Research progress of artificial intelligence in medical imaging field

Kaiyan Wang [1, a]

[1]Medical School, Tianjin University, China

[a] U3021004152@tju.edu.cn

Abstract. With the rapid development of medical equipment, medical imaging has entered the era of big data. How to process massive data quickly and accurately extract critical information for disease diagnosis and treatment is a major challenge facing clinical practice. Artificial intelligence has unique advantages in real-time processing, prediction and analysis, transfer learning, etc., which provides a new breakthrough for solving this problem. This review reviews the artificial intelligence models applied to medical image processing in recent years, shows their wide application scenarios, briefly discusses their unique advantages from three aspects of image recognition, image enhancement and image registration, and finally summarizes the current research status and puts forward some prospects.

Keywords: Artificial intelligence medical imaging

1. Introduction

With the advances in graphics processors in recent years, new potential has been activated by Artificial intelligence (AI) [1]. At present, Deep learning (DL) is the main research direction in the field of image processing. Depth usually refers to the number of layers and is used to describe how a Deep neural network (DNN) can reuse a given feature in multiple hidden layers. Compared with shallow neural networks, DNNS can learn and extract feature representations of data through multi-layer neural networks, thus minimizing loss functions and reducing generalization errors with fewer weights and units [2].Medical images are of great significance for disease diagnosis and analysis, and are usually stored in DICOM, JPEG, PNG and other formats. Different from the image processing fields of other disciplines, the object of medical image is human, so researchers need to accurately extract the image information that human eyes can distinguish from the data set. AI is able to automatically learn and extract features specified by people in an image without manually designing a feature extractor. Moreover, DLs multi -layer structure enables it to learn hierarchical feature representations and analyze image information in more detail.

2. Research status

2.1 Main application scenarios

Traditional machine learning algorithms are hampered by the high cost of annotating todays medical data sets [3]. In 2012, Alex Krizhevskys team proposed the Alex Net model in a competition, which successfully promoted the large-scale use of DL. Today, DL application scenarios are often divided into supervised learning and unsupervised learning.

When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use. Supervised learning refers to training by using training data with known outputs. Sahin et al. [4] selected 4000 lung computed tomography (CT) images from Yozgat Bozok University to automatically quantify the severity of coronavirus disease in CT images. At the same time, they also counted the classification losses for each region of interest, the classification losses and total training losses in the region proposal network, and the average accuracy proved the superior performance of the model. Similarly for the processing of lung CT images, in order to avoid the difficulty of manually annotating the nodule region in nodule detection,

V. Bishnoi and N. Goel [5] established a DL classification model based on complete CT images instead of marker annotation selection.The model was trained on 19,419 open lung CT slices, which was more complete and comprehensive. The framework includes K-means clustering for preprocessing and segmentation, weighted VGG deep network for model tuning and regularization, and TensorRT library for model acceleration and optimization. Both of them are study cases based on supervised learning, which reflects the characteristics of accurate and rapid processing of medical data.

Unsupervised learning is more suitable for dealing with situations where the target variables of training data are insufficient, such as incomplete database due to insufficient research on related diseases, and data sets are not disclosed for the protection of patient privacy. Faced with medical data with few samples, Guo et al. [6] generated adversarial model MedGAN based on adaptive alternate training strategy to enhance images. They trained the model on just 20 samples from each of the four disease types, blister, anthropozoosis, parakeratosis, and mollusk, taking a total of 1,350 seconds over tens of thousands of iterations. The generated images they presented all had clear lesion features. In addition, the idea of unsupervised learning is not limited to two-dimensional images. Y.Li et al. [7] chose a special neural network structure autoencoder VAE model to reconstruct 3D lung nodule images. The training data was very small, only 36 CT images of lesions were selected from the Stanford University data set for training. Unsupervised learning is also suitable for processing a sufficient number of unlabeled images. J.Li et al. [8] proposed a self-supervised deep adaptive regularization clustering framework. The researchers first input unlabeled histopathological images into the neural network to extract the representation, then cluster them, and can adaptively update the objective function throughout the training process. These examples illustrate the wide range of applications of unsupervised learning, which is why this area is becoming a major research area in deep learning today.

