

UWFound, A domain-adapted foundation model for hierarchical posterior uveitis diagnosis from ultra-widefield fundus images

Ruiping Gu

EYE & ENT Hospital of Fudan University

Jiajie Zhao

EYE & ENT Hospital of Fudan University

Lin Wang

EYE & ENT Hospital of Fudan University

Ming Hu

Monash University

Lie Ju

University College London

Boya Lei

EYE & ENT Hospital of Fudan University

Xingtao Zhou

EYE & ENT Hospital of Fudan University

Jiahao Wang

Tianjin Medical University

Mengjun Fu

EYE & ENT Hospital of Fudan University

Zesheng Liu

EYE & ENT Hospital of Fudan University

Zongyuan Ge

Monash University

Imran Razzak

Mohamed bin Zayed University of Artificial Intelligence

Xiaoying Wang

EYE & ENT Hospital of Fudan University

Yang Shen

EYE & ENT Hospital of Fudan University

Min Zhou

`dr_minzhou@163.com`

Article

Keywords:

Posted Date: July 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10143846/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Accurate diagnosis and aetiological classification of posterior uveitis remains challenging because of heterogeneous manifestations, overlapping phenotypes, and severe class imbalance, yet existing artificial intelligence models are limited by modality mismatch and poor generalisability to rare subtypes. Here, we developed UWFound by domain-adaptive self-supervised fine-tuning of RETFound-DINOv2 using 21 234 unlabelled UWF images. UWFound was integrated into a hierarchical framework reflecting clinical diagnostic reasoning from uveitis detection to infectious versus non-infectious classification and aetiological subtyping. Validated on an internal cohort (617 images from 373 individuals), the model achieved near-perfect uveitis screening (AUROC 0.9999) and strong infectious differentiation (mean AUROC 0.9496). Furthermore, external validation across three independent datasets—comprising normal and uveitis images strictly unseen during training—demonstrated robust preserved screening performance (AUROC 0.7638–0.9767). Notably, UWFound consistently outperformed conventional and generic retinal foundation models, particularly in medium- and few-shot scenarios, establishing a robust and generalisable tool for automated posterior uveitis screening and complex aetiological identification.

Introduction

Uveitis includes a group of vision-threatening inflammatory disorders affecting the uveal layer and mostly involving the posterior segment of the eye. The global incidence of uveitis is estimated at approximately 50 cases per 100 000 person-years, and the disease accounts for 10–15% of blindness in developed countries and up to 25% in developing regions.^{1,2} Posterior uveitis is particularly challenging because inflammatory lesions often involve the peripheral retina and choroid, where subtle or atypical findings may be missed on conventional fundus photography. Delayed or incorrect aetiological diagnosis can lead to inappropriate treatment, preventable disease progression, and irreversible visual loss.^{1,3,4}

Aetiological classification is central to clinical management for posterior uveitis. Infectious and non-infectious uveitis require fundamentally different therapeutic strategies, and inappropriate corticosteroid treatment in infectious disease can worsen ocular and systemic outcomes.^{3,5} However, posterior uveitis encompasses diverse entities with overlapping imaging features, variable media opacity, and highly imbalanced disease frequencies, making automated diagnosis difficult.

Ultra-widefield (UWF) fundus imaging provides panoramic visualisation of the posterior and peripheral retina, enabling improved assessment of inflammatory lesions that extend beyond the central fundus.⁶ Compared with conventional fundus photography, UWF imaging can capture broader lesion distributions, peripheral vasculitis, necrotic lesions, and vitreoretinal inflammatory changes.^{7–9} Despite these advantages, most artificial intelligence studies in ophthalmology have focused on common retinal diseases or conventional fundus photographs, and automated UWF-based diagnosis of posterior uveitis remains underexplored.^{10–14}

Foundation models offer a potential solution by learning transferable visual representations from large-scale data and adapting them to downstream tasks with limited labelled samples.¹⁵ However, existing retinal foundation models are predominantly trained on conventional colour fundus photographs, which differ from UWF scanning laser ophthalmoscopy in imaging mechanism, field of view, colour characteristics, and peripheral lesion appearance. This modality gap may limit direct transfer to UWF-based inflammatory disease diagnosis. Moreover, flat classification models do not reflect the hierarchical nature of uveitis diagnosis, in which clinicians first identify ocular inflammation, then distinguish infectious from non-infectious disease, and finally consider specific aetiologies.¹⁵

To address these challenges, we developed UWFound, a UWF-adapted foundation model obtained by domain-adaptive self-supervised fine-tuning on large-scale unlabelled UWF images. We integrated UWFound into a hierarchical diagnostic framework designed to emulate clinical reasoning from posterior uveitis detection to infectious versus non-infectious classification and aetiological subtyping. We hypothesised that UWF-specific foundation-model adaptation combined with hierarchical inference would improve clinically aligned posterior uveitis diagnosis, particularly under class imbalance and limited labelled data.

Results

The internal labelled dataset comprised 617 UWF images from 373 individuals, including 378 images from 265 patients with posterior uveitis and 239 images from 108 healthy controls (Fig. 2). The dataset was divided at the patient level into training, validation, and internal test sets. The final internal test set included 181 UWF images, comprising 62 normal images and 119 posterior uveitis images. Clinical and subtype characteristics of the internal cohort are summarised in the appendix Table 2 (appendix p10).

