AMERICAN ACADEMY OF OPHTHALMOLOGY®

A Convolutional Neural Network Using Multimodal Retinal Imaging for Differentiation of Mild Cognitive Impairment from Normal Cognition

C. Ellis Wisely, MD, MBA,[1,2,*] Alexander Richardson,[1,2,3,4,*] Ricardo Henao, PhD,[2,3,4] Cason B. Robbins, MD,[1,2] Justin P. Ma, MD,[1,2] Dong Wang, PhD,[2,3,4] Kim G. Johnson, MD,[5] Andy J. Liu, MD,[5] Dilraj S. Grewal, MD,[1,2] Sharon Fekrat, MD, FASRS[1,2,5]

Purpose: To develop a machine learning tool capable of differentiating eyes of subjects with normal cognition from those with mild cognitive impairment (MCI) using OCT and OCT angiography (OCTA).

Design: Evaluation of a diagnostic technology.

Participants: Subjects with normal cognition were compared to subjects with MCI.

Methods: A multimodal convolutional neural network (CNN) was built to predict likelihood of MCI from ganglion cell-inner plexiform layer (GC-IPL) thickness maps, OCTA images, and quantitative data including patient characteristics.

Main Outcome Measures: Area under the receiver operating characteristic curve (AUC) and summaries of the confusion matrix (sensitivity and specificity) were used as performance metrics for the prediction outputs of the CNN.

Results: Images from 236 eyes of 129 cognitively normal subjects and 154 eyes of 80 MCI subjects were used for training, validating, and testing the CNN. When applied to the independent test set using inputs including GC-IPL thickness maps, OCTA images, and quantitative OCT and OCTA data, the AUC value for the CNN was 0.809 (95% confidence interval [CI]: 0.681−0.937). This model achieved a sensitivity of 79% and specificity of 83%. The AUC value for GC-IPL thickness maps alone was 0.681 (95% CI: 0.529−0.832), for OCTA images alone was 0.625 (95% CI: 0.466−0.784) and for both GC-IPL maps and OCTA images was 0.693 (95% CI: 0.543−0.843). Models using quantitative data alone were also tested, with a model using quantitative data derived from images, 0.960 (95% CI: 0.902−1.00), outperforming a model using demographic data alone, 0.580 (95% CI: 0.417−0.742).

Conclusions: This novel CNN was able to identify an MCI diagnosis using an independent test set comprised of OCT and OCTA images and quantitative data. The GC-IPL thickness maps provided more useful decision support than the OCTA images. The addition of quantitative data inputs also provided significant decision support to the CNN to identify individuals with MCI. Quantitative imaging metrics provided superior decision support than demographic data.

Financial Disclosure(s): Proprietary or commercial disclosure may be found after the references.

Mild cognitive impairment (MCI), often considered the clinical precursor to Alzheimer's disease (AD), is an intermediate state where cognitive impairment is present, but the ability to perform activities of daily living is preserved.[1] Mild cognitive impairment can be further classified into amnestic and nonamnestic types, with amnestic MCI referring to a state where memory is affected more significantly than other domains of cognitive function (e.g., executive, language, or visuospatial).[2,3] Most individuals diagnosed with MCI progress to AD[1]; however, patients with amnestic MCI are more likely to progress to AD than individuals with nonamnestic disease.[4] Some estimates indicate amnestic patients with MCI may progress to AD at rates of approximately 20% per year.[5] Thus, early identification of MCI is important for effective intervention, particularly as new therapies become available. Given the existing unmet need of accurately establishing an MCI diagnosis, especially in those who will progress to AD, in the setting of an aging worldwide population, reliable biomarkers that are also noninvasive and widely available are needed to support the clinical diagnosis of MCI.

Currently, a definitive diagnosis of AD is made through brain tissue histopathology at autopsy; however, antemortem testing modalities, including neuroimaging with brain positron emission tomography or magnetic resonance imaging (MRI),[6] cerebrospinal fluid sampling, and serologic

https://doi.org/10.1016/j.xops.2023.100355
ISSN 2666-9145/23

testing,[7] are becoming surrogate diagnostic tools. The expense, invasiveness, and sensitivity of these diagnostic tools have limited their widespread use in clinical practice.[6,7] By comparison, retinal imaging is noninvasive, has a low recurring cost, and is widely available.[8,9] Prior studies have indicated that structural changes in the neurosensory retina and its microvasculature can be identifiable in patients with AD using OCT[8,10–19] and OCT angiography (OCTA).[11,16,17,19–22] Similarly in patients with MCI, such findings on OCT, including changes in retinal nerve fiber layer (RNFL) thickness[8,19,23–25] and ganglion cell–inner plexiform layer (GC-IPL) thickness,[12,14,19] have been demonstrated, with less pronounced thinning of the RNFL and GC-IPL than in individuals with AD. Retinal microvascular abnormalities demonstrable via OCTA imaging, particularly decreased vessel density (VD) in the parafoveal superficial capillary plexus (SCP),[26] have also been demonstrated in patients with MCI. Additionally, studies have documented a clinical correlation between OCT retinal thickness parameters and cognitive testing such as Mini Mental State Examination (MMSE) scores in patients with MCI.[27] Similarly, OCTA VD and perfusion density (PD) parameters have been correlated with MMSE scores in patients with AD[17] and MCI.[19] Our prior work using traditional statistical analyses did not demonstrate significant differences between cognitively normal persons and those with MCI, except for temporal RNFL thickness; however, that MCI cohort was relatively small.[19]

