



OPEN

# Enhancing foveal avascular zone analysis for Alzheimer's diagnosis with AI segmentation and machine learning using multiple radiomic features

Je Moon Yoon<sup>1,13</sup>, Chae Yeon Lim<sup>2,13</sup>, Hoon Noh<sup>3</sup>, Seung Wan Nam<sup>3,4</sup>, Sung Yeon Jun<sup>1</sup>, Min Ji Kim<sup>1</sup>, Mi Yeon Song<sup>1</sup>, Hyemin Jang<sup>5,6,7,8</sup>, Hee Jin Kim<sup>5,6</sup>, Sang Won Seo<sup>5,6,7</sup>, Duk L. Na<sup>5,6,9</sup>, Myung Jin Chung<sup>11,12</sup>, Don-Il Ham<sup>1,13</sup>✉ & Kyungsu Kim<sup>10,11,13</sup>✉

We propose a hybrid technique that employs artificial intelligence (AI)-based segmentation and machine learning classification using multiple features extracted from the foveal avascular zone (FAZ)—a retinal biomarker for Alzheimer's disease—to improve the disease diagnostic performance. Imaging data of optical coherence tomography angiography from 37 patients with Alzheimer's disease and 48 healthy controls were investigated. The presence or absence of brain amyloids was confirmed using amyloid positron emission tomography. In the superficial capillary plexus of the angiography scans, the FAZ was automatically segmented using an AI method to extract multiple biomarkers (area, solidity, compactness, roundness, and eccentricity), which were paired with clinical data (age and sex) as common correction variables. We used a light-gradient boosting machine (a light-gradient boosting machine is a machine learning algorithm based on trees utilizing gradient boosting) to diagnose Alzheimer's disease by integrating the corresponding multiple radiomic biomarkers. Fivefold cross-validation was applied for analysis, and the diagnostic performance for Alzheimer's disease was determined by the area under the curve. The proposed hybrid technique achieved an area under the curve of  $72.2 \pm 4.2\%$ , outperforming the existing single-feature (area) criteria by over 13%. Furthermore, in the holdout test set, the proposed technique exhibited a 14% improvement compared to single features, achieving an area under the curve of  $72.0 \pm 4.8\%$ . Based on these facts, we have demonstrated the effectiveness of our technology in achieving significant performance improvements in FAZ-based Alzheimer's diagnosis research through the use of multiple radiomic biomarkers (area, solidity, compactness, roundness, and eccentricity).

<sup>1</sup>Department of Ophthalmology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea. <sup>2</sup>Department of Medical Device Management and Research, SAIHST, Sungkyunkwan University, Seoul 06351, Republic of Korea. <sup>3</sup>Hangil Eye Hospital, 35 Bupyeong-daero, Bupyeong-gu, Incheon 21388, Republic of Korea. <sup>4</sup>Department of Ophthalmology, Catholic Kwandong University College of Medicine, 35 Bupyeong-daero, Bupyeong-gu, Incheon 21388, Republic of Korea. <sup>5</sup>Alzheimer's Disease Convergence Research Center, Samsung Medical Center, Seoul, Republic of Korea. <sup>6</sup>Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea. <sup>7</sup>Neuroscience Center, Samsung Medical Center, Seoul, Republic of Korea. <sup>8</sup>Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Republic of Korea. <sup>9</sup>Happymind Clinic, Seoul, Republic of Korea. <sup>10</sup>Medical AI Research Center, Research Institute for Future Medicine, Samsung Medical Center, Seoul 06351, Republic of Korea. <sup>11</sup>Department of Data Convergence and Future Medicine, Sungkyunkwan University School of Medicine, Suwon 16419, Republic of Korea. <sup>12</sup>Department of Radiology and AI Research Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea. <sup>13</sup>These authors contributed equally: Je Moon Yoon, Chae Yeon Lim, Don-Il Ham and Kyungsu Kim. ✉email: oculus@naver.com; kskim.doc@gmail.com

