

**THE EYE AS A "WINDOW TO THE BRAIN": A COMPREHENSIVE REVIEW OF
MACHINE LEARNING FOR OCULAR-BASED NEUROLOGICAL DISEASE
DETECTION**

Deep Narayan ¹, Bibhuti Kumbhakar ¹, Dr. Pratik Patel ², Prashant Pradhan ³, Sonu Kumar ¹, Ram Kumar Thakur ⁴, Ritesh kumar Jha ¹, Dr Ranjan Kumar Mishra ^{1*}

¹Department of Information Technology- Netaji Subhas University, Jamshedpur, Jharkhand, India.

E-mail: ndeepwaz@gmail.com

bibhulikumbhakar@gmail.com

Qskumar7274@gmail.com

pujariteshjha@gmail.com

ranjanlnmi4u@gmail.com (corresponding author)

²Department of School of Forensic Science - National Forensic Sciences University, Gandhinagar Gujarat.

e-mail: Pratik.patel@nfsu.ac.in

³Department of Engineering –NTTF, Tata Steel Technical Institute, Jamshedpur, Jharkhand

e-mail: pradhanprashant619@gmail.com

⁴Department of Information Technology –AL-Karim University, Katihar, Bihar, India

e-mail: ramtestprofssional@gmail.com

Abstract

Neurological disorders are increasingly recognized as a significant global health challenge, thereby necessitating the development of diagnostic tools that are non-invasive, easily deployable, and affordable. This review proposes a novel approach: utilizing machine learning (ML) to identify neurological diseases via ocular imaging. Based on the neuroanatomical connection between the retina and the central nervous system (CNS), the retina is conceptualized as a dynamic "window to the brain," where systemic neuropathology's manifest visibly. The subsequent analysis identifies retinal biomarkers associated with particular neurological conditions: amyloid-beta ($A\beta$) and phosphorylated-tau (pTau) aggregates in Alzheimer's disease (AD); dopaminergic neuronal loss and retinal thinning in Parkinson's Disease (PD); and axonal degeneration following optic neuritis in Multiple Sclerosis (MS). The computational domain is subsequently delineated across the machine learning spectrum, encompassing fundamental image segmentation techniques utilizing U-Net architectures and progressing to sophisticated network paradigms like Vision Transformers (ViT) and Graph Neural Networks (GNNs), which are adept at capturing

intricate morphological and relational characteristics. Contemporary models have exhibited remarkable effectiveness, as evidenced by certain Alzheimer's disease detection systems attaining accuracies surpassing 98%. Nevertheless, the practical application of these models in clinical settings is presently impeded by considerable challenges, including the opacity of algorithmic processes, the presence of biases within data environments, and the ethical implications associated with patient privacy. Therefore, the future necessitates a multifaceted strategy Explainable AI to improve interpretability, federated learning to enable secure data sharing, and multimodal fusion to ensure comprehensive diagnostic integration—thereby progressing the field towards a truly transparent and clinically relevant area within neurodiagnostics.

.Key Words: SVM, CNN, Transformers, U-Net, GNN, etc.

Introduction

The Emerging Paradigm of Ocular Biomarkers

1.1 The Global Burden of Neurological Disease

Neurological disorders represent a primary contributor to global disability and illness. The Global Burden of Disease (GBD) Study provides standardized data that highlights this significant impact, thereby illustrating the considerable societal and economic burdens associated with ailments like Alzheimer's disease (AD), Parkinson's disease (PD), and multiple sclerosis (MS). Specifically, Alzheimer's and related dementias are responsible for a considerable proportion of new cases and disability-adjusted life years (DALYs) globally. Consequently, the economic consequences are substantial, stemming not only from direct healthcare costs but also from the considerable societal losses linked to reduced productivity and the persistent requirements of long-term care. The escalating nature of this crisis is exacerbated by a persistent diagnostic deadlock. While current gold-standard techniques, such as positron emission tomography (PET) and cerebrospinal fluid (CSF) analysis obtained through lumbar puncture, offer high diagnostic accuracy, they are nonetheless invasive, expensive, and generally employed only in the later stages of disease, well after irreversible neural damage has taken place. This situation underscores a critical and unmet clinical need: the development of diagnostic methods that are precise, scalable, readily available, and fundamentally non-invasive, thus facilitating the early, even preclinical, identification of neurodegenerative diseases [1-5].

1.2 The Retina: A "Window to the Brain"

The human eye might offer a solution to this diagnostic problem. The retina, beyond its role as a sensory organ, constitutes a distinctive and readily accessible segment of the central nervous system (CNS) [7]. This fundamental biological principle, frequently referred to as the "eye-brain connection," is predicated on several crucial tenets:

1. Embryological Development: The retina and the optic nerve arise as direct extensions of the diencephalon during embryogenesis, thereby representing tangible extensions of the

brain [7].

2. **Structural Parallels:** The retina is comprised of identical cell types (neurons and glial cells) and exhibits a comparable structural arrangement to the brain [6]. Furthermore, it features a blood-retinal barrier, which functions similarly to the blood-brain barrier [7].

3. **Vascular Access:** The retinal microvascular network is unique in the human body because it allows for the *in vivo*, non-invasive observation of both microvasculature and neural tissue [7].

Because of this close relationship, it's believed that brain diseases often show similar signs in the retina [6]. These signs can be broadly categorized into two main types of biomarkers. Neuro-specific pathology represents the first category, wherein the proteinopathies or cell-specific deficits characteristic of a brain disease, such as β deposits in Alzheimer's disease or dopaminergic cell loss in Parkinson's disease, are reflected within the retinal structure [6]. The second category encompasses systemic-vascular pathology, wherein shared risk factors, including hypertension and diabetes, precipitate concurrent damage to the microvasculature of both the brain and the retina [7]. Consequently, the eye offers a unique and valuable source of biomarkers for the assessment of both cerebrovascular and neurodegenerative diseases.

