WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 3, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v3n1.07
Journal homepage: https://wepub.org/index.php/IJCSIT/index

A Review of Artificial Intelligence in Tumor Pathology Image Analysis

Saisai Feng, Mingchuan Zhang

Henan University of Science and Technology, Luoyang 471000, China

ABSTRACT

Accurate diagnosis of tumors is crucial to the treatment and prognosis of patients. Pathological diagnosis is regarded as the "gold standard" of tumor diagnosis, which helps to detect the disease at an early stage and formulate precise treatment plans for patients. However, traditional pathology diagnosis relies heavily on the expertise and diagnostic experience of physicians, making the quality and accuracy of pathology diagnosis largely dependent on their individual capabilities. With the popularization of Whole Slide Image (WSI) technology, the application of AI in pathology has gained significant momentum. With its powerful analyzing ability, AI has been widely used in computational pathology, especially in pathology-assisted diagnosis, showing great potential. This paper first explores two core tasks of AI in the field of pathology image analysis - image segmentation and image classification. Finally, it looks at the challenges and opportunities facing the field.

KEYWORDS

Artificial intelligence; Image segmentation; Image classification; Pathology image analysis

1. INTRODUCTION

Cancer, as a serious disease, poses a great threat to both individuals and society and has become a global health problem. According to the International Agency for Research on Cancer (IARC), in 2020, there were 19.29 million new cancer cases and as many as 9.96 million deaths worldwide. Histopathology is an important branch in the field of diagnostic pathology, which is mainly used to examine the changes in the morphology of patient tissues, and by observing the information of tissue structure and morphology, doctors can further determine the type and grade of the tumor, so histopathological diagnosis is the gold standard for clinical diagnosis [1-2]. In the traditional pathological diagnosis process, pathologists usually need to examine pathological sections with the help of a microscope. By observing the presence of lesions in the sections and their distribution, the doctor can make a pathologic diagnosis and prognostic assessment. However, this approach is highly dependent on the doctor's professional knowledge and diagnostic experience, making the quality and accuracy of the pathology diagnosis entirely dependent on the doctor's personal ability. At present, the pathology departments of major hospitals are facing the problems of a long training cycle for doctors and a huge gap in the talent pool, and at the same time, with the increasing number of patients, the pathology departments are generally overloaded. This situation is likely to lead to omissions and misdiagnosis in the diagnostic process, which will have a negative impact on the treatment and prognosis of patients.

In recent years, with the continuous improvement of computing power and medical data, computer-aided diagnosis (CAD) systems have seen rapid development in the field of pathology image analysis [3-5]. By utilizing CAD systems, doctors' work efficiency has been significantly improved, and the problem of inconsistent diagnostic results due to subjective judgment has also been reduced to a certain extent [6].

The purpose of this paper is to provide an in-depth and systematic overview of the state-of-the-art research on the application of artificial intelligence techniques in the field of pathology, especially in pathology-assisted diagnosis. The paper first introduces the importance of whole slide digital scanning technology, and then describes the current research status of image segmentation and classification in pathology image analysis. Finally, it summarizes the full paper and looks forward to the future challenges and opportunities in this field.

2. WHOLE SLIDE DIGITAL SCANNING TECHNOLOGY

Whole slide digital scanning technology is a technology that organically combines computer technology with the traditional field of pathology. As shown in Fig. 1, it first scans the tissue section through a scanner, converts it into a digital image, and then applies a computer to seamlessly stitch multiple pathology images together, thus obtaining high-quality whole-slide pathology image data to form a digital section or virtual section.

Figure 1. Process of digitizing pathology slides [7]

The advent of WSI technology has gradually transformed pathology slides, which could only be viewed through a microscope, into data that can be stored on a hard disk. This change has made it easier to store and transport pathology slides, and has also opened up the possibility of using computers to assist physicians in pathology diagnosis. Pathology images usually have complex and diverse features, and in order to better display tissue structure and cellular morphology, the most common method used today is to process pathology sections using hematoxylin eosin (H&E) staining, which is simple to produce and less expensive [8]. However, H&E-stained sections from different batches or between different institutions can suffer from staining differences, so the traditional image processing methods are not effective in coping with complex pathology images and cannot be used for practical auxiliary diagnosis.

