IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 July 2024, accepted 22 July 2024, date of publication 29 July 2024, date of current version 6 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3434670

SURVEY

Automated Diagnosis of Alzheimer's Disease Using OCT and OCTA: A Systematic Review

YASEMIN TURKAN[1], (Member, IEEE), F. BORAY TEK[2], FATİH ARPACI[3], OZAN ARSLAN[4], DEVRİM TOSLAK[3], MEHMET BULUT[5], AND AYLİN YAMAN[4]

[1]Department of Computer Engineering, Işık University, 34398 Istanbul, Turkey
[2]Department of Artificial Intelligence and Data Engineering, Istanbul Technical University, 34469 Istanbul, Turkey
[3]Department of Ophthalmology, Antalya Training and Research Hospital, 07100 Antalya, Turkey
[4]Department of Neurology, Antalya Training and Research Hospital, 07070 Antalya, Turkey
[5]Department of Ophthalmology, Anatolia Hospital, Antalya Akev University, 07230 Antalya, Turkey

Corresponding author: Yasemin Turkan (turkan.yasemin@gmail.com)

This study was supported by Scientific and Technological Research Council of Turkey (TUBITAK) under the Grant Number 122E509. The authors thank to TUBITAK for their supports.

ABSTRACT Retinal optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA) have emerged as promising, non-invasive, and cost-effective modalities for the early diagnosis of Alzheimer's disease (AD). However, a comprehensive review of automated deep learning techniques for diagnosing AD or mild cognitive impairment (MCI) using OCT/OCTA data is lacking. We addressed this gap by conducting a systematic review using the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines. We systematically searched databases, including Scopus, PubMed, and Web of Science, and identified 16 important studies from an initial set of 4006 references. We then analyzed these studies through a structured framework, focusing on the key aspects of deep learning workflows for AD/MCI diagnosis using OCT-OCTA. This included dataset curation, model training, and validation methodologies. Our findings indicate a shift towards employing end-to-end deep learning models to directly analyze OCT/OCTA images in diagnosing AD/MCI, moving away from traditional machine learning approaches. However, we identified inconsistencies in the data collection methods across studies, leading to varied outcomes. We emphasize the need for longitudinal studies on early AD and MCI diagnosis, along with further research on interpretability tools to enhance model accuracy and reliability for clinical translation.

INDEX TERMS Alzheimer's disease, cognitive impairment, deep learning, dementia, neural networks, optical coherence tomography, optical coherence tomography angiography, retinal imaging.

I. INTRODUCTION

Alzheimer's disease (AD) is an irreversible and progressive brain disorder characterized by a decline in cognitive function, and is the most prevalent type of dementia. Currently, there is no known cure, and it is marked by a significant reduction in brain size (neurodegeneration) caused by the accumulation of proteins (amyloid-beta and tau) in neurons [1]. As the retina and brain originate from the same neural tube, the eyes are often regarded as an extension of the brain [2]. Postmortem studies have shown that amyloid-beta and tau proteins accumulate in the retinas of individuals with AD [3]. The slow progression of the disease begins almost 20 years before the first symptoms arise and later turn into a stage called Mild Cognitive Impairment (MCI) [4]. High-resolution visual imaging technologies, such as optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA) have recently been proposed to examine the structural and vascular changes in the retinas of patients with AD.

In a recent study, Vij and Arora [5] reviewed advances in retinal modalities for diagnosing Alzheimer's disease.

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

Their review covered a range of structural, functional, and multimodal/paired imaging modalities. Among these, OCT is the most commonly studied tool used in 57% of diagnoses primarily due to its non-invasive, cost-efficient, and user-friendly nature. Another relatively new imaging technique that has gained attention in recent years is OCTA, which visualizes retinal blood flow computed from consecutive OCT scans without the need for contrast agents. Several studies have combined OCT and OCTA to enhance AD diagnosis [6], [7], [8], and other techniques have also shown promising results in diagnosing other cognitive impairments [9], [10], [11], [12], [13]. Overall, OCT and OCTA are practical and cost-effective retinal imaging tools that show considerable promise for the diagnosis of AD and cognitive impairment.

However, interpreting high-dimensional and multimodal retinal scans generated by OCT and OCTA is a difficult and time-consuming process. Consequently, automated approaches based on machine learning (ML) or deep learning (DL) have been developed to improve the efficiency of diagnosing retinal abnormalities. Although deep learning methods present several challenges for retinal disease analysis [14], they have been used to improve image quality and segment retinal layers, as well as to diagnose and predict the disease course [15], [16], [17], [18].

Bourkhime et al. [19] recently reviewed 13 studies on machine learning and novel ophthalmologic biomarkers (not specific to OCT or OCTA) for AD screening. However, the use of DL for AD diagnosis based on OCT and OCTA scans or biomarkers is yet to be surveyed - a significant gap, given the potential benefits of these techniques for early detection. Our systematic review fills a critical gap in the understanding of DL applications for diagnosing AD and MCI using OCT and OCTA scans. We focus on methodologies that directly impact the accuracy, reliability, and, ultimately, the clinical value of DL-based diagnosis using these modalities. This review incorporates the following key aspects:

Recent Medical Findings: We identified recent relevant advancements linking AD/MCI to retinal abnormalities detectable using OCT/OCTA imaging, which provides a foundation for DL-based diagnosis studies.

Dataset Curation Practices: We analyzed data collection practices, including dataset size, inclusion/exclusion criteria, and data source consistency, to identify potential model performance limitations.

Models and Training Techniques: We explored the transition from using hand-crafted features to directly processing images with DL models. This analysis also examines the strategies that address the limitations of small datasets.

Validation Strategies: We investigated the methods researchers used to validate the performance and interpret the output of their DL models.

By analyzing these aspects, our review highlights the limitations of current data collection practices and the growing trend towards image-based DL analysis on AD/MCI diagnosis. We identified the need for public datasets and longitudinal early AD/MCI studies. We also identified the need for further research on interpretability tools to improve their accuracy and reliability, ultimately enabling their reliable and interpretable clinical use.

The remainder of this paper is organized as follows. Section II establishes the relevant medical background and recent findings linking AD/MCI to retinal abnormalities that are detectable using OCT/OCTA imaging. Section III introduces a framework for "Deep Learning Flow in AD/MCI diagnosis using OCT/OCTA," which allows us to analyze the strengths, weaknesses, and prevailing trends in this field. Section IV details the methodology of this systematic review, adhering to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA [20]) guidelines, and describes the selection process of 16 relevant studies from an initial pool of 4006 candidates. Section V provides a comprehensive examination of these studies through the lens of the deep learning design flow. Finally, Sections VI and VII present a detailed discussion and conclusions of our study.

II. RELATED MEDICAL FINDINGS

The World Health Organization (WHO) classifies dementia as the seventh leading cause of death among older people globally and recognizes it as a public health priority [1]. Dementia is not a single disease [21]; it is a general term similar to heart disease and covers a wide range of specific medical conditions. WHO defines dementia as follows:

"Dementia is a syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgment" [22].

Alzheimer's disease is the most common form of dementia and is characterized by the accumulation of amyloid-beta and tau proteins in neurons [21] causing degenerative changes and mass decreases in the brain, especially in the hippocampus, cerebral cortex, and ventricles [23], [24], [25].

During embryonic development, the retina and optic nerve grow from the neural tube and are thus considered parts or extensions of the central nervous system [26]. The eye is the only organ in which we can see the cardiovascular blood vessels and an inner blood-retinal barrier (a tight protective layer of cells and capillaries that prevent larger molecules from entering the retina) similar to the blood-brain barrier [3]. Because the eye is more accessible than the brain, techniques used to examine the eye are more practical, generally non- or less invasive, and more cost-effective than brain imaging alternatives such as magnetic resonance imaging (MRI), positron emission tomography(PET), and computed tomography (CT) [27], [28]. The neurodegenerative process that affects brain neurons also affects the retinal neurons. Specifically, postmortem AD studies have highlighted the accumulation of Aβ and NFTs in retinal and optic nerve tissues. Moreover, animal [29], [30], [31] and human studies [32] have conclusively revealed retinal Aβ

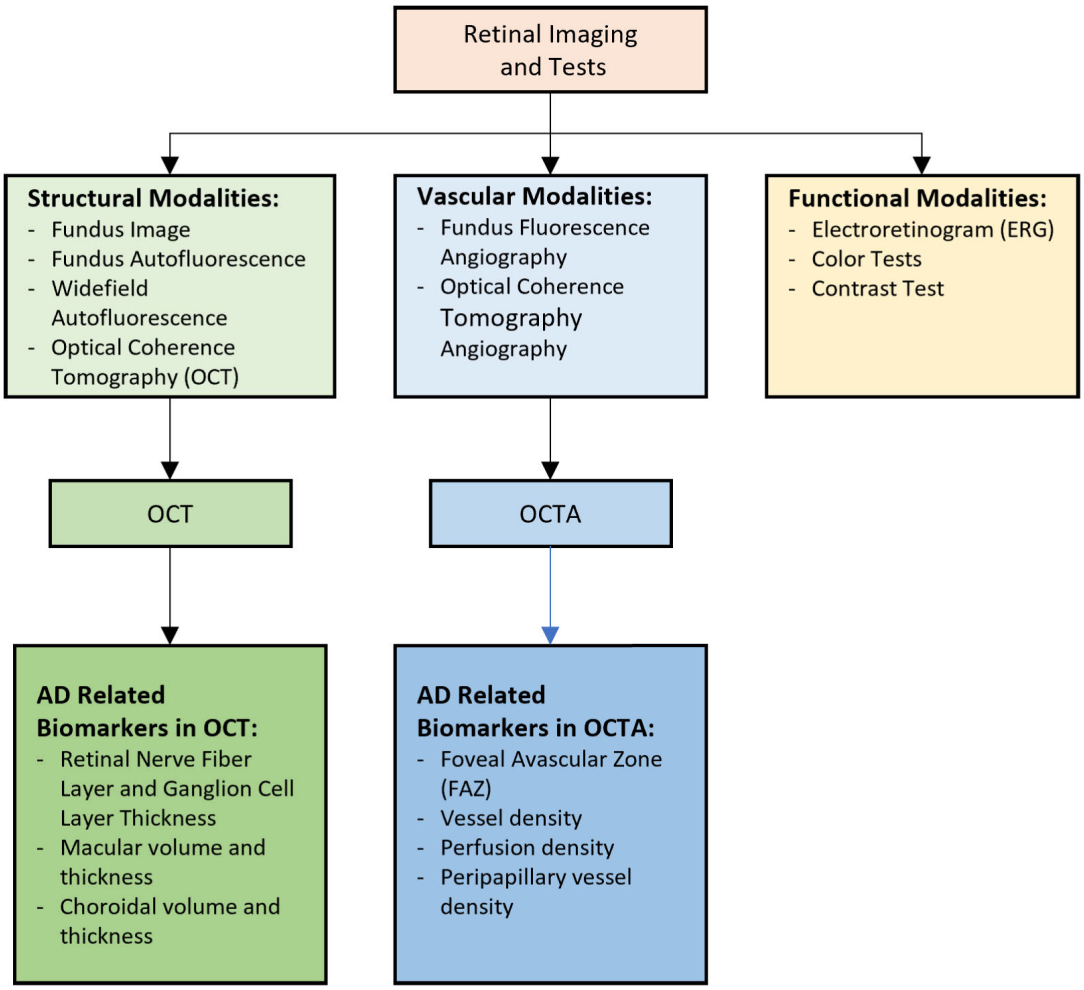

FIGURE 1. Ophthalmologists employ three groups of modalities to assess various eye diseases, including drusen, age-related macular degeneration (AMD), and glaucoma: Structural, Vascular, and Functional assessments. The diagnosis of neurodegenerative conditions such as Alzheimer's disease (AD) and Parkinson's disease using OCT and OCTA exhibits potential. Section II discusses major AD-related biomarkers identified through OCT and OCTA.

accumulation. This raises the question: Can ophthalmological assessments detect Alzheimer's disease and related disorders? Ophthalmologists commonly use multiple modalities to assess various diseases. Figure 1 shows the modalities grouped into three main categories [33].

1) Structural modalities: Fundus image, OCT, fundus autofluorescence [33], and widefield autofluorescence [34].
2) Vascular modalities: Fundus fluorescence angiography and OCTA [33].
3) Functional modalities: Electroretinogram (ERG) [35], and various color and contrast tests [3].

In this study, we focused on a survey of AD diagnosis studies performed via deep learning using two modalities: OCT and OCTA images, as highlighted in Figure 1. Examples of the OCT and OCTA scans are shown in Figure 2.

OCT is a non-invasive method used to obtain views of retinal structures in two-dimensional (2D), cross-sectional, and three-dimensional (3D) volumetric images [36]. OCT provides extensive information regarding retinal morphology and assists in the diagnosis of several diseases. A common theme in these studies is the use of quantitative parameters derived from OCT scans of the retina in addition to directly examining 3D scans.

OCTA is a relatively new technique that has been developed as a utility for use with OCT devices. It is computed as the difference between two consecutive OCT scans to visualize the relative motion in the retinal region, which reveals blood flow and vascular details at a micrometer resolution. It offers an effective imaging method for diagnosing eye diseases by unveiling microvascular changes or abnormalities in the blood flow patterns. In addition to being unavailable in all OCT devices, it is susceptible to most OCT-induced distortions such as projection, motion, segmentation, and signal loss [33].