1.3 The Role of Machine Learning (ML)

Although ocular biomarkers present considerable potential, their observable characteristics are frequently quite subtle. Small pathological alterations, including micron-level thinning within specific retinal layers or undetectable changes in capillary density, often escape identification during routine clinical assessments. It is within this realm of invisibility that machine learning (ML) and artificial intelligence (AI) become essential tools. Deep learning (DL) frameworks, especially, demonstrate a strong capacity for unraveling the complex, high-dimensional data inherent in retinal imaging techniques like fundus photography and optical coherence tomography (OCT). These models, by moving beyond manually designed feature extraction, independently identify latent, higher-order representations that elucidate the relationships between retinal structure and disease manifestation. Consequently, ML-driven analyses provide the computational precision necessary to convert the retina from a passive biological indicator of neuropathology into an active, measurable diagnostic tool, thereby connecting molecular pathology with clinical application.

1.4 Scope and Structure of this Review

This review provides a comprehensive analysis of machine learning methodologies applied to the identification of neurological disorders via ocular imaging. Initially, it establishes the biological underpinnings of this domain, specifically addressing the retinal manifestations of Alzheimer's Disease, Parkinson's Disease, and Multiple Sclerosis, which are indicative of the eye-brain connection. Following this, the review explores the key imaging techniques, including fundus photography and optical coherence tomography, alongside the development of extensive public datasets that have facilitated research advancements and performance evaluation.

The subsequent analysis delves into the computational spectrum, encompassing initial image segmentation methodologies and evolving into modern, sophisticated classification architectures that utilize deep and hybrid learning approaches. Performance metrics, representing the current pinnacle of achievement, are rigorously evaluated to gauge diagnostic dependability and applicability across diverse patient groups. Ultimately, the discourse broadens to encompass the obstacles hindering clinical implementation, highlighting the significance of transparency through Explainable AI, collaborative intelligence facilitated by federated learning, and multimodal integration as the forthcoming frontier for developing clinically useful, reliable AI systems within neuro-ophthalmic diagnostics.

2. The Biological Nexus: Retinal Manifestations of Neurological Disease

2.1 Alzheimer's Disease (AD): Retinal Amyloid and Tau Pathologies

The primary biomarkers for AD are the hallmark proteinopathies of the brain: extracellular amyloid-beta ($A\beta$) plaques and intracellular neurofibrillary tangles (NFTs) composed of hyperphosphorylated-tau (τ) protein[6]. A growing body of evidence indicates that these same pathologies accumulate in the retina.

Post-mortem studies have reported the presence of $A\beta$ deposits, including the toxic $A\beta_{40}$ and $A\beta_{42}$ alloforms, in the retinas of confirmed AD patients [6]. This accumulation is not an isolated finding; studies in transgenic mouse models (e.g., APP^{swe}/PS1^{dE9}) show a direct correlation between retinal $A\beta$ levels and $A\beta$ burden in the hippocampus [14]. Crucially, retinal $A\beta$ may follow a similar time course to cerebral $A\beta$, potentially accumulating decades before the onset of cognitive symptoms, positioning it as a powerful preclinical biomarker.[6] Similarly, τ has been identified in the inner plexiform layer and the ganglion cell layer (GCL), where its presence is linked to inflammation and the loss of retinal ganglion cells (RGCs)[6].

However, the literature is not without conflict. While numerous studies report $A\beta$ deposits [16], others have explicitly stated that they "did not observe $A\beta$ plaques in the AD retinas"[16]. This same study, however, *did* observe diffuse τ in the peripheral retina of AD patients[16]. This apparent contradiction suggests the retinal pathology may not be identical to the brain's. The $A\beta$ may be present in more soluble, non-plaque forms (oligomers) or as diffuse aggregates that are methodologically difficult to detect.

This is precisely why ML models are so valuable. An algorithm does not need to "see" a discrete plaque. Regardless of the precise form of the proteinopathy, its neurotoxic presence leads to clear *downstream structural consequences*. These include the measurable loss of RGCs, a reduction in the thickness of the retinal nerve fiber layer (RNFL), and thinning of the GCL [6]. An ML model can be trained to detect the unique *spatial signature* of this combined neurodegeneration and inflammation, effectively acting as an indirect sensor for the underlying AD pathology.

2.2 Parkinson's Disease (PD): Dopaminergic Depletion and Neuroretinal Degeneration

The pathology of PD is not confined to the substantia nigra; the retina itself contains a

population of dopaminergic amacrine cells that are implicated in the disease [8]. The key retinal biomarker for PD is the degeneration of these specific cells [8]. Post-mortem studies have confirmed this finding, demonstrating "reduced tyrosine hydroxylase immunoreactivity" in the retinas of PD patients [8]. Tyrosine hydroxylase is the rate-limiting enzyme in dopamine synthesis, making its absence a direct proxy for the loss of dopaminergic neurons [8].

This specific neuronal loss, much like in AD, has clear and quantifiable structural consequences that are detectable with Optical Coherence Tomography (OCT). A major systematic review and meta-analysis of OCT measurements in PD patients versus healthy controls provided robust quantitative evidence for this degeneration [9]. The study found significant reductions in:

- **Peripapillary RNFL (pRNFL) thickness:** Weighted Mean Difference (WMD) of $-3.51 \mu\text{m}$
- **Macular Ganglion Cell Complex (GCC) thickness:** WMD of $-4.18 \mu\text{m}$
- **Macular Volume:** WMD of -0.21mm^3

This evidence strongly supports a *disease-specific biomarker hypothesis*. The thinning observed in PD is not just "general atrophy," which occurs in many retinal and neurological conditions [19]. It is a specific pattern of degeneration linked to the loss of a specific cell type. This implies that an ML model is not just learning "thinning"; it is plausibly learning the *unique spatial signature of atrophy* that results from dopaminergic depletion, making it distinct from the signature of AD.

Furthermore, this pathology manifests as functional deficits. The loss of retinal dopamine impacts visual processing, leading to impaired saccadic eye movements [21]. These abnormal movements can be captured using eye-tracking technology and classified by ML models, offering a functional biomarker that complements the structural OCT data [22].

2.3 Multiple Sclerosis (MS): Optic Neuritis and Axonal Loss

The retinal link in MS represents a different class of biomarker entirely. In AD and PD, the eye acts as a *passive window* reflecting a systemic brain disease. In MS, the optic nerve and retina are an *active target* of the disease's autoimmune attack [23].

