ISSN: 2996-0940
Case Report
Science Set Journal of Ophthalmology and Eye Care

Automated Detection of Pathological Myopia from Colour Fundus Photographs using Code-free Deep Learning

Carolyn Yu Tung Wong[1, 2, 3*], Tin Lik Wong[3], Timing Liu[1, 2], and Henry Hing Wai Lau[4]

[1]*Institute of Ophthalmology, University College London, London, United Kingdom*
[2]*Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom*
[3]*Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China*
[4]*Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China*

[*]Corresponding author: Carolyn Yu Tung Wong, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, New Territories, Hong Kong.

Submitted: 18 June 2024 **Accepted:** 24 June 2024 **Published:** 02 July 2024

https://doi.org/10.63620/MKSSJOEC.2024.1014

Citation: *Wong, C. Y. T., Wong, T. L., Liu, T., & Lau, H. H. W. (2024). Automated Detection of Pathological Myopia from Colour Fundus Photographs using Code-free Deep Learning. Sci Set J of Ophthalmology & Eye Care, 3(3), 01-05.*

Abstract

Introduction: *The maculopathy observed in highly myopic eyes is intricate. Clinical diagnosis of this condition poses a significant workload and is subjective, particularly in Asian countries such as China. To streamline and expedite the classification of pathologic myopia (PM), a code-free deep learning (CFDL) model was created utilizing a novel platform, specifically the Huawei Cloud, which is accessible to regions with restricted availability of other commercial platforms. Subsequently, the performance of the CFDL model was assessed against a deep learning (DL) algorithm designed for screening PM lesions using color fundus photographs (CFPs).*

Methods: *This research involved analyzing 12,000 CFPs obtained from 1200 individuals in the PALM dataset. Out of these images, 10,800 were utilized for training purposes, while the remaining 1200 were allocated for validation. A CFDL algorithm was developed, validated, and tested on the Huawei Cloud platform to screen for PM in CFPs. Additionally, a custom DL model was trained using transfer learning and the EfficientNet-B8 architecture on the training dataset. The study compared the performance metrics, including the area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, and accuracy, of the CFDL model with those of the bespoke DL model.*

Results: *In the validation dataset, the CFDL model exhibited AUROCs ranging from 0.95 to 0.98 for detecting PM and normal fundus, sensitivities between 92.90% and 93.10%, specificities from 93% to 97.60%, and an overall accuracy of 92.90%. Conversely, the custom traditional DL model achieved AUROCs of 0.94 to 0.96, sensitivities ranging from 91.2% to 92%, specificities between 92.5% and 94%, and an overall accuracy of 91.8%.*

Conclusion: *We created a CFDL model on the Huawei cloud platform to detect and screen for PM using CFPs. Our model demonstrated superior sensitivities, specificities, and consistent accuracies when compared to conventional DL methods.*

Keywords: Artificial Intelligence, Code-free Deep Learning, Pathological Myopia, Colour Fundus Photograph

Introduction

The prevalence of myopia is increasing globally, posing a significant public health challenge. It is estimated that approximately 2 billion people have myopia (≥–0.5 D) currently, with a projected increase to 4.76 billion (49.8% of the world population) by 2050 [1]. High myopia (≥–6.0 D) is also on the rise, particularly in Asian countries, starting at a younger age [2, 3]. Pathologic myopia (PM) is a severe form of myopia associated with excessive axial elongation and structural changes in the posterior segment of the eye (e.g. myopic macular degeneration (MMD) and posterior staphyloma), leading to visual impairment [3]. It affects around 3% of the global population, resulting in po-

tential productivity loss and a significant economic burden on healthcare systems [4]. Timely identification and prevention of PM progression are crucial. Nevertheless, there is a significant shortage of ophthalmologists to care for high-risk populations, particularly in less developed nations [5, 6]. The use of artificial intelligence (AI) for automated detection of PM holds promise in addressing this challenge.

AI systems, particularly deep learning (DL) algorithms, have shown high accuracy, sensitivity, and specificity in ophthalmology imaging for various conditions, including myopia [7]. AI applications in myopia encompass diagnosis, classification of PM, prediction of progression, and guidance for refractive surgery [7]. The use of imaging modalities has evolved from fundus images to optical coherence tomography (OCT) images [7]. Moreover, a novel type of deep learning, referred to as code-free deep learning (CFDL), has emerged to empower healthcare providers with limited coding expertise to create their own AI algorithms for screening purposes [8]. The introduction of a CFDL construction platform by Huawei Cloud, in addition to existing platforms like Google, Apple, and Microsoft Azure, offers a code-free approach for users to develop AI algorithms [9].