Recently, machine learning models have been developed to provide neurologists adjunctive measures to establish clinical diagnoses of AD and MCI. Models using brain MRI[28,29] or combinations of MRI and positron emission tomography images[30,31] have demonstrated the ability to successfully differentiate individuals with normal cognition from those with AD, as well as from individuals with MCI. One model has also demonstrated the ability to predict which patients with MCI will progress to AD using analysis of baseline MRIs.[32] Recently, machine learning models using ophthalmic images have demonstrated promise for differentiation of AD and control subjects. Tian et al[33] developed a model using fundus photos with a focus on the retinal vasculature to differentiate AD and control subjects. Our group has previously shared a potential diagnostic support machine learning tool for AD,[34] which utilized easily obtainable OCT and OCTA imaging inputs along with clinical data instead of relying on more invasive and costly brain MRI and positron emission tomography image inputs. In the present study, we sought to create a convolutional neural network (CNN) algorithm capable of differentiating patients with MCI from those with normal cognition.

Methods

The images obtained for CNN input were from an ongoing clinical trial (Clinicaltrials.gov, NCT03233646) approved by Duke University Health System's institutional review board. This study followed the tenets of the Declaration of Helsinki. Written informed consent was obtained for each study subject.

Participants

A detailed description of patient eligibility and selection criteria has been previously published.[19] In brief, patients were excluded if they had dementia, diabetes mellitus, uncontrolled hypertension, demyelinating disorders, high myopia or hyperopia, glaucoma, age-related macular degeneration, other vitreoretinal pathology, history of ocular surgery apart from cataract surgery, and corrected ETDRS visual acuity $<$ 20/40 at the time of image acquisition. Patients were screened for vitreoretinal or optic nerve pathology on review of ultra-widefield scanning laser ophthalmoscopy color fundus images (Optos, California).

Eligible subjects with MCI and normal cognition were selected based on clinical diagnoses, as outlined in our prior published work.[34] Subjects with MCI were diagnosed based on clinical history, cognitive testing, and any available neuroimaging by an expert neurologist (A.L.) with specialization in memory disorders. All diagnoses were made in accordance with National Institute on Aging and Alzheimer's Association guidelines.[35] Cognitively normal control subjects were adults $\geq$ 50 years old who volunteered from the local community or who were participants in the Duke Alzheimer's Disease Research Center's registry for cognitively normal community subjects.

In order to account for age as a potential confounding variable between the MCI and control groups, stochastic age-matching for the test dataset was employed before each model run; 80% of the MCI eyes in the test dataset were randomly selected, and, for each MCI eye, the control eye with the nearest age was also added to the age-matched test dataset. After age-matching selection, a t test was conducted to ensure that the age differences between these MCI and control groups were statistically insignificant ($\alpha = 0.05$) before the age-matched dataset was used in testing. This age-matching procedure was repeated 10 times for each model, and the results were reported in aggregate. This procedure resulted in a balanced test dataset with no significant differences in age between the MCI and control groups.

Imaging

All subjects underwent undilated imaging using the Zeiss Cirrus HD-5000 Spectral-Domain OCT with AngioPlex OCTA (Carl Zeiss Meditec, Version 11.0.0.29946).[19] Color maps depicting GC-IPL thickness were generated after from the macular cube 512 $\times$ 128 scan with automated segmentation using the Zeiss OCT software. OCTA *en face* 3- $\times$ 3-mm and 6- $\times$ 6-mm images centered on the fovea were also acquired. OCTA images herein refer to the SCP *en face* 6- $\times$ 6-mm images. Images were manually assessed by experienced study staff, and poor quality images were excluded based on low signal strength ($<$ 7/10), shadow artifact, segmentation and motion artifacts, and poor centration.[34]

Quantitative Data

The CNNs described herein incorporated quantitative (structured) data including patient characteristics of age, sex, ETDRS visual acuity converted to the logarithm of the minimal angle of resolution, and years of education, as well as quantitative OCT and OCTA parameters. Quantitative data not derived from images, such as age, sex, corrected visual acuity, and years of education, referred to in the Results as "Demographic," was excluded from some analyses as noted. Briefly, quantitative OCT parameters included subfoveal choroidal thickness, central subfield thickness, average GC-IPL thickness, and average RNFL thickness; 3- $\times$ 3-mm OCTA quantitative parameters included area of the foveal avascular zone, foveal avascular zone circularity index, SCP PD in the 3-mm ETDRS circle and 3-mm ring,[19] and SCP

