Neurosymbolic AI approach for dry and wet AMD classification using OCT images

Aleksandar Miladinović[1], Alessandro Biscontin[1], Miloš Ajčević[2], Agostino Accardo[2], Dario Marangoni[3], Daniele Tognetto[3], and Leandro Inferrera[3]

[1] Institute for Maternal and Child Health IRCCS "Burlo Garofolo", Trieste, Italy
[2] Department of Engineering and Architecture, University of Trieste, Trieste, Italy
[3] Department of Medicine, Surgery and Health Sciences, Ophthalmology Clinic, University of Trieste, Trieste, Italy

Abstract. Accurate classification between dry and wet age-related macular degeneration (AMD) is crucial for determining appropriate treatment strategies and improving patient outcomes. However, deep learning convolutional neural networks (CNNs) often face challenges when working with small datasets, particularly in the context of rare pathologies, which can hinder their robustness and generalizability. To address these issues, we employed a neurosymbolic approach that integrates medical knowledge in the form of symbolic AI, enhancing the model's interpretability and reasoning capabilities. The aim of this study was to improve the classification of retinal conditions, specifically dry AMD, wet AMD, and normal retinas. Our results demonstrated an overall accuracy of 93%, indicating the effectiveness of this methodology in accurately classifying retinal diseases. These findings suggest that the neurosymbolic approach holds promise for advancing diagnostic support in ophthalmology while providing a transparent decision-making framework.

Keywords: Neurosymbolic AI, Optical Coherence Tomography, OCT,Retinal imaging

1 Introduction

Neurosymbolic learning represents a significant advancement in artificial intelligence, particularly in the medical domain, where decision-making processes must often mirror human reasoning. This approach integrates the strengths of deep learning, specifically convolutional neural networks (CNNs) and Deep Learning (DL), with symbolic reasoning, thereby creating a more interpretable and transparent framework for classification tasks.

Until today, DL is the most common type of architecture used for the classification and management of various medical conditions due to its ability to identify pathological features from images effectively [1, 2]. Specifically, such algorithms serve as robust tools for the automatic identification and quantification of retinal abnormality singns observable in optical coherence tomography (OCT) images [3–5]. Recently, a variety of ML models have been developed

and extensively utilized for analyzing OCT images from patients suffering from significant ocular conditions, including diabetic retinopathy (DR), age-related macular degeneration (AMD), central serous chorioretinopathy (CSC), epiretinal membranes (ERM), and glaucoma [6–16].

In terms of OCT image classification, the most commonly employed convolutional neural network (CNN) architectures, such as VGG, ResNet, Inception and their derivatives, have demonstrated highly promising performance [17–21]. For example, a study conducted by Inferrera et al. reported impressive accuracy rates (between 93% and 99%) when identifying conditions like epiretinal membranes, intraretinal fluid, and macular neovascularization [22].

While CNNs excel at detecting patterns and abnormalities, they often fall short in providing a transparent model [23]. This limitation can be particularly pronounced when dealing with rare conditions, where CNNs may struggle due to small datasets and insufficient training examples. In contrast, the neurosymbolic framework allows for the integration of medical knowledge in the form of rules, which enhances the model's performance and explanability of the model [24]. This is particularly important in ophthalmology, where certain retinal conditions may present infrequently, making it challenging for standard CNNs to learn effectively from limited examples [25] and incorporating additional medical knowledge into the model may enhance its performance.

Accurate differentiation between dry AMD and wet AMD is crucial for effective patient management and treatment planning. Wet AMD is characterized by the development of macular neovascularization (MNV), which can lead to rapid vision loss due to fluid accumulation and subsequent retinal damage [26]. In contrast, dry AMD typically progresses more slowly and is associated with the accumulation of drusen (D) and retinal pigment epithelium changes, which may not lead to immediate vision impairment but can result in geographic atrophy over time [27]. By analyzing the presence or absence of specific retinal abnormalities, clinicians can accurately determine the final pathology. This method aligns with the principles of neurosymbolic learning, where integrating domain-specific knowledge allows for more informed decision-making and improved diagnostic outcomes [22, 28].

Therefore, the aim of this study is to demonstrate how the innovative integration of deep learning and symbolic reasoning (neurosymbolic approach) can enhance the classification of dry AMD, wet AMD, and normal retinas, not only improving model performance but also increasing explainability and interpretability, thereby providing deeper insights into the decision-making process.