Despite the diverse range of AI models for myopia screening developed through various methods and platforms, there is a lack of comprehensive evaluation on the performance of AI models created using the Huawei platform. This evaluation is crucial, especially for Asian countries without unrestricted access to other commercial platforms like Google, as it presents an opportunity to explore the potential of the Huawei platform. Assessing the performance of AI models derived from the Huawei platform can provide valuable insights to ophthalmologists worldwide, particularly in Asian countries, seeking accessible and effective AI solutions to meet the growing demand for timely myopia screenings.

This paper emphasizes the potential of the PM screening model created with the Huawei Cloud CFDL platform and evaluates its effectiveness against a custom DL model trained on the same dataset.

Methods

Data Acquisition

This research utilized color fundus photographs (CFP) from the publicly accessible PALM dataset, which was made available as part of the Pathological Myopia challenge. Approval for data usage was granted by the dataset's data manager on the corresponding webpage. We adhered to the guidelines outlined by the data manager on the webpage and obtained informed consent from all participants.

The PALM dataset comprises retinal images obtained retrospectively from a myopic examination cohort at the Zhongshan Ophthalmic Center (ZOC), Sun Yat-sen University, China [10]. Each colour fundus photograph (CFP) was taken with the optic disc (OD) and macula as the centre in either a single field of view or a dual field of view. The image retrieval protocol was approved by the ZOC ethics board. CFPs were included based on criteria such as single or dual field of view acquisition and absence of significant quality issues that could impact image clarity. Images were excluded if they displayed treatment marks, exposure abnormalities, refractive interstitial opacities, contaminations, or lacked origin information. The final dataset consisted of CFPs from the left eyes of 720 subjects, with 1-3 high-quality images retained per eye, totaling 1200 CFPs. These images were captured simultaneously during the same examination. Among the subjects, 48.1% were male, with an average age of 37.5±15.91 years, and all subjects were of Chinese ethnicity. The dataset comprised 1047 images taken with a Zeiss Visucam 500 camera (2124×2056 pixels resolution) and 153 images taken with a Canon CR-2 camera (1444×1444 pixels resolution). Figure 1 and 2 show examples of the normal CFP and PM CFP used for model training.

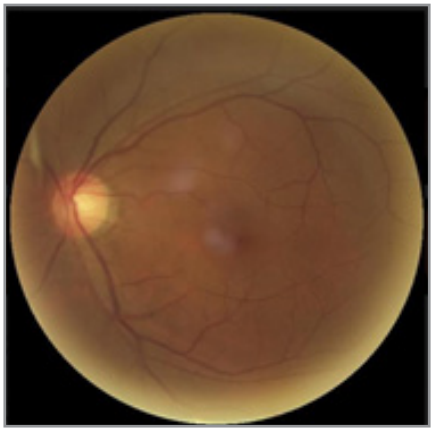

Figure 1: CFP example of a normal fundus

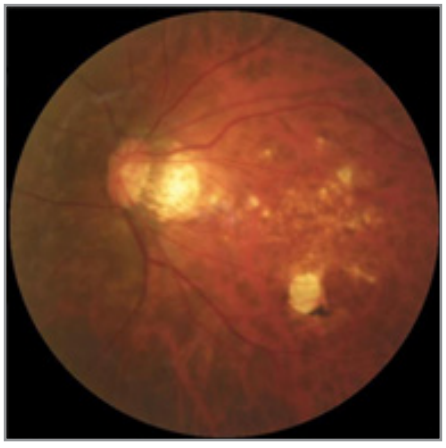

Figure 2: CFP example of a PM fundus

The manual delineation of the optic disc, fundus lesions, and fovea localization involved seven ophthalmologists from the ZOC staff, with an average of 8 years of experience in the field (ranging from 5 to 10 years), and one senior ophthalmologist with over 10 years of experience. Each ophthalmologist independently annotated the structures without access to patient information or knowledge of disease prevalence in the data. The criteria outlined by the International Myopia Institute 3 were followed to identify subjects with PM based on structural changes in the posterior segment of the eye associated with excessive

axial elongation in myopia, such as posterior staphyloma, myopic maculopathy, and high myopia-associated glaucoma-like optic neuropathy. These changes were identified through clinical examination using various imaging modalities, including OCT, fluorescein angiography (FA), and OCT angiography (OCTA). A total of 1200 CFPs with visible macula from 720 patients were collected and used for model development.