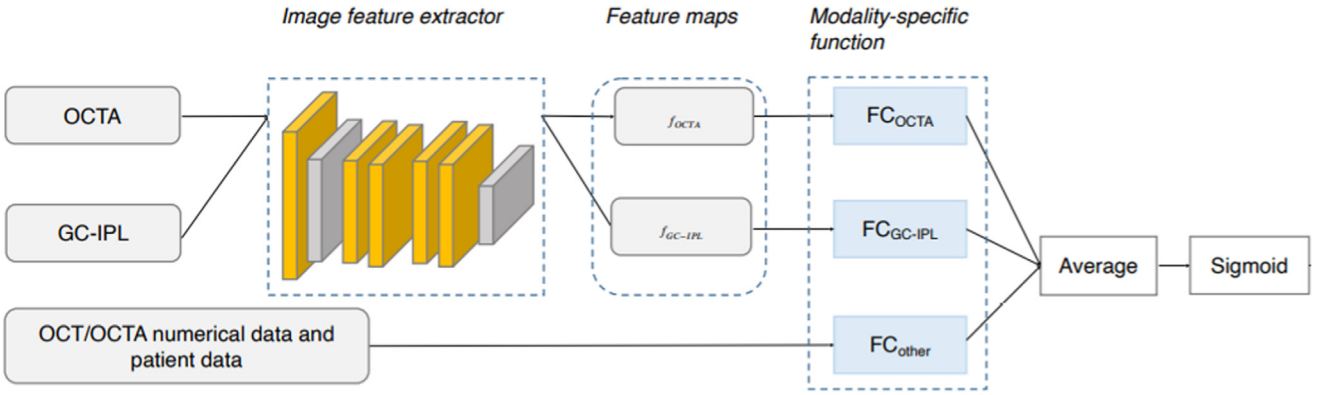

Figure 1. Illustration of the structure of the convolutional neural network for differentiation of subjects with mild cognitive impairment from cognitively normal controls. The convolutional neural network uses a convolutional encoder for the image inputs (OCT angiography [OCTA] superficial capillary plexus 6 × 6 mm *en face* images and ganglion cell-inner plexiform layer [GC-IPL] thickness maps), image modality specific feature transformations (f_{OCTA} and $f_{GC\text{-}IPL}$), prediction heads for all modalities (FC_{OCTA}, $FC_{GC\text{-}IPL}$ and FC_{other}) whose outputs are aggregated using averages and then processed through a sigmoid activation function to yield a model score.

VD in the 3-mm circle and 3-mm ring; 6- × 6-mm OCTA quantitative parameters included PD and VD in the 6-mm circle, 6-mm outer ring, and 6-mm inner ring.[34] All parameters were calculated by the Zeiss software apart from subfoveal choroidal thickness, which was manually measured by 2 graders as a linear measurement from the hyper-reflective line of the outer border of the retinal pigment epithelium perpendicularly to the hyper-reflective sclerochoroidal junction on the enhanced depth imaging foveal scan (captured using the HD 21 line raster protocol), with discrepancies adjudicated by a third grader.

Model Development

The model was trained to receive inputs of GC-IPL thickness color maps, OCTA SCP 6 × 6 mm *en face* images, and quantitative data as outlined above to produce a probability score that would suggest whether a patient carried a clinical diagnosis of MCI. The model illustrated in Figure 1 consists of shared convolutional encoder for the image inputs (GC-IPL and OCTA), image modality specific feature transformations (f_{OCTA} and $f_{GC\text{-}IPL}$), prediction heads for all modalities (FC_{OCTA}, $FC_{GC\text{-}IPL}$, and FC_{other}) whose outputs get aggregated using simple averages and then processed through a sigmoid activation function to produce a total model output. The model is similar in structure to that previously introduced to differentiate AD subjects from controls,[34] both of which leveraged transfer learning from a ResNet18[36] convolutional encoder trained with ImageNet data. The feature transformations are specified as single fully connected layers. The model also employs 2 separate dropout layers, 1 for quantitative features and 1 for image features, to reduce overfitting to the training dataset. The model was trained for 100 epochs using the adaptive moment estimation optimizer[37] with a weight decay of 0.01 and using the binary cross entropy loss function. As a quantitative estimate of performance, we considered the receiver operating characteristic, the area under the receiver operating characteristic curve (AUC), and summaries of the confusion matrix, namely, accuracy, sensitivity, and specificity. In order to convert the model scores into a dichotomous decision (MCI vs. control), we thresholded the model scores with the label set to optimize the Youden index.[38] Given the complexity of the model relative to the size of the data used for training, we considered simplified versions in which only 1 or both imaging modalities or quantitative data only were used as inputs, which we contrasted with the full model that also leveraged quantitative data described above. DeLong tests were used to compare the

Table 1. Demographic Data for MCI and Controls

	Total Eyes	Sex (Male)		Age (Yrs)	Yrs of Education	logMAR VA	MMSE Score*
MCI							
N	154	69	Mean	72.5	15.6	0.14	25.3
%	39.5	44.8	SD	7.1	2.3	0.10	4.3
Controls							
N	236	71	Mean	68.6	17.2	0.11	29.2
%	60.5	30.1	SD	7.7	2.3	0.10	1.2
P value†		0.018†		< 0.001‡	< 0.001‡	0.044§	< 0.001§

logMAR = logarithm of the minimum angle of resolution; MCI = mild cognitive impairment; MMSE = Mini Mental State Examination; SD = standard deviation; VA = visual acuity.
*MMSE scores were not included in the convolutional neural network.
†Calculated using Fisher exact test.
‡Calculated using Wilcoxon rank sum test.
§*P* value based on generalized estimating equations analysis of difference between means accounting for the correlation between eyes of the same subject, adjusted for age and sex.

Table 2. Quantitative OCT Angiography Summary Data

	FAZ Area (mm^2)	FAZ Circle	PD 3-mm Circle	PD 3-mm Ring	VD 3-mm Circle (mm/mm^2)	VD 3-mm Ring (mm/mm^2)	PD 6-mm Circle	PD 6-mm Inner ring	PD 6-mm Outer ring	VD 6-mm Circle (mm/mm^2)	VD 6-mm Inner Ring (mm/mm^2)	VD 6-mm Outer Ring (mm/mm^2)
MCI (N = 154)												
Mean	0.23	0.62	0.36	0.30	19.90	20.98	0.43	0.42	0.45	17.68	17.63	17.97
SD	0.10	0.13	0.03	0.10	1.72	1.69	0.03	0.04	0.03	1.32	1.62	1.34
Controls (N = 236)												
Mean	0.25	0.66	0.36	0.39	20.15	21.24	0.44	0.43	0.45	17.91	17.90	18.24
SD	0.12	0.08	0.03	0.03	1.56	1.53	0.03	0.03	0.02	1.00	1.35	0.96
P value*	0.133	0.063	0.661	<0.001	0.548	0.509	0.575	0.500	0.411	0.772	0.662	0.433

FAZ = foveal avascular zone; MCI = mild cognitive impairment; PD = perfusion density; SD = standard deviation; VD = vessel density.
**P* value based on generalized estimating equations analysis of difference between means accounting for the correlation between eyes of the same subject, adjusted for age and sex.

performance metrics for each set of inputs, using the input combination of GC-IPL maps, OCTA images, and image-based quantitative data as the reference value.