In recent years there has been an explosion in AD/MCI biomarker-related research studies and reviews [5], [27], [37], [38], [39] on OCT and OCTA. Costanzo et al. [39]

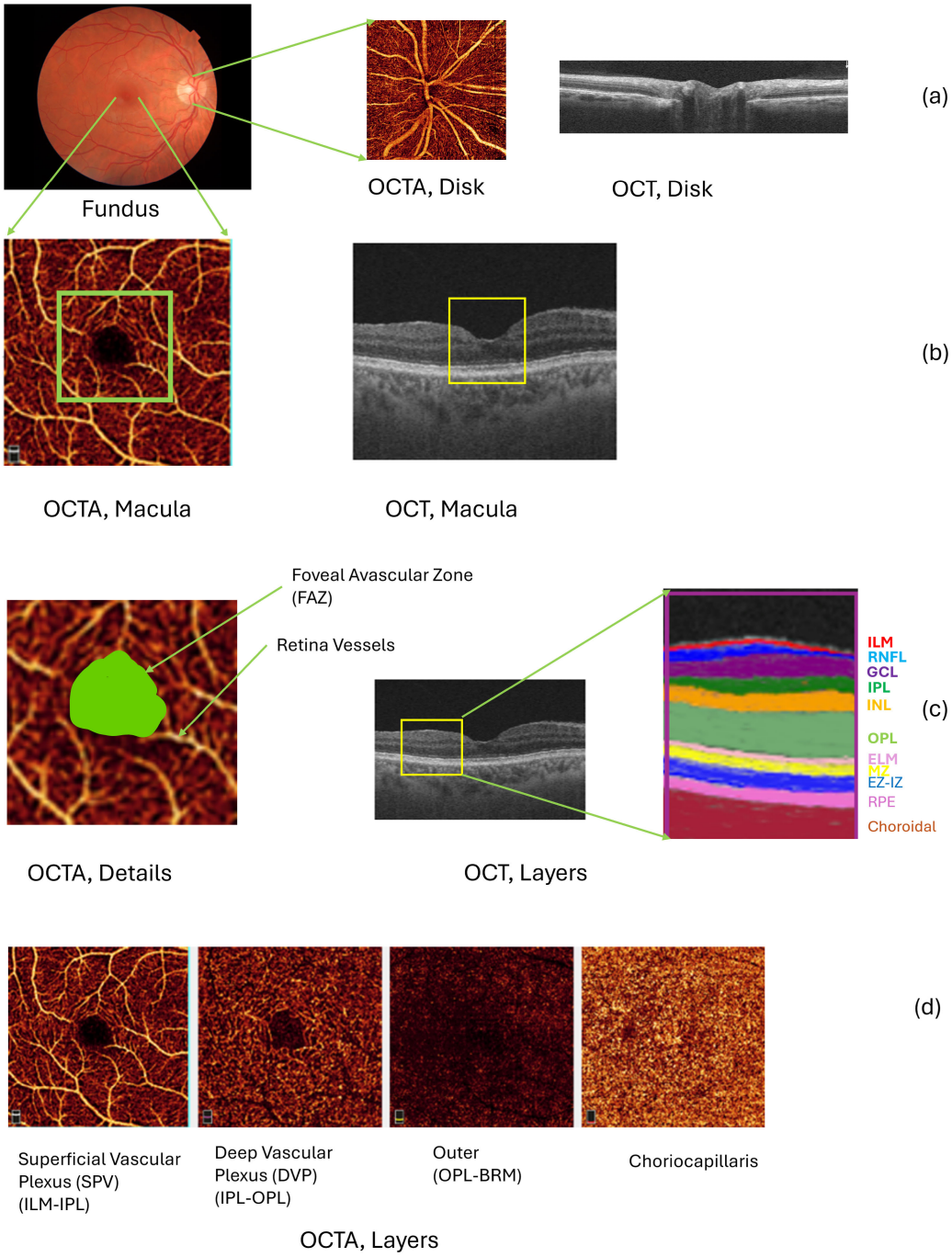

FIGURE 2. OCT and OCTA Modalities in detail. (a) Linkage of a fundus and its disk area shown by OCTA and OCT. (b) Enface OCTA image and one of the OCT slices of the macula area. (c) Enface OCTA image showing the Foveal Avascular Area (FAZ) in the middle with retinal blood and peripapillary vessels. The retina layers in the yellow box colored and explained in detail. (d) OCTA enface images of four layers of the retina. Superficial Vascular Plexus (SVP): the first layers of the retina (ILM, RNFL, GCL, and IPL). Deep Vascular Plexus (DVP): deeper layers of retina (IPL,INL, OPL). The outer: avascular layers of the retina. Choriocapillaris: bottom layers of the retina (Choroidal layer in OCT).

performed an umbrella review of all systematic reviews and meta-analyses on ocular biomarkers of AD between 2000-2021.

The following are the most commonly studied OCT-based biomarkers we compiled from these reviews:

- The most significant biomarker observation noted in most surveys was the thickness of the retinal layers [39]. Figure 2(c) shows the retinal layers. Retinal thickness studies often refer to the inner layers such as the Retinal Nerve Fiber Layer (RNFL), Ganglion Cell Layer (GCL),

and Inner Plexiform Layer. These three layers are known as the ganglion cell complexes. Ganglion cells are vulnerable to neurodegeneration because of their related mitochondrial dysfunction and unique composition of axons without myelination [30]. Recent studies have shown that the GCL thickness is more likely to be associated with AD and MCI. [40], [41].

- Secondly, the highlighted observations were on macular volume and thickness changes [27]. The macula provides sharp, clear, straight-ahead vision. It is responsible for central and color vision [42]. Figure 2(b) shows the macula region of the eye.
- Some studies have reported changes in choroidal volume and thickness [27], wherein the choroidal layer supplies nutrients to the retina and maintains the temperature and volume of the eye [42].

The following are the most commonly studied OCTA-based biomarkers we compiled from these studies:

- Changes in the Foveal Avascular Zone (FAZ) [27], [38]. The fovea, which marks the center of the retina [42], [43] is shown in Figure 2(c).
- Changes in vessel density [27].
- Changes in perfusion and vessel density were also observed. Vessel perfusion density is defined as the total area of perfused vasculature per unit area [43], [44].
- Changes in peripapillary vessel density, another parameter that requires further research, were discussed in [27]. Radial peripapillary capillaries are distinctive vascular networks within the RNFL around the optic disc. The retinal layers are shown in Figure 2c.

In their comprehensive systematic review conducted in 2023, Ibrahim et al. [38] examined retinal biomarkers for AD and MCI diagnosis using OCT/OCTA imaging techniques. They observed that the inconsistency in findings across different studies can be primarily attributed to the small sizes of the study cohorts and the reliance on certain OCT/OCTA device brands. These factors led to inconsistent results and also showed that some parameters were significant only when specific OCT machines were used.

AD investigations have included studies using mice, that exhibit a genetic similarity of 99% to humans [45]. The utilization of mice in research holds notable importance as it enables longitudinal and age-related inquiries into AD, with approximately nine mouse days equating to one human year [45]. Specifically, triple-transgenic mouse models (3xTg-AD or TMM) emulate crucial human AD characteristics, such as the aggregation of amyloid-β and tau proteins. A systematic review by Sánchez-Puebla et al. [46] focused on OCT imaging in mouse models of AD. This review examined OCT scans of healthy wild-type (WT) mice and TMM of AD. Their analysis of 14 eligible studies revealed a consistent focus on retinal layer thicknesses. Similar to human studies, inconsistent results were observed mainly due to the localization of retinal layers, particularly in the macula and peripapillary regions. Besides, six of these studies used the same mouse model.

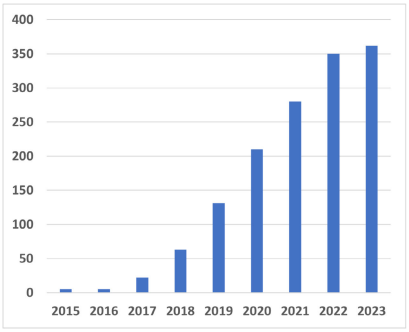

FIGURE 3. PubMed statistics of DL, OCT and OCTA applications in years.

III. DEEP LEARNING-DRIVEN FLOW IN OCT/OCTA STUDIES

Classical machine learning techniques require four common steps: feature extraction, dimensionality reduction, algorithm/model selection, and feature-based learning. Artificial neural networks are a subset of AI inspired by the simplification of neurons and their connections in the brain. DL is a multi-layer structure of neural networks that learns features and high-dimensional non-linear functions by mimicking neuron connections and interactions in the human brain. After the 2010s, DL gained momentum from Machine Learning [47] because it can learn directly from raw data by implicitly learning feature extraction, representation, and classification tasks.

Tian et al. [18] reviewed the applications of the DL-based approach to different biomedical optical imaging fields such as microscopy, fluorescence lifetime imaging, in vivo microscopy, widefield endoscopy, photoacoustic imaging and sensing, diffuse tomography, and functional optical brain imaging. Unsurprisingly, DL has also become the latest trend in analyzing OCT and OCTA scans [15], [16], [17]. Figure 3 shows the number of DL studies on OCT and OCTA images reported in PubMed over the last seven years. Enhancements in image quality, segmentation, clinical diagnosis, and prognosis detection represent key areas where deep learning has been significantly applied in the OCT/OCTA field [15], [16], [18]. Despite this, research connecting DL with AD and Mild Cognitive Impairment (MCI) diagnosis remains minimal, accounting for only 1% of the studies.

Deep learning models require a substantial amount of data for generalization. Despite the challenges associated with limited datasets, lack of standardized image collection and evaluation metrics, and computational limitations as outlined by Yanagihara et al. [14], deep learning models have recently been favored over traditional ML counterparts in OCT and OCTA applications. We structured our analysis of the reviewed studies within the framework of Deep Learning Flow for AD/MCI diagnosis in OCT-OCTA as shown in Figure 4. This framework enabled us to categorize DL studies into key stages, such as dataset curation practices, model training techniques, and validation methods. Using this framework, we extracted valuable insights into the strengths, weaknesses, and prevailing trends of this domain.

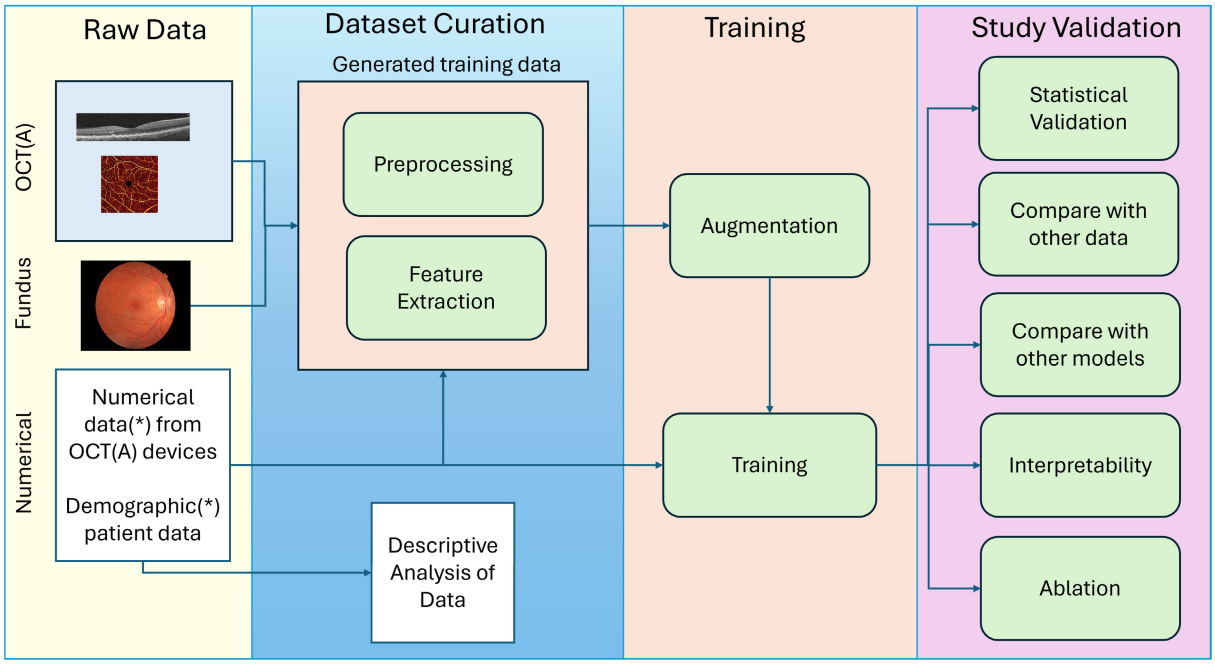

FIGURE 4. The framework of Deep Learning-driven flow for AD/MCI diagnosis in OCT/OCTA Studies showing common processes in dataset curation, training, and validation phases.

A. DATASET CURATION

Dataset curation is a critical step in deep learning research, which requires considerable time and effort. Medical studies should comply with global (tenants of the Declaration of Helsinki) and local ethical regulations and guidelines. Written informed consent is obtained from all participants. To determine their level of AD/MCI, all participants are subjected to clinical neurological and/or neuropsychological evaluations. Control groups are formed using various data inclusion criteria, such as matching age, sex, and education level. Not all individuals diagnosed with AD/MCI are included in deep learning studies. Some are excluded if they have a history or signs of other neurological or psychiatric conditions or different forms of dementia. Particularly in OCT/OCTA research, the ophthalmologists screen out individuals with additional conditions, such as diabetes mellitus, systemic arterial hypertension, cardiovascular diseases, retinal disorders, retinal diseases (such as glaucoma), or eye issues (including trauma, surgery, or inflammation). Even if data collection is not part of the study, the dataset used must be evaluated for eligibility based on inclusion and/or exclusion criteria, collection methodology and standards, and how well the data represents the entire population. The dataset distribution (percentages and p-values) is usually reviewed in terms of eyes used (left/right), age, sex, years of education, and AD/MCI scores in descriptive analysis [48].

OCT and OCTA scans consist of slices (B-scans) of two-dimensional images that form the three-dimensional volumetric data. Scanning devices can also generate numerical measurements based on these scans. Therefore, these studies provide a rich source of alternatives for feature extraction. Some later studies preferred directly using images, whereas others applied various feature extraction processes. Similar to clinical diagnosis, demographics and patients' other medical details, such as sex, age, and test results, are also included in the model datasets.

Raw OCT slice images are susceptible to speckle noise due to coherent light scattering, while low signal levels in some devices result in other types of noise. Because the retina is non-rigid, a moving layer of tissue inside a mobile body, non-standard frames, motion blur, and artifacts may occur. Moreover, patients with AD/MCI may not follow the scanning procedure instructions effectively, resulting in frequent motion artifacts in the images. In response, researchers have utilized various image improvement techniques [15], [16], [18]. Preprocessing, such as enhancing the image quality using computer vision (CV) techniques before ML or DL, is a popular topic in medical imaging research.

