Automated early detection of acute retinal necrosis from ultra-widefield color fundus photography using deep learning

Yuqin Wang[1†], Zijian Yang[1†], Xingneng Guo[1†], Wang Jin[1], Dan Lin[1], Anying Chen[2] and Meng Zhou[1*]

Abstract

Background Acute retinal necrosis (ARN) is a relatively rare but highly damaging and potentially sight-threatening type of uveitis caused by infection with the human herpesvirus. Without timely diagnosis and appropriate treatment, ARN can lead to severe vision loss. We aimed to develop a deep learning framework to distinguish ARN from other types of intermediate, posterior, and panuveitis using ultra-widefield color fundus photography (UWFCFP).

Methods We conducted a two-center retrospective discovery and validation study to develop and validate a deep learning model called DeepDrARN for automatic uveitis detection and differentiation of ARN from other uveitis types using 11,508 UWFCFPs from 1,112 participants. Model performance was evaluated with the area under the receiver operating characteristic curve (AUROC), the area under the precision and recall curves (AUPR), sensitivity and specificity, and compared with seven ophthalmologists.

Results DeepDrARN for uveitis screening achieved an AUROC of 0.996 (95% CI: 0.994–0.999) in the internal validation cohort and demonstrated good generalizability with an AUROC of 0.973 (95% CI: 0.956–0.990) in the external validation cohort. DeepDrARN also demonstrated excellent predictive ability in distinguishing ARN from other types of uveitis with AUROCs of 0.960 (95% CI: 0.943–0.977) and 0.971 (95% CI: 0.956–0.986) in the internal and external validation cohorts. DeepDrARN was also tested in the differentiation of ARN, non-ARN uveitis (NAU) and normal subjects, with sensitivities of 88.9% and 78.7% and specificities of 93.8% and 89.1% in the internal and external validation cohorts, respectively. The performance of DeepDrARN is comparable to that of ophthalmologists and even exceeds the average accuracy of seven ophthalmologists, showing an improvement of 6.57% in uveitis screening and 11.14% in ARN identification.

Conclusions Our study demonstrates the feasibility of deep learning algorithms in enabling early detection, reducing treatment delays, and improving outcomes for ARN patients.

Keywords Acute retinal necrosis, Ultra-widefield fundus photography, Uveitis, Deep learning

†Yuqin Wang, Zijian Yang and Xingneng Guo contributed equally as first authors and should be considered co-first authors.

*Correspondence:
Meng Zhou
zhoumeng@wmu.edu.cn
[1] National Clinical Research Center for Ocular Diseases, Eye Hospital, Wenzhou Medical University, Wenzhou 325027, China
[2] The Affiliated Ningbo Eye Hospital of Wenzhou Medical University, Ningbo 315042, China

Background

Acute retinal necrosis syndrome (ARN) is a relatively rare but highly damaging and potentially sight-threatening type of uveitis caused by human herpesvirus infection [1]. ARN initially presents as acute panuveitis, characterized by inflammation around the retinal arteries, and rapidly progresses to extensive necrotizing retinitis, often leading to rhegmatogenous retinal detachment (RRD) [2]. ARN accounts for a small

Wang *et al. Eye and Vision*   (2024) 11:27

Page 2 of 11

proportion of uveitis cases, ranging from 0.1% to 1.3% [3–8], with an annual incidence rate of approximately 0.5 to 0.63 per million individuals [9–11]. The primary treatment approach for ARN is systemic antiviral therapy, often supplemented by intravitreal antiviral injections, effectively managing the disease [1]. However, a substantial proportion of treated eyes, ranging from 20% to 73%, still develop secondary RRD, which is the leading cause of poor visual outcomes in ARN [1, 12, 13]. Patients diagnosed with ARN who experience an average delay of 5.2 days from symptom onset to treatment are 2.3 times more likely to experience severe visual loss compared to those who receive prompt treatment within one day of symptom onset [14]. Therefore, timely and accurate diagnosis of ARN plays a critical role in ensuring effective clinical intervention and reducing the risk of permanent vision loss.

The diagnostic criteria for ARN were initially established by the Executive Committee of the American Uveitis Society in 1994, focusing on specific clinical manifestations [15]. Subsequent advances in molecular techniques have made polymerase chain reaction (PCR) testing more accessible, demonstrating high sensitivity and specificity for detecting ARN by identifying viral DNA in vitreous and aqueous specimens [10, 12, 16–18]. The Japanese ARN Study Group and the Standardization of Uveitis Nomenclature (SUN) Working Group incorporated virological testing of intraocular fluids into their classification criteria for ARN. However, these test results were not considered essential for diagnosis [19, 20]. Meanwhile, the collection of intraocular fluid is an invasive procedure with potential risks of infection. Furthermore, patients with characteristic clinical features of ARN should receive immediate antiviral treatment without waiting for the results of the PCR test. Thus, early disease detection relies on clinical expertise and subjective assessment, a significant challenge for ophthalmologists, especially in primary care settings.

Recent developments in deep learning have shown promising potential in medical image analysis [21–25]. The unique advantage of deep learning lies in its ability to discern complex and subtle features within images, enabling the identification of minute retinal changes that may escape human observation. Ultra-widefield fundus photography (UWFCFP) has been shown to be more effective than conventional fundus cameras in capturing the peripheral circumferential extension of disease [26]. To address the urgent clinical need for improved early diagnosis of ARN, we propose a deep learning model based on the Swin Transformer architecture to distinguish ARN from other types of intermediate, posterior, and panuveitis using UWFCFP. This model aims to enable computer-assisted early diagnostic tools for ARN,

facilitating more accurate and timely identification of this vision-threatening disease.

Methods
Two-center patient cohorts
This study adhered to the tenets of the Declaration of Helsinki and was approved by the ethics committees of the Eye Hospital of Wenzhou Medical University (2023–025-K-20–01) and Ningbo Eye Hospital (2023–26(K)-C2).

A total of 1,112 subjects and 11,508 corresponding UWFCFPs [580 from normal eyes, 2,884 with ARN, and 8,044 with non-ARN uveitis (NAU)] were included in this two-center retrospective study, conducted between June 2015 to March 2023, at Eye Hospital of Wenzhou Medical University (WMUEH) and Ningbo Eye Hospital (NEH). All ophthalmic diagnoses were made by experienced uveitis and retina specialists. Normal eyes were classified based on the absence of any uveal or vitreoretinal disease, except for mild vitreous opacities or white without pressure, with no history of vitreoretinal surgery, retinal photocoagulation, and exhibiting normal fundus findings. ARN diagnosis adhered to the SUN classification criteria. Non-ARN uveitis refers to other commonly observed conditions such as intermediate, posterior, and panuveitis. Electronic medical records, multimodal imaging data, and laboratory results for each subject were independently reviewed by two ophthalmologists. Disagreements were resolved by a third uveitis specialist. A comprehensive list of disease entities and their inclusion criteria is provided in Additional file 4. Enrollment criteria for subjects required any eye to meet the outlined criteria in one of the three groups.