Data Augmentation

In order to enhance the diversity of the dataset and mitigate the risk of overfitting, various image augmentation techniques were employed on the training data set [11]. These techniques included horizontal and vertical flipping, rotation up to 60°, brightness adjustment within the range of 0.8–1.2, and contrast modification within the range of 0.9–1.1. This augmentation process expanded the dataset to ten times its original size, resulting in a final dataset comprising 12,000 CFPs with a balanced distribution of PM and non-PM classes at a 1:1 ratio. Subsequently, the dataset was divided into a training set (90% of the images, i.e.,10,800) and a validation set (10% of the images, i.e., 1200) for model validation purposes.

CFDL Model Training and Development on Huawei Cloud

We organized the 12,000 CFPs into separate sub-directories based on their class names. Images labeled as 'normal' (N) were placed in a sub-directory named 'N', while those labeled as 'PM' (Pathological Myopia) were placed in a directory named 'PM'. These images were then uploaded to the 'Object Storage Service' within the cloud storage module. The dataset was created using this uploaded image repository. The AutoML training process commenced by initiating the ExeML project, which automatically progressed through stages such as 'dataset labeling', 'dataset version release', 'data validation', 'image classification', 'model registration', and 'service deployment'. Subsequently, an AI model was generated, and its performance metrics, including the area under the receiver operator curve (AUROC), precision, and recall for the overall classifier, were automatically computed by the platform. Additionally, the platform provided confusion matrices, indicating the frequency (and percentages) of each label predicted for every label in the testing set.

Bespoke Traditional DL Model Training and Development

Our training platform was developed using the PyTorch framework with Python 3.6 and CUDA 10.0. The training setup included a 2.60 GHz Intel(R) CPU and a Tesla V100-SXM2 GPU. We employed the EfficientNet-B8 architecture, a convolutional neural network optimized for processing large input images [12]. The model was initially pretrained on ImageNet weights and fine-tuned with our dataset of 12,000 CFPs [13]. During training, we utilized cross entropy as the objective function, set an initial learning rate of 10^-2, applied a weight decay coefficient of 1e-5 for L2 regularization, and incorporated a dropout rate of 0.5 for the output layer. Stochastic gradient descent (SGD) optimizer was used for 80 epochs on the training data, with validation performed at each epoch to determine the final weights. To prevent overfitting, we implemented early stopping and the sharpness-aware minimization (SAM) optimizer. Training ceased if the validation loss did not improve over 20 consecutive epochs, and the model state with the lowest validation loss was saved as the final model. Figure 3 shows the EfficientNet-B8 architecture used for the development of the deep learning model.

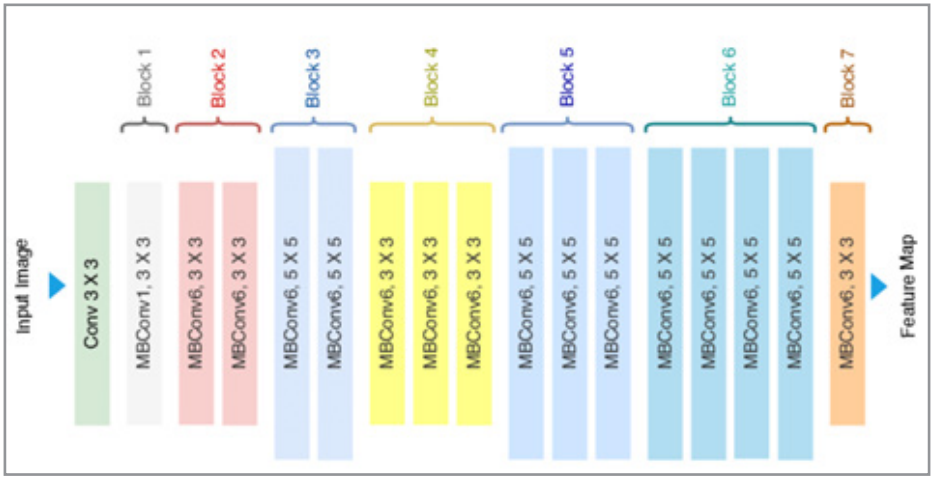

Figure 3: The EfficientNet-B8 architecture

Statistical Analysis

For assessing the model's performance, AUROC curves were employed and analyzed using Python software. The classification model results were used to evaluate the average sensitivity, specificity, and overall accuracy for the two groups. Sensitivity was determined by the formula: true positive / (true positive + false negative), while specificity was calculated using the formula: true negative / (true negative + false positive) [14]. Additionally, per-class accuracy was computed using the formula: (true positive + true negative) / (true positive + false negative + true negative + false positive) [14].

