HHS Public Access

Author manuscript

Mayo Clin Proc Digit Health. Author manuscript; available in PMC 2025 January 02.

Published in final edited form as:

Mayo Clin Proc Digit Health. 2024 December ; 2(4): 548–558. doi:10.1016/j.mcpdig.2024.08.005.

Color Fundus Photography and Deep Learning Applications in Alzheimer Disease

Oana M. Dumitrascu, MD, MSc,
Department of Neurology, Mayo Clinic, Scottsdale, AZ; Department of Ophthalmology, Mayo Clinic, Scottsdale, AZ

Xin Li, MS,
School of Computed and Augmented Intelligence, Arizona State University, Tempe, AZ

Wenhui Zhu, MS,
School of Computed and Augmented Intelligence, Arizona State University, Tempe, AZ

Bryan K. Woodruff, MD,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Simona Nikolova, PhD,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Jacob Sobczak,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Amal Youssef, MD,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Siddhant Saxena,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Janine Andreev,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Richard J. Caselli, MD,
Department of Neurology, Mayo Clinic, Scottsdale, AZ

Correspondence: Address to Oana M. Dumitrascu, MD, Mayo Clinic College of Medicine and Science, 13400 East Shea Boulevard, Scottsdale, AZ 85259 (dumitrascu.oana@mayo.edu; Twitter: @OanaDumitrascu5).

POTENTIAL COMPETING INTERESTS

Dr Dumitrascu and Wang report pending patents on D24-186 A BERT-Style Self-Supervised Learning CNN for Disease Identification from Retinal Images, D24-170 Context-Aware Optimal Transport Learning for Retinal Color Fundus Photograph Enhancement, and D23-176 Systems and methods for enhancing retinal color fundus images for retinopathy analysis. Dr Dumitrascu is a board member of AHA Greater Phoenix Chapter. The authors report no competing interests.

SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at https://www.mcpdigitalhealth.org/. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

ETHICS STATEMENT

The study received exempt approval from an institutional review board, with waiver of informed consent owing to exclusive use of retrospective and deidentified data.

Data Previously Presented: A subaim of this work was presented at the American Academy of Neurology Annual Meeting, April 2024, Denver, CO.

John J. Chen, MD, PhD,
Department of Ophthalmology, Mayo Clinic Rochester, MN; Department of Neurology, Mayo Clinic Rochester, MN

Yalin Wang, PhD
School of Computed and Augmented Intelligence, Arizona State University, Tempe, AZ

Abstract

Objective: To report the development and performance of 2 distinct deep learning models trained exclusively on retinal color fundus photographs to classify Alzheimer disease (AD).

Patients and Methods: Two independent datasets (UK Biobank and our tertiary academic institution) of good-quality retinal photographs derived from patients with AD and controls were used to build 2 deep learning models, between April 1, 2021, and January 30, 2024. ADVAS is a U-Net–based architecture that uses retinal vessel segmentation. ADRET is a bidirectional encoder representations from transformers style self-supervised learning convolutional neural network pretrained on a large data set of retinal color photographs from UK Biobank. The models' performance to distinguish AD from non-AD was determined using mean accuracy, sensitivity, specificity, and receiving operating curves. The generated attention heatmaps were analyzed for distinctive features.

Results: The self-supervised ADRET model had superior accuracy when compared with ADVAS, in both UK Biobank (98.27% vs 77.20%; $P<.001$) and our institutional testing data sets (98.90% vs 94.17%; $P=.04$). No major differences were noted between the original and binary vessel segmentation and between both eyes vs single-eye models. Attention heatmaps obtained from patients with AD highlighted regions surrounding small vascular branches as areas of highest relevance to the model decision making.

Conclusion: A bidirectional encoder representations from transformers style self-supervised convolutional neural network pretrained on a large data set of retinal color photographs alone can screen symptomatic AD with high accuracy, better than U-Net–pretrained models. To be translated in clinical practice, this methodology requires further validation in larger and diverse populations and integrated techniques to harmonize fundus photographs and attenuate the imaging-associated noise.

Without new advancements in Alzheimer disease (AD) care, the cost associated with AD and other dementias' care is projected to reach nearly $1 trillion in 2050.[1] Accurate and widely accessible AD screening tools are crucial in this emerging global health crisis.[2] Current AD biomarkers' availability is limited by their prohibitive cost, invasiveness, or need for additional validation.[3,4] The retina is an unshielded extension of the brain that offers the opportunity to investigate multiple central nervous system disorders.[5-7] Postmortem histopathologic investigations of eyes and brains from patients with AD[8,9] and clinical evaluations of patients with AD[10,11] have revealed pathologic alterations in the neuro-sensory retina that precede and correlate with brain AD changes.[5,12] Retina accumulates amyloid β-protein plaques and abnormal tau proteins[13-15] and exhibits neurodegeneration, vascular amyloidosis, and increased inflammation.[5,10-12,16,17] Retinal changes can be captured via color fundus photography, a widely accessible technology

in eye care, primary care, and underresourced community settings, carrying the promise of a noninvasive and cost-effective biomarker for AD.[17-20] Machine learning tools were developed to overcome the subjectivity and low efficiency associated with manual retinal photographs analysis for disease and biomarker identification and to automate their interpretation in multiple ocular disorders.[21-23] Moreover, machine learning applications using retinal imaging alone demonstrate good prediction accuracy of the individuals' biological sex, blood pressure, and smoking status.[24,25] Given the retinal changes in Alzheimer's mirrors the Alzheimer's brain pathology,[5,12,25,26] deep learning applications have been expanded[27-31] to enable the automatic classification of AD and identification of retinal AD biomarkers that are not visible to the human eye.[29,30,32] Although promising, we noted certain limitations of the traditional convolutional neural networks (CNNs) that rely on large numbers of images and expert labeling. Modern technologies, such as vision transformers and self-supervised learning, provide a pretraining strategy that uses easier attainable unlabeled data, overcoming the challenge of label acquisition and expanding the span of images utilization. These are based on the advancement of natural language processing, in the form of self-supervised learning models of bidirectional encoder representations from transformers (BERT)[33] and generative pretrained transformers.[34] Here, we evaluate the value of both traditional CNNs and the modern generative self-supervised learning. We report the technological development and clinical application of ADVAS, a U-Net–based architecture employing retinal vessel segmentation[35] and ADRET, a BERT-style self-supervised learning CNN pretrained on retinal color photographs from UK Biobank.[36] We determined the 2 models' performance to classify AD in 2 real-world patient populations and attempted to identify the retinal regions that were highlighted as AD discriminators by the models' decision making.