B. MODELS AND TRAINING

In machine learning and deep learning, the development and optimization of models play a crucial role in achieving accurate and reliable outcomes. Convolutional neural networks (CNN) have become the standard for feature extraction, allowing DL models to undergo end-to-end training directly from the input images. Recent methods, such as attention mechanisms [49] and transformers [50], are leading the way in training these models, using innovations to improve their performance. Attention-based models, initially designed to improve recurrent neural networks (RNN) for machine

translation, have broadened their applications across various domains, including language, speech, and vision. They enhance the deep network performance in numerous pattern recognition tasks by enabling easier visualization and interpretation of network outputs. For instance, attention maps provide a visual representation of the relative importance of different areas in the input data, aiding in understanding the decision-making process of the model. Furthermore, the introduction of the transformer architecture [51], [52] for images has marked a significant leap forward. These self-attention-powered models can process input-output representations without requiring conventional CNNs or RNNs, boosting efficiency and interpretability in pattern and image recognition tasks.

ML and particularly DL models require large amounts of data (examples) to generalize learned tasks. This requirement makes overfitting a significant concern when using small datasets. The most widespread solution to this issue is to increase the size of the image datasets by producing new examples using data augmentation. Another widely used method is transfer learning, which involves copying and importing weight values from the lower layers of a network previously trained on a large generic image database, such as ImageNet [53] or alternatively (or additionally) on a similar larger dataset of another condition. After transferring the weights, the network was fine-tuned on the study data for a few iterations to complete training. However, using ImageNet pre-trained networks imposes some restrictions, such as the requirement for three-channel (Red, Green, Blue) input images of a certain size, which are not fully compatible with the 3D volumes generated by OCT/OCTA scanners.

C. STUDY VALIDATION

A rigorously and methodologically validated machine learning model can provide the reliability required for practical deployment in medical settings. The main concerns are the generalizability and repeatability of the results [54]. Numerous factors influence the generalizability of the findings, including the dataset size, composition, diversity, and methodology employed for partitioning data into training and test sets. Several studies have presented accuracy metrics by using separate (hold-out) test datasets. A more reliable method often employed is N-fold cross-validation, where the dataset is divided into N random folds. Subsequently, the model undergoes training on N-1 folds for each iteration, with the remaining fold designated for testing. The aggregated test results from all iterations are then averaged and reported, along with their mean and standard deviation.

Importantly, patient-based splitting is necessary to prevent data leakage. For example, including slices from the same OCT scan in both the training and testing sets or augmenting the entire dataset before splitting would compromise validation [55]. Similarly, using one eye's OCT scan for training and the other for testing from the same patient could be considered leakage.

Data imbalance, prevalent in medical scenarios in which one or more target conditions are less frequent, presents another challenge beyond the limited dataset size. Thus, the distribution of classes within the test dataset becomes critical. For diagnostic tasks such as Alzheimer's vs. healthy person, metrics such as precision (1-specificity), recall (sensitivity), and F1-score provide views of performance from the perspective of both classes. A common measure, the Area Under the Receiver Operating Characteristic Curve (AUROC), reflects the ranking ability of the method, independent of the decision thresholds. Additionally, Cohen's Kappa [56] offers a measure of agreement between the predicted labels and expert annotations (or agreement between multiple experts). Adjusting for chance agreement enhances the interpretability of model accuracy for categorical outcomes [57]. In multi-class scenarios, confusion matrices summarize the frequency of incorrect classifications between different classes.

Furthermore, statistical methods, such as t-tests, are used to compare metrics across different models or conditions. The associated p-value indicates whether the observed differences in the means are statistically significant, which is essential for validating the effectiveness of new diagnostic tools or algorithms against existing standards or competing technologies [58].

Despite their high performance, DL architectures are often regarded as black-box models. Ensuring trust in predictions is crucial for clinical decision-making. In response to this challenge, researchers are actively developing and employing interpretability and explainability tools, such as Deep Taylor Decomposition [59] and Grad-CAM [60]. Interpretability and explainability, although often used interchangeably in the context of deep learning (DL) models, have distinct meanings. Interpretability focuses on understanding the inner workings of the models, whereas explainability focuses on explaining the decisions made. Consequently, interpretability demands a deeper level of detail than explainability does. In medical imaging, explainability refers to the extent to which a healthcare professional can understand the cause of a classification decision made by using a model. Accurate and comprehensive explanations of DL model decisions can enhance trust and adoption among healthcare professionals. In AD studies, explainability in terms of highlighting the most salient discriminative regions in the input could aid in the localization of AD-related changes and abnormalities in the retina. Furthermore, this could facilitate the discovery of new AD-related biomarkers. Although these techniques provide insights into model reasoning, they are limited in highlighting relevant regions within the input data [61]. With the advancement of large language models, there is a growing trend towards the generation of human-readable text-based explanations [62].

Recent DL studies are pushing the boundaries by integrating complex features into advanced layers and architectures. To understand the contribution of these elements, ablation

studies are conducted, where certain features or components are intentionally omitted. This helps to verify their impact on the model's performance by comparing the outcomes with and without these specific aspects [63], [64].

IV. METHODS

To review ML/DL-based approaches for AD or MCI diagnosis in OCT and/or OCTA scans, we followed the guidelines of the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) [20]. Figure 5 shows the PRISMA flow chart for the systematic review.

Information Sources: We performed an exhaustive search of PubMed, Web of Science, Scopus, Google Scholar, Semantic Scholar, and CrossRef databases for relevant studies published between 2015-2024.

Search Strategy: We surveyed the databases above using the following combinations of terms:

("Alzheimer's" OR "dementia" OR "cognitive impairment") AND ("optical coherence tomography" OR "optical coherence tomography angiography" OR "retinal imaging") AND ("neural networks" OR "machine learning" OR "deep learning")

Eligibility Criteria: Articles meeting all the criteria below were selected for analysis:

- The article was available in English.
- An English abstract was available for non-English language articles.
- It was published as a primary research paper in a peer-reviewed journal. Conference papers, duplicates, datasets, book chapters, and articles that provided only statistical analyses were excluded.
- It described an ML/DL model for AD detection, screening, or prediction using only OCT/OCTA scan images or data derived from these images. Articles related only to segmentation or image quality improvement were excluded.
- It focused solely on AD and/or MCI (not other diseases such as age-related macular degeneration, drusen, etc.).

Data Extraction, Quality Assessment, and Bias Analysis: We searched the databases (PubMed, Web of Science, Scopus, Google Scholar, Semantic Scholar, and CrossRef) to identify studies that matched our search strategy. We then cross-checked and identified the studies that met our eligibility criteria. For eligible studies, a detailed analysis was performed to obtain critical information such as the number of participants, year of publication, algorithms applied and their characteristics, model prediction parameters, and performance metrics (including accuracy, discrimination sensitivity, and specificity rates). We did not use QUADAS-2 [65] for bias and validity analysis in our study because its construct validity and the reliability of its items for measuring study quality have not been confirmed [66], particularly in Alzheimer's disease diagnostic accuracy analyses where sample sizes are typically small.

V. RESULTS

Our search initially retrieved 4006 references. After applying the elimination steps shown in Figure 5, we identified 16 studies that met our inclusion criteria. Two of these were previously identified by Bourkhime et al. [19]. In our survey, we identified five animal (mouse) and eleven human studies. Four animal studies were conducted by a group of researchers at Coimbra University, utilizing the same data. These investigations were longitudinal studies involving OCT and Fundus imaging of mice aged between 1 and 16 months. However, there are extensive inclusion and exclusion criteria for patients in human studies. Additional checks are required to ensure the sample dataset correctly represents the total population. Due to these design differences between human and mouse studies, we analyzed mouse studies in a separate section.

A. DATASET CURATION

Tables 1 and 2 present a detailed overview of each study, including the data sources, study designs, participant demographics, scanning devices, data exclusion rules, descriptive data analysis, main variables studied, feature selection processes, and data preprocessing steps.

We observed that clinics collected local datasets from a limited number of patients for their research. There is a lack of openly available OCT/OCTA datasets for patients with AD and MCI. Although the data collected increased in later studies, the numbers were still very low for the deep learning models. Individuals with MCI are at an increased risk of developing AD or other forms of dementia, making early detection and intervention particularly important. Studies conducted after 2023 [67], [68], [69] included MCI patients in their work as well.

We encountered a lack of consistency in data collection standards. The datasets originated from various clinics, each with distinct AD, MCI, and Healthy Cohort (HC) numbers. Furthermore, the exclusion criteria detailed in Section III-A varied between studies, with some studies not mentioning these criteria altogether. All studies listed their data sources and inclusion criteria, while five failed to explain their data exclusion rules [69], [70], [71], [72], [73]. Numerous studies enhanced their analyses by incorporating demographic information to alleviate the biases arising from small dataset sizes. However, such descriptive analyses were absent in these five studies [69], [70], [71], [72], [73], where exclusion rules were also omitted. OCT and/or OCTA-driven numerical data were also incorporated in the descriptive analysis in some studies [67], [68], [72], [74]. Inaddition, OCT and OCTA scans were obtained using different devices across studies. These discrepancies resulted in disparate data domains across the studies.

The limited number of cases in small OCT datasets poses challenges for DL networks in identifying features

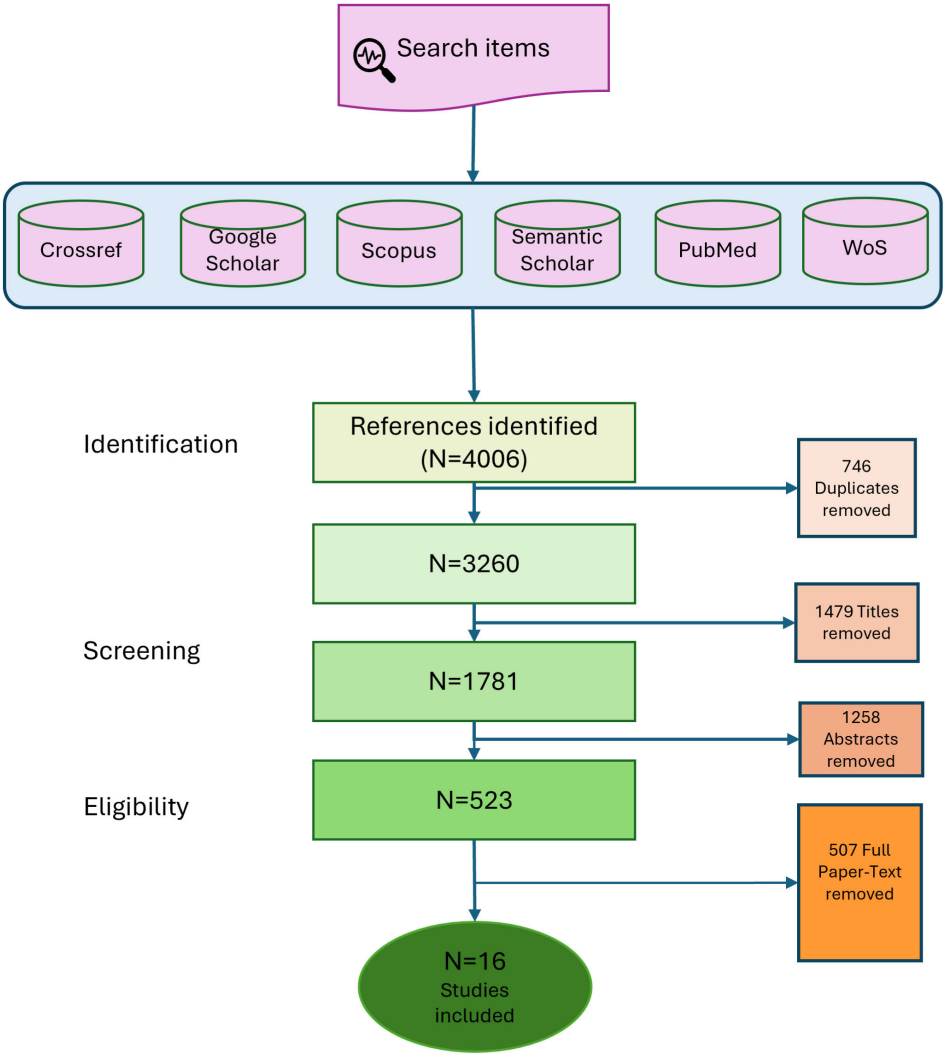

FIGURE 5. Flow diagram of eligible study selection from a search across PubMed, Web of Science, Scopus, Google Scholar, Semantic Scholar, and CrossRef databases yielded 4006 deep learning (DL) publications related to OCT and OCTA applications between 2015 to 2024 (End of Feb. 2024).

associated with complex conditions such as AD or MCI. Moreover, researchers needed to exclude some of the scans further due to low signal quality, noise, and motion artifacts. Therefore, some studies integrated additional information, including descriptive patient data and quantitative data derived from OCT and/or OCTA scans, such as Retinal Nerve Fiber Layer (RNFL) thickness, into their models to address this challenge [72]. Wisely et al. [6], [67] compared various input combinations, with the best single input identified as the GC-IPL thickness image obtained from OCT. Combining multimodal retinal images, OCT and OCTA quantitative data, and patient data only marginally improved performance. However, the best performance was obtained using only the numerical data obtained from OCT devices. Yoon et al. [75] applied another strategy by combining multiple quantitative radiomic features and patient data with OCTA images.

Earlier studies [72], [73], [76] predominantly used numerical data derived from OCT and OCTA images. Similar to animal studies, Nunes et al. [76] generated Mean Value Fundus (MVF) [77] images from each layer, followed by the computation of a feature vector based on the gray-level co-occurrence matrix (GLCM) [78]. In later studies, OCTA images were exclusively utilized by Wang et al. [79], Liu et al. [70], and Gao et al. [69], suggesting a potential trend towards relying solely on image inputs.

All studies after 2022, except those by Wang et al. [71] and Gao et al. [69], successfully leveraged OCTA images in their AD detection studies. While Wang et al. [79] used Superficial (SVP or SVC) and Deep (DVP or DVC) Vascular Plexus images of the FAZ area directly, Gao et al. [69] had a different approach by using the polar transformed SVC and DVC images. Zhao et al. [68] and Yoon et al. [75] used automated segmentation of the FAZ area. Yoon et al. [75]

TABLE 1. Dataset curation protocol details of selected studies. (Part 1).

