biomedicines

Article

AI-Based Classification of Multiple Sclerosis Using OCT Retinal Layer Thickness Across Two Centers

Miguel Ortiz [1], Javier Dongil-Moreno [2], Gema Rebolleda [3,4], Naiara Artiaga [5,6], Luciano Boquete [2], Juan M. Miguel-Jimenez [2], Maria J. Rodrigo [5,6], Almudena López-Dorado [2], Rosario Zamora [6], Eduardo García Vicente [3,4], Eva M. Sánchez-Morla [7,8], Beatriz Andres-Luna [3], Francisco J. Muñoz Negrete [3,4] and Elena Garcia-Martin [5,6,*]

[1] School of Physics, University of Melbourne, Melbourne, VIC 3010, Australia
[2] Biomedical Engineering Group, Department of Electronics, University of Alcalá, 28801 Alcalá de Henares, Spain; jmanuel.miguel@uah.es (J.M.M.-J.); almudena.lopez@uah.es (A.L.-D.)
[3] Ophthalmology Service, Ramon y Cajal University Hospital, IRYCIS, 28034 Madrid, Spain
[4] Department of Surgery, Medical and Social Sciences, University of Alcalá, 28801 Alcalá de Henares, Spain
[5] Department of Ophthalmology, Miguel Servet University Hospital, 50009 Zaragoza, Spain; mariajesusrodrigo@hotmail.es (M.J.R.)
[6] Miguel Servet Ophthalmology Research Group (GIMSO), Aragon Health Research Institute (IIS Aragon), Biotech Vision Center, Quirón Ophthalmology Institute, University of Zaragoza, 50009 Zaragoza, Spain
[7] Institute of Psychiatry and Mental Health, Gregorio Marañón University General Hospital, IiSGM, 28007 Madrid, Spain; emsmorla@gmail.com
[8] Faculty of Medicine, Complutense University of Madrid, 28040 Madrid, Spain
* Correspondence: egmvivax@yahoo.com; Tel.: +0034-976765558; Fax: +0034-976566234

Abstract

Background: The latest revision of the McDonald criteria for diagnosis of multiple sclerosis (MS) establishes that the optic nerve can serve as a fifth anatomical location within the central nervous system for diagnosis. Optical coherence tomography (OCT) images can serve as evidence for this purpose. **Objective:** To assess the accuracy of automated artificial-intelligence-based classification of MS patients using OCT data obtained from two different centers. **Methods:** OCT data were collected from two centers using standardized APOSTEL-based protocols and similar equipment. Retinal layer thicknesses—mean and standard deviation (STD) values—were analyzed in four layers and in six regions per layer per eye. A support vector machine classifier with recursive feature elimination and Shapley additive explanations value analysis was applied to identify the most relevant features and maximize classification accuracy between control subject and MS patient eyes. **Results:** The database drawn from two hospitals comprised 112 eyes with MS without prior history of optic neuritis and 193 eyes of control subjects. The classifier achieved maximum accuracy (0.8459) using 20 input features. The mean and STD metrics had similar importance, with the most influential layers being the ganglion cell layer, inner plexiform layer, and the inner retinal layer complex. Key regions included the papillomacular bundle and the superior temporal perimacular area. **Conclusions:** OCT data facilitates highly accurate MS diagnosis across different centers. Artificial intelligence assessment could facilitate automated classification. These findings provide evidence of the important role of the optic nerve in MS diagnosis.

Keywords: multiple sclerosis; optic nerve; optical coherence tomography; artificial intelligence; support vector machine classifier; Shapley additive explanations

Check for updates

Academic Editors: Víctor M. Rivera and Lorenzo Sanesi

Received: 24 March 2026
Revised: 4 June 2026
Accepted: 24 June 2026
Published: 17 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Multiple sclerosis (MS) diagnosis is based on the McDonald criteria, which require evidence of dissemination in space (DIS) and time (DIT) of demyelinating lesions in the central nervous system (CNS). These disseminations are detected by clinical evaluation, imaging, and laboratory tests.

The recent 2024 revisions of the McDonald criteria establish that the optic nerve—assessed by MRI, visual evoked potentials, or optical coherence tomography (OCT)—can serve as a fifth anatomical location within the CNS for diagnosis, if no better explanation exists for optic nerve pathology [1].

Improvements in MS diagnostic criteria have increased sensitivity and specificity, thereby reducing the length of time between disease onset and definitive diagnosis [2]. Earlier diagnosis and consequent treatment of MS patients positively influence long-term outcomes [3,4]. It is recommended that the timeline between referral and completion of diagnostic investigation in suspected MS cases should not exceed 12 [5] or even 7 weeks [6]. However, diagnostic delays in MS are common, often due to early misdiagnosis, and are associated with greater disease burden [7], older age at onset, lower education level, and motor symptoms at onset [8].

The frequency of misdiagnosis (incorrect assignment of an MS diagnosis) ranges from 5% to 41% and that of underdiagnosis (unrecognized MS) ranges from 3% to 58% [9]. Misdiagnosis can lead to unnecessary exposure to immunosuppressive or immunomodulatory disease-modifying therapies.

OCT allows clinicians to quantify retinal layer atrophy due to neurodegeneration in MS. It is non-invasive, well tolerated by patients, performed quickly, uses relatively affordable equipment, and follows widely accepted protocols [10].

Several studies have demonstrated structural retinal alterations in MS patients both with [11,12] and without prior episodes of optic neuritis (ON). The prior literature also describes artificial-intelligence-based MS diagnosis decision support systems [13–16], with some publications that recommend the use of OCT as an additional tool to support MS diagnosis [17–21].

Validation of any automated diagnostic tool requires multicenter data to ensure reproducibility and comparability. Few studies to date have used multicenter OCT data for MS diagnosis [21].

The aim of this study—conducted across two hospitals—was to evaluate the effectiveness of an MS diagnosis system based on artificial intelligence (AI) in a multicenter framework. To this end, we ensured full comparability of OCT devices (same model, acquisition protocols, and review and calibration procedures) and analyzed the combined data.