PATIENTS AND METHODS

Participants and Retinal Color Photograph Acquisition

UK Biobank Data Set.—We used 178,803 unlabeled color fundus images from the UK Biobank (http://www.ukbiobank.ac.uk/about-biobank-uk)[36] involving 87,245 participants, including patients with AD (1136 images, 553 patients) and without AD (176,392 images, 86,069 patients). During the ADRET model pretraining on these images, we first randomly masked 60% of the input image, where the size of each mask is equal to the (mask's height [H] adjusted by the downsampling ratio [D], mask's width [W] adjusted by the downsampling ratio [D] of the network) (HD, WD). Then, we processed only the unmasked visible regions of the input image. The size of the output image was (224, 224), and the downsampling ratio was 32, making our mask size (7, 7). We used the feature maps learned from the pretraining module for classification studies. For quality control, 2 trained graders (A.Y. and S.S.) excluded the retinal images exhibiting blur, low contrast, poor illumination, or artifacts (Supplemental Figure 1, available online at https://www.mcpdigitalhealth.org/). After the quality control, we selected 362 good-quality images (169 left eyes and 193 right eyes) from 230 patients with AD. The reference group included 389 good-quality images (170 left eyes and 219 right eyes) from 282 patients without AD. The AD label was based on International Classification of Diseases (ICD) codes from hospital admission and death records, indicating a definitive clinical diagnosis of dementia caused by AD (Data-field

42021).[37] The non-AD label was based on absent neurodegenerative conditions and other dementias. We excluded patients with Parkinson disease (G20), secondary parkinsonism (G21), other parkinsonism (G22), other degenerative disorders of nervous system (G32), vascular syndromes in cerebrovascular disease (G46), vascular dementia (F01), dementia associated with other diseases (F02), unspecified dementia (F03), and organic amnestic syndrome (F04). We also excluded patients with ocular conditions, including age-related macular degeneration, glaucoma, and diabetic retinopathy. For each input image, we first detected the retina using the Hough circle transform and then cropped the mask region to minimize the effect of the black background. Afterward, the images were resized to 224 × 224 and normalized to (−1, 1).[38]

Institutional Data Set.—The study received exempt approval from our institutional review board, with waiver of informed consent owing to exclusive use of retrospective and deidentified data (Protocol #21-013272). We collected 45° digital color fundus photographs from 118 patients with a clinical diagnosis of AD from our academic multisite tertiary institution. The patients were retrospectively identified through an electronic medical record search using ICD9 (331) and ICD10 (G30) codes from January 1, 2011, to June 30, 2021; 129 patients from the EyePACS database created the control data set.[39] After image quality control, the ADVAS model included 318 binary vessel segmentation images from 113 patients with AD and 338 original vessel segmentation images from 116 patients with AD. The control group comprised 259 vessel segmentation images (original and binary) from 129 participants. ADRET used 283 images from 76 patients with AD. We detected the retina using the Hough circle transform and cropped the mask region to minimize the effect of the black background. The images were then resized to 512 × 512 and normalized to (−1, 1).[38] This resolution was chosen to increase the attention to detail in this small data set because the requirements for the number of images in a single batch and the computational efficiency were lower.[40,41]

Deep Learning Models Training and Testing (April 1, 2021, Through January 30, 2024) ADVAS Model.—The framework consisted of 2 main steps. In the first step, we used a U-Net–based architecture[42] with 2 parts, encoder (downsampling) and decoder (upsampling). We first segmented the retinal vasculature in the retinal images from the UK Biobank and our institutional data set and then inputted the segmented vessel results into the U-Net encoder for feature extraction. For vessel segmentation, we used the Digital Retinal Images for Vessel Extraction database[43] to train the model and obtain its optimized weight parameters. Two outputs were generated: the original vessel segmentation image, with detailed vessel structure (model 1); and binary vessel segmentation (denoted either 0 or 1), which enhanced the clear parts of the vessel and ignored the faint parts, facilitating further analyses (model 2). Subsequently, these segmentation results were used as inputs for further extraction of vascular features using U-Net encoders,[44] which performed initialization using weight from the first segmentation stage. This process focused on extracting key information from the segmented images that could help with the disease diagnosis. Finally, the extracted features were fed to a new linear classifier, a fully connected layer, and a Softmax function (pipeline illustrated in Figure 1). A fully connected layer is a neural network layer in which every neuron is connected to all activation units from the preceding layer. This

layer typically resides at the network's end and maps the learned nonlinear features to the sample's output space. The Softmax function is an activation function widely used for multiclass classification problems.[45] It transforms a real-value vector into a probability distribution where each element's value is between 0 and 1, and the sum of all elements equals 1. The output from the Softmax function can be interpreted as the probability distribution over various classes, representing the likelihood that the sample belongs to each class. The classifier predicted whether the individual represented by the image is an AD or control patient. The heatmaps were generated in the last layer of the U-Net classifier using Gradient-weighted Class Activation Mapping (GRAD-CAM).[46,47]

ADRET Model.—The backbone was our recently developed nn-MobileNet[48] that reported improved network performance by the following: (1) adjusting the order of channel configurations for the inverted linear residual bottleneck in the MobileNetV2 network[49]; (2) using a heavy data augmentation strategy through Mixup,[50] CutMix[51] image cropping, flipping, contrast adjustment, brightness adjustment, and sharpening; and (3) adding spatial-dropout[52] modules at various locations within the network to identify their optimal placement in an attempt to address the over-fitting. We adopted a BERT-style self-supervised learning[53] method (Figure 2), masking the image and then reducing it to pretrain the encoder to obtain the representative features.[54] To improve the computational efficiency, we resized the input image to 224 × 224. This image resolution was chosen given ADRET was pretrained on many fundus images, to ensure computational efficiency and the number of samples processed in each single batch, while guaranteeing that sufficient features can be acquired. Previous reports also reported a successful use of the 224 × 224 resolution to balance the computational performance with the number of features.[55-58] Next, we performed random masking with a masking rate of 60%. We adopted the hierarchical design principle of the Spark framework[59] and combined it with our nn-MobileNet[48] to generate feature maps with different resolutions. Then, the feature maps used sparse convolution and a lightweight U-Net[42] decoder to perform upsampling, which serves as an image reconstruction self-supervised learning task. We set up 1600 epochs for pretraining. All experiments were conducted on a single 4 NVIDIA A100 80GB GPUs with an AMD EPYC 7413 24-Core Processor.