Author	Source Data	Study Design	Scanned Device	Data Exclusion Rule
Wisely, 2024 [67]	Duke Alzheimer's Disease Research Center	129 HC 80 MCI	Zeiss Cirrus HD-5000 pectral-Domain OCT with AngioPlex OCTA.	**Neurological Conditions:** dementia, demyelinating disorders, **Ophthalmological Conditions:** glaucoma, AMD, other vitreoretinal pathology, history of ocular surgery apart from cataract surgery, high myopia or hyperopia, **Medical Conditions:** diabetes mellitus, uncontrolled hypertension, **Visual Acuity and Sensory Impairments:** corrected ETDRS visual acuity < 20/40 at the time of image acquisition.
Yoon, 2024 [75]	Samsung Medical Center	48 HC 37 AD	DRI OCT Triton Topcon	**Neurological Conditions:** Dementia types other than AD, traumatic brain injury, normal pressure hydrocephalus, territorial infarction, white matter hyperintensities due to etiologies non-vascular etiologies
Zhao, 2023 [68]	Hospital of Ningbo Uni.(1), West China Hosp.(2), Zhejiang Uni.(3), Ningbo Uni.(4)	ROAD: 631 HC, 239 AD, ROMCI: 455 HC, 96 MCI	N/A	ROAD (1)(2): **Neurological Conditions:**brain disorders (both neurologic and psychiatric, except for seizures thought to be related to EOAD or headaches), MRI evidence of infection, focal lesions such as strokes, multiple or strategic lacunae, and/or space-occupying lesions, **Psychiatric Conditions:** History of substance abuse or suicidal behaviors within the past 12 months, history of participation in therapeutic trials, **Miscellaneous:** Pregnancy or lactating status ROMCI (3)(4): **Psychiatric Conditions:** major depression or other psychiatric disorders. **Visual Acuity and Sensory Impairments:** significant sensory impairment such as verbal or hearing impairment.
Gao, 2023 [69]	1) Army Medical University 2) cross-sectional study (2021-086-K-73-01)	1)50 HC 38 AD 29 NC, 2)133 HC, 140 MCI	RetiCam 3100. VG 200 SS-OCT Topcon	N/A
Liu, 2023 [70]	N/A	291 HC 114 AD	VG200S SVision Imaging	N/A
Wang, 2022 [79]	West China Hospital	286 HC 114 AD	SS-OCTA (VG200 SVision Imaging)	N/A
Wang, 2022 [71]	Xiangya Hospital	299 HC 159 AD	Cirrus HD-OCT 4000; Carl Zeiss Meditec	**Neurological Conditions:** Unclearly specified neurological conditions, **Psychiatric Conditions:** Unclearly specified psychiatric conditions, **Ophthalmological Conditions:** cataract, glaucoma, uveitis, epiretinal membrane, macular hole, AMD, history of eye trauma, any eye surgery, best-corrected visual acuity (BCVA) <0.5, refractive spherical equivalent >6.00 D, astigmatism >3.00 D, intraocular pressure (IOP) >21 mmHg **Medical Conditions:** diabetes mellitus, uncontrolled hypertension
Wisely, 2022 [6]	Duke University (NCT03233646)	123 HC 36 AD	Zeiss Cirrus HD-5000 Spectral-Domain OCT with AngioPlex OCTA, UWF SLO (Optos)	**Neurological Conditions:** dementia, demyelinating disorders, **Medical Conditions:** diabetes mellitus, uncontrolled hypertension, **Ophthalmological Conditions:** high myopia or hyperopia, glaucoma, AMD, other vitreoretinal pathology, ocular surgery apart from cataract surgery, **Visual Acuity and Sensory Impairments:** corrected ETDRS visual acuity < 20/40 at the time of image acquisition.
Lemmens, 2020 [72]	University Hospitals UZ Leuven	22 HC 10 AD 7 amyloid AD	RT-vue XR Avanti (Optovue)	N/A
Sandeep, 2019 [73]	College of Engineering Trivandrum Kerala	25 AD 25 HC	N/A	N/A
Nunes, 2019 [76]	University of Coimbra	N/A	Cirrus SD-OCT 5000	**Ophthalmological Conditions:** Retinal injuries, Retinopathies, optic neuropathies. glaucoma, AMD, severe visual impairment, **Medical Conditions:** diabetes

derived "randomic features" of the FAZ area in their study, and Zhao et al. [68] added microvasculature segmentation as well to the features of their training.

B. MODELS AND TRAINING

Table 3 presents the models and training methodologies employed in each study, describing the augmentation techniques, data-splitting strategies, and ML/DL algorithms used in each study. Additionally, it highlights the utilization of transfer learning techniques, where applicable.

We observed a change in the training models and setup that aligned with the developments in the field. Early studies mainly incorporated numerical and/or derived numerical features to train ML models [72], [73], [76]. In contrast, later

TABLE 2. Dataset curation protocol details of selected studies.(Part 2).

Author	Descriptive Analysis of Data	Image Modalities	Feature Selection	Data Preprocessing
Wisely, 2024 [67]	Demographic Data for MCI and HC, Quantitative OCTA Summary Data, Quantitative OCT Summary Data	**Images:** OCTA, GC-IPL, **Quantitative:** OCT/OCTA, **Demographic:** Patient data	**Images:** GC-IPL thickness color maps, OCTA, **Quantitative:** OCT/OCTA Data, **Demographic:** Sex, Age, Years of Education, logMAR Vai, MMSE Score, **Numerical**	Manually low-quality images excluded signal/strength (<7/10), shadow artifact, segmentation and motion artifacts, and poor centration
Yoon, 2024 [75]	Demographic Data for AD and HC (Age, Sex, MMSE, Study period, Clinical dementia rating, Patients w/hypertension, Patients/w diabetes)	**Images:** OCTA, **Demographic:** Patient data	**Images:** OCTA entered on the FAZ, Manual and AI-based segmentation, **Quantitative:** Radiomic features: Area, Soldicity, Compact, Round, Eccent, **Demographic:** Sex, Age	Manually low-quality and noisy images were excluded.
Zhao, 2023 [68]	Demographic data for AD, HC, and MCI (Age, sex, hypertension, diabetes, education level, MMSE, MoCA), Global and regional statistical analysis of the OCTA parameters	**Images:** OCTA	**Images:**OCTA: 3 enface images of the volume taken from FAZ	Manually low-quality images excluded. Patients with eye disease or ophthalmic surgery history are excluded.
Gao, 2023 [69]	N/A	**Images:** Fundus OCT	**Images:** Fundus: Preprocessed, OCT: Centered on the corneal apex, **Quantitative:** Modality Feature Extraction (MFE)	Fundus: Background mask, Image resize, CLAHE OCT: Background mask, noise reduction hybrid filtering (bilateral and medial)
Liu, 2023 [70]	N/A	**Images:** OCTA images	**Images:** OCTA: SVC, DVC and CC projections of 3x3 mm2 area centered on the fovea	Transform the images into polar coordinates with the FAZ as the origin
Wang, 2022 [79]	N/A	**Images:** OCTA	**Images:** OCTA: Projection layers of OCTA Scans (SVC -DVC)	Resize the images
Wang, 2022 [71]	**Demographic:** (Age, sex, education, disease duration, MMSE, MoCa, CDR, IOP, BCVA, AL)	N/A	**Quantitative:** OCT, CSF biomarkers, MRI Analysis, APOE genotyping, **Demographic:** Patient data	Manually low-quality images excluded. Patients with eye retinal abnormalities and preexisting systemic diseases that may affect vision excluded
Wisely, 2022 [6]	N/A	**Images:** OCTA, UWF, FAF, **Quantitative:** OCT/OCTA, **Demographic:** patient data	**Images:** OCTA, GC-IPL, FAF, UWF, **Quantitative:** OCT/OCTA data OCT:thicknesses of mean GC-IPL, mean RNFL, Subfoveal Choroidal, Central Subfield, **Demographic:** patient data	Manually low-quality images excluded based on signal/strength (<7/10), shadow artifact, segmentation and motion artifacts and poor centration images excluded
Lemmens, 2020 [72]	Demographical and clinical data: Age, Sex, body mass index, eye, MMSE, BCVA, IOP, Phakic, Quantitative OCT data	**Images:** Hyperspectral, **Quantitative:** OCT	**Quantitative:** Standardization vector from ROI (Region of interest from Hyperspectral image in 4 quadrants), OCT- Numerical Data	Manually low-quality images excluded.
Sandeep, 2019 [73]	N/A	**Images:** OCT	**Quantitative:** Morphological Features from segmented RNFL	Resize, Noise reduction, distortion removal, segmentation
Nunes, 2019 [76]	Demographical: Age, Sex, Right/Left Eye	**Images:** Fundus, **Quantitative:** OCT Layers	**Quantitative:** MVF Image calculated for each OCT layers, GLCM (80 calculated features from MVF)	The segmentation of layers generated by the OCT Explorer Software was manually corrected.

studies relied on deep learning models despite the associated challenges listed by Yanagihara et al. [14].

After 2022, the studies started to use images directly in their training and feature extraction with convolutional networks. Wisely et al. [6], [67] in both of their studies, employed a convolutional neural network-based architecture processing a combination of OCTA images and other data to predict AD. Recent studies [6], [68], [69], [70], [79] have demonstrated the use of attention in OCT/OCTA-based AD diagnosis. Although transfer learning from ImageNet was more popular for the feature extraction part of the networks [6], [67], [69], Xie et al. [74] used transfer learning from other OCTA segmentation networks called OCTA-NET and FAZ-Net. However, we did not include this study in our list of eligible studies because the results were not used for AD classification. The only example of a UNet used for segmentation was by Yoon et al. [75]

Training with small data sets was the biggest challenge in these studies. Our review revealed that the most widespread solution to this issue was to increase the size of the datasets by producing new instances using data augmentation [6], [70], [79]. Although there were studies in our survey that transferred their weights from a range of pre-trained networks [6], [67], [69], [75], none of these transfer learning examples could use 3D OCT scans directly as inputs because the source networks were trained on generic two-dimensional

TABLE 3. Details of the models and training procedures of the selected models. Acc: Accuracy, Sens: Sensitivity, Spec=Specificity, AUC: Area under ROC curve.

Author Year	Augmentation	Data-Split	ML/DL Algorithms	Transfer Learning	Classification Results
Wisely 2024	N/A	HC: Train 64%, Val 10%, Test 25% MCI: Train 68%, Val 13%, Test 19%	Feature Extraction: Convolutional Networks (ResNet18), Classification: FC	ImageNet	AUC= 0.96
Yoon 2024	N/A	Train 80%, Val 20%, Holdout test 64%	Feature Extraction: Segmentation (nnUNET), Classification: Gradient Boost Decision Tree	N/A	Acc=0.65, Sens=0.54, Spec=0.94, AUC=0.72
Zhao 2023	Random horizontal & vertical flip	80% Train, 20% Val	Feature Extraction: Eye-AD: Convolutional networks (ResNet18), Attention: Importance-aware graph neural networks, Classification: MLP	N/A	Acc = 0.89, Pre=0.89 F1 score=0.89 Kappa=0.7 AUC=0.94
Gao 2023	Random cropping, reversing, Standardization (mean, std)	80% Train, 20% Test	Feature Extraction: DuCAN: Convolutional networks, Attention: Self Attention & Cross attention, Classification: MLP	Train with 1st data set and transfer learning with the 2nd data set	Acc =0.92, AUC=0.96
Liu 2023	Drift cropping, random rotation, resize	N/A	Feature Extraction: POLARNET: Convolutional Based (PFEM), Attention: PRIM	N/A	Acc = 0.85, AUC =0.85, Kappa = 0.58
Wang 2022	Random rotation, flipping, contrast adjustment, adding Gaussian Noise	Train 80%, Test 20%	Feature Extraction: MUCONe, Convolutional Layers, Attention: CsCp Consistency and Complementary Attention, Classification: Cross View Fusion	N/A	Acc=0.86, F1 score=0.86, Kappa 0.67
Wang 2022	N/A	Train 70%, Test 30%	Feature Extraction: XGBoost, Light GBM, KNN, Random Forest, Gradient Boost, AdaBoost, Classification: Logistic Regression	N/A	Acc=0.74, AUC=0.69. F1 score=0.7, recall 0.74
Wisely 2022	Rotating, translating, scaling, flipping, distorting, adding Gaussian noise	Train 73%, Test 27%	Feature Extraction: Convolutional Networks (ResNet18), Classification: FC	ImageNet	AUC=0.841
Lemmens 2020	N/A	N/A	Classification: Linear Discriminant Analysis	N/A	AUC=0.74
Sandeep 2019	N/A	N/A	Feature Extraction: Fixed-Grid Wavenet Network, Classification: Radial Basis Function of Neural Networks	N/A	False Acceptance Rate=1% , False Rejection Rate=1%
Nunes 2019	N/A	N/A	Classification: SVM	N/A	Sens=0.89, Spec=0.84, Acc=0.82

(RGB) color images, which comprise ImageNet-like databases.

The classification results of recent deep learning (DL) studies were promising when each individual study was considered. Eight studies reported Area Under the Curve (AUC) results. Wisely et al. [67] and Gao et al. [69] reported the highest AUCs of 0.96, whereas Lemmens et al. [72] reported the lowest AUC of 0.74. Among the seven studies that reported accuracy results, Gao et al. [69] achieved the highest accuracy of 0.92, and Yoon et al. [75] had the lowest accuracy of 0.65. Only Nunes et al. [76] and Yoon et al. [75] reported their sensitivity and specificity results, while Sandeep et al. [73] reported inverted measures of false acceptance and false rejection rates. Moderate to high Kappa results were observed in studies by Zhao et al. [68], Liu et al. [70], and Wang et al. [79]. F1-Scores of 0.89, 0.86, and 0.71 were reported in three respective studies [68], [71], [79].

C. VALIDATION OF THE STUDY

We detailed the validation methods and scores used to assess the accuracy and reliability of each study's findings, along with any external validation efforts that further verified the robustness of the models outside the initial testing environments in Table 4.

The only common practice observed in these studies was that the training results were calculated using a five-fold cross-validation technique. In their last work, only

TABLE 4. Validation strategies of the selected studies.