Optic neuritis (ON), an acute inflammatory demyelinating event of the optic nerve, is a "central manifestation" of MS and is frequently the *initial* clinical presentation of the disease [23]. This inflammatory attack causes "frequent retinal ganglion cell damage" [24]. While some recovery occurs, the event leaves behind a permanent "scar" of chronic axonal degeneration, which is measurable as a significant and progressive thinning of the GCL and RNFL [24].

This retinal thinning is a powerful biomarker for both diagnosis and prognosis. Its presence and severity are used to monitor disability progression [24]. Machine learning algorithms, including traditional models like Support Vector Machines (SVMs), have been successfully employed to analyze these OCT-derived variables (e.g., layer thicknesses) to classify MS

patients against healthy controls [28]. This application demonstrates a clear, clinically relevant use case where ML can quantify the damage from a demyelinating event.

3. Imaging Modalities and Data Ecosystems

The translation of the biological hypotheses from Section 2 into functional diagnostic models is entirely dependent on two factors: the imaging technologies to capture the biomarkers and the large-scale datasets to train the algorithms. The alignment of these two factors has created a "perfect storm" for rapid advancement in computational neurology.

3.1 Non-Invasive Retinal Imaging

Three primary imaging modalities provide the raw data for ML analysis:

- **Retinal Fundus Photography:** This is the most common and accessible modality, providing a 2D color photograph of the posterior eye [30]. It is excellent for visualizing the optic disc, the retinal surface, and the *macro-vascular* architecture [7]. ML models have used fundus images to detect conditions like diabetic retinopathy, anemia, and to predict Parkinson's disease [11].
- **Optical Coherence Tomography (OCT):** This is the key technology for neurodegenerative disease detection [33]. Often described as an "optical biopsy," OCT uses light waves to generate high-resolution (micron-scale) 3D cross-sectional images of the retina's layered structure [7]. Its essential function is to enable the precise segmentation and *thickness measurement* of individual neural layers (e.g., RNFL, GCL), which are the primary structural biomarkers for AD, PD, and MS [9].
- **OCT Angiography (OCTA):** This is a functional extension of OCT. By detecting the motion of red blood cells, OCTA generates 3D, depth-resolved maps of the retinal capillary networks *without* requiring dye injection [23]. This allows for the analysis of *micro-vascular* changes, such as reduced vessel density or capillary drop-out, which are associated with neurodegeneration and the inflammation of optic neuritis [7].

3.2 Public and Large-Scale Datasets: The Computational Catalyst

For years, research in this area was limited to small, single-center studies. The advent of deep learning, which requires vast amounts of data to learn subtle patterns, was bottlenecked by data availability. The recent catalyst for the entire field has been the curation and release of massive, population-scale, multimodal datasets [38].

- **UK Biobank (UKB):** This is arguably the single most important dataset for this field [32]. The UKB is a prospective cohort study with over 500,000 participants. Critically, a large subset of these participants (over 67,321) underwent comprehensive ocular imaging [40]. The dataset includes both retinal fundus photography [32] and spectral-domain OCT scans [36]. This ocular data is linked to the participants' full medical histories, genetic data, and neurological disease diagnoses [34]. The UKB has even made derived data, such as automated retinal thickness measurements, available to researchers [34]. This dataset creates a perfect testbed, explicitly designed to link ocular biomarkers to "systemic disease (diabetes and neurodegenerative diseases)" [34].

- Alzheimer's Disease Neuroimaging Initiative (ADNI):** This is a foundational dataset for AD research. While its primary focus is on "gold standard" neuroimaging like MRI and PET [43]. ADNI provides the critical clinical and imaging data needed to *validate* ocular biomarkers against established brain pathologies [38].

The availability of these datasets, particularly UKB, has enabled the shift from small-scale hypothesis testing to large-scale deep learning, serving as the primary enabler of the field's recent progress.

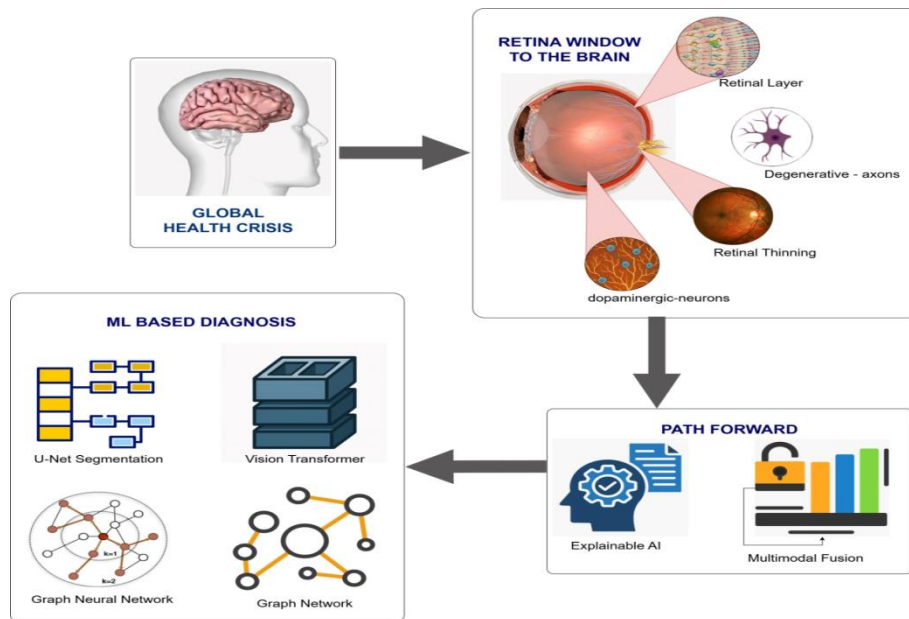


Image:1

4. Computational Methods I: Retinal Layer Segmentation via Deep Learning

4.1 The Critical Role of Segmentation

Before a biomarker like "RNFL thinning" can be used for classification, it must be precisely quantified. This requires a foundational ML task: **image segmentation**. In the context of OCT, segmentation is the process of automatically drawing the precise boundaries of each retinal layer in a 3D scan [19]. This is a non-trivial challenge due to low image contrast, variable anatomy, and signal noise from blood flow [19].

