Order-ViT: Order Learning Vision Transformer for Cancer Classification in Pathology Images

Ju Cheon Lee, Jin Tae Kwak
School of Electrical Engineering, Korea University

INTRODUCTION

- Cancer grading plays an important role in the treatment and management of patients
- It has been primarily studied as a categorical classification problem where the relationship among cancer grades is, by and large, ignored
- We propose an order learning vision transformer (Order-ViT) for cancer grading in pathology images that can learn both the histopathological patterns of individual cancer grades and the ordering relationship among cancer grades

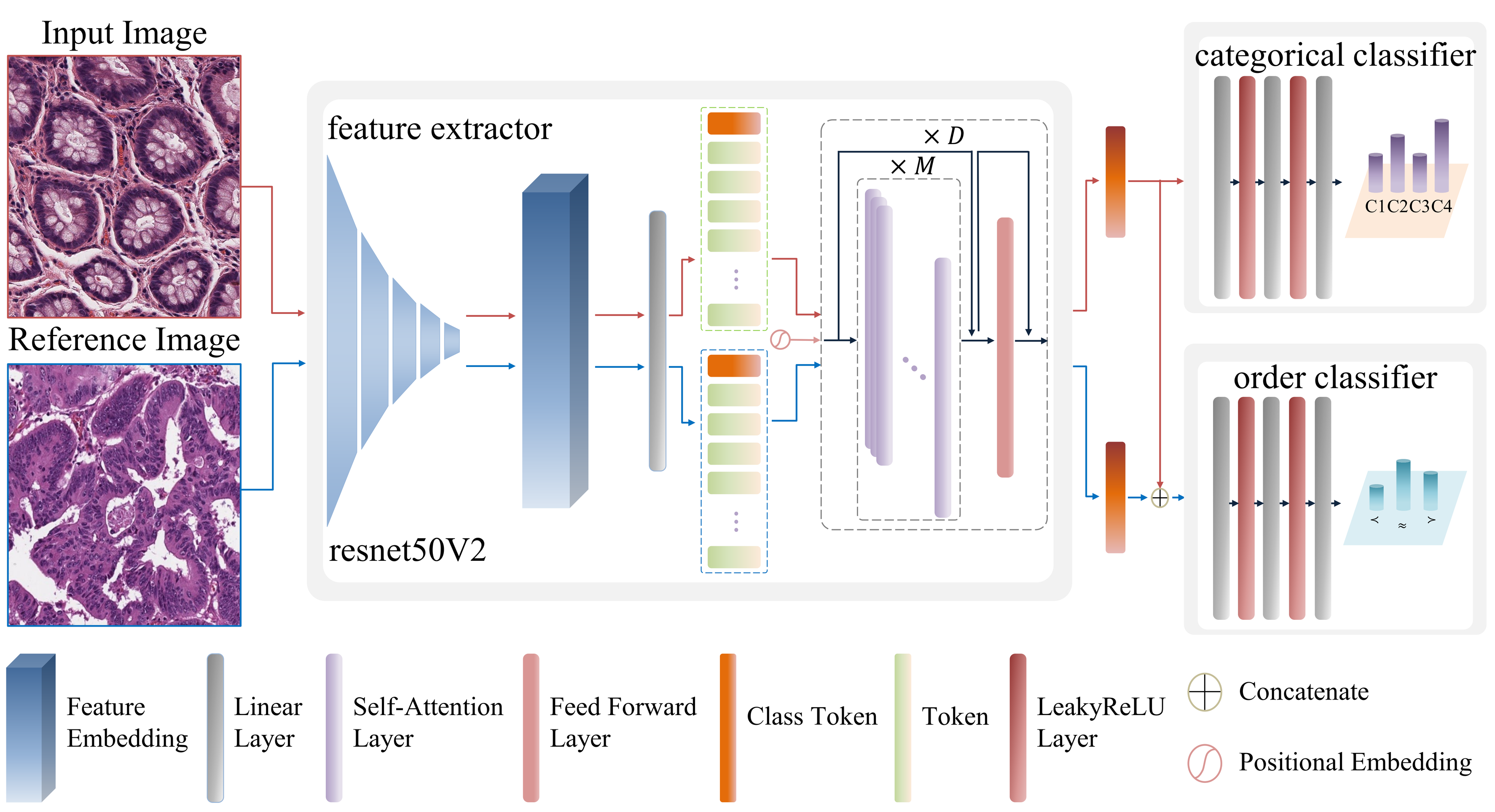
METHODS

Dataset

- Public colorectal cancer datasets [1] and a gastric cancer dataset [2] are employed in this study