Alzheimer's disease (AD) is the most common type of dementia, affecting an estimated 5.4 million people in the United States<sup>1</sup>. With an estimated 13.8 million people affected by AD by 2050, health and societal expenses are predicted to notably increase<sup>2</sup>. The identification of biomarkers for early diagnosis and recruitment in interventional clinical trials has become a priority given the increasing prevalence of AD and lack of efficient therapeutic options. Current AD diagnosis is constrained by its expensive equipment (e.g., magnetic resonance imaging, positron emission tomography [PET]), invasiveness (e.g., extraction of cerebrospinal fluid [CSF]), insufficient specificity and sensitivity (e.g., genetic markers, serum amyloid), length of examination, accessibility to specialists, and neuropsychological testing<sup>3</sup>. For effective risk screening, the demand for faster, more accessible, and less invasive diagnostic tests is largely unmet.

AD is characterized by neuronal death, brain atrophy, extracellular deposition of beta-amyloid plaques, and intracellular accumulation of neurofibrillary tangles<sup>4,5</sup>. In addition, changes to the cerebral vasculature, such as cerebral amyloid angiopathy, atherosclerosis, and arteriosclerosis, reduced capillary density, and changed capillary morphology, have also been observed<sup>6–9</sup>. However, current *in vivo* imaging modalities fail to detect cerebral microvascular alterations.

The retina is considered an extension of the central nervous system, because the retina originates as outgrowth of the developing brain<sup>10</sup>. Hence, the retina may be a mirror of the brain. As sources of biomarkers for AD, retinal imaging modalities such as optical coherence tomography and fundus photography have been the subject of systematic reviews and meta-analyses<sup>11,12</sup>. Furthermore, unlike *in vivo* imaging of cerebral microvasculature, retinal microvasculature can be detected by optical coherence tomography angiography (OCTA), which provides high-resolution images of the retinal microvasculature and choroid<sup>13</sup>.

The foveal avascular zone (FAZ) is a region surrounding the fovea and devoid of retinal capillaries, and it can be imaged using OCTA. Although a meta-analysis has revealed an increase in the FAZ area in AD<sup>14</sup>, heterogeneity and conflicting results have been observed among studies. Recent research has shown that the FAZ shape in OCTA images is a reliable indicator of retinal disorders. For instance, the FAZ circularity and axial ratio are considerably different in eyes with diabetic retinopathy and normal eyes<sup>15,16</sup>, and the FAZ circularity is significantly lower in eyes with glaucoma presenting central visual field deficits than in eyes with peripheral visual field defects<sup>17</sup>. However, previous studies on AD using OCTA have mainly focused on the FAZ area, but its shape has mostly been neglected<sup>14</sup>.

In this study, we analyzed multiple FAZ features using radiomics-based machine learning (ML) for AD diagnosis. In addition, we investigated the diagnostic ability of other features when combined with the FAZ area. Finally, we developed a diagnosis method for AD that combines artificial-intelligence (AI)-based FAZ segmentation and an ML model for processing multiple radiomic features.

Our contributions are summarized as follows:

- The validity of existing representative techniques for AD diagnosis considering the FAZ area is verified with data from Korean patients collected at our hospital, showing an area under the curve (AUC) of 60%.
- Existing techniques do not use features other than the FAZ area, but *our technique includes other key FAZ features*, thus improving the AD diagnostic performance by more than 10% in AUC up to 72%.
- Unlike existing techniques that require manual annotations from specialists to extract the FAZ region, we apply *automatic AI-based segmentation that can promote the diagnostic performance*. Hence, a fully automatic technique for AD diagnosis from OCTA scans is obtained.
- We demonstrate that the proposed technique outperforms representative AI models for the differentiation of AD using OCTA scans as input. This demonstrates the usefulness of our *hybrid* diagnosis technique that combines AI-based segmentation and ML-based classification for Alzheimer's using multiple radiomic features.