Statistical Analyses—We randomly divided the data into training and validation sets in the ratio of 8:2. We used 5-fold stratified cross-validation on the data set to evaluate the performance and generalization ability of our models. The data set was divided equally into 5 subsets, 4 of which were used at a time for training and the remaining one for testing. The process was repeated 5 times, with each subset used once each as a testing set. The performance metrics of the 5 tests were averaged to obtain an overall performance evaluation of the model. The 5-fold cross-validation uses all data points for training and validation, thus providing a more robust estimate of the model performance. Models were evaluated based on quadratic-weighted κ, area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, and accuracy. AUROC quantifies the overall ability of the model to discriminate between positive and negative classes. Additionally, for ADVAS, we explored models using all available images and images derived from 1 single eye. We separately trained and tested all groups and obtained the model performance on the

testing set. We compared the performance metrics of both eyes vs single-eye models and original vs binary vessel segmentation models, using the Pearson χ^2 test. A P value of <.05 was considered statistically significant.

Codes Availability

Our study's implementation was based on the Python (version 3.10) and PyTorch (version 2.0) environments. In the image preprocessing stage, we used OpenCV (version 4.6). In the data visualization stage, our study was implemented by Grad-CAM (version 1.5, https://github.com/jacobgil/pytorch-grad-cam) and Matplotlib (version 3.8).

RESULTS

ADVAS Accuracy for AD Classification

In the UK Biobank, the overall original vessel segmentation model (362 AD and 389 non-AD images) had 77.2% accuracy, which was similar with the binary vessel segmentation model (73.73% accuracy; P=.57) (Table 1, Figure 3). Left eye images (169 AD and 170 non-AD) had 86.7% accuracy in original vessel and 79.41% accuracy in binary vessel segmentation. Right eye images (193 AD and 219 non-AD) had 72.29% accuracy in original vessel and 69.88% accuracy in binary vessel segmentation. In our institutional data set, the overall original vessel segmentation group (120 test, 476 train images) had nonsignificantly greater accuracy than the binary vessel segmentation model (116 test, 460 train images; 94.17% vs 90.17%; P=.29). In single-eye models, the maximum (97.73%) accuracy was reached by the original vessel segmentation of the right eye, followed by the binary vessel segmentation in the left eye (Table 1). Nonsignificant differences were noted between both eyes and single-eye (right or left) models (P>.05 in all models).

ADRET Accuracy for AD Classification

In the UK Biobank data set, the model achieved 98.22% accuracy, a κ score of 0.9652, and an AUROC of 0.9967 for AD classification. In our institutional data set, the model achieved 98.90% accuracy, a κ score of 0.9777, and an AUROC of 0.9978 for AD classification (Table 2, Figure 3).

All models generated attention heatmaps using GRAD-CAM. The AD-derived attention maps highlighted areas surrounding retinal small vascular branches as regions of highest relevance to the model decision making (Supplemental Figure 2, available online at https://www.mcpdigitalhealth.org/).

DISCUSSION

We report 2 different CNNs trained exclusively on 45° retinal color photographs, which reached good performance to discriminate AD from non-AD in 2 different populations. We trained and tested these automated models to overcome the subjectivity and inefficiency of retinal photographs manual analysis. First, we aimed to further investigate the potential of a formerly reported AD classification tool based on retinal vessel segmentation.[30] In this scope, we developed ADVAS, a novel U-Net–based CNN using original retinal vessel

segmentation. We also explored binary vessel segmentation, which pays more attention to clearly delineated vessels and less attention to the faint vessels. All ADVAS models achieved a classifying accuracy of at least 69.8%, without notable differences between both eyes vs single-eye models. The highest accuracy (97%) was reached when images from right eye and original vessel segmentation, and left eye and binary vessel segmentation, were used, suggesting that any of them may be chosen for optimization in future studies. The fact that multiple distinctive groups had similarly high performance reinforces our findings accuracy and applicability. Our framework identified regions surrounding smaller retinal vessels in AD-derived heatmaps, which was similarly highlighted in AD imaging and pathologic studies.[10,15,16,60-62] These hypothesis-generating findings will require confirmation in future machine learning studies. ADVAS reported different performance metrics in the 2 data sets. Previous literature similarly reports considerable differences between UK Biobank and other data sets (such as EyePACS) when performing the same task.[63,64] This effect is explained by several factors impacting diverse data sets, such as different sampling equipment, sample diversity, data quality, or condition severity and is particularly pronounced in large-scale data sets.

Our novel BERT-style self-supervised ADRET model had superior accuracy when compared with ADVAS, in both UK Biobank (98.27% vs 77.20%; $P<.001$) and our institutional testing data sets (98.90% vs 94.17%; $P=.04$). This underscores the increased precision of modern self-supervised techniques for disease classification. ADRET was implemented through a lightweight CNN architecture, nn-MobileNet.[48] This method fully leverages the advantages of CNNs in medical image processing and incorporates the pre-training mechanism of self-supervised learning, aiming to efficiently process large-scale unlabeled medical image data sets. The performance of the model in data representation and feature extraction was remarkably enhanced by employing a sparse convolution technique to process the masked regions of the image, while preserving the original CNN hierarchical structure. To validate the effectiveness of the proposed method, we conducted pretraining on the UK Biobank data set. Subsequently, we applied the pretrained model to the AD classification task in both UK Biobank and our institutional data set, and ADRET achieved outstanding performance on these downstream tasks. This demonstrates the effectiveness of self-supervised pretraining in improving the model's understanding of retinal color photographs and provides support for applying these novel techniques to other areas of medical image analysis.