- Colorectal cancer; Benign (BN), well-differentiated (WD), moderate-differentiated (MD), and poorly-differentiated (PD) tumors
- Gastric cancer; Benign (BN), tubular well-differentiated (WD), tubular moderate-differentiated (MD), and tubular poorly-differentiated (PD) tumors

Tissue Type	Class	Training	Validation	Test I	Test II
Colorectal Tissue	BN	773	374	453	27,896
	WD	1,866	264	192	8,394
	MD	2,997	370	738	61,985
	PD	1,391	234	205	11,896
Gastric Tissue	BN	20,883	8,398	7,955	-
	TW	14,251	2,239	1,795	-
	TM	20,815	2,370	2,458	-
	TP	27,689	2,374	3,579	-

Model Architecture

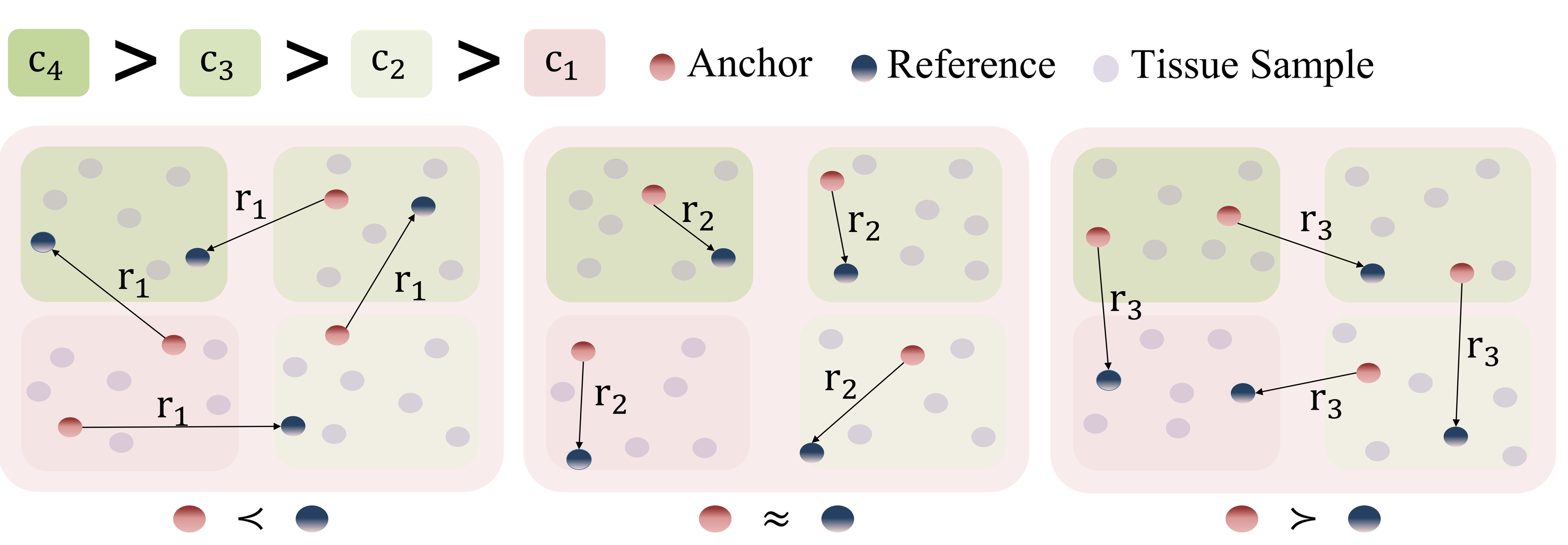


- Input image undergoes ResNet50V2 and ViT for feature representation
- Categorical classifier predicts the class label of input images and order classifier predicts the relationship between two images