The accuracy of the final diagnostic model is fundamentally *bottlenecked* by the accuracy of this initial segmentation step [35]. An error of just a few microns in a layer boundary can be the difference between a "healthy" and "pathological" thickness measurement. Thus, segmentation is not "preprocessing"; it is the *core feature extraction* step that enables the entire quantitative pipeline.

4.2 U-Net: The Dominant Architecture

In the domain of biomedical image segmentation, the **U-Net** architecture is the *de facto* industry and academic standard [45]. Its design is uniquely suited for this task.

A U-Net consists of two paths: an **encoder** (contracting path) and a **decoder** (expanding

path) [19]. The encoder path uses successive convolutional and pooling layers to capture the *context* and *semantic features* of the image (i.e., "what" it is looking at). The decoder path uses upsampling and convolutional layers to gradually reconstruct a high-resolution segmentation map, localizing the features (i.e., "where" they are) [45].

The key innovation of the U-Net is the use of **skip connections**, which concatenate feature maps from the encoder path directly with the corresponding layer in the decoder path [45]. This allows the network to combine deep, semantic information with shallow, high-resolution feature maps. This fusion is essential for precisely delineating the boundaries of retinal layers [45].

4.3 Evolutions of U-Net

The field continues to build upon the standard U-Net. Many modern approaches integrate **attention mechanisms**, which help the model learn to "pay attention" to the most salient image regions and suppress irrelevant noise. More recent innovations include hybrid models like **TransUnet** or **LightReSeg**, which integrate Transformer blocks into the U-Net structure [19]. This leverages the Transformer's power for global reasoning, allowing the model to better understand the relationships between distant parts of the image, often leading to improved segmentation performance with fewer parameters [19].

This segmentation stage is the foundation of the dominant workflow in the field: the **"Segmentation-then-Classification" (StC) paradigm**. In this pipeline, a U-Net variant first segments the raw 3D OCT scan, which is then used to generate a quantitative thickness map[35].or tabular data (e.g., "pRNFL thickness = 90.5 μm "). This *derived data* is then fed into a second, often simpler, ML model (like an SVM[29] or Random Forest[49]) for the final disease classification.

5. Computational Methods II: Advanced Architectures for Disease Classification

While the STC pipeline is robust and interpretable, a second stream of research focuses on "end-to-end" models that learn directly from the raw images, as well as novel data representations that capture different aspects of the pathology.

5.1 Transformer-Based Models (ViT, Swin)

The application of Transformers, which were originally designed for natural language processing, to image analysis (Vision Transformer, or ViT) has yielded powerful results [50]. Unlike CNNs, which excel at learning local features, Transformers use a self-attention mechanism to model *global relationships* and long-range dependencies in the data [51].

This is highly relevant for retinal analysis. A Transformer can learn a *holistic pattern* of atrophy across the entire macula, rather than just isolated local thinning.

- The **Swin Transformer**, a hierarchical model that builds representations at multiple scales, has been applied directly to OCT images to distinguish AD patients from healthy controls, achieving state-of-the-art accuracy [52].
- TransNetOCT, another Transformer-based architecture, was also developed

specifically for OCT classification and demonstrated SOTA performance in AD detection [52].

These models represent a shift toward true end-to-end learning, where the model itself learns the most salient features without relying on a prior segmentation step.

5.2 Graph Neural Networks (GNNs): Modeling Relational Biomarkers

One of the most advanced frontiers in the field involves fundamentally changing the data representation. Instead of treating the retina as a 2D or 3D grid of pixels, this approach models it as a **graph**, composed of nodes and edges [37].

This method is primarily applied to the retinal *vasculature* using fundus or OCTA images [37]. In this representation:

- Vessel bifurcations or junctions are treated as **nodes** [57].
- The vessel segments connecting them are treated as **edges** [57].

A Graph Neural Network (GNN) is then used to learn from this "relational data" [53]. GNNs are designed to capture complex topological patterns, such as connectivity, tortuosity, and branching angles, that are lost in standard pixel-based analysis [56]. The "Eye-AD" model, for example, uses a "multilevel graph representation" to assess not just individual features (like vessel diameter) but also the *relationship* between structures (e.g., how vascular topology relates to the fovea)[37]. This allows the model to capture the intricate anatomical and vascular relationships that may be a more robust biomarker of disease than any single measurement [37].

5.3 Convolutional (CNN) and Multimodal Fusion Approaches

This analysis reveals two parallel and complementary research pipelines. The first is the **Quantitative-Atrophy Pipeline**, which uses OCT images, applies U-Net for segmentation, and classifies the resulting *neural layer thickness data*[9]. The second is the **Relational-Vascular Pipeline**, which uses fundus or OCTA images, applies GNNs to model the *vascular tree*, and classifies the resulting *graph topology* [37].

These two streams are not competing; they are analyzing different aspects of the disease's manifestation (neural atrophy and vascular pathology). The most powerful models are now beginning to *fuse* them. **Multimodal fusion** is a key trend where models combine multiple data types [49]. A prime example comes from MS detection, where a model that used "early fusion" to *combine* tabular OCT features (from the atrophy pipeline) with a CNN analysis of the raw OCT images *outperformed* all models that used only one data type[49]. This strongly suggests the future of the field lies in integrating these complementary approaches.

6. State-of-the-Art (SOTA) Performance in Neurological Disease Detection

The application of the aforementioned ML models has yielded highly promising performance metrics, particularly in the classification of Alzheimer's Disease. These results are summarized in Table 1.

6.1 Alzheimer's Disease (AD)

Research in AD detection has reported exceptionally high accuracies, demonstrating strong proof-of-concept.

- The **TransNetOCT** model achieved an average accuracy of **98.18%** across 5-fold cross-validation for classifying AD from OCT images [52].
- A **Swin Transformer** model, evaluated on the Dryad OCT dataset, achieved a nearly identical accuracy of **98.11%** [52].
- Other studies using hybrid VGG19 models with feature fusion and optimization have reported accuracies of **98%** and even **99%** [62].
- A multi-modal AI system on the RetinalOCT dataset also achieved **98.0%** accuracy [63].