The photographs used in this study were obtained using a commercially available ultra-widefield (UWF) scanning laser ophthalmoscope (Daytona, Optos PLC, Dunfermline, UK) with a fixed aspect ratio of 256:325. The dataset included multiple images per patient across multiple visits, with several different images taken at different eye positions at each visit. UWFCFPs showing active inflammatory conditions were specifically selected from this dataset for the ARN and NAU cohorts. Active inflammatory conditions were identified by visual indicators such as retinal necrotic lesions, choroidal or chorioretinal lesions, and exudative retinal detachment. Exclusion criteria included the absence of inflammatory conditions and factors hindering fundus lesion observation, such as significant media opacities, intravitreal implants, retinal photocoagulation scars, and poor patient coordination during the examination. Two ophthalmologists independently reviewed each UWFCFP to ensure accurate inclusion. Disagreements were resolved through consultation with a third uveitis specialist. Specifically, 6,384 UWFCFPs were excluded

Wang *et al. Eye and Vision*   (2024) 11:27

Page 3 of 11

based on criteria including severe media opacity (any retinal structure is completely invisible in the image, $n = 109$), presence of retinal photocoagulation scars ($n = 1,831$), quiet inflammatory period ($n = 3,505$) and interference with fundus lesion observation due to vitreous implants used in uveitis treatment ($n = 939$). Eligible photographs were then divided into four sub-cohorts for training and testing. The workflow and details of the UWFCFP collection and cohort division are illustrated in Additional file 1. Dataset volumes for each disease entity are detailed in Additional file 4.

The architecture of deep learning algorithms

The scheme of our proposed hierarchical framework, DeepDrARN, is shown in Fig. 1. DeepDrARN consists of two stages, namely uveitis screening and ARN detection. In the first stage, a deep learning model was trained to discriminate between uveitis and normal. In the second
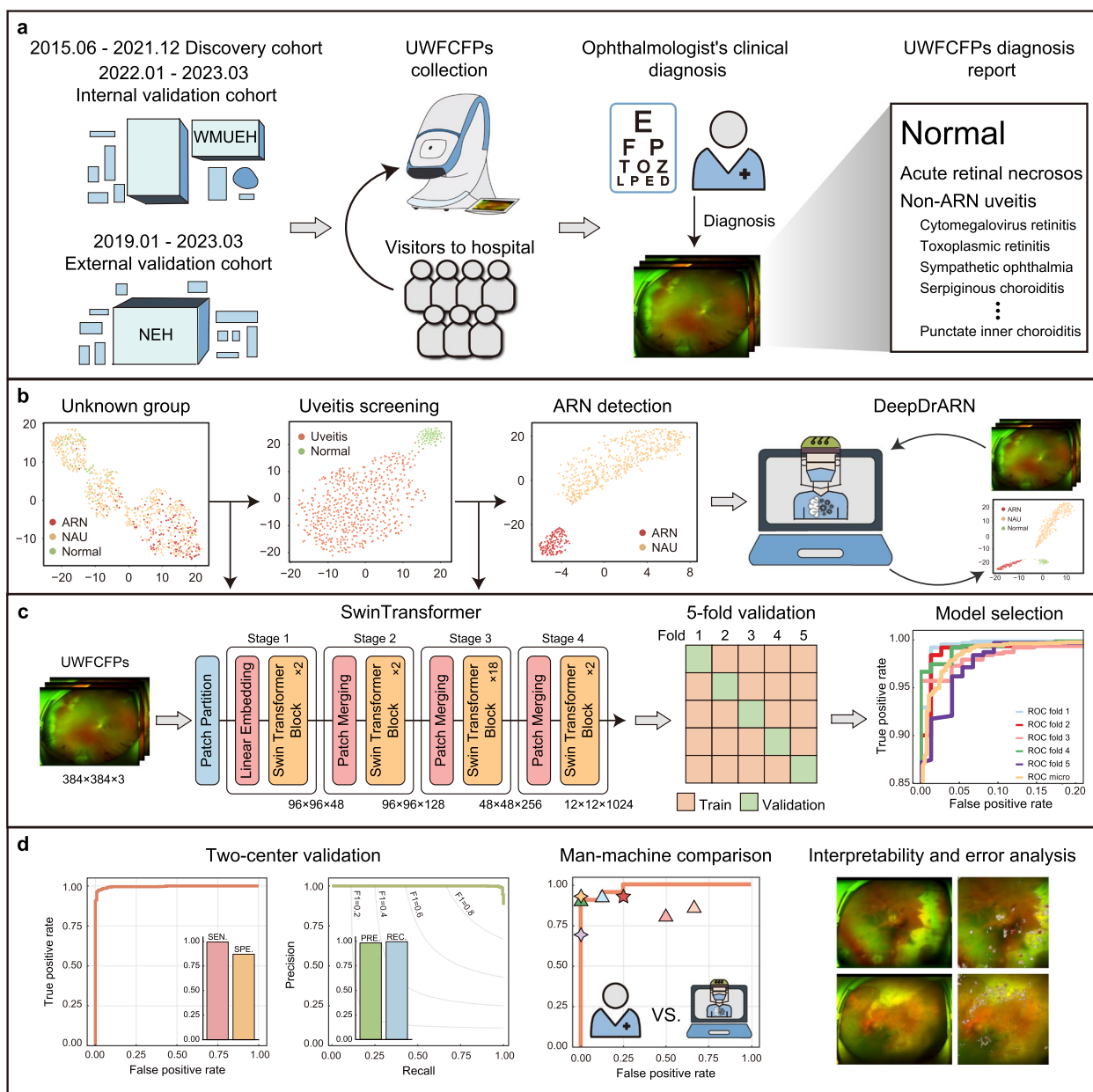
Fig. 1 Schematic workflow of DeepDrARN. **a** Data acquisition from two ophthalmic centers in China. **b** and **c** Schematic diagram and workflow of DeepDrARN for uveitis screening and ARN identification. **d** Multi-perspective evaluation and analysis. UWFCFPs, ultra-widefield color fundus photographs; ARN, acute retinal necrosis; NAU, non-ARN uveitis; SEN, sensitivity; SPE, specificity; PRE, precision; REC, recall

stage, the model was refined for detailed stratification, focusing on the accurate detection of ARN from NAU. We proposed a deep learning model with the Swin Transformer [27] as its backbone, which incorporates the self-attention mechanism, allowing for a comprehensive investigation of features related to ARN phenotypes. We implemented data enhancement techniques to counter color deviation and resolution disparities in UWFCFPs. UWFCFPs were resized and cropped to 384×384 dimensions.