We first compared the performance of proposed UWFound models with conventional convolutional and transformer-based models and generic retinal foundation-model baselines across high-, medium-, and few-support categories (Table 1, appendix p11-12). The high-support category comprised normal eyes; medium-support categories comprised Behçet's disease, tuberculosis-related uveitis, ocular sarcoidosis, and acute retinal necrosis; and few-support categories comprised endogenous bacterial endophthalmitis, endogenous fungal endophthalmitis, endophthalmitis following cataract surgery, and traumatic endophthalmitis. In the high-support category, all foundation-model approaches achieved high discrimination, with AUROC values close to 1·0. In medium-support categories, UWFound achieved the highest AUROC and AUPRC among the compared models, with an AUROC of 0·8577 and an AUPRC of 0·4681. In few-support categories, overall sensitivity and F1 scores were low across all models, reflecting the difficulty of rare subtype recognition. Nevertheless, the UWFound-based hierarchical model achieved the highest AUPRC in the few-support setting, suggesting an improved precision–recall trade-off under severe class imbalance.

Table 1

Performance comparison of different models for ultra-widefield posterior uveitis etiological classification across many-, medium-, and few-shot categories

Model	High-support AUROC / AUPRC	Medium-support AUROC / AUPRC	Few-support AUROC / AUPRC	Few-support sensitivity	Few-support F1
ResNet50	1.0000 / 1.0000	0.8003 / 0.4050	0.8178 / 0.1040	0.9573	0.1458
ViT	0.9965 / 0.9943	0.8033 / 0.4608	0.8305 / 0.1760	0.9901	0.115
VisionFM	0.9974 / 0.9959	0.8295 / 0.4777	0.8362 / 0.2012	0.9844	0.2964
RetFound	0.9908 / 0.9846	0.7425 / 0.3243	0.6745 / 0.1594	0.9887	0.1786
RetFound-DINOv2	1.0000 / 1.0000	0.7682 / 0.3236	0.7183 / 0.0562	0.9690	0.0788
UWF Found	1.0000 / 1.0000	0.8577 / 0.8577	0.8081 / 0.2294	0.9815	0.1779
UWF Found-Hierarchy	0.9999 / 0.9997	0.8206 / 0.4583	0.7537 / 0.3439	0.9759	0.1762
High-support category included normal eyes. Medium-support categories included Behçet's disease, tuberculosis-related uveitis, ocular sarcoidosis, and acute retinal necrosis. Few-support categories included endogenous bacterial endophthalmitis, endogenous fungal endophthalmitis, endophthalmitis following cataract surgery, and traumatic endophthalmitis.					
Bold values indicate the best performance among all compared models within each category, and underlined values indicate the second-best performance.					

We then evaluated the diagnostic performance in the internal test set (appendix p5-6, 13). The UWFound-based hierarchical model achieved strong performance for first-level posterior uveitis detection, which reached an AUROC of 0.9999 and an AUPRC of 0.9998, with mean sensitivity and specificity of 0.9958. The Brier score was 0.0050, indicating good probabilistic calibration for this binary screening task. At the second diagnostic level, the model maintained strong discrimination for infectious versus non-infectious uveitis. Non-infectious uveitis achieved an AUROC of 0.9391 and an AUPRC of 0.8971, whereas infectious uveitis achieved an AUROC of 0.9098 and an AUPRC of 0.7853. Mean sensitivity and specificity across second-level classes were 0.8500 and 0.9377, respectively. These results indicate that the hierarchical model can support clinically relevant identification between infectious and non-infectious posterior uveitis. At the third level, aetiology-specific performance varied substantially across disease subtypes. AUROC values remained high for acute retinal necrosis and endogenous bacterial endophthalmitis, reaching 0.9232 and 0.9718, respectively. However, mean sensitivity and mean F1 score across third-level classes were 0.4039 and 0.3690, respectively, with particularly limited sensitivity

in several few-support categories. Specificity remained high across most subtypes, with a mean specificity of 0.9465. These findings suggest that the model provides supportive fine-grained aetiological information, while rare-subtype classification remains challenging.

To evaluate the generalisability of the UWFound-based hierarchical model, we performed external validation using three independent UWF datasets: WEI-UWF, ZPPH-UWF, and the clinical reference dataset (appendix p14). For first-level normal versus uveitis classification, the model preserved discriminative ability across all external datasets, with AUROC values of 0.9767, 0.7638, and 0.7975 respectively. Sensitivity and specificity varied across datasets when the validation-derived operating threshold was applied without recalibration. In WEI-UWF, uveitis sensitivity was 0.5338 with specificity of 1.0000, whereas in ZPPH-UWF, uveitis sensitivity was 0.9000 with specificity of 0.3300. These threshold-dependent differences indicate substantial external domain shift and highlight the need for calibration and prospective threshold selection before clinical deployment. Second- and third-level external validation were evaluated in the clinical reference dataset. At the second level, infectious uveitis achieved an AUROC of 0.8899 and an AUPRC of 0.7382, whereas non-infectious uveitis showed lower discrimination, with an AUROC of 0.5293 and an AUPRC of 0.1149. When external results were grouped by data availability, the model achieved higher performance in the high-support category than in medium- and few-support categories. Few-support categories had low sensitivity and AUPRC, confirming that rare-subtype recognition remains vulnerable to class imbalance and external domain shift.