Author Year	Validation methods scores	External Validation	Compared with other methods	Interpretability	Ablation	Major outcomes
Wisely 2024	Median Performance of 10 runs: AUC, Acc., Sens, Spec, Confusion Matrix, p-values	N/A	N/A	N/A	Testing the model with 6 different feature combinations	The best prediction performance is obtained with image-based quantitative data rather than combining any other image features. GC-IPL has better performance than OCTA
Yoon 2024	5 Folds- Cross Val., Holdout Test: AUC, Acc., Sens, Spec, p-values	N/A	1) Baseline 1 & 2 Against proposed 2) XGBoost, Random Forest Against Proposed	N/A	Testing results with 4 different Feature combinations	The best prediction was obtained from derived randomic features of the FAZ area from AI-segmented OCTA images.
Zhao 2023	5 Folds- Cross Val.: AUC, Acc., Precision, F1-Score, Kappa	External datasets ROAD -II and ROMCI-II	ViT, SwinV2-T, Early fusion, Middle fusion, Late fusion, MCC, MUCO, GCN, GAT, UG-GAT	Pixel Level: Grad-CAM Image Level and Region Level	N/A	FAZ and its surrounding microvasculature are more sensitive biomarkers for Early-onset AD diagnosis
Gao 2023	5 Fold-Cross Val.: AUC, Acc., Sens, Spec	Transfer learning and test with the second data set.	1)ResNet34 as baseline 2)ViT, EfficientFormarV2, SwinTransformer	Grad-CAM	1)Compare dual stream to single stream of images 2) Compare the steams with the second data set.	Using both Fundus and OCT Images increases the performance of MCI detection. Interpretability results show that the network focuses on the last layer of the OCT images.
Liu 2023	5 Folds- Cross Val.: AUC, AUROC, Kappa	OCTA 500 dataset	ResNet-34, EfficientNet-B3, ConvNeXt-S, HorNet-SGF, VAN-B6, ViT-Base, SwinV2-T, MUCO-Net	GradCAM, Importance map regions within projection, Importance map in different projections.	(with polar data) ResNet-34, EfficientNet-B3, ViT-Base, MUCO-Net	AD can be detected using FAZ-centered Polar Images of OCTA projections. (SVP, DVP, and CC)
Wang 2022	5-Folds-Cross Val.: Acc, F1 score, Kappa	OCTA 500	Resnet 50, Early, Late, Middle fusion, MCC, CRBM	Grad-CAM	Gradually adding attention layers	SCV and DVC of FAZ projections can be used to detect dementia
Wang 2022	5 Folds-Cross Val.: Acc, AUC, f1 score, recall	N/A	N/A	N/A	N/A	OCT measurements are significantly correlated with MOCA and MMSE, MTA and PCA scores (cerebral atropy, tau and Abeta) and APEO genotypes.
Wisely 2022	AUC	N/A	N/A	GradCAM	Testing the model with 11 different feature combinations	Best performance obtained from GC-IPL with quantitative data and patient data
Lemmens 2020	Leave-one-out cross validation: AUC	N/A	N/A	N/A	Testing with 4 different regions and adding RNFL thickness later	L2 region with RNFL thickness values are the most significant.
Sandeep 2019	5 Folds-Cross Val. Segmentation: Acc, Precision, Sens, Spec Classification: False Acceptance Rate (FAR)	N/A	FGWN with NN and GVF	N/A	N/A	Segmentation of layers can be used for AD and PD classification
Nunes 2019	5 Folds-Cross Val.: Acc, Sensitivity, Specificity	N/A	N/A	N/A	Results are obtained for 9 different regions	GLCM parameters derived from OCT images can help to diagnose AD and PD

Wisely et al. [67] ran their models 10 repetitions and used the median performance to calculate their results. All the studies used additional metrics apart from accuracy. The later studies [67], [68], [69], [75] incorporated the most common metrics: Accuracy, Area Under Curve (AUC), Sensitivity, and Specificity. Wisley et al. [67] and Yoon et al. [75] added confidence calculations (p-values) to their results.

Another common practice was to verify the studies using various ablation methodologies. All of the studies apart from Wang et al. [71] and Sandeep et al. [73] applied these techniques.

The three studies by Wang et al. [79], Liu et al. [70], and Gao et al. [69], which used images directly, also verified their results using other datasets. These studies also used interpretability techniques and applied various ablation

TABLE 5. Animal study details of ML/DL models using OCT and OCTA data.

Author	Year	Dataset	Input Data	ML/DL Algorithms	Validation method	Major outcomes
Batista [84]	2023	57 3xTg-AD 57 WT mice	MVF	Segmentation: U Net Classification: Convolutional Neural Networks (Inception-v3)	Trained with mice aged 3-48 months and tested with 12 and 12-months	High accuracy of AD detection (>0.8) obtained from MVF images regardless of retinal layers
Guimaraes [81]	2022	60 3xTg-AD 57 WT mice	MVF	Classification: Dense Net pre-trained on ImageNet	Hold out and statistical validation	Age could be predicted from retinal images but there were differences in retinal aging in WT and 3xTg-AD mice
Ferreira [82]	2022	57 3xTg-AD 57 WT mice	MVF	Classification: Convolutional Neural Networks (Inception-v3)	Trained with mice aged 3-48 months and tested with 12 and 12-months	The best performance was obtained with the MVF images computed from the RNFL-GCL complex.
Sayeed [83]	2022	24 3xTg-AD 2256 WT mice	GLCM features from MVF	Classification: Decision Tree Neural network Random Forest SVM	N/A	SVM gave the best performance of AD classification from derived features.
Bernardes [77]	2017	22 3xTg-AD 23 WT mice	GLCM features from MVF	Classification: SVM	k-fold cross-validation	AD detection performance was better in 8-month-old mice than 4 months old mice

techniques. In our analysis, we observed that these three studies were well-designed examples of deep learning flows.

It is difficult to collect AD-specific data in isolation, therefore Liu et al. [70] and Wang et al. [79] used the public OCTA-500 dataset [80], whereas Nunes et al. [76] included Parkinson's disease in their original dataset. Zhao et al. [68] and Gao et al. [69] used additional AD-specific datasets to verify their results. They both observed relatively good results with transfer learning and fine-tuning in other domains.

The most frequently employed methods in terms of interpretability were Grad-CAM [6], [68], [69], [70], [79]. Singh et al. [59] evaluated 13 deep learning explainability methods for OCT scans and found that the Deep Taylor method outperformed others in diagnosing choroidal neovascularization (CNV), diabetic macular edema (DME), and Drusen. Interestingly, none of the AD-related studies in our review applied this method to explain classification decisions. However, Zhao et al. [68] stood out as they applied an additional interpretability technique called importance maps. Lie et al. [70] divided the image into early treatment of diabetic retinopathy study (ETDRS) grids. They analyzed the importance of each grid for diagnosis. Zhao et al. [68] extracted eight parameters from the images characterizing both the retinal microvasculature and foveal avascular zone (FAZ), such as vascular length density and vascular area density. Then, they constructed two additional interpretability analyses to show the importance of these parameters in diagnosis, first at the image level with different OCTA layers and second at the region level on enface images. Gao et al. [69] used mixed modalities (Fundus and OCT images) to detect AD. Their interpretability results showed that the network focused on the last OCT image layers (ELM to Chorodial).

D. ANIMAL STUDIES

In our analysis, we identified five mouse studies. Although they do not follow a standard framework like human studies, they still apply dataset curation, model training, and validation processes. We grouped these findings in Table 5 according to the study design, main variables used in the specified ML/DL algorithms, validation methods, and major outcomes.

All mice studies involved OCT and Fundus imaging of mice aged between 1 and 16 months. Although the dataset selection process was not as complicated as that in human studies, image quality was one of the biggest challenges. Researchers have utilized various image improvement techniques, including normalization, [81], [82], [83] histogram equalization, [82] contrast enhancement, [83] as well as Gaussian [82] and median filters [73] to enhance the quality of OCT scans. Subsequently, in later studies [84], deep learning-based (U-Net) models were employed for automated segmentation. In these studies, the MVF images and their GLCM features were generated. The GLCM features were then classified using a support vector machine (SVM) model, demonstrating the feasibility of AD detection in mice using OCT scans [77]. Remarkably, early detection was achievable with over 80% accuracy in 4-month-old mice, and over 90% accuracy in 8-month-old mice. Ferreira et al. [82] used MVF images for each layer directly to re-train separate convolutional networks (Inception v3) for AD classification. They performed best with MVF images computed using the RNFL-GCL complex. The only study that worked directly from OCT-scanned images was conducted by Guimares et al. [81]. They re-trained two separate DenseNet (ImageNet) models to predict the age of TMM and WT mice. They observed the possibility of age prediction for a mouse in both networks with reasonable accuracy.

Sayeed et al. [83] also used 24 3xTg-AD and 2256 WT mice to compare the classification accuracy of various ML algorithms, including decision trees, neural networks, random forests, and SVM. They used GLCM from a normalized covariance matrix of the fundus and OCT images. The SVM algorithm exhibited the best performance. The decision tree algorithm registered the second-best performance, whereas the neural network algorithm registered the worst.

VI. DISCUSSION

Neuroimaging tools, such as magnetic resonance imaging (MRI), positron emission tomography (PET), computerized tomography (CT), and electroencephalography (EEG) help identify Alzheimer's disease (AD). Interest has grown in optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA), which are promising tools for early AD diagnosis. These tools are more accessible, non-invasive, and cost-effective than alternatives (MRI, PET, CT, and EEG). OCT/OCTA generates high-dimensional, multimodal imaging data. It is time-consuming and challenging to interpret and classify. Therefore, the integration of computer-aided machine learning (ML) and deep learning (DL) approaches is required to aid reading and classification.

There are a few reviews in this field, each with a limited scope. Richardson et al. [85] briefly reviewed artificial intelligence in dementia, while Nahid et al. [86] explored imaging biomarkers for the early detection of AD using Deep Learning in neuroimaging and ophthalmology based on a limited number of studies. Bourkhime et al. [19] recently reviewed studies on machine learning and novel ophthalmologic biomarkers for AD screening and identified only two studies related to OCT images.

This systematic review used the PRISMA framework [20] to evaluate the effectiveness of machine learning (ML) and deep learning (DL) approaches in detecting or diagnosing Alzheimer's disease (AD) and Mild Cognitive impairment (MCI) using OCT and OCTA scans. A key difference from other related reviews is our structured approach through a ''Deep Learning Flow Framework'' explicitly tailored for AD/MCI diagnosis using OCT/OCTA, as depicted in Figure 4. This structure allowed us to classify DL studies into critical phases, including dataset curation practices, model training techniques, and validation methods. Unlike previous reviews that broadly focused on structural and functional imaging modalities, our review emphasizes the systematic application of deep learning techniques to OCT/OCTA data. Additionally, we examined both human and animal studies to provide a more comprehensive understanding of the current research landscape and highlight the translational potential of animal models in clinical practice. Our study identified 16 studies, including six that used mice as subjects. We reviewed the human and mouse studies separately.

Table 6 visually summarizes Tables 1, 2, 3 and 4, outlining our key findings on how each study conformed to the framework for ML and DL studies. We did not include the classification results in Table 6 for several reasons that hindered a direct comparison of their reported performances. (1) The models were trained and tested on highly distinct private datasets acquired using various devices. (2) Different methodologies were used to distribute instances between training and testing partitions. (3) The final classification decision might be based on the results from either one or both eyes. (4) Performance was reported using different metrics, such as accuracy, AUC, sensitivity-specificity, F-1 score, and Cohen's Kappa.

Building on this unique framework, we discuss our findings from the perspectives of performance, datasets, overall performance, deep learning models, and interpretability.

Datasets: A major issue with all the human-based studies we analyzed was the small sample size and lack of public data. The use of distinct private datasets in different studies prevents direct comparisons and makes it difficult to assess the generalizability of the classification performances achieved. Our review also highlights significant challenges in the dataset curation process, such as different scanning devices and inconsistent data inclusion/exclusion criteria for the cohorts. These variations resulted in inconsistent study domains.

We observed OCTA datasets and studies [68], [70], [75] in the years following 2023. They were aligned with medical protocols and concentrated on FAZ biomarkers for AD detection.

Overall Performances: The inconsistencies between the metrics and datasets used in different studies made it impossible to compare and validate the classification accuracies. We observed that all recent DL studies used AUC metrics in addition to other metrics such as (accuracy, Kappa, Sensitivity, Specificity, and F1-Score). Wisely et al. [6], [67] and Lemmens et al. [72] preferred to use a single metric (AUC). The best overall performance was achieved by Gao et al. [69] with high accuracy (0.92) and AUC (0.96). Most studies used additional features (image-based or quantitative) to increase performance. All studies, for those except Wang et al. [71] and Lemmens et al. [72] used both eyes (left and right) to double the dataset size. Nunes et al. [76] found that classification accuracy improved from 82% to 96% when both eyes received the same classification. Some studies, particularly those of Wang et al. [79], Liu et al. [70], and Gao et al. [69], demonstrated a more comprehensive benchmarking approach using external datasets, different models, interpretability, and ablation studies. Yoon et al. [75] tested their data using other models, and Wisely et al. [6] included interpretability results.

Deep Learning Models: Our analysis revealed a shift from classical ML pipelines to fully automated DL models for AD detection using OCT/OCTA images. We observed a shift from numerical features such as layer thickness measurements and textural metrics, to the direct processing of images as inputs. This transition was evident in recent mouse [81], [83], [87] and human studies [6], [67], [68], [69], [70], [75], [79] built on end-to-end DL architectures. Furthermore, we observed a correlation between the increasing complexity

TABLE 6. High-level summary of Tables 1, 2, 3 and 4, showing how the studies comply with the framework details explained in deep learning-driven flow for AD/MCI diagnosis in OCT/OCTA. Green: the full process is observed in the study, Red: the process is not mentioned or observed in the paper. Amber: no evidence of any exclusion criteria is observed even with inclusion criteria explained in the paper.