We have implemented random resize cropping, augmentation, and erasing for the training set. Additionally, all RGB channels of the UWFCFPs were standardized and normalized. In the training phase, the cross-entropy loss function was used as the objective function, and the Adam optimizer was used to optimize the model. The deep learning models for uveitis screening and ARN detection were trained with a batch size of 32, a weight decay of 0.05, and learning rates of $1e-5$ and $1.25e-4$, respectively. Transfer learning was used to initialize self-attention-based deep learning architectures with parameters pre-trained on ImageNet. Five-fold cross-validation was used to ensure model robustness in the discovery cohort. Each fold was subjected to 100 training rounds, with the most accurate model saved as the best. The model with the highest accuracy among those saved in the five-fold cross-validation is selected for subsequent internal and external validation.

Model interpretability

The integrated gradient method was used to generate pixel-level saliency maps and visual explanations for the key class-discriminative regions in the UWFCFPs as follows:

$$\text{IG}_i(x) = (x - x\prime) \times \int_{\alpha=0}^{1} \frac{\partial F(x\prime + \alpha \times (x - x\prime))}{\partial x} d\alpha \qquad (1)$$

where F(x) is the deep learning model, IG_i refers to the integrated gradient of pixel i, x is the input UWFCFP and $x\prime$ is the baseline image which is a black image of the same size as the UWFCFPs.

Statistical analysis

Statistical analyses were conducted using R software (v.4.2.2) and Python (v.3.6). Model performance was evaluated by calculating the positive predictive value (PPV), negative predictive value (NPV), accuracy, precision, recall, sensitivity, and specificity with the 'sklearn' package (v.0.24.2) at a threshold of 0.5. The area under the receiver operating characteristic curve (AUROC) and the area under the precision and recall curves (AUPR) were also calculated to assess the model's performance. The 95% confidence intervals (CIs) for the AUROCs and

AUPRs were calculated using the non-parametric bootstrap method with 2,000 resamplings with the 'pROC' package (v.1.18.0). Means and standard deviations (SDs) were used to summarize characteristics for continuous variables and percentages for categorical variables.

Results

Baseline characteristics of subjects and study design

A total of 5,124 UWFCFPs from 908 subjects (mean age of 42.3 ± 15.0 years; 487 men and 421 women) from two medical centers (from June 2015 to March 2023) were used to develop and validate the proposed deep learning model. Additional file 4 shows the detailed inclusion/exclusion criteria for the patients enrolled in this study. The 5,124 UWFCFPs included 580 from normal eyes, 1,000 from ARN, and 3,544 from NAU cases. Patients were divided into four sub-cohorts: (i) Discovery cohort (WMUEH-I cohort) comprising 3,533 UWFCFPs from 587 subjects at WMUEH, collected between June 2015 and December 2021 for model development; (ii) Internal validation cohort (WMUEH-II), which included 978 UWFCFPs from 235 subjects at WMUEH collected from January 2022 to March 2023; (iii) External validation cohort (NEH-I), consisting of 513 UWFCFPs from 159 subjects at NEH, collected between January 2019 and March 2023; (iv) Comparison cohort (NEH-II), consisting of the remaining 100 UWFCFPs from 66 subjects at NEH, was used for model and ophthalmologist diagnostic comparison. Demographic characteristics and clinical information of the sub-cohorts are shown in Additional file 3.

Development and performance of the DeepDrARN

The Swin Transformer, initialized with ImageNet-trained weights, was used as the default backbone for training the DeepDrARN to effectively screen various uveitis conditions and accurately detect ARN from UWFCFPs through a five-fold cross-validation (CV) in the discovery cohort. The workflow of DeepDrARN is illustrated in Fig. 1. First, we evaluated the performance of the DeepDrARN in identifying uveitis conditions from UWFCFPs and demonstrated that it achieved AUROC, AUPRC, PPV, and NPV values of 0.996 ± 0.002, 0.999 ± 0.000, 99.1% and 94.5%, respectively (Fig. 2a and b). DeepDrARN was also evaluated for its ability to discriminate ARN from NAU. The results indicated that DeepDrARN achieved an AUROC of 0.997 ± 0.002, AUPRC of 0.993 ± 0.005, PPV of 99.3%, and NPV of 99.1% (Fig. 2c and d). Overall accuracy, precision, recall, and F1 score were analyzed for each fold (Additional file 2). The performance of DeepDrARN was consistent across data variations. These results confirmed the robustness and effectiveness in diagnosing uveitis and detecting ARN.

Wang *et al. Eye and Vision*   (2024) 11:27

Page 5 of 11
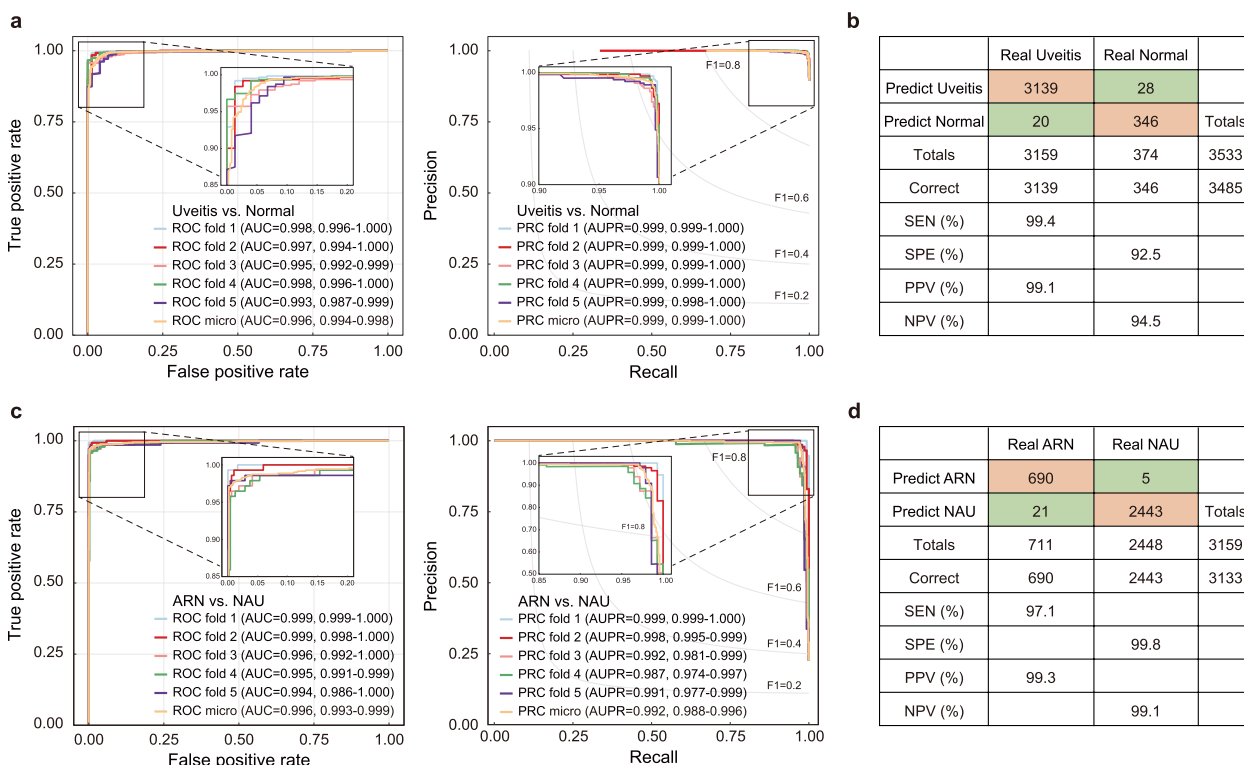
Fig. 2 Performance of DeepDrARN in the discovery cohort. Uveitis detection with five-fold CV, ROC and PRC curves (**a**), and confusion matrix (**b**). ARN detection with five-fold CV, ROC and PRC curves (**c**), and confusion matrix (**d**). CV, cross validation; ROC, receiver operating characteristic curve; PRC, precision and recall curve; ARN, acute retinal necrosis; NAU, non-ARN uveitis; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value

