Supplementary

Response to: “Role of large language models in medical image processing: issues that should be considered”

Dianzhe Tian^, Xin Lu, Yiyao Xu

Department of Liver Surgery, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Correspondence to: Xin Lu, MD; Yiyao Xu, MD. Department of Liver Surgery, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No.1 Shuaifuyuan, Wangfujing, Dongcheng District, 100730 Beijing, China. Email: luxin@pumch.cn; xuyiyao@pumch.cn.

Response to: Daungsupawong H, Wiwanitkit V. Role of large language models in medical image processing: issues that should be considered. Quant Imaging Med Surg 2024. doi: 10.21037/qims-24-89.

We acknowledge the valid point about the potential bias introduced by limiting publications to English. We agree that inclusivity is critical, and there may indeed be valuable studies in other languages. However, we made this choice because the frontiers of LLM research are currently concentrated in the United States and China, and the most recent highly cited articles are also published in the English literature (1). Simultaneously, since the application of LLMs in the field of medical image processing is brand new, the quantity of the most recent literatures is relatively small, which are mainly concentrated in English literatures. Therefore, the non-English literatures are almost negligible.

Regarding the timeframe and methodology used in the identification of the articles mentioned by the authors of the comment, the authors have discussed in detail in the original paper (2). From the knowledge of the LLMs and the related convolutional neural network (CNN) model, we set the timeframe of the literature search from 2013 to 2023 to achieve a more comprehensive screening (3).

The proposal to justify the use of Web of Science (WoS) and PubMed databases as primary sources for literature searches will be considered. As we mentioned in the original paper, recent publications in the arXiv database were also reviewed to catch up with the recent research breakthrough in LLMs in medical image segmentation. However, for papers on arXiv that have not undergone peer review, it is even more crucial for authors to conduct thorough scrutiny and selection before inclusion. We agree with the suggestions that other professional databases should be considered. We thoroughly search another professional database in medical image segmentation, IEEE Xplore, which is a preeminent electronic resource offered by the Institute of Electrical and Electronics Engineers (IEEE), providing a comprehensive and authoritative platform for a wide range of engineering and technology disciplines, including the critical area of medical imaging. However, after the literature search with the search strategy (“All Metadata”: GPT) AND (“All Metadata”: Medical image), only 13 results are shown, 2 results are already included in PubMed, 2 results are completely irrelevant to the title, 8 results are almost irrelevant, and only 1 document matches the topic the article wishes to address (4). Technical databases such as IEEE Xplore, SPIE Digital Library can be an effective complement to PubMed and WoS in the field of medical imaging, but the volume of literature is small and can be almost negligible in the early stages of research.

While our study emphasizes the advantages and possibilities of LLMs, we also recognize the importance of discussing potential drawbacks and areas for improvement in the discussion part, including 4 main concerns: interpretability of LLMs, quality of hardware infrastructure, real-time performance of LLMs, and ethical considerations (2). The potential enhancements are

^ ORCID: 0009-0006-3618-9054.

 https://dx.doi.org/10.21037/qims-24-89

elaborated in the original article.

There are indeed some special types that may not be suitable for LLMs when it comes to medical image data (5). Firstly, highly sensitive personal information contains extremely sensitive medical information, such as images of certain diseases, which may be subject to privacy and compliance constraints. Secondly, images in certain medical fields may require specialized domain knowledge to interpret correctly. LLMs may lack this specialized knowledge, resulting in inaccurate interpretation of images. Thirdly, LLMs may not be able to fully understand or accurately process medical images of rare diseases since there are relatively little data of rare cases. Specialized medical imaging techniques may also be too complex for LLMs and require more specialized interpretation and understanding.

Improvements and variations in medical image processing using LLMs may vary between datasets and healthcare facilities (6). Improvements in LLMs for medical imaging may depend on the diversity of the datasets used. If the dataset covers a wide range of diseases, image qualities, and sources, then the model may be more generalizable and applicable to different contexts. Differences in the level and equipment of healthcare facilities can also affect the effectiveness of LLMs in real-world applications (7). Simultaneously, there may be differences in treatment processes and practices between healthcare facilities that may affect the practical application of LLMs. The performance of the model may need to be adapted to the specific healthcare environment.

Many factors have a marked impact on the functionality and performance of LLMs in real-world situations (8). The performance of LLMs is directly influenced by the quality and diversity of the training data. If the training data are well represented and cover a wide range of contexts, especially domain expertise and continuous updating, the model may be more generalizable. From a technical perspective, the choice of model architecture and parameter tuning can make the model more adaptable to specific tasks and environments (9). Optimized hardware and software environments can also improve the efficiency of model execution. Finally, privacy and other ethical implications must be considered, as LLMs may face ethical and social issues in real-world applications, such as privacy disclosure, algorithmic bias, and social impact of the models.

Therefore, while improvements in LLMs for medical imaging can have a positive impact on a variety of datasets and healthcare facilities, the above factors must be considered in specific applications to ensure the effectiveness and repeatability of the model in each environment (10).

References

1. HAI_AI-Index-Report_2023.pdf [Internet]. [cited 2024 Feb 2]. Available online: https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
2. Tian D, Jiang S, Zhang L, Lu X, Xu Y. The role of large language models in medical image processing: a narrative review. Quant Imaging Med Surg 2024;14:1108-21.
3. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Jianming Liang. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Trans Med Imaging 2016;35:1299-312.
4. Ali H, Bulbul MF, Shah Z. Prompt Engineering in Medical Image Segmentation: An Overview of the Paradigm Shift. 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), Mount Pleasant, MI, USA, 2023:1-4.
5. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, Löffler CML, Schwarzkopf SC, Unger M, Veldhuizen GP, Wagner SJ, Kather JN. The future landscape of large language models in medicine. Commun Med (Lond) 2023;3:141.
6. Kim S, Lee CK, Kim SS. Large Language Models: A Guide for Radiologists. Korean J Radiol 2024;25:126-33.
7. Meskó B. The Impact of Multimodal Large Language Models on Health Care's Future. J Med Internet Res 2023;25:e52865.
8. Park SH, Sul AR, Ko Y, Jang HY, Lee JG. Radiologist's Guide to Evaluating Publications of Clinical Research on AI: How We Do It. Radiology 2023;308:e230288.
9. Gupta N. A Pre-Trained Vs Fine-Tuning Methodology in

https://dx.doi.org/10.21037/qims-24-89

Transfer Learning. J Phys Conf Ser 2021;1947:012028.
10. Daungsupawong H, Wiwanitkit V. Role of large language models in medical image processing: issues that should be considered. Quant Imaging Med Surg 2024. doi: 10.21037/qims-24-89.

https://dx.doi.org/10.21037/qims-24-89