		Data curation				Training				Validation with					
	Author, Year	Data Selection Rules	Descriptive Analysis	Feature Selection	Preprocessing	Augment.	Feature Extraction via model	Attention	Transfer Learning	Other metrics	Other Data	Other Models	Interpret.	Ablation with Features	Ablation with Model parts
DEEP LEARNING	Wisely 2024	✓	✓	✓	✓	✓	✓	✓	✓	✓	⛔	⛔	⛔	✓	✓
	Yoon 2024	✓	✓	✓	✓	⛔	✓	⛔	✓	✓	⛔	✓	⛔	✓	⛔
	Zhao 2023	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	⛔	✓
	Gao 2023	⚠	⛔	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Liu 2023	⚠	⛔	✓	✓	✓	✓	✓	⛔	✓	✓	✓	✓	⛔	✓
	Wang Xingyu 2022	⚠	⛔	✓	✓	✓	✓	✓	⛔	✓	✓	✓	✓	⛔	✓
	Wisely 2022	✓	✓	✓	✓	✓	✓	✓	✓	✓	⛔	⛔	✓	✓	✓
MACHINE LEARNING	Wang Xin 2022	✓	✓	✓	✓	⛔	⛔	⛔	⛔	✓	⛔	✓	⛔	⛔	⛔
	Lemmens 2020	⚠	✓	✓	✓	⛔	⛔	⛔	⛔	✓	⛔	⛔	⛔	✓	✓
	Sandeep 2019	⚠	⛔	✓	✓	⛔	✓	⛔	⛔	✓	⛔	⛔	⛔	⛔	⛔
	Nunes 2019	✓	✓	✓	✓	⛔	⛔	⛔	⛔	✓	⛔	⛔	⛔	✓	⛔

of the models that incorporate advanced DL techniques. Attention mechanisms were applied in all DL studies except that by Yoon et al. [75]. Transformers were not used directly in any of the proposed models; however, Zhao et al. [68], Gao et al. [69], and Liu et al. [70] compared their model's classification performance with well-known transformer architectures. These DL networks are considered complex black boxes. Therefore, in addition to interpretability analysis, more tools are required to test their robustness. All DL studies applied various ablation mechanisms to test the contribution of the features and/or parts to the performance of their models.

We also reviewed how these studies addressed the challenges of applying data-hungry DL models to OCT images on small datasets. To address this issue, many researchers applied techniques such as data augmentation [6], [67], [68], [69], [70], [79] and transfer learning [6], [67], [68], [69], [75]. Another challenge involved OCT image quality and standardization issues, which were mostly addressed by manually removing low-quality images. Studies used various image transformation [69], [73], [79], enhancement [69], and artifact removal techniques [69], [73].

All reviewed DL studies except those by Liu et al. [70] and Wang et al. [79] utilized transfer learning from generic networks pre-trained on ImageNet-like databases. However, OCT scan slice images differ significantly from images in general databases because the pixel values represent the measured depth of the retinal tissue at a resolution of a few micrometers, which may limit the benefits of transfer learning. Additionally, because pre-trained networks are often trained on generic RGB color image datasets, their application requires alignment with the input format. This requirement confines the analysis to a fixed-size single 2D image input, thereby limiting the utilization of the entire 3D retinal volume. New OCT fine-tuned foundational models have become available. However, they imposed the same limitations because they were also fine-tuned on generic 2D ImageNet pre-trained foundation models. RETFound [88], for instance, was trained on a substantial collection of 736K OCT images from various datasets aimed at both segmentation and classification tasks. OCTA-NET [89] and FAZ-NET [90] were developed using the ROSE dataset [89] for microvasculature and FAZ segmentation. Despite these developments, none of the analyzed studies employed OCT/OCTA-specific pre-trained networks.

Interpretability and Explainability: Popular DL explainability methods were designed for standard images. However, retinal images exhibit different characteristics, making it important to identify the most suitable method for explaining OCT/OCTA images. Despite reports that the Deep Taylor method performs better for OCT scans [59], some recent studies [6], [68], [69], [70], [79] utilized the most popular explainability technique called GradCAM.

New interpretability and explainability methodologies have been developed for transformer-based architectures [91]. RETFound [88], a pre-trained transformer model for OCT scans, employs a technique specifically designed for transformer networks that utilizes Deep Taylor mechanisms [59]. However, further studies are required to investigate the most effective interpretability methods for OCT/OCTA images, particularly when trained using transformer-based networks.

Liu et al. [70] and Zhao et al. [68] utilized an advanced interpretability technique known as importance maps. This

approach involved generating additional parameters from images and conducting interpretability analyses to assess the significance of these parameters in the model's classification process. These studies were notable for their focus on identifying novel biomarkers rather than merely classifying data. Liu et al. [70] discovered that, in diagnosing Alzheimer's disease (AD), the choriocapillaris (CC) layer is more crucial than the deep vascular complex (DVC) layer in optical coherence tomography angiography (OCTA), with the parafoveal region being the most critical part of the retina. In contrast, Zhao et al. [68] determined that the DVC is the most significant layer for distinguishing between AD and mild cognitive impairment. They found that five out of eight parameters in the DVC showed significant differences between early-onset AD and controls, whereas only two parameters were significant between mild cognitive impairment and controls. These studies also demonstrated that utilizing more tailored interpretability techniques can enhance the identification of new biomarkers, as these tools can reveal additional informative details.

All eligible studies compared their explainability results with those of medical biomarkers discussed in Section II. However, none of these studies verified their findings with ophthalmologists or through any annotations of test data. Comparing the explainability results with regions annotated by ophthalmologists on the same images could offer more accurate metrics for assessing the performance of the trained models. Specifically, the Intersection over Union (IoU) metric can be used to quantify the overlap between highlighted regions of the model and the expert-annotated regions, providing a robust measure of agreement and accuracy.

Animal Studies and Their Translational Potential: Our analysis of animal (mouse) studies revealed their pivotal role in Alzheimer's disease (AD) research, providing insights that complement human investigations. Mouse models, which share 99% genetic similarity with humans, enable longitudinal inquiries into AD, with triple-transgenic mouse models (3xTg-AD or TMM) replicating key characteristics of human AD. Studies focusing on OCT and Fundus imaging in mice aged between 1 and 16 months mainly focused on the early stages of cognitive impairment and disease progression. Since one human year is roughly equivalent to nine mice days, mouse studies are crucial for longitudinal and age detection research into early AD progression.

Limitations and Advantages Compared to Other Modalities: OCT/OCTA studies are relatively new compared with other neuroimaging research areas. Consequently, the datasets and publicly available data are limited. The scarcity of large standardized datasets hinders the ability to test and compare the robustness of OCT/OCTA-based models across different studies. Despite being more accessible, non-invasive, and cost-effective compared to alternatives such as MRI, PET, CT, and EEG, the challenges in dataset curation limit the number of publications in this field. This scarcity makes it difficult to conduct meta-analyses or validate findings against established benchmarks.

Furthermore, the presence of numerous other eye diseases, such as age-related macular degeneration and glaucoma, complicates the identification of AD-related changes in the eye. Currently, no research or data collection includes multiple eye diseases alongside AD in patients. The lack of comprehensive data makes it difficult to distinguish AD-specific biomarkers from those associated with other ocular conditions. Detailed structural and vascular information regarding the retina found on OCT/OCTA may reflect other neurodegenerative changes in the brain. For example, changes in the retinal nerve fiber layer were also reported in Parkinson's Disease, Amyotrophic lateral Sclerosis, and Huntington's disease [92].

Future Directions: Developing standardized protocols for data collection and image acquisition is essential to ensure consistency and comparability across studies. The American Alzheimer's Association has initiated the collaborative Atlas of Retinal Imaging in Alzheimer's Study (ARIAS) [93], This study defined a ''minimum dataset'' framework for SD-OCT retinal image acquisition and analysis and an ongoing data collection initiative that adheres to the framework [34]. Similarly, Sampson et al. [94] discussed the requirements for standardizing the retinal OCTA imaging.

Increasing the availability of large, high-quality public datasets is critical for supporting the training and validation of deep learning models. Several initiatives have been undertaken to address this issue. For example, the UK Biobank offers a database containing thousands of OCT scans along with a comprehensive set of health records. Wagner et al.'s AlzEye initiative [95] currently represents the most extensive dataset of ophthalmic images but is not publicly available at the time of writing.

It is essential to conduct longitudinal studies to track disease progression and validate the predictive capabilities of deep learning models over time. Although two recent studies by Wisely et al. [67] and Zhao et al. [68] worked with cohorts with mild cognitive impairment, longitudinal studies in humans have yet to be conducted.

The integration of OCT/OCTA imaging with other diagnostic modalities, such as genetic and neuropsychological data, will enable the development of more comprehensive and accurate diagnostic models. Large-scale global datasets, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) [96], which include magnetic resonance imaging (MRI), positron emission tomography (PET), and related data from thousands of subjects are essential for neuroimaging-based AD research. Nevertheless, data availability remains a major obstacle to progress in the OCT/OCTA field, and there is an urgent need to compile more comprehensive datasets.

Advancing the use of 3D OCT/OCTA imaging techniques to capture more detailed structural and vascular changes in the retina could improve the accuracy and robustness of deep learning models for AD diagnosis. All reviewed studies were based on 2D OCT/OCTA slice images, although OCT/OCTA scans are 3D volumes of multiple slices. 3D OCT scans have the potential to explore larger retinal areas

and accelerate the search for novel biomarkers. However, working with 3D volumetric data poses additional challenges, including the requirement for larger training datasets and computational complexities. Advances in computational power and cloud services have enabled the analysis of 3D volumetric retinal images. However, a substantial increase in input dimensionality requires a significant amount of data to facilitate the learning of the representation and downstream tasks of AD detection. Recent research has begun to focus on entire 3D OCT scans [97], [98], [99], [100].

State-of-the-art DL model explainability is limited to highlighting salient or important regions in an input image. They do not provide higher-level observational explanations, such as thinning, thickening, rotation, deformation of specific anatomic features, or interrelations of various regions. Consequently, the current state-of-the-art tools do not support differential comparisons across norms of healthy individuals or patients with AD to fully support research or early diagnosis. Therefore, further studies are required to investigate more effective explainability and interpretability methods for OCT and OCTA images.

Limitations of the Current Survey: The primary limitation of this study is that despite reviewing an extensive collection of 4006 papers, only 16 were eligible for comparison and analysis. This study covers only OCT/OCTA-related ML and DL studies published between 2015 and 2024 (end of February 2024). This survey only considered the studies written in English.

VII. CONCLUSION

This study reviewed the use of OCT and OCTA retinal imaging in deep learning to diagnose Alzheimer's disease and mild cognitive impairment. We examined recent deep learning approaches and OCT/OCTA datasets targeting the diagnosis of these disorders and analyzed 16 relevant ML-/DL-based studies. The relatively small number of studies indicates the early stages of this research field. Our analysis was structured around a Deep Learning Flow framework that facilitated the analysis of key aspects within the studies, including data curation protocols, models, training methodologies, and validation strategies.

Our review revealed significant challenges in the dataset curation process, including the limited size of datasets, inconsistent criteria for including or excluding data, and variations in the standards of scanning devices, leading to discrepancies across studies. Nevertheless, research on AD using both OCT and OCTA modalities is gaining momentum, potentially leading to the discovery of new AD-specific biomarkers, particularly as large-scale public datasets become available.

ACKNOWLEDGMENT

We thank the Küçükarslan Family Fund, particularly Yusuf Kenan Küçükarslan, for donating computing hardware to this project and study.

REFERENCES

[1] WHO. (2022). *World Health Organization: Dementia Fact Sheet*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

[2] A. Grzybowski and P. Barboni, *The Eye As a Window to the Brain OCT and Imaging in Central Nervous System Diseases*. Cham, Switzerland: Springer, 2020.

[3] A. London, I. Benhar, and M. Schwartz, "The retina as a window to the brain—From eye research to CNS disorders," *Nature Rev. Neurol.*, vol. 9, no. 1, pp. 44–53, Jan. 2013.

[4] B. James, "Alzheimer's disease facts and figures," *Alzheimer's Dement.*, vol. 18, no. 4, pp. 700–789, 2022.

[5] R. Vij and S. Arora, *A Systematic Survey of Advances in Retinal Imaging Modalities for Alzheimer's Disease Diagnosis*. Cham, Switzerland: Springer, 2022.

[6] C. E. Wisely, D. Wang, R. Henao, D. S. Grewal, A. C. Thompson, C. B. Robbins, S. P. Yoon, S. Soundararajan, B. W. Polascik, J. R. Burke, A. Liu, L. Carin, and S. Fekrat, "Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging," *Brit. J. Ophthalmology*, vol. 106, no. 3, pp. 388–395, Mar. 2022.

[7] E. Salobrar-García, R. de Hoz, A. I. Ramírez, I. López-Cuenca, P. Rojas, R. Vazirani, C. Amarante, R. Yubero, P. Gil, M. D. Pinazo-Durán, J. J. Salazar, and J. M. Ramírez, "Changes in visual function and retinal structure in the progression of Alzheimer's disease," *PLoS One*, vol. 14, no. 8, Aug. 2019, Art. no. e0220535.

[8] G. Querques, E. Borrelli, R. Sacconi, L. De Vitis, L. Leocani, R. Santangelo, G. Magnani, G. Comi, and F. Bandello, "Functional and morphological changes of the retinal vessels in Alzheimer's disease and mild cognitive impairment," *Sci. Rep.*, vol. 9, no. 1, pp. 1–20, Jan. 2019.

[9] J. Hui, Y. Zhao, S. Yu, J. Liu, K. Chiu, and Y. Wang, "Detection of retinal changes with optical coherence tomography angiography in mild cognitive impairment and Alzheimer's disease patients: A meta-analysis," *PLoS One*, vol. 16, no. 8, Aug. 2021, Art. no. e0255362.

[10] S. Fereshetian, J. S. Agranat, N. Siegel, S. Ness, T. D. Stein, and M. L. Subramanian, "Protein and imaging biomarkers in the eye for early detection of Alzheimer's disease," *J. Alzheimer's Disease Rep.*, vol. 5, no. 1, pp. 375–387, May 2021.

[11] D. Romaus-Sanjurjo, U. Regueiro, M. López-López, L. Vázquez-Vázquez, A. Ouro, I. Lema, and T. Sobrino, "Alzheimer's disease seen through the eye: Ocular alterations and neurodegeneration," *Int. J. Mol. Sci.*, vol. 23, no. 5, p. 2486, Feb. 2022.