Independent validation of DeepDrARN in two-center cohorts

DeepDrARN was tested in two independent cohorts from different medical centers. For uveitis screening, DeepDrARN showed similar predictive performance in two cohorts, with AUROCs of 0.996 (95% CI: 0.994–0.999) and 0.973 (95% CI: 0.956–0.990), AUPRCs of 0.999 (95% CI: 0.999–1.000) and 0.994 (95% CI: 0.986–0.998), PPVs of 98.3% and 93.8%, and NPVs of 96.2% and 90.0% for WMUEH-II and NEH-I cohorts, respectively (Fig. 3a to d). Furthermore, DeepDrARN also performed well in discriminating ARN from NAU, with AUROCs of 0.960 (95% CI: 0.943–0.997) and 0.971 (95% CI: 0.956–0.986), AUPRCs of 0.902 (95% CI: 0.864–0.934) and 0.923 (95% CI: 0.880–0.957), PPVs of 83.9% and 92.1%, and NPVs of
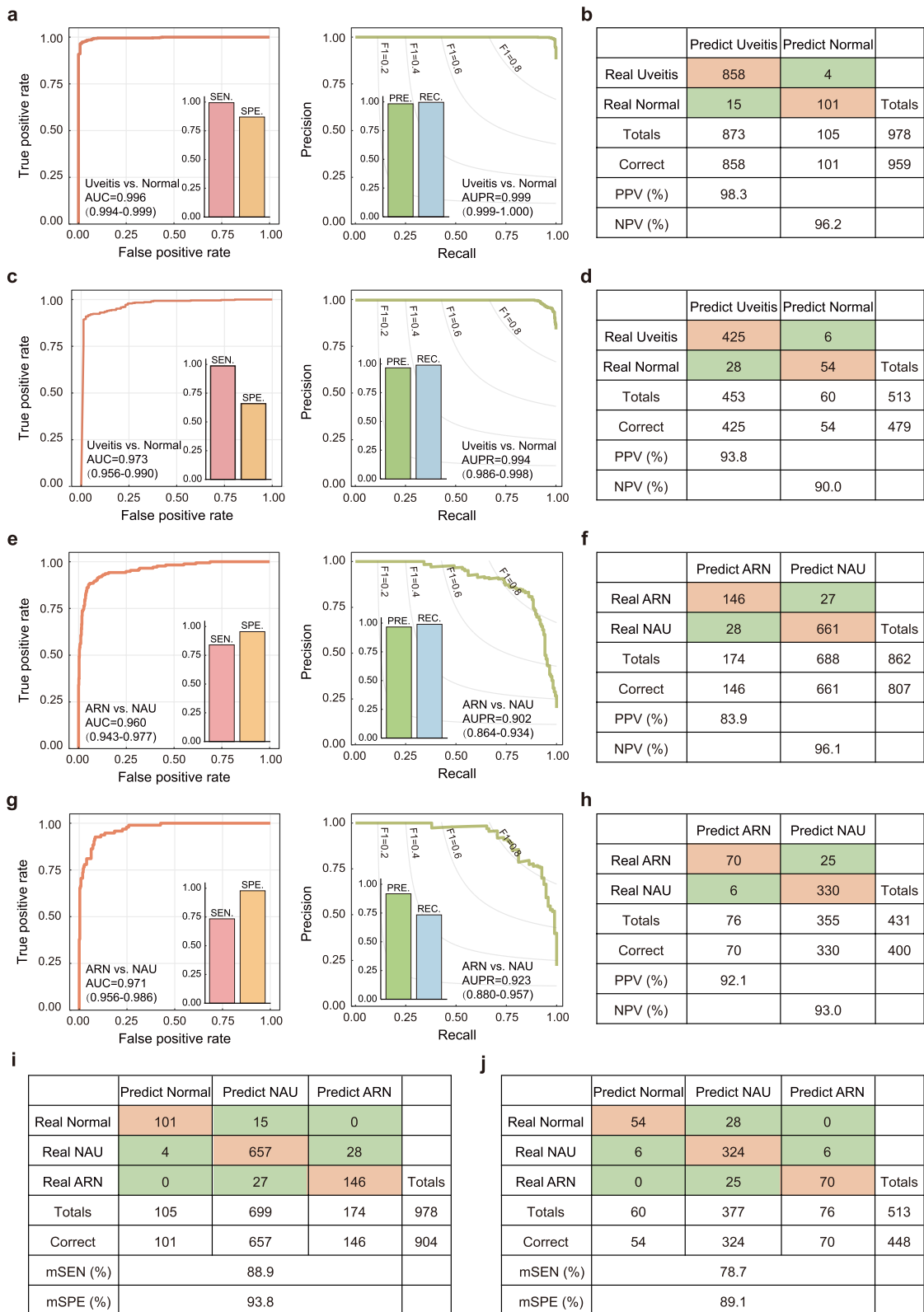
96.1% and 93.0% in the WMUEH-II and NEH-I cohorts, respectively (Fig. 3e to h). In addition, we tested the performance of DeepDrARN in differentiating ARN, NAU, and normal subjects from an unknown population. As shown in Fig. 3, DeepDrARN demonstrated sensitivities of 88.9% and 78.7% and specificities of 93.8% and 89.1% in the WMUEH-II and NEH-I cohorts, respectively.