2. Material and Methods

Patients with MS diagnosis based on the McDonald criteria were included in the study. All study participants were Caucasian. Eyes with a clinical history of optic neuritis were excluded because the aim was to find a diagnostic method for cases that are difficult to detect or for early detection. When a patient had one eye with a history of optic neuritis and the other without, only the eye without optic neuritis was included.

The exclusion criteria were clinical history of optic neuritis in this eye, axial length longer than 25.2 mm, refractive errors $\geq$5 diopters (D) of equivalent spherical or $\geq$3 D of astigmatism. The eyes had no concomitant ocular diseases, nor any previous history of retinal pathology, glaucoma, amblyopia, or systemic conditions that could affect the visual system. In order to expand the database, in most cases both subject eyes were included; therefore, the principle of independence was not strictly adhered to in the statistical tests performed.

The study was approved by the local ethics committees (C.I. PI21/113) and was conducted in accordance with the Declaration of Helsinki. Both protocols include provisions for the use of data in multicenter studies. Written informed consent was obtained from all participants prior to enrollment.

Visual acuity was measured in all eyes using the Snellen chart. The anterior segment, including the lens, was examined and intraocular pressure was measured using applanation. The fundus was also examined using OCT. In this way, it was confirmed that the eyes met the criteria and did not have any other conditions that would prevent participation in the study.

2.1. OCT Acquisition

The OCT data acquisition procedures at both centers adhered to the APOSTEL recommendations [10], employing similar OCT devices, standardized protocols, and identical calibration processes.

OCT imaging was performed by experienced neuro-ophthalmologists at both hospitals using a Heidelberg Engineering Spectralis spectral-domain device (Heidelberg Engineering, Heidelberg, Germany). The system employed automatic real-time averaging, active eye tracking, and an anatomical positioning system in a dark room environment without pupil dilation. Data were acquired using a macular volume scan with eye tracking, covering a 25° × 30° field, comprising 61 vertical B-scans with 768 A-scans per B-scan. The posterior pole asymmetry analysis protocol used represented the retinal thickness on an 8 × 8 grid [22].

2.2. Retinal Thickness Evaluation

Retinal layer segmentation was automated using the proprietary software algorithm (HEYEX v2.5.4, HRA version 6.9.a) and was manually corrected when necessary. All scans underwent quality control based on the latest OSCAR-IB criteria [23] to ensure sufficient image quality.

Consistent with previous studies [16,24] demonstrating that MS primarily affects the inner retinal layers, thickness measurements focused on the retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), and their combined structure, the inner retinal layer (IRL) complex (IRL: RNFL + GCL + IPL). These layers were analyzed within the anatomically defined regions illustrated in Figure 1.

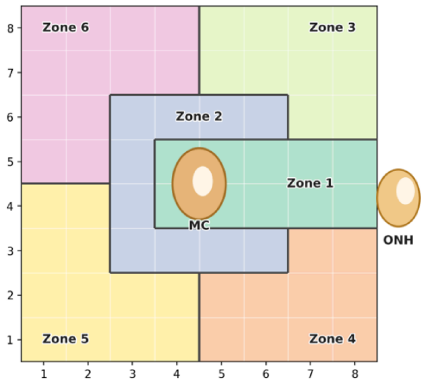


Figure 1. Representation of the areas analyzed, with ONH being the position of the optic nerve head, MC being the position of the macula, and the 6 defined areas being Zone 1 (papillomacular bundle that connects the area of maximum vision of the eye—the macula—to the optic nerve exit), Zone 2 (paramacular), Zone 3 (upper nasal perimacular), Zone 4 (lower nasal perimacular), Zone 5 (lower temporal perimacular), and Zone 6 (upper temporal perimacular). Abbreviations: ONH: optical nerve head. MC: macula.

The unit of analysis was the eye. Mean and standard deviation (STD) retinal layer thickness values were calculated for each of the 6 predefined anatomical zones and the 4 retinal structures, resulting in a total of 48 variables (6 zones × 2 metrics × 4 layers).

2.3. Statistical Analysis

Continuous variables were expressed as mean ± standard deviations (normal distribution) or median and range (nonnormal distribution), and categorical variables as numbers and percentages. Normality was assessed using the Shapiro–Wilk test for sample sizes <50 and the Kolmogorov–Smirnov test for sample sizes of ≥50. Statistical significance for continuous variables was evaluated using the non-parametric Mann–Whitney U test, while Pearson's chi-square test was applied to categorical variables. A p-value < 0.05 was considered statistically significant. OCT thickness differences between controls and MS were represented graphically using the area under the receiver operating characteristic curve (AUC).

2.4. AI Classification

OCT data from control subjects and people with multiple sclerosis (PwMS) from both hospitals were combined into a single dataset for AI classification.

Five classifier models coupled with recursive feature elimination (RFE) to enhance model reliability by removing less relevant variables and improving explainability were employed [25]. The RFE process optimized accuracy using leave-one-out (LOO) cross validation. In each iteration, feature importance was ranked using Shapley additive explanations (SHAP) [26] to calculate the mean absolute SHAP value for each feature; features with lower mean values were considered less important and subsequently removed.

The procedure was as follows:

1. Each classifier was trained using the initial set of 48 features ($F_N = 48$) and applying LOO cross validation. At each step, the model was trained on all eyes except one, which was used for validation, and SHAP values for each input feature were computed. After completing the LOO cycle, a confusion matrix and overall accuracy were calculated, along with mean SHAP values per feature.
2. The least important feature (lowest mean SHAP value) was eliminated, reducing the number of features by one ($F_N = F_N - 1$), and Step 1 was repeated until only one feature remained.

This RFE approach identified the feature subset that yielded the highest classification accuracy.