3. APPLICATION OF ARTIFICIAL INTELLIGENCE IN PATHOLOGY IMAGE ANALYSIS

In recent years, there have been a number of researchers attempting to analyze digitized pathology images using AI techniques [9], which has improved the diagnostic efficiency and accuracy of doctors to a certain extent. In this section, we will delve into the research progress related to the two algorithms of classification and segmentation in the task of pathology image analysis.

3.1. Application of Classification Algorithms in Pathology Image Analysis

Image classification is a fundamental task in the field of computer vision, which aims to automatically label or classify images into predefined categories based on the semantic information in the images. In recent years, deep learning has been widely used in pathology image classification. Deep learning models can automatically extract features by learning a large amount of pathology image data and construct more complex classifiers. This method can greatly improve the accuracy and efficiency of classification and provide more reliable support for the classification and analysis of pathology images.

For lesion classification of cervical precancerous pathology images, AlMubarak et al [10] used an approach that fuses deep learning and traditional image processing. They first used a distance transform method to identify the mid-axis of the cervical epithelium, followed by dividing the cervical epithelium into 10 vertical segments using the mid-axis as a baseline. Then, by using the modeling approach proposed by the authors, they were able to extract depth features and hand-designed features for each vertical segment. Finally, by integrating these features, they trained Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) classifiers to achieve accurate classification of cervical precancerous pathology images. Araújo et al [11], in a study on breast cancer tissue classification, in order to improve the sensitivity to the cancer category, the authors divided the original image into 12 consecutive, non-overlapping image chunks, and then used a CNN+SVM scheme to calculate the probability of each image chunk. Finally, by employing a probabilistic fusion method, the authors effectively fused the probabilistic information from different image chunks to produce the final classification results of the images.

3.2. Application of Segmentation Algorithms in Pathology Image Analysis

Segmentation algorithms play a crucial role in the field of pathology-assisted diagnosis. These algorithms are able to accurately segment key basic elements and tissue regions from complex pathology images, providing doctors with important reference information to assist them in accurate diagnosis and condition assessment. Taking colorectal cancer as an example, by applying advanced segmentation algorithms, we are able to accurately extract glandular regions and further analyze the morphological features of these glands, such as size, number and other key indicators. This information not only helps doctors better understand the lesion process, but also effectively assess the disease progression of colorectal cancer, providing strong support for the development of accurate treatment plans.

During the traditional semantic segmentation period, in order to be able to accurately segment the glandular region, Wu et al [12] designed a segmentation algorithm based on binary images, which used an adaptive thresholding approach to more accurately separate the nucleus and cytoplasm. After completing the initial segmentation, the authors also optimized the final segmentation results by removing pseudo-glands. With the continuous improvement of medical data, more and more researchers have begun to apply deep learning techniques to the field of medical image segmentation, which has led to a significant improvement in the accuracy and speed of medical image segmentation [13-14].Feng et al [15], in order to accelerate the training speed of the model, they used a migration learning strategy during the training process, which used the pre-trained VGG-16 as the model's encoder, which can not only accelerate the training speed of the model, but also effectively improve the segmentation performance of the model. In addition, to address the problem of the imbalance in the number of benign and malignant samples in the dataset, the authors also designed a specific loss function, which can better deal with the data with imbalance in categories in the segmentation task. Yu [16] et al. took into account that the annotation process of medical data is difficult, so they proposed an enhancement method of cell nucleus segmentation based on the diffusion model. Firstly, they trained an unconditional diffusion model to synthesize cell nucleus structures defined as pixel-level semantic and distance transformed representations, and these synthesized cell nucleus structures will be used as constraints for the synthesis of histopathology images; then, they trained a conditional diffusion model to synthesize histopathology images based on cell nucleus structures. Finally, the synthesized constraints are paired with the synthesized histopathology images and added to the real dataset for training the segmentation model. The experimental results show that segmentation results comparable to those of fully supervised networks can be obtained using datasets composed of synthetic samples.

4. SUMMARY AND OUTLOOK

Pathological diagnosis, as the "gold standard" of clinical diagnosis, is an important reference for doctors to make accurate diagnosis for patients, but with the increasing number of patients, the pathology department is generally overloaded with work. At the same time, traditional pathology diagnosis is also highly dependent on the doctor's professional knowledge and diagnostic experience, which makes the quality and accuracy of pathology diagnosis depend largely on the doctor's personal ability, and this situation can easily lead to the doctor's diagnosis in the diagnostic process of omission and misdiagnosis. This paper firstly describes the application prospect of artificial intelligence technology in the field of pathology assisted diagnosis, then introduces the development of whole slide digital scanning technology for the application value of pathology assisted diagnosis, and finally introduces the current research status of classification algorithms and segmentation algorithms in pathology assisted diagnostic models respectively.