While some of these high-performing models were trained on MRI data [62], the consistent SOTA results from OCT-based models[52] are particularly noteworthy.

6.2 Multiple Sclerosis (MS)

Performance in MS detection is also robust, though reported accuracies are more conservative.

- A study developing a multimodal fusion model for MS classification achieved a top accuracy of **90%**, with an F1-score of 0.87 and an AUC of 0.90[49]. This model used an "early fusion" approach, combining tabular OCT features with raw image data, and notably outperformed unimodal models [49].

6.3 Parkinson's Disease (PD)

For PD, the literature confirms the successful application of deep learning to classify the disease using fundus images from the UK Biobank dataset [32]. and from eye-tracking data[22].The focus has been on establishing the viability of these modalities and the strong correlation of the underlying biomarkers[9].

Table 1: Summary of SOTA Machine Learning Models for Ocular-Based Neurological Disease Detection

Model Architecture	Disease	Data Modality	Dataset	Reported Performance	Reference(s)
TransNetOCT	Alzheimer's Disease	OCT		Accuracy: 98.18% (5-fold CV)	52

Swin Transformer	Alzheimer's Disease	OCT	Dryad	Accuracy: 98.11%	52
Multi-modal AI System	Alzheimer's Disease	OCT	RetinalOCT	Accuracy: 98.0%	63
VGG19 + Feature Optimization	Alzheimer's Disease	MRI		Accuracy: 99%	62
Convolutional Mixer	Alzheimer's Disease	MRI	ADNI	Accuracy: 98.87%	64
Early Fusion (CNN + SVC)	Multiple Sclerosis	OCT (Image + Tabular)	[Private Cohort]	Accuracy: 90%, AUC: 0.90	49
Deep Learning	Parkinson's Disease	Fundus	UK Biobank	Classification study	32

6.4 Critical Analysis of SOTA Performance

The 98-99% accuracy figures for AD are remarkable, but they must be interpreted with critical academic caution. These results are often achieved on smaller, curated, and relatively "clean" academic datasets [65]. This can lead to a **"generality gap,"** where models exhibit high performance in the lab but fail to generalize to the "messy" data of a diverse, real-world clinical population.

In this context, the **90%** accuracy achieved by the MS model [49]. may be a *more impressive and realistic* result. This model was explicitly tested on "diagnostically complex cases" [49], representing a more significant clinical challenge than a simple binary classification of severe AD vs. healthy controls. The true test of all SOTA models will be their performance on large-scale, heterogeneous population data, such as the UK Biobank [32].

7. Critical Challenges and Pathways to Clinical Translation

Despite the promising SOTA results, these models are not yet in routine clinical use. The transition from lab to clinic is blocked by several critical, interconnected challenges.

7.1 The "Black Box" Problem and Explainable AI (XAI)

A primary barrier to clinical and regulatory acceptance is the "black box" nature of deep learning [66]. A model that simply outputs a diagnosis ("AD: 98% probability") with no justification is clinically unacceptable. Clinicians and regulators must be able to audit the

model's "reasoning" to ensure its decisions are based on sound biological principles and not on spurious correlations or dataset artifacts [68].

The solution to this problem is **Explainable AI (XAI)**, a suite of techniques designed to make model decisions interpretable [66]. A case study on PD diagnosis from OCT provides a clear blueprint for XAI implementation [17]:

- **Global Interpretability (e.g., SHAP):** Techniques like SHapley Additive exPlanations (SHAP) are used to provide *global* explanations. They quantify which biomarkers the model weighed most heavily across *all* predictions (e.g., "The model bases 26% of its decision on the superior RNFL quadrant") [17].
- **Local Interpretability (e.g., LIME):** Techniques like Local Interpretable Model-agnostic Explanations (LIME) provide *local*, case-by-case explanations. They explain *why* a *single patient* received their diagnosis (e.g., "This patient was flagged as high-risk *because* their superior RNFL thickness was below the 120 μm threshold") [17].

This transparency is essential for building clinical trust and is increasingly becoming a prerequisite for regulatory approval.

7.2 Data Privacy, Security, and Algorithmic Bias

The challenges of data privacy and algorithmic bias are not separate issues; they are part of an interconnected "vicious cycle" that currently stalls research.

1. **Privacy:** Medical images are highly sensitive protected health information, governed by strict regulations like HIPAA in the U.S. and GDPR in Europe [67].
2. **Data Siloing:** As a result, this sensitive data is locked in institutional "silos" (individual hospitals or research centers) and cannot be easily or legally pooled into a single, large database [72]
3. **Bias:** This forces researchers to train models on the data they have access to, which is often small, local, and homogeneous (e.g., from a single city or health system) [67]. When a model is trained on such non-diverse data, it may learn spurious correlations specific to that group and fail dramatically when deployed on a different patient population, thus creating or exacerbating health inequities [70].
4. **The "Black Box":** The lack of transparency in "black box" models *hides* this underlying bias, making it impossible to detect until the model fails in deployment [66].

Breaking this cycle—which flows from essential privacy laws to dangerous algorithmic bias—is the central challenge for the clinical translation of these AI models.

8. Future Frontiers: Federated Learning and Multimodal Integration

The solutions to the field's most significant challenges lie in new computational paradigms that can systematically break the "vicious cycle" of privacy, siloing, and bias.

8.1 Federated Learning (FL) for Privacy-Preserving Collaboration

Federated Learning (FL) is an ML paradigm specifically designed to solve the data-sharing

problem [73]. It *directly* breaks the link between privacy and data siloing.

The FL process works as follows:

1. A central server holds a "global" ML model.
2. This model is *sent to* participating institutions (e.g., multiple hospitals worldwide) [71].
3. Each hospital then trains the model *locally* on its own private patient data, which *never leaves the hospital's secure server* [71].
4. Only the updated model *weights*—anonymous mathematical adjustments, not the data itself—are sent back to the central server [71].
5. These weights are aggregated to improve the "global" model, which is then sent back out for another round of training.