Several machine learning algorithms and proof-of-concept studies had found that retinal photographs alone could estimate the risk of AD.[29,30] An algorithm developed using the UK Biobank retinal images, U-Net for vessel segmentation, and a support vector machine-based classifier reported a binary accuracy of 82.4% for AD classification after matching for age.[30] Another study that used EfficientNET,[65] an unsupervised domain adaptation network, and retinal color photographs from multiethnic cohorts reported an accuracy ranging from 79.6%-92.1% to differentiate AD dementia from non-AD.[29] A machine learning study using optical coherence tomography (OCT) images had identified the XGBoost algorithm with the best diagnostic performance for AD (0.74 accuracy), and macular thickness with the greatest importance for guiding the algorithm to the AD diagnosis.[66] Another study that used multimodal retinal imaging (OCT, OCT angiography, fundus photography, autofluorescence, and metadata) to identify symptomatic AD[67] reported AUROC of 0.836

(CI, 0.729-0.943) using all inputs. As technology for automated imaging analysis had evolved, we explored innovative techniques and our BERT-style self-supervised model reported superior performance, with a sensitivity of 98.9% and specificity of 98.05% to discriminate AD from non-AD in a real-world institutional data set. This is in line with recent studies reporting self-supervised models' superiority over supervised learning-based transfer learning, even when the self-supervised models were built using limited data sets[68,69] and when tested on new data.[68,70]

There are some limitations of our study. Few patients had amyloid-positron emission tomography or cerebrospinal fluid biomarker confirmation of their cerebral amyloid status; however, both data sets used the clinical expert established diagnosis of symptomatic AD, which included neuropsychometric testing and fluorodeoxyglucose-positron emission tomography. Because the control EyePACS data set did not report the patients' ages, we cannot exclude the impact of age on the models' performance. However, UK Biobank AD and non-AD groups had similar mean ages, implying that the classification accuracy was not confounded by age.

The use of 45° images may have deprived the model of relevant biomarker from the peripheral retina; however, this model reached an excellent accuracy, and 45° images are easier obtainable in nonophthalmic care settings. Nonmydriatic retinal photography is susceptible to various sources of noise leading to suboptimal image quality. We excluded a considerable number of retinal photographs after quality control because suboptimal image quality can hinder the CNN model development.[30,32] The lack of decent quality retinal photographs obtained in the real-world constitute an obstacle in the implementation and validation of these CNN models in routine practice. Our generative image enhancement models tested on various retinopathies[35,71-73] have shown potential to increase the adaptability and resilience of retinal images across diverse distributions, making them suitable for clinical settings.

These results suggest that deep learning–assisted color fundus photography analysis could become an accurate AD risk stratification tool in optometry, community eye care, or established screening programs for other retinopathies (eg, diabetic retinopathy). Future studies should assess the performance of ADRET in large cohorts with biomarker-confirmed AD and evaluate its cost-effectiveness in comparison with current AD screening methods, especially in communities with limited infrastructure and access to specialized care. Incoming studies should also investigate deep learning–assisted retinal imaging tools in AD clinical trials.

CONCLUSION

A BERT-style self-supervised neural network pretrained on a large data set of retinal color photographs alone can screen symptomatic AD with high accuracy, better than U-Net–pretrained models. With further validation in diverse populations, as well as integrated methods to attenuate imaging-associated noise, this methodology has the potential for application in clinical practice for point-of-care AD screening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Dr Dumitrascu and Mr Li contributed equally to this work.

Grant Support:

This work has been partially supported by the NIH (R01EY032125, R01DE030286, R01AG069453, and P30AG072980) and the state of Arizona via the Arizona Alzheimer's Consortium.

Abbreviations and Acronyms:

AD	Alzheimer disease
CNN	convolutional neural network
BERT	bidirectional encoder representations from transformers
ICD	International Classification of Diseases
GRAD-CAM	Gradient-weighted Class Activation Mapping

REFERENCES

1. Alzheimer's association: 2024 Alzheimer's disease facts and figures report: executive summary. https://alz.org/media/Documents/Facts-And-Figures-2024-Executive-Summary.pdf. Accessed June 20, 2024.
2. Goldman DP, Fillit H, Neumann P. Accelerating Alzheimer's disease drug innovations from the research pipeline to patients. Alzheimers Dement. 2018;14(6):833–836. 10.1016/j.jalz.2018.02.007. [PubMed: 29680407]
3. Fatima H, Rangwala HS, Riaz F, Rangwala BS, Siddiq MA. Break-throughs in Alzheimer's research: a path to a more promising future? Ann Neurosci. 2024;31(1):63–70. 10.1177/09727531231187235. [PubMed: 38584978]
4. Garcia MJ, Leadley R, Ross J, et al. Prognostic and predictive factors in early Alzheimer's disease: a systematic review. J Alzheimers Dis Rep. 2024;8(1):203–240. 10.3233/ADR-230045. [PubMed: 38405341]
5. Koronyo Y, Rentsendorj A, Mirzaei N, et al. Retinal pathological features and proteome signatures of Alzheimer's disease. Acta Neuropathol. 2023;145(4):409–438. 10.1007/s00401-023-02548-2. [PubMed: 36773106]
6. Hinton DR, Sadun AA, Blanks JC, Miller CA. Optic-nerve degeneration in Alzheimer's disease. N Engl J Med. 1986;315(8):485–487. 10.1056/NEJM198608213150804. [PubMed: 3736630]
7. Blanks JC, Torigoe Y, Hinton DR, Blanks RH. Retinal degeneration in the macula of patients with Alzheimer's disease. Ann N Y Acad Sci. 1991;640:44–46. 10.1111/j.1749-6632.1991.tb00188.x. [PubMed: 1776758]
8. La Morgia C, Ross-Cisneros FN, Koronyo Y, et al. Melanopsin retinal ganglion cell loss in Alzheimer disease. Ann Neurol. 2016;79(1):90–109. 10.1002/ana.24548. [PubMed: 26505992]
9. Lee CS, Larson EB, Gibbons LE, et al. Associations between recent and established ophthalmic conditions and risk of Alzheimer's disease. Alzheimers Dement. 20l9;15(1):34–41. 10.1016/j.jalz.2018.06.2856. [PubMed: 30098888]
10. Dumitrascu OM, Rosenberry R, Sherman DS, et al. Retinal venular tortuosity jointly with retinal amyloid burden correlates with verbal memory loss: a pilot study. Cells. 2021;10(11):2926. 10.3390/cells10112926. [PubMed: 34831149]