3. Results

Table 1 reports the main demographic and clinical characteristics of the PwMS and healthy controls. In the PwMS cohort, the number of patients in the MSH (Miguel Servet Hospital) group was higher than in the RCH (Ramón y Cajal Hospital) group, and there was a significant difference in both disease duration and Expanded Disability Status Scale (EDSS) score. In the control group, there was also an imbalance in the number of subjects, but no significant difference in the male-to-female ratio or the age of the participants.

Table 1. Study participant demographics and disease characteristics.

	PwMS			**Controls**		
	MSH	**RCH**	**Statics**	**MSH**	**RCH**	**Statics**
N (eyes)	98	14	-	135	58	-
Male/female eyes	21/77	2/12	$p = 0.53$	34/101	8/50	$p = 0.078$

Table 1. *Cont.*

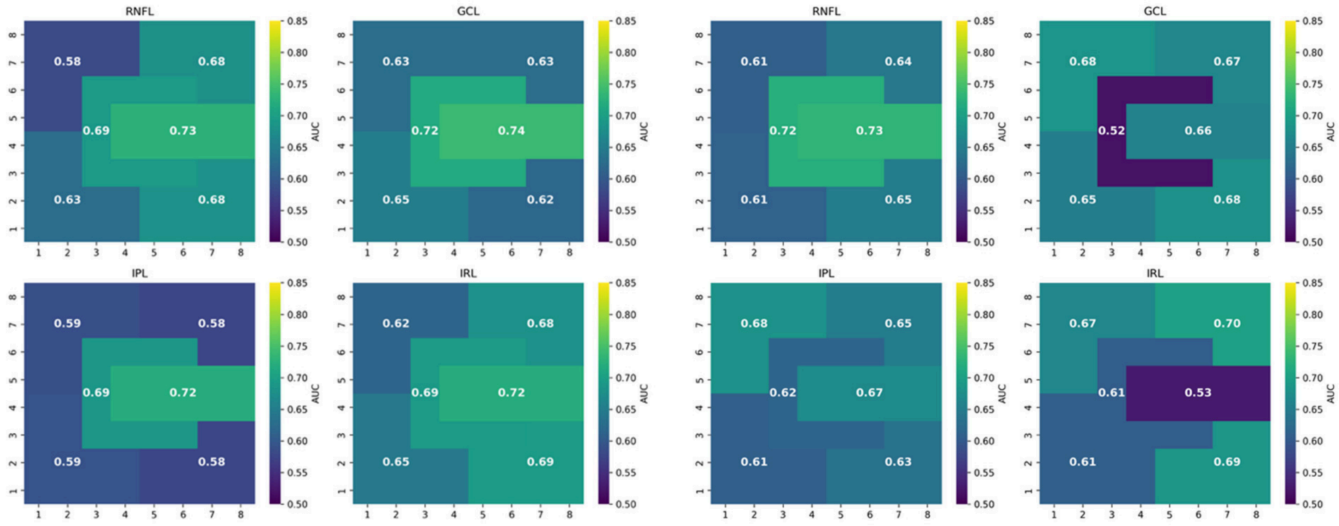
	PwMS			Controls		
	MSH	**RCH**	**Statics**	**MSH**	**RCH**	**Statics**
Age (years)	42.54 ± 10.12	35.14 ± 9.43	0.019	48.20 ± 12.22	44.77 ± 11 .89	0.051
Disease duration (years)	1.42 ± 0.72	6.82 ± 2.57	<0.001	NA	NA	NA
EDSS (median) [range]	1.28 (0–3)	2.40 (1–6.5)	<0.001	NA	NA	NA

PwMS: people with multiple sclerosis. **MSH:** Miguel Servet Hospital. **RCH:** Ramón y Cajal Hospital. **EDSS:** Expanded Disability Status Scale. **NA:** not applicable.

As indicated, OCT data from both hospitals were combined into a single dataset for classification.

3.1. OCT Thickness Comparison

Figure 2 shows the AUC values of the mean and STD variables for controls and patients (MS).



(**a**) (**b**)

Figure 2. AUC values of thickness values. (**a**) Mean variable. (**b**) STD variable.

For the mean variable, the highest AUC values in the four retinal structures analyzed are found in Zone 1, presenting values $\geq$ 0.72. The next most affected zone by AUC value is Zone 2, presenting values between 0.69 and 0.72.

For the STD variable, the greatest difference between controls and PwMS occurs in the RNFL (Zone 1 and Zone 2), followed by Zones 3 and 4 of the IRL complex.

Table 2 shows a detailed analysis of the AUC values.

Table 2. Detailed analysis of the AUC values.

	Variable: MEAN				Variable: STD				
	RNFL	**GCL**	**IPL**	**IRL**	**RNFL**	**GCL**	**IPL**	**IRL**	**Median of Each Zone**
Zone 1	0.73	0.74	0.72	0.72	0.73	0.66	0.67	0.53	0.72
Zone 2	0.69	0.72	0.69	0.69	0.72	0.52	0.62	0.61	0.69
Zone 3	0.68	0.63	0.58	0.68	0.64	0.67	0.65	0.70	0.66

Table 2. *Cont.*

	Variable: MEAN				Variable: STD				
	RNFL	**GCL**	**IPL**	**IRL**	**RNFL**	**GCL**	**IPL**	**IRL**	**Median of Each Zone**
Zone 4	0.68	0.62	0.58	0.69	0.65	0.68	0.63	0.69	0.67
Zone 5	0.63	0.65	0.59	0.65	0.61	0.65	0.61	0.61	0.62
Zone 6	0.58	0.63	0.59	0.62	0.61	0.68	0.68	0.67	0.63
Median of variable per layer	RNFL: 0.68	GCL: 0.64	IPL: 0.59	IRL: 0.69	RNFL: 0.65	GCL: 0.67	IPL: 0.64	IRL: 0.64	
Median of variable	MEDIAN: 0.67				STD: 0.65				
Median values per layer, considering variables MEAN and STD									
RNFL	0.66								
GCL	0.65								
IPL	**0.63**								
IRL	**0.68**								

