

Comment on “Retinal Disease Diagnosis Using Deep Learning on Ultra-Wide-Field Fundus Images”

Dr. Minnan Al-Khafaji, BMBS*

Barts Health NHS Trust, UK

Citation: Al-Khafaji M. Comment on “Retinal Disease Diagnosis Using Deep Learning on Ultra-Wide-Field Fundus Images”. *Medi Clin Case Rep J* 2026;4(1):1569-1570. DOI: doi.org/10.51219/MCCRJ/Minnan-Al-Khafaji/432

Received: 08 January, 2026; **Accepted:** 20 January, 2026; **Published:** 22 January, 2026

***Corresponding author:** Dr Minnan Al-Khafaji, BMBS, Barts Health NHS Trust, UK

Copyright: © 2026 Al-Khafaji M., This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

To the Editor,

We read with great interest the article by Nguyen et al., “Retinal Disease Diagnosis Using Deep Learning on Ultra-Wide-Field Fundus Images,” published in *Diagnostics* in 2024. The authors present a deep-learning (DL) model trained on 4,697 ultra-wide-field (UWF) fundus images and report an impressive test AUC of 96.47%. This work represents an important step toward automated detection of retinal abnormalities using UWF imaging. However, we would like to respectfully highlight several methodological issues regarding image quality and preprocessing, which may have implications for the interpretability and generalisability of the reported results.

UWF imaging is inherently susceptible to variability illumination, peripheral distortion, media opacity effects, motion artefacts and non-uniform focus. DL systems trained on such images are known to be highly sensitive to these quality parameters, sometimes learning quality associated features rather than true pathological features. Although Nguyen et al. describe a preprocessing pipeline involving brightness/contrast “enhancement,” cropping, horizontal flipping and resizing, the paper does not provide information on the distribution of image-quality characteristics in the dataset, nor whether quality metrics were balanced between normal and abnormal images. This omission leaves open the possibility that the models reported performance may be influenced by quality-related cues.

Several factors contribute to this risk. There is no objective image-quality assessment (IQA) reported. Metrics such as sharpness, illumination uniformity, motion blur or peripheral artefact burden were not quantified or compared across groups. No exclusion or stratification based on image quality appears to have been performed. If low-quality images are more prevalent in eyes with pathology, this may unintentionally aid the classifier. Pre-processing steps may affect normal and abnormal images differently. Changes in brightness/contrast or resizing can alter the visibility of peripheral lesions or artefacts in ways that may bias the classifier without reflecting true diagnostic capability. There is no sensitivity analysis presented to demonstrate that performance remains stable when controlling for quality variation.

Given the increasing reliance on automated tools for retinal disease screening, it is important to establish that DL systems trained on UWF images identify pathology itself, not indirectly correlated quality features. We suggest that future iterations of this work incorporate: objective quantitative image-quality analysis; reporting of quality distributions across groups; quality matched training and testing sets and evaluation of a dedicated image quality assessment module prior to classification.

Nguyen et al. have contributed meaningfully to DL development in UWF imaging and we offer these suggestions in the spirit of strengthening the evidence for real world deployment. Ensuring that performance metrics are robust to image quality variability is essential for safe and effective clinical translation.