2 Methods

This retrospective observational study was conducted at the University Eye Clinic of Trieste, following the ethical guidelines of the Declaration of Helsinki. The study protocol received approval from the Regional Ethics Committee (CEUR) of Friuli Venezia Giulia, Italy (protocol number 17094/2022). All participants gave their informed consent for the use of their data.

We retrospectively evaluated anonymized optical coherence tomography (OCT) scans obtained using an A-line scan protocol with a length of 9.0 mm. The images were captured using the Spectralis OCT system (Heidelberg Engineering, Heidelberg, Germany), which operates at a central wavelength of 815 nm and offers an axial resolution of 3.9 µm/pixel, a lateral resolution of 5.7 µm/pixel, and an image dimension of 768x496 pixels. The dataset included both horizontal and vertical line scans centered on the fovea, collected from individuals aged 18 to 95 years between January 2017 and September 2022.

Finally, 1407 OCT images of dry AMD, wet AMD affected and Normal retinas (469 for each class) were collected.

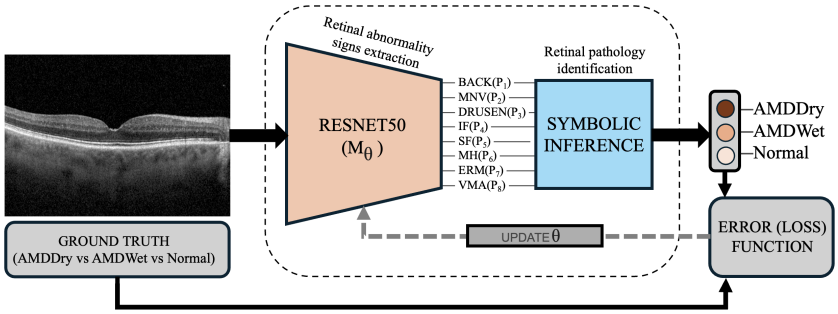

Fig. 1. Illustration of the learning process in the neurosymbolic framework.

2.1 Neurosymbolic AI Reasoning

In this study, for the neurosymbolic AI reasoning, we employed the Scallop programming framework [29]. Scallop allows an integration of logical rules (i.e. expert knowledge) with subsymbolic machine learning approaches (i.e. neural networks, in our case CNN), in a structured manner. This capability is particularly beneficial in medical applications, where domain-specific knowledge can significantly enhance the interpretability and performance of AI decision support systems.

Scallop utilizes provenance tracking to support differentiable reasoning by associating weights, probabilities, or gradients with facts and rules. For our implementation (see Figure 1), we utilized Differentiable Max-Min Probability Provenance (DMMP), designed to support gradient-based optimization in probabilistic reasoning systems. Unlike traditional provenance systems, DMMP employs dual-number representation, where each provenance tag is augmented with a derivative component. This allows the system to track changes in probability values while enabling efficient backpropagation in neurosymbolic learning models [30].

2.2 Expert knowledge representation

In Scallop, expert knowledge is represented in the form of logical rules. These rules, defined by ophthalmologists, encapsulate the symbolic reasoning behind the classification of dry AMD, wet AMD, and Normal retinas. For this work, we define the following rules that capture domain-specific knowledge:

$$\begin{aligned}\text{DryAMD}\big((P_3 \wedge \neg P_4 \wedge \neg P_2) \wedge (P_3 \wedge \neg P_5 \wedge \neg P_2) \wedge (P_1 \vee P_7)\big)\text{:-}\\ \text{BS}(P_1),\quad \text{MNV}(P_2),\quad \text{D}(P_3),\quad \text{IF}(P_4),\quad \text{SF}(P_5),\\ \text{MH}(P_6),\quad \text{ERM}(P_7),\quad \text{VMA}(P_8).\end{aligned} \tag{1}$$

$$\begin{aligned}\text{WetAMD}\big(((P_3 \wedge P_4) \vee (P_3 \wedge P_5) \vee (P_3 \wedge \neg P_7)) \wedge (P_2 \wedge P_1)\big)\text{:-}\\ \text{BACK}(P_1),\quad \text{MNV}(P_2),\quad \text{D}(P_3),\quad \text{IF}(P_4),\quad \text{SF}(P_5),\quad \text{MH}(P_6),\\ \text{ERM}(P_7),\quad \text{VMA}(P_8).\end{aligned} \tag{2}$$

$$\begin{aligned}\text{Normal}\big(\neg P_1 \wedge \neg P_2 \wedge \neg P_3 \wedge \neg P_4 \wedge \neg P_5 \wedge \neg P_7\big)\text{:-}\\ \text{BACK}(P_1),\quad \text{MNV}(P_2),\quad \text{D}(P_3),\quad \text{IF}(P_4),\quad \text{SF}(P_5),\quad \text{MH}(P_6),\\ \text{ERM}(P_7),\quad \text{VMA}(P_8).\end{aligned} \tag{3}$$