Results

Classification Performance of the Huawei CFDL Model in the Validation Data Set

In the validation set, the CFDL model accurately distinguished between normal and PM cases with an AUROC of 0.98, a sensitivity of 93.10%, and a specificity of 97.60%. Furthermore, the

model effectively differentiated PM fundus from normal fundus with an AUROC of 0.95, a sensitivity of 92.90%, and a specificity of 93%. The overall accuracy of the model was 92.90%.

Classification Performance of the Traditional DL Model in the Validation Data Set

The conventional deep learning model, trained on the identical dataset, distinguished between normal fundus and PM with an AUROC of 0.96, a sensitivity of 92.00%, and a specificity of 94.90%. When differentiating PM from normal fundus, the model achieved an AUROC of 0.94, a sensitivity of 91.2%, and a specificity of 92.5%. The overall accuracy of the DL system was 91.80%.

Discussion

A DL algorithm capable of accurately screening and evaluating PM has the potential to offer significant advantages by improving the accessibility and affordability of PM screening for a large population at risk. This can lead to enhanced access to care and substantial cost reductions, particularly in remote and underserved communities. The importance of remote medical systems has been further emphasized by the COVID-19 outbreak in 2019 [15, 16]. An AI-integrated telemedicine platform is expected to play a crucial role in the future of myopia healthcare post-COVID-19 [17]. In this research, we developed both a codeless model using the Huawei Cloud platform and an efficient deep learning model based on EfficientNet-B8 using 12,000 CFPs. The study demonstrated the potential of these models in PM screening. Both AI models exhibited excellent performance in distinguishing between normal fundus and PM, with the Huawei CFDL model showing superior overall performance in terms of per-class AUROC, sensitivity, specificity, and accuracy compared to the custom deep learning model.

In addition to the effective performance of the Huawei CFDL model in identifying PM, the Huawei Cloud platform offers various advantages for developing AI-based PM screening models. Firstly, conducting CFDL pilot studies on the Huawei platform is cost-effective in terms of time and money compared to platforms like Google, allowing researchers to assess the feasibility of solving research questions using AI and DL techniques. CFDL empowers clinicians to conduct initial proof-of-concept studies using private or well-labeled data, producing preliminary results without the need for collaboration with computational experts and optimizing resource allocation. For instance, CFDL can help determine the utility and practicality of specific sample sizes or image quality levels for developing PM screening models, offering a structured approach instead of inefficient trial and error methods that consume expertise, time, and resources. Researchers can leverage promising interim results from CFDL to engage external technical collaborators, enhancing the chances of the PM screening project success. Furthermore, positive pilot study outcomes from CFDL can bolster applications for research funding.

Furthermore, the Huawei platform itself provides unique user-friendly benefits. These include automated image quality analysis and filtering, as well as a detailed examination of how the model's performance correlates with various characteristics of the color fundus photographs (e.g., image resolution, brightness, etc.). These functionalities allow ophthalmologists to consistently assess the model development datasets and training procedures, facilitating adjustments to enhance model performance. Our research demonstrates that CFDL could serve as a promising alternative to traditional DL for developing AI screening models for pathological myopia, particularly in resource-constrained regions. The Huawei platform shows promise as a valuable tool for deriving CFDL models in such contexts.

While our study has highlighted the benefits of utilizing a larger training dataset through data augmentation to enhance the robustness of an AI model, it is important to note that the CFPs used were sourced from a public database. Real-world datasets offer greater complexity and original disease information compared to public databases. Additionally, the size of the training and testing datasets used in the development process was relatively small. Although the metadata indicated that ophthalmologists or senior retinal specialists conducted the image labeling, details regarding the annotation consensus procedure or the qualifications of the labelers were not provided. External validation of the models was not performed, a crucial step in ensuring the reproducibility of the developed models upon deployment.