Decision curve analysis showed positive net benefit for first-level uveitis detection across a broad range of threshold probabilities and variable net benefit for second- and third-level tasks (appendix p7-8). Norm-Layer class activation maps showed a progressive shift from broad retinal attention at the screening level to more localised lesion-associated activation at finer diagnostic levels. Representative maps highlighted clinically plausible regions, including vascular and peripheral inflammatory lesions in non-infectious uveitis and focal peripheral necrotic or inflammatory regions in infectious entities (Fig. 3). These visualisations support the interpretability of the hierarchical framework, while remaining qualitative rather than definitive evidence of causal model reasoning.

Discussion

In this multicentre retrospective diagnostic study, we developed UWFound, a UWF-adapted foundation model for hierarchical posterior uveitis diagnosis. Three findings are clinically relevant. First, domain-adaptive self-supervised fine-tuning improved UWF representation learning compared with conventional models and generic retinal foundation-model baselines, particularly in medium- and few-support categories. Second, the hierarchical framework achieved strong performance for posterior uveitis detection and infectious versus non-infectious classification. These two diagnostic steps are closely aligned with clinical decision-making. Third, fine-grained aetiological identification remained difficult for rare subtypes, especially under external domain shift, indicating that UWFound is best interpreted as a triage and decision-support system rather than a standalone diagnostic tool.

The clinical management of posterior uveitis depends on stepwise diagnostic reasoning. Ophthalmologists normally first determine whether ocular inflammation is present, then assess whether infection is likely, and finally integrate ocular signs with systemic history, laboratory testing, microbiological evidence, and multimodal imaging to infer the most likely aetiology.²⁰ The proposed hierarchy was designed to reflect this workflow. The near-perfect internal performance for normal versus uveitis classification suggests potential utility for UWF-based screening and referral support. Accurate differentiation between infectious and non-infectious uveitis is also critical because treatment strategies are fundamentally different, and the strong discrimination of UWFound could provide meaning clinical instruction.^{1,21}

UWFound extends previous ophthalmic foundation-model work by addressing the modality gap between conventional colour fundus photography and UWF scanning laser ophthalmoscopy. UWF images differ in field of view, illumination mechanism, colour distribution, peripheral distortion, and lesion appearance. These differences are especially relevant for posterior uveitis, in which peripheral vasculitis, necrotic lesions, and vitreoretinal inflammatory changes often carry diagnostic information.⁷⁻⁹ By adapting a retinal foundation model using large-scale unlabelled UWF images, UWFound learned modality-specific representations while retaining general retinal priors. The performance gains in medium- and few-support categories support the value of UWF-specific domain adaptation for specialised inflammatory disease tasks.

The results also show the limitations of image-only fine-grained aetiological classification. Although AUROC values were high for some rare but clinically important entities, third-level sensitivity and F1 scores were modest overall. This pattern is expected in posterior uveitis, where subtypes may share overlapping fundus appearances and where some diagnoses depend on systemic manifestations, laboratory evidence, microbiology, or longitudinal treatment response.²⁰ Therefore, the model output should not be used to replace specialist assessment. Its most realistic role is to provide structured diagnostic support: identifying eyes likely to have posterior uveitis, prioritising infectious versus non-infectious pathways, and suggesting possible aetiologies that warrant targeted clinical investigation.

There have been several AI approaches for grading inflammatory activity, such as anterior chamber cell counts, vitreous signal intensity, and choriocapillaris flow deficits based on OCT.²²⁻²⁴ However, these methods are primarily limited to severity assessment and cannot reliably distinguish uveitis aetiologies. Similarly, recent attempts using large language models achieved limited diagnostic accuracy for uveitis classification.²⁵ In contrast, the UWFound framework directly leverages visual retinal patterns captured by UWF imaging, enabling both detection and aetiological differentiation, which is essential for clinical management.

External validation demonstrated preserved discrimination for first-level uveitis detection but also revealed substantial threshold-dependent variation in sensitivity and specificity across datasets. This finding is important because it reflects real-world differences in case mix, image acquisition, device settings, disease severity, and annotation standards. The observed increase in Brier scores and reduced

rare-subtype performance under external validation indicate that calibration and threshold selection should be prospectively assessed before deployment. Future clinical use should therefore include site-specific monitoring, calibration assessment, and human-in-the-loop review, especially for rare disease subtypes.

Interpretability analysis provided biologically plausible visual evidence for the hierarchical framework. Norm-Layer class activation maps showed a transition from broad retinal attention during screening to more localised lesion-associated activation during aetiological refinement. Representative examples highlighted clinically relevant regions, including vascular and peripheral inflammatory lesions, peripheral necrotic areas, and diffuse inflammatory changes in endophthalmitis. These findings support the plausibility of the model's visual reasoning. However, activation maps remain qualitative and should not be interpreted as definitive causal explanations. In cases with severe media opacity or vitreous haze, the model may attend to visibility-related features rather than specific retinal lesions.

This study has several strengths. It used a clinically motivated hierarchy, domain-adaptive foundation-model training, patient-level internal splitting, independent external validation, and a comprehensive evaluation framework including discrimination, precision–recall performance, calibration, decision curve analysis, and visual interpretability. The inclusion of rare posterior uveitis subtypes also reflects the clinical complexity encountered in tertiary referral settings.

