Computers in Biology and Medicine 193 (2025) 110327

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Automatic transformer-based grading of multiple retinal inflammatory signs in uveitis on fluorescein angiography

Victor Amiot [a,b,#], Oscar Jimenez–del–Toro [c,#], Yan Guex–Crosier [a,d], Muriel Ott [a,d], Teodora-Elena Bogaciu [e,f], Shalini Banerjee [g,h], Jeremy Howell [g,h], Christoph Amstutz [g,h], Christophe Chiquet [e,f], Ciara Bergin [a], Ilenia Meloni [a,b], Mattia Tomasoni [a,b,1], Florence Hoogewoud [a,d,1], André Anjos [c,*,1]

[a] Department of Ophthalmology, University of Lausanne, Fondation Asile des Aveugles, Jules Gonin Eye Hospital, Lausanne, Switzerland
[b] Platform for Research in Ocular Imaging, Fondation Asile des Aveugles, Jules Gonin Eye Hospital, Lausanne, Switzerland
[c] Idiap Research Institute, Martigny, Switzerland
[d] Ocular Immune-infectiology Unit Jules-Gonin Eye Hospital, FAA, University of Lausanne, Switzerland
[e] Grenoble Alpes University, Grenoble, France
[f] Department of Ophthalmology, Grenoble Alpes University Hospital, Grenoble, France
[g] Department of Ophthalmology, Cantonal Hospital Lucerne, Lucerne, Switzerland
[h] Faculty of Health Sciences and Medicine, University of Lucerne, Lucerne, Switzerland

ARTICLE INFO

Keywords:
Fluorescein angiography
Uveitis
Vasculitis
Deep learning
Optic disease grading
Transformers
Intergrader agreement

ABSTRACT

Background: Grading fluorescein angiography (FA) for uveitis is complex, often leading to the oversight of retinal inflammation in clinical studies. This study aims to develop an automated method for grading retinal inflammation.

Methods: Patients from Jules-Gonin Eye Hospital with active or resolved uveitis who underwent FA between 2018 and 2021 were included. FAs were acquired using a standardized protocol, anonymized, and annotated following the Angiography Scoring for Uveitis Working Group criteria, for four inflammatory signs of the posterior pole. Intergrader agreement was assessed by four independent graders. Four deep learning transformer models were developed, and performance was evaluated using the Ordinal Classification Index, accuracy, F1 scores, and Kappa scores. Saliency analysis was employed to visualize model predictions.

Findings: A total of 543 patients (1042 eyes, 40987 images) were included in the study. The models closely matched expert graders in detecting vascular leakage (F1-score = 0·87, 1-OCI = 0·89), capillary leakage (F1-score = 0·86, 1-OCI = 0·89), macular edema (F1-score = 0·82, 1-OCI = 0·86), and optic disc hyperfluorescence (F1-score = 0·72, 1-OCI = 0·85). Saliency analysis confirmed that the models focused on relevant retinal structures. The mean intergrader agreement across all inflammatory signs was F1-score = 0·79 and 1-OCI = 0·83.

Interpretation: We developed a vision transformer-based model for the automatic grading of retinal inflammation in uveitis, utilizing the largest dataset of FAs in uveitis to date. This approach provides significant clinical benefits for the evaluation of uveitis and paves the way for future advancements, including the identification of novel biomarkers through the integration of clinical data and other modalities.

1. Introduction

Uveitis is a group of complex, sight-threatening inflammatory disorders that affect the uvea and retina, with significant consequences for patients, healthcare systems, and society. The annual incidence of uveitis is estimated to range from 17 to 52 cases per 100,000 people [1]. This debilitating condition can lead to severe visual impairment and is a major cause of blindness in industrialized countries [2]. This study focuses on the most severe forms of uveitis, including intermediate, posterior, and panuveitis, where inflammation compromises key structures

* Corresponding author. rue Marconi 19, Martigny, 1920, Switzerland.
E-mail address: andre.anjos@idiap.ch (A. Anjos).
# Victor Amiot and Oscar Jimenez-del-Toro contributed equally as first authors.
[1] Mattia Tomasoni, Florence Hoogewoud and Andre Anjos contributed equally as co-last authors.

https://doi.org/10.1016/j.compbiomed.2025.110327
Received 17 January 2025; Received in revised form 4 April 2025; Accepted 30 April 2025
Available online 21 May 2025

such as the retina, choroid, and optic disc.

Fluorescein angiography (FA) is the gold standard for evaluating inflammation in the retina, as it is the only clinical method that assesses the function and integrity of the blood-retinal barrier. However, interpreting FA in the context of uveitis is inherently challenging due to the dynamic nature of the examination and the presence of multiple confounding factors. Accurate interpretation requires extensive clinical experience. While imaging characteristics vary across different uveitis diagnoses, inflammation of key retinal structures, including the macula, optic disc, and retinal vasculature, is consistently associated with disease severity.

Precise monitoring of retinal inflammation is essential, as it often triggers the initiation or escalation of treatment by immunomodulatory therapy (IMT) [3]. Despite its clinical importance, retinal inflammation is rarely considered in clinical trials due to the time-consuming and partly subjective nature of FA interpretation. Most randomized controlled trials (RCT) base treatment outcome on visual acuity (VA), macular thickness on optical coherence tomography (OCT) and/or clinical flare up [4–6]. These parameters show a weak correlation with retinal inflammation on FA [7], highlighting the value of a standardized scoring system for assessing treatment response.

The Angiography Scoring for Uveitis Working Group (ASUWOG) [8] developed a semi-quantitative angiographic scoring system that has been validated and used in previous clinical studies. Its use requires an experienced grader and it is however very time-consuming hence the necessity for the development of a fully automatic grading of fluorescein angiography.

In recent years, deep learning (DL), a subset of artificial intelligence (AI), has transformed medical applications by facilitating advanced data analysis, pattern recognition, and predictive modeling, resulting in unprecedented accuracy in disease diagnosis, treatment planning, and personalized healthcare [9]. DL has also significantly impacted ophthalmology by enabling the creation of highly accurate models for diagnosing and managing ocular diseases, including diabetic retinopathy, glaucoma, and macular degeneration, through sophisticated imaging analysis and predictive algorithms [10].