Categorical and Order Classifier

- Linear - LeakyReLU - Linear - LeakyReLU – Linear
- Categorical classifier receives CLS token from the input image
- Order classifier receives CLS tokens from the input and reference images

Order Classifier



Order relationship between two samples x_i and x_j

$$x_j \prec x_j \quad \text{if} \quad y_j - y_j < 0$$
$$x_j \approx x_j \quad \text{if} \quad y_j - y_j = 0$$
$$x_j \succ x_j \quad \text{if} \quad y_j - y_j > 0$$

Order label for two samples x_i and x_j

$$z_{i,j} = \begin{cases} r_1 & \text{if} \quad y_i < y_j \\ r_2 & \text{if} \quad y_i = y_j \\ r_3 & \text{if} \quad y_i > y_j \end{cases}$$

TWO-STAGE INFERENCE

- Given an input image, categorical classifier predicts the class label and computes confidence score u
- If $u < \tau$, then we invoke adaptive voting for final prediction

Confidence Score

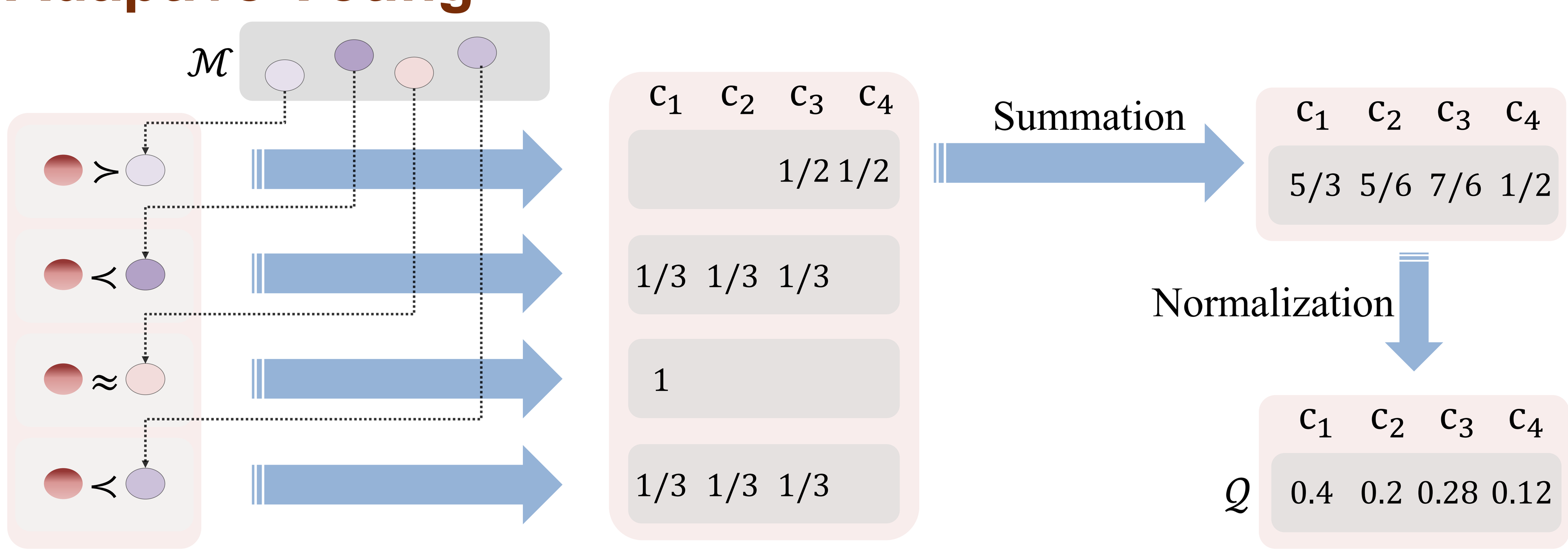
- Confidence score is computed as the difference between the highest probability ($\mathcal{P}_{1st}$) and the second highest probability ($\mathcal{P}_{2nd}$)

$$u = \mathcal{P}_{1st} - \mathcal{P}_{2nd}$$

Memory Bank($\mathcal{M}$)

