Cell Reports Medicine

CellPress
OPEN ACCESS

Preview

Solving data quality issues of fundus images in real-world settings by ophthalmic AI

Zhongwen Li[1,2] and Wei Chen[1,2,*]
[1]Ningbo Eye Hospital, Wenzhou Medical University, Ningbo, China
[2]School of Ophthalmology and Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, China
*Correspondence: chenwei@eye.ac.cn
https://doi.org/10.1016/j.xcrm.2023.100951

Liu et al.[1] develop a deep-learning-based flow cytometry-like image quality classifier, DeepFundus, for the automated, high-throughput, and multidimensional classification of fundus image quality. DeepFundus significantly improves the real-world performance of established artificial intelligence diagnostics in detecting multiple retinopathies.

In recent years, artificial intelligence (AI), a revolutionary technology that comes out of computer science, has received widespread attention in the field of medicine as the initial results of applying AI to medical image classification are very promising.[2] Ophthalmology is uniquely suited for AI because of the massive image data ophthalmologists collect digitally, especially in the ocular fundus. These digital images provide an AI algorithm with a large amount of information that it can utilize to learn and evolve.

Ocular fundus diseases, such as diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma, affect millions of people worldwide. These diseases can cause irreversible vision impairment, visual field defect, and even blindness, having severe influence on the living quality of patients. Most vision loss caused by fundus diseases is avoidable via early diagnoses and treatment strategies. However, in the current era of relative shortage of ophthalmologists, early identification and timely referral for treatment of fundus diseases are challenging, especially in resource-limited settings.[3] The use of AI to automatically screen for fundus diseases on a large scale has the potential to address this issue.

Numerous studies have shown that medical AI algorithms achieve high accuracies in detecting multiple fundus diseases from fundus images, often at a human expert level.[4,5] Most of the studies trained and evaluated their AI models using only good-quality images. Nevertheless, in real-world clinical practice, images have different qualities, making the performance of these established models uncertain. The generation of poor-quality fundus images often results from operator errors, camera imperfections, patient noncompliance, and miscellaneous factors, which potentially lead to a loss of diagnostic information and negatively influence AI-based image analyses.[6] Therefore, it is essential to perform automated image quality assessment to filter out poor-quality images at the initial stage, ensuring only good-quality images would be used in the subsequent AI-based diagnostic analyses.

Several investigators established an automated image quality assessment system for the detection of poor-quality fundus images and advised that cases with poor-quality images should be referred to eye doctors for further evaluation.[7] However, in real-world clinical settings, especially in community hospitals, approximately half of fundus images were of low quality due to multiple socio-environmental factors, indicating that the referral of cases with poor-quality images would pose a significant burden on the healthcare system since most of the cases were found to be disease-free.[1]

To address this issue, Liu et al.[1] developed DeepFundus, a deep-learning-based flow cytometry-like image quality classifier, based on 12,710 images from 4 different cohorts and evaluated it using 53,141 images from another 22 different cohorts nationwide and the Kaggle dataset from the United States. DeepFundus can not only detect poor-quality images but also identify reasons that cause the generation of these images in real time, which can guide a technician to take a precise solution to acquire good-quality images immediately rather than direct referral of cases with poor-quality images. Notably, in addition to identifying technical factors (e.g., out of focus, underexposure, and positional deviation), DeepFundus can discern obscured images resulting from ocular media opacity (e.g., corneal edema, cataracts, and vitreous hemorrhage), which cannot be corrected by image recaptures. This feature helps avoid unnecessary recaptures and increases the efficiency of AI in clinical practice.

At present, most existing AI-based image quality assessment systems are used to detect poor-quality images corresponding to specific fundus diseases. For example, in glaucoma detection, a poor-quality image is defined if vessels within the optic disk area cannot be identified,[8] while in AMD detection, a poor-quality image is defined if vessels within the macular area cannot be recognized or over half of the macular area is obscured.[9] The generalizability of these image quality assessment systems is limited as the system constructed for one specific disease might not be adapted to another. Since an AI-based diagnostic system has already been employed to detect multiple fundus diseases, the construction of an image quality assessment system that can be applied in different clinical scenarios is more practical. Based on this consideration, Liu et al.[1] established DeepFundus consisting of 13 AI models to evaluate images in terms of multidimensional quality properties including overall quality, clinical quality factors, affected retinal structures, and refractive media opacity. This design ensures that DeepFundus can be widely used in various clinical scenarios to enhance model performance for detecting multiple retinopathies. In both internal and external test datasets, DeepFundus achieved areas under the receiver operating characteristic curves over 0.9 for all quality aspects, substantiating the effectiveness and generalizability of the system.