where *AMDDry*, *AMDWet*, and *Normal* represent retinal pathologies; *BACK*, *MNV*, *D*, *IF*, *SF*, *MH*, *ERM*, and *VMA* denote retinal abnormality signs; and P_n represents the probability of the presence of a corresponding retinal abnormality. In particular, the rule (1) for dry AMD is predicated on the identification of Drusen (D) as a hallmark characteristic of this condition. The absence of Intraretinal Fluid (IF), Subretinal Fluid (SF) and Macular Neovascularization (MNV), are more commonly associated with wet AMD. The presence of Choroidal Backscattering (BACK) may occur in later stages AMD [31].

In contrast, the rule (2) for wet AMD is constructed to reflect the condition's defining characteristics, which include the presence of Drusen (D) and Macular Neovascularization (MNV) in conjunction with either Intraretinal Fluid (IF) or Subretinal Fluid (SF) [28].

For the classification of a normal retina, the rule is explicitly defined by the absence of all key pathological features [22].

Table 1 presents an illustrative example demonstrating how the rule engine operates using rules (1)–(2) and Differentiable Max-Min Probability Provenance (DMMP) for classification. The table showcases how different probabilities of retinal abnormality signs influence the classification outcomes for dry AMD, wet AMD, and Normal categories.

The objective of deep-learning CNNs in a neurosymbolic framework is to estimate the probability of the presence of these retinal signs abnormalities, even though they are not explicitly trained or supervised on them.

2.3 CNN Deep-Learning Architecture

For the deep learning component of our model, we utilized the ImageNet pretrained ResNet50 architecture. The final layer of the model was modified to provide independent probabilities for each of the eight retinal abnormality signs.

Table 1. Illustrative example demonstrating how the inference engine operates using rules (1)–(2) and Differentiable Max-Min Probability Provenance (DMMP) for classification.

P(Retinal Abnormality Signs)								**P(Classification Outcome)**		
BS	MNV	D	IF	SF	MH	MER	VMA	AMDDry	AMDWet	Normal
0.80	0.20	0.90	0.10	0.10	0.30	0.80	0.20	**0.75**	0.42	0.01
0.60	0.70	0.60	0.70	0.80	0.20	0.10	0.20	0.14	**0.71**	0.00
0.10	0.00	0.10	0.00	0.00	0.10	0.20	0.10	0.08	0.02	**0.65**

This adjustment allows the model to perform multi-label classification, enabling it to predict the presence of multiple retinal abnormalities simultaneously. The ResNet50 model was fine-tuned to classify dry AMD, wet AMD, and Normal retinas, rather than relying solely on direct supervision of the retinal signs (see Figure 1). The gradient of the pre-trained ResNet50 was updated using the symbolic rules defined earlier, which represent the medical knowledge encoded in the system. The total OCT dataset was partitioned into 80% for training, 10% for validation, and 10% for testing to ensure a balanced evaluation of the model's performance. During training, the following learning parameters were employed: a learning rate of 0.0001, the Binary Cross-Entropy (BCE) loss function, and a total of 15 epochs.

3 Results

The performance of the neurosymbolic learning model was evaluated on the test set. The confusion matrix (Table 2) shows the classification performance for each category: dry AMD, wet AMD, and Normal.

Table 2. Confusion matrix on the test set, expressed as percentages relative to the total number of samples in each class.

	Predicted Classification		
	dry AMD	**wet AMD**	**Normal**
dry AMD	1.00	0.00	0.00
wet AMD	0.21	0.79	0.00
Normal	0.00	0.00	1.00

The detailed classification metrics, including precision, recall, and F1-score, are summarized in Table 3.

The model achieved an overall accuracy of 93% across all classes. The precision for dry AMD was 0.80, reflecting that 80% of the instances predicted as dry AMD were correct. The recall for dry AMD was at 1.00, indicating that all actual dry AMD cases were identified. For wet AMD, the precision was 1.00,

Table 3. Summary of classification metrics, including precision, recall, and F1-score for each class.

Classification Metrics Summary			
Class	**Precision**	**Recall**	**F1-Score**
dry AMD	0.80	1.00	0.89
wet AMD	1.00	0.79	0.88
Normal	1.00	1.00	1.00

while the recall was lower at 0.79, suggesting that some wet AMD cases were misclassified as dry AMD. Normal cases were classified with precision, recall and F1 of 1.0. None of the pathological retinas were exchanged for healthy ones.

