TransRP: Transformer-based PET/CT feature extraction incorporating clinical data for recurrence-free survival prediction in oropharyngeal cancer

Baoqiang Ma [1], Jiapan Guo [1 2], Lisanne V. van Dijk [1], Peter M.A. van Ooijen [1 2], Stefan Both [1], Nanna M. Sijtsema [1]
Department of Radiation Oncology [1], Data Science Centre in Health (DASH) [2]
University of Groningen, University Medical Centre Groningen

MIDL Nashville 2023

Introduction

Oropharyngeal cancer (OPC) patients are treated with chemo(radiotherapy) with/without surgery. There are large treatment outcome variations between patients. Outcome prediction models can select patients for personalized treatment. Although convolutional neural networks (CNN) based models showed promising predictive performance in OPC, they are limited in learning global-context features. In contrast, transformer is superior in learning global features and combing multi-modality features while it is data-hungry.

Purpose

1.Propose a CNN-transformer combined model for OPC recurrence-free survival prediction based on both clinical and imaging data. 2.Test whether transformer is better in combing multi modality features.

Materials & Methods

HECKTOR 2022: 489 OPC patients with planning CT, PET images and Gross Tumor Volume (GTV) in seven centers.

362 for 5-fold cross-validation; 120 for test.

TransRP model

A CNN (ResNet18 /DenseNet121) to extract detailed features.

A Transformer (ViT-base) to extract global context features.

Incorporating clinical data

Using Transformer in Figure 2(b) (c) ; m1, m2.

Using a fully connected layer (FCN) in Figure 2(d) ; m3.

Experiment setting:

SGD optimizer with a momentum 0.90; learning rate 0.0002; early stopping; data augmentation.

Evaluation metric: C-index.

Figure 1. Example of CT, PET and GTV.

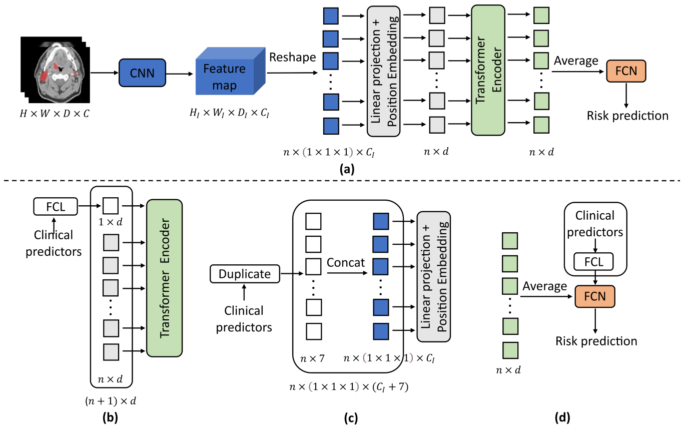

Figure 2. TransRP model and three methods of incorporating clinical data.

Table 1. The C-index in the validation and test sets.

Model	Validation sets	Test
DeepSurv [1]	0.614 ± 0.065	0.601
ResNet18-m3	0.680 ± 0.070	0.650
DenseNet121-m3 [2]	0.679 ± 0.071	0.676
TransRP(ResNet18)-m1	0.658 ± 0.047	0.624
TransRP(ResNet18)-m2	0.695 ± 0.081	0.644
TransRP(ResNet18)-m3	0.674 ± 0.069	0.686
TransRP(DenseNet121)-m1	0.693 ± 0.055	0.693
TransRP(DenseNet121)-m2	0.694 ± 0.068	0.671
TransRP(DenseNet121)-m3	0.686 ± 0.063	**0.698**

Table 2. Ablation studies

Clinical	CNN	Transformer	Model	Test
✓			DeepSurv	0.601
	✓	✓	TransRP(DenseNet121)	0.686
✓		✓	ViT-m3	0.656
✓		✓	DenseNet121-m3	0.676
✓	✓	✓	TransRP(DenseNet121)-m3	**0.698**

Results

- TransRP(DenseNet121)-m3 achieved highest C-index (0.698) in the test set
- TransRP-m3 beat corresponding CNNs-m3
- A fully connected layer (m3) is better than transformer (m1, m2) in combing multi-modality features

Ablation

Removing any parts caused a C-index decrease

TransRP(DenseNet121) beats DenseNet121-m3

Conclusion

TransRP models performed better than CNNs for OPC outcome prediction. A fully connected layer is better than transformer in incorporating multi-modality features

References

1. Katzman et al., 2018 2. Wahid et al., 2021 (b). Saeed et al., 2022 (c). Zheng et al., 2022

Acknowledgments

China Scholarship Council
DASH Data Science Center in Health
rijksuniversiteit groningen