The latest advancement in DL are Transformer models, which have shown significant promise in natural language processing, and have been adapted for computer vision tasks and medical image analysis, achieving notable results [11]. Unlike traditional convolutional networks, the convolution-free transformer leverages the self-attention mechanism to relate different positions in a sequence, offering richer feature representation [12]. Transformers, with their large effective receptive fields, excel in capturing long-range dependencies, thereby improving their understanding of contextual information compared to convolutional neural networks (CNNs) [11].

The Swin transformer is a hierarchical vision transformer with a shifted windows approach [11]. It can reduce the computational complexity of other transformer models, such as Vision transformer (ViT), thanks to the window self-attention mechanism, which can lead to high efficiency [13]. To mitigate the challenge of training transformer models with large-scale datasets, we rely on a pre-trained Swin transformer backbone. The pre-training for this backbone was performed on the larger ImageNet-22K dataset, which contains 14.2 million images and 22K classes.

One of the key drawbacks of deep learning (DL) is the difficulty in interpreting how models make decisions, often referred to as the “black box” problem. To address this issue and improve model interpretability, saliency maps can be used to highlight the regions of the input data that are most relevant to the predicted class. In this work, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) to create heatmaps, which visually illustrate where transformer models concentrated their attention within fluorescein angiography (FA) images.

Our primary contribution is the development of a vision transformer-based approach to automatically grade retinal inflammation using fluorescein angiography (FA). Specifically [1]: we collected and manually graded a large clinical dataset of FA examinations acquired in routine clinical practice [2]; we evaluated the intergrader agreement using four different graders [3]; we fine-tuned a pre-trained vision transformer model to accurately grade four key inflammatory signs in the posterior pole; and [4] we developed an ordinal classification evaluation framework to comprehensively compare the performance of the vision transformer models against three independent uveitis experts. This approach aims to improve both the precision and efficiency of retinal inflammation assessment.

2. Methods

2.1. Data collection

The study was conducted in accordance with the principles of the Declaration of Helsinki and received approval from the Institutional Review Board (CER-VD n° 2018–02161) for retrospective analysis up to 2018. For patients included after this period, informed consent was obtained, allowing the use of their data for research purposes (CER-VD PB 2016−00868).

We included patients from the ocular immune-infectiology department of Jules-Gonin Eye Hospital who were treated between 2018 and 2021 and underwent fluorescein angiography (FA) using the Spectralis-Heidelberg device (Heidelberg Engineering, Heidelberg, Germany). Both patients with active uveitis and those with resolved or no inflammation were included. Medical experts evaluated the image quality, and images deemed uninterpretable or of low quality (e.g., due to eyelid interference, advanced cataracts, or severe vitritis) were excluded. Additionally, patients with conditions that could interfere with angiography interpretation, such as diabetic retinopathy or vascular occlusions, were also excluded from the dataset.

The exam was performed according to a standardized protocol: images of the posterior pole were captured with a 55-degree field of view at approximately 1, 2, 5, and 10 min after intravenous injection of fluorescein (fluorescein 10 %, 5 mL) in both eyes. To capture the periphery, the fundus was divided into quadrants, and four images with a 102-degree field of view were taken between 3 and 5 min after injection. The image resolution was 768 × 768 pixels. As it is standard practice, the photographer manually adjusted parameters such as gain and exposure to optimize image quality. Gain control is crucial for enhancing the visibility of retinal structures. This manual adjustment was necessary due to the variability in fluorescence levels across different patients and stages of the angiography.

Data were extracted in DICOM format from the machine database of the Spectralis-Heidelberg angiograph. When the patient had multiple FAs, only the first was used. For input into the deep learning models, the last late frame (acquired 5–15 min after injection) of the 55-degree posterior pole was selected.

2.2. Data annotation, grading and reference standard

All images from the FA examination were anonymized and made available for grading. No clinical information was accessible to the graders. A pipeline was deployed on the Swiss Ophthalmic Imaging Network (SOIN) [14], allowing graders and computer scientists to collaborate seamlessly from their respective remote locations.

Data annotation was performed using Discovery® software by RetinAI, an Image Management System and Image Viewer for ophthalmology, which allows the attachment of Electronic Case Report Form (eCRF) to patient data. For manual grading of uveitis in FA, we used the scoring system validated by the ASUWOG [8]. In this study, a senior uveitis expert graded all the FA images. To assess inter-grader agreement, 111 eyes from the test set (see experimental setup below) were annotated by three additional experts from three different hospitals, and varying training backgrounds.

This study focuses on the four key inflammatory signs in the

posterior pole: macular edema, optic disc hyperfluorescence, retinal vascular staining and/or leakage in the posterior pole arcades, and capillary leakage in the posterior pole.

2.3. Experimental setup

The data were split into three subsets for training, validation and testing the machine learning models in a 70/15/15 % ratio. To mitigate potential biases, both eyes from the same patient were always included in the same subset. All left-eye images were flipped vertically to align their retinal structure with that of the right eyes. Stratified random splitting was employed to balance the representation of each of the four inflammatory signs. More specifically, we computed an optimal split by iteratively shuffling and selecting data sets until the prevalence of each inflammatory sign was balanced (see Supplementary Table 1).

2.4. Deep learning system development

For each of the targeted inflammatory signs, an encoder with the Swin transformer backbone was initialized using pre-trained weights, and a task-specific classification head was re-initialized based on the number of grades from inflammatory signs. A 12 × 12 pixel patch configuration was selected with non-overlapping local windows for computational efficiency. The Swin transformer's shifted windows approach facilitated information exchange between these non-overlapping windows. To match the input requirements of the Swin backbone, FA images were resized to 384 × 384 pixels.

The model was then fine-tuned over 15 epochs. The batch-size set to two, and a stock Adam optimizer with default parameters (via PyTorch) was used. An optimization function that penalizes larger errors, specifically the Ordinal Cross-Entropy Loss [15], was used during training and selected the best model. The final model was chosen based on performance on the validation set, measured using the Ordinal Classification Index (OCI, see Statistical Analyses below).

2.5. Interpretability: saliency map visualizations

Grad-CAM analysis was performed to generate visual explanations for the transformer model on individual FA images (Selvaraju et al., 2017). The implementation of Grad-CAM was sourced from https://github.com/jacobgil/pytorch-grad-cam.

2.6. Statistical Analyses

Intergrader agreement and algorithm performance were assessed using a traditional Kappa score, and the Ordinal Classification Index (OCI). The Ordinal Classification Index (OCI) [15] is a metric designed for measuring the performance of ordinal data classification (i.e., categories that can be ranked). Compared to more mainstream metrics such as the F1-score or the Kappa-score, this metric additionally penalizes the largest ordinal errors. This makes them more suited to the task at hand, since the scoring system used to grade the inflammatory signs includes relationships of superiority and inferiority between the grades. For completeness of algorithmic performance, and comparison with previously published models, we also report the F1-score, Kappa-score and accuracy.

2.7. Role of the funding source

This work was supported by the Swiss National Science Foundation (2018DRI13 to Thomas J. Wolfensberger), which provided access to the SOIN infrastructure. The Claire and Selma Kattenburg Foundation (Kattenburg Prize 2021 to Florence Hoogewoud and Mattia Tomasoni) and the Schmieder-Bohrisch Foundation (Schmieder-Bohrisch Prize 2023 to Mattia Tomasoni) who financially supported the development and analysis of the AI model. Additionally, the AURIS Foundation, the W. & E. Grand d'Hauteville Foundation, the Ingvar Kamprad Fund, the Fleurette Wagemakers Foundation, the Kononchuk Family grant, the Blatter Family grant, and the Rhumatismes-Enfants-Suisse Foundation supported the acquisition of clinical data through the JIR-cohort.

None of the funding agencies influenced the conceptualisation, data curation, formal analysis, investigation, methodology, supervision, validation, and visualizations documented in this study.

3. Results

The study design is illustrated in Fig. 1. A total of 543 patients (248 male, 295 female), with an average age of 48·35 years (SD 19), were included in the analysis. Of the 1042 eyes examined, 89 were excluded due to the unavailability of late-frame images, 53 were excluded due to poor image quality, and 149 were excluded due to alternative diagnoses.

In the study population (comprising 751 eyes), the most common inflammatory sign was optic disc hyperfluorescence (present in 304 eyes, 40·5 %), followed by macular edema (223 eyes, 29·7 %), vascular leakage (173 eyes, 23·0 %), and capillary leakage (141 eyes, 18·8 %), as depicted in Fig. 2. Overall, manual grading revealed that the prevalence of at least one inflammatory sign in the posterior pole of the retina was 49·2 %.

The four inflammatory signs were highly correlated, as shown in Supplementary Fig. 1. After binarizing each sign (present vs. absent) and calculating Pearson R correlation coefficients, the values ranged from 0·40 to 0·60, with the strongest correlation found between macular edema and capillary leakage.

The performance of the fine-tuned transformer models on the test set for each inflammatory sign is summarized in Table 1. The models demonstrated good generalization, with an average score across all signs of 1-OCI = 0·87 (SD 0·02), accuracy = 0·82 (SD 0·07), and F1-score = 0·82 (SD 0·07). Specifically, the individual model performances were comparable to the senior uveitis expert in detecting vascular leakage (1-OCI = 0·89, accuracy = 0·87, F1-score = 0·87), capillary leakage (1-OCI = 0·89, accuracy = 0·87, F1-score = 0·86), macular edema (1-OCI = 0·86, accuracy = 0·81, F1-score = 0·82), and optic disc hyperfluorescence (1-OCI = 0·85, accuracy = 0·72, F1-score = 0·72). The agreement between the transformer models and the senior uveitis expert for the four inflammatory signs on the test set is further illustrated in the confusion matrices shown in Fig. 3.

The inter-grader agreement for all inflammatory signs, based on grading provided by three independent experts, is also summarized in Table 1. For the four inflammatory signs, the average inter-grader agreement was 1-OCI = 0·83 (SD = 0·02), accuracy = 0·78 (SD = 0·03), and F1-score = 0·79 (SD = 0·03). Confusion matrices from the grading by the three independent experts are provided in Supplementary Figs. 2, 3, 4. The performance of the transformer models is comparable to the inter-grader agreement across all four inflammatory signs.

Qualitative results from the saliency map analysis are presented in Fig. 4, showing four representative images, one for each inflammatory sign, with GradCAM colorized heatmaps superimposed. Visual inspection of these saliency maps demonstrated that the transformer models accurately highlighted regions corresponding to the affected retinal structures. A manual review of cases revealed that only five eyes had errors exceeding two grades, primarily due to the presence of chorioretinal scars and diffuse capillaropathy (see examples in Supplementary Fig. 5). Two large errors were associated with vascular leakage grading, and two were related to macular edema. No evidence of systemic bias was detected upon visual inspection.

4. Discussion

Uveitis comprises a group of rare diseases with diverse clinical presentations and numerous associated diagnoses. However, inflammatory signs in the posterior pole are universal markers of disease severity, with a critical impact in determining treatment initiation and response

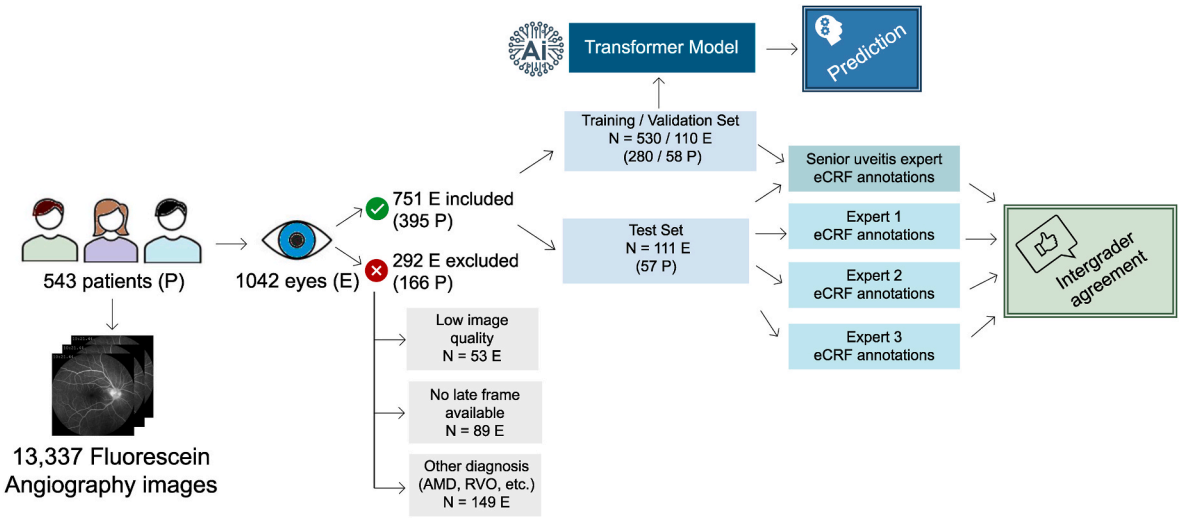

Fig. 1. Schematic depiction of the study design. Flow diagram of the study population from the ocular immune-infectiology department of Jules-Gonin Eye Hospital that met all inclusion criteria. Fluorescein angiographies were obtained from 543 patients. Of the 1042 eyes examined, 89 were excluded due to unavailability of late-frame images, 53 for poor image quality, and 149 for alternative diagnoses, resulting in a study population composed of 751 eyes in total. The data were split into training, validation, and testing subsets in a 70/15/15 % ratio. Grading was performed by a senior uveitis expert to train the algorithm, with intergrader agreement calculated based on additional grading by three experts from different hospitals.

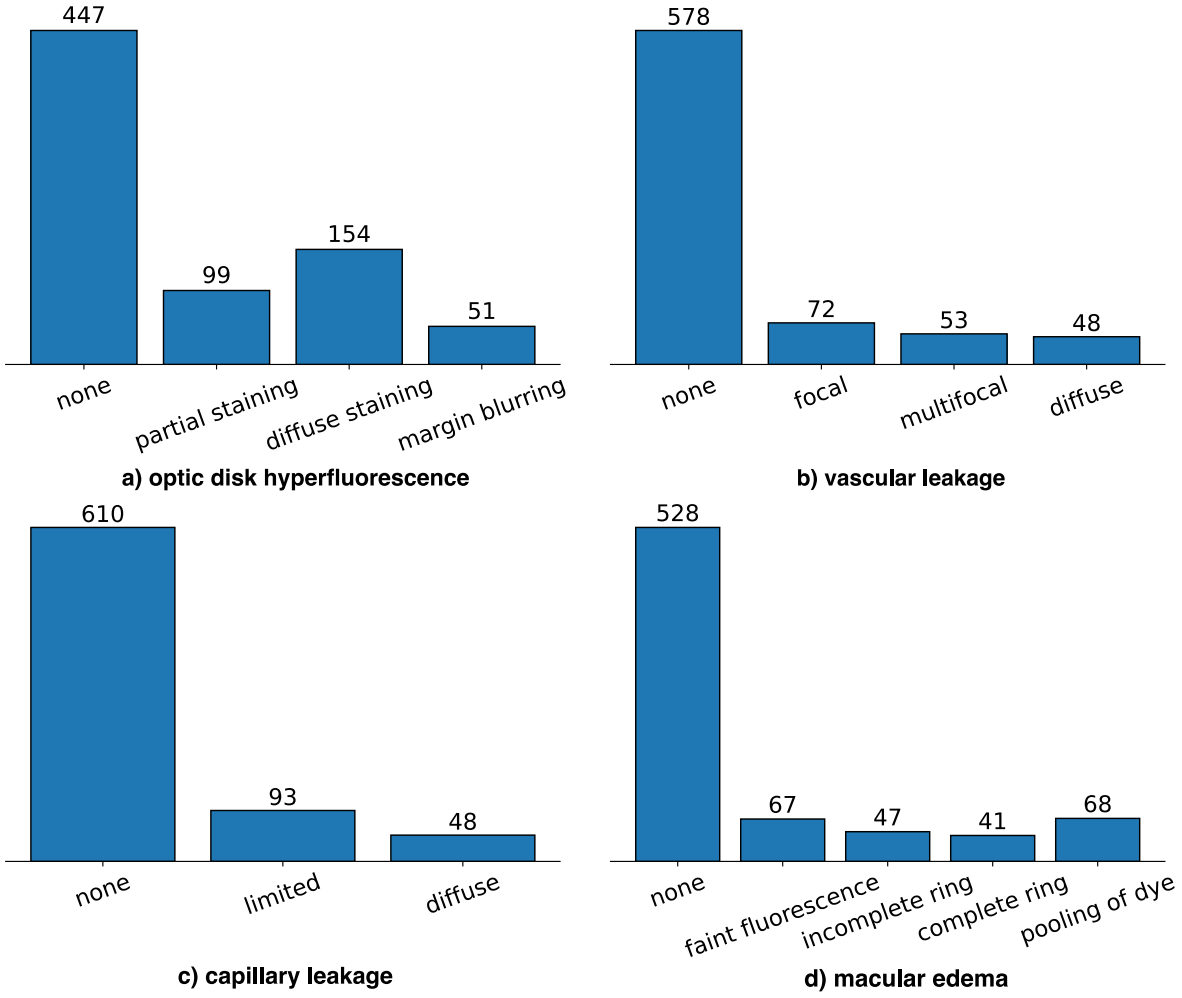

Fig. 2. Prevalence of inflammatory signs according to human grading. The study population was composed of 751 eyes in total. According to grading performed by a senior uveitis expert **a.** Optic disk hyperfluorescence was present in 304 eyes (40·5 %), not present in 447 eyes). **b.** Vascular leakage was present in 173 eyes (23·0 %), not present in 578. **c.** Capillary leakage was present in 141 eyes (18·8 %) not present in 610. **d.** Macular edema was present in 223 eyes (29·7 %), not present in 528.

Table 1
Performance of the fine-tuned transformer models on the test set for each inflammatory sign. Performance was evaluated against the senior uveitis expert using several metrics (1-OCI, accuracy, F1-score, and Kappa) for the four inflammatory signs: Vascular Leakage, Capillary Leakage, Macular Edema, and Optic Disc Hyperfluorescence. The top performance for each inflammatory sign and each metric is highlighted in bold. For each inflammatory sign, results are reported for the corresponding transformer model and each of the three independent uveitis experts.

	1-OCI	Acc	F1-w	Kappa
Vascular Leakage				
Transformer model	**0.888**	**0.865**	**0.869**	**0.621**
Expert 1	0.811	0.784	0.769	0.319
Expert 2	0.848	0.739	0.781	0.403
Expert 3	0.830	0.802	0.788	0.409
Expert avg.	0.830	0.775	0.779	0.377
Capillary Leakage				
Transformer model	**0.889**	**0.865**	0.860	**0.644**
Expert 1	0.873	**0.865**	**0.869**	0.642
Expert 2	0.836	0.775	0.792	0.455
Expert 3	0.846	0.829	0.826	0.524
Expert avg.	0.852	0.823	0.829	0.540
Macular Edema				
Transformer model	0.859	**0.811**	**0.824**	0.629
Expert 1	**0.861**	0.809	0.812	**0.636**
Expert 2	0.822	0.774	0.780	0.591
Expert 3	0.807	0.757	0.736	0.501
Expert avg.	0.830	0.780	0.776	0.576
Optic Disk Hyperfluorescence				
Transformer model	0.851	0.721	0.724	0.508
Expert 1	0.751	0.703	0.681	0.434
Expert 2	0.797	0.703	0.780	0.591
Expert 3	**0.870**	**0.829**	**0.834**	**0.701**
Expert avg.	0.806	0.745	0.765	0.575

assessment in clinical trials. In this study, we introduced the largest existing dataset of fluorescein angiography images in uveitis patients, with a manual grading according to a semi-quantitative angiographic scoring system. We developed transformer models to automatically assess four major retinal inflammatory signs, and compared their performance with the intergrader agreement from three experts. The proposed transformer models mirrored the intergrader agreement on all of the targeted inflammatory signs.

While previous studies have attempted to automate the detection of retinal inflammation, few have addressed this challenge using FA images. To the best of our knowledge, only four studies have tackled automated inflammation detection in uveitis [16–19]. A key limitation of these studies is their reliance on hand-picked images from selected patient groups, which limits the applicability in realistic clinical settings. Furthermore, they do not distinguish between different retinal structures, such as the capillaries and the optic disk, and use scoring systems that differ from those used in routine clinical practice [20]. This study, by contrast, includes a large, unselected database of 543 patients from daily clinical practice, making it the largest FA dataset for uveitis to date.

In [21] we developed an automated pipeline to detect (absence vs. presence) vascular leakage on individual FA images. We trained and tested several machine learning models to detect vascular leakage. The most promising results were obtained by leveraging a Convolutional Neural Network (CNN) optimized for medical imaging, with which, an area under the receiver operating characteristic curve (AUROC) of 0·86 for binary detection of vascular leakage was obtained. In an equivalent binary setting, the performance of the proposed current model based on a visual transformer greatly improved the AUROC of 0·96. Moreover, we now classify not one but four inflammatory signs and can additionally output severity grades. Finally, the performance of the ensemble of models mirrored that of three independent uveitis expert graders, with large errors only being observed in cases with important confounding factors (such as atrophic macular lesions).

The capabilities of the proposed model can be contextualized by comparing it to developments in diabetic retinopathy (DR). While several algorithms have been developed to detect DR, few have achieved high accuracy in classifying disease severity. A recent model [22] achieved high accuracy, but it does not differentiate between distinct retinal manifestations of DR, providing only a single output. In contrast, the proposed model offers an individual grade for each inflammatory sign, which could have clinical significance, particularly as certain treatments may be more effective for specific types of inflammation (e.g. tocilizumab and macular edema).

The intergrader agreement of FA grading systems has been examined in smaller studies [23,24]. For example, Moon et al. reported an intra-class correlation coefficient of 0·874 in standard FA and 0·928 in ultra-wide-field FA in a cohort of 21 patients with Behçet disease. The fine-tuned transformer models achieved a comparable average 1-OCI score of 0·872. Unlike prior studies that use hand-picked, smaller datasets, the proposed work is based on data collected on more realistic conditions. The inherent variability in translating continuous clinical signs into discrete variables explains why perfect intergrader agreement (1·0) constitutes a rather challenging goal. The use of a continuous inflammation scale could relax this constraint. To address the issue of ordinal misclassification, we trained models using an Ordinal cross-entropy loss, which strongly penalizes larger misclassifications, agreeing with clinically significant. For instance, the clinical impact of misclassifying "faint hyperfluorescence" versus "incomplete ring hyperfluorescence" is not as severe as misclassifying "faint hyperfluorescence" as "pooling of dye into cystic spaces." Classical metrics like F1 or Kappa do not capture this distinction, making the use of ordinal figures-of-merit more appropriate for this study.

This study presents some limitations. While the dataset used is the largest existing FA dataset for uveitis to date, it remains both modest and specific compared to those used for more common conditions like diabetic retinopathy. Although the model demonstrated good generalizability, including for capillary leakage, the least prevalent inflammatory sign in the dataset, external validation with data from other centers will be necessary to confirm broader applicability across different FA acquisition protocols and diverse patient populations. The integration of wide-field retinal imaging is planned in the development of our algorithm. Additionally, prospective clinical validation in larger cohorts will be crucial to further establish the utility of the proposed AI models.

One of the main advantages of transformer models is their flexibility in integrating multimodal data, including clinical information and other imaging modalities. This could open doors for future research, where image-based biomarkers could be used to predict treatment efficacy, adverse effects, and disease recurrence, as has been explored in diabetic retinopathy and age-related macular degeneration [25–27]. In uveitis, the identification of robust imaging biomarkers remains an unmet need [28,29].

Another important aspect is the potential for AI-driven tools to democratize access to expert-level uveitis care, which is often concentrated in academic medical centers in wealthier countries. By using AI to provide automated grading, we could improve access to high-quality diagnostics in regions with limited resources, contributing to a broader understanding of the disease's variability across different populations [30].

5. Conclusion

In this study, we developed a vision transformer-based model for the automatic grading of retinal inflammation, leveraging the largest

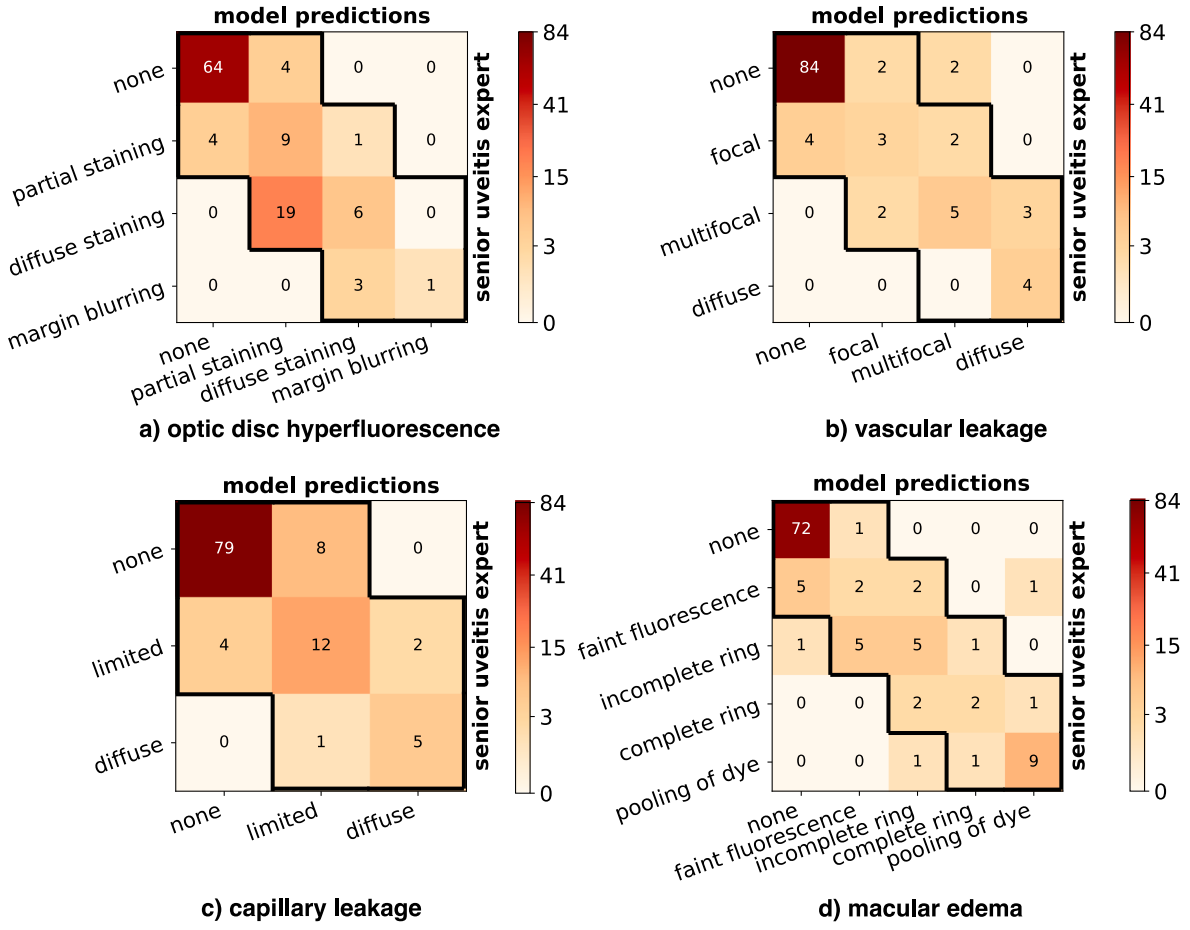

Fig. 3. Confusion matrices showing the performance from four transformer models in comparison to the grading from the senior uveitis expert. A test set of 111 eyes was used to evaluate the grading of the transformer models for a. Optic disk hyperfluorescence, b. Vascular leakage, c. Capillary leakage, d. Macular Edema.

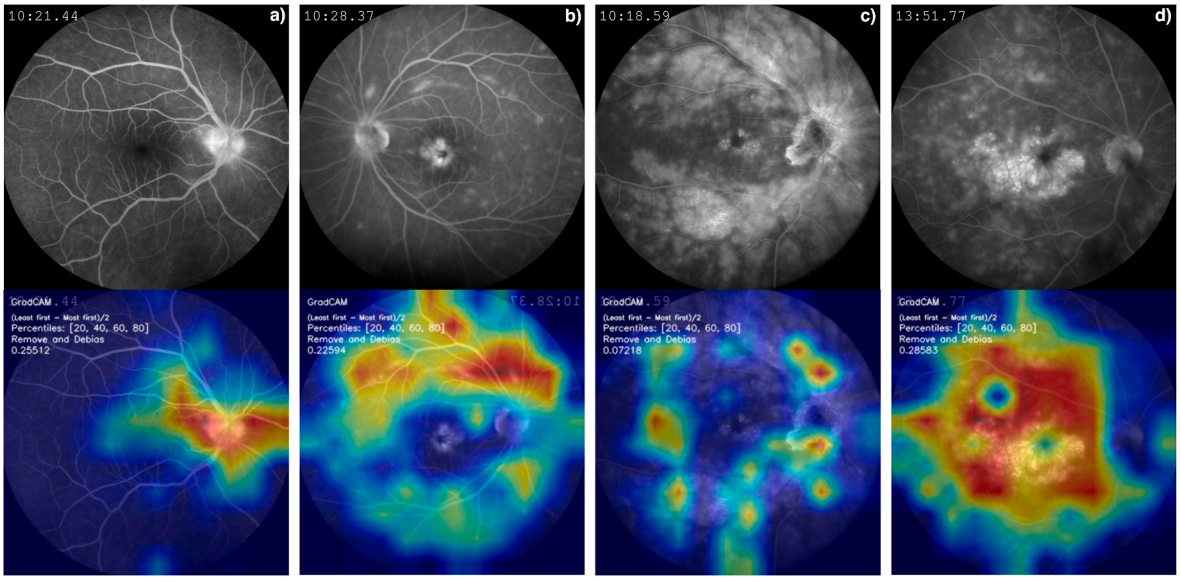

Fig. 4. Representative test set images and GradCAM saliency maps for each of the studied retinal inflammatory signs. a. Optic disk hyperfluorescence: Diffuse staining **b.** Vascular leakage: Focal **c.** Capillary leakage: Diffuse **d.** Macular Edema: Pooling dye in cystic spaces. **Bottom row:** Composite colorized Grad-CAM saliency maps use red or orange to highlight the areas that are, on average, most important for transformer models in grading the representative image. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

dataset of fluorescein angiography (FA) images in uveitis to date. The model replicated the scoring system of the Angiography Scoring for Uveitis Working Group with high accuracy and consistency when compared to expert uveitis graders. Notably, it outperformed expert agreement in detecting vascular and capillary leakage and showed significant promise in grading macular edema. Saliency map visualizations confirmed the model's precise focus on key retinal structures associated with inflammation. Future work will focus on integrating additional modalities, such as optical coherence tomography (OCT), to enhance the model's capabilities. By automating the grading process, this tool holds the potential to significantly improve diagnostic consistency, streamline clinical workflows, and enable more personalized uveitis treatment while facilitating the discovery of novel biomarkers.

CRediT authorship contribution statement

Victor Amiot: Writing – review & editing, Visualization, Validation, Software, Investigation, Data curation. **Oscar Jimenez–del–Toro:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Project administration, Methodology, Investigation. **Yan Guex–Crosier:** Resources, Project administration, Formal analysis. **Muriel Ott:** Investigation, Data curation. **Teodora-Elena Bogaciu:** Investigation, Data curation. **Shalini Banerjee:** Investigation, Data curation. **Jeremy Howell:** Investigation, Formal analysis, Data curation. **Christoph Amstutz:** Resources, Methodology, Data curation, Conceptualization. **Christophe Chiquet:** Supervision, Resources, Investigation, Data curation. **Ciara Bergin:** Methodology, Funding acquisition, Conceptualization. **Ilenia Meloni:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Mattia Tomasoni:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Florence Hoogewoud:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **André Anjos:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Data sharing

Models, network weights and code will be made available to researchers for non-commercial purposes at https://github.com/JulesGoninRIO/uveitis_transformers. Original images, and other patient data from Jules Gonin Eye Hospital cannot be shared due to confidentiality agreements and/or ethical approval for disclosing patient data.

Ethics statement

We declare the submitted study was conducted in accordance with the principles of the Declaration of Helsinki and received approval from the Institutional Review Board (CER-VD n° 2018–02161) for retrospective analysis up to 2018. For patients included after this period, informed consent was obtained, allowing the use of their data for research purposes (CER-VD PB 2016−00868).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ciara Bergin reports financial support was provided by Swiss Personalized Health Network. Florence Hoogewoud reports financial support was provided by Claire and Selma Kattenburg Foundation. Mattia Tommasoni reports financial support was provided by AURIS Foundation. Florence Hoogewoud reports financial support was provided by W. & E. Grand d'Hauteville Foundation. Mattia Tommasoni reports financial support was provided by Ingvar Kamprad Fund. Mattia Tommasoni reports financial support was provided by Fleurette Wagemakers Foundation. Florence Hoogewoud reports financial support was provided by Konochuk Family grant. Mattia Tommasoni reports financial support was provided by Batter Family grant. Florence Hoogewould reports administrative support was provided by Rhumatismes-Enfants-Suisse Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2025.110327.

References

[1] T. Tsirouki, A. Dastiridou, C. Symeonidis, O. Tounakaki, I. Brazitikou, C. Kalogeropoulos, et al., A focus on the epidemiology of uveitis, Ocul. Immunol. Inflamm. 26 (1) (2018 Jan 2) 2–16.
[2] O.M. Durrani, N.N. Tehrani, J.E. Marr, P. Moradi, P. Stavrou, P.I. Murray, Degree, duration, and causes of visual loss in uveitis, Br. J. Ophthalmol. 88 (9) (2004 Sep) 1159–1162.
[3] A.D. Dick, J.T. Rosenbaum, H.A. Al-Dhibi, R. Belfort, A.P. Brézin, S.P. Chee, et al., Guidance on noncorticosteroid systemic immunomodulatory therapy in noninfectious uveitis: fundamentals of care for UveitiS (FOCUS) initiative, Ophthalmology 125 (5) (2018 May) 757–773.
[4] Q.D. Nguyen, P.T. Merrill, G.J. Jaffe, A.D. Dick, S.K. Kurup, J. Sheppard, et al., Adalimumab for prevention of uveitic flare in patients with inactive non-infectious uveitis controlled by corticosteroids (VISUAL II): a multicentre, double-masked, randomised, placebo-controlled phase 3 trial, Lancet 388 (10050) (2016 Sep 17) 1183–1192.
[5] S.R. Rathinam, J.A. Gonzales, R. Thundikandy, A. Kanakath, S.B. Murugan, R. Vedhanayaki, et al., Effect of corticosteroid-sparing treatment with mycophenolate mofetil vs methotrexate on inflammation in patients with uveitis: a randomized clinical trial, JAMA 322 (10) (2019 Sep 10) 936–945.
[6] Z. Zhong, L. Dai, Q. Wu, Y. Gao, Y. Pu, G. Su, et al., A randomized non-inferiority trial of therapeutic strategy with immunosuppressants versus biologics for Vogt-Koyanagi-Harada disease, Nat. Commun. 14 (1) (2023 Jun 24) 3768.
[7] M.A. Sadiq, M. Hassan, R. Afridi, M.S. Halim, D.V. Do, Y.J. Sepah, et al., Posterior segment inflammatory outcomes assessed using fluorescein angiography in the STOP-UVEITIS study, Int J Retina Vitr 6 (1) (2020 Dec) 47.
[8] The Angiography Scoring for Uveitis Working Group (ASUWOG), I. Tugal-Tutkun, C.P. Herbort, M. Khairallah, Scoring of dual fluorescein and ICG inflammatory angiographic signs for the grading of posterior segment inflammation (dual fluorescein and ICG angiographic scoring system for uveitis), Int. Ophthalmol. 30 (5) (2010 Oct) 539–552.
[9] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, et al., A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017 Dec 1) 60–88.
[10] U. Schmidt-Erfurth, A. Sadeghipour, B.S. Gerendas, S.M. Waldstein, H. Bogunović, Artificial intelligence in retina, Prog. Retin. Eye Res. 67 (2018 Nov) 1–29.
[11] Z. Liu, Q. Lv, Z. Yang, Y. Li, C.H. Lee, L. Shen, Recent progress in transformer-based medical image analysis, Comput. Biol. Med. 164 (2023 Sep 1) 107268.
[12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, et al., Attention is all you need, in: Advances in Neural Information Processing Systems, Curran Associates, Inc., 2017 [cited 2025 Apr 4]. Available from: https://dl.acm.org/doi/pdf/10.5555/3295222.3295349.
[13] L. Huang, Y. Xu, S. Wang, L. Sang, H. Ma, SRT: Swin-residual transformer for benign and malignant nodules classification in thyroid ultrasound images, Med. Eng. Phys. 124 (2024 Feb) 104101.
[14] C. Bergin, I. Mantel, R.O. Schlingemann, M. Tomasoni, T.J. Wolfensberger, SOIN – MI data lab: personalized ophthalmology through collaborative data collection and dynamic patient consent, in: B. Séroussi, P. Weber, F. Dhombres, C. Grouin, J. D. Liebe, S. Pelayo, et al. (Eds.), Studies in Health Technology and Informatics, IOS Press, 2022, https://doi.org/10.3233/SHTI220457 [Internet].
[15] J.S. Cardoso, R. Sousa, Measuring the performance of ordinal classification, Int. J. Pattern Recogn. Artif. Intell. 25 (8) (2011 Dec) 1173–1195.
[16] M. Karampelas, D.A. Sim, C. Chu, E. Carreno, P.A. Keane, J. Zarranz-Ventura, et al., Quantitative analysis of peripheral vasculitis, ischemia, and vascular leakage in uveitis using ultra-widefield fluorescein angiography, Am. J. Ophthalmol. 159 (6) (2015 Jun 1) 1161–1168.e1.
[17] H. Keino, T. Wakitani, W. Sunayama, Y. Hatanaka, Quantitative analysis of retinal vascular leakage in retinal vasculitis using machine learning, Appl. Sci. 12 (24) (2022 Jan) 12751.
[18] A. Sood, S. Sharma, K. Baynes, E. Fisher, N. Karthik, J. Welsh, et al., [10] automated quantification of vascular leakage in ultra-widefield fluorescein angiography in patients treated with fluocinolone acetonide implant 0.59mg, Investig. Ophthalmol. Vis. Sci. 60 (9) (2019 Jul 22) 1575.

[19] L.H. Young, J. Kim, M. Yakin, H. Lin, D.T. Dao, S. Kodati, et al., Automated detection of vascular leakage in fluorescein angiography – a proof of concept, Transl. Vis. Sci. Technol. 11 (7) (2022 Jul 25) 19.
[20] J.P. Ehlers, K. Wang, A. Vasanji, M. Hu, S.K. Srivastava, Automated quantitative characterisation of retinal vascular leakage and microaneurysms in ultra-widefield fluorescein angiography, Br. J. Ophthalmol. 101 (6) (2017 Jun 1) 696–699.
[21] V. Amiot, O. Jimenez-Del-Toro, P. Eyraud, Y. Guex-Crosier, C. Bergin, A. Anjos, et al., Fully automatic grading of retinal vasculitis on fluorescein angiography time-lapse from real-world data in clinical settings, in: 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), IEEE, L'Aquila, Italy, 2023, pp. 689–693, https://doi.org/10.1109/CBMS58004.2023.00301 [Internet].
[22] U. Bhimavarapu, N. Chintalapudi, G. Battineni, Automatic detection and classification of diabetic retinopathy using the improved pooling function in the convolution neural network, Diagnostics 13 (15) (2023 Aug 5) 2606.
[23] S.W. Moon, B.H. Kim, U.C. Park, H.G. Yu, Inter-observer variability in scoring ultra-wide-field fluorescein angiography in patients with Behçet retinal vasculitis, Ocul. Immunol. Inflamm. 25 (1) (2017 Jan 2) 20–28, https://doi.org/10.1080/09273948.2016.1175640.
[24] I. Tugal-Tutkun, C.P. Herbort, M. Khairallah, A. Mantovani, Interobserver agreement in scoring of dual fluorescein and ICG inflammatory angiographic signs for the grading of posterior segment inflammation, Ocul. Immunol. Inflamm. 18 (5) (2010 Oct) 385–389.
[25] Y. Chen, H. Li, A. Janowczyk, P. Toro, G. Corredor, J. Whitney, et al., Computational pathology improves risk stratification of a multi-gene assay for early stage ER+ breast cancer, Npj Breast Cancer 9 (1) (2023 May 17) 1–10.
[26] C. Jin, H. Yu, J. Ke, P. Ding, Y. Yi, X. Jiang, et al., Predicting treatment response from longitudinal images using multi-task deep learning, Nat. Commun. 12 (1) (2021 Mar 25) 1851.
[27] H.M. Kang, S.C. Lee, Long-term progression of retinal vasculitis in Behçet patients using a fluorescein angiography scoring system, Graefes Arch. Clin. Exp. Ophthalmol. 252 (6) (2014 Jun 1) 1001–1008.
[28] D.D. Sevgi, S.K. Srivastava, J. Whitney, M. O'Connell, S.S. Kar, M. Hu, et al., Characterization of ultra-widefield angiographic vascular features in diabetic retinopathy with automated severity classification, Ophthalmol Sci 1 (3) (2021 Sep 1) 100049.
[29] Y. Zhao, I.J.C. MacCormick, D.G. Parry, N.A.V. Beare, S.P. Harding, Y. Zheng, Automated detection of vessel abnormalities on fluorescein angiogram in malarial retinopathy, Sci. Rep. 5 (1) (2015 Jun 8) 11154.
[30] Development of classification criteria for the uveitides, Am. J. Ophthalmol. 228 (2021 Aug) 96–105.