11. Hart NJ, Koronyo Y, Black KL, Koronyo-Hamaoui M. Ocular indicators of Alzheimer's: exploring disease in the retina. Acta Neuropathol. 2016;132(6)767–787. 10.1007/s00401-016-1613-6. [PubMed: 27645291]
12. Mirzaei N, Shi H, Oviatt M, et al. Alzheimer's retinopathy: seeing disease in the eyes. Front Neurosci. 2020;14:921. 10.3389/fnins.2020.00921. [PubMed: 33041751]
13. Dumitrascu OM, Doustar J, Fuchs DT, et al. Retinal peri-arteriolar versus peri-venular amyloidosis, hippocampal atrophy, and cognitive impairment: exploratory trial. Acta Neuropathol Commun. 2024;12(1):109. 10.1186/s40478-024-01810-2. [PubMed: 38943220]
14. Gaire BP, Koronyo Y, Fuchs DT, et al. Alzheimer's disease pathophysiology in the retina. Prog Retin Eye Res. 2024;101:101273. 10.1016/j.preteyeres.2024.101273. [PubMed: 38759947]
15. Dumitrascu OM, Doustar J, Fuchs DT, et al. Distinctive retinal peri-arteriolar versus peri-venular amyloid plaque distribution correlates with the cognitive performance. Preprint. bioRxiv. Published online February 29, 2024. 10.1101/2024.02.27.580733.
16. Shi H, Koronyo Y, Rentsendorj A, et al. Identification of early pericyte loss and vascular amyloidosis in Alzheimer's disease retina. Acta Neuropathol. 2020;139(5):813–836. 10.1007/s00401-020-02134-w. [PubMed: 32043162]
17. Jiang H, Wang J, Levin BE, et al. Retinal microvascular alterations as the biomarkers for Alzheimer disease: are we there yet? J Neuroophthalmol. 2021;41(2):251–260. 10.1097/WNO.0000000000001140. [PubMed: 33136677]
18. Dumitrascu OM, Qureshi TA. Retinal vascular imaging in vascular cognitive impairment: current and future perspectives. J Exp Neurosci. 2018;12:1179069518801291. 10.1177/1179069518801291. [PubMed: 30262988]
19. Sasaki M. [Retinal imaging as potential biomarkers for dementia]. Brain Nerve. 2021;73(11):1209–1216. 10.11477/mf.1416201919. [PubMed: 34759057]
20. Cheung CY, Chan VTT, Mok VC, Chen C, Wong TY. Potential retinal biomarkers for dementia: what is new? Curr Opin Neurol. 2019;32(1):82–91. 10.10097/WCO.0000000000000645. [PubMed: 30566412]
21. Li T, Bo W, Hu C, et al. Applications of deep learning in fundus images: a review. Med Image Anal. 2021;69:101971. 10.1016/j.media.2021.101971. [PubMed: 33524824]
22. Dai L, Wu L, Li H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun. 2021;12(1):3242. 10.1038/s41467-021-23458-5. [PubMed: 34050158]
23. Nadeem MW, Goh HG, Hussain M, Liew SY, Andonovic I, Khan MA. Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions. Sensors (Basel). 2022;22(18):6780. 10.3390/s22186780. [PubMed: 36146130]
24. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng. 2018;2(3):158–164. 10.1038/s41551-018-0195-0. [PubMed: 31015713]
25. Wagner SK, Fu DJ, Faes L, et al. Insights into systemic disease through retinal imaging-based oculomics. Transl Vis Sci Technol. 2020;9(2):6. 10.1167/tvst.9.2.6.
26. Snyder PJ, Alber J, Alt C, et al. Retinal imaging in Alzheimer's and neurodegenerative diseases. Alzheimers Dement. 2021;17(1):103–111. 10.1002/alz.12179. [PubMed: 33090722]
27. Ng WY, Cheung CY, Milea D, Ting DSW. Artificial intelligence and machine learning for Alzheimer's disease: let's not forget about the retina. Br J Ophthalmol. 2021;105(5):593–594. 10.1136/bjophthalmol-2020-318407. [PubMed: 33495160]
28. Bahr T, Vu TA, Tuttle JJ, lezzi R. Deep learning and machine learning algorithms for retinal image analysis in neurodegenerative disease: systematic review of datasets and models. Transl Vis Sci Technol. 2024;13(2):16. 10.1167/tvst.13.2.16.
29. Cheung CY, Ran AR, Wang S, et al. A deep learning model for detection of Alzheimer's disease based on retinal photographs: a retrospective, multicentre case-control study. Lancet Digit Health. 2022;4(11):e806–e815. 10.1016/S2589-7500(22)00169-8. [PubMed: 36192349]
30. Tian J, Smith G, Guo H, et al. Modular machine learning for Alzheimer's disease classification from retinal vasculature. Sci Rep. 2021;11(1):238. 10.1038/s41598-020-80312-2. [PubMed: 33420208]