4 Discussion

The implementation of a neurosymbolic approach in this study highlights the critical need for advanced AI methodologies in the medical field, particularly in ophthalmology. Differentiating between wet AMD and dry AMD is essential from a clinical perspective, as these two conditions require distinct treatment strategies and have different prognoses. Accurate classification of these conditions can significantly impact patient outcomes and treatment efficacy.

The neurosymbolic approach combines the strengths of deep learning with symbolic reasoning, allowing for a more robust and interpretable classification system. By integrating domain knowledge into the model, we can create classifiers that are not only accurate but also capable of reasoning in a manner similar to that of an ophthalmologist. This modularized learning process enables the convolutional neural network (CNN) to identify structural abnormalities, even when not explicitly supervised.

The aim of this work was to demonstrate how a neurosymbolic approach can improve the classification of retinal conditions, specifically dry AMD, wet AMD, and Normal, using a dataset of 469 images per class. The results indicated an overall accuracy of 93%, with high precision and recall for each class, underscoring the effectiveness of this methodology.

Moreover, the neurosymbolic approach is particularly advantageous for addressing rare pathologies where datasets may not be large enough to train robust deep learning CNNs. In such cases, the ability to incorporate expert knowledge into the model can enhance its performance and reliability, providing a pathway to better diagnose conditions that are infrequently encountered in clinical practice.

The neurosymbolic innovative approach contrasts sharply with traditional black-box models, which often lack transparency and interpretability. By providing a clear rationale for its decisions, the neurosymbolic model fosters trust and understanding among clinicians, potentially leading to higher adoption rates in clinical settings.

However, a key limitation of this study lies in the simplicity of the defined rules, which may not always capture the full complexity of retinal abnormality signs. While the proposed approach effectively integrates algorithmic supervision for feature extraction, further validation is needed to ensure that the extracted features are accurately identified. In this study, this limitation was mitigated by the availability of both final pathology labels and individual abnormality signs, allowing for cross-verification.

As this is a preliminary study, future work could explore how refining and expanding the rule set might enhance the localization of retinal abnormalities. More sophisticated rule definitions, such as assigning greater weight to certain signs when making classifications, could improve the model's ability to discern pathological features. Additionally, if multiple signs are assigned equal importance for a given pathology, the model might arbitrarily attribute the classification to one of them. While this simplification was intentional for the initial study, it remains an area for future refinement.

Overall, this work serves as an initial demonstration of the potential neurosymbolic approaches hold for medical imaging, balancing the advantages of deep learning with enhanced interpretability through symbolic reasoning.

In conclusion, the integration of symbolic reasoning with deep learning through a neurosymbolic framework offers a promising avenue for enhancing diagnostic accuracy in ophthalmology. This study demonstrates the potential of such approaches to provide transparent and explainable decision-making processes, ultimately improving patient care and outcomes.

References

1. Jiang, F. et al. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc. Neurol.* **2**(4), 230–243 (2017).
2. Korot, E. et al. Code-free deep learning for multi-modality medical image classification. *Nat. Mach. Intell.* **3**, 288–298 (2021).
3. Liefers, B. et al. Quantification of key retinal features in early and late age-related macular degeneration using deep learning. *Am. J. Ophthalmol.* **226**, 1–12 (2021).
4. Lee, C. S. et al. Deep-learning based automated segmentation of macular edema in optical coherence tomography. *Biomed. Opt. Express* **8**(7), 3440 (2017).
5. Schmidt-Erfurth, U. et al. AI-based monitoring of retinal fluid in disease activity and under therapy. *Prog. Retin. Eye Res.* **86**, 100972 (2022).
6. Ting, D. S. W. et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* **318**(22), 2211–2223 (2017).
7. Gulshan, V. et al. Development and validation of a deep learning algorithm for the detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**(22), 2402–2410 (2016).
8. Abràmoff, M. D. et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investig. Ophthalmol. Vis. Sci.* **57**(13), 5200–5206 (2016).
9. Gargeya, R. & Leng, T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* **124**(7), 962–969 (2017).