Performance comparison (DeepDrARN vs. ophthalmologists)

To further validate the diagnostic competence of Deep-DrARN, a comparative analysis was conducted to assess its performance against seven ophthalmologists (four junior, two intermediate, and one senior). This evaluation was carried out on an independent comparison cohort

(See figure on next page.)
Fig. 3 Independent evaluation of DeepDrARN. **a–d** ROC and PRC curves, confusion matrices for uveitis screening. **e–h** ROC and PRC curves, confusion matrices for ARN diagnosis. **i**, **j** Confusion matrices for differentiation of ARN, NAU, and normal subjects. ROC, receiver operating characteristic curve; PRC, precision and recall curve; ARN, acute retinal necrosis; NAU, non-ARN uveitis; AUC, area under the receiver operating characteristic curve; AUPR, area under the precision and recall curve; mSEN, mean sensitivity; mSPE, mean specificity; PPV, positive predictive value; NPV, negative predictive value; PRE, precision; REC, recall

Wang *et al. Eye and Vision*   (2024) 11:27

Page 6 of 11

a

b

	Predict Uveitis	Predict Normal	
Real Uveitis	858	4	
Real Normal	15	101	Totals
Totals	873	105	978
Correct	858	101	959
PPV (%)	98.3		
NPV (%)		96.2	

c

d

	Predict Uveitis	Predict Normal	
Real Uveitis	425	6	
Real Normal	28	54	Totals
Totals	453	60	513
Correct	425	54	479
PPV (%)	93.8		
NPV (%)		90.0	

e

f

	Predict ARN	Predict NAU	
Real ARN	146	27	
Real NAU	28	661	Totals
Totals	174	688	862
Correct	146	661	807
PPV (%)	83.9		
NPV (%)		96.1	

g

h

	Predict ARN	Predict NAU	
Real ARN	70	25	
Real NAU	6	330	Totals
Totals	76	355	431
Correct	70	330	400
PPV (%)	92.1		
NPV (%)		93.0	

i

	Predict Normal	Predict NAU	Predict ARN	
Real Normal	101	15	0	
Real NAU	4	657	28	
Real ARN	0	27	146	Totals
Totals	105	699	174	978
Correct	101	657	146	904
mSEN (%)	88.9			
mSPE (%)	93.8			

j

	Predict Normal	Predict NAU	Predict ARN	
Real Normal	54	28	0	
Real NAU	6	324	6	
Real ARN	0	25	70	Totals
Totals	60	377	76	513
Correct	54	324	70	448
mSEN (%)	78.7			
mSPE (%)	89.1			

Fig. 3 (See legend on previous page.)

Wang *et al. Eye and Vision*   (2024) 11:27

Page 7 of 11

(NEH-II), in which UWFCFPs had not previously been examined by either DeepDrARN or ophthalmologists.

Ophthalmologists independently and anonymously made diagnoses without patient-specific clinical information. Assessments were performed in a quiet environment without time constraints. Results of the comparative analysis are shown in Fig. 4 and Table 1. The performance of DeepDrARN was comparable to that of the ophthalmologists, and even exceeded the average accuracy of seven ophthalmologists, showing an improvement of 6.57% and 11.14% in uveitis screening and ARN identification, respectively. In contrast, considerable variation in
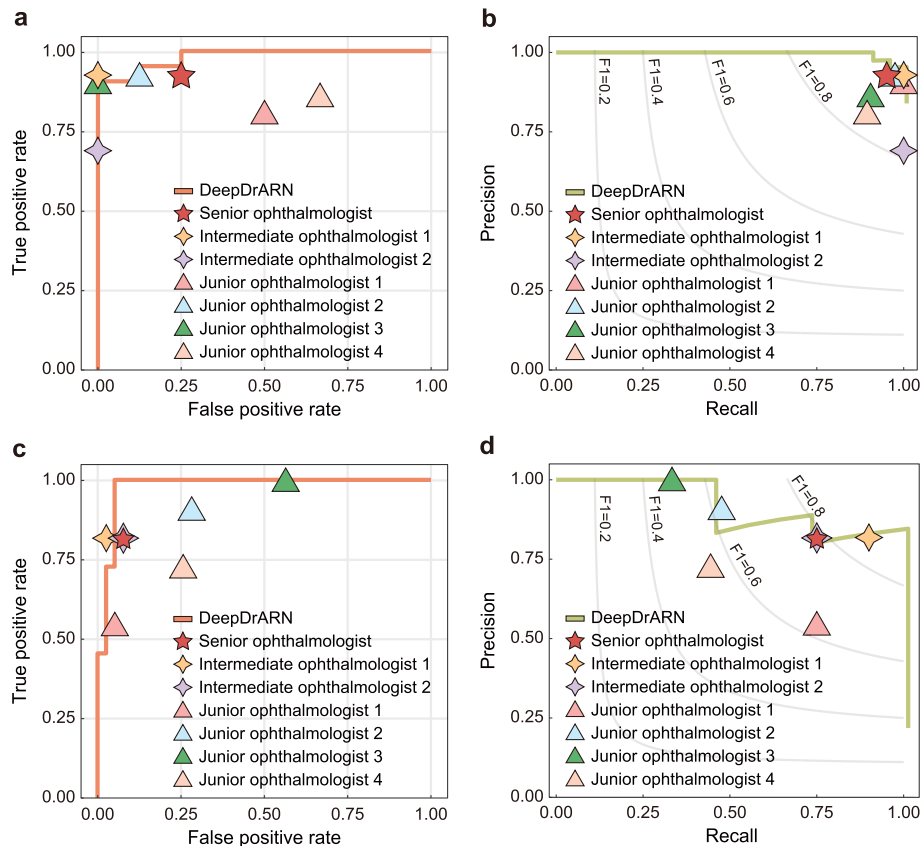
Fig. 4 Comparison between DeepDrARN and human ophthalmologists. ROC curve (**a**) and PRC curve (**b**) for uveitis screening. ROC curve (**c**) and PRC curve (**d**) for ARN identification. ROC, receiver operating characteristic curve; RPC, precision and recall curve; ARN, acute retinal necrosis