[12] T.-C. Yeh, C.-T. Kuo, and Y.-B. Chou, "Retinal microvascular changes in mild cognitive impairment and Alzheimer's disease: A systematic review, meta-analysis, and meta-regression," *Frontiers Aging Neurosci.*, vol. 14, pp. 1–23, Apr. 2022.

[13] H. M. Kim, J. W. Han, Y. J. Park, J. B. Bae, S. J. Woo, and K. W. Kim, "Association between retinal layer thickness and cognitive decline in older adults," *JAMA Ophthalmology*, vol. 140, no. 7, p. 683, Jul. 2022.

[14] R. T. Yanagihara, C. S. Lee, D. S. W. Ting, and A. Y. Lee, "Methodological challenges of deep learning in optical coherence tomography for retinal diseases: A review," *Translational Vis. Sci. Technol.*, vol. 9, no. 2, p. 11, Feb. 2020.

[15] Y. Tong, W. Lu, Y. Yu, and Y. Shen, "Application of machine learning in ophthalmic imaging modalities," *Eye Vis.*, vol. 7, no. 1, pp. 1–32, Dec. 2020.

[16] A. Ran and C. Y. Cheung, "Deep learning-based optical coherence tomography and optical coherence tomography angiography image analysis: An updated summary," *Asia–Pacific J. Ophthalmology*, vol. 10, no. 3, pp. 253–260, May 2021.

[17] R. S. Latha, G. R. Sreekanth, B. Bizu, K. Suvalakshmi, and R. E. Selvaraj, "Analysis of ophthalmic disorders for retinal images using deep learning: A review," *Natural Volatiles Essential Oils*, vol. 8, no. 5, pp. 656–673, 2021.

[18] L. Tian, B. Hunt, M. A. Bell, J. Yi, J. T. Smith, M. Ochoa, X. Intes, and N. J. Durr, "Deep learning in biomedical optics," *Lasers Surgery Med.*, vol. 53, no. 6, pp. 748–775, Aug. 2021.

[19] H. Bourkhime, N. Qarmiche, M. Omari, N. Bahra, N. Tachfouti, S. E. Fakir, and N. Otmani, "Machine learning and novel ophthalmologic biomarkers for Alzheimer's disease screening: Systematic review," in *Proc. ITM Web Conf.*, vol. 43, 2022, p. 01009.

[20] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Syst. Rev.*, vol. 10, no. 1, pp. 1–11, Dec. 2021.

[21] T. Dening and M. B. Sandilyan, "Dementia: Definitions and types," *Nursing Standard*, vol. 29, no. 37, pp. 37–42, May 2015.

[22] WHO. (2014). *International Statistical Classification of Diseases and Related Health Problems 10th Revision*. [Online]. Available: https://icd.who.int/browse10/2014/en#/F00F09

[23] G. H. Havoutis, "Analysis of diagnostic, preventive, and disease-modifying therapeutic measures of alzheimer's disease," Capstone, Nova Southeastern Univ., 2017. [Online]. Available: https://nsuworks.nova.edu/cnso_stucap/334

[24] Q. Mu, J. Xie, Z. Wen, Y. Weng, and Z. Shuyun, "A quantitative MR study of the hippocampal formation, the amygdala, and the temporal horn of the lateral ventricle in healthy subjects 40 to 90 years of age," *Amer. J. Neuroradiology*, vol. 20, pp. 207–211, Aug. 1999.

[25] K. Juottonen, M. P. Laakso, K. Partanen, and H. Soininen, "Comparative MR analysis of the entorhinal cortex and hippocampus in diagnosing Alzheimer disease," *Amer. J. Neuroradiology*, vol. 20, no. 1, pp. 139–144, 1999.

[26] M. Blazes and C. S. Lee, "Understanding the brain through aging eyes," *Adv. Geriatric Med. Res.*, vol. 3, no. 2, pp. 1–24, 2021.

[27] A. Song, N. Johnson, A. Ayala, and A. C. Thompson, "Optical coherence tomography in patients with Alzheimer's disease: What can it tell us?" *Eye Brain*, vol. 13, pp. 1–20, Jan. 2021.

[28] L. P. Cunha, L. A. Pires, M. M. Cruzeiro, A. L. M. Almeida, L. C. Martins, P. N. Martins, N. Shigaeff, and T. C. Vale, "Optical coherence tomography in neurodegenerative disorders," *Arquivos De Neuro-Psiquiatria*, vol. 80, no. 2, pp. 180–191, 2022.

[29] X. Hadoux et al., "Non-invasive in vivo hyperspectral imaging of the retina for potential biomarker use in Alzheimer's disease," *Nature Commun.*, vol. 10, no. 1, pp. 1–19, Sep. 2019.

[30] V. Carelli, C. La Morgia, F. N. Ross-Cisneros, and A. A. Sadun, "Optic neuropathies: The tip of the neurodegeneration Iceberg," *Hum. Mol. Genet.*, vol. 26, no. 2, pp. R139–R150, Oct. 2017.

[31] M. R. Gardner, V. Baruah, G. Vargas, M. Motamedi, T. E. Milner, and H. G. Rylander, "Scattering angle resolved optical coherence tomography detects early changes in 3xTg Alzheimer's disease mouse model," *Transl. Vis. Sci. Technol.*, vol. 9, no. 5, p. 18, Apr. 2020.

[32] J. den Haan, T. H. J. Morrema, F. D. Verbraak, J. F. de Boer, P. Scheltens, A. J. Rozemuller, A. A. B. Bergen, F. H. Bouwman, and J. J. Hoozemans, "Amyloid-beta and phosphorylated tau in post-mortem Alzheimer's disease retinas," *Acta Neuropathologica Commun.*, vol. 6, no. 1, p. 147, Dec. 2018.

[33] S. Sadda, Y. Attiku, Y. He, and M. Nittala, "Current status and future possibilities of retinal imaging in diabetic retinopathy care applicable to low- and medium-income countries," *Indian J. Ophthalmology*, vol. 69, no. 11, p. 2968, 2021.

[34] J. Alber, E. Arthur, S. Sinoff, D. C. DeBuc, E. Y. Chew, L. Douquette, W. V. Hatch, C. Hudson, A. Kashani, C. S. Lee, S. Montaquila, S. Mozdbar, L. P. Cunha, F. Tayyari, G. Van Stavern, and P. J. Snyder, "A recommended-minimum data set," *Alzheimer's Dementia, Diagnosis, Assessment Disease Monitor.*, vol. 12, no. 1, pp. 1–29, Jan. 2020, doi: 10.1002/dad2.12119.

[35] S. H. Tsang and T. Sharma, *Electroretinography*. Cham, Switzerland: Springer, 2018.

[36] P. J. Snyder, J. Alber, C. Alt, L. J. Bain, B. E. Bouma, F. H. Bouwman, D. Cabrera Debuc, M. C. W. Campbell, M. C. Carrillo, E. Y. Chew, M. F. Cordeiro, M. R. Dueñas, B. M. Fernández, M. Koronyo-Hamaoui, and H. M. Snyder, "Retinal imaging in Alzheimer's and neurodegenerative diseases," *Alzheimer's Dementia*, vol. 1, no. 1, pp. 1–28, 2020.

[37] Y. J. Ge, W. Xu, Y. N. Ou, Y. Qu, Y. H. Ma, Y. Y. Huang, X. N. Shen, S. D. Chen, L. Tan, Q. H. Zhao, and J. T. Yu, "Retinal biomarkers in Alzheimer's disease and mild cognitive impairment: A systematic review and meta-analysis," *Ageing Res. Rev.*, vol. 69, no. 1, pp. 1–29, Aug. 2021.

[38] Y. Ibrahim, J. Xie, A. Macerollo, R. Sardone, Y. Shen, V. Romano, and Y. Zheng, "A systematic review on retinal biomarkers to diagnose dementia from OCT/OCTA images," *J. Alzheimer's Disease Rep.*, vol. 7, no. 1, pp. 1201–1235, Nov. 2023.

[39] E. Costanzo, I. Lengyel, M. Parravano, I. Biagini, M. Veldsman, A. Badhwar, M. Betts, A. Cherubini, D. J. Llewellyn, I. Lourida, T. Macgillivray, T. Rittman, S. Tamburin, X. Y. Tai, and G. Virgili, "Ocular biomarkers for Alzheimer disease dementia: An umbrella review of systematic reviews and meta-analyses," *JAMA*, vol. 141, no. 1, pp. 84–91, 2023.

[40] C. Moran, Z. Y. Xu, H. Mehta, M. Gillies, C. Karayiannis, R. Beare, C. Chen, and V. Srikanth, "Neuroimaging and cognitive correlates of retinal optical coherence tomography (OCT) measures at late middle age in a twin sample," *Sci. Rep.*, vol. 12, no. 1, pp. 1–9, Jun. 2022, doi: 10.1038/s41598-022-13662-8.

[41] A. Yuan and C. S. Lee, "Retinal biomarkers for Alzheimer disease: The facts and the future," *Asia–Pacific J. Ophthalmology*, vol. 11, no. 2, pp. 140–148, Mar. 2022.

[42] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience, Exploring the Brain*. Chennai, India: Wolters Kluwer, 2016.

[43] A. J. Augustin and J. Atorf, "The value of optical coherence tomography angiography (OCT–A) in neurological diseases," *Diagnostics*, vol. 12, no. 2, p. 468, Feb. 2022.

[44] G. Triolo, A. Rabiolo, N. D. Shemonski, A. Fard, F. D. Matteo, R. Sacconi, P. Bettin, S. Magazzeni, G. Querques, L. E. Vazquez, P. Barboni, and F. Bandello, "Optical coherence tomography angiography macular and peripapillary vessel perfusion density in healthy subjects, glaucoma suspects, and glaucoma patients," *Investigative Opthalmology Vis. Sci.*, vol. 58, no. 13, p. 5713, Nov. 2017.

[45] S. Dutta and P. Sengupta, "Men and mice: Relating their ages," *Life Sci.*, vol. 152, pp. 244–248, May 2016, doi: 10.1016/j.lfs.2015.10.025.

[46] L. Sánchez-Puebla, I. López-Cuenca, E. Salobrar-García, A. I. Ramírez, J. A. Fernández-Albarral, J. A. Matamoros, L. Elvira-Hurtado, J. J. Salazar, J. M. Ramírez, and R. de Hoz, "OCT imaging in murine models of Alzheimer's disease in a systematic review: Findings, methodology and future perspectives," *Biomedicines*, vol. 12, no. 3, p. 528, Feb. 2024.

[47] E. Akleman, "Deep learning," *Computer*, vol. 53, no. 9, pp. 1–17, Sep. 2020.

[48] L. Woodrow, *Presenting Descriptive Statistics*. London, U.K.: Palgrave Macmillan, 2014, pp. 49–60.

[49] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. 3rd Int. Conf. Learn. Represent., (ICLR)*, 2015, pp. 1–15.

[50] A. Vaswani, "Attention is all you need," in *Proc. 31st Conf. Neural Inf. Process. Syst.*, vol. 30, Dec. 2017, pp. 5999–6009.

[51] N. J. Parmar, A. Vaswani, J. Uszkoreit, L. Kaiser, N. Shazeer, A. Ku, and D. Tran, "Image transformer," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2018, pp. 1–11.

[52] A. Kolesnikov, A. Dosovitskiy, D. Weissenborn, G. Heigold, J. Uszkoreit, L. Beyer, M. Minderer, M. Dehghani, N. Houlsby, S. Gelly, T. Unterthiner, and X. Zhai, "An image is worth 16×16 words: Transformers for image recognition at scale," 2021, *arXiv:2010.11929*.

[53] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[54] F. Maleki, N. Muthukrishnan, K. Ovens, C. Reinhold, and R. Forghani, "Machine learning algorithm validation: From essentials to advanced applications and implications for regulatory certification and deployment," *Neuroimaging Clinics*, vol. 30, no. 4, pp. 433–445, 2020.

[55] F. Maleki, K. Ovens, R. Gupta, C. Reinhold, A. Spatz, and R. Forghani, "Generalizability of machine learning models: Quantitative evaluation of three methodological pitfalls," *Radiology, Artif. Intell.*, vol. 5, no. 1, pp. 1–38, Jan. 2023.

[56] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.

[57] D. M. Powers, "Evaluation: From precision, recall and F-factor to R," *J. Mach. Learn. Technol.*, vol. 1, pp. 37–64, Nov. 2011.

[58] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci. Rep.*, vol. 14, no. 1, pp. 1–12, Mar. 2024.

[59] A. Singh, J. Jothi Balaji, M. A. Rasheed, V. Jayakumar, R. Raman, and V. Lakshminarayanan, "Evaluation of explainable deep learning methods for ophthalmic diagnosis," *Clin. Ophthalmology*, vol. 15, pp. 2573–2581, Jun. 2021.

[60] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020.

[61] W. Yang, Y. Wei, H. Wei, Y. Chen, G. Huang, X. Li, R. Li, N. Yao, X. Wang, X. Gu, M. B. Amin, and B. Kang, "Survey on explainable AI: From approaches, limitations and applications aspects," *Hum.-Centric Intell. Syst.*, vol. 3, no. 3, pp. 161–188, Aug. 2023.