31. Corbin D, Lesage F. Assessment of the predictive potential of cognitive scores from retinal images and retinal fundus metadata via deep learning using the CLSA database. Sci Rep. 2022;12(1):5767. 10.1038/s41598-022-09719-3. [PubMed: 35388080]
32. Dumitrascu OM, Zhu W, Qiu P, Wang YL. Automated retinal imaging analysis for Alzheimer's disease screening. Ann Neurol. 2022:S106.
33. Devlin J, Chang MW, Lee K, Toutanova K. Bert: pre-training of deep bidirectional transformers for language understanding. *Preprint.* Posted online October 11, 2018. arXiv:1810.04805.
34. Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. Adv Neural Inform Process Syst. 2022;33:1877–1901.
35. Zhu WQP, Lepore N, Dumitrascu OM, Wang Y. NNMobile-Net: rethinking CNN design for deep learning-based retinopathy research. *Preprint.* Published online June 2, 2023. arXiv:2306.01289.
36. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–209. 10.1038/s41586-018-0579-z. [PubMed: 30305743]
37. Bush K, Wilkinson T, Schnier C, Nolan J, Sudlow C. Definitions of Dementia and the Major Diagnostic Pathologies. UK Biobank Phase I Outcomes Adjudication; 2018.
38. Fu H, Wang B, Shen J, et al. Evaluation of Retinal Image Quality Assessment Networks in Different Color-Spaces. Springer International Publishing; 2019.
39. Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. J Diabetes Sci Technol. May 1 2009;3(3):509–516. 10.1177/193229680900300315. [PubMed: 20144289]
40. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol. 2017;135(11):1170–1176. 10.1001/jamaophthalmol.2017.3782. [PubMed: 28973096]
41. 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. 10.1016/j.cell.2018.02.010. [PubMed: 29474911]
42. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015. Springer;2015:234–241.
43. DRIVE Digital Retinal Images for Vessel Extraction. Kaggle; 2020.
44. Zhuang J. LadderNet: Multi-path networks based on U-Net for medical image segmentation. *Preprint.* Posted online October 17, 2018. arXiv:1810.07810.
45. Bishop C. Pattern Recognition and Machine Learning. Springer;2006. ISBN 0-387-31073-8.
46. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-Cam: visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017:618–626.
47. Pytorch library for cam methods. 2021: https://github.com/jacobgil/pytorch-grad-cam. Accessed September 10, 2024.
48. Zhu WQP, Chen X, Li X, et al. *nn*MobileNet rethinking CNN for retinopathy research. Presented at: Data Curation and Augmentation in Enhancing Medical Imaging Applications Workshop (DCAMI). Seattle, WA: CVPR; June 2024. https://arxiv.org/abs/2306.01289. Accessed September 10, 2024.
49. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. Mobile-netv2: Inverted Residuals and Linear Bottlenecks. IEEE. 2018:4510–4520.
50. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D. mixup: beyond empirical risk minimization. *Preprint.* Posted online October 25, 2017. arXiv:1710.09412.
51. Yun S, Han D, Oh SJ, Chun SJ, Junsuk Choe J, Yoo Y. CutMix: regularization strategy to train strong classifiers with localizable features. Presented at 2019 IEEE International Conference on Computer Vision, November 2019, Seoul, Korea (South), https://arxiv.org/abs/1905.04899. Accessed September 10, 2024.

52. Tompson J, Goroshin R, Jain A, LeCun Y, Bregler C. Efficient Object Localization Using Convolutional Networks. Boston, MA: IEEE Conference on Computer Vision and Pattern Recogntion; 2015.
53. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. North American Chapter of the Association for Computational Linguistics; 2019. https://arxiv.org/abs/1810.04805. Accessed September 10, 2024.
54. He K, Chen X, Xie S, Li Y, Dollar P, Girshick R. Masked autoencoders are scalable vision learners. 2022 IEEE/Conference on Computer Vision and Pattern Recognition, New Orleans, June 2022, https://arxiv.org/abs/2111.06377v3. Accessed September 10, 2024.
55. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. Banff, Canada: International Conference on Learning Representations; 2014. 10.48550/arXiv.1409.1556.
56. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. In: 1:1097–1105.
57. 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. IEEE; 2016:770–778.
58. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE; 2015:1–9.
59. Tian K, Jiang Y, Diao Q, et al. Designing bert for convolutional networks: sparse and hierarchical masked modeling. Kigali, Rwanda: International Conference on Learning Representations; 2023. arXiv:2301.03580v2.
60. Cheung CY, Ong YT, Ikram MK, et al. Microvascular network alterations in the retina of patients with Alzheimer's disease. Alzheimers Dement. 2014;10(2):135–142. 10.1016/j.jalz.2013.06.009. [PubMed: 24439169]
61. Williams MA, McGowan AJ, Cardwell CR, et al. Retinal microvascular network attenuation in Alzheimer's disease. Alzheimers Dement (Amst). 2015;1(2):229–235. 10.1016/j.dadm.2015.04.001. [PubMed: 26634224]
62. Jiang H, Wei Y, Shi Y, et al. Altered macular microvasculature in mild cognitive impairment and Alzheimer disease. J Neuroophthalmol. 2018;38(3):292–298. 10.1097/WNO.0000000000000580. [PubMed: 29040211]
63. Vaghefi E, Squirrell D, Yang S, et al. Development and validation of a deep-learning model to predict 10-year atherosclerotic cardiovascular disease risk from retinal images using the UK Biobank and EyePACS 10K datasets. Cardiovasc Digit Health J. 2024;5(2):59–69. 10.1016/j.cvdhj.2023.12.004. [PubMed: 38765618]
64. Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. Nature. 2023;622(7981):156–163. 10.1038/s41586-023-06555-x. [PubMed: 37704728]
65. Tan MLQ. EfficientNet: rethinking model scaling for convolutional neural networks. Proc Mach Learning. 2019;97:6105–6114.
66. Wang X, Jiao B, Liu H, et al. Machine learning based on optical coherence tomography images as a diagnostic tool for Alzheimer's disease. CNS Neurosci Ther. 2022;28(12):2206–2217. 10.1111/cns.13963. [PubMed: 36089740]
67. Wisely CE, Wang D, Henao R, et al. Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. Br J Ophthalmol. 2022;106(3):388–395. 10.1136/bjophthalmol-2020-317659. [PubMed: 33243829]
68. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: Daume H, Singh A, eds. Proceedings of the 37th International Conference on Machine Learning. 2020;149:1597–1607.
69. Chen K, Ghisays V, Luo J, et al. Improved comparability between measurements of mean cortical amyloid plaque burden derived from different PET tracers using multiple regions-of-interest and machine learning. Paper presented at 2021 Alzheimer's Association International Conference, Denver, CO.. https://alz.confex.com/alz/2021/meetingapp.cgi/Paper/51419.

70. Kea H. Masked autoencoders are scalable vision learners. In: Dana K, ed. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE; 2022:16000–16009.