Table 1 Comparison between human ophthalmologists and DeepDrARN

Parameter	Uveitis vs. Normal				ARN vs. NAU			
	Accuracy (%)	Precision (%)	Recall (%)	F1 score	Accuracy (%)	Precision (%)	Recall (%)	F1 score
Senior ophthalmologist	90.00	95.12	92.86	93.98	90.00	94.74	92.31	93.51
Intermediate ophthalmologist 1	94.00	100.00	92.86	96.30	94.00	95.00	97.44	96.20
Intermediate ophthalmologist 2	74.00	100.00	69.05	81.69	90.00	94.74	92.31	93.51
Junior ophthalmologist 1	92.00	100.00	90.84	95.00	86.00	88.10	94.87	91.36
Junior ophthalmologist 2	92.00	97.50	92.86	95.12	76.00	96.55	71.79	82.35
Junior ophthalmologist 3	80.00	86.36	90.48	88.37	56.00	100.00	43.59	60.71
Junior ophthalmologist 4	76.00	89.47	80.95	85.00	74.00	90.63	74.36	81.69
Mean ophthalmologists	85.43	95.49	87.13	90.78	80.86	94.25	80.95	85.62
DeepDrARN	92.00	91.30	100.00	95.54	92.00	81.82	81.82	81.82

ARN = acute retinal necrosis; *NAU* = non-ARN uveitis

precision and recall was observed among ophthalmologists, reflecting differences in experience and expertise, with a wide range of accuracy for both uveitis detection and ARN identification.

Interpretability and misdiagnosis analysis of DeepDrARN
The misdiagnosis of DeepDrARN was analyzed using integrated gradients to gain a more comprehensive understanding of DeepDrARN. Figure 5 shows representative cases and their corresponding saliency maps of DeepDrARN. The saliency maps show that Deep-DrARN focuses on specific UWFCFP regions, including the optic disc, retinal blood vessels, and lesion areas (Fig. 5). Specifically, for ARN, DeepDrARN focuses primarily on critical areas such as the optic disc, necrotic lesions, vascular occlusions, and inflammatory vitreous haze (Fig. 5a and c). Notable areas of concern include retinal lesions, vasculitis or sheathing, and inflammatory vitreous haze, particularly in cases of toxoplasma retinochoroiditis (TR) (Fig. 5b top), cytomegalovirus retinitis (CMVR) (Fig. 5b bottom and 5d bottom), and idiopathic retinal vasculitis (IRV) (Fig. 5d top). These

results indicate that DeepDrARN has acquired significant features that match the clinically relevant knowledge of uveitis experts, suggesting that DeepDrARN has developed the ability to prioritize key fundus areas that are critical for uveitis diagnosis. In the WMUEH-II and NEH-I cohorts, the characteristics of the misinterpreted UWFCFPs by DeepDrARN were summarized in Additional file 5. False negatives in uveitis screening were consistently observed in cases with minor or mild lesions, while false positives correlated with mild vitreous opacity (5 photographs, 33.3%) in the WMUEH-II cohort and camera lens reflections (11 photographs, 39.3%) in the NEH-I cohort. Misclassification was particularly evident during the regression phase of retinal necrosis, with 66.7% in WMUEH-II and 64.0% in NEH-I. DeepDrARN tended to misclassify specific uveitis subtypes, such as IRV and CMVR, as ARN.

Discussion
ARN, a potentially devastating ocular disease, has a low incidence rate, leading to its underdiagnosis and misdiagnosis due to limited familiarity among many
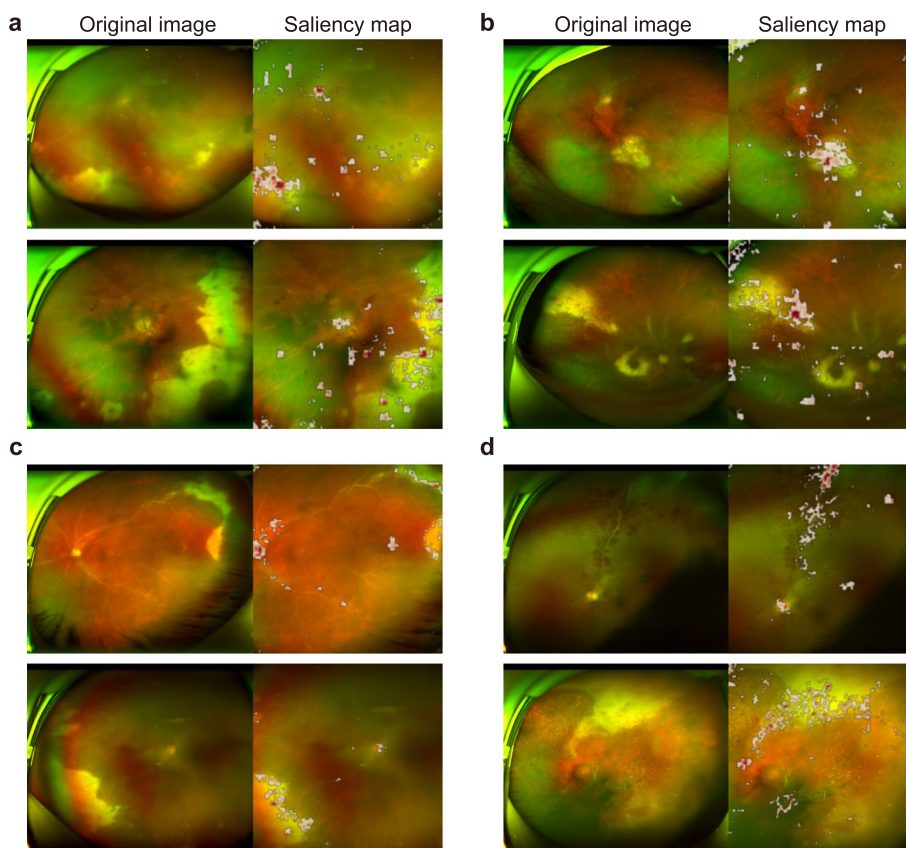
Fig. 5 Visualization of DeepDrARN decision. Original UWFFP and saliency maps for ARN (**a**) and NAU (**b**) in uveitis screening. Original UWFFP and saliency maps for ARN (**c**) and NAU (**d**) for ARN identification. UWFFP, ultra-widefield color fundus photograph; ARN, acute retinal necrosis; NAU, non-ARN uveitis

Wang *et al. Eye and Vision*   (2024) 11:27

Page 9 of 11

ophthalmologists in clinical practice. This study develops and validates a clinical-level deep learning model using UWFCFPs to automatically detect uveitis conditions and further differentiate ARN from other uveitis types. Early detection of ARN is crucial for preventing irreversible vision loss. Traditionally, the diagnosis of ARN relies heavily on clinical expertise, subjective assessment, and sometimes invasive procedures. These limitations highlight the urgent need for objective, non-invasive diagnostic tools, particularly in primary care. Delays caused by traditional diagnostic methods in primary care can lead to significant consequences. Growing evidence shows that integrating deep learning algorithms into clinical practice could revolutionize healthcare by improving disease diagnosis, treatment selection, and clinical laboratory testing [28–33]. Existing deep-learning retinal disease screening is primarily based on fundus images with a limited 45° to 55° field of view [34–36]. Uveitis, particularly ARN, presents unique challenges as early lesions may occur in the peripheral retina. The 200° coverage of UWFCFPs, which does not require pupil dilation, overcomes the limitations of traditional systems, making it ideal for large-scale screening. We designed a hierarchical vision transformer architecture to accurately identify disease-specific discriminative features, including subtle abnormalities in early-stage UWFCFPs.