Several limitations should be acknowledged. First, despite multicentre data collection, the number of samples in several rare aetiological categories remained small, limiting sensitivity and confidence interval precision. Second, this was a retrospective study and may be affected by referral bias and spectrum bias, thus may not fully represent real-world deployment conditions. Third, the model used UWF images alone, whereas real-world aetiological diagnosis of posterior uveitis depends on clinical history, slit-lamp findings, OCT, OCTA, fluorescein or indocyanine green angiography, laboratory testing, microbiology, and systemic evaluation.²⁶ Fourth, other common ocular diseases that cause vitreous opacities, such as haemorrhagic vitreous opacities and vitreous degeneration, were not included in the model training, it is necessary to rule out these diseases before clinical application, which can be easily differentiated based on the patient's medical history and ophthalmic manifestations. Finally, prospective validation is needed to assess workflow integration, threshold calibration, reader interaction, and patient-level clinical benefit.

In conclusion, UWFound provides a UWF-adapted foundation-model framework for hierarchical posterior uveitis diagnosis. The model achieved strong performance for posterior uveitis screening and infectious versus non-infectious classification, while providing supportive but imperfect aetiological subtype information for rare diseases. These findings support further development of UWFound as an interpretable triage and decision-support tool for complex inflammatory retinal disease. Future work should focus on prospective multicentre validation, multimodal integration with clinical and laboratory data, and calibration strategies for safe real-world deployment.

Methods

Study design and ethics

This retrospective multicentre diagnostic study consisted of three sequential phases: domain-adaptive self-supervised fine-tuning of a retinal foundation model using large-scale unlabelled UWF images; supervised development and internal testing of a hierarchical posterior uveitis diagnostic framework using a labelled cohort; and external validation using three independent UWF datasets. The study was conducted in accordance with the Declaration of Helsinki. The study was approved by the Ethics Committee of the Eye and Ear, Nose, and Throat Hospital (EENT) of Fudan University, Shanghai, China (approval numbers: 2024126 and 2024-07-225). The requirement for written informed consent was waived by the ethics committee because all data were retrospective and anonymised.

Data sources and study population

For domain-adaptive foundation-model fine-tuning, 21 234 unlabelled UWF images were collected from routine clinical examinations and publicly available UWF datasets (appendix p1, 9). These images represented a broad range of UWF appearances, imaging devices, and retinal conditions, and were used only for self-supervised adaptation without diagnostic labels.

The internal labelled dataset comprised 617 UWF images from 373 individuals at the EENT Hospital, including 378 images from 265 patients with posterior uveitis and 239 images from 108 healthy controls (appendix p1). Consecutive patients with posterior uveitis were retrospectively identified over a 5-year period. Individuals with uncertain diagnoses, absence of UWF imaging, or inadequate image quality were excluded. Healthy controls underwent UWF imaging during the same period and had no evidence of ocular inflammation. The labelled dataset was randomly divided at the patient level into training, validation, and internal test sets to minimise information leakage between model development and evaluation.

Posterior uveitis diagnoses were established by experienced uveitis specialists (RG and MZ) on comprehensive clinical evaluation of clinical history, ophthalmic examination, laboratory testing, microbiological evidence when available, and multimodal imaging findings according to established international diagnostic criteria (appendix p1). Non-infectious posterior uveitis included Behçet's disease, tuberculosis-related uveitis, and ocular sarcoidosis. Infectious posterior uveitis included acute retinal necrosis, endogenous bacterial endophthalmitis, endogenous fungal endophthalmitis, post-traumatic bacterial endophthalmitis, and endophthalmitis after cataract surgery.

External validation was performed using three independent UWF datasets: the WEI-UWF dataset from Weifang Eye Institute, the ZPPH-UWF dataset from Zhejiang Provincial People's Hospital, and a clinical reference dataset constructed from peer-reviewed publications and publicly available normal UWF images (appendix p1-2).

UWFound development

Conventional task-specific training relies on limited labelled datasets and is often constrained by poor generalisability, particularly for rare subtypes (Fig. 1A). UWFound was therefore developed by domain-adaptive self-supervised fine-tuning of RETFound-DINOv2 on 21 234 unlabelled UWF images (Fig. 1B). Because UWF scanning laser ophthalmoscopy is acquired primarily through red and green laser channels, only the informative red and green channels were retained, and images were resized to 784 × 784 pixels to preserve peripheral retinal and choroidal lesion morphology. Domain adaptation was designed to reduce the modality gap between conventional colour fundus photographs used for generic retinal foundation-model pretraining and UWF scanning laser ophthalmoscopy.

Hierarchical diagnostic framework

Posterior uveitis diagnosis was formulated as a hierarchical classification task reflecting clinical reasoning (Fig. 1C). The first level distinguished normal eyes from posterior uveitis. The second level classified uveitis into infectious and non-infectious categories. The third level predicted specific aetiological subtypes. UWFound was used as the image encoder, and hierarchy-specific prediction heads generated diagnostic probabilities at each level.

To support coarse-to-fine inference, the prediction from each hierarchy level was transformed into a class embedding and used to condition the subsequent refinement stage through a class-conditioned attention module. This module combined channel modulation and spatial attention to emphasise features relevant to the next diagnostic decision while preserving global retinal context. To improve robustness under class imbalance and intra-class heterogeneity, each hierarchy level used a prototype-based metric classifier with learnable class prototypes and momentum-updated prototype memory. Full architectural details are provided in the appendix (appendix p2-4).

Model training and comparison

The hierarchical model was trained using cross-entropy losses summed across hierarchy levels. Teacher forcing was applied during training to reduce error accumulation between hierarchy levels, whereas inference was fully automated and used only model-predicted probabilities from preceding levels. Class-prior logit adjustment was used to mitigate class imbalance. Model selection was based on validation-set performance before final evaluation on the internal test set and external validation datasets.