2.2 Flexible application based on DL

DL framework has many flexible applications in medical image processing, and its understanding should not be limited to the two categories of supervised and unsupervised. DNNS based on semi-supervised learning can learn meaningful representations or features from large-scale unlabeled data. Rafiei et al. [9] successfully combined the supervised learning model with the unsupervised learning model to analyze a large number of unlabeled electroencephalograms. Transformer is originally used for processing natural language. Pu et al. [10] have learned how its self-attention mechanism captures long-distance dependence of images to understand complex medical image content and make up for the lack of information captured by convolutional kernel. In addition,Hu et al. [11] argued that reinforcement learning ADAPTS to complex application scenarios by interacting with the environment and maximizing expected cumulative rewards. For example, tumor detection is a dynamic decision-making process, and reinforcement learning can dynamically adjust the decision-making according to the feedback of the environment, even in the case of scarce data, good results can be achieved.

Researchers sometimes need to use multiple algorithms at the same time to improve training efficiency based on DL. Venkatapathy et al. [12] used an integrated Graph Neural Network GNN model based on multiple basic models. Based on resting state functional magnetic resonance imaging (FMRI) data, they used this model to construct a graph structure that simulates brain functional networks and effectively distinguishes different types of depression.

3. Advantages of AI in the field of image processing

With the popularity of deep learning, AI has shown many outstanding performances in processing image extraction features.

In terms of image reconstruction and data enhancement, AI can precisely learn complex mapping relationships from input to output, thereby extracting key information from limited data to help reconstruct high-quality images, enriching medical datasets. It can also be further trained on large amounts of data to build models with excellent generalization ability. Reader and Pan[13] analyzed the data of positron emission tomography images by convolutional coding layer learning, and then learned multiple advanced feature combinations by convolutional layer analysis.Chen et al. [14] learned how to generate a large number of medical images based on the label map of fundus and skin images, combined with a large number of real images, and finally applied to the adversarial training of downstream models.

In terms of image recognition, unlike traditional methods, DNN models can eliminate the need for experts to manually design features, simplifying the development process and improving overall performance. At the same time, each layer is responsible for extracting different levels of features, which can accurately distinguish and locate individual objects in the image. In addition, the backpropagation algorithm in DNN is based on a large amount of training, so that the model is constantly adjusted to adapt to different task requirements.Dolz et al. [15] used progressively expanded convolutional neural networks to segment MRI images in a progressive manner, so as to accurately identify and segment bladder walls and tumors. Zhang et al. [16] designed an automatic marking method to replace manual marking while also improving accuracy.

In the aspect of image registration, DL framework can learn complex nonlinear mapping relationships. Researchers can optimize the loss function in the training process to achieve more accurate image alignment and registration, and can also simplify the step-by-step optimization process to reduce the risk of error caused by human intervention.For example, in the field of image registration, the usual operations are rigid registration: translation, rotation, scaling. Based on DL, Bouza et al. [17] applied non-rigid registration technology to deal with the image registration problem of Diffusion Magnetic Resonance Images (DMRI) of the brain, improving the efficiency of processing complex brain images such as DMRI. Fang et al. [18] divided the process of brain image registration into different modules. At the same time, constraints are introduced to train the model so that the highly ill-conditioned inverse problem can be decomposed into simpler subproblems.Liu et al. [19] combined the probabilistic model based on CNN and the contrastive learning mechanism under the DL framework. The contrast learning mechanism can help the network learn more distinguishing feature representations, better retain image details when processing 3D brain MRI images, and reduce the blur effect during registration.

4. Summary

In recent years, the development of artificial intelligence technology is extremely rapid, which has deeply influenced and helped the development direction of the medical image field. With the excellent learning ability of deep learning, medical image data sets have been greatly enriched, and more comprehensive and accurate features have been extracted from traditional image omics images. However, we should also be aware that there are many challenges to research in this area. Including but not limited to improving model generalization ability, reducing model training costs, intellectual copyright disputes and patient privacy disclosure. The development direction of artificial intelligence technology in medical imaging should pay more attention to the following aspects: 1. How to avoid doctors ignoring image quality problems due to the improvement of image quality by artificial intelligence. 2. How to improve the ability to explain the results generated by AI to avoid over-explanation or insufficient basis. 3 While updating more specialized components to develop more complex models, it is also necessary to reduce the time and economic costs required for large model computational analysis by designing chips that are more suitable for AI computing or by promoting computing infrastructure and cloud services. 4 How to provide a more personalized AI medical system. It is believed that in the future, with more researchers investment and financial support, AI technology will have more fruitful research and application results in the field of medical image processing.