Several recent studies have developed diagnostic models for various forms of uveitis based on clinical cases [37–40]. However, these models are unsuitable for comprehensive screening as they rely heavily on extensive clinical data for diagnosis. Conversely, existing fundus image screening models for uveitis are limited by single-center focus, small dataset size, and lack of external validation [35, 41]. This two-center retrospective study utilized the largest ARN UWFCFP dataset to date, thereby increasing the reliability and applicability of DeepDrARN. To ensure data representativeness, our study followed current international standards for uveitis diagnosis and classification, distinguishing it from previous studies that relied solely on labeling by ophthalmologists at their respective centers. Our comparative study showed that DeepDrARN matched expert performance and outperformed primary ophthalmologists in ARN detection. The reduced sensitivity and specificity of primary eye care practitioners in detecting ARN reflect their limited familiarity with this rare condition. Saliency maps allowed DeepDrARN to identify critical features, demonstrating its reliability as a diagnostic tool.

Although our study has made some efforts, certain challenges and avenues for future investigation should be acknowledged. First, further validation studies in diverse cohorts and populations are needed. Second, the retrospective nature of the study highlights the need for prospective cohorts to verify the reliability of Deep-DrARN in real-world clinical settings. Furthermore, although the deep learning model has shown promise as a diagnostic aid, it is undeniable that uveitis diagnosis requires a combination of medical history, laboratory findings, and multimodal imaging rather than solely relying on a single imaging modality. Especially in cases of highly atypical or media opacity, the effectiveness of the deep learning model may be limited. Therefore, the integrated decision-making of clinical ophthalmologists and PCR testing remain indispensable and central components of the diagnostic process.

Conclusions

This study introduces DeepDrARN, a deep learning model for automated early detection of ARN using UWFCFPs. The robust performance and non-invasive nature establish DeepDrARN as a valuable screening tool for uveitis and ARN, aiding clinical decision-making, especially for junior ophthalmologists. The potential implementation of DeepDrARN across clinical platforms shows promise in enabling early referrals, reducing treatment delays, and improving outcomes for ARN patients.

Abbreviations

ARN	Acute retinal necrosis
AUROC	Area under the receiver operating characteristic curve
AUPRC	Area under the precision and recall curve
CV	Cross validation
CMVR	Cytomegalovirus retinitis
IG	Integrated gradient
IRV	Idiopathic retinal vasculitis
NAU	Non-ARN uveitis
NEH	Ningbo Eye Hospital
NPV	Negative predictive value
PCR	Polymerase chain reaction
PPV	Positive predictive value
RRD	Rhegmatogenous retinal detachment
SD	Standard deviation
SUN	Standardization of uveitis nomenclature
TR	Toxoplasma retinochoroiditis
UWFCFP	Ultra-widefield color fundus photography
WMUEH	Eye Hospital of Wenzhou Medical University

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40662-024-00396-z.

Additional file 1. Workflow of ultra-widefield color fundus photograph (UWFCFP) collection and cohort division. ARN, acute retinal necrosis; NAU, non-ARN uveitis; WMUEH, Eye Hospital of Wenzhou Medical University; NEH, Ningbo Eye Hospital.

Additional file 2. Overall accuracy, precision, recall, and F1 score for 5-fold cross validation in uveitis screening (**a**) and ARN identification (**b**).

Additional file 3. Demographic characteristics and clinical information of the four sub-cohorts for training and testing of DeepDrARN.

Additional file 4. Inclusion criteria and data volumes of all enrolled disease entities.

Additional file 5. Characteristics of the misinterpreted ultra-widefield color fundus photographs (UWFCFPs) by DeepDrARN.