We compared UWFound and the UWFound-based hierarchical model with conventional and foundation-model baselines, including ResNet50, ViT, VisionFM, RETFound, and RETFound-DINOv2.^{16–18} All models were trained and evaluated using the same patient-level data splits. For long-tailed evaluation, classes were grouped by data availability. The high-support category comprised normal eyes. Medium-support categories comprised Behçet's disease, tuberculosis-related uveitis, ocular sarcoidosis, and acute retinal necrosis. Few-support categories comprised endogenous bacterial endophthalmitis, endogenous fungal

endophthalmitis, endophthalmitis following cataract surgery, and traumatic endophthalmitis. Training hyperparameters, augmentation strategy, and implementation details are provided in the appendix (appendix p2-4).

Outcomes and statistical analysis

The primary outcome was diagnostic performance across the predefined hierarchy: normal versus posterior uveitis, infectious versus non-infectious uveitis, and aetiological subtype classification. Model performance was assessed using sensitivity, specificity, F1 score, area under the receiver operating characteristic curve (AUROC), area under the precision–recall curve (AUPRC), and Brier score. Because of substantial class imbalance, AUPRC and class-wise sensitivity were emphasised for medium- and few-shot categories. Ninety-five percent confidence intervals were estimated using non-parametric bootstrap resampling with 1000 iterations.

Receiver operating characteristic and precision–recall curves were plotted for each diagnostic level. Decision curve analysis was used to assess net clinical benefit across threshold probabilities compared with treat-all and treat-none strategies. For threshold-dependent metrics, operating thresholds were selected using the validation set and then applied without modification to the internal test and external validation datasets. Model interpretability was assessed using Norm-Layer class activation mapping to visualise retinal regions contributing to predictions at different hierarchy levels.¹⁹

Declarations

Data and code availability

The UWF image data from EENT dataset are available upon reasonable request from the corresponding author but will not be publicly available because they contain sensitive information that could compromise patient privacy. The code used in the current study for developing the UWFound is available upon reasonable request from the corresponding author.

Acknowledgements

The study was supported by Project of Shanghai Shenkang Hospital Development Centre (SHDC12024148 to XWang), National Natural Science Foundation of China (Grant No.82171095 to XWang, Grant No.82101183 to YShen), Shanghai Xuhui District Health Commission (XHLHGG202104 to XZhou), Project of Shanghai Science and Technology (Grant No.23XD1400500 to XWang).

Contribution

RG, LW, XZ, XW, YS and MZ designed the study. RG, BL, JW, MF and MZ collected the data. RG, JZ, BL, ZL and YS verified the raw underlying data and made data annotation. LW, MH and LJ developed the UWFound model. RG, JZ and LW had full access to all the data and analysed the results. RG, JZ, LW and

YS co-wrote the manuscript. XZ, ZG, IR, XW and MZ critically revised the manuscript. All authors approved the final manuscript. All authors had final responsibility for the decision to submit for publication.

Declaration of interests

The authors declare no competing interests.

References

1. Maghsoudlou P, Epps SJ, Guly CM, Dick AD. Uveitis in Adults: A Review. *Jama* 2025; 334(5): 419–34.
2. Tomkins-Netzer O, Talat L, Bar A, et al. Long-term clinical outcome and causes of vision loss in patients with uveitis. *Ophthalmology* 2014; 121(12): 2387–92.
3. Paez-Escamilla M, Caplash S, Kalra G, et al. Challenges in posterior uveitis-tips and tricks for the retina specialist. *J Ophthalmic Inflamm Infect* 2023; 13(1): 35.
4. Papaliodis GN, Rosner BA, Dreger KA, et al. Incidence of and Risk Factors for Cataract in Anterior Uveitis. *Am J Ophthalmol* 2023; 254: 221–32.
5. Sève P, Cacoub P, Bodaghi B, et al. Uveitis: Diagnostic work-up. A literature review and recommendations from an expert committee. *Autoimmun Rev* 2017; 16(12): 1254–64.
6. Gupta V, Al-Dhibi HA, Arevalo JF. Retinal imaging in uveitis. *Saudi J Ophthalmol* 2014; 28(2): 95–103.
7. Campbell JP, Leder HA, Sepah YJ, et al. Wide-field retinal imaging in the management of noninfectious posterior uveitis. *Am J Ophthalmol* 2012; 154(5): 908–11.e2.
8. Leder HA, Campbell JP, Sepah YJ, et al. Ultra-wide-field retinal imaging in the management of non-infectious retinal vasculitis. *J Ophthalmic Inflamm Infect* 2013; 3(1): 30.
9. Kumar V, Surve A, Kumawat D, et al. Ultra-wide field retinal imaging: A wider clinical perspective. *Indian J Ophthalmol* 2021; 69(4): 824–35.
10. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. *Prog Retin Eye Res* 2018; 67: 1–29.
11. Ipp E, Liljenquist D, Bode B, et al. Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy. *JAMA Netw Open* 2021; 4(11): e2134254.
12. Danieleescu C, Dabija MG, Nedelcu AH, et al. Automated Retinal Vessel Analysis Based on Fundus Photographs as a Predictor for Non-Ophthalmic Diseases-Evolution and Perspectives. *J Pers Med* 2023; 14(1).
13. Zhao X, Gu X, Teng D, et al. Ultra-wide-field fundus photography and AI-based screening and referral for multiple ocular fundus diseases. *Cell Rep Med* 2025; 6(6): 102187.
14. Abellanas M, Elena MJ, Keane PA, Balaskas K, Grewal DS, Carreño E. Artificial Intelligence and Imaging Processing in Optical Coherence Tomography and Digital Images in Uveitis. *Ocul Immunol Inflamm* 2022; 30(3): 675–81.

15. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. *Nature* 2023; 616(7956): 259–65.
16. Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. *Nature* 2023; 622(7981): 156–63.
17. Qiu J, Wu J, Wei H, et al. Development and Validation of a Multimodal Multitask Vision Foundation Model for Generalist Ophthalmic Artificial Intelligence. *NEJM AI* 2024; 1(12): A10a2300221.
18. Hou Q, Zhou Y, Lin Goh JH, et al. Can a Natural Image-Based Foundation Model Outperform a Retina-Specific Model in Detecting Ocular and Systemic Diseases? *Ophthalmol Sci* 2026; 6(1): 100923.
19. Jiang PT, Zhang CB, Hou Q, Cheng MM, Wei Y. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Trans Image Process* 2021; 30: 5875–88.
20. Heiligenhaus A, Rothaus K, Pleyer U. [Development of classification criteria for uveitis by the standardization of uveitis nomenclature (SUN) working group]. *Ophthalmologe* 2021; 118(9): 913–8.
21. Agarwal A, Aggarwal K, Gupta V. Infectious uveitis: an Asian perspective. *Eye (Lond)* 2019; 33(1): 50–65.
22. Tailor PD, Xu TT, Fortes BH, et al. Appropriateness of Ophthalmology Recommendations From an Online Chat-Based Artificial Intelligence Model. *Mayo Clin Proc Digit Health* 2024; 2(1): 119–28.
23. Mihalache A, Huang RS, Popovic MM, et al. Accuracy of an Artificial Intelligence Chatbot's Interpretation of Clinical Ophthalmic Images. *JAMA Ophthalmol* 2024; 142(4): 321–6.
24. Mutawa AM, Alzuwawi MA. Multilayered rule-based expert system for diagnosing uveitis. *Artif Intell Med* 2019; 99: 101691.
25. Jamilloux Y, Romain-Scelle N, Rabilloud M, et al. Development and Validation of a Bayesian Network for Supporting the Etiological Diagnosis of Uveitis. *J Clin Med* 2021; 10(15).
26. Gupta V, Davis JL, Gangaputra S, et al. Evidence and Consensus-Based Imaging Guidelines in Noninfectious Posterior Uveitis-Methodology of the Multimodal Imaging in Uveitis (MUV) Project Report 2. *Am J Ophthalmol* 2025; 278: 257–65.

