DOI: 10.5582/bst.2023.01148

Editorial

Advances in deep learning: From diagnosis to treatment

Tianqi Huang[1], Longfei Ma[1], Boyu Zhang[2], Hongen Liao[1,*]

[1] Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China;
[2] Research Center for Industries of the Future, and Key Laboratory of 3D Micro/Nano Fabrication and Characterization of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, Zhejiang, China.

SUMMARY Deep learning has brought about a revolution in the field of medical diagnosis and treatment. The use of deep learning in healthcare has grown exponentially in recent years, achieving physician-level accuracy in various diagnostic tasks and supporting applications such as electronic health records and clinical voice assistants. The emergence of medical foundation models, as a new approach to deep learning, has greatly improved the reasoning ability of machines. Characterized by large training datasets, context awareness, and multi-domain applications, medical foundation models can integrate various forms of medical data to provide user-friendly outputs based on a patien's information. Medical foundation models have the potential to integrate current diagnostic and treatment systems, providing the ability to understand multi-modal diagnostic information and real-time reasoning ability in complex surgical scenarios. Future research on foundation model-based deep learning methods will focus more on the collaboration between physicians and machines. On the one hand, developing new deep learning methods will reduce the repetitive labor of physicians and compensate for shortcomings in their diagnostic and treatment capabilities. On the other hand, physicians need to embrace new deep learning technologies, comprehend the principles and technical risks of deep learning methods, and master the procedures for integrating them into clinical practice. Ultimately, the integration of artificial intelligence analysis with human decision-making will facilitate accurate personalized medical care and enhance the efficiency of physicians.

Keywords deep learning, foundation model, integrative diagnosis and treatment, human-machine collaboration

Deep learning (DL) is a novel research direction in the field of machine learning. Deep learning involves inputting raw data into a machine, which then develops the representations required for pattern recognition through multiple network layers. By combining these multilayer neural networks, computers can learn highly complex functions to perform tasks (*1*). DL methods have achieved significant success in various medical imaging tasks (*2*) and medical natural language processing (NLP) tasks (*3*). In medical image analysis, deep learning methods such as R-CNN and U-Net have achieved physician-level accuracy in various diagnostic tasks, including detection of diabetic retinopathy, classification of liver lesions on CT, detection of breast lesions on X-rays, assessment of hepatic diseases on ultrasound, and analysis of the spine on MRI (*4,5*). In these tasks, deep learning technologies can assist physicians with diagnosis by providing second opinions and annotating regions in images. For medical NLP tasks, deep learning methods have been used to support various applications, such as electronic health records and clinical voice assistants. Recurrent neural networks, for instance, can model the time series of structured events cited in patient records to predict future clinical events (*6*). Automatic speech recognition technology can convert patient-doctor conversations into transcribed text records and clinical voice assistants can be developed to accurately record patient visits (*7*).

Technological advancements in medical foundation models

Previous research on deep learning has primarily focused on low-level tasks in image processing and NLP, such as segmentation of medical images, identification of lesion features, classification of disease, and semantic analysis of medical information. Advancements in deep learning technology have enabled these methods to achieve the accuracy level of professional physicians in these low-level tasks. These deep learning methods can effectively assist physicians in their work by reducing the time spent on repetitive labor in these tasks and improving medical efficiency.

In recent years, research on foundation models has brought new breakthroughs. The characteristics of foundation models include large training datasets, context awareness, multi-domain applications, and the ability to perform more advanced tasks by integrating information and performing reasoning (*8*). In the medical field, ChatGPT has approached the passing threshold for the United States Medical Licensing Examination without any specialized training, demonstrating its potential in medical diagnosis and clinical decision support (*9*).

How can medical foundation models benefit treatment?

Medical foundation models possess significant value due to their exceptional reasoning ability, enabling the transition from low-level to high-level tasks. Compared to current deep learning methods, medical foundation models have the potential to perform more diverse and challenging tasks. Current foundation models, such as GPT-4, are now capable of comprehending multiple forms of information, including language and images (*10*). Future research can utilize diverse and extensive datasets to enable medical foundational models to flexibly integrate various forms of medical data, including medical images, electronic health records, genomics, charts, and medical texts. These models can provide user-friendly outputs based on multiple forms of patient information, such as diagnostic suggestions or image annotations. By utilizing these capabilities, medical professionals can make more informed decisions and provide better care to patients.