## Materials and methods

### Ethical approval

All authors of this study confirm that all methods or experiments were performed in accordance with the Declaration of Helsinki and the relevant guidelines and regulations provided by the policies of the Nature Portfolio journals. This study was approved by the Institutional Review Board of the Samsung Medical Center (IRB number: SMC 2021-05-073). Written informed consent from the patients was waived by the Institutional Review Board (Samsung Medical Center, Seoul, Republic of Korea) because we used anonymized retrospective data.

### Study participants

All participants underwent amyloid PET and brain MRI at the memory clinic in the Department of Neurology at SMC in Seoul, South Korea<sup>18</sup>. As previously described<sup>19</sup>, all participants underwent comprehensive dementia evaluation, including a standardized neuropsychological test (Seoul Neuropsychological Screening Battery, 2nd edition<sup>20</sup>), blood tests including APOE genotyping, and brain MRI. We excluded participants who had any of the following conditions: (1) white matter hyperintensities due to etiologies other than vascular pathology, including radiation injury, multiple sclerosis, leukodystrophy, or metabolic/toxic disorders; (2) traumatic brain injury; (3) normal pressure hydrocephalus; (4) territorial infarction; (5) neurodegenerative disorders other than AD or ischemic etiologies such as progressive supranuclear palsy, corticobasal syndrome, frontotemporal dementia, or Lewy body/Parkinson disease dementias; or (6) rapidly progressive dementias and treatable dementias. All participants with normal controls (NCs) fulfilled the following criteria: (1) subjective memory complaints by participants or caregivers; (2) no objective cognitive dysfunction, as assessed by scores from evaluations on any cognitive domain; (3) no history of medical diseases likely to affect cognitive function; and (4) no significant

| Data type   | Patient-wise training set ( <i>n</i> = 55) |                        |                 | Patient-wise holdout test set ( <i>n</i> = 30) |                       |                 |
|---|--|------------------------|-----------------|--|-----------------------|-----------------|
| Data collection period                            | November 2021–September 2022               |                        |                 | July 2022–February 2023                        |                       |                 |
| Disease type                                      | AD ( <i>n</i> = 20)                        | NC ( <i>n</i> = 35)    | <i>P</i> -value | AD ( <i>n</i> = 17)                            | NC ( <i>n</i> = 13)   | <i>P</i> -value |
| Age (years)                                       | 65.7 ± 7.9                                 | 69.8 ± 7.1             | 0.186           | 66.8 ± 4.9                                     | 73.7 ± 5.5            | < 0.005         |
| Sex (Male, %)                                     | 45% ( <i>n</i> = 9)                        | 28.5% ( <i>n</i> = 10) | 0.185           | 58.8% ( <i>n</i> = 10)                         | 46.1% ( <i>n</i> = 6) | 0.164           |
| Mini-mental state examination score <sup>23</sup> | 18.4 ± 5.5                                 | 28.7 ± 1.6             | < 0.005         | 20.1 ± 1.3                                     | 28.0 ± 2.4            | < 0.005         |
| Study period (months)                             | 12.0 ± 3.7                                 | 12.0 ± 5.0             | 0.996           | 13.5 ± 4.7                                     | 11.0 ± 4.9            | 0.107           |
| Clinical dementia rating <sup>24</sup>            | 0.9 ± 0.9                                  | 0.1 ± 0.3              | < 0.005         | 0.6 ± 0.2                                      | 0.5                   | < 0.005         |
| Patients w/hypertension (%)                       | 30% ( <i>n</i> = 6)                        | 37.1% ( <i>n</i> = 13) | 0.789           | 29.4% ( <i>n</i> = 5)                          | 35.7% ( <i>n</i> = 5) | 0.218           |
| Patients w/diabetes (%)                           | 15% ( <i>n</i> = 3)                        | 20% ( <i>n</i> = 7)    | 0.787           | 29.4% ( <i>n</i> = 5)                          | 7.1% ( <i>n</i> = 1)  | 0.139           |