Figures

Figure 1

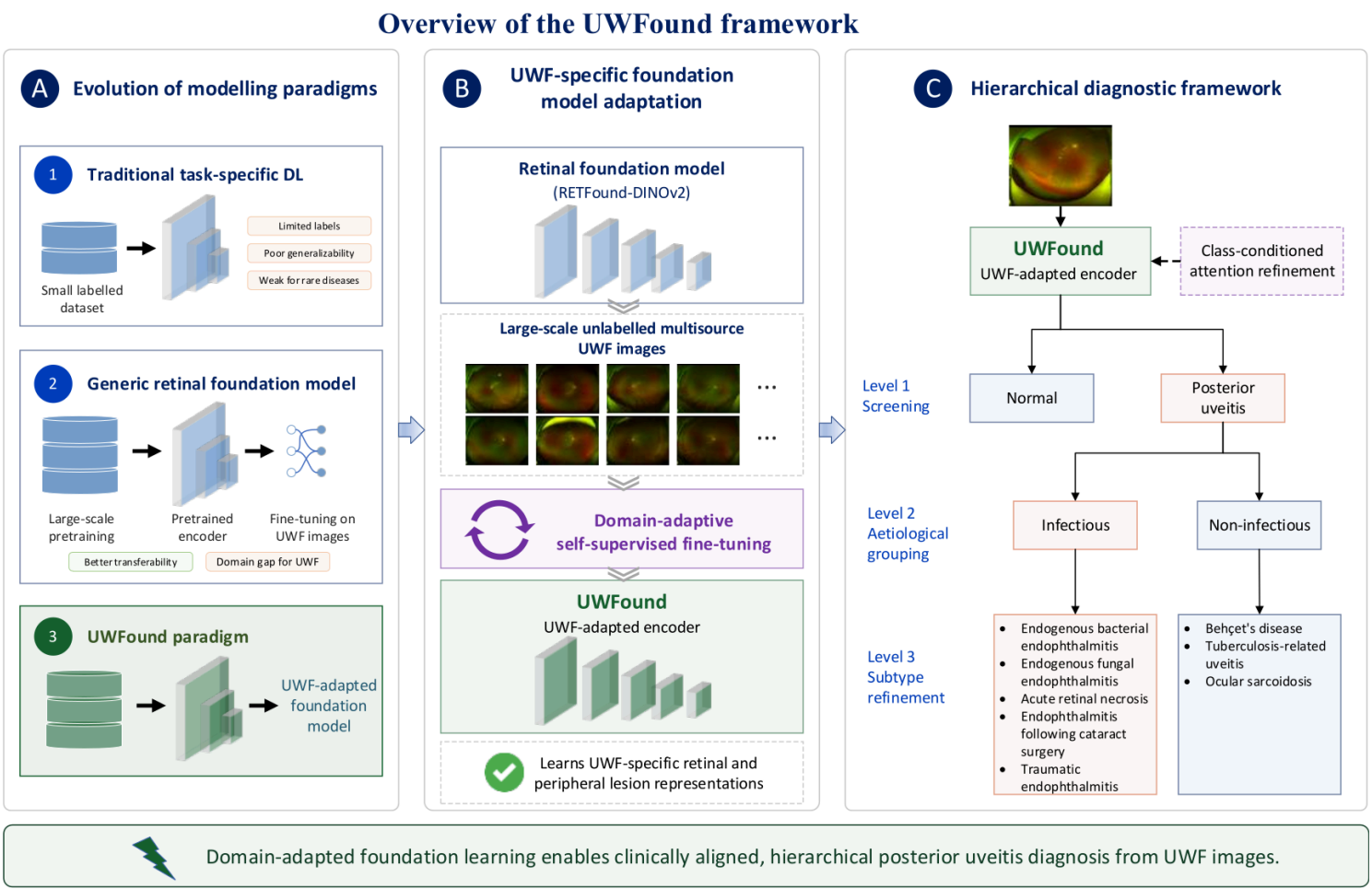


Figure 1

Evolution of modelling paradigms and overview of the UWFound-based hierarchical diagnostic framework

(A) Evolution of modelling paradigms. Traditional task-specific deep learning relies on small labelled datasets and has limited generalisability, particularly for rare posterior uveitis subtypes. Generic retinal foundation models improve transferability but remain affected by the domain gap between conventional colour fundus photography and ultra-widefield (UWF) imaging. The UWFound paradigm adapts foundation representations specifically to UWF images.

(B) UWFound-specific foundation model adaptation. UWFound was developed by domain-adaptive self-supervised fine-tuning of RETFound-DINOv2 using 21 234 unlabelled multisource UWF images, enabling the encoder to learn UWF-specific retinal and peripheral lesion representations.

(C) Hierarchical diagnostic framework. UWF images are encoded by UWFound and analysed through a clinically aligned hierarchy: level 1 distinguishes normal eyes from posterior uveitis; level 2 stratifies posterior uveitis into infectious and non-infectious categories; and level 3 refines predictions to specific

aetiological subtypes. Class-conditioned attention refinement uses predictions from preceding hierarchy levels to guide subsequent feature refinement, supporting coarse-to-fine diagnostic reasoning.