FL allows for the creation of a single, highly robust global model trained on a massive, diverse, multi-institutional dataset *without* any patient data ever being shared or exposed [72]. This approach respects patient privacy and regulatory boundaries while directly addressing the data-silo problem that leads to algorithmic bias. While FL introduces its own technical hurdles, such as the need for data and scanner harmonization and managing statistical heterogeneity between sites [75], it represents the most viable path forward for collaborative medical AI.

8.2 Multimodal Deep Learning

The future of diagnostic accuracy lies not in finding a single "silver bullet" biomarker, but in **multimodal data fusion** [59]. As the SOTA results demonstrated, models that combine data types (e.g., images + tabular data [49]) or different modalities (e.g., OCTA + quantitative data [65]) are more robust and accurate.

The most powerful future diagnostic systems will be multimodal, integrating data from many sources to build a comprehensive patient profile [67].

This includes:

- **Ocular Data:** (Fundus photography, 3D OCT, and OCTA) [23]
- **Genetic Data:** (e.g., APOE status for AD) [65]
- **Clinical Data:** (e.g., cognitive scores, or NLP of electronic health records) [69]
- **Other Imaging Data:** (e.g., validation against MRI) [12]

9. Conclusion

This review has synthesized the extensive evidence supporting the retina as a viable, non-invasive "window to the brain" for the detection of neurological disease. This paradigm is built on a solid biological foundation: the direct manifestation of specific pathologies, such as β and τ in AD, dopaminergic cell loss in PD, and axonal degeneration in MS,

within the retinal layers.

We have analyzed the computational methodologies that translate this biology into practice, from the foundational "Segmentation-then-Classification" pipeline enabled by U-Net to advanced end-to-end Transformers and relational GNNs. The state-of-the-art performance of these models, with accuracies exceeding 98% in AD classification, provides a powerful proof-of-concept.

However, we have critically identified that SOTA results are necessary but not sufficient for clinical translation. The path to real-world deployment is currently blocked by an interconnected system of challenges: the "black box" trust deficit, the data-siloing effect of essential privacy laws, and the dangerous algorithmic bias that results.

The future of the field is not in developing a slightly more accurate algorithm. It is in building integrated, trustworthy systems. The ultimate solution will be a **Federated, Multimodal, and Explainable** AI. This approach will leverage **Federated Learning** to train models on diverse, global data in a privacy-preserving manner; it will be **Multimodal**, integrating ocular, genetic, and clinical data for maximum robustness; and it will be **Explainable**, using XAI to provide the transparency and audibility required for clinical and regulatory trust. By focusing on these three pillars, the field can successfully bridge the gap from a promising research concept to a transformative clinical reality.

References

- [1] Effects of Neurological Disorders on Health Expenditure and Economic Output: Dynamic Panel Analysis for OECD Countries - MDPI, accessed November 1, 2025, <https://www.mdpi.com/2079-8954/13/7/521>
- [2] Incidence and lifetime prevalence of neurological disorders in a ..., accessed November 1, 2025, <https://academic.oup.com/brain/article/123/4/665/281572>
- [3] Global, regional, and national burden of four major neurological diseases in women from 1990 to 2021 - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12014452/>
- [4] Global, regional, and national burden of four major neurological diseases in women from 1990 to 2021 - Frontiers, accessed November 1, 2025, <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2025.1561216/full>
- [5] Global, regional, and national burden of neurological disorders ..., accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC6459001/>
- [6] Potential Retinal Biomarkers in Alzheimer's Disease - PMC, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10649108/>
- [7] The Role of AI in Early Detection of Life-Threatening Diseases: A Retinal Imaging Perspective - arXiv, accessed November 1, 2025, <https://arxiv.org/html/2505.20810v1>
- [8] Retinal Alterations Predict Early Prodromal Signs of Neurodegenerative Disease - MDPI,

- accessed November 1, 2025, <https://www.mdpi.com/1422-0067/25/3/1689>
- [9] Changes in Visual Evoked Potential and Optical Coherence ..., accessed November 1, 2025, https://www.researchgate.net/publication/395823835_Changes_in_Visual_Evoked_Potential_and_Optical_Coherence_Tomography_in_Parkinson's_Disease_A_Systematic_Review_and_Meta-Analysis
- [10] Review of Machine Learning Applications Using Retinal Fundus Images - PMC, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC8774893/>
- [11] Deep Learning and Machine Learning Algorithms for Retinal Image Analysis in Neurodegenerative Disease: Systematic Review of Datasets and Models - PMC - NIH, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10893898/>
- [12] An Explainable Transformer Model for Alzheimer's Disease Detection Using Retinal Imaging, accessed November 1, 2025, <https://arxiv.org/html/2507.04259v1>
- [13] Retina pathology as a target for biomarkers for Alzheimer's disease: Current status, ophthalmopathological background, challenges, and future directions - PMC, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10917008/>
- [14] Retinal Levels of Amyloid Beta Correlate with Cerebral Levels of Amyloid Beta in Young APPswe/PS1dE9 Transgenic Mice before Onset of Alzheimer's Disease - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC7532376/>
- [15] Role of Retinal Amyloid- β in Neurodegenerative Diseases: Overlapping Mechanisms and Emerging Clinical Applications - PMC - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC7956232/>
- [16] Retinal imaging in Alzheimer's and neurodegenerative diseases - PMC - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC8062064/>
- [17] (PDF) Decoding Parkinson's diagnosis: An OCT-based explainable ..., accessed November 1, 2025, https://www.researchgate.net/publication/392697615_Decoding_Parkinson's_Diagnosis_An_OCT-Based_Explainable_AI_with_SHAPLIME_Transparency_from_the_Persian_Cohort_Study
- [18] Retinal Alterations Predict Early Prodromal Signs of ... - MDPI, accessed November 1, 2025, <https://mdpi.com/1422-0067/25/3/1689>
- [19] Lightweight Retinal Layer Segmentation With Global Reasoning - PMC - NIH, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11759324/>
- [20] CTS-Net: A Segmentation Network for Glaucoma Optical Coherence Tomography Retinal Layer Images - ResearchGate, accessed November 1, 2025,

https://www.researchgate.net/publication/368407104_CTS-Net_A_Segmentation_Network_for_Glaucoma_Optical_Coherence_Tomography_Retinal_Layer_Images