3.2. Classification

To validate the choice of classifier, several machine learning methods were benchmarked using the same evaluation protocol: LOO cross-validation with RFE, applied to the combined dataset of 305 eyes from two hospitals. Five classifiers were compared: Support Vector Machine with RBF kernel (SVM-RBF), Support Vector Machine with linear kernel (SVM-Linear), Logistic Regression (LR), Linear Discriminant Analysis (LDA), and K-Nearest Neighbors (KNN). For each method, hyperparameters were tuned via grid search, and the feature subset yielding the highest overall accuracy was selected. As shown in Table 3, the SVM-RBF classifier achieved the best accuracy (84.59%) and the highest sensitivity (71.43%), confirming its suitability for this classification task. The hyperparameters of the SVM-RBF model, selected via grid search, are C = 5 and γ = 0.1.

Table 3. Performance of the 5 classifiers tested.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
SVM-RBF	84.59	71.43	92.23
SVM-Linear	79.34	52.68	94.82
LR	77.38	52.68	91.71
LDA	78.03	50.89	93.78
KNN	81.97	64.29	92.23

Figure 3 shows application of the RFE algorithm to the combined OCT dataset from both hospitals using an SVM as classifier. The graph illustrates how classification accuracy changes as the number of features decreases from the initial 48 down to 1.

The highest accuracy value (0.8459) is achieved when the classifier is fed with 20 or 17 features, resulting in the confusion matrix shown in Figure 4.

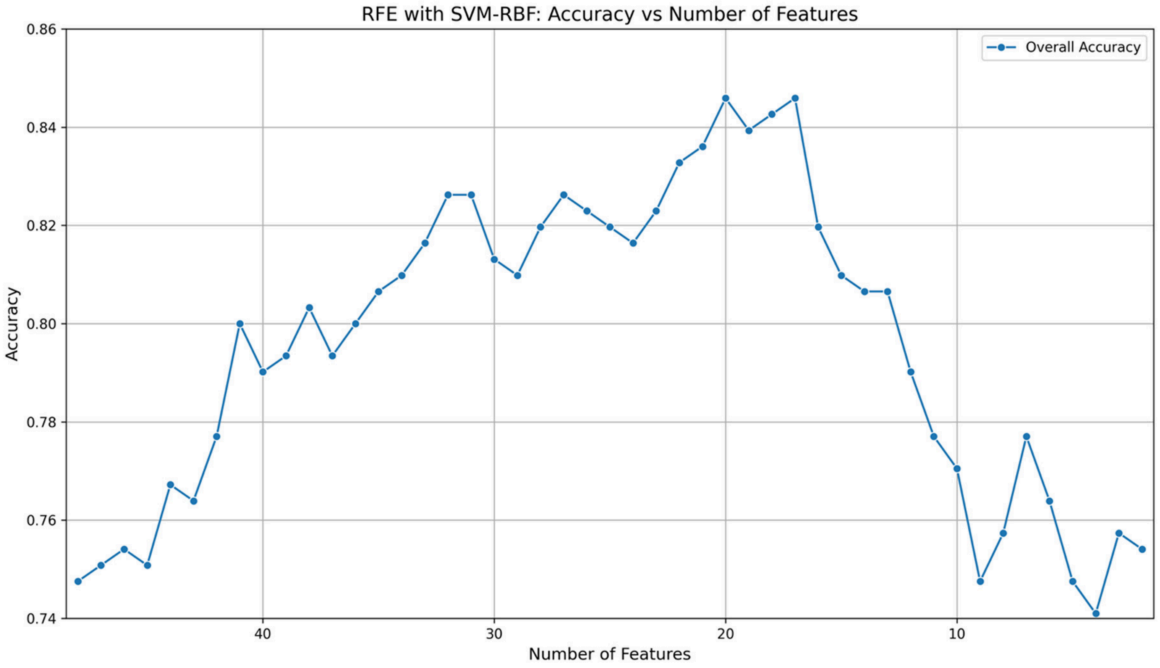


Figure 3. Representation of application of the RFE algorithm.

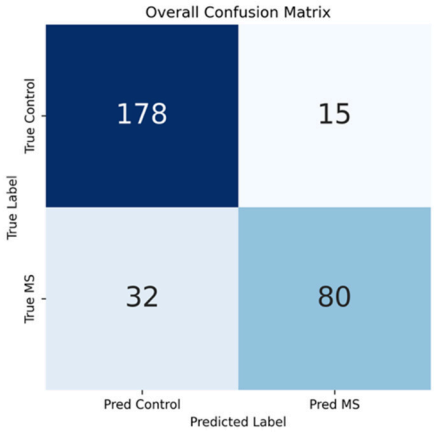


Figure 4. Confusion matrix obtained in the SVM-RBE classifier with the 20 best features selected by RFE.

Figure 5 displays the 20 features that yielded the highest classifier accuracy, ranked by SHAP importance. A cumulative analysis leads to three main conclusions:

- Metric importance: mean thickness contributes slightly more to model performance than standard deviation (weight 0.2448 vs. 0.2135).
- The most influential retinal layers, in order of importance, are the GCL (weight = 0.134), IPL (weight = 0.1207), IRL complex (weight = 0.1135), and RNFL (weight = 0.0901).
- Key anatomical regions: Zone 1 (weight = 0.1661) is the most discriminative, followed by Zone 6 (superotemporal quadrant, weight = 0.1293). The weights for the other zones are as follows: Zone 4 = 0.0755, Zone 2 = 0.0459, Zone 5 = 0.0255, and Zone 3 = 0.016.

Since, according to the SHAP values, the GCL is the most influential in the classification, in Figure 6 the box plot is presented, which allows us to appreciate how, in MS patients, this structure is thinner than in the control subjects.