To assess the role of DeepFundus in the clinical deployment of medical AI, this image quality classifier, as a prescreening tool, was integrated into a certified AI diagnostic system to filter out poor-quality images prior to the detection of DR, AMD, and optic disc edema in real-world prospective cohorts.[1] The results showed that DeepFundus significantly enhanced the real-world performance of established AI diagnostic systems in detecting multiple fundus diseases, demonstrating that DeepFundus had a positive impact on the whole life cycle of medical AI practice.

Despite the remarkable performance of medical AI systems in good-quality images, they often have low accuracy in poor-quality images. However, the performance of human experts in poor-quality images is better than that of medical AI systems, displaying the vulnerability of these systems.[10] Because poor-quality images are inevitable in real-world clinical practice, approaches that can improve the performance of medical AI systems in poor-quality images deserve further investigation to bring the systems to the same level as human experts.

To accelerate the translation of AI into clinical settings, apart from avoiding garbage in AI and developing robust AI models, it is also important to focus on other key factors related to medical AI, including the overall healthcare strategies, standards, privacy protection, legislation, policy, regulation, and compliance with clinical and ethical guidelines. Even though this field is not completely matured yet, recent research progress shows that AI holds great promise to play a pivotal role in the future of ophthalmology, making diagnostics more efficient and accurate and assisting ophthalmologists in meeting the needs of a growing number of patients.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Liu, L., Wu, X., Lin, D., Zhao, L., Li, M., Yun, D., Lin, Z., Pang, J., Li, L., Wu, Y., et al. (2023). DeepFundus: a flow cytometry-like image quality classifier for boosting the whole life-cycle of medical artificial intelligence. Cell Rep. Med. *4*, 100912.
2. Rajpurkar, P., Chen, E., Banerjee, O., and Topol, E.J. (2022). AI in health and medicine. Nat. Med. *28*, 31–38.
3. Resnikoff, S., Lansingh, V.C., Washburn, L., Felch, W., Gauthier, T.M., Taylor, H.R., Eckert, K., Parke, D., and Wiedemann, P. (2020). Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): Will we meet the needs? Br. J. Ophthalmol. *104*, 588–592.
4. Cen, L.P., Ji, J., Lin, J.W., Ju, S.T., Lin, H.J., Li, T.P., Wang, Y., Yang, J.F., Liu, Y.F., Tan, S., et al. (2021). Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. Nat. Commun. *12*, 4828.
5. Son, J., Shin, J.Y., Kim, H.D., Jung, K.H., Park, K.H., and Park, S.J. (2020). Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. Ophthalmology *127*, 85–94.
6. Ruamviboonsuk, P., Tiwari, R., Sayres, R., Nganthavee, V., Hemarat, K., Kongprayoon, A., Raman, R., Levinstein, B., Liu, Y., Schaekermann, M., et al. (2022). Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: A prospective interventional cohort study. Lancet. Digit. Health *4*, e235–e244.
7. Ting, D.S.W., Cheung, C.Y.L., Lim, G., Tan, G.S.W., Quang, N.D., Gan, A., Hamzah, H., Garcia-Franco, R., San Yeo, I.Y., Lee, S.Y., et al. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA *318*, 2211–2223.
8. Li, Z., He, Y., Keel, S., Meng, W., Chang, R.T., and He, M. (2018). Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology *125*, 1199–1206.
9. Keel, S., Li, Z., Scheetz, J., Robman, L., Phung, J., Makeyeva, G., Aung, K., Liu, C., Yan, X., Meng, W., et al. (2019). Development and validation of a deep-learning algorithm for the detection of neovascular age-related macular degeneration from colour fundus photographs. Clin. Exp. Ophthalmol. *47*, 1009–1018.
10. Li, Z., Jiang, J., Qiang, W., Guo, L., Liu, X., Weng, H., Wu, S., Zheng, Q., and Chen, W. (2021). Comparison of deep learning systems and cornea specialists in detecting corneal diseases from low-quality images. iScience *24*, 103317.