**Table 1.** Characteristics of data sets.

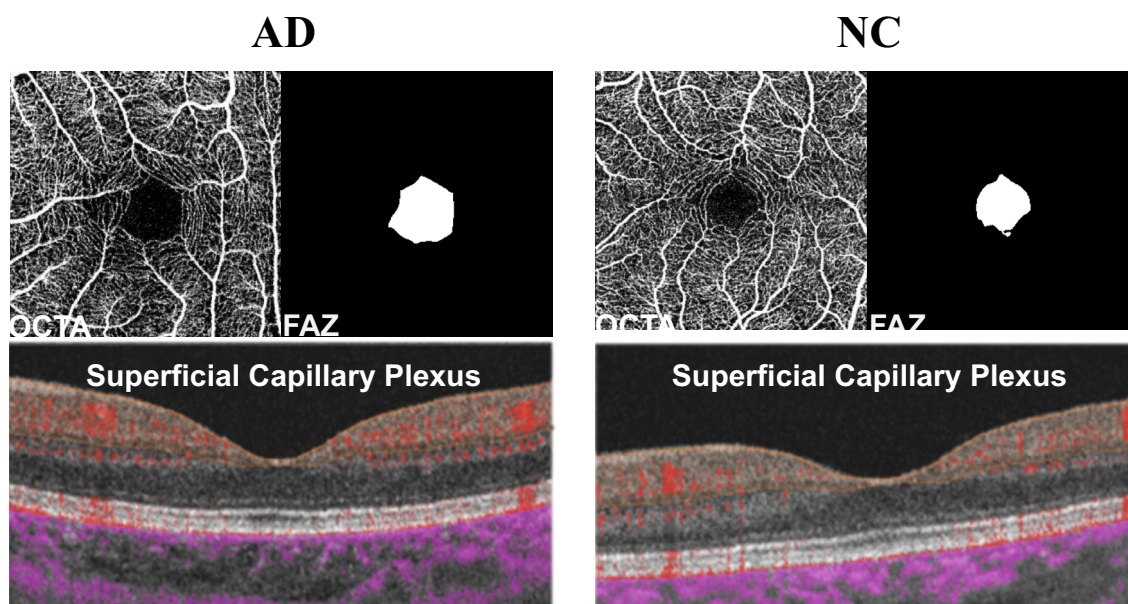
impairment in activities of daily living. All patients diagnosed with MCI fulfilled Petersen criteria for MCI<sup>21</sup>. Patients with dementia satisfied diagnostic criteria for dementia according to the DSM-IV<sup>22</sup>.

### Optical coherence tomography angiography acquisition and clinical data

All participants underwent OCTA scans<sup>25</sup> of the superficial capillary plexus layer<sup>26</sup> at Samsung Medical Center, Gangnam-Gu, Seoul, South Korea, between November 2021 and February 2023. OCTA was performed by an expert technician. The OCTA scanning protocol used was a 3 × 3 mm<sup>2</sup> volume scan centered on the fovea (DRI OCT Triton, Topcon, Japan). During data collection, we acquired clinical information of patients including age, sex, presence of hypertension or diabetes, education level, mini-mental state examination score<sup>23</sup>, and clinical dementia rating<sup>24</sup>, as specified in Table 1. From the clinical information, we used age and sex, which can affect the FAZ shape<sup>27</sup> and measured without any additional cost, along with OCTA scans as inputs for the proposed diagnosis technique. We obtained a total of 170 OCTA scan sets from 85 participants, as shown in Fig. 1. In addition, the 170 OCTA scans are divided into the training set and the holdout test set. From the scan sets, 25 scans were excluded from the training set, and 15 were excluded from the holdout test set. The exclusion criterion was considerable noise or low image quality, depending on the scanning environment<sup>28</sup>. Hence, the OCTA training set includes a total of 85 scans, with 31 scans from AD cases and 54 scans from NC cases, while the holdout test set consists of 45 scans, with 29 scans from AD cases and 16 scans from NC cases. In addition, we aimed to maximize the separation between the training dataset and the holdout dataset by collecting them at different times, although both were from a single institution.