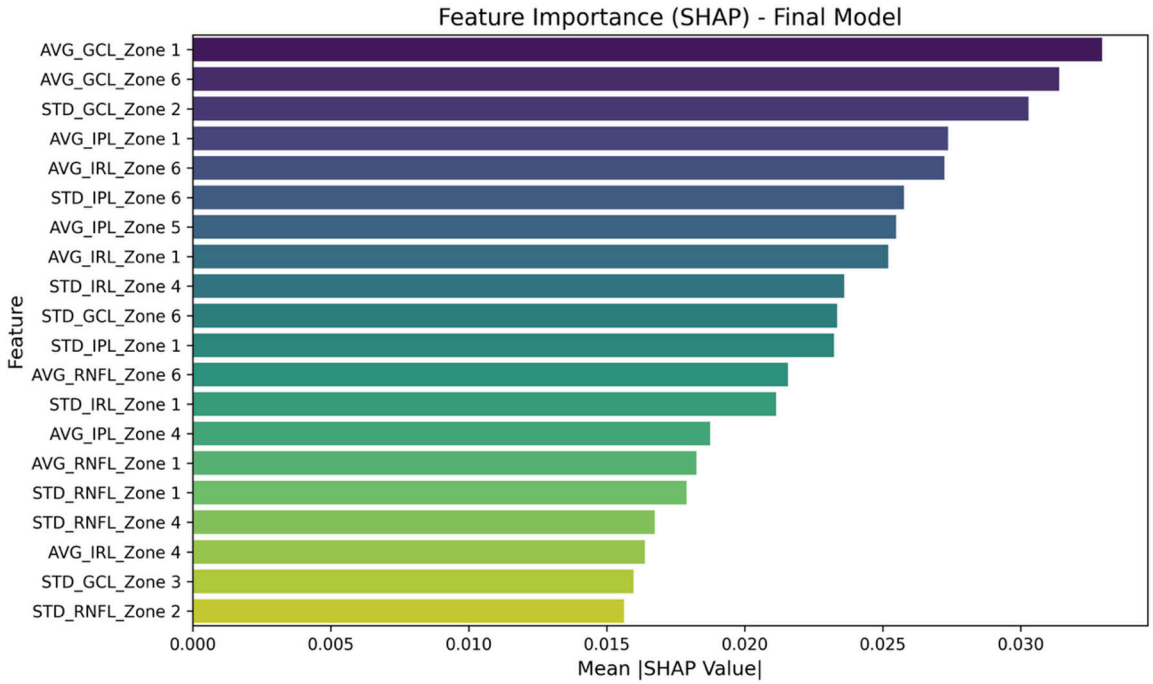


Figure 5. Ranking by SHAP value of the 20 characteristics that yield the best accuracy value.

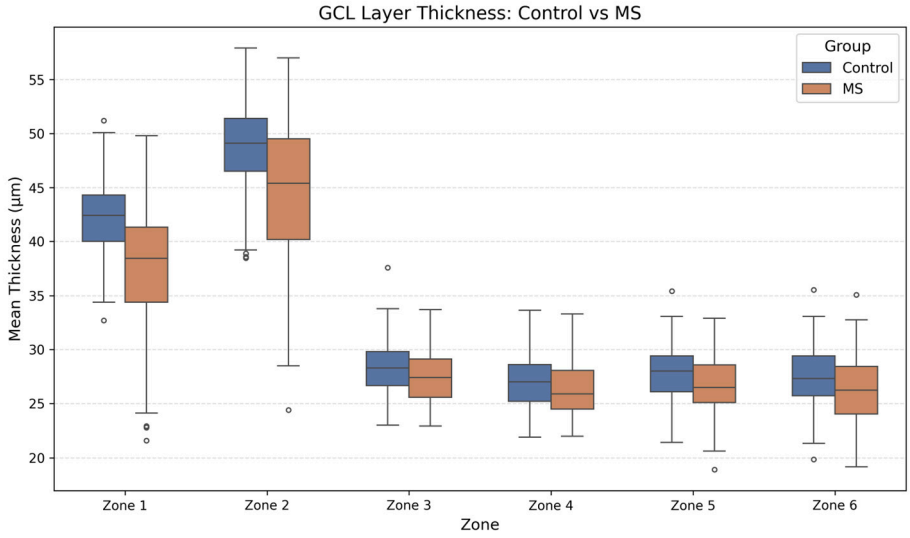


Figure 6. Representation of GCL thickness between Control and MS groups.

4. Discussion

A growing body of evidence supports the use of OCT as a reliable and non-invasive tool in MS diagnosis [18,21]. Previous studies have suggested that the retina itself may represent an early target of neurodegenerative and inflammatory processes in MS, leading to detectable neuronal damage from disease onset. In line with this, several authors have argued that OCT measurements may be particularly valuable for detecting neuroaxonal loss in early MS when conventional MRI brain atrophy metrics are less sensitive [27]. Based on the evidence, the revised 2024 McDonald criteria recognize the optic nerve as a fifth topographical location for dissemination in space [1].

In this multicenter study, we developed and validated an AI-based MS diagnosis support system using OCT data collected at two independent hospitals. Our approach achieved overall accuracy of 0.8459 while maintaining classifier decision interpretability

through SHAP analysis. These findings highlight explainable AI's potential to provide transparent diagnostic support in clinical practice.

Previous single-center studies have demonstrated that mean retinal thickness and inter-eye differences are particularly informative for detecting MS in patients without ON [16,24]. However, in our cohort, inter-eye difference could not be assessed because, in patients with unilateral ON, only the unaffected eye was included in the analysis. Interestingly, our results indicate that variability within regions, measured as the STD of retinal thickness, had similar importance to mean values (weights: 0.2448 vs. 0.2135). Although this observation requires confirmation in larger datasets, it may suggest that PwMS exhibit greater variability in retinal thickness compared with controls.