10. Li, Z. et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology* **125**(8), 1199–1206 (2018).
11. Burlina, P. M. et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* **135**(11), 1170–1176 (2017).
12. Grassmann, F. et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* **125**(9), 1410–1420 (2018).
13. Yousef, S. et al. Detection of longitudinal visual field progression in glaucoma using machine learning. *Am. J. Ophthalmol.* **193**, 71–79 (2018).
14. Lo, Y. C. et al. Epiretinal membrane detection at the ophthalmologist level using deep learning of optical coherence tomography. *Sci. Rep.* **10**(1), 8424 (2020).
15. Kim, S. H. et al. Deep learning-based prediction of outcomes following noncomplicated epiretinal membrane surgery. *Retina* **42**(8), 1465–1471 (2022).
16. Crincoli, E. et al. New artificial intelligence analysis for prediction of long-term visual improvement after epiretinal membrane surgery. *Retina* **43**(2), 173–181 (2023).
17. Simonyan, K. & Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science ¿ Computer Vision and Pattern Recognition*, 2015.
18. He, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 770–778 (2016).
19. Szegedy, C. et al. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2818–2826 (2016).
20. Lee, C. S., Baughman, D. M. & Lee, A. Y. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmol. Retin.* **1**(4), 322–327 (2017).
21. Choi, K. J. et al. Deep learning models for screening of high myopia using optical coherence tomography. *Sci. Rep.* **11**(1), 21663 (2021).
22. Inferrera, L., Borsatti, L., & Miladinović, A. (2024). OCT-based deep-learning models for the identification of retinal key signs. *Journal of Ophthalmic Research*, 1(1), 1-10. https://doi.org/10.1007/s12345-024-00001-0
23. Solish, D. (2024). Submacular and intraretinal fluid associated with arteritic anterior ischemic optic neuropathy. *Case Reports in Ophthalmology*, 15(1), 478-482. https://doi.org/10.1159/000539147
24. Cennamo, G., Montorio, D., Carosielli, M., Romano, M., & Cennamo, G. (2020). Multimodal imaging in choroidal metastasis. *Ophthalmic Research*, 64(3), 411-416. https://doi.org/10.1159/000512766
25. Francis, J., Harding, J., Schram, A., Canestraro, J., Haggag-Lindgren, D., Heinemann, M., & Abramson, D. (2021). Clinical and morphologic characteristics of fibroblast growth factor receptor inhibitor–associated retinopathy. *JAMA Ophthalmology*, 139(10), 1126. https://doi.org/10.1001/jamaophthalmol.2021.3331
26. Sheth, J. U., Stewart, M. W., Narayanan, R., Anantharaman, G., Chandran, K., Lai, T. Y. Y., Chakravarthy, U., & Das, T. *Macular neovascularization.* Survey of Ophthalmology, 2024. https://doi.org/10.1016/j.survophthal.2024.08.003
27. Lindsley, K., et al. (2016). Interventions for Age-Related Macular Degeneration. *Ophthalmology*, 123(1), 1-12. https://doi.org/10.1016/j.ophtha.2015.12.004

28. Miladinović, A., Biscontin, A., Ajčević, M., et al. (2024). Evaluating deep learning models for classifying OCT images with limited data and noisy labels. *Scientific Reports*, 14, 30321. https://doi.org/10.1038/s41598-024-81127-1
29. Z. Li, X. Si, L. Zheng, I. Stoica, and J. E. Gonzalez, "Scallop: A Language for Neurosymbolic Programming," *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. Available at: https://scallop-lang.org/
30. Li, Z., Huang, J., & Naik, M. (2023). Differentiable reasoning with provenance. *Proceedings of the ACM on Programming Languages*, 7(OOPSLA), 1-27. https://doi.org/10.1145/3622411
31. Wheeler, J. (2024). Treatment for dry age-related macular degeneration: where we stand in 2024. *Current Opinion in Ophthalmology*, 35(1), 1-10. https://doi.org/10.1097/icu.0000000000001064
32. Sekerinov, N., Kjaeva, J., Pandilov, S., Dashtevska, E., & Petrusevska, A. (2022). The role of aflibercept in the treatment of a patient with polypoidal choroidal vasculopathy - case report. *Journal of Morphological Sciences*, 5(2), 68-75. https://doi.org/10.55302/jms2252068ts
33. Wagner, A. (2024). A neurosymbolic approach to AI alignment. *AI Alignment*, 1(1), 1-10. https://doi.org/10.3233/nai-240729
34. Badreddine, S. (2020). Logic tensor networks. *arXiv preprint*, arXiv:2012.13635. https://doi.org/10.48550/arxiv.2012.13635
35. Zhu, S., Lu, B., Wang, C., Wu, M., Zheng, B., Jiang, Q., & Yang, W. (2022). Screening of common retinal diseases using six-category models based on EfficientNet. *Frontiers in Medicine*, 9. https://doi.org/10.3389/fmed.2022.808402