



REVIEW

Deep Learning Applications to Classification and Detection of Age-Related Macular Degeneration on Optical Coherence Tomography Imaging: A Review

Neslihan Dilruba Koseoglu · Andrzej Grzybowski · T. Y. Alvin Liu

Received: March 29, 2023 / Accepted: July 14, 2023 / Published online: July 26, 2023
© The Author(s) 2023

ABSTRACT

Age-related macular degeneration (AMD) is one of the leading causes of blindness in the elderly, more commonly in developed countries. Optical coherence tomography (OCT) is a non-invasive imaging device widely used for the diagnosis and management of AMD. Deep learning (DL) uses multilayered artificial neural networks (NN) for feature extraction, and is the cutting-edge technique for medical image analysis for diagnostic and prognostication purposes. Application of DL models to OCT image analysis has garnered significant interest in recent years. In this review, we aimed to summarize studies focusing on DL models used in classification and detection of AMD. Additionally, we provide a brief introduction to other DL applications in AMD, such as segmentation, prediction/prognostication, and models trained on multimodal imaging.

Keywords: Age-related macular degeneration; Optical coherence tomography; Deep learning; Classification; Prediction

Key Summary Points

Deep learning (DL) is the cutting-edge machine learning for medical image analysis. In recent years, DL applications to age-related macular degeneration (AMD) have shifted from color fundus photographs to optical coherence tomography (OCT) imaging.

DL techniques for AMD OCT analysis can be applied for different purposes, including classification, segmentation, and prediction.

Classification DL models utilized for the detection of AMD, differentiation of AMD from other macular pathologies, and identification of different stages of AMD typically have robust performance.

N. D. Koseoglu · T. Y. A. Liu (✉)
Wilmer Eye Institute, Johns Hopkins University, 600
N. Wolfe St., Maumenee 726, Baltimore, MD 21287,
USA
e-mail: tliu25@jhmi.edu

A. Grzybowski
Institute for Research in Ophthalmology,
Foundation for Ophthalmology Development,
Poznan, Poland

INTRODUCTION

Age-related macular degeneration (AMD) is the most common cause of central vision loss in developed countries [1]. Of people 40 years and

older in the USA, it was estimated that 18.34 million (11.64%) were living with early-stage AMD, while 1.49 million (0.94%) were living with late-stage AMD in 2019. By 2040, AMD is expected to affect 288 million people worldwide, with 10% of this population with intermediate or late-stage AMD [1, 2]. Optical coherence tomography (OCT) is the most commonly utilized and arguably the most important imaging modality for the management of AMD. Deep learning (DL), a subtype of machine learning (ML), has garnered a lot of attention in recent years as the cutting-edge ML technique for medical image analysis. The most common architecture used in DL models is deep convolutional neural network (DCNN), which is modeled after how the human brain processes visual information and typically contains many layers that perform different functions such as convolution and pooling. A newer architecture, called vision transformers (ViTs), has also emerged recently. In contrast to DCNN, ViTs decouple feature aggregation and transformation, and are composed of a series of transformer blocks that contain a feed-forward layer and self-attention layer [3]. To date, DL has been widely applied to ophthalmic image analysis, including images from eyes with AMD. In this article, we plan to perform an in-depth review of published studies on DL applications to OCT and AMD. Specifically, we focus on applications that detect or classify AMD.

METHODS

We performed a systematic search of the PubMed database, including studies published between January 2017 and October 2022. The search criteria were “deep learning” AND “age-related macular degeneration” AND “optical coherence tomography.” After a list of relevant publications was generated, we included studies that focused on classification or detection of AMD on OCT images. Studies that did not involve DL or OCT imaging were excluded.

This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Detection of AMD

In this section, we summarize studies that used DL to detect AMD on OCT images or to differentiate between OCT images with and without AMD pathologies.

He et al. developed a two-stage DL model for the detection of AMD (neovascular and non-neovascular) in OCT volume scans. In the first stage, a ResNet-50 deep convolutional neural network (CNN) was used for classification. In the second stage, the image feature vector set from healthy controls and test image feature vector set were used as inputs for the local outlier factor (LOF) algorithm. The model was initially trained on the USCD dataset [4], consisting of 250 neovascular AMD, 250 non-neovascular AMD (nnAMD), and 250 healthy control OCT volumes. When tested on the external Duke dataset (723 AMD and 1407 healthy control volumes) [5], the model achieved a performance of area under the receiver operating characteristic curve (AUC) of 0.99, sensitivity of 95.0%, and specificity of 95.0% [6].

Lee et al. used a modified version of visual geometry group 16 (VGG-16) CNN model to distinguish between AMD OCT images and healthy controls. The model was trained on 80,839 OCT volume scans (41,074 AMD and 39,765 controls) and validated on 20,163 images (11,616 AMD and 8547 controls). The proposed model achieved an AUC of 92.7% with an accuracy of 87.6% on an image level [7].