Using the mean and inter-eye difference variables, a single-center analysis [16,24] showed that the layers and regions most influencing the classifier were the GCL and Zones 2, 5, and 6. In this study, mean and STD variable analysis revealed that the layers with greatest impact on classifier performance were the GCL, IPL, and IRL, all with very similar influence, whereas the RNFL contributed less. Among these three most relevant structures, the primary regions of interest for the classifier were Zone 1 (papillomacular bundle) and Zone 6 (upper temporal perimacular).

Our findings are generally consistent with other studies. For example, [28] reported atrophy of the mRNFL and mGCIPL in early relapsing–remitting multiple sclerosis (RRMS) (mean disease duration = 23 months), while [29] identifying GCIPL (GCL+IPL) atrophy as an early feature of MS, predominantly affecting the perifoveal zone. However, our conclusions do not fully coincide, likely due to differences in patient cohorts, instrumentation, and OCT acquisition protocols. Importantly, previous studies reached their conclusions through univariate statistical analysis of each variable, whereas our results are derived from overall data analysis performed by the classifier.

The comparison between AUC values (Figure 2) and SHAP-based feature ranking (Figure 5) also deserves comment. Although there is some overlap between the variables with the highest AUC values and those selected by the classifier (Figure 5), important differences also exist. These discrepancies may arise because AUC values reflect the relationship between controls and PwMS at the individual level. More detailed analysis of AUC values (Table 2) shows that:

- The median AUC value of the four analyzed structures, considering thickness measurements (mean, STD) across the six zones, ranged between 0.63 (IPL, mean variable) and 0.68 (IRL, mean variable).
- The median AUC value of the six anatomical zones, considering mean and STD for the four structures, ranged between 0.63 (Zone 5, Zone 6) and 0.69 (Zone 1).
- When analyzing mean and STD variables globally, the median AUC values were 0.66 and 0.65, respectively.

In conclusion, AUC values are useful to identify differences between controls and PwMS in individual variables; however, their average values across metrics (mean and STD) are quite similar. SHAP analysis is preferable, since the variables are analyzed jointly and selected recursively, ranking the input variables that maximize classifier accuracy. This highlights the value of explainable AI in extracting clinically meaningful insights beyond conventional univariate analysis.

Other studies using multicenter OCT datasets have also endorsed the value of this technique in diagnosing early-stage MS. For example, in a multicenter study with participants from 28 centers [28], it was demonstrated that significant thinning of the IRL was observed in individuals with MS (disease duration > 1 year) versus controls. In an international multicenter study spanning 11 sites, inter-eye differences of 5 μm in the RNFL and of 4 μm in the macular ganglion cells + IPL were found to be a very promising criteria

for identifying unilateral optic nerve lesions [29]. In [28] the longitudinal stability of using inter-eye differences in the pRNFL and GCIPL as reliable markers of optic nerve lesions in PwMS is demonstrated. In [22], the authors used datasets from two medical centers, applying a dimensionality reduction method for feature selection combined with an SVM classifier. They identified the GCL and INL as the most influential layers; however, in one dataset, patient disease duration was not reported and, in the other, mean duration was 7.67 ± 1.37 years.

Cohort-wise evaluation showed stronger performance on the MSH dataset than on the RCH dataset. In particular, the model retained reasonable discrimination for control cases but showed reduced sensitivity for MS cases in the RCH cohort. This suggests that part of the performance degradation may be associated with inter-cohort variability, potentially including differences in acquisition protocol, image quality, population characteristics, or annotation distribution. These results highlight the importance of external validation and suggest that future work should focus on improving robustness through cohort harmonization, domain adaptation, or fine-tuning using additional data from under-represented acquisition settings.

Nevertheless, our study has limitations. The sample sizes from each center were not balanced, and there were differences in the disease durations between hospitals that may have influenced the results, and mismatched EDSS scores indicate clinical heterogeneity across centers. In most cases, both eyes of each subject have been considered, so the principle of independence of the statistical tests used is not strictly met. Although clinical histories were carefully reviewed, the possibility of subclinical ON episodes in the included eyes cannot be excluded. Moreover, our analysis was restricted to structural OCT data acquired with a single device and protocol.

Future studies should explore whether combining structural metrics (mean, STD, and inter-eye differences) with complementary modalities such as OCT angiography and visual evoked potentials could further enhance diagnostic accuracy. Additionally, if cohort size allows, it would be of interest to investigate the combined behavior of these three structural variables (mean, STD, and inter-eye difference) in AI-based diagnostics [30].

In conclusion, our results suggest that OCT, when analyzed using explainable AI methods, can support highly accurate MS diagnosis in a multicenter setting.

Author Contributions: Conceptualization, L.B. and E.G.-M. Methodology, R.Z., M.J.R., G.R., E.G.V., L.B., M.O., J.D.-M., J.M.M.-J., B.A.-L., N.A. and A.L.-D. Formal analysis, L.B., M.O., J.D.-M., J.M.M.-J. and A.L.-D. Investigation, G.R., E.G.V., N.A. and B.A.-L. Resources, E.G.-M. and F.J.M.N. Writing—original draft, L.B., E.M.S.-M. and E.G.-M. Writing—review & editing, L.B., F.J.M.N. and E.G.-M. Funding acquisition, E.G.-M. and F.J.M.N. Data curation, M.O., J.D.-M., J.M.M.-J. and A.L.-D. Software, L.B., M.O., J.D.-M., J.M.M.-J. and A.L.-D. Validation, L.B., E.M.S.-M., N.A. and B.A.-L. Visualization, M.O., J.D.-M. and J.M.M.-J. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by JR22/00057, PI23/00935, RD21/0002/0050, RD21/0002/0018, RD24-0007-0022, and RD24-0007-0020 (Carlos III Health Institute); by PID2023-148219OB-C22 (MINECO/AEI/ERDF, EU) funded by MCIN/AEI/10.13039/501100011033; and by the Government of Aragon through group B23_23R and PROY_B50_24.