Moreover, the application of CFDL in ophthalmological image analysis and screening necessitates adherence to rigorous academic standards, including benchmark standardization, reproducibility assurance in analysis and outcome interpretability, and compliance with research and reporting guidelines such as those outlined for Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence (DECIDE-AI) [18]. The technical constraints of the CFDL platform may pose challenges in meeting these standards, but despite this, valuable results have been and will continue to be generated through CFDL technology.

Conclusion

Considering the exceptional performance achieved by our CFDL algorithm, it is possible for ophthalmologists lacking coding expertise to employ CFDL for PM screenings using CFPs. The distinct advantages offered by CFDL and the Huawei platform could enable cost-effective PM mass screenings in resource-constrained community settings and marginalised communities. This could democratize AI for healthcare professionals, particularly in regions where access to other commercial CFDL platforms like Google is limited, such as China.

Reference

1. Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., & al., e. (2016). Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. Ophthalmology, 123, 1036-1042.
2. Sankaridurg, P., Tahhan, N., Kandel, H., Naduvilath, T., Zou, H., & al., e. (2021). IMI Impact of Myopia. Invest Ophthalmol Vis Sci, 62, 2.
3. Flitcroft, D. I., He, M., Jonas, J. B., Jong, M., Naidoo, K., & al., e. (2019). IMI - Defining and Classifying Myopia: A Proposed Set of Standards for Clinical and Epidemiologic Studies. Invest Ophthalmol Vis Sci, 60, 20-30.
4. Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., & al., e. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ Digit Med, 4, 65.

5. Vela, C., Samson, E., Zunzunegui, M. V., Haddad, S., Aubin, M.-J., & al., e. (2012). Eye care utilization by older adults in low, middle, and high-income countries. BMC Ophthalmol, 12, 5.
6. Resnikoff, S., Lansingh, V. C., Washburn, L., Felch, W., Gauthier, T.-M., & al., e. (2020). Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs?. Br J Ophthalmol, 104, 588-592.
7. Du, R., & Ohno-Matsui, K. (2022). Novel Uses and Challenges of Artificial Intelligence in Diagnosing and Managing Eyes with High Myopia and Pathologic Myopia. Diagnostics (Basel), 12, 1210.
8. Wong, C. Y. T., O'Byrne, C., Taribagil, P., Liu, T., Antaki, F., & al., e. (2024). Comparing code free and bespoke deep learning approaches in ophthalmology. Graefes Arch Clin Exp Ophthalmol, 01-14.
9. Korot, E., Guan, Z., Ferraz, D., Wagner, S. K., Zhang, G., & al., e. (2021). Code-free deep learning for multi-modality medical image classification. Nature Machine Intelligence, 3, 288-298.
10. Fang, H., Li, F., Wu, J., Fu, H., Sun, X., & al., e. (2024). Open Fundus Photograph Dataset with Pathologic Myopia Recognition and Anatomical Structure Annotation. Sci Data, 11, 99.
11. Li, Z., Guo, C., Nie, D., Lin, D., Zhu, Y., & al., e. (2020). Development and Evaluation of a Deep Learning System for Screening Retinal Hemorrhage Based on Ultra-Wide-field Fundus Images. Transl Vis Sci Technol, 9, 3.
12. Tan, M., & Le, Q. V. (2019). Efficient Net: Rethinking Model Scaling for Convolutional Neural Networks. 1905, 119-146.
13. Zhang, K., Liu, X., Xu, J., Yuan, J., Cai, W., & al., e. (2021). Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. Nat Biomed Eng, 5, 533-545.
14. Trevethan, R. (2020). Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. Front Public Health, 7, 408.
15. Wong, C. W., Tsai, A., Jonas, J. B., Ohno-Matsui, K., Chen, J., & al., e. (2021). Digital Screen Time During the COVID-19 Pandemic: Risk for a Further Myopia Boom?. Am J Ophthalmol, 223, 333-337.
16. Wang, J., Li, Y., Musch, D. C., Wei, N., Qi, X., & al., e. (2021). Progression of Myopia in School-Aged Children After COVID-19 Home Confinement. JAMA Ophthalmol, 139, 293-300.
17. Zhang, C., Zhao, J., Zhu, Z., Li, Y., Li, K., & al., e. (2022). Applications of Artificial Intelligence in Myopia: Current and Future Directions. Front Med, 9, 840498.
18. Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., & al., e. (2022). Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med, 28, 924-933.