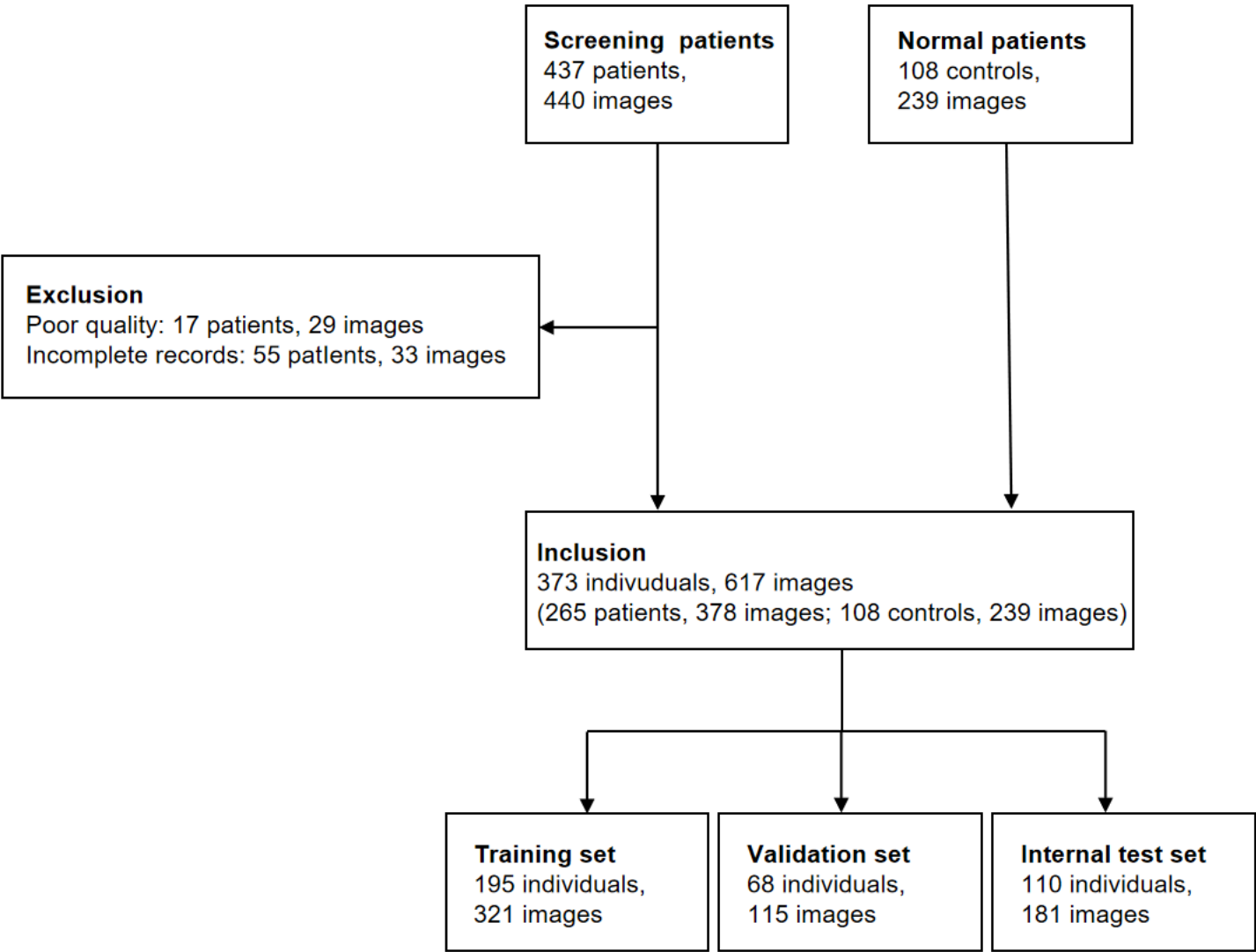


Figure 2

Study cohort selection and internal dataset split.

A total of 337 uveitis patients with 440 ultra-widefield fundus images were screened. After exclusions, 265 patients with 378 images were included. 108 normal controls with 239 images were randomly selected, and the combined dataset were divided at the patient level into training, validation, and internal test sets. UWF=ultra-widefield.

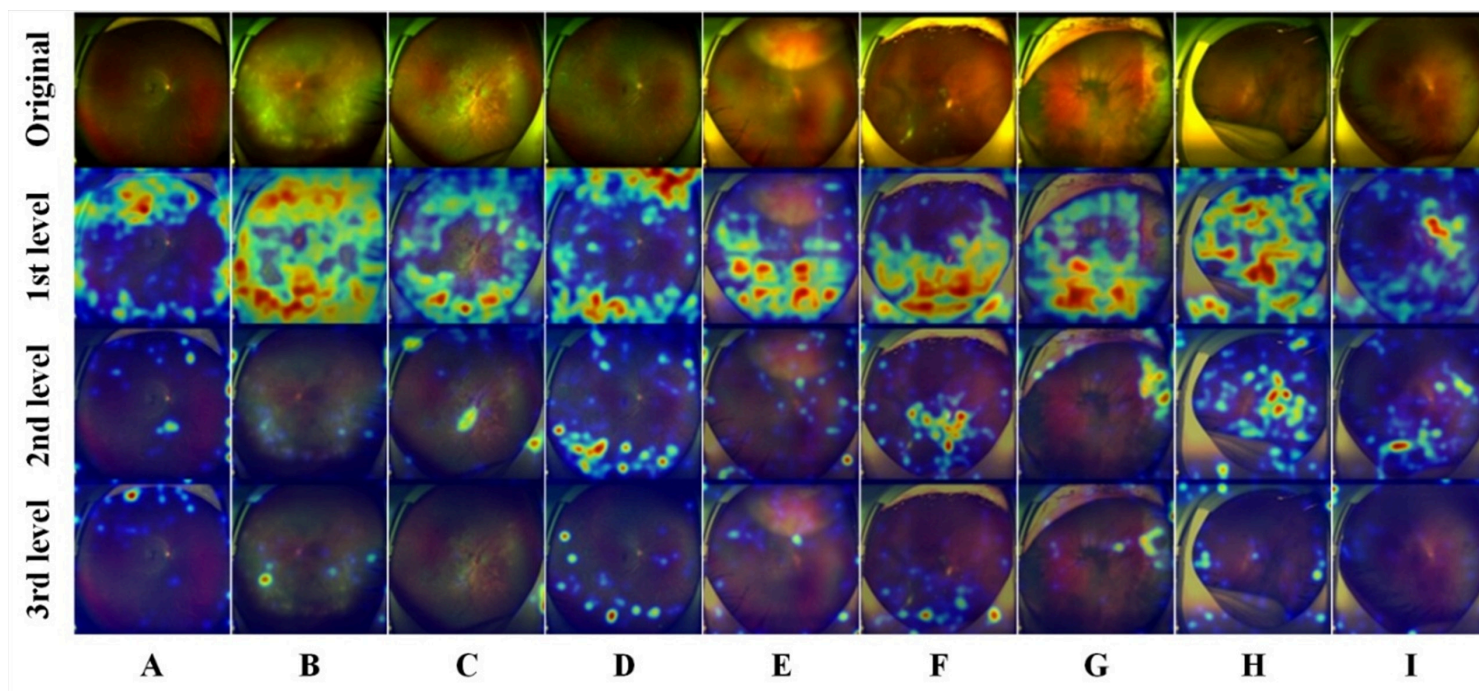


Figure 3

Hierarchical visual interpretability of posterior uveitis classification using Norm-Layer CAM

Norm-Layer class activation maps (CAMs) were generated to visualise discriminative retinal regions across different diagnostic hierarchy levels for representative posterior uveitis subtypes. (A) Normal, (B) Behçet's disease, (C) Tuberculosis-related uveitis, (D) Ocular sarcoidosis, (E) Endogenous bacterial endophthalmitis, (F) Endogenous fungal endophthalmitis, (G) Acute retinal necrosis, (H) Endophthalmitis following cataract surgery, (I) Traumatic endophthalmitis. Across hierarchy levels, attention progressively shifts from global retinal context to more localised, disease-specific lesion patterns, illustrating that the hierarchical framework enables coarse-to-fine visual reasoning consistent with clinical diagnostic practice.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryAppendix.docx](#)