### Dataset split

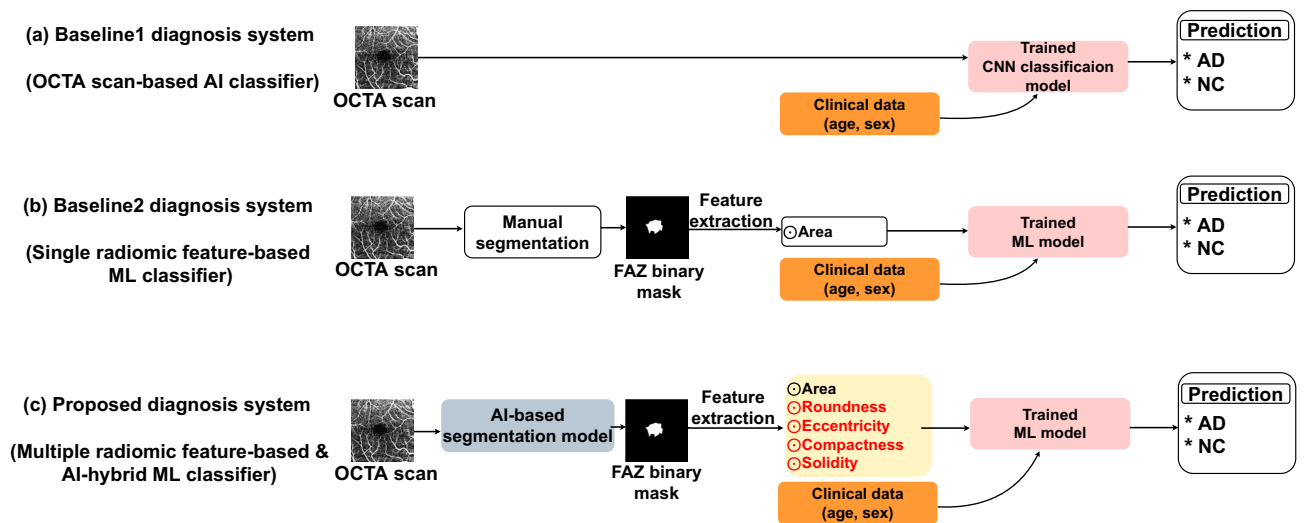
The dataset was divided into training and holdout test sets for each image before model selection. The training set was divided on a per-image basis for the fivefold cross-validation experiment, and the holdout test set was



**Figure 1.** OCTA scans in superficial capillary plexus layer for AD and NC. We aim to determine whether various FAZ features extracted from these scans contribute to AD diagnosis.

| Method    | Fivefold cross-validation |      | Holdout test              |
|-----------|---------------------------|------|---------------------------|
| Data type | Training set (n = 85)     |      | Holdout test set (n = 45) |
| Split     | Train                     | Test | Test                      |
| Fold 1    | 70                        | 15   | 45 (64%)                  |
| Fold 2    | 68                        | 17   |                           |
| Fold 3    | 69                        | 16   |                           |
| Fold 4    | 67                        | 18   |                           |
| Fold 5    | 66                        | 19   |                           |

**Table 2.** Summary of fivefold cross-validation and holdout test the data distribution of FAZ binary mask images.



**Figure 2.** Comparison of AD diagnosis techniques using FAZ analysis on OCTA scans. (a) Baseline 1—AI classifier with OCTA input, (b) baseline 2—ML classifier using single radiomic FAZ feature with manual FAZ segmentation, and (c) proposed technique—ML classifier with AI-based automatic FAZ segmentation using multiple radiomic FAZ features.

used exclusively for holdout testing without overlap with the training set. A summary of the data splitting used in the fivefold cross-validation and the holdout test is shown in Table 2.

### Existing and proposed techniques for AD diagnosis

In this section, we describe the inference of current AD diagnosis techniques by FAZ analysis using OCTA scans and compare these techniques with our proposal, as shown in Fig. 2. Existing AD diagnosis techniques using FAZ features include deep learning classification using OCTA scans (baseline 1) and radiomics based on a single FAZ feature (baseline 2).