[62] E. Cambria, L. Malandri, F. Mercorio, M. Mezzanzanica, and N. Nobani, "A survey on XAI and natural language explanations," *Inf. Process. Manage.*, vol. 60, no. 1, Jan. 2023, Art. no. 103111.
[63] S. Sheikholeslami, "Ablation programming for machine learning," M.S. thesis, School Elect. Eng. Comput. Sci., KTH, 2019
[64] R. Meyes, M. Lu, C. Waubert de Puiseau, and T. Meisen, "Ablation studies in artificial neural networks," 2019, *arXiv:1901.08644*.
[65] P. F. Whiting, "QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies," *Ann. Internal Med.*, vol. 155, no. 8, p. 529, Oct. 2011, doi: 10.7326/0003-4819-155-8-201110180-00009.
[66] A. Venazzi, W. Swardfager, B. Lam, J. D. O. Siqueira, N. Herrmann, and H. Cogo-Moreira, "Validity of the QUADAS-2 in assessing risk of bias in Alzheimer's disease diagnostic accuracy studies," *Frontiers Psychiatry*, vol. 9, pp. 1–28, May 2018.
[67] C. E. Wisely, A. Richardson, R. Henao, C. B. Robbins, J. P. Ma, D. Wang, K. G. Johnson, A. J. Liu, D. S. Grewal, and S. Fekrat, "A convolutional neural network using multimodal retinal imaging for differentiation of mild cognitive impairment from normal cognition," *Ophthalmology Sci.*, vol. 4, no. 1, Jan. 2024, Art. no. 100355.
[68] Y. Zhao, J. Hao, W. Kwapong, Y. Xu, T. Shen, H. Fu, Q. Lu, S. Liu, J. Zhang, Y. Liu, Y. Zhao, Y. Zheng, A. Frangi, and J. Liu, "Eye-AD: A graph-based model for early-onset alzheimer's disease and mild cognitive impairment detection based on retinal octa images," ResearchSquare, 2023, doi: 10.21203/rs.3.rs-3680789/v1.
[69] H. Gao, S. Zhao, G. Zheng, X. Wang, R. Zhao, Z. Pan, H. Li, F. Lu, and M. Shen, "Using a dual-stream attention neural network to characterize mild cognitive impairment based on retinal images," *Comput. Biol. Med.*, vol. 166, Nov. 2023, Art. no. 107411.
[70] S. Liu, J. Hao, Y. Xu, H. Fu, X. Guo, J. Liu, Y. Zheng, Y. Liu, J. Zhang, and Y. Zhao, "Polar-net: A clinical-friendly model for Alzheimer's disease detection in OCTA images," 2023, *arXiv:2311.06009*.
[71] X. Wang, B. Jiao, H. Liu, Y. Wang, X. Hao, Y. Zhu, B. Xu, H. Xu, S. Zhang, X. Jia, Q. Xu, X. Liao, Y. Zhou, H. Jiang, J. Wang, J. Guo, X. Yan, B. Tang, R. Zhao, and L. Shen, "Machine learning based on optical coherence tomography images as a diagnostic tool for Alzheimer ' s disease," *CNS Neurosci. Therapeutics*, vol. 1, pp. 1–12, May 2022.
[72] S. Lemmens, T. Van Craenendonck, J. Van Eijgen, L. De Groef, R. Bruffaerts, D. A. de Jesus, W. Charle, M. Jayapala, G. Sunaric-Mégevand, A. Standaert, J. Theunis, K. Van Keer, M. Vandenbulcke, L. Moons, R. Vandenberghe, P. De Boever, and I. Stalmans, "Combination of snapshot hyperspectral retinal imaging and optical coherence tomography to identify Alzheimer's disease patients," *Alzheimer's Res. Therapy*, vol. 12, no. 1, pp. 1–28, Dec. 2020.
[73] C. S. Sandeep, A. S. Kumar, K. Mahadevan, and P. Manoj, "Analysis of retinal OCT images for the early diagnosis of Alzheimer's disease," *Adv. Intell. Syst. Comput.*, vol. 749, pp. 509–520, Oct. 2019.
[74] J. Xie, Q. Yi, Y. Wu, Y. Zheng, Y. Liu, A. Macerollo, H. Fu, Y. Xu, J. Zhang, A. Behera, C. Fan, A. F. Frangi, J. Liu, Q. Lu, H. Qi, and Y. Zhao, "Deep segmentation of OCTA for evaluation and association of changes of retinal microvasculature with Alzheimer's disease and mild cognitive impairment," *Brit. J. Ophthalmology*, vol. 108, no. 3, pp. 432–439, Mar. 2024.
[75] J. M. Yoon, C. Y. Lim, H. Noh, S. W. Nam, S. Y. Jun, M. J. Kim, M. Y. Song, H. Jang, H. J. Kim, S. W. Seo, D. L. Na, M. J. Chung, D.-I. Ham, and K. Kim, "Enhancing foveal avascular zone analysis for Alzheimer's diagnosis with AI segmentation and machine learning using multiple radiomic features," *Sci. Rep.*, vol. 14, no. 1, pp. 1–29, Jan. 2024.
[76] A. Nunes, G. Silva, C. Duque, C. Januário, I. Santana, A. F. Ambrósio, M. Castelo-Branco, and R. Bernardes, "Retinal texture biomarkers may help to discriminate between Alzheimer's, Parkinson's, and healthy controls," *PLoS One*, vol. 14, no. 6, Jun. 2019, Art. no. e0218826.
[77] R. Bernardes, G. Silva, S. Chiquita, P. Serranho, and A. F. Ambrósio, "Retinal biomarkers of Alzheimer's disease: Insights from transgenic mouse models," in *Artificial Intelligence*, vol. 10317. Cham, Switzerland: Springer, 2017, pp. 541–550.
[78] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man, Cybern.*, vols. SMC-3, no. 6, pp. 610–621, Nov. 1973.
[79] X. Wang, H. Li, Z. Xiao, H. Fu, Y. Zhao, R. Jin, S. Zhang, W. R. Kwapong, Z. Zhang, H. Miao, and J. Liu, *Screening Dementia OCTA Images via Multi-Projection Consistency Complementarity*. Cham, Switzerland: Springer, 2022.
[80] M. Li, K. Huang, Q. Xu, J. Yang, Y. Zhang, Z. Ji, K. Xie, S. Yuan, Q. Liu, and Q. Chen, "OCTA-500: A retinal dataset for optical coherence tomography angiography study," *Med. Image Anal.*, vol. 93, Apr. 2024, Art. no. 103092.
[81] P. Guimarães, P. Serranho, J. Martins, P. I. Moreira, A. F. Ambrósio, M. Castelo-Branco, and R. Bernardes, "Retinal aging in 3× Tg-AD mice model of Alzheimer's disease," *Frontiers Aging Neurosci.*, vol. 14, pp. 1–6, Jun. 2022.
[82] H. Ferreira, P. Serranho, P. Guimarães, R. Trindade, J. Martins, P. I. Moreira, A. F. Ambrósio, M. Castelo-Branco, and R. Bernardes, "Stage-independent biomarkers for Alzheimer's disease from the living retina: An animal study," *Sci. Rep.*, vol. 12, no. 1, pp. 1–7, Aug. 2022, doi: 10.1038/s41598-022-18113-y.
[83] F. Sayeed, K. Rafeeq Ahmed, M. S. Vinmathi, A. I. Priyadarsini, C. B. Gundupalli, V. Tripathi, W. Shishah, and V. P. Sundramurthy, "Classification of transgenic mice by retinal imaging using SVMS," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–8, Jun. 2022.
[84] A. Batista, P. Guimarães, P. Serranho, A. Nunes, J. Martins, P. I. Moreira, A. F. Ambrósio, M. Morgado, M. Castelo-Branco, and R. Bernardes, "Retinal imaging in animal models: Searching for biomarkers of neurodegeneration," *Frontiers Ophthalmology*, vol. 3, pp. 1–39, Apr. 2023.
[85] A. Richardson, C. B. Robins, C. E. Wisely, R. Henao, and D. S. Grewal, "Artificial intelligence in dementia," *Current Opinion Ophthalmology*, vol. 33, no. 9, pp. 1–27, 2022.
[86] S. Nahid, M. Aaisha, M. Farid, E. K. C. Myra, M. Aaisha, C. Myra, V. Antoine, A. E. K. Anas, B. Mohamed-Rafik, and H. Ravindra, "Exploring imaging biomarkers for early detection of Alzheimer's disease using deep learning: A comprehensive analysis," in *Recent Trends in Image Processing and Pattern Recognition*. Cham, Switzerland: Springer, 2024, pp. 197–206.
[87] R. Bernardes, H. Ferreira, P. Guimarães, and P. Serranho, "Shedding light on early central nervous system changes for Alzheimer's disease through the retina: An animal study," in *Proc. 2nd Int. Conf. Image Process. Vis. Eng.*, 2022, pp. 247–258.
[88] Y. Zhou, M. A. Chia, S. K. Wagner, M. S. Ayhan, D. J. Williamson, R. R. Struyven, T. L. M. Xu, M. G. Lozano, P. Woodward-Court, Y. Kihara, A. Altmann, A. Y. Lee, E. J. Topol, A. K. Denniston, D. C. Alexander, and P. A. Keane, "A foundation model for generalizable disease detection from retinal images," *Nature*, vol. 622, no. 7981, pp. 156–163, 2023.
[89] Y. Ma, H. Hao, J. Xie, H. Fu, J. Zhang, J. Yang, Z. Wang, J. Liu, Y. Zheng, and Y. Zhao, "ROSE: A retinal OCT-angiography vessel segmentation dataset and new model," *IEEE Trans. Med. Imag.*, vol. 40, no. 3, pp. 928–939, Mar. 2021.
[90] J. Hao, T. Shen, X. Zhu, Y. Liu, A. Behera, D. Zhang, B. Chen, J. Liu, J. Zhang, and Y. Zhao, "Retinal structure detection in OCTA image via voting-based multitask learning," *IEEE Trans. Med. Imag.*, vol. 41, no. 12, pp. 3969–3980, Dec. 2022.
[91] H. Chefer, S. Gur, and L. Wolf, "Transformer interpretability beyond attention visualization," 2020, *arXiv:2012.09838*.
[92] A. Suh, J. Ong, S. A. Kamran, E. Waisberg, P. Paladugu, N. Zaman, P. Sarker, A. Tavakkoli, and A. G. Lee, "Retina oculomics in neurodegenerative disease," *Ann. Biomed. Eng.*, vol. 51, no. 12, pp. 2708–2721, Dec. 2023.
[93] J. Alber, E. Arthur, L. I. Thompson, D. Goldfarb, B. M. Fernandez, S. P. Salloway, S. Sinoff, and P. J. Snyder, "The atlas of retinal imaging in Alzheimer's study (ARIAS): Study design and objectives," *Alzheimer's Dementia*, vol. 16, no. 5, p. 20, Dec. 2020.
[94] D. M. Sampson, A. M. Dubis, F. K. Chen, R. J. Zawadzki, and D. D. Sampson, "Towards standardizing retinal optical coherence tomography angiography: A review," *Light, Sci. Appl.*, vol. 11, no. 1, pp. 1–45, Mar. 2022.
[95] S. K. Wagner, F. Hughes, M. Cortina-Borja, N. Pontikos, R. Struyven, X. Liu, H. Montgomery, D. C. Alexander, E. Topol, S. E. Petersen, K. Balaskas, J. Hindley, A. Petzold, J. S. Rahi, A. K. Denniston, and P. A. Keane, "AlzEye: Longitudinal record-level linkage of ophthalmic imaging and hospital admissions of 353 157 patients in London, U.K.," *BMJ Open*, vol. 12, no. 3, Mar. 2022, Art. no. e058552.
[96] ADNI. (2022). *Alzheimer's Disease Neuroimaging Initiative (ADNI)*. Accessed: Aug. 13, 2022. [Online]. Available: http://adni.loni.usc.edu/
[97] S.-J. Park, T. Ko, C.-K. Park, Y.-C. Kim, and I.-Y. Choi, "Deep learning model based on 3D optical coherence tomography images for the automated detection of pathologic myopia," *Diagnostics*, vol. 12, no. 3, p. 742, Mar. 2022.

[98] H. Danesh, K. Maghooli, A. Dehghani, and R. Kafieh, "Synthetic OCT data in challenging conditions: Three-dimensional OCT and presence of abnormalities," *Med. Biol. Eng. Comput.*, vol. 60, no. 1, pp. 189–203, Jan. 2022, doi: 10.1007/s11517-021-02469-w.
[99] J. J. Xu, Y. Zhou, Q. J. Wei, K. Li, Z. P. Li, T. Yu, J. C. Zhao, D. Y. Ding, X. R. Li, G. Z. Wang, and H. Dai, "Three-dimensional diabetic macular edema thickness maps based on fluid segmentation and fovea detection using deep learning," *Int. J. Ophthalmology*, vol. 15, no. 3, pp. 495–501, 2022.
[100] M. Elsharkawy, A. Sharafeldeen, A. Soliman, F. Khalifa, M. Ghazal, E. El-Daydamony, A. Atwan, H. S. Sandhu, and A. El-Baz, "A novel computer-aided diagnostic system for early detection of diabetic retinopathy using 3D-OCT higher-order spatial appearance model," *Diagnostics*, vol. 12, no. 2, p. 461, Feb. 2022.

YASEMIN TURKAN (Member, IEEE) received the B.S. and M.S. degrees in computer engineering and information sciences from Bilkent University. She is currently pursuing the Ph.D. degree with the Department of Computer Engineering, Işik University, with a focus on deep learning-based analysis of retinal optical coherence tomography (OCT) scans for detecting Alzheimer's disease.

F. BORAY TEK received the Ph.D. degree from the School of Electronics and Computer Science, University of Westminster, in February 2007, with a focus on the microscopy diagnosis of malaria. He is currently with the Department of Artificial Intelligence and Data Engineering, Istanbul Technical University. His research interests include computer vision, (bio)image analysis, and machine learning.

FATİH ARPACI received the degree from the Cerrahpaşa Faculty of Medicine, Istanbul University, in 2021. He is currently a Medical Doctor and Ophthalmology Resident at Antalya Training and Research Hospital.

OZAN ARSLAN received the degree from the Faculty of Medicine, Pamukkale University. He is currently a specialized Researcher and a Physician Assistant with the Department of Neurology, Antalya Training and Research Hospital.

DEVRİM TOSLAK is currently an Ophthalmologist, has a strong interest in optical design and the prototype construction of retinal imaging devices. Previously, he was a Research Specialist with the Biomedical Optics and Functional Imaging Laboratory, Lions of Illinois Eye Research Institute, University of Illinois at Chicago, from 2015 to 2020. He is also with the Department of Ophthalmology, Antalya Training and Research Hospital.

MEHMET BULUT received the degree from Istanbul Faculty of Medicine, Istanbul University, in 1996. He is currently a Medical Doctor and Ophthalmologist. He is also with Antalya Anatolia Hospital and Belek University. He has articles published in international journals on OCTA and OCT findings, especially in Alzheimer's patients.

AYLİN YAMAN received the degree from the Faculty of Medicine, Hacettepe University. She specialized in neurology with the Faculty of Medicine, Ankara University. She is currently an Associate Professor of neurology with Antalya Training and Research Hospital, her specific fields of interest and practice are dementia and neuromuscular diseases.