Within treatment scenarios, deep learning technology is primarily used for surgical planning and decision-making. However, there is a growing interest in automated operations by surgical robots. The autonomous operation of robots during surgery using deep learning presents a challenging yet promising direction. Reinforcement learning methods in particular have shown potential in enabling robots to autonomously perform simple surgical tasks (*11,12*). However, more complex surgical scenarios necessitate more complex reasoning ability so that robot operations can be modified based on intraoperative diagnostic information. Researchers have developed various surgical devices to integrate diagnosis and treatment (*13,14*). Medical foundation model technology can provide such devices with the capabilities to effectively analyze medical information and perform diagnostic decision-making. Integrating different diagnostic modes and various forms of medical information in medical foundation models can ensure reliable, timely, and intelligent diagnostic conclusions, thereby improving the efficiency and quality of treatment.

Collaboration between physicians and machines

Currently, research on foundation models in the medical field is still in its early stages. However, we believe that these models have great potential for future applications. With advancements in technology and computing power, the performance of foundation models, including their understanding of information and task accuracy, will continue to improve. Ultimately, creating a fully autonomous diagnosis and treatment system that only requires patients to lie down on the machine for automatic diagnosis and accurate treatment may be possible. At this stage of research, however, autonomous robot systems that can be used in clinical practice without the involvement of physicians is still a challenge. The primary reasons for this include: 1). the complexity of medical data, which are often unstructured, in multiple forms, and individualized; 2). the non-repeatability and safety requirements of medical procedures; and 3. the diagnosis and treatment of diseases involves social and humanistic factors that machines alone cannot fully incorporate.

Researchers should prioritize enhancing physicians' diagnostic and treatment capabilities and reducing physicians' repetitive labor through deep learning technology, rather than replacing them. The development of deep learning technology should aim to address the inadequacies of human skills by assisting with diagnosis and treatment. At the same time, physicians must embrace artificial intelligence and deep learning technology, as they have with other tools such as medical imaging and computer-assisted surgical navigation over the past few decades. To do so, they must comprehend the principles and technical risks of deep learning methods and they must master the procedures for integrating them into clinical practice. Only then can they adapt to the development of deep learning technologies in this era, maintain their independence amidst technological progress, and promote personalized healthcare. Ultimately, the integration of artificial intelligence analysis with human decision-making will facilitate accurate personalized medical care and enhance the efficiency of physicians.

Funding: None.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436-444.
2. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Image Anal. 2017; 42:60-88.
3. Hahn U, Oleynik M. Medical information extraction in the age of deep learning. Yearb Med Inform. 2020; 29:208-220.
4. Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K,

Mannel RS, Liu H, Zheng B, Qiu Y. Recent advances and clinical applications of deep learning in medical image analysis. Med Image Anal. 2022; 79:102444.
5. Ma LF, Wang R, He Q, Huang LJ, Wei XY, Lu X, Du YN, Luo JW, Liao HE. Artificial intelligence-based ultrasound imaging technologies for hepatic diseases. iLIVER. 2022; 1:252-264.
6. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. JMLR Workshop Conf Proc. 2016 ; 56:301-318.
7. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE J Biomed Health Inform. 2017; 22:1589-1604.
8. Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, Rajpurkar P. Foundation models for generalist medical artificial intelligence. Nature. 2023; 616:259-265.
9. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Madriaga M, Aggabao R, Diaz-Candido G, Maningo J, Tseng V. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023; 2:e0000198.
10. Bubeck S, Chandrasekaran V, Eldan R, *et al*. Sparks of artificial general intelligence: Early experiments with gpt-4. ArXiv230312712. 2023. https://doi.org/10.48550/arXiv.2303.12712
11. Ning G, Zhang X, Liao H. Autonomic robotic ultrasound imaging system based on reinforcement learning. IEEE Trans Biomed Engineer. 2021; 68: 2787-2797.
12. Saeidi H, Opfermann JD, Kam M, Wei S, Leonard S, Hsieh MH, Kang JU, Krieger A. Autonomous robotic laparoscopic surgery for intestinal anastomosis. Sci Robot. 2022; 7:eabj2908.
13. Liao H. Integrated diagnostic and therapeutic techniques: Toward an intelligent medical system. Computerized Med Imaging Graphics. 2014; 38:421-422.
14. Liao H, Noguchi M, Maruyama T, Muragaki Y, Kobayashi E, Iseki H, Sakuma I. An integrated diagnosis and therapeutic system using intra-operative 5-aminolevulinic-acid-induced fluorescence guided robotic laser ablation for precision neurosurgery. Med Image Anal. 2012; 16:754-66.

Received June 7, 2023; Accepted June 28, 2023.

**Address correspondence to:*
Hongen Liao, Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China.
E-mail: liao@tsinghua.edu.cn

Released online in J-STAGE as advance publication June 30, 2023.