0.832
Resnet50		69.2	68.1	0.582	0.539	73.6	77.3	0.599	0.601	82.9	64.0	0.713	0.890
MaxViT		67.2	69.2	0.606	0.562	69.2	70.9	0.525	0.631	83.6	65.9	0.749	0.931
SwinV2		65.5	66.9	0.531	0.542	65.8	69.2	0.487	0.553	81.7	65.0	0.739	0.923
ViT		67.2	66.4	0.544	0.579	68.8	72.8	0.598	0.629	81.7	64.2	0.710	0.909
GIT	E_{TAG}	65.9	67.2	0.538	0.476	68.3	71.7	0.467	0.616	80.7	63.7	0.727	0.867
CLIP,OPT		62.0	63.3	0.598	0.587	63.4	62.2	0.521	0.575	81.6	63.4	0.726	0.912
GPC		70.4	71.9	0.628	0.612	76.9	**79.0**	**0.641**	**0.700**	83.7	**69.3**	0.768	0.925



CONCLUSION

The generative models hold great potential for pathology image analysis, paving the way for developing a universal model for multiple tasks of computational pathology.

Quantitative Imaging and Informatics Laboratory

This work was supported by the grant of the National Research Foundation of Korea (NRF) (No. 2021R1A2C2014557)