Wang *et al. Eye and Vision*   (2024) 11:27

Page 10 of 11

Declarations

Ethics approval and consent to participate
This study was approved by the Ethics Committees at the Eye Hospital of Wenzhou Medical University (2023–025-K-20–01) and Ningbo Eye Hospital (2023–26(K)-C2) in accordance with the Declaration of Helsinki. Patients gave informed consent allowing their retinal images to be used.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Schoenberger SD, Kim SJ, Thorne JE, Mruthyunjaya P, Yeh S, Bakri SJ, et al. Diagnosis and treatment of acute retinal necrosis: a report by the American Academy of Ophthalmology. Ophthalmology. 2017;124(3):382–92.
2. Uruyama A, Yamada N, Sasaki T. Unilateral acute uveitis with periarteritis and detachment. Jpn J Clin Ophthalmol. 1971;25:607–19.
3. Henderly DE, Genstler AJ, Smith RE, Rao NA. Changing patterns of uveitis. Am J Ophthalmol. 1987;103(2):131–6.
4. Yang P, Zhang Z, Zhou H, Li B, Huang X, Gao Y, et al. Clinical patterns and characteristics of uveitis in a tertiary center for uveitis in China. Curr Eye Res. 2005;30(11):943–8.
5. Khairallah M, Yahia SB, Ladjimi A, Messaoud R, Zaouali S, Attia S, et al. Pattern of uveitis in a referral centre in Tunisia, North Africa. Eye (London). 2007;21(1):33–9.
6. Grajewski RS, Caramoy A, Frank KF, Rubbert-Roth A, Fätkenheuer G, Kirchhof B, et al. Spectrum of uveitis in a German tertiary center: review of 474 consecutive patients. Ocul Immunol Inflamm. 2015;23(4):346–52.
7. Jones NP. The Manchester uveitis clinic: the first 3000 patients–epidemiology and casemix. Ocul Immunol Inflamm. 2015;23(2):118–26.
8. Llorenç V, Mesquida M, Sainz de la Maza M, Keller J, Molins B, Espinosa G, et al. Epidemiology of uveitis in a Western urban multiethnic population. The challenge of globalization. Acta Ophthalmol. 2015;93(6):561–7.
9. Winterhalter S, Stuebiger N, Maier AK, Pleyer U, Heiligenhaus A, Mackensen F, et al. Acute retinal necrosis: diagnostic and treatment strategies in Germany. Ocul Immunol Inflamm. 2016;24(5):537–43.
10. Cochrane TF, Silvestri G, McDowell C, Foot B, McAvoy CE. Acute retinal necrosis in the United Kingdom: results of a prospective surveillance study. Eye (Lond). 2012;26(3):370–7; quiz 378.
11. Muthiah MN, Michaelides M, Child CS, Mitchell SM. Acute retinal necrosis: a national population-based study to assess the incidence, methods of diagnosis, treatment strategies and outcomes in the UK. Br J Ophthalmol. 2007;91(11):1452–5.
12. Hillenkamp J, Nölle B, Bruns C, Rautenberg P, Fickenscher H, Roider J. Acute retinal necrosis: clinical features, early vitrectomy, and outcomes. Ophthalmology. 2009;116(10):1971–5.e2.
13. Meghpara B, Sulkowski G, Kesen MR, Tessler HH, Goldstein DA. Long-term follow-up of acute retinal necrosis. Retina. 2010;30(5):795–800.
14. Baltinas J, Lightman S, Tomkins-Netzer O. Comparing treatment of acute retinal necrosis with either oral valacyclovir or intravenous acyclovir. Am J Ophthalmol. 2018;188:173–80.
15. Holland GN. Standard diagnostic criteria for the acute retinal necrosis syndrome. Executive Committee of the American Uveitis Society. Am J Ophthalmol. 1994;117(5):663–7.
16. Wong R, Pavesio CE, Laidlaw DA, Williamson TH, Graham EM, Stanford MR. Acute retinal necrosis: the effects of intravitreal foscarnet and virus type on outcome. Ophthalmology. 2010;117(3):556–60.
17. Sugita S, Shimizu N, Watanabe K, Mizukami M, Morio T, Sugamoto Y, et al. Use of multiplex PCR and real-time PCR to detect human herpes virus genome in ocular fluids of patients with uveitis. Br J Ophthalmol. 2008;92(7):928–32.
18. Pendergast SD, Werner J, Drevon A, Wiedbrauk DL. Absence of herpesvirus DNA by polymerase chain reaction in ocular fluids obtained from immunocompetent patients. Retina. 2000;20(4):389–93.
19. Takase H, Okada AA, Goto H, Mizuki N, Namba K, Ohguro N, et al. Development and validation of new diagnostic criteria for acute retinal necrosis. Jpn J Ophthalmol. 2015;59(1):14–20.
20. Jabs DA, Belfort R Jr, Bodaghi B, Graham E, Holland GN, Lightman SL, et al. Classification criteria for acute retinal necrosis syndrome. Am J Ophthalmol. 2021;228:237–44.
21. Peng Y, Dharssi S, Chen Q, Keenan TD, Agrón E, Wong WT, et al. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. Ophthalmology. 2019;126(4):565–75.
22. Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. Ophthalmology. 2018;125(9):1410–20.
23. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316(22):2402–10.
24. Xu K, Huang S, Yang Z, Zhang Y, Fang Y, Zheng G, et al. Automatic detection and differential diagnosis of age-related macular degeneration from color fundus photographs using deep learning with hierarchical vision transformer. Comput Biol Med. 2023;167:107616.
25. Yang Z, Zhang Y, Xu K, Sun J, Wu Y, Zhou M. DeepDrRVO: a GAN-auxiliary two-step masked transformer framework benefits early recognition and differential diagnosis of retinal vascular occlusion from color fundus photographs. Comput Biol Med. 2023;163:107148.
26. Lei B, Zhou M, Wang Z, Chang Q, Xu G, Jiang R. Ultra-wide-field fundus imaging of acute retinal necrosis: clinical characteristics and visual significance. Eye (Lond). 2020;34(5):864–72.
27. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021. p. 10012–22.
28. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23(1):689.
29. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med. 2021;13(1):152.
30. Zhang Y, Yan C, Yang Z, Zhou M, Sun J. Multi-omics deep-learning prediction of homologous recombination deficiency-like phenotype improved risk stratification and guided therapeutic decisions in gynecological cancers. IEEE J Biomed Health Inform. 2023;PP. https://doi.org/10.1109/JBHI.2023.3308440.
31. Zhang Z, Chen H, Yan D, Chen L, Sun J, Zhou M. Deep learning identifies a T-cell exhaustion-dependent transcriptional signature for predicting clinical outcomes and response to immune checkpoint blockade. Oncogenesis. 2023;12(1):37.

Wang *et al. Eye and Vision*   *(2024) 11:27*

Page 11 of 11

32. Yang Z, Zhang Y, Zhuo L, Sun K, Meng F, Zhou M, et al. Prediction of prognosis and treatment response in ovarian cancer patients from histopathology images using graph deep learning: a multicenter retrospective study. Eur J Cancer. 2024;199:113532.

33. Zhang Y, Yang Z, Chen R, Zhu Y, Liu L, Dong J, et al. Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. NPJ Digit Med. 2024;7(1):15.

34. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. Lancet Digit Health. 2022;4(4):e235–44.

35. Parra R, Ojeda V, Vázquez Noguera JL, García Torres M, Mello Román JC, Villalba C, et al. Automatic diagnosis of ocular toxoplasmosis from fundus images with residual neural networks. Stud Health Technol Inform. 2021;281:173–7.

36. Parra R, Ojeda V, Vázquez Noguera JL, García-Torres M, Mello-Román JC, Villalba C, et al. A trust-based methodology to evaluate deep learning models for automatic diagnosis of ocular toxoplasmosis from fundus images. Diagnostics (Basel). 2021;11(11):1951.

37. González-López JJ, García-Aparicio ÁM, Sánchez-Ponce D, Muñoz-Sanz N, Fernandez-Ledo N, Beneyto P, et al. Development and validation of a Bayesian network for the differential diagnosis of anterior uveitis. Eye (Lond). 2016;30(6):865–72.

38. Mutawa AM, Alzuwawi MA. Multilayered rule-based expert system for diagnosing uveitis. Artif Intell Med. 2019;99:101691.

39. Tugal-Tutkun I, Onal S, Stanford M, Akman M, Twisk JWR, Boers M, et al. An algorithm for the diagnosis of Behçet disease uveitis in adults. Ocul Immunol Inflamm. 2021;29(6):1154–63.

40. Jamilloux Y, Romain-Scelle N, Rabilloud M, Morel C, Kodjikian L, Maucort-Boulch D, et al. Development and validation of a bayesian network for supporting the etiological diagnosis of uveitis. J Clin Med. 2021;10(15):3398.

41. Feng L, Zhou D, Luo C, Shen J, Wang W, Lu Y, et al. Clinically applicable artificial intelligence algorithm for the diagnosis, evaluation, and monitoring of acute retinal necrosis. J Zhejiang Univ Sci B. 2021;22(6):504–11.