#### Baseline 1: OCTA AI-based classifier

As shown in Fig. 2a, baseline 1<sup>29</sup> receives an OCTA scan as input for AI-based classification and learns to diagnose AD by binary classification between AD and NC. We use convolutional neural networks (CNNs)<sup>30</sup>, which are the gold standard models, for classification and design of a multimodal AI network that uses clinical and image information by expanding the input vector to receive two additional clinical datapoints in a fully connected network within the target CNN.

Baseline 1 learns to classify AD and NC through a CNN classifier using OCTA scan  $O$  and two clinical datapoints as follows:

$$\text{Baseline}_1(O, C) = \text{CNN}(O, C) \in \{AD, NC\} \quad (1)$$

where  $\text{CNN}$  represents the CNN (e.g., ResNet<sup>31</sup>, DenseNet<sup>32</sup>, EfficientNet<sup>33</sup>, Inception<sup>34</sup>), and  $C \in \mathbb{R}^2$  denotes two clinical datapoints (i.e., age and sex), which are commonly used for AD diagnosis. We concatenate information  $C$  to the fully connected network input vector obtained from global average pooling. The CNN output is a two-dimensional probability vector, and the CNN is trained using the cross-entropy loss to determine the diagnostic probabilities for AD and NC.

| Method     | Segmentation | Feature                 | Classification | Comp model |
|------------|--------------|-------------------------|----------------|------------|
| Baseline 1 | None         | Automatically extracted | AI             | AI         |
| Baseline 2 | Manual       | Only area               | ML             | ML         |
| Proposed   | AI           | Area + other features   | ML             | AI + ML    |

**Table 3.** Factor-specific differences between proposed and existing techniques.

*Baseline 2: ML classifier using single radiomic feature*

Baseline 2<sup>35</sup> uses the FAZ area extracted from OCTA scans to diagnose AD. This is a radiomic method<sup>36</sup> that uses only the FAZ area, which is only one of the available radiomic features. Consequently, baseline 2 uses a single radiomic feature, as illustrated in Fig. 2b. We train an ML model for binary classification of AD and NC, with the output using a 5D vector integrating the FAZ area and two clinical datapoints into the input vector for the ML model.

When using baseline 2, ophthalmologists should perform manual segmentation of the FAZ region on OCTA scan  $O$  to obtain FAZ binary mask  $S_O^{ma} \in \{0, 1\}^{h \times w}$  (i.e., 0 and 1 for the outer and inner FAZ, respectively). The FAZ area is obtained by multiplying the number of nonzero pixels in  $S_O$  (i.e., pixels located inside the FAZ) by constant  $c$  for the area per pixel as follows:

$$Area(S_O^{ma}) = \sum_{ij} c \times S_O^{ma}[i, j] \quad (2)$$

Then, baseline 2 is trained to classify AD and NC through an ML algorithm using the FAZ area and two clinical datapoints as follows:

$$Baseline_2(O, C) = ML(Area(S_O^{ma}), C) \in \{AD, NC\} \quad (3)$$

where  $ML(\cdot)$  represents an ML model (e.g., XGBoost<sup>37</sup>, random forest<sup>38</sup>, LGBM<sup>39</sup>).

*Proposed technique (AI-based segmentation with ML classifier using multiple radiomic features)*

In contrast to baseline 2, as shown in Fig. 2c, the proposed technique simultaneously uses five representative radiomic features<sup>40</sup> (i.e., area, roundness, eccentricity, compactness, and solidity) rather than simply considering the FAZ area. We experimentally demonstrate that multiple radiomic features increase the classifier diversity, thereby improving the AD diagnostic performance. In addition, unlike baseline 1, the proposed technique does not use an AI-based classifier but AI-based FAZ area segmentation. To obtain the binary mask for the FAZ area from an OCTA scan, AI-based segmentation is applied rather than the manual segmentation required for baseline 2. Hence, the proposed technique performs automatic AD diagnosis without pretreatment of OCTA scans, like in baseline 1. Moreover, AI-based FAZ segmentation mitigates annotation errors that may occur during manual FAZ segmentