Check for updates

OPEN ACCESS

EDITED BY Haoyu Chen, The Chinese University of Hong Kong, Hong Kong SAR, China

REVIEWED BY Gilbert Yong San Lim, SingHealth, Singapore Ling-Ping Cen, Guangdong Medical University, China Md Mahmudul Hasan, University of New South Wales, Australia

*CORRESPONDENCE Lina Huang ⊠ lina_h@126.com Shiqi Ling ⊠ lingshiqi123@163.com

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 04 January 2025 ACCEPTED 14 April 2025 PUBLISHED 07 May 2025

CITATION

Gu S, Bao T, Wang T, Yuan Q, Yu W, Lin J, Zhu H, Cui S, Sun Y, Jia X, Huang L and Ling S (2025) Multimodal AI diagnostic system for neuromyelitis optica based on ultrawide-field fundus photography. *Front. Med.* 12:1555380. doi: 10.3389/fmed.2025.1555380

COPYRIGHT

© 2025 Gu, Bao, Wang, Yuan, Yu, Lin, Zhu, Cui, Sun, Jia, Huang and Ling. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Multimodal AI diagnostic system for neuromyelitis optica based on ultrawide-field fundus photography

Simin Gu^{1†}, Tiancheng Bao^{1†}, Tao Wang¹, Qiting Yuan², Weiguang Yu², Jiayi Lin¹, Haocheng Zhu¹, Shihai Cui¹, Yi Sun¹, Xiuhua Jia¹, Lina Huang³* and Shiqi Ling¹*

¹Department of Ophthalmology, The Third Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China, ²Zeemo Technology Company Limited, Shenzhen, China, ³Shenzhen Eye Hospital, Jinan University, Shenzhen, China

Purpose: While deep learning (DL) has demonstrated significant utility in ocular diseases, no clinically validated algorithm currently exists for diagnosing neuromyelitis optica (NMO). This study aimed to develop a proof-of-concept multimodal artificial intelligence (AI) diagnostic model that synergistically integrates ultrawide field fundus photographs (UWFs) with clinical examination data for predicting the onset and stage of suspected NMO.

Methods: The study utilized the UWFs of 330 eyes from 285 NMO patients and 1,288 eyes from 770 non-NMO participants, along with clinical examination reports, to develop an AI model for predicting the onset or stage of suspected NMO. The performance of the AI model was evaluated based on the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity.

Results: The multimodal AI diagnostic model achieved an AUC of 0.9923, a maximum Youden index of 0.9389, a sensitivity of 97.0% and a specificity of 96.9% in predicting the prevalence of NMO on test data set.

Conclusion: Our study demonstrates the feasibility of DL algorithms in diagnosing and predicting of NMO.

KEYWORDS

neuromyelitis optic, neuromyelitis optica spectrum disorders, deep learning, artificial intelligence, diagnostic model

1 Introduction

Neuromyelitis optica (NMO), also known as Devic syndrome, is an idiopathic neuroinflammatory disorder. This rare relapsing clinical syndrome characterized by astrocytopathy of the central nervous system, with a predilection for the optic nerves and spinal cord (1), disease occurs globally and affects individuals of all ethnicities (2). Its primarily manifestations include acute optic neuritis (ON) (3) and transverse myelitis (TM) (4), both of which cause blindness and paralysis. NMO has a poor prognosis and has long been considered a clinical variant of multiple sclerosis (MS). However, the disease is now studied as a prototypic autoimmune disorder on the basis of the discovery of a novel and pathogenic anti-astrocytic serum autoantibody that targets aquaporin-4 (AQP4-Ab) and IgG autoantibodies to myelin oligodendrocyte glycoprotein (MOG-IgG). Pathogenetic AQP4-IgG is responsible for more than 80% of NMO cases (5), while approximately 10–40% of individuals with NMO lack

AQP4-IgG and instead exhibit pathogenetic MOG-IgG (6). Although the AQP4 antibody is highly specific for this disorder, some patients harboring this antibody may present with isolated ON or TM. Magnetic resonance imaging (MRI) is commonly used to identify and characterize lesions in suspected NMO cases, helping to distinguish between NMO and MS. To encompass the broader clinical spectrum, including atypical or incomplete presentations, the term 'neuromyelitis optica spectrum disorders' (NMOSD) was subsequently introduced to refer to NMO and its formes frustes (7).

A subset of patients with NMO exhibit a variety of different symptoms indicating brain or brainstem involvement; however, ON and TM are the predominant manifestations of the disease. Ocular complications frequently include reduced visual acuity, a decline in high-contrast visual acuity, color desaturation (8), scotoma, and ocular pain during eye movement (9). These visual disturbances frequently serve as critical diagnostic indicators of NMO. If left untreated, NMO can lead to severe and irreversible visual impairment as well as significant motor dysfunction owing to incomplete recovery from acute attacks. Hence, advancements in early and accurate diagnosis are crucial, as they would facilitate prompt therapeutic intervention and significantly improve long-term clinical outcomes for patients with NMO.

Currently, the clinical diagnosis of NMO is conducted by qualified ophthalmologists or neurologists based on clinical presentation, disease course, and the detection of specific autoantibodies, supplemented by MRI scanning (10). Notably, the confirmation process requires specialized clinical expertise, posing a difficult challenge in resource-limited settings such as developing countries or rural communities where access to trained ophthalmologists or neurologists is scarce. Taken together, existing approaches for detecting NMO have substantial limitations in terms of accessibility and widespread implementation. Given the increasing global health burden associated with this disease, there is an urgent need to develop innovative diagnostic strategies that can overcome these constraints, enhance early detection, and ultimately improve patient outcomes.

Advancements in artificial intelligence (AI), particularly in machine learning (ML) and deep learning (DL), has driven remarkable breakthroughs in different areas of research (11-13). In ophthalmology, the widespread availability of fundus digital imaging has sparked growing interest in leveraging AI for early diagnosis and treatment in retinal and optic nerve disorders (14). Among these imaging modalities, ultrawide field fundus photographs (UWFs) have gained prominence in medical applications, particularly for ocular disease detection. UWFs offers various merits, including their suitability for nondilated pupils, the speed of acquisition, and the wide retinal imaging ranges (up to 200°) (15). Consequently, a variety of studies have explored the integration of AI with UWFs for diagnosing and treating of various ocular conditions such as diabetic retinopathy (16), age-related macular degeneration (17), cataracts (18), glaucoma (19, 20) and myopia (21). DL methodologies based on UWFs have become instrumental in reshaping diagnostic strategies in ophthalmology. Notably, DL methodologies based on UWFs have emerged as transformative tools, reshaping diagnostic paradigms in ophthalmology and paving the way for more efficient, accurate, and accessible disease detection strategies.

Standard color fundus photographs provides a 30 to 50-degree image whereas UWFs provide an encompassing view of the retina, allowing examination of not only the central retinal area but also the peripheral zones, which able to detect predominantly peripheral lesions in eyes with its wide coverage (22). The analysis of ultrawide field fundus could be of value in screening, given the prognostic importance of peripheral lesions in distinguishing NMOSD from other fundus disease progression. However, despite NMOSD being a major cause of visual impairment worldwide, the application of AI in its diagnosis remains relatively unexplored. This raises an intriguing question: could NMOSD be accurately diagnosed solely through UWF imaging, without the need for multiple diagnostic modalities? Exploring this possibility could pave the way for more accessible and efficient diagnostic strategies in clinical practice.

2 Methods

2.1 Subject characteristics

This clinical study adhered to the tenets of the Declaration of Helsinki and was approved by the research ethics committee of the Third Affiliated Hospital of Sun Yat-sen University.

The study utilized the UWFs of 330 eyes from 285 NMO patients and 1,288 eyes from 770 non-NMO participants, along with clinical examination reports were initially selected for the study. All the enrolled patients were referred to the Third Affiliated Hospital of Sun Yat-sen University from January 2022 to April 2024. The whole process was performed at the hospital, and informed consent was obtained from all study participants. The inclusion criteria included a diagnosis of NMO confirmed by a qualified neurosurgeon from the Third Affiliated Hospital of Sun Yat-sen University, based on diagnostic criteria established in previous population-based studies. Patients excluded from the possibility of NMO were classified as non-NMO, however, they could present with other ocular pathologies, such as diabetic retinopathy, glaucoma, retinal detachment, or retinal degeneration.

2.2 Multimodal dataset preparation

The dataset used in this study contains the data from two modalities: UWF fundus images, captured using scanning laser ophthalmoscopy (SLO, Optos Daytona) and clinical examination reports. To ensure the quality of the research, stringent image quality control criteria were implemented. All images were captured by extensively trained and specialized technicians, ensuring that the scanning quality of each image adheres to the standards of clarity and visual interpretability, absented from significant motion artifacts, and with precise centration on the optic nerve head or macula. Images failing to meet these criteria were excluded. During data preprocessing, all images were assessed for quality using both automated algorithms and manual inspection by experienced ophthalmologists. Images with severe artifacts that could compromise feature extraction were excluded from analysis. All UWFs analyzed in the present study were restricted to a single, highest-quality capture per patient, as determined by standardized ophthalmic imaging quality metrics (focus clarity \geq 70%, illumination uniformity, and absence of artifacts). Six ophthalmologists with clinical experience were recruited to evaluate the images and clinical examination reports. To ensure privacy, all images were first deidentified to remove any patientrelated information. Subsequently, the images were preprocessed using a pretrained deep learning segmentation model to remove extraneous "image borders," including the eyelids and the edges of the scanning machine, which are irrelevant to the current classification task. The segmentation model removes these image borders by generating a predicted mask of the detected border region while retaining the retina region. The output image of the segmentation model is padded using black pixels to produce a square shape with the retina area centered vertically for model input. This approach is well-justified, particularly in tasks that prioritize critical retinal regions, such as the optic disc and macula. By employing this method, interference such as eyelids, and machine edges can be minimized, data quality can be significantly improved, and overall model performance can be enhanced. This represents a widely accepted and validated practice in the field. Figure 1 shows the original SLO image and the corresponding preprocessed "borderless" image.

Three types of clinical examination reports, including those with data from MRI of the optic nerve, MRI of the spine and AQP4 antibody tests, were collected and analyzed. The model prediction probabilities and clinical data are integrated using a conditional decision approach. Specifically, the image score, obtained from the DL module, and the clinical score, derived from the clinical criteria, are weighted and input into the model to generate the final NMO prediction. The reports for the optic nerve MRI scans and the AQP4 antibody test were considered either positive or negative. Categorical variables were first converted into two dummy variables (1 for positive and 0 for negative) through manual coding, followed by one-hot encoding for model input. The record for spine MRI is an unstructured text description of the MRI result. Thus, the spine MRI results were preprocessed into dummy variables as follows: (a) the results were encoded as 1 (positive) if demyelinating disease was observed across three consecutive vertebrae; (b) the results were encoded as 0 (negative) if demyelinating disease was not observed or if it did not affect three consecutive vertebrae. Finally, we created a new dummy variable named "clinical result," which was set to 1 if any of the dummy variables above was positive and 0 otherwise. The composite clinical outcome was coded as positive (1) if any one of the following



criteria were present: AQP4-IgG seropositivity, orbital MRI demonstrating characteristic features diagnostic of NMO, or spinal cord MRI showing lesions spanning \geq 3 vertebral segments. Certain cases exhibiting characteristic optic nerve abnormalities, including optic disc hyperemia, edema, indistinct margins, and a partially increased cup-to-disc (C/D) ratio suggestive of optic nerve atrophy (Figure 2) were classified as high-risk for NMO. For all 330 included fundus images, matched clinical data were available. Cases with missing data were classified as positive if at least one available criterion was met; if all criteria were missing, the AI prediction alone determined the outcome.

The image dataset comprised 1,618 images in total (330 or 20.4% NMO positive and 1,288 or 79.6% NMO negative), among which 330 were associated with clinical records. The dataset was split into training, validation and test datasets at a ratio of 80:10:10 on the basis of a stratified-group-split approach, i.e., stratification by NMO labels and grouping by subject ID.

2.3 Model design

We developed a multimodal AI diagnosis system for NMO (MAiDS-NMO), the structure of which is depicted in Figure 3. MAiDS-NMO takes inputs from two modalities: fundus imaging and clinical results. The fundus images are fed to a deep neural network model for NMO likelihood inference, and the clinical result is used

together with the predicted NMO probability from the fundus images to make a positive or negative prediction of NMO. The deep model is Inception-V3 with Log-Sum-Exp (LSE) pooling (replacing the original adaptive average pooling method) (23), enabling pixels with similar scores to have similar weights in the pooling process during training.

$$x_{LSE_p} = \frac{1}{r} \log \left[\frac{1}{w * h} \sum_{(i,j)} rx_{ij} \right], \text{ where } r \text{ is a}$$

$$hyperparameter, (h, w) \text{ is input shape}$$

2.4 Model development

We use the PyTorch implementation of the Inception-V3 model (LSE pooling) with the pretrained weight file "inception_v3_google," which was developed by Google using the ImageNet dataset (ILSVRC-2012-CLS dataset for image classification) under the TensorFlow framework. One NVIDIA GeForce RTX 3090 Ti (24 GB) GPU was used for model development.

The training dataset was used to fine-tune the ImageNet pretrained Inception-V3 model, while the validation dataset was used to select the best weights (those which yielded the minimum validation loss).



FIGURE 2

Representative "positive" NMO UWF images. (A1,B1) The right eye of a 34-year-old female with NMO demonstrated a partially increased cup-to-disc (C/D) ratio. (A2,B2) The left eye of a 31-year-old female with NMO exhibited optic disc hyperemia, edema, and indistinct margins.



We adopted a three-step "warm-up" procedure during training, i.e., we froze all the layers (blocks) except for the fully connected dense layer of the model for the first several epochs of training and gradually unfroze Block 7, then Blocks 5 & 6, and finally the remaining blocks of the model. This method effectively prevents the training process from being trapped by a local minimum before it finds a global minimum.

The diagnosis of NMO can be reliably established through serological antibody testing (AQP4-IgG/MOG-IgG) and characteristic neuroimaging findings on MRI. Our diagnostic model developed exclusively using UWFs parameters demonstrated comparable diagnostic accuracy (AUROC 0.97, specificity 0.92). After training the DL model with the dataset of 1,618 fundus photos, including 330 NMO images, our diagnostic model developed exclusively using UWFs parameters demonstrated high confidence in the results, with an area under the curve (AUC) > 0.97 and a specificity > 92% in distinguishing NMO from non-NMO fundus photos. Notably, the model incorporating comprehensive clinical metadata showed enhanced diagnostic performance, with an area under the curve (AUC) > 0.99 and a specificity > 96%, though comparative analysis revealed no statistically significant difference between the two models.

Five-fold cross-validation was performed to rigorously evaluate the robustness and generalizability of the model. This methodological approach entails partitioning the dataset into five equal subsets, wherein the model is iteratively trained on four subsets and validated on the remaining one. By systematically rotating the test set across all five folds, this technique facilitates a comprehensive assessment, mitigating potential biases associated with a single training-test split.

2.5 Comparative test

A panel of six ophthalmologists with different levels of experience (comprising two neuro-ophthalmology specialists, two attending ophthalmologists, and two resident ophthalmologists) were recruited to participated in a diagnostic validation study involving 200 UWFs at the Third Affiliated Hospital of Sun Yat-sen University between January 2022 to April 2024. While the fundus photos were being presented, the AI model identified whether the UWFs depicted NMO. Simultaneously, the ophthalmologists were asked to independently complete the same test as the AI model without access to the clinical examination reports. In the subsequent validation phase, both the MAiDS-NMO and human experts repeated the evaluation with full access to multimodal clinical parameters. The diagnostic performances of the AI model and ophthalmologists were recorded and subjected to comparative statistical analysis. The functional architecture and training pipeline of the AI model and the development process are represented in Figure 4.

3 Results

3.1 Definitions of NMOSD

The diagnostic criteria for possible NMOSD were based on the following published international consensus diagnostic criteria: (10) an AQP4-IgG-positive serostatus, the presence of at least one of the "core characteristics," and the exclusion of alternative diagnoses. The diagnostic and exclusion criteria are detailed in Box 1, 2, respectively.

3.2 Image datasets and patient clinical characteristics

We established a dataset composed of UWFs and clinical examination reports collected in Guangzhou, Third Affiliated Hospital of Sun Yat-sen University, from January 2022 to April 2024. The demographic and clinical information of the study participants is summarized in Table 1.



FIGURE 4

Functional architecture and training pipeline of multimodal AI diagnosis system for NMO. (A) Study dataset comprising two modalities: ultrawide-field fundus images (UWFs) and corresponding clinical examination reports. (B) Schematic of the NMO diagnostic algorithm workflow. (C) Development of the deep learning (DL) system using annotated datasets for training and validation. (D) Comparative performance evaluation of MAiDS-NMO for NMO diagnosis against clinical standards.

BOX 1 2015 IPND criteria for NMOSD with AQP4-IgG.

Criterion A: Aquaporin 4 (AQP4)-IgG-positive serostatus

Criterion B: (At least one of the following 'core characteristics' (which may be the result of one or more clinical attacks))

- 1. Clinical evidence for acute opticneuritis
- 2. Clinical evidence for acute myelitis
- 3. Clinical evidence for acute area postrema syndrome
- Clinical evidence for acute brainstem encephalitis other than area postrema syndrome
- Clinical evidence for acute diencephalitis or symptomatic narcolepsy plus MRI evidence of
- · a periependymal lesion at the level of the third ventricle, or
- a lesion in the thalamus or hypothalamus
- 6. Clinical evidence for acute (tel) encephalitis plus MRI evidence of
- an extensive periependymal lesion at the level of the lateral ventricles, or
- a large/confluent deep or subcortical white matter lesion (often with gadolinium enhancement), or
- a longitudinally extensive (≥1/2 of its length), diffuse, heterogeneous or oedematous corpus callosum lesion, or
- a longitudinally extensive (contiguously from the internal capsule to the cerebral peduncles) corticospinal tract lesion

Criterion C: Exclusion of alternative diagnoses

First, we developed a model to diagnose possible NMOSD on the basis of 1,618 UWFs from 1,055 individuals seen at the Third Affiliated Hospital of Sun Yat-sen University ophthalmology department. The images were split into training, validation and test datasets at a ratio of 80:10:10 on the basis of a stratified-group-split approach. Among these images, 330 images were associated with clinical records.

3.3 Diagnostic performance of MAiDS-NMO based on UWFs captured by SLO

Several performance metrics for selected fine-tuned weights during the training, validation and test processes without use of the clinical results. Weight No. 1 showed the best overall BOX 2 Exclusion criteria for NMOSD based on 2015IPND criteria.

Criterion A: Clinical Exclusion Criteria

- 1. Isolated Syndromes Without Core Features
- Pure brainstem symptoms (e.g., isolated vertigo)
- Isolated cortical syndromes (e.g., seizures)
- Chronic progressive myelopathy (>3 months progression)
- 2. Atypical Presentations
- · Peripheral nervous system involvement (e.g., radiculopathy)
- · Non-specific encephalopathy without MRI lesions

Criterion B: Neuroimaging Exclusion Criteria: Location Exclusionary Findings

- 1. Spinal Cord: Short-segment myelitis (<3 vertebral segments)
- 2. Optic Nerve: Unilateral involvement with <50% nerve length affected
- 3. Brain: Dawson's fingers, ovoid periventricular lesions (MS-like)
- 4. Orbit: Isolated anterior optic nerve involvement (MOGAD-pattern)
- Criterion C: Laboratory Exclusion Criteria 1. Positive MOG-IgG (by cell-based assay, CBA) with compatible
- clinical/MOGAD features
- 2. CSF-Specific Oligoclonal Bands (OCBs) (strongly suggestive of MS)
- 3. Seropositivity for Alternative Diagnoses:
- Infectious (HIV, syphilis, HTLV-1)
- Paraneoplastic (anti-CRMP5, anti-amphiphysin)
- Metabolic (vitamin B12 deficiency, copper deficiency)

Criterion D: Disease-Specific Exclusions

- 1. Multiple Sclerosis (MS):
- Meeting 2017 McDonald MRI criteria for dissemination in space/time
- "Central vein sign" on high-resolution MRI (>40% perivenous lesions)
- 2. MOG Antibody-Associated Disease (MOGAD):
- Bilateral optic neuritis with optic disc edema
- Short-segment or conus-predominant myelitis
- 3. Other Mimics:
- Spinal dural arteriovenous fistula (MRI: flow voids, cord edema)
- Neurosarcoidosis (hilar lymphadenopathy, elevated ACE)

performance among the four selected weights, with an AUC of 0.9751 and a maximum Youden's index of 0.8783 on the test dataset (Figure 5 shows the ROC curve of the model on the test

Group	NMO			non-NMO		
	Training	Validation	Test	Training	Validation	Test
Participants	219	33	33	798	127	130
Eyes	264	33	33	1,031	127	130
Images	264	33	33	1,031	127	130
Male/female	59/160	7/26	8/25	231/567	38/89	40/90
Ages	38.11 (±11.8)	41.4 (±12.1)	40.5 (±12.7)	39.9 (±13.1)	40.2 (±13.9)	41.2 (±12.5)
AQP4-IgG seropositivity	183	27	29	/	/	/
Optic nerve MRI (consistent with NMO)	171	20	23	/	/	/
Spinal cord MRI (consistent with NMO)	219	33	33	/	/	/
Underlying conditions (N)						
Hypertension	9 (4.1%)	1 (3.0%)	2 (6.0%)	81 (10.2%)	11 (8.6%)	15 (11.5%)
Diabetes	12 (5.4%)	2 (6.0%)	2 (6.0%)	90 (11.3%)	10 (7.8%)	14 (14.7%)
Ocular comorbidities (eyes)						
Glaucoma	6 (2.3%)	2 (6.0%)	3 (9.0%)	51 (4.9%)	10 (7.8%)	14 (10.7%)
Diabetic retinopathy	7 (2.6%)	2 (6.0%)	2 (6.0%)	91 (8.8%)	11 (8.6%)	15 (11.5%)
Retinal detachment	1 (0.3%)	1 (0.3%)	2 (6.0%)	33 (3.2%)	5 (3.9%)	6 (4.6%)
Retinal degeneration	30 (11.4%)	4 (12.1%)	3 (9.0%)	67 (6.5%)	13 (10.2%)	17 (13.0%)
Retinal vein obstruction	8 (3.0%)	2 (6.0%)	1 (3.0%)	27 (2.6%)	2 (1.6%)	4 (3.0%)

TABLE 1 Baseline characteristics of the study participants in the data sets.



dataset). This weight was therefore selected as the final weight for the MAiDS-NMO system based on maximization of the Youden index. The MAiDS-NMO system was then evaluated in the test dataset using the optimum cut-off threshold value (0.84) and with multimodal input data. The system achieved an AUC of 0.9923 (ROC curve shown



in Figure 6), a maximum Youden index of 0.9389, a sensitivity of 97.0% and a specificity of 96.9%.

The results obtained from five-fold cross-validation provide a more robust and reliable assessment of the MAiDS-NMO system overall predictive performance, mitigating potential biases that may arise from a single training-test split. By employing a systematic approach to data partitioning and iterative validation, this technique ensures the model's stability and generalizability across diverse data distributions, thereby enhancing its applicability and scalability in realworld clinical settings. The detailed results are presented in Table 2.

3.4 Heatmap

Heatmaps were generated and superimposed onto the fundus images to visualize the relative contributions of each pixel in predicting the grading for each image. In-NMO-positive cases, the regions exhibiting the highest signals intensity—indicating the most significant influence on the grading prediction—were predominantly localized to the optic nerve and peripapillary areas, aligning with known pathological sites of the disease. In contrast, NMO-negative cases displayed no areas of high signal intensity in the optic nerve and its surrounding regions, suggesting an absence of characteristic NMO-related features. Representative examples of these heatmaps are shown in Figure 7.

3.5 Comparative test

To verify the performance of MAiDS-NMO in assessing NMO, a comparative test involving 200 UWFs was conducted between the model and six ophthalmologists. In the initial assessment, the precision rates of the proposed system in the two tests were 90.2 and 97.8% across two tests, demonstrating stable and reliable performance of the neural network. Indicating that the neural network achieved relatively stable training results. In contrast, the accuracy rates of experts, attending physicians, and residents in the first test were 78.3,

TABLE 2 Results of five-fold cross-validation.

Fold ID	Training	Validation	Test	
Fold 1	Sensitivity: 87.8%	AUC: 0.971	AUC: 0.9734	
		Youden: 0.838	Youden: 0.8249	
	Specificity: 89.5%	Sensitivity: 88.1%	Sensitivity: 87.9%	
		Specificity: 93.1%	Specificity: 94.6%	
Fold 2	Sensitivity: 85.3%	AUC: 0.960	AUC: 0.9774	
		Youden: 0.8298	Youden: 0.8331	
	Specificity: 89.0%	Sensitivity: 75.0%	Sensitivity: 84.8%	
		Specificity: 93.8%	Specificity: 96.9%	
Fold 3	Sensitivity: 92.2%	AUC: 0.9311	AUC: 0.9641	
		Youden: 0.7208	Youden: 0.8471	
	Specificity: 93.4%	Sensitivity: 75.8%	Sensitivity: 78.9%	
		Specificity: 94.3%	Specificity: 96.9%	
Fold 4	Sensitivity: 92.4%	AUC: 0.9318	AUC: 0.9834	
		Youden: 0.726	Youden: 0.8702	
	Specificity: 94.3%	Sensitivity: 83.6%	Sensitivity: 84.8%	
		Specificity: 87.7%	Specificity: 94.6%	
Fold 5	Sensitivity: 88.0%	AUC: 0.9554	AUC: 0.9753	
		Youden: 0.782	Youden: 0.8322	
	Specificity: 90.7%	Sensitivity: 80.9%	Sensitivity: 84.8%	
		Specificity: 90.2%	Specificity: 93.8%	

56.8, and 33.3%, respectively. However, when provided with clinical information the accuracy rates improved to 98, 85.7, and 70.2%, respectively. Results of the comparative test were shown in Figure 6. These results highlight the potential of MAiDS-NMO in assisting clinical diagnosis, particularly in settings where access to specialized expertise is limited. The detailed results of the comparative test are presented in Figure 8.



FIGURE 7

Original SLO image (right) and heatmap (left). (A1) Positive case. (A2) Negative case. (B1) In the NMO-positive case, the areas of high signal intensity closely corresponded to the anatomical distribution of the optic nerve and peripapillary region. (B2) In the negative case, no areas of high signal intensity were observed in these anatomical regions.



4 Discussion

NMOSD occurs worldwide and affects individuals of all ethnic backgrounds (24). This severe and widespread disorder has a profound impact on global health, imposing a heavy global burden. Without timely treatment, approximately 50% of NMOSD patients may become disabled and blind, and one-third die within five years of their initial attack (25). Diagnosing NMOSD and monitoring treatment is challenging, burdensome, and clinically time-consuming due to the disease's variable and often insidious manifestations. Furthermore, access to essential diagnostic tools such as MRI scanning and tests for AQP4-IgG and MOG-IgG is limited in many regions, and some patients are seronegative. Thus, it may be challenging to distinguish NMOSD solely based on brain MRI at disease onset. Misdiagnosis of NMOSD is common, leading to delays in appropriate treatment and potentially resulting in severe or irreversible vision loss. Additionally, NMOSD lesions may be missed if patients are evaluated outside the acute phase or present with atypical clinical symptoms. Differences among ethnic populations, selection bias, and the availability of expert knowledge further complicate accurate diagnosis and analysis.

Screening for NMOSD, coupled with timely referral and treatment, is a widely recognized strategy for preventing blindness. Thus, there is high clinical demand for an efficient and reliable AI model that capable of overcoming the abovementioned obstacles and aiding in the early detection of NMOSD. Such a model could facilitate prompt intervention and improve patient outcomes. Deep learning algorithms have been widely applied in the diagnosis of ophthalmic diseases, demonstrating outstanding diagnostic performance (26, 27). However, to our knowledge, few studies have explored the effectiveness of DL- or AI-based methods in NMOSD diagnosis.

In this study, our AI model achieved exceptional performance in diagnosing NMOSD using UWFs as the input. First, the training dataset was constructed entirely based on the ophthalmologist's grading of all color fundus images, enabling the development of a DL algorithm with robust accuracy in identifying individuals suspected of having NMOSD at various stages. The model achieved an AUC > 0.97, sensitivity > 90%, and specificity > 92% across all datasets with available data. Second, beyond using UWFs as input, the AI model is well-suited for integrating additional clinical examinations and MRI reports of the optic nerve and spinal cord. This allows for NMOSD subtype classification and disease progression tracking, incorporating multimodal data to enhance predictive performance. Third, the model was trained on both high-quality and low-quality UWFs, ensuring its applicability to images affected by eye movement, media opacities, or other conditions that may obscure image clarity. Lastly, in a comparative test between the model and ophthalmologists, results indicated that while the neural network achieved relatively stable training results, ophthalmologists demonstrated superior diagnostic performance when provided with clinical data for reference.

The shortage of ophthalmologists and neurologists in rural areas has led to delays in NMOSD diagnosis and treatment, leaving many patients without timely medical care. UWFs is noninvasive and userfriendly tool that playing a significant role in telemedicine, proving invaluable in regions with limited access to specialized ophthalmological services. By enabling remote diagnostics and facilitating prompt medical interventions (28), UWF helps bridge healthcare gaps in underserved areas. Our AI-driven algorithm and platform, designed to diagnose NMOSD using UWF images, streamline the diagnostic process without requiring direct involvement from ophthalmologists or neurologists. This innovation has the potential to assist governments in providing more accurate and timely medical support to economically disadvantaged populations, improving healthcare accessibility and outcomes.

There are several limitations to this study. First, the dataset included a relatively small number of patients with NMSOD, largely due to the rarity of the disease in the general population. This limited sample size may have affected the model's training and its overall predictive performance. Second, all the data were derived exclusively from a single-center, hospital-based Chinese population, without representation from broader communitybased or multi-ethnic cohorts. This homogeneity, coupled with the lack of an external validation dataset, restricts the generalizability and external applicability of our findings. Validation using data from multiple centers and more diverse ethnic populations will be essential to confirm the robustness and adaptability of the model. Third, the current application is designed specifically for the early identification or onset prediction of NMOSD but does not address disease progression or recurrence was limited to onset prediction and cannot predict progression, the latter being an essential part of NMOSD management. Despite this limitation, the model still provides clinical value, especially in assisting with the challenging differential diagnosis during the early stages of the disease. Future research with larger, more diverse datasets and longitudinal follow-up will be essential to develop a fully automated and comprehensive deep learning framework capable of both diagnosing and monitoring NMOSD.

5 Conclusion

The emergence and advancement of AI have provided new opportunities for improving novel systems and strategies for detecting NMOSD. Our study demonstrated that the proposed DL system exhibits high sensitivity and specificity in identifying NMOSD patients, while also showing potential for predicting disease onset and progression. With improved computing capabilities and advanced database technologies, this model could potentially be developed into a comprehensive virtual screening system for NMOSD in clinical practice.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

SG: Writing – original draft, Conceptualization, Formal analysis. TB: Data curation, Writing – review & editing. TW: Software, Visualization, Writing – review & editing. QY: Formal analysis, Writing – review & editing. WY: Software, Writing – review & editing. JL: Software, Writing – review & editing. HZ: Methodology, Writing – review & editing. SC: Data curation, Writing – review & editing. YS: Methodology, Writing – review & editing. XJ: Methodology, Writing – review & editing. LH: Supervision, Writing – review & editing. SL: Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was funded

by Guangzhou Huangpu International Science and Technology Collaboration Project, grant number "2021GH08," Guangzhou Science and Technology Planning Project, Grant number "2024A03J0278."

Conflict of interest

QY and WY were employed by Zeemo Technology Company Limited.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Jarius S, Paul F, Weinshenker BG, Levy M, Kim HJ, Wildemann B. Neuromyelitis optica. *Nat Rev Dis Primers*. (2020) 6:85. doi: 10.1038/s41572-020-0214-9

2. Salama S, Marouf H, Ihab Reda M, Mansour AR. Clinical and radiological characteristics of neuromyelitis optica spectrum disorder in the north Egyptian Nile Delta. *J Neuroimmunol.* (2018) 324:22–5. doi: 10.1016/j.jneuroim.2018.08.014

3. Jarius S, Wildemann B. The history of neuromyelitis optica. Part 2: 'Spinal amaurosis', or how it all began. *J Neuroinflammation*. (2019) 16:280. doi: 10.1186/s12974-019-1594-1

4. Jarius S, Wildemann B. Devic's index case: a critical reappraisal - AQP4-IgGmediated neuromyelitis optica spectrum disorder, or rather MOG encephalomyelitis? *J Neurol Sci.* (2019) 407:116396. doi: 10.1016/j.jns.2019.07.014

5. Lennon VA, Wingerchuk DM, Kryzer TJ, Pittock SJ, Lucchinetti CF, Fujihara K, et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *Lancet.* (2004) 364:2106–12. doi: 10.1016/S0140-6736(04)17551-X

6. Mader S, Gredler V, Schanda K, Rostasy K, Dujmovic I, Pfaller K, et al. Complement activating antibodies to myelin oligodendrocyte glycoprotein in neuromyelitis optica and related disorders. *J Neuroinflammation*. (2011) 8:184. doi: 10.1186/1742-2094-8-184

7. Wingerchuk DM, Lennon VA, Lucchinetti CF, Pittock SJ, Weinshenker BG. The spectrum of neuromyelitis optica. *Lancet Neurol.* (2007) 6:805–15. doi: 10.1016/S1474-4422(07)70216-8

8. Nakajima H, Hosokawa T, Sugino M, Kimura F, Sugasawa J, Hanafusa T, et al. Visual field defects of optic neuritis in neuromyelitis optica compared with multiple sclerosis. *BMC Neurol.* (2010) 10:45. doi: 10.1186/1471-2377-10-45

9. Chen JJ, Flanagan EP, Jitprapaikulsan J, López-Chiriboga ASS, Fryer JP, Leavitt JA, et al. Myelin oligodendrocyte glycoprotein antibody-positive optic neuritis: clinical characteristics, radiologic clues, and outcome. *Am J Ophthalmol.* (2018) 195:8–15. doi: 10.1016/j.ajo.2018.07.020