- [21] Bridging Cognitive and Deep Learning Models of Attention, accessed November 1, 2025, <https://d-nb.info/1373153598/34>
- [22] GazeNet: End-to-end eye-movement event detection with deep neural networks | Request PDF - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/328360324_gazeNet_End-to-end_eye-movement_event_detection_with_deep_neural_networks
- [23] Imaging biomarkers in optic neuritis: current tools and ... - Frontiers, accessed November 1, 2025, <https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2025.1666835/pdf>
- [24] Prior optic neuritis detection on peripapillary ring scans using deep learning - PMC, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC9639624/>
- [25] Globally optimal OCT surface segmentation using a constrained IPM optimization, accessed November 1, 2025, https://www.researchgate.net/publication/357633892_Globally_optimal_OCT_surface_segmentation_using_a_constrained_IPM_optimization
- [26] 1-s2.0-S0010482525007632-main | PDF | Machine Learning | Retina - Scribd, accessed November 1, 2025, <https://www.scribd.com/document/904129847/1-s2-0-S0010482525007632-main>
- [27] (PDF) Towards label-free 3D segmentation of optical coherence tomography images of the optic nerve head using deep learning - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/343754573_Towards_label-free_3D_segmentation_of_optical_coherence_tomography_images_of_the_optic_nerve_head_using_deep_learning
- [28] Comparison of Machine Learning Methods Using Spectralis OCT for Diagnosis and Disability Progression Prognosis in Multiple Sclerosis - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC9001622/>
- [29] Computer-Aided Diagnosis of Multiple Sclerosis Using a Support Vector Machine and Optical Coherence Tomography Features - MDPI, accessed November 1, 2025, https://www.mdpi.com/1424-8220/19/23/5323?type=check_update&version=1
- [30] Deep Learning in Medical Image Processing and Analysis [39] 9781839537936, 9781839537943 - DOKUMEN.PUB, accessed November 1, 2025, <https://dokumen.pub/deep-learning-in-medical-image-processing-and-analysis-39-9781839537936-9781839537943.html>
- [31] (PDF) Efficient Deep Retinal Fundus Image-Based Network for Alzheimer's Disease

- Diagnosis Using Mobile Device Applications - ResearchGate, accessed November 1, 2025,
https://www.researchgate.net/publication/380310250_Efficient_Deep_Retinal_Fundus_Image-Based_Network_for_Alzheimer's_Disease_Diagnosis_Using_Mobile_Device_Applications
- [32] Evaluation of Parkinson Disease from MRI Images Using Deep ..., accessed November 1, 2025,
https://www.researchgate.net/publication/396029187_Evaluation_of_Parkinson_Disease_from_MRI_Images_Using_Deep_Learning_Techniques
- [33] Deep Learning in Retinal Optical Coherence Tomography (OCT): A comprehensive survey | Request PDF - ResearchGate, accessed November 1, 2025,
https://www.researchgate.net/publication/362562831_Deep_Learning_in_Retinal_Optical_Coherence_Tomography_OCT_A_comprehensive_survey
- [34] Return 1875 - UK Biobank, accessed November 1, 2025,
<https://biobank.ctsu.ox.ac.uk/ukb/dset.cgi?id=1875>
- [35] Deep Learning based retinal layer segmentation in optical coherence tomography scans of patients with inherited retinal diseases | medRxiv, accessed November 1, 2025,
<https://www.medrxiv.org/content/10.1101/2023.09.26.23295852v2.full-text>
- [36] Imaging Data - UK Biobank Community, accessed November 1, 2025,
<https://community.ukbiobank.ac.uk/hc/en-gb/articles/24618819821981-Imaging-Data>
- [37] The Eye as a Window to Brain Health: Can Retinal Imaging and AI Modeling Predict Alzheimer's Disease? - PMC - PubMed Central, accessed November 1, 2025,
<https://pmc.ncbi.nlm.nih.gov/articles/PMC12484699/>
- [38] CarveMix: A Simple Data Augmentation Method for Brain Lesion Segmentation | Request PDF - ResearchGate, accessed November 1, 2025,
https://www.researchgate.net/publication/369306169_CarveMix_A_Simple_Data_Augmentation_Method_for_Brain_Lesion_Segmentation
- [39] Accessed November 1, 2025, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetAuthorizedRequestDownload.cgi?study_id=phs000179.v5.p2
- [40] Optical Coherence Tomography in the UK Biobank Study – Rapid Automated Analysis of Retinal Thickness for Large Population-Based Studies - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC5055325/>
- [41] Optical Coherence Tomography in the UK Biobank Study – Rapid Automated Analysis of Retinal Thickness for Large Population-Based Studies | PLOS One, accessed November 1, 2025,
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0164095>
- [42] Insights into human health from phenome- and genome-wide analyses of UK Biobank