Institutional Review Board Statement: Ethics approval and consent to participate: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee (CEI) (protocol code PI21/113, date of approval: 20 April 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available upon reasonable request to the corresponding author, due to ethical reasons.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Montalban, X.; Lebrun-Frénay, C.; Oh, J.; Arrambide, G.; Moccia, M.; Amato, M.P.; Amezcua, L.; Banwell, B.; Bar-Or, A.; Barkhof, F.; et al. Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *Lancet Neurol.* **2025**, *24*, 850–865, Correction in *Lancet Neurol.* **2025**, *24*, E13. https://doi.org/10.1016/S1474-4422(25)00355-2. Erratum in *Lancet Neurol.* **2025**, *24*, 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4.. [CrossRef] [PubMed]
2. Tintore, M.; Cobo-Calvo, A.; Carbonell, P.; Arrambide, G.; Otero-Romero, S.; Río, J.; Tur, C.; Comabella, M.; Nos, C.; Arévalo, M.J.; et al. Effect of Changes in MS Diagnostic Criteria Over 25 Years on Time to Treatment and Prognosis in Patients with Clinically Isolated Syndrome. *Neurology* **2021**, *97*, e1641–e1652. [CrossRef] [PubMed]
3. Cobo-Calvo, A.; Tur, C.; Otero-Romero, S.; Carbonell-Mirabent, P.; Ruiz, M.; Pappolla, A.; Alvarez, J.V.; Vidal-Jordana, A.; Arrambide, G.; Castilló, J.; et al. Association of Very Early Treatment Initiation with the Risk of Long-term Disability in Patients With a First Demyelinating Event. *Neurology* **2023**, *101*, e1280–e1292. [CrossRef] [PubMed]
4. Comi, G.; Radaelli, M.; Sørensen, P.S. Evolving concepts in the treatment of relapsing multiple sclerosis. *Lancet* **2017**, *389*, 1347–1356. [CrossRef] [PubMed]
5. NICE. Multiple Sclerosis in Adults: Management NICE Guideline Reference Number: NG220. 2022. Available online: https://www.nice.org.uk/Guidance/Ng220 (accessed on 20 March 2026).
6. Hobart, J.; Bowen, A.; Pepper, G.; Crofts, H.; Eberhard, L.; Berger, T.; Boyko, A.; Boz, C.; Butzkueven, H.; Celius, E.G.; et al. International consensus on quality standards for brain health-focused care in multiple sclerosis. *Mult. Scler. J.* **2019**, *25*, 1809–1818. [CrossRef] [PubMed]
7. Uher, T.; Adzima, A.; Srpova, B.; Noskova, L.; Maréchal, B.; Maceski, A.M.; Krasensky, J.; Stastna, D.; Andelova, M.; Novotna, K.; et al. Diagnostic delay of multiple sclerosis: Prevalence, determinants and consequences. *Mult. Scler. J.* **2023**, *29*, 1437–1451. [CrossRef] [PubMed]
8. Patti, F.; Chisari, C.G.; Arena, S.; Toscano, S.; Finocchiaro, C.; Fermo, S.L.; Judica, M.L.; Maimone, D. Factors driving delayed time to multiple sclerosis diagnosis: Results from a population-based study. *Mult. Scler. Relat. Disord.* **2021**, *57*, 103361. [CrossRef] [PubMed]
9. Zürrer, W.E.; Cannon, A.E.; Ilchenko, D.; Gaitán, M.I.; Granberg, T.; Piehl, F.; Solomon, A.J.; Ineichen, B.V. Misdiagnosis and underdiagnosis of multiple sclerosis: A systematic review and meta-analysis. *Mult. Scler. J.* **2024**, *30*, 1409–1422. [CrossRef] [PubMed]
10. Aytulun, A.; Cruz-Herranz, A.; Aktas, O.; Balcer, L.J.; Balk, L.; Barboni, P.; Blanco, A.A.; Calabresi, P.A.; Costello, F.; Sanchez-Dalmau, B.; et al. APOSTEL 2.0 Recommendations for Reporting Quantitative Optical Coherence Tomography Studies. *Neurology* **2021**, *97*, 68–79. [CrossRef] [PubMed]
11. Mauceri, V.A.; Puthenparampil, M.; Gallo, P.; Petzold, A. Optic nerve as a 5th location in the revised McDonald diagnostic criteria for multiple sclerosis: Limitations of OCT in the acute phase. *BMJ Neurol. Open* **2025**, *7*, e001118. [CrossRef] [PubMed]
12. Savastano, M.C.; Nociti, V.; Giannuzzi, F.; Cestrone, V.; Carlà, M.M.; Fossataro, C.; Biagini, I.; Rizzo, C.; Kilian, R.; Bisurgi, M.; et al. Optical Coherence Tomography Advanced Parameters in Patients with Multiple Sclerosis: Ophthalmological and Neurological Assessments. *Am. J. Ophthalmol.* **2024**, *267*, 41–49. [CrossRef] [PubMed]
13. Cavaliere, C.; Vilades, E.; Alonso-Rodríguez, M.C.; Rodrigo, M.J.; Pablo, L.E.; Miguel, J.M.; López-Guillén, E.; Morla, E.M.S.; Boquete, L.; Garcia-Martin, E. Computer-Aided Diagnosis of Multiple Sclerosis Using a Support Vector Machine and Optical Coherence Tomography Features. *Sensors* **2019**, *19*, 5323. [CrossRef] [PubMed]
14. López-Dorado, A.; Ortiz, M.; Satue, M.; Rodrigo, M.J.; Barea, R.; Sánchez-Morla, E.M.; Cavaliere, C.; Rodríguez-Ascariz, J.M.; Orduna-Hospital, E.; Boquete, L.; et al. Early Diagnosis of Multiple Sclerosis Using Swept-Source Optical Coherence Tomography and Convolutional Neural Networks Trained with Data Augmentation. *Sensors* **2021**, *22*, 167. [CrossRef] [PubMed]
15. Ortiz, M.; Mallen, V.; Boquete, L.; Sánchez-Morla, E.M.; Cordón, B.; Vilades, E.; Dongil-Moreno, F.J.; Miguel-Jiménez, J.M.; Garcia-Martin, E. Diagnosis of multiple sclerosis using optical coherence tomography supported by artificial intelligence. *Mult. Scler. Relat. Disord.* **2023**, *74*, 104725. [CrossRef] [PubMed]
16. Ortiz, M.; Pueyo, A.; Dongil, F.J.; Boquete, L.; Sánchez-Morla, E.M.; Barea, R.; Miguel-Jimenez, J.M.; López-Dorado, A.; Vilades, E.; Rodrigo, M.J.; et al. New Method of Early RRMS Diagnosis Using OCT-Assessed Structural Retinal Data and Explainable Artificial Intelligence. *Transl. Vis. Sci. Technol.* **2025**, *14*, 14. [CrossRef] [PubMed]
17. Brownlee, W.J.; Miszkiel, K.A.; Tur, C.; Barkhof, F.; Miller, D.H.; Ciccarelli, O. Inclusion of optic nerve involvement in dissemination in space criteria for multiple sclerosis. *Neurology* **2018**, *91*, e1130–e1134. [CrossRef] [PubMed]
18. Bsteh, G.; Hegen, H.; Altmann, P.; Auer, M.; Berek, K.; Di Pauli, F.; Kornek, B.; Krajnc, N.; Leutmezer, F.; Macher, S.; et al. Diagnostic Performance of Adding the Optic Nerve Region Assessed by Optical Coherence Tomography to the Diagnostic Criteria for Multiple Sclerosis. *Neurology* **2023**, *101*, e784–e793. [CrossRef] [PubMed]