10. Wingerchuk DM, Banwell B, Bennett JL, Cabre P, Carroll W, Chitnis T, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology*. (2015) 85:177–89. doi: 10.1212/WNL.000000000001729

11. Ankolekar A, Eppings L, Bottari F, Pinho IF, Howard K, Baker R, et al. Using artificial intelligence and predictive modelling to enable learning healthcare systems (LHS) for pandemic preparedness. *Comput Struct Biotechnol J.* (2024) 24:412–9. doi: 10.1016/j.csbj.2024.05.014

12. Chen X, Fan X, Meng Y, Zheng Y. AI-driven generalized polynomial transformation models for unsupervised fundus image registration. *Front Med.* (2024) 11:1421439. doi: 10.3389/fmed.2024.1421439

13. Keenan TDL, Chen Q, Agrón E, Tham YC, Goh JHL, Lei X, et al. DeepLensNet: deep learning automated diagnosis and quantitative classification of cataract type and severity. *Ophthalmology*. (2022) 129:571–84. doi: 10.1016/j.ophtha. 2021.12.017

14. Goh JHL, Lim ZW, Fang X, Anees A, Nusinovici S, Rim TH, et al. Artificial intelligence for cataract detection and management. *Asia Pac J Ophthalmol.* (2020) 9:88–95. doi: 10.1097/01.APO.0000656988.16221.04

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15. Zhang W, Zhao X, Chen Y, Zhong J, Yi Z. DeepUWF: an automated ultra-wide-field fundus screening system via deep learning. *IEEE J Biomed Health Inform*. (2021) 25:2988–96. doi: 10.1109/JBHI.2020.3046771

16. 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) 316:2402–10. doi: 10.1001/jama.2016.17216

17. Deng Y, Qiao L, Du M, Qu C, Wan L, Li J, et al. Age-related macular degeneration: epidemiology, genetics, pathophysiology, diagnosis, and targeted therapy. *Genes Dis.* (2022) 9:62–79. doi: 10.1016/j.gendis.2021.02.009

 Zhang H, Niu K, Xiong Y, Yang W, He Z, Song H. Automatic cataract grading methods based on deep learning. *Comput Methods Prog Biomed*. (2019) 182:104978. doi: 10.1016/j.cmpb.2019.07.006

19. Hood DC, De Moraes CG. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology.* (2018) 125:1207–8. doi: 10.1016/j.ophtha.2018.04.020

20. Hasan MM, Phu J, Wang H, Sowmya A, Meijering E, Kalloniatis M. Predicting visual field global and local parameters from OCT measurements using explainable machine learning. *Sci Rep.* (2025) 15:5685. doi: 10.1038/s41598-025-89557-1

21. Tan TE, Anees A, Chen C, Li S, Xu X, Li Z, et al. Retinal photograph-based deep learning algorithms for myopia and a blockchain platform to facilitate artificial intelligence medical research: a retrospective multicohort study. *Lancet Digit Health*. (2021) 3:e317–29. doi: 10.1016/S2S89-7500(21)00055-8

22. Aiello LP, Odia I, Glassman AR, Melia M, Jampol LM, Bressler NM, et al. Comparison of early treatment diabetic retinopathy study standard 7-field imaging with Ultrawide-field imaging for determining severity of diabetic retinopathy. *JAMA Ophthalmol.* (2019) 137:65–73. doi: 10.1001/jamaophthalmol.2018.4982

23. Calafiore GC, Gaubert S, Possieri C. Log-sum-Exp neural networks and Posynomial models for convex and log-log-convex data. *IEEE Trans Neural Netw Learn Syst.* (2020) 31:827–38. doi: 10.1109/TNNLS.2019.2910417

24. Mori M, Kuwabara S, Paul F. Worldwide prevalence of neuromyelitis optica spectrum disorders. *J Neurol Neurosurg Psychiatry.* (2018) 89:555–6. doi: 10.1136/jnnp-2017-317566

25. Huda S, Whittam D, Bhojak M, Chamberlain J, Noonan C, Jacob A. Neuromyelitis optica spectrum disorders. *Clin Med.* (2019) 19:169–76. doi: 10.7861/clinmedicine. 19-2-169

26. Li F, Su Y, Lin F, Li Z, Song Y, Nie S, et al. A deep-learning system predicts glaucoma incidence and progression using retinal photographs. *J Clin Invest.* (2022) 132:e157968. doi: 10.1172/JCI157968

27. Hasan MM, Phu J, Wang H, Sowmya A, Kalloniatis M, Meijering E. OCT-based diagnosis of glaucoma and glaucoma stages using explainable machine learning. *Sci Rep.* (2025) 15:3592. doi: 10.1038/s41598-025-87219-w

28. Nguyen TD, Le DT, Bum J, Kim S, Song SJ, Choo H. Retinal disease diagnosis using deep learning on ultra-wide-field fundus images. *Diagnostics*. (2024):14. doi: 10.3390/diagnostics14010105