- retinal optical coherence tomography phenotypes - PMC - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10246137/>
- [43] NeuroNet-AD: A Multimodal Deep Learning Framework for Multiclass Alzheimer's Disease Diagnosis - NIH, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12561078/>
- [44] Accessed November 1, 2025, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetAuthorizedRequestDownload.cgi?study_id=phs000429.v1.p1
- [45] A Hybrid Attention Mechanism-Based U-Net Model for Precise Segmentation of Sub-Retinal Layers in OCT Images - PubMed Central, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10967828/>
- [46] Comparative Analysis of Deep Learning Architectures for Macular ..., accessed November 1, 2025, <https://www.mdpi.com/2313-433X/11/2/53>
- [47] DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images - Optica Publishing Group, accessed November 1, 2025, <https://opg.optica.org/boe/fulltext.cfm?uri=boe-9-7-3244>
- [48] Automatic segmentation of retinal layers in OCT images with intermediate age-related macular degeneration using U-Net and DexiNed | PLOS One - Research journals, accessed November 1, 2025, <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0251591>
- [49] Multimodal Machine Learning for Diagnosis of Multiple Sclerosis ..., accessed November 1, 2025, <https://www.medrxiv.org/content/10.1101/2025.09.12.25335649v1.full-text>
- [50] Horizon Academic Research Journal Vol. 4 No. 1 - Scribd, accessed November 1, 2025, <https://www.scribd.com/document/681580046/Horizon-Academic-Research-Journal-Vol-4-No-1>
- [51] Improving Alzheimer's Disease Detection for Speech Based on Feature Purification Network, accessed November 1, 2025, <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2021.835960/full>
- [52] Alzheimer's Disease Classification using Retinal OCT: TransNetOCT and Swin Transformer Models - arXiv, accessed November 1, 2025, <https://arxiv.org/html/2503.11511v1>
- [53] Beyond the Eye: A Relational Model for Early Dementia Detection Using Retinal OCTA Images - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/383036867_Beyond_the_Eye_A_Relational_Model_for_Early_Dementia_Detection_Using_Retinal_OCTA_Images
- [54] VGA-Net: Vessel graph based attentional U-Net for retinal vessel segmentation, accessed November 1, 2025, https://www.researchgate.net/publication/379782726_VGA-

[Net Vessel graph based attentional U-Net for retinal vessel segmentation](#)

- [55] Interpretable Retinal Disease Prediction Using Biology-Informed Heterogeneous Graph Representations - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/389316149_Interpretable_Retinal_Disease_Prediction_Using_Biology-Informed_Heterogeneous_Graph_Representations
- [56] Improving Hypertensive Retinopathy Diagnosis Using Deep Learning and Graph Networks - Journal of Electrical Systems, accessed November 1, 2025, <https://journal.esrgroups.org/jes/article/download/8518/5726/15496>
- [57] Publications: International Conferences - VARPA Group, accessed November 1, 2025, <http://www.varpa.es/publications/intconference.php>
- [58] Beyond the Eye: A Relational Model for Early Dementia Detection Using Retinal OCTA Images - arXiv, accessed November 1, 2025, <https://arxiv.org/html/2408.05117v2>
- [59] Early Disease Detection Using AI: A Deep Learning Approach to Predicting Cancer and Neurological Disorders - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/390465855_Early_Disease_Detection_Using_AI_A_Deep_Learning_Approach_to_Predicting_Cancer_and_Neurological_Disorders
- [60] A Review on Deep Learning Approaches in Healthcare Systems: Taxonomies, Challenges, and Open Issues | Request PDF - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/347222110_A_Review_on_Deep_Learning_Approaches_in_Healthcare_Systems_Taxonomies_Challenges_and_Open_Issues
- [61] IEEE embc 2024 - conference proceedings, accessed November 1, 2025, https://embc.embs.org/2024/wp-content/uploads/sites/102/2024/07/EMBC-2024_Proceedings.pdf
- [62] Alzheimer's Disease Prediction Using Deep Feature Extraction and Optimization - MDPI, accessed November 1, 2025, <https://www.mdpi.com/2227-7390/11/17/3712>
- [63] Multi-Modal AI for Multi-Label Retinal Disease Prediction Using OCT and Fundus Images: A Hybrid Approach - MDPI, accessed November 1, 2025, <https://www.mdpi.com/1424-8220/25/14/4492>
- [64] A Convolutional Mixer-Based Deep Learning Network for ... - MDPI, accessed November 1, 2025, <https://www.mdpi.com/2075-4418/15/11/1318>
- [65] (PDF) Artificial Intelligence in Retinal Imaging for Early Alzheimer's Disease Detection: A Systematic Review - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/392431764_Artificial_Intelligence_in_Retinal_Imaging_for_Early_Alzheimer's_Disease_Detection_A_Systematic_Review
- [66] Explainable and Efficient Machine Learning using Meta Inverse Entailment Dany Varghese, accessed November 1, 2025,

https://openresearch.surrey.ac.uk/view/pdfCoverPage?instCode=44SUR_INST&filePid=13203700820002346&download=true

- [67] Harnessing Artificial Intelligence in Early Detection and Diagnosis of ..., accessed November 1, 2025, https://www.researchgate.net/publication/390532278_Harnessing_Artificial_Intelligence_in_Early_Detection_and_Diagnosis_of_Alzheimer's_Disease_Current_and_Future_Applications
- [68] Reliable image classification based on evidence and GCN - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/386902258_Reliable_image_classification_based_on_evidence_and_GCN
- [69] AI in Disease Detection - bibis.ir, accessed November 1, 2025, [https://download.bibis.ir/Books/Artificial-Intelligence/Artificial-Intelligence-Medical/2025/AI%20in%20Disease%20Detection%20-%20Advancements%20and%20Applications%20\(Rajesh%20Singh,%20Anita%20Gehlot,%20Navjot%20Rathour\)_bibis.ir.pdf](https://download.bibis.ir/Books/Artificial-Intelligence/Artificial-Intelligence-Medical/2025/AI%20in%20Disease%20Detection%20-%20Advancements%20and%20Applications%20(Rajesh%20Singh,%20Anita%20Gehlot,%20Navjot%20Rathour)_bibis.ir.pdf)
- [70] Stelia's Hyperband™ accelerates the AI Healthcare revolution, accessed November 1, 2025, <https://newsroom.stelia.ai/ai-in-healthcare-the-2025-transformation/>
- [71] Current status and practical considerations of artificial intelligence ..., accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10399944/>
- [72] (PDF) Federated Learning: A new frontier in the exploration of multi-institutional medical imaging data - ResearchGate, accessed November 1, 2025, https://www.researchgate.net/publication/390213953_Federated_Learning_A_new_frontier_in_the_exploration_of_multi-institutional_medical_imaging_data
- [73] Medical image analysis using deep learning algorithms - Frontiers, accessed November 1, 2025, <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2023.1273253/full>
- [74] AI in Disease Detection, accessed November 1, 2025, <https://content.e-bookshelf.de/media/reading/L-25845994-b8ff10ceba.pdf>
- [75] Retinal Scans and Data Sharing: The Privacy and Scientific Development Equilibrium - PMC, accessed November 1, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11975763/>