Shi et al. developed an interpretable DL model (Med-XAI-Net) to specifically detect the presence or absence of geographic atrophy (GA) in OCT images and to interpret the localization of GA in B-scans. The model was trained and tested on 1284 OCT volume scans (321 with GA and 963 without GA). Med-XAI-Net achieved a performance of 93.5% AUC, 82.8% sensitivity, and 94.6% specificity [8].

In a recent study, Srivastava et al. worked on a DL model that utilized the choroid section in OCT scans to improve the diagnostic accuracy for detecting AMD. A ResNet50 model was trained to classify AMD vs. normal images on a publicly available dataset of 384 subjects (total of 14,560 images out of which 7890 with signs

of AMD and 6670 normal). Their proposed method achieved an AUC of 96.7% for the detection of AMD [9].

Treder et al. developed a DL framework that was trained with 1012 cross-sectional OCT images (701 AMD and 311 healthy) and later tested in 100 images (50 AMD and 50 healthy). The algorithm achieved 92% specificity, 100% sensitivity, and 96% accuracy for detecting AMD [10].

Differentiation of AMD from Other Macular Pathologies

This section summarizes studies that aimed to differentiate AMD from other macular pathologies on OCT.

Baharlouei et al. used a wavelet scattering network to identify normal retina vs. four macular pathologies, namely central serous retinopathy (CSR), macular hole (MH), diabetic retinopathy (DR), and AMD. The open-access OCTID dataset [11] containing 572 OCT images was used for training and testing. The extracted features were fed to a principal component analysis (PCA) classifier. The accuracy for detecting normal images was 97.4% and 82.5% for the classification of the four pathologies [12].

Sun et al. extracted the global features of an OCT volume for classification of macular diseases. B-scan features were extracted and stacked together to generate a two-dimensional (2-D) feature map to be used as input. A pre-trained ResNet50 CNN was used as the backbone. The Duke [5] and Tsinghua datasets [448 AMD and 462 diabetic macular edema (DME)] were used in this study, and OCT volumes were labeled as normal, AMD, and DME. Fivefold cross-validation showed an average accuracy, sensitivity, and specificity of 98.1%, 99.2%, and 95.6%, respectively, for classification [13].

Vellakani and Pushbam proposed a DL model to differentiate between normal, AMD, and DME. The Kaggle dataset [14] was used (1750 training images, 250 validation images, and 8 testing images for each diagnosis). The model involved CNN and long short-term memory (LSTM) networks. The model using

DenseNet201 and LSTM performed best with an overall accuracy of 96.9%, positive predictive value of 97.2%, and true-positive rate of 96.9% using OCT images enhanced by generative adversarial network (GAN) [15].

In a recent study, Alqudah presented a novel CNN (AOCT-Net) for a multiclass classification system based on OCT. Publicly available datasets from USCD [4] (108,312 OCT images) and Duke [5] (26,900 OCT images) were utilized in this study. Images were labeled as normal, AMD, choroidal neovascularization (CNV), DME, or drusen. The model correctly identified 100% of cases with AMD, 98.8% of cases with CNV, 99.1% cases with DME, 98.9% cases with drusen, and 99.1% cases of normal with an overall accuracy of 95.3% [16].

Finally, Li et al. used transfer learning and the VGG-16 network to classify AMD and DME in OCT images. A total of 109,312 retinal OCT images from five different institutions were included study (37,456 with CNV, 11,599 with DME, 8867 with drusen, and 51,390 normal). The algorithm reached a high accuracy of 98.6%, with a sensitivity of 97.8%, a specificity of 99.4%, and AUC of 1 [17].

Fine-Grained Classification Within AMD

This section focuses on studies that differentiated between different stages of AMD or between different features of AMD. For example, polypoidal choroidal vasculopathy (PCV) is a choroidal vascular disease that was initially described by Yanuzzi et al. as a part of the spectrum of manifestations of AMD [18]. However, PCV and typical nAMD present differently on retinal imaging and may show different responses to different treatment modalities. Therefore, it is important clinically to distinguish between the two entities.

Wongchaisuwat et al. developed a ResNet-based DL model to distinguish PCV from nAMD, using fundus fluorescein angiography (FFA) and indocyanine green angiography (ICGA) as gold standard. The model was trained on 2334 SD-OCT images and validated on 1171 images. An AUC of 0.81 with 60% sensitivity

and 71% specificity was achieved when tested on the external validation dataset [19].

In a study by Ma et al., 3704 B-scans from 42 eyes with nAMD, 4307 B-scans from 31 eyes with PCV, and a ResNet34 backbone were utilized. Several versions of the model were trained, achieving an AUC that ranged from 0.95 to 0.98. Clinical interpretability analysis was performed through projection of normalized aggregation of B-scan-based saliency maps onto the en face plane, with good correspondence seen between pathological features detected from the projected en face saliency map and ICGA [20].

In another cross-sectional study by Hwang et al. a deep CNN was built to detect nAMD and distinguish retinal angiomatous proliferation (RAP) and PCV. A total of 3951 SD-OCT images from 314 participants (229 AMD, 85 controls) were analyzed. When distinguishing between AMD (PCV or RAP) and normal cases, the proposed model had 99.1% accuracy, with a sensitivity and specificity of 99.2% and 99.1%, respectively. While the model had 89.1% accuracy in distinguishing between RAP and PCV, the sensitivity and specificity of the proposed model for RAP detection were 89.4% and 88.8%, respectively, and its AUC was 0.95 [95% confidence interval (CI) 0.72–0.85] [21].

Hwang et al. trained three different CNN backbones, ResNet50, InceptionV3, and VGG-16, to accurately diagnose AMD aimed for telemedicine. Using data augmentation, they generated 35,900 OCT images (7924 normal, 2906 dry AMD, 11,304 active wet AMD, and 13,766 inactive wet AMD). The models were trained on 28,720 images, validated on 7180 images, and finally clinically verified on 3872 images. The detection accuracy for the different stages of AMD was generally higher than 90% and was significantly superior ($p < 0.001$) to that of medical students (69.4% and 68.9%) and equal ($p = 0.99$) to that of retina specialists (92.7% and 91.9%) [22].

Sotoudeh-Paima et al. proposed a multiscale CNN, based on a feature pyramid network (FPN), for the classification of normal, dry (defined as drusen), and wet AMD (defined as CNV) in OCT images. The authors utilized VGG-16 as the backbone network. The training set

consisted of 12,649 retinal OCT images from 441 patients (120 normal, 160 drusen, and 161 CNV) and 108,312 OCT images from 4686 patients (3548 normal, 713 drusen, 791 CNV, and 709 DME) [4]. The testing set comprised of 1000 retinal OCT images (250 from each class) from the USCD dataset [4], and the best model achieved a performance of 93.4% accuracy [23].

Yan et al. used ResNet34, integrated with a convolutional block attention module (CBAM) block, to classify OCT images as normal, dry AMD (drusen), inactive wet AMD (inactive CNV), and active wet AMD (active CNV). The respective precision and recall for detection of drusen were 84.3% and 87.3%, 81.2% and 80.0% for inactive CNV, 97.7% and 90.2% for active CNV, and 93.7% and 96.5% for normal. The GradCam heatmap also indicated a high level of correspondence for ROI between the model and the ophthalmologists [24].

A CapsNet-based DL architecture was trained on a total of 726 spectral domain OCT (SD-OCT) images in a study by Celebi et al. to improve the accuracy of diagnosing AMD. The images were labeled as drusen ($n = 159$), dry AMD ($n = 145$), wet AMD ($n = 156$), and normal ($n = 266$). Images were pre-processed by identifying the region of interest (ROI) and speckle noise reduction based on optimized Bayesian non-local mean (OBNLM) filter. The proposed algorithm was validated on the public Kaggle dataset [14] and achieved an accuracy of 98.0%, sensitivity of 96.7%, and specificity of 99.9% [25].

An et al. utilized 185 normal OCT images, 535 OCT images of AMD with fluid, and 514 OCT images of AMD without fluid for training. During testing, 49 images from 25 normal eyes, 188 AMD OCT images with fluid, and 154 AMD images without any fluid from 77 AMD eyes were used. A two-stage classification process was developed. In the first stage, a VGG-16 model was trained to differentiate between normal and AMD OCT images. In the second stage, the VGG-16 model was then fine-tuned to detect fluid in AMD OCT images. During the first stage, normal and AMD OCT images were classified with 99.2% accuracy. In the second stage, fluid detection was achieved with a 95.1% accuracy [26].

Similarly, Motozawa et al. developed a two-stage process: first to distinguish healthy eyes from AMD and second to determine AMD with and without exudative changes. The first stage of the CNN model was able to classify AMD and normal OCT images with 100% sensitivity, 91.8% specificity, and 99.0% accuracy. The second stage was trained and validated with 721 AMD images with exudative changes and 661 AMD images without any exudation, achieving a performance of 98.4% sensitivity, 88.3% specificity, and 93.9% accuracy [27].

Potapenko et al. developed a CNN to detect retinal edema associated with AMD in OCT images. The algorithm was initially trained on a dataset consisting of 50,439 OCT images and associated treatment information retrieved from the Department of Ophthalmology, Rigshospitalet, Copenhagen, Denmark database, and later trained on a subset of the same dataset relabeled by three ophthalmologists. There was moderate intergrader agreement on the presence of edema in the relabeled data (76.4%). The proposed CNN performed on par with intergrader agreement in detecting edema on OCT images with an AUC of 0.97 and an accuracy rate of 90.9% [28].

Saha et al. developed a DL model for the automated detection and classification of AMD OCT biomarkers, including hyperreflective foci (HF), hyporeflective foci, and subretinal drusenoid deposits. A total of 1050 OCT B-scans with subretinal drusenoid deposits, 326 with intraretinal HF, and 206 with hyporeflective drusen were included in the training dataset. The algorithm achieved an overall accuracy of 87% for identifying the presence of these AMD biomarkers [29] (Table 1).

Clinical Implications of Classification/Detection Models

To date, multiple studies have demonstrated the robust performances of DL models in detecting the presence of AMD or classifying different stages of AMD on OCT images. Given the widespread availability of OCT technology, this development could lead to the decentralization of management for AMD. For example,

OCT machines can be set up in the community setting to perform screening on a population level, using DL models to automatically detect the presence of AMD-related pathologies in patients with no known history of AMD. Alternatively, for patients with known dry AMD, OCT imaging could be performed in a community, primary care or at-home setting on a regular basis, and these patients could be promptly referred to retinal specialists for further management, if signs of conversion to exudative AMD are detected by DL models.

Other Applications and Future Directions

While the focus of the current review is on DL applications to classification and detection of AMD in OCT images, other important applications exist. In the following sections we highlight some of these other applications and approaches, such as segmentation, prognostication, and incorporation of multimodal imaging.

Segmentation

DL techniques can be used to segment OCT images in order to detect and quantify biomarkers. Schlegl et al. developed a DL method to automatically detect and quantify intraretinal fluid (IRF) and subretinal fluid (SRF) in different retinal vascular diseases. The clinical dataset consisted of 1200 OCT volumes of patients with nAMD ($n = 400$), DME ($n = 400$), and retinal vein occlusion ($n = 400$). The proposed method achieved an AUC of 0.94 (range 0.91–0.97), a mean precision of 0.91, and a mean recall of 0.84 in detection of IRF and AUC of 0.92 (range 0.86–0.98), a mean precision of 0.61, and a mean recall of 0.81 in detection of SRF [30]. In a subsequent study by the same research group, Schmidt-Erfurth et al. used this algorithm to accurately measure the fluid response to anti-VEGF therapy in AMD using the HARBOR clinical trial dataset. The study utilized 24,362 3-D volumetric OCT scans from 1095 patients, who were treated over a period of 24 months, and biomarkers including IRF, SRF, and pigment epithelial detachment (PED) were

Table 1 Summary of the studies included in the current review article

Study and year	Clinical context	Neural network backbone	Performance
Detection of AMD			
He et al. [6], 2022	AMD vs. healthy controls	ResNet-50 and local outlier factor	AUC 0.99 Sensitivity 95.02% Specificity 95.02%
Lee et al. [7], 2017	AMD vs. healthy controls	VGG-16	AUC 0.92 Accuracy 87.63%
Shi et al. [8], 2021	Presence or absence of GA	Med-XAI-Net	AUC 0.93 Sensitivity 82.8% Specificity 94.6%
Srivastava et al. [9], 2020	AMD vs. normal	ResNet-50	AUC 0.96
Treder et al. [10], 2018	AMD vs. healthy controls	Multilayer deep CNN accelerated with TensorFlow	Sensitivity 100% Specificity 92% Accuracy 96%
Differentiation of AMD from other macular pathologies			
Baharlouei et al. [12], 2022	AMD vs. central serous retinopathy vs. diabetic retinopathy vs. macular hole vs. healthy controls	Wavelet scattering network; extracted features were fed to a principal component analysis (PCA) classifier	Accuracy for the detection of normal images 97.4% Accuracy for classification of pathologies 82.5%
Sun et al. [13], 2020	AMD vs. DME	ResNet-50	Sensitivity 99.2% Specificity 95.6% Accuracy 98.1%
Vellakani and Pushbam [15], 2020	AMD vs. DME vs. healthy controls	DenseNet201 and long short-term memory (LSTM) using OCT images enhanced by GAN	Accuracy 0.96 Positive predictive value 0.97 True-positive rate 0.96

Table 1 continued

Study and year	Clinical context	Neural network backbone	Performance
Alqudah [16], 2020	Drusen vs. AMD vs. CNV vs. DME vs. normal	AOCT-Net	Overall accuracy 95.3% Accuracy for AMD identification 100% Accuracy for CNV identification 98.6% Accuracy for drusen identification 98.9% Accuracy for DME identification 99.1% Accuracy for identification of controls 99.1%
Li et al. [17], 2019	AMD vs. DME	VGG-16	AUC 1 Sensitivity 97.8% Specificity 99.4% Accuracy 98.6%
Wongchaisuwat et al. [19], 2022	nAMD vs. PCV	ResNet	AUC 0.81 Sensitivity 60% Specificity 71%
Ma et al. [20], 2022	nAMD vs. PCV	ResNet34	AUC 0.95–0.98 (several versions of the model were trained and tested)
Hwang et al. [21], 2021	nAMD (PCV or RAP) vs. normal; PCV vs. RAP	VGG-19	nAMD vs. normal Accuracy 99.1% Sensitivity 99.2% Specificity 99.1% AUC 99.9% PCV vs. RAP Accuracy 89.1% Sensitivity 89.4% Specificity 88.8% AUC 0.953

Table 1 continued

Study and year	Clinical context	Neural network backbone	Performance
Hwang et al. [22], 2019	Dry AMD (drusen deposits) vs. active wet AMD (presence of subretinal fluid, intraretinal cysts, or hyporeflective space and subretinal hyperreflective exudate) vs. inactive wet AMD (lack of active wet AMD characteristics) vs. normal	ResNet50 InceptionV3 VGG-16	Detection accuracy for the different stages of AMD was higher than 90%
Sotoudeh-Paima et al. [23], 2022	Dry AMD (drusen) vs. wet AMD (CNV) vs. DME vs. normal	FPN VGG-16 ResNet50 DenseNet121	Best performing model (FPN-VGG-16) achieved 93.4% accuracy for the classification of dry vs. wet AMD
Yan et al. [24], 2021	Dry AMD (drusen) vs. inactive wet AMD (inactive CNV) vs. active wet AMD (active CNV) vs. normal	ResNet34	AUC 0.93–0.99
Celebi et al. [25], 2023	Drusen vs. dry AMD vs. wet AMD vs. healthy controls	CapsNet	Sensitivity 96.7% Specificity 99.9% Accuracy 98.07%
An et al. [26], 2019	2-step classification: 1—normal vs. AMD; 2—AMD with fluid vs. AMD without fluid	VGG-16	Accuracy for AMD detection 99.2% Accuracy for fluid detection 95.1%
Motozawa et al. [27]—2019	2-step classification: 1—normal vs. AMD; 2—AMD with exudative changes vs. AMD without exudative changes	Custom CNN	Normal vs. AMD Sensitivity 100% Specificity 91.8% Accuracy 99.0% Exudative vs. non-exudative AMD Sensitivity 98.4% Specificity 88.3% Accuracy 93.9%
Potapenko et al. [28]—2022	AMD with edema vs. AMD without edema	Custom CNN	AUC 0.97 Accuracy 90.9%

Table 1 continued

Study and year	Clinical context	Neural network backbone	Performance
Saha et al. [29]—2019	Hyperreflective foci vs. hyporefective foci vs. subretinal drusenoid deposits	Inception-v3 ResNet-50 InceptionresNet50	Accuracy for identifying the presence of biomarkers 86–89%

AMD age-related macular degeneration, *AUC* area under the receiver operating characteristic curve, *CNN* convolutional neural network, *CNV* choroidal neovascularization, *DME* diabetic macular edema, *FPN* feature pyramid network, *GA* geographic atrophy, *GAN* generative adversarial network, *LSTM* long short-term memory, *OCT* optical coherence tomography, *PCV* polypoidal choroidal vasculopathy, *RAP* retinal angiomatous proliferation, *VGG-16* visual geometry group 16

identified and quantified. The study revealed that mean central 1-mm residual fluid volumes were significantly lower in monthly treatment regimen when compared to pro re nata treatment regimen. The difference in residual IRF volume was detected as -0.55 nl ($p < 0.001$) and -1.11 nl ($p = 0.031$) for residual SRF volume, between treatment regimens by the end of the trial [31]. Gerendas et al. validated the performance of the same segmentation algorithm on real-world data, utilizing the database of Vienna Imaging Biomarker Eye Study (VIBES, 2007–2018); 1127 treatment-naïve AMD eyes with at least one follow-up visit within 5 years were included in the study. When compared to human experts, the algorithm achieved an AUC of 0.91 in detecting the presence of IRF and an AUC of 0.87 in detecting the presence of SRF in the entire scan [32].

Prediction/Prognostication

DL models have also been developed for prediction and prognostication purposes, such as predicting conversion to nAMD and treatment response.

Yim et al. developed a model to predict progression to nAMD in the fellow eye of patients, who had been diagnosed with nAMD in the first eye. Their model was trained with both raw OCT volumes and segmented tissue maps; 2795 patients across seven different sites were included. On a volumetric scan level, the model was able to predict imminent conversion to nAMD within 6 months with a sensitivity of

80% and specificity of 55% (sensitive mode) and sensitivity of 34% and specificity of 90% (specific mode) [33].

In another study, Pfau et al. first used DL (Deeplabv3 model with a ResNet-50 backbone) to segment OCT biomarkers (the different retinal layers and retinal pigment epithelium drusen complex), then used various classic ML techniques to predict the anti-VEGF treatment frequency within the next 12 months. In total, 99 nAMD eyes from 96 patients (138 visits) were included. The best performing model used random forest regression, and had a mean absolute error of 2.60 injections/year (2.25–2.96) [$R^2 = 0.39$] [34].

Fu et al. investigated the usefulness of OCT biomarkers (retinal pigment epithelium, IRF, SRF, PED, subretinal hyperreflective material, and HF) quantified by DL in predicting treatment response to anti-VEGF injections. Their study included 926 treatment-naïve eyes from 926 patients. The model that included baseline OCT biomarkers, baseline visual acuity (VA), and OCT/VA changes after injections had a mean absolute error of 5.6, 5.0, and 7.2 letters in predicting VA after injection #2, after injection #3, and at month 12, respectively [35].

Multimodal Imaging

Although OCT is the most common and important imaging modality for the management of AMD, incorporation of other imaging modalities into model training could further improve the models' performance. In this

section, we highlight studies that combined OCT images with other commonly used imaging modalities to train DL models.

Thakoor et al. developed a DL model to differentiate among non-AMD vs. nnAMD vs. nAMD. Nine model variations were trained, using different permutations of imaging modalities and configurations, including 2-D OCT B-scan flow images, 3-D OCT volumes, 3-D OCT volumes consisting of high-definition B-scans, and 3-D optical coherence tomography angiography (OCTA) volumes. The dataset consisted of 501 eye images (104 non-AMD, 247 nnAMD, and 150 nAMD) from 305 patients. The data was then split into training ($n = 301$), validation ($n = 100$), and testing ($n = 100$) sets. The models trained with multimodal images consistently outperformed models trained with only OCT and OCTA. The best performing model trained with multimodal imaging achieved an accuracy of $70.8 \pm 1.12\%$ for the testing dataset [36].

Similarly, Jin et al. also aimed to determine the efficacy of a DL model, trained with OCT and OCTA images, in the activation of CNV in nAMD. The model was developed with a feature level fusion (FLF) method. The training dataset included 924 images from 181 eyes ($n = 176$ patients) with AMD. The external datasets were used for testing: Ningbo dataset (30 images from 15 patients with AMD) and Jinhua dataset (42 images from 21 patients with AMD). During internal testing, the model achieved an accuracy of 95.5% (AUC = 0.97) comparable to human experts (accuracy ranging between 90.3% and 97.8%). The proposed model also achieved favorable results for both external dataset with an AUC of 1.0 for Ningbo and an AUC of 0.97 for Jinhua datasets, respectively [37].

Instead of combining OCT with OCTA, Chen et al. combined OCT with infrared reflectance (IR) images to develop a DL model for the classification of AMD. They collected 2006 pairs of IR and OCT images to develop the algorithms using vertical plane feature fusion (VFFF). The results were validated using an independent external dataset containing 506 pairs of IR and OCT images. The best performing model achieved an AUC of 0.95 for detection of

normal eyes, an AUC of 0.91 for detection of dry AMD, and 0.89 for detection of wet AMD in the external dataset. When compared to human observers the algorithm also performed similar results (accuracy of 92.5% and 91.5% for senior ophthalmologist and the algorithm, respectively) [38].

CONCLUSIONS AND FUTURE DIRECTIONS

In the current review, we focused on DL applications for classification and detection of AMD and OCT images. The vast majority of the published studies on DL/AMD/OCT pertained to “classification” tasks, such as detection of AMD, differentiation of AMD from other macular pathologies, and fine-grained classification of different AMD stages or features. In general, these studies demonstrated robust performances by their models, which typically involved a pre-trained CNN and transfer learning. More technical studies that focused primarily on DL-based segmentation, prediction/prognostication, and studies using multimodal images were beyond the scope of the current review, but were briefly mentioned as highlights. For each included study, we provided details for the clinical context, datasets, and DL technical information. In terms of future directions, studies combining OCT images and other imaging modalities will likely produce better-performing models. Prospective, randomized clinical trials will be needed to determine whether utilizing DL-based predictive models as clinical decision support tools can provide value by improving patient outcomes.

ACKNOWLEDGEMENTS

Author Contributions. Conceptualization: T.Y. Alvin Liu. Material preparation and data collection: Neslihan Dilruba Koseoglu. Writing: Neslihan Dilruba Koseoglu, T.Y. Alvin Liu. Writing-review and editing: Andrzej Grzybowski.

Funding. No funding or sponsorship was received for this study or publication of this article.

Data Availability. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Ethical Approval. The article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Conflict of Interest. Neslihan Dilruba Koseoglu and Andrzej Grzybowski have no conflict of interest/disclosures. T.Y. Alvin Liu received funding from Research to Prevent Blindness Career Advancement Award.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Wong WL, Su X, Li X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health*. 2014;2(2):e106–16. [https://doi.org/10.1016/s2214-109x\(13\)70145-1](https://doi.org/10.1016/s2214-109x(13)70145-1).
2. <https://www.cdc.gov/visionhealth/vehss/estimates/amd-prevalence.html>. Accessed 27 Jan 2023.
3. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: a survey. <https://arxiv.org/abs/2101.01169>. Accessed 28 June 2023.
4. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172(5):1122–1131.e9. <https://doi.org/10.1016/j.cell.2018.02.010>.
5. Farsiu S, Chiu SJ, O'Connell RV, et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. *Ophthalmology*. 2014;121(1):162–72. <https://doi.org/10.1016/j.ophtha.2013.07.013>.
6. He T, Zhou Q, Zou Y. Automatic detection of age-related macular degeneration based on deep learning and local outlier factor algorithm. *Diagnostics (Basel)*. 2022. <https://doi.org/10.3390/diagnostics12020532>.
7. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. *Ophthalmol Retina*. 2017;1(4):322–7. <https://doi.org/10.1016/j.oret.2016.12.009>.
8. Shi X, Keenan TDL, Chen Q, et al. Improving interpretability in machine diagnosis: detection of geographic atrophy in OCT scans. *Ophthalmol Sci*. 2021;1(3):100038. <https://doi.org/10.1016/j.xops.2021.100038>.
9. Srivastava R, Ong EP, Lee BH. Role of the choroid in automated age-related macular degeneration detection from optical coherence tomography images. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020;2020:1867–70. <https://doi.org/10.1109/embc44109.2020.9175809>.
10. Treder M, Lauer mann JL, Eter N. Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning. *Graefes Arch Clin Exp Ophthalmol*. 2018;256(2):259–65. <https://doi.org/10.1007/s00417-017-3850-3>.
11. Gholami P, Roy P, Parthasarathy MK, Lakshminarayanan V. OCTID: optical coherence tomography image database. <https://arxiv.org/abs/1812.07056>. Accessed 27 Jan 2023.
12. Baharlouei Z, Rabbani H, Plonka G. Detection of retinal abnormalities in OCT images using wavelet scattering network. *Annu Int Conf IEEE Eng Med Biol Soc*. 2022;2022:3862–5. <https://doi.org/10.1109/embc48229.2022.9871989>.

13. Sun Y, Zhang H, Yao X. Automatic diagnosis of macular diseases from OCT volume based on its two-dimensional feature map and convolutional neural network with attention mechanism. *J Biomed Opt.* 2020. <https://doi.org/10.1117/1.Jbo.25.9.096004>.
14. <https://www.kaggle.com/paultimothymooney/kermany2018>. 2020.
15. Vellakani S, Pushbam I. An enhanced OCT image captioning system to assist ophthalmologists in detecting and classifying eye diseases. *J Xray Sci Technol.* 2020;28(5):975–88. <https://doi.org/10.3233/xst-200697>.
16. Alqudah AM. AOCT-NET: a convolutional network automated classification of multiclass retinal diseases using spectral-domain optical coherence tomography images. *Med Biol Eng Comput.* 2020;58(1):41–53. <https://doi.org/10.1007/s11517-019-02066-y>.
17. Li F, Chen H, Liu Z, Zhang X, Wu Z. Fully automated detection of retinal disorders by image-based deep learning. *Graefes Arch Clin Exp Ophthalmol.* 2019;257(3):495–505. <https://doi.org/10.1007/s00417-018-04224-8>.
18. Yannuzzi LA, Sorenson J, Spaide RF, Lipson B. Idiopathic polypoidal choroidal vasculopathy (IPCVC). *Retina.* 1990;10(1):1–8.
19. Wongchaisuwat P, Thamphithak R, Jitpukdee P, Wongchaisuwat N. Application of deep learning for automated detection of polypoidal choroidal vasculopathy in spectral domain optical coherence tomography. *Transl Vis Sci Technol.* 2022;11(10):16. <https://doi.org/10.1167/tvst.11.10.16>.
20. Ma D, Kumar M, Khetan V, et al. Clinical explainable differential diagnosis of polypoidal choroidal vasculopathy and age-related macular degeneration using deep learning. *Comput Biol Med.* 2022;143:105319. <https://doi.org/10.1016/j.compbimed.2022.105319>.
21. Hwang DD, Choi S, Ko J, et al. Distinguishing retinal angiomatous proliferation from polypoidal choroidal vasculopathy with a deep neural network based on optical coherence tomography. *Sci Rep.* 2021;11(1):9275. <https://doi.org/10.1038/s41598-021-88543-7>.
22. Hwang DK, Hsu CC, Chang KJ, et al. Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics.* 2019;9(1):232–45. <https://doi.org/10.7150/thno.28447>.
23. Sotoudeh-Paima S, Jodeiri A, Hajizadeh F, Soltanian-Zadeh H. Multi-scale convolutional neural network for automated AMD classification using retinal OCT images. *Comput Biol Med.* 2022;144:105368. <https://doi.org/10.1016/j.compbimed.2022.105368>.
24. Yan Y, Jin K, Gao Z, et al. Attention-based deep learning system for automated diagnoses of age-related macular degeneration in optical coherence tomography images. *Med Phys.* 2021;48(9):4926–34. <https://doi.org/10.1002/mp.15002>.
25. Celebi ARC, Bulut E, Sezer A. Artificial intelligence based detection of age-related macular degeneration using optical coherence tomography with unique image preprocessing. *Eur J Ophthalmol.* 2023;33(1):65–73. <https://doi.org/10.1177/11206721221096294>.
26. An G, Akiba M, Yokota H, et al. Deep learning classification models built with two-step transfer learning for age related macular degeneration diagnosis. *Annu Int Conf IEEE Eng Med Biol Soc.* 2019;2019:2049–52. <https://doi.org/10.1109/embc.2019.8857468>.
27. Motozawa N, An G, Takagi S, et al. Optical coherence tomography-based deep-learning models for classifying normal and age-related macular degeneration and exudative and non-exudative age-related macular degeneration changes. *Ophthalmol Ther.* 2019;8(4):527–39. <https://doi.org/10.1007/s40123-019-00207-y>.
28. Potapenko I, Kristensen M, Thiesson B, et al. Detection of oedema on optical coherence tomography images using deep learning model trained on noisy clinical data. *Acta Ophthalmol.* 2022;100(1):103–10. <https://doi.org/10.1111/aos.14895>.
29. Saha S, Nassisi M, Wang M, et al. Automated detection and classification of early AMD biomarkers using deep learning. *Sci Rep.* 2019;9(1):10990. <https://doi.org/10.1038/s41598-019-47390-3>.
30. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology.* 2018;125(4):549–58. <https://doi.org/10.1016/j.ophtha.2017.10.031>.
31. Schmidt-Erfurth U, Vogl WD, Jampol LM, Bogunović H. Application of automated quantification of fluid volumes to anti-VEGF therapy of neovascular age-related macular degeneration. *Ophthalmology.* 2020;127(9):1211–9. <https://doi.org/10.1016/j.ophtha.2020.03.010>.
32. Gerendas BS, Sadeghipour A, Michl M, et al. Validation of an automated fluid algorithm on real-world data of neovascular age-related macular degeneration over five years. *Retina.* 2022;42(9):

- 1673–82. <https://doi.org/10.1097/iae.0000000000003557>.
33. Yim J, Chopra R, Spitz T, et al. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med*. 2020;26(6):892–9. <https://doi.org/10.1038/s41591-020-0867-7>.
 34. Pfau M, Sahu S, Rupnow RA, et al. Probabilistic forecasting of anti-VEGF treatment frequency in neovascular age-related macular degeneration. *Transl Vis Sci Technol*. 2021;10(7):30. <https://doi.org/10.1167/tvst.10.7.30>.
 35. Fu DJ, Faes L, Wagner SK, et al. Predicting incremental and future visual change in neovascular age-related macular degeneration using deep learning. *Ophthalmol Retina*. 2021;5(11):1074–84. <https://doi.org/10.1016/j.oret.2021.01.009>.
 36. Thakoor KA, Yao J, Bordbar D, et al. A multimodal deep learning system to distinguish late stages of AMD and to compare expert vs. AI ocular biomarkers. *Sci Rep*. 2022;12(1):2585. <https://doi.org/10.1038/s41598-022-06273-w>.
 37. Jin K, Yan Y, Chen M, et al. Multimodal deep learning with feature level fusion for identification of choroidal neovascularization activity in age-related macular degeneration. *Acta Ophthalmol*. 2022;100(2):e512–20. <https://doi.org/10.1111/aos.14928>.
 38. Chen M, Jin K, Yan Y, et al. Automated diagnosis of age-related macular degeneration using multimodal vertical plane feature fusion via deep learning. *Med Phys*. 2022;49(4):2324–33. <https://doi.org/10.1002/mp.15541>.