71. Zhu W, Qiu P, Chen X, et al. Beyond MobileNet An Improved MobileNet for Retinal Diseases. Springer; 2023:56–65.

72. Zhu W, Qiu P, Dumitrascu OM, et al. OTRE: where optimal transport guided unpaired image-to-image translation meets regularization by enhancing. Inf Process Med Imaging. 2023;13939:4l5–427. 10.1007/978-3-031-34048-2_32.

73. Zhu W, Qiu P, Farazi M, Nandakumar K, Dumitrascu OM, Wang Y. Optimal transport guided unsupervised learning for enhancing low-quality retinal images. Proc IEEE Int Symp Biomed Imaging. 2023;2023. 10.1109/isbi53787.2023.10230719.

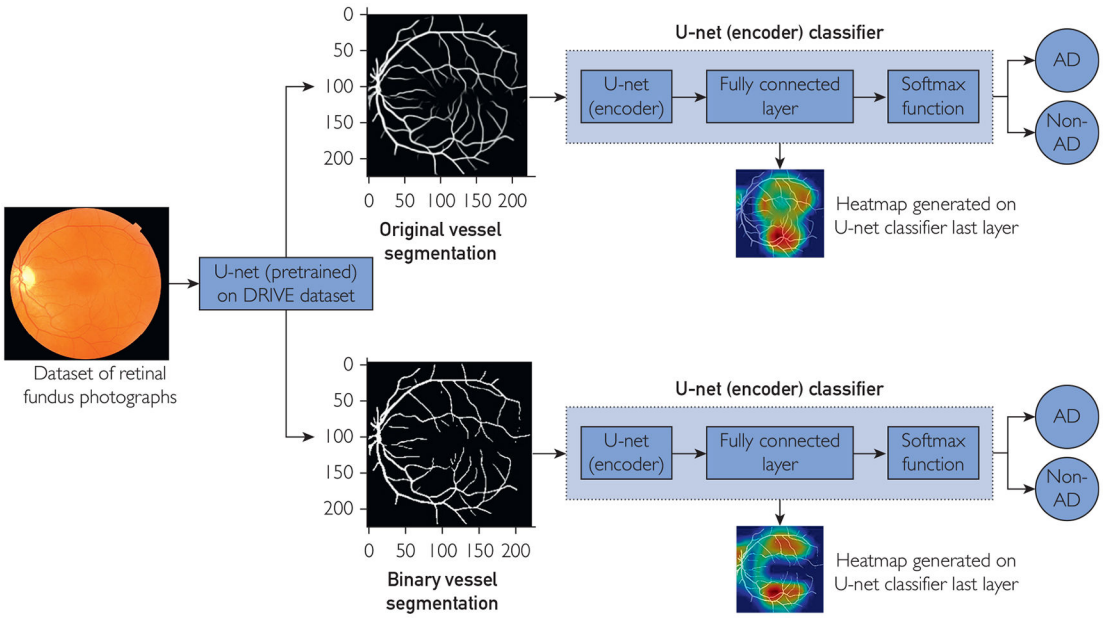

FIGURE 1.
Pipeline for the ADVAS model.

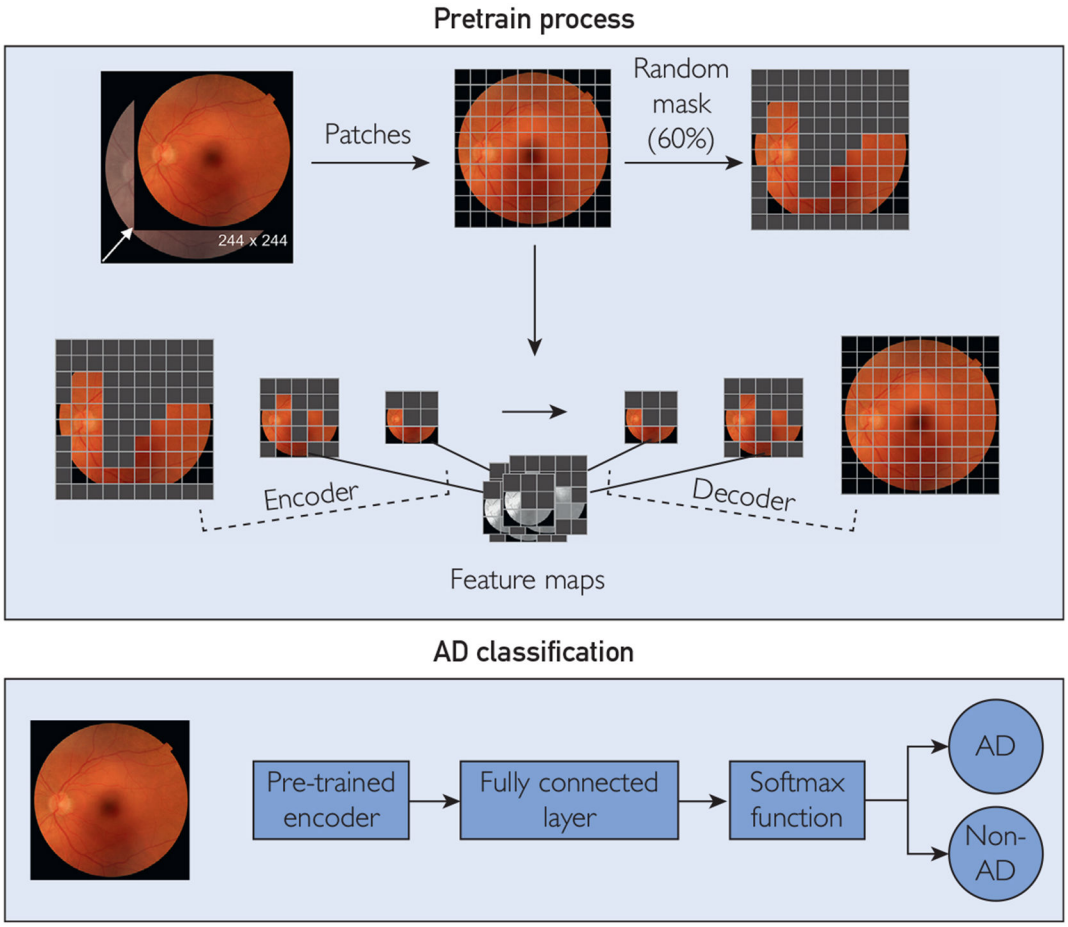

FIGURE 2.
Pipeline for the ADRET model.