References

[1] S. Raschka, J. Patterson, and C. Nolet, "Machine Learning in Python: MAIn Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence," Information, vol. 11, no. 4, Art. no. 4, pp. 193, April 2020.

[2] N. Kriegeskorte and T. Golan, "Neural network models and deep learning," Current Biology, vol. 29, no. 7, pp. R231-R236, April 2019.

[3] M. Sharma, R. K. Patel, A. Garg, R. SanTan and U. R. Acharya, "Automated detection of schizophrenia using deep learning: a review for the last decade," Physiological Measurement, vol. 44, no. 3, page 03TR01, March 2023.

[4] M. E. Sahin, H. Ulutas, E. Yuce and M. F. Erkoc, "Detection and classification of COVID-19 by using faster R-CNN and mask R-CNN on CT images," Neural Computing and Applications, vol. 35, no. 18, pp. 13597-13611, June 2023.

[5] V. Bishnoi and N. Goel, "Tensor-RT-Based Transfer Learning Model for Lung Cancer Classification," Journal of Digital Imaging, vol. 36, no. 4, pp. 1364-1375, August 2023.

[6] K. Guo et al., "MedGAN: An adaptive GAN approach for medical image generation," Computational Biology and Medicine, vol. 163, page 107119, September 2023.

[7] Y. Li, C. Y. Sadee, F. Carrillo-Perez, H. M. Selby, A. H. Thieme and O. Gevaert, "A 3D lung lesion variational autoencoder," Cell Reports. Methods, pp. 100695, January 2024.

[8] J. Li et al., "DARC: deep adaptive regularized clustering for histopathological image classification," Medical Image Analysis, vol. 80, p. 102521, August 2022.

[9] M. H. Rafiei, L. V. Gauthier, H. Adeli and D. Takabi, "Self-Supervised Learning for Electroencephalography," IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 2, pp. 1457-1471, February 2024.

[10] Q. Pu, Z. Xi, S. Yin, Z. Zhao and L. Zhao, "Advantages of transformer and its application for medical image segmentation: a survey," Biomedical Engineering Online, vol. 23, no. 1, page 14, February 2024.

[11] M. Hu, J. Zhang, L. Matkovic, T. Liu and X. Yang, "Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions," Journal of Applied Clinical Medical Physics, vol. 24, no. 2, page e13898, February 2023.

[12] S. Venkatapathy et al., "Ensemble graph neural network model for classification of major depressive disorder using whole brAIn Functioning Connectivity," Frontiers in Psychiatry, vol. 14, p. 1125339, 2023.

[13] A. J. Reader and B. Pan, "AI for PET image reconstruction," British Journal of Radiology, vol. 96, no. 1150, p. 20230292, October 2023.

[14] Y. Chen et al., "Generative Adversarial Networks in Medical Image augmentation: A review," Computational Biology and Medicine, vol. 144, pp. 105382, May 2022.

[15] J. Dolz et al., "Multiregion segmentation of bladder cancer structures in MRI with progressive dilated convolutional networks," Medical Physics, vol. 45, no. 12, pp. 5482-5493, December 2018.

[16] J. Zhang, C. Santos, C. Park, M. A. Mazurowski and R. Colglazier, "Improving Image Classification of Knee Radiographs: An Automated Image Labeling Approach," Journal of Digital Imaging, vol. 36, no. 6, pp. 2402-2410, December 2023.

[17] J. J. Bouza, C.-H. Yang and B. C. Vemuri, "Geometric Deep Learning for Unsupervised Registration of Diffusion Magnetic Resonance Images," Lecture Notes in Computer Science, vol. 13939, pp. 563-575, 2023.

[18] J. Fang et al., "Decoupled learning for brAIn image registration," Frontiers in Neuroscience, vol. 17, p. 1246769, 2023.

[19] L. Liu, A. I. Aviles-Rivero and C.-B. Schonlieb, "Contrastive Registration for Unsupervised Medical Image Segmentation," IEEE Transactions on Neural Networks and Learning Systems, vol. PP, pp. 1-13, 2024.