- $\mathcal{M}$ serves as a repository for reference images to be compared with the uncertain input images

Adaptive Voting



- An uncertain input image is compared against every reference image in $\mathcal{M}$ using order classifier to vote for the final prediction:
  - If order classification result is $\succ$, add $\frac{1}{N_c - c}$ for label higher than c
  - If order classification result is $\approx$, add 1 to c
  - If order classification result is $\prec$, add $\frac{1}{N_c - c}$ for label lower than c

  N_c is number of classes and c is reference image class in $\mathcal{M}$
- Final prediction: $\text{argmax}(\mathcal{P} \times W_1 + \mathcal{Q} \times W_2)$

$$W_1 = \frac{\text{entropy}(\mathcal{P})}{\text{entropy}(\mathcal{Q}) + \text{entropy}(\mathcal{P})};\ W_2 = \frac{\text{entropy}(\mathcal{Q})}{\text{entropy}(\mathcal{Q}) + \text{entropy}(\mathcal{P})}$$

RESULTS

- Result of Colorectal and gastric cancer classification

Model	Colon Test I Acc(%)	Colon Test I F1	Colon Test I K_w	Colon Test II Acc(%)	Colon Test II F1	Colon Test II K_w	Gastric Test I Acc(%)	Gastric Test I F1	Gastric Test I K_w
ResNet 50	86.65	0.825	0.935	68.24	0.864	0.818	84.04	0.775	0.926
DenseNet 121	85.77	0.820	0.930	70.03	0.671	0.843	83.44	0.772	0.928
EfficientNet B0	86.71	0.821	0.926	64.60	0.627	0.794	82.81	0.766	0.919
MSBP-Net	86.21	0.824	0.933	74.67	0.708	0.860	84.53	0.770	0.928
M_{MSE-CE_o}	87.28	**0.838**	0.940	75.95	0.710	0.846	84.34	0.771	0.925
M_{MAE-CE_o}	86.65	0.826	0.937	75.44	0.702	0.861	84.10	0.768	0.922
ViT	86.27	0.829	0.931	77.54	0.712	0.874	84.06	0.772	**0.931**
Swin	85.26	0.820	0.931	77.10	0.721	0.868	83.71	0.759	0.919
DeiT III	76.76	0.673	0.794	48.42	0.396	0.271	77.05	0.656	0.847
Order-ViT	**87.66**	0.834	**0.942**	**83.21**	**0.740**	**0.899**	**84.89**	**0.783**	0.930

- Ablation study for colorectal and gastric cancer classification

Cat	Ord	Vot	Vot_w	Colon Test I Acc(%)	Colon Test I F1	Colon Test I K_w	Colon Test II Acc(%)	Colon Test II F1	Colon Test II K_w	Gastric Test I Acc(%)	Gastric Test I F1	Gastric Test I K_w
v				86.27	0.829	0.931	77.54	0.712	0.874	84.06	0.772	**0.931**
v	v			87.41	0.829	0.941	82.77	0.737	0.897	84.58	0.776	0.929
v	v	v		87.47	0.831	0.941	82.98	0.739	0.898	84.61	0.779	0.929
v	v	v	v	**87.66**	**0.834**	**0.942**	**83.21**	**0.740**	**0.899**	**84.89**	**0.793**	0.930

CONCLUSIONS

- Order-ViT demonstrates superior cancer classification performance
- Order-ViT utilizes the individual histological patterns of pathology images by categorical classification and the relationship among different pathological images by order learning

1. Le Vuong, Trinh Thi, et al. "Joint categorical and ordinal learning for cancer grading in pathology images." Medical image analysis 73 (2021): 102206.
2. ABBAS, Syed Farhan, et al. Multi-cell type and multi-level graph aggregation network for cancer grading in pathology images. *Medical Image Analysis*, 2023, 102936.

Quantitative Imaging and Informatics Laboratory

This work was supported by the grant of the National Research Foundation of Korea (NRF) (No. 2021R1A2C2014557) and Institute of Information & communications Technology Planning & evaluation (IITP) (No. RS-2022-00167143), funded by the Korea government (MSIT)