Currently, the application of artificial intelligence in the field of pathology-assisted diagnosis is still in the budding stage, and although many scholars have already achieved preliminary results through image processing technology, a complete set of intelligent assisted diagnosis process has not yet been formed. However, with the continuous enrichment and improvement of pathology data, more and more researchers are actively engaging in the cutting-edge exploration of this field, with a view to promoting the intelligent development of pathology diagnosis. Meanwhile, the deep integration of artificial intelligence and pathology diagnosis will greatly promote innovation and progress in the field of computer-aided diagnosis (CAD). This combination not only improves the accuracy and efficiency of diagnosis, but also makes the application of CAD technology in the medical field more extensive and in-depth. With the continuous development and improvement of the technology, we have reason to believe that the union of artificial intelligence and pathology diagnosis will play an even more important role in the future of smart healthcare.

ACKNOWLEDGEMENTS

This work was supported in part by the Key Technologies R & D Program of Henan Province under Grant No. 232102210028, 232102211008 and 232102210048.

REFERENCES

[1] Sun Z C, Liu L X. Progress and prospect of artificial intelligence in diagnosis and treatment of liver cancer [J]. Medical Review, 2020, 26(22): 4421-4424.

[2] Xue K. Digital pathology+artificial intelligence to promote new development of precise diagnosis [N]. China Science News, 2017-11-23(006).

[3] AlZubaidi A K, Sideseq F B, Faeq A, et al. Computer aided diagnosis in digital pathology application: Review and perspective approach in lung cancer classification [C]. 2017 annual conference on new trends in information & Communications technology applications (NTICT). IEEE, 2017: 219-224.

[4] Nam S, Chong Y, Jung C K, et al. Introduction to digital pathology and computer-aided pathology [J]. Journal of Pathology and Translational Medicine, 2020, 54(2): 125.

[5] Bengtsson E, Danielsen H, Treanor D, et al. Computer-aided diagnostics in digital pathology [J]. Cytometry Part A, 2017, 91(6): 551-554.

[6] Shafi S, Parwani A V. Artificial intelligence in diagnostic pathology [J]. Diagnostic Pathology, 2023, 18(1): 109.

[7] Farahani N, Parwani A V, Pantanowitz L. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives [J]. Pathology and Laboratory Medicine International, 2015: 23-33.

[8] Suda R, Sakai N, Matsushita K, et al. Prediction of mismatch repair deficient biliary tract cancer: Role of morphological features and host immune response detected by routine hematoxylin-eosin staining [J]. Journal of Hepato-Biliary-Pancreatic Sciences, 2021, 28(8): 680-691.

[9] Wang Y, Crookes D, Eldin O S, et al. Assisted diagnosis of cervical intraepithelial neoplasia (CIN) [J]. IEEE Journal of Selected Topics in Signal Processing, 2009, 3(1): 112-121.

[10] AlMubarak H A, Stanley J, Guo P, et al. A hybrid deep learning and handcrafted feature approach for cervical cancer digital histology image classification [J]. International Journal of Healthcare Information Systems and Informatics, 2019, 14(2): 66-87.

[11] Araújo T, Aresta G, Castro E, et al. Classification of breast cancer histology images using convolutional neural networks [J]. PloS one, 2017, 12(6): e0177544.

[12] Wu H S, Xu R, Harpaz N, et al. Segmentation of intestinal gland images with iterative region growing [J]. Journal of Microscopy, 2005, 220(3): 190-204.

[13] Wang R, Chen S, Ji C, et al. Boundary-aware context neural network for medical image segmentation [J]. Medical Image Analysis, 2022, 78: 102395.

[14] Lin J, Lin J, Lu C, et al. CKD-TransBTS: Clinical knowledge-driven hybrid transformer with modality-correlated cross-attention for brain tumor segmentation [J]. IEEE Transactions on Medical Imaging, 2023, 42(8): 2451-2461.

[15] Feng R, Liu X, Chen J, et al. A deep learning approach for colonoscopy pathology WSI analysis: Accurate segmentation and classification [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(10): 3700-3708.

[16] Yu X, Li G, Lou W, et al. Diffusion-based data augmentation for nuclei image segmentation [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2023: 592-602.