EBioMedicine 73 (2021) 103671

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Commentary

Pathological diagnosis and prognosis of Gastric cancer through a multi-instance learning method

Runnan Cao[a,b,1], Lixin Gong[c,1], Di Dong[a,b,*]

[a] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[b] CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[c] College of Medicine and Biological Information Engineering School, Northeastern University, Shenyang 110016, China

ARTICLE INFO

Article History:
Received 20 October 2021
Accepted 20 October 2021

Gastric cancer (GC) is one of the most common malignant tumors with poor prognostic results. Endoscopic examination is mainly utilized for early detection, while pathological confirmation and CT scanning are suggested for further treatment. Machine learning and deep learning methods have been widely applied to explore automatic GC diagnosis and prognosis analysis. Endoscopic images based deep learning models for early GC detecting have achieved a performance equal with experienced endoscopists [1] and radiomic methods on CT images have proposed clinically significant imaging biomarkers with diagnostic and prognostic values [2–4]. In this article of Ebiomedicine, Huang et al. [5] proposed a simple multi-instance learning (MIL) approach for GC diagnosis and prognosis analysis on whole slide imaging (WSI) pathological images. Experiments on three datasets achieved good performances for both tasks.

The researchers collected a total of 2508 pathological images from 1128 patients and cropped the images into small tiles (224 × 224 pixels). Image-level diagnosis by pathologists was treated as labels for tiles. RegNetY, an improved model from ResNet, extracted features from each tile, followed by aggregation models to fuse features from the most significant tiles. They adopted a recurrent neural network (RNN) to merge features from different tiles for the diagnostic model, achieving an accuracy of 0.976 and 0.920 in the internal and external validation datasets, respectively. Multi-layer perception (MLP) was trained and validated for prognosis with a *C*-index of 0.671 and 0.657. Besides, the predicted risk was a strong predictor for survival prediction.

The traditional diagnosis of pathological images, relying on pathologists' visual observation, is time-consuming and labor-intensive. Multiple explorations for GC prediction on pathological images have emerged nowadays. In practice, the WSI image is cropped into small tiles and trained via a weakly supervised training method to deal with its large size. During training, the challenges are as follows: extracting features from tiles and merging these tile-level features to make image-level predictions. For GC detection [6,7], researchers applied ResNet for feature extraction at tile-level and semantic segmentation models to detect GC regions at image-level, which requires delineating the margins of the tumor region on WSIs by pathologists. Muti et al. [8] used ShuffleNet for tile-level prediction and the fraction of predicted positive tiles as the patient-level prediction score for GC classification. Besides, Wang et al. [9] developed a framework including segmentation, classification, and risk quantification to predict the prognosis of GC using lymph node pathological images. Compared with previous related works, Huang et al. adopted a relatively simple MIL strategy with RNN and MLP but demonstrated powerful performance in GC diagnosis and prognosis analysis.

Determining the best practice of MIL in pathological images is still an open question. There are semantic segmentation models such as DeepLab to detect tumor regions, the fraction of predicted positive tiles for molecular subtyping, merge top-*k* features using sequential models or other simple models, and advanced MIL approaches [10]. Choosing the MIL strategy for pathological images is task-determinant, requiring previous experience and explorative experiments. Huang et al. used RNN for diagnosis but MLP for prognosis analysis. As explained by the researchers, a sequential model is more applicable to eliminate the effect of outliers, while MLP could be more efficient in the prediction based on all tiles. Interestingly, when researchers tried RNN algorithms in prognosis, the model could hardly converge, resulting in poor performance with a nearly 0.5 *C*-index. These results are consistent with the intuitive idea that the general observation on most tiles could predict diagnosis, but prognosis might require considering extreme tiles. Moreover, the results showed the potential of simple models to achieve good performance, guiding us to reconsider the model design for MIL practices in WSI images.

The research of Huang et al. provided a new solution for diagnosis and prognosis analysis of GC based on pathological images, and also raised several open questions for future research. First, the number *k*

* Corresponding author at: CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, No.95 Zhongguancun East road, Beijing 100190, China.
E-mail addresses: caorunnan2019@ia.ac.cn (R. Cao), 1810454@stu.neu.edu.cn (L. Gong), di.dong@ia.ac.cn (D. Dong).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ebiom.2021.103671

of selected top-*k* significant tiles could largely influence model performance. However, it lacked a determined method for choosing this value. There is a trade-off between sufficient coverage and superfluous insignificant tiles. Generally, it relates to the tumor type, WSI magnification, tile size, fusion model, and specific tasks, thus no one-fit-all solution exists. Comparative experiments and discussions for different parameter settings might provide a deeper insight into the problem. Also, it could be treated as a tuning parameter to be selected by cross-validation. Second, RNN for diagnosis and MLP for prognosis analysis were applied in the research. Nevertheless, its power in other datasets and tasks is open for exploration. What about the general ability of the approach to other datasets and tasks? Can it outperform advanced MIL techniques? What are the similarities or differences between the tasks that facilitate or hinder the utilization? These explorations might guide rethinking the complexities and generality of the model in the MIL scenario, and may also provide more insights into tumor diagnosis and prognosis.

Contributors

All authors contributed to conceptualization, writing, reviewing, editing and have read and agreed to the published version of the manuscript.

Declaration of Competing Interest

The authors declare no conflict of interest.

Role of the Funding Source

This commentary was supported by National Natural Science Foundation of China (82022036, 91959130, 81971776, 81771924, 91959205) and Beijing Natural Science Foundation (L182061, Z20J00105).

References

[1] Hu H, Gong L, Dong D, et al. Identifying early Gastric cancer under magnifying narrow-band images with deep learning: a multicenter study. Gastrointest Endosc 2021;93(6):1333–41 e3.
[2] Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced Gastric cancer: an international multicenter study. Ann Oncol 2020;31(7):912–20.
[3] Dong D, Tang L, Li ZY, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced Gastric cancer. Ann Oncol 2019;30(3):431–8.
[4] Zhang L, Dong D, Zhong L, et al. Multi-focus network to decode imaging phenotype for overall survival prediction of Gastric cancer patients. IEEE J Biomed Health Inf 2021;25(10):3933–42.
[5] Huang B, Tian S, Zhan N, et al. Accurate diagnosis and prognosis prediction of Gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. EBioMedicine 2021;71:103631. doi: 10.1016/j.ebiom.2021.103631.
[6] Hu Y, Su F, Dong K, et al. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. Gastric Cancer 2021;24:868–77.
[7] Song Z, Zou S, Zhou W, et al. Clinically applicable histopathological diagnosis system for Gastric cancer detection using deep learning. Nat Commun 2020;11(1):4294.
[8] Muti HS, Heij LR, Keller G, et al. Development and validation of deep learning classifiers to detect epstein-barr virus and microsatellite instability status in Gastric cancer: a retrospective multicentre cohort study. Lancet Dig Health 2021;3(10):e654–64.
[9] Wang X, Chen Y, Gao Y, et al. Predicting Gastric cancer outcome from resected lymph node histopathology images using deep learning. Nat Commun 2021;12(1):1637.
[10] Wang S, Zhu Y, Yu L, et al. RMDL: recalibrated multi-instance deep learning for whole slide Gastric image classification. Med Image Anal 2019;58:101549.