Results

Two hundred thirty-six eyes of 129 control subjects and 154 eyes of 80 MCI subjects were used for development and testing of the model. One hundred fifty-two (64%) control eyes were used for training, 24 (10%) were used for validation, and 60 (25%) were used for testing. One hundred four (68%) MCI eyes were used for training, 20 (13%) were used for validation, and 30 (19%) were used for testing. There was no patient overlap between the training, validation, and test groups, and patients were randomly assigned to each group. When model evaluation was performed on the test groups, individual eyes were randomly sampled so that the difference in ages between the MCI and control groups was not statistically significant, as previously outlined herein. Performance characteristics are reported as the median performance of 10 runs of the model.

Demographic information, visual acuity data, and MMSE scores appear in Table 1. Tables 2 and 3 summarize the quantitative OCTA and OCT data, respectively.

Table 4 demonstrates AUC values for 4 different CNNs. Each used the structure outlined in Figure 1, but with various input combinations as outlined. The GC-IPL thickness maps as a single input outperformed OCTA images. The best-performing model used all available inputs, including GC-IPL maps, OCTA images, and quantitative data. We report an AUC value of 0.809 (95% confidence interval: 0.681−0.937) for this input combination. To ensure consistency of this result, we reassigned eyes randomly to the training, validation, and testing groups in 10 variations and ran the CNN for each grouping, resulting in a mean AUC of 0.828 and standard deviation of 0.035. Thus, our reported AUC of 0.809 represents a conservative (37th percentile) description of CNN performance.

Figure 2 demonstrates model scores for each eye in the MCI and control test sets evaluated by the best-performing model. Figure 3 is the receiver operating characteristic curve for the best-performing model on the test set. The values in Figure 3 were used to calculate a Youden index threshold value of 0.095, as displayed in Figure 2. This Youden threshold was used as the decision point for determining sensitivity and specificity. Our best-performing model achieved sensitivity of 79% and specificity of 83%. More specifically, Figure 4 shows the confusion matrix, demonstrating how each eye in the test set was classified

Table 3. Quantitative OCT Summary Data

	Mean GC-IPL Thickness (μm)	Mean RNFL Thickness (μm)	Subfoveal Choroidal Thickness (μm)	Central Subfield Thickness* (μm)
MCI (N = 154)				
Mean	75.84	89.24	256.49	263.99
SD	8.62	9.62	76.49	36.93
Controls (N = 221)				
Mean	77.40	89.87	276.95	266.48
SD	7.31	10.18	92.39	24.56
P value†	0.778	0.731	0.332	0.345

GC-IPL = ganglion cell-inner plexiform layer; MCI = mild cognitive impairment; RNFL = retinal nerve fiber layer; SD = standard deviation.
†*P* value based on generalized estimating equations analysis of difference between means accounting for the correlation between eyes of the same subject, adjusted for age and sex.

Table 4. Predictive Capabilities of Each Convolutional Neural Network

	GC-IPL Thickness	OCTA	Quant Data (Image-Based)	Quant Data (Demographic)	AUC on Test Set [95% CI]*	*P* Value†
Quantitative data				X	0.580 [0.417, 0.742]	0.0543
OCTA		X			0.625 [0.466, 0.784]	0.0458
GC-IPL	X				0.681 [0.529, 0.832]	0.0881
OCTA and GC-IPL	X	X			0.693 [0.543, 0.843]	0.0852
Quantitative data			X		0.960 [0.902, 1.00]	0.0120
GC-IPL, OCTA, and quant data	X	X	X		0.809 [0.681, 0.937]	N/A

AUC = area under the receiver operating characteristic curve; CI = confidence interval; GC-IPL = ganglion cell-inner plexiform layer; OCTA = OCT angiography; quant = quantitative.
AUC on validation and test set figures describe the performance of each model on predicting the probability of mild cognitive impairment diagnosis for each individual eye in the independent sets. An "X" indicates that the input(s) were included in the model described in that row.
*Confidence intervals obtained using the DeLong method.
†*P* values for DeLong tests comparing the performance of the model for each set of inputs to the input set (GC-IPL, OCTA, and image-based quantitative data [excludes demographic data]).

(either as MCI or as control) by the model using the threshold value selected via the Youden index.

Discussion

The CNN model developed herein is the first to use retinal OCT and OCTA images to differentiate subjects with MCI from individuals with normal cognition. Our best performing model using images, including GC-IPL thickness color maps and OCTA SCP 6 × 6 mm *en face* images, as well as quantitative data (including only data derived from the OCT and OCTA images), achieved an AUC of 0.809 when applied to an independent test set. This performance quality was similar to the performance demonstrated in our previously developed CNN for differentiation of control and AD subjects.[34]

Both OCTA images and GC-IPL maps were selected as image inputs for this model since these images were found to be useful modalities in our prior machine learning work that distinguished individuals with AD from control subjects with normal cognition.[34] As demonstrated in Table 4, GC-IPL thickness maps were the most useful image input evaluated in our MCI model, in alignment with our prior work.[34] Of note, the performance of OCTA image inputs alone in distinguishing MCI subjects was similar to the observed performance of these images in our previously published model that identified individuals with AD.

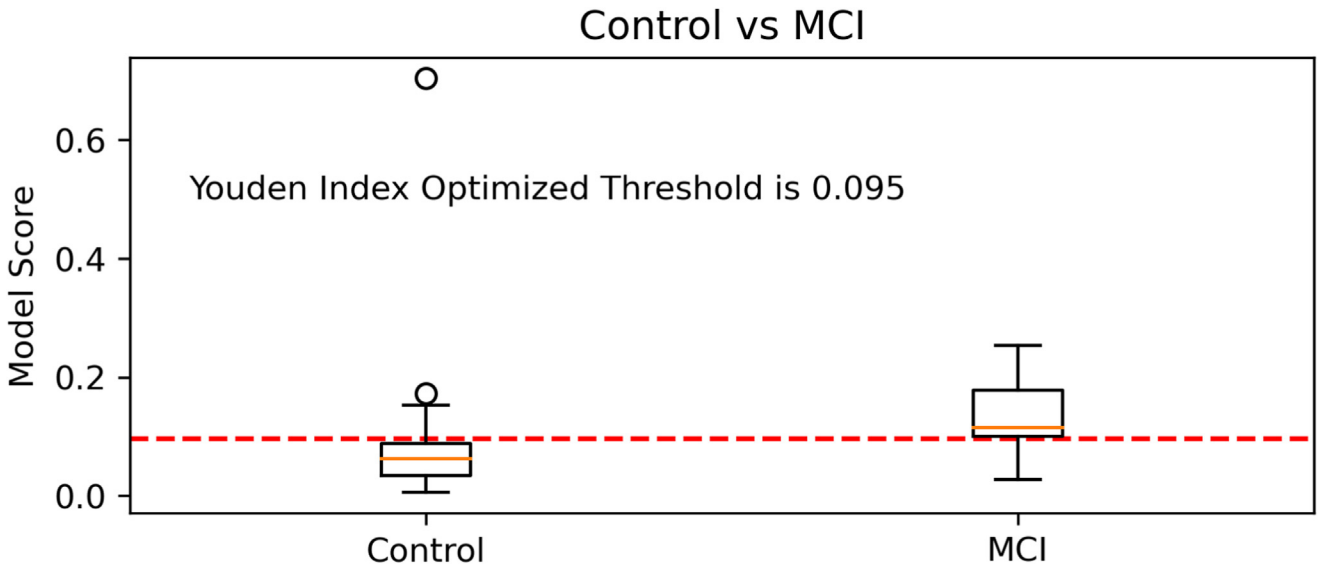

Figure 2. Box and whisker plot of model predictions for the best-performing model using ganglion cell-inner plexiform layer thickness maps, OCT angiography 6 × 6 mm *en face* images, and quantitative data inputs for the cognitively normal control and mild cognitive impairment (MCI) groups in the test dataset. The control median is 0.063 (interquartile range: 0.034–0.088) and MCI median is 0.115 (interquartile range: 0.100–0.178). The red dotted line represents the Youden index optimized threshold value. Model scores above this red line are classified as MCI, while model scores below are classified as control. Orange lines present the median model score values for the MCI and control patients in the test dataset. Circles represent outliers, defined as scores outside of the ([Q1-1.5 IQR], [Q3+1.5 IQR]) margin.

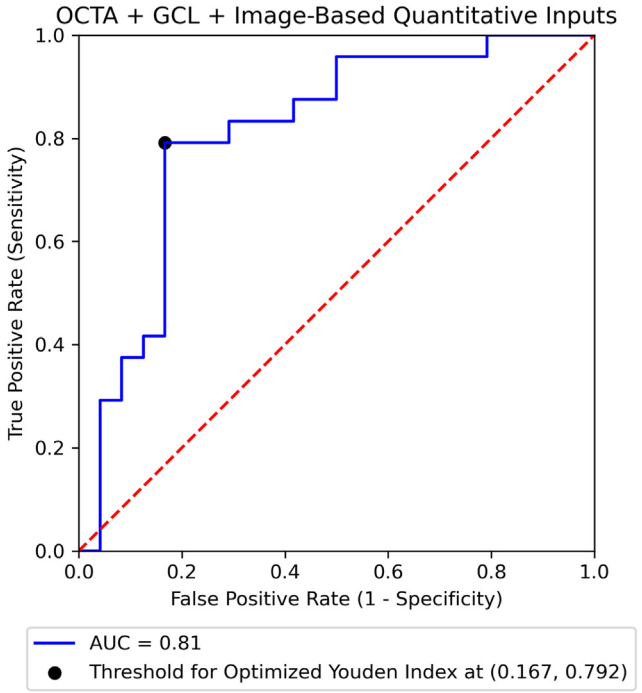

Figure 3. Receiver operating characteristic curve for model scores for each eye in the test set (blue line) using the ganglion cell-inner plexiform layer (GCL) thickness maps, OCT angiography (OCTA) 6 × 6 mm *en face* images, and quantitative data model. Area under the receiver operating characteristic curve (AUC) values for the validation and test sets along with sensitivity and specificity values for the test set are reported. The diagonal dotted line represents the performance of a completely random classifier (1/2 probability assigned to either classification). An ROC curve above the dotted diagonal line represents better classification performance compared to random chance.

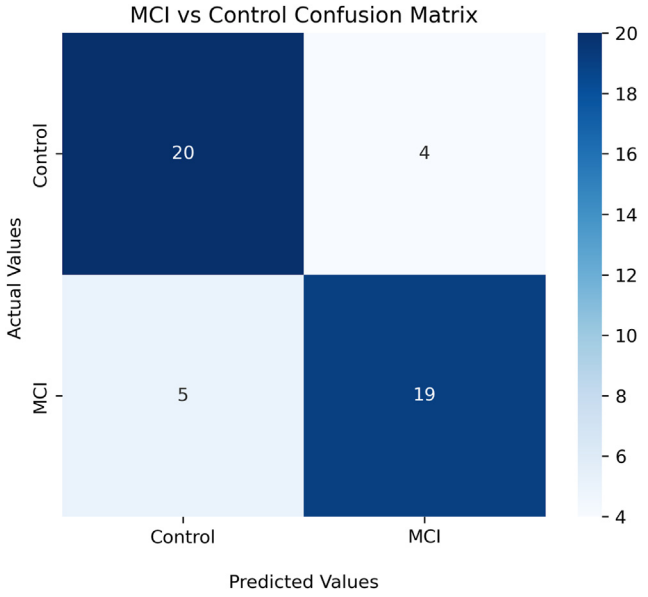

Figure 4. Confusion matrix demonstrating the classification of each test set eye by the ganglion cell-inner plexiform layer thickness maps, OCT angiography 6 x 6 mm *en face* images, and quantitative data model. MCI = mild cognitive impairment.

Traditional statistical analysis of only OCTA image quantitative metrics to differentiate MCI from controls has been significant in some, but not all, published studies,[26,39,40] suggesting that OCTA images alone as model inputs may have utility in MCI diagnosis, particularly if the images are obtained from individuals with amnestic MCI as opposed to nonamnestic MCI.[39] However, the findings in our present study and prior work[19] demonstrate that the utility of such OCTA images as sole model inputs currently remains an open question which may be, in part, attributed to variations in the MCI population being studied.[19,22,41] The addition of image-based quantitative metrics to the images themselves as model inputs has value (Table 4) and resulted in a marked improvement in performance when compared to our images-only model.

We hypothesize that the model's reliance on quantitative data suggests that we may require larger training sets to achieve optimal performance with images alone. The size of our training set limited the complexity of our model, since employing a bigger (more expressive) neural network with more or wider convolutional layers would risk overfitting. If the neural network has too many parameters (neurons) compared to the dataset it is training on, the neural network will tend to overfit, i.e., 'memorize' the relatively small dataset, limiting its ability to generalize to new (test) data. CNNs are complex neural network architectures that require large training datasets to fully extract features from images, especially with subtle features such as the retinal thickness and microvasculature changes between a control subject and a subject with MCI. Using quantitative features extracted from the image data as an additional input, utilizing a pre-trained neural network as a foundation, and reducing the pretrained neural network size were all measures taken to circumvent the challenges posed by the subtle biomarker differences and the dataset size constraints.

The CNN developed herein demonstrated a clear benefit in performance from the use of quantitative OCT and OCTA data inputs in differentiation of eyes of control and MCI subjects. In developing this CNN, we used a dataset that included subjects from our prior work[19,41] where we found no significant differences in OCTA parameters between the control and MCI groups. Additionally, we also found no differences in GC-IPL thickness parameters, despite the apparent utility of thickness maps to the CNN model (Table 4). The only quantitative OCT parameter demonstrating a significant difference between control and MCI groups in our prior work was temporal quadrant RNFL thickness.[19] Table 1 also demonstrates the similarity of the quantitative OCT and OCTA imaging parameters between the control and MCI groups. Thus, this demonstrates the effectiveness of the CNN in detecting signals differentiating the 2 datasets that may not be apparent with traditional statistical methods, and this may partially be attributed to the CNN's ability to utilize the entire available image rather than being limited to certain regions, as is used for generating quantitative data alone.

Sharing our work thus far stimulates this burgeoning field as we continue to recruit and image cognitively normal subjects and MCI subjects to further improve performance of this CNN. Our relatively smaller sample size was a study limitation that contributed to the lack of an age-matched control cohort. As a result, we employed an accepted sampling technique for post hoc age-matching for our test dataset outlined herein to limit any potential confounding effect of age on the model's performance. We purposefully excluded patients with known ocular and systemic diseases such as glaucoma, diabetes, and vitreoretinal pathology. Doing so allowed us to provide a proof of concept and build a solid starting point that would spark additional efforts in this space around the world. Because we as a field are still exploring the creation of a CNN for neurodegenerative diseases such as MCI, there is value in using strictly-controlled, high-quality images from individuals without these concomitant ocular or systemic diseases (potential confounders) for early work, and, then, with this foundation, we can begin to better understand how these other ocular and systemic conditions contribute. Future work will incorporate individuals with these and other ocular and systemic diseases that will ultimately expand the generalizability of such a CNN to more broadly differentiate those with MCI from those with normal cognition. Additionally, given the initial success of this CNN in the current pilot dataset, we will identify partner institutions with similar OCT and OCTA datasets that may further enlarge and complement our dataset. By standardizing imaging platforms and protocols with peer institutions, validation of our CNN with external datasets will be facilitated. The incorporation of federated learning may also provide a path forward to obviate associated challenges and concerns regarding the sharing of protected health information comprising image-based datasets with other institutions. A longitudinal dataset from individuals who have progressed from normal cognition to MCI to AD is a particularly exciting area of future study. Of note, our most recent work has demonstrated greater velocity of decline in OCTA PD and VD in patients with MCI compared to cognitively normal individuals that occur prior to significant MMSE changes.[41] This provides further optimism around our primary goal: development of a model that can predict MCI or Alzheimer's onset prior to the manifestation of clinical symptoms.

Footnotes and Disclosures

Originally received: June 30, 2022.
Final revision: June 8, 2023.
Accepted: June 14, 2023.
Available online: June 25, 2023. Manuscript no. XOPS-D-22-00154

[1] Department of Ophthalmology, Duke University School of Medicine, Durham, North Carolina.

[2] iMIND Study Group, Duke University School of Medicine, Durham, North Carolina.

[3] Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina.

[4] Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina.

[5] Department of Neurology, Duke University School of Medicine, Durham, North Carolina.

*C.E.W. and A.R. contributed equally to this work.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

A.R.: Support for attending meetings and/or travel — Vit-Buckle Society Academic Award, YoungMDConnect Live Travel Grant.

A.J.L.: Grants or contracts from any entity — Ann B. Bussel Award 2022-2024; Consulting fees — Darwin Research Group; Payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events —University of Kansas ADRC presentation; Leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid —Junior Associate clinical core lead Duke-UNC ADRC.

S.F.: Grants or contracts from any entity — Optos, Genentech; Royalties or licenses — Alcon—patent royalties; Consulting fees — Glaukos; Participation on a Data Safety Monitoring Board or Advisory Board — Bausch Surgical; Receipt of equipment, materials, drugs, medical writing, gifts or other services — Optos California machine loaner.

Funded in part by the Alzheimer's Drug Discovery Foundation.

HUMAN SUBJECTS: The images obtained for CNN input was from an ongoing clinical trial (Clinicaltrials.gov, NCT03233646) and was approved by the Duke University Health System investigational review board. This study followed the tenets of the Declaration of Helsinki. Written informed consent was obtained for each study subject. No animal subjects were used in this study.

Author Contributions:

Conception and design: Wisely, Richardson, Henao, Wang, Johnson, Liu, Grewal, Fekrat

Data collection: Wisely, Richardson, Henao, Robbins, Ma, Wang, Grewal, Fekrat

Analysis and interpretation: Wisely, Richardson, Henao, Robbins, Ma, Wang, Grewal, Fekrat; Obtained funding: N/A

Overall responsibility: Wisely, Richardson, Henao, Robbins, Ma, Wang, Johnson, Liu, Grewal, Fekrat

Abbreviations and Acronyms:
AD = Alzheimer's disease; **AUC** = area under the receiver operating characteristic curve; **CNN** = convolutional neural network; **GC-IPL** = ganglion cell-inner plexiform layer; **MRI** = magnetic resonance imaging; **MCI** = mild cognitive impairment; **MMSE** = Mini Mental State Examination; **OCTA** = OCT angiography; **PD** = perfusion density; **RNFL** = retinal nerve fiber layer; **SCP** = superficial capillary plexus; **VD** = vessel density.

Keywords:
Convolutional neural network, Machine learning, Mild cognitive impairment, OCT, OCT angiography.

Correspondence:
Sharon Fekrat, MD, FASRS, Duke Eye Center, 2351 Erwin Road, Durham, NC 27710. E-mail: sharon.fekrat@duke.edu.

References

1. Jack Jr CR, Albert MS, Knopman DS, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:257–262.
2. Ward A, Arrighi HM, Michels S, Cedarbaum JM. Mild cognitive impairment: disparity of incidence and prevalence estimates. *Alzheimers Dement*. 2012;8:14–21.
3. Petersen RC, Smith G, Waring SC, et al. Mild cognitive impariment: clinical characterization and outcome. *Arch Neurol*. 1999;56:303–308.
4. Bradfield NI, Ellis KA, Savage G, et al. Baseline amnestic severity predicts progression from amnestic mild cognitive impairment to Alzheimer disease dementia at 3 years. *Alzheimer Dis Assoc Disord*. 2018;32:190–196.
5. Davis M, Connell TO, Johnson S, et al. Estimating Alzheimer's disease progression rates from normal cognition through mild cognitive impairment and stages of dementia. *Curr Alzheimer Res*. 2018;15:777–788.
6. Liao H, Zhu Z, Peng Y. Potential utility of retinal imaging for Alzheimer's disease: a review. *Front Aging Neurosci*. 2018;10:188.
7. Blennow K, Zetterberg H. Biomarkers for Alzheimer's disease: current status and prospects for the future. *J Intern Med*. 2018;284:643–663.
8. Cheung CY, Ong YT, Hilal S, et al. Retinal ganglion cell analysis using high-definition optical coherence tomography in patients with mild cognitive impairment and Alzheimer's disease. *J Alzheimers Dis*. 2015;45:45–56.
9. Williams MA, McGowan AJ, Cardwell CR, et al. Retinal microvascular network attenuation in Alzheimer's disease. *Alzheimers Dement (Amst)*. 2015;1:229–235.
10. Cipollini V, Abdolrahimzadeh S, Troili F, et al. Neurocognitive assessment and retinal thickness alterations in Alzheimer disease: is there a correlation? *J Neuroophthalmol*. 2019;40:370–377.
11. Mirzania D, Thompson AC, Robbins CB, et al. Retinal and choroidal changes in men compared with women with Alzheimer's disease. *Ophthalmol Sci*. 2022;2:100098.
12. Chan VTT, Sun Z, Tang S, et al. Spectral-domain OCT measurements in Alzheimer's disease: a systematic review and meta-analysis. *Ophthalmology*. 2019;126:497–510.
13. Lad EM, Mukherjee D, Stinnett SS, et al. Evaluation of inner retinal layers as biomarkers in mild cognitive impairment to moderate Alzheimer's disease. *PLoS One*. 2018;13:e0192646.
14. Choi SH, Park SJ, Kim NR. Macular ganglion cell-inner plexiform layer thickness is associated with clinical progression in mild cognitive impairment and Alzheimers disease. *PLoS One*. 2016;11:e0162202.
15. Lu Y, Li Z, Zhang X, et al. Retinal nerve fiber layer structure abnormalities in early Alzheimer's disease: evidence in optical coherence tomography. *Neurosci Lett*. 2010;480:69–72.
16. Grewal DS, Polascik BW, Hoffmeyer GC, Fekrat S. Assessment of differences in retinal microvasculature using OCT angiography in Alzheimer's disease: a twin discordance report. *Ophthalmic Surg Lasers Imaging Retina*. 2018;49:440–444.
17. Bulut M, Kurtulus F, Gozkaya O, et al. Evaluation of optical coherence tomography angiographic findings in Alzheimer's type dementia. *Br J Ophthalmol*. 2018;102:233–237.
18. Uchida APJ, Bermel R, Bonner-Jackson A, et al. Outer retinal assessment using spectral-domain optical coherence tomography in patients with Alzheimer's and Parkinson's disease. *Vis Neurosci*. 2018;59:2768–2777.
19. Yoon SP, Grewal DS, Thompson AC, et al. Retinal microvascular and neurodegenerative changes in Alzheimer's disease and mild cognitive impairment compared with control participants. *Ophthalmol Retina*. 2019;3:489–499.
20. Rifai OM, McGrory S, Robbins CB, et al. The application of optical coherence tomography angiography in Alzheimer's disease: a systematic review. *Alzheimers Dement (Amst)*. 2021;13:e12149.
21. Yoon SP, Thompson AC, Polascik BW, et al. Correlation of OCTA and Volumetric MRI in mild cognitive impairment and Alzheimer's disease. *Ophthalmic Surg Lasers Imaging Retina*. 2019;50:709–718.
22. Robbins CB, Grewal DS, Stinnett SS, et al. Assessing the retinal microvasculature in individuals with early and late-onset Alzheimer's disease. *Ophthalmic Surg Lasers Imaging Retina*. 2021;52:336–344.
23. Coppola G, Di Renzo A, Ziccardi L, et al. Optical coherence tomography in Alzheimer's disease: a meta-analysis. *PLoS One*. 2015;10:e0134750.
24. Paquet C, Boissonnot M, Roger F, et al. Abnormal retinal thickness in patients with mild cognitive impairment and Alzheimer's disease. *Neurosci Lett*. 2007;420:97–99.
25. Kesler A, Vakhapova V, Korczyn AD, et al. Retinal thickness in patients with mild cognitive impairment and Alzheimer's disease. *Clin Neurol Neurosurg*. 2011;113:523–526.
26. Zhang YS, Zhou N, Knoll BM, et al. Parafoveal vessel loss and correlation between peripapillary vessel density and cognitive performance in amnestic mild cognitive impairment and early Alzheimer's disease on optical coherence tomography angiography. *PLoS One*. 2019;14:e0214685.
27. Almeida ALM, Pires LA, Figueiredo EA, et al. Correlation between cognitive impairment and retinal neural loss assessed by swept-source optical coherence tomography in patients with mild cognitive impairment. *Alzheimers Dement (Amst)*. 2019;11:659–669.
28. Amoroso NDD, Fanizzi A, La Rocca M, et al. For the Alzheimer's disease Neuroimaging Initiative. Deep learning reveals Alzheimer's disease onset in MCI subjects: results from an international challenge. *J Neurosci Methods*. 2018;302:3–9.
29. Basaia S, Agosta F, Wagner L, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin*. 2019;21:101645.
30. Liu M, Cheng D, Wang K, Wang Y. Alzheimer's disease neuroimaging I. Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis. *Neuroinformatics*. 2018;16:295–308.
31. Shi J, Zheng X, Li Y, Zhang Q. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. *IEEE J Biomed Health Inform*. 2018;22:173–183.
32. Gullett JM, Albizu A, Fang R, et al. Baseline neuroimaging predicts decline to dementia from amnestic mild cognitive impairment. *Front Aging Neurosci*. 2021;13:758298.
33. Tian J, Smith G, Guo H, et al. Modular machine learning for Alzheimer's disease classification from retinal vasculature. *Sci Rep*. 2021;11:238.
34. Wisely CE, Wang D, Henao R, et al. Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. *Br J Ophthalmol*. 2020;106:388–395.

35. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:263−269.
36. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: Institute of Electrical and Electronics Engineers (IEEE); 2016: 770−778.
37. Kingma DP, Ba JL. ADAM: a method for stochastic optimization. In: *International Conference on Learning Representations*. San Diego, CA: Conference paper at the 3rd International Conference for Learning Representations; 2015: 1−15.
38. Youden W. Index for rating diagnostic tests. *Cancer*. 1950;3: 32−35.
39. Robbins CB, Akrobetu D, Ma JP, et al. Assessment of retinal microvascular alterations in individuals with amnestic and nonamnestic mild cognitive impairment using optical coherence tomography angiography. *Retina*. 2022;42:1338−1346.
40. Criscuolo C, Cennamo G, Montorio D, et al. Assessment of retinal vascular network in amnestic mild cognitive impairment by optical coherence tomography angiography. *PLoS One*. 2020;15:e0233975.
41. Robbins CB, Hemesath A, Kundu A, et al. Longitudinal differences in retinal structure and microvasculature in individuals with mild cognitive impairment. In: *American Society of Retina Specialists Annual Meeting*. Seattle, WA: American Society of Retina Specialists Conferece; 2023.