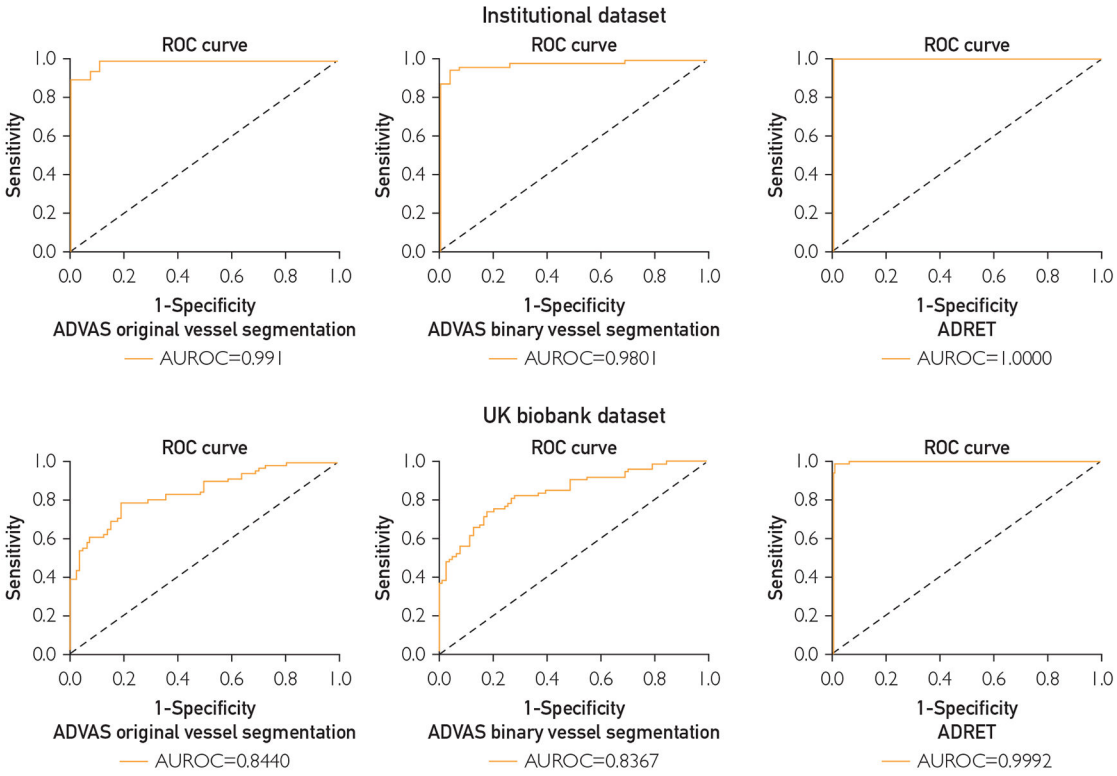

FIGURE 3.

Graphs illustrating the performance of the ADVAS and ADRET models in both data sets and the corresponding receiver operating characteristic (ROC) curves.

TABLE 1.

ADVAS Models (Original and Binary Vessel Segmentation) Performance for Alzheimer Disease Prediction Using Images Derived From Both Eyes vs Single Eye

UK Biobank data set					Institutional data set				
Original vessel segmentation		**Binary vessel segmentation**		**Original vs binary**	**Original vessel segmentation**		**Binary vessel segmentation**		**Original vs binary**
Both eyes images		Both eyes images		*P*=.57	Both eyes images		Both eyes images		*P*=.29
601 training		601 training			476 training		460 training		
150 testing		150 testing			120 testing		116 testing		
Accuracy = 0.772		Accuracy = 0.7373			Accuracy = 0.9417		Accuracy = 0.9017		
Right eye	Original	Right eye	Binary		Right eye	Original	Right eye	Binary	
329 training	Both vs R	329 training	Both vs R		175 training	Both vs R	169 training	Both vs R	
83 testing	*P*=.17	83 testing	*P*=.21		44 testing	*P*=.34	43 testing	*P*=.97	
Accuracy = 0.7229		Accuracy = 0.6988			Accuracy = 0.9773		Accuracy = 0.9070		
Left eye	Original	Left eye	Binary		Left eye	Original	Left eye	Binary	
271 training	Both vs L	271 training	Both vs L		271 training	Both vs L	169 training	Both vs L	
68 testing	*P*=.22	68 testing	*P*=.73		68 testing	*P*=.08	43 testing	*P*=.12	
Accuracy = 0.8676		Accuracy = 0.7941			Accuracy = 0.8636		Accuracy = 0.9767		

L, left; R, right.

TABLE 2.

ADRET and ADVAS Models Performance for Alzheimer Disease Prediction in the UK Biobank and Our Institutional Data Set

	ADVAS		ADVAS			
	Original vessel segmentation		Binary vessel segmentation		ADRET	
	UK Biobank testing data set (n=150)	Institutional testing data set (n=120)	UK Biobank testing data set (n=150)	Institutional testing data set (n=116)	UK Biobank testing data set (n=150)	Institutional testing data set (n=109)
Accuracy	0.772 (0.0179)	0.9417 (0.0243)	0.7373 (0.0278)	0.9017 (0.0326)	0.9827 (0.0059)	0.989 (0.0077)
κ score	0.5377 (0.0376)	0.8785 (0.0412)	0.4691 (0.0635)	0.7984 (0.0618)	0.9652 (0.012)	0.9777 (0.0157)
AUROC	0.8137 (0.0189)	0.9800 (0.0077)	0.7809 (0.0372)	0.9588 (0.0177)	0.9967 (0.0022)	0.9978 (0.0019)
Sensitivity	0.6554 (0.0906)	0.9549 (0.0310)	0.6823 (0.0683)	0.9065 (0.0540)	0.9827 (0.0059)	0.989 (0.0077)
Specificity	0.8775 (0.0517)	0.9176 (0.0393)	0.7839 (0.0719)	0.8992 (0.0797)	0.9817 (0.0122)	0.9805 (0.0137)

The data are mean (SD). Five-fold cross-validation method was applied in each testing data set.

AUROC, area under the receiver operating characteristic curve.