19. Foster, M.A.; Pontillo, G.; Davagnanam, I.; Collorone, S.; Prados, F.; Kanber, B.; Yiannakas, M.C.; Ogunbowale, L.; Burke, A.; Wheeler-Kingshott, C.A.M.G.; et al. Improving criteria for dissemination in space in multiple sclerosis by including additional regions. *Ann. Clin. Transl. Neurol.* **2024**, *11*, 2572–2582. [CrossRef] [PubMed]
20. Galetta, S.; Brownlee, W. Assignment Completed. The Optic Nerve Should Graduate to Be the Fifth Lesion Site for the Diagnosis of Multiple Sclerosis. *Neurology* **2024**, *102*, e207913. [CrossRef] [PubMed]
21. Vidal-Jordana, A.; Rovira, A.; Calderon, W.; Arrambide, G.; Castilló, J.; Moncho, D.; Rahnama, K.; Collorone, S.; Toosy, A.T.; Ciccarelli, O.; et al. Adding the Optic Nerve in Multiple Sclerosis Diagnostic Criteria. *Neurology* **2024**, *102*, e200805, Correction in *Neurology* **2024**, *102*, e209214. https://doi.org/10.1212/WNL.0000000000209214. [CrossRef] [PubMed]
22. Khodabandeh, Z.; Rabbani, H.; Ashtari, F.; Zimmermann, H.G.; Motamedi, S.; Brandt, A.U.; Paul, F.; Kafieh, R. Discrimination of multiple sclerosis using OCT images from two different centers. *Mult. Scler. Relat. Disord.* **2023**, *77*, 104846. [CrossRef] [PubMed]
23. Asrani, S.; Rosdahl, J.A.; Allingham, R.R. Novel software strategy for glaucoma diagnosis: Asymmetry analysis of retinal thickness. *Arch. Ophthalmol.* **2011**, *129*, 1205–1211. [CrossRef] [PubMed]
24. Petzold, A.; Albrecht, P.; Balcer, L.; Bekkers, E.; Brandt, A.U.; Calabresi, P.A.; Deborah, O.G.; Graves, J.S.; Green, A.; Keane, P.A.; et al. Artificial intelligence extension of the OSCAR-IB criteria. *Ann. Clin. Transl. Neurol.* **2021**, *8*, 1528–1542. [CrossRef] [PubMed]
25. Dongil-Moreno, F.J.; Ortiz, M.; Pueyo, A.; Boquete, L.; Sánchez-Morla, E.M.; Jimeno-Huete, D.; Miguel, J.M.; Barea, R.; Vilades, E.; Garcia-Martin, E. Diagnosis of multiple sclerosis using optical coherence tomography supported by explainable artificial intelligence. *Eye* **2024**, *38*, 1502–1508. [CrossRef] [PubMed]
26. Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* **2002**, *46*, 389–422. [CrossRef]
27. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS2017), Long Beach, CA, USA, 4–9 December 2017.
28. Paul, F.; Calabresi, P.A.; Barkhof, F.; Green, A.J.; Kardon, R.; Sastre-Garriga, J.; Schippling, S.; Vermersch, P.; Saidha, S.; Gerendas, B.S.; et al. Optical coherence tomography in multiple sclerosis: A 3-year prospective multicenter study. *Ann. Clin. Transl. Neurol.* **2021**, *8*, 2235–2251. [CrossRef] [PubMed]
29. Pietroboni, A.M.; Carandini, T.; Dell'Arti, L.; Bovis, F.; Colombi, A.; De Riz, M.A.; Casazza, E.; Scola, E.; Fenoglio, C.; Arighi, A.; et al. Evidence of retinal anterograde neurodegeneration in the very early stages of multiple sclerosis: A longitudinal OCT study. *Neurol. Sci.* **2020**, *41*, 3175–3183. [CrossRef] [PubMed]
30. Nolan-Kenney, R.C.; Liu, M.; Akhand, O.; Calabresi, P.A.; Paul, F.; Petzold, A.; Balk, L.; Brandt, A.U.; Martínez-Lapiscina, E.H.; Saidha, S.; et al. Optimal intereye difference thresholds by optical coherence tomography in multiple sclerosis: An international study. *Ann. Neurol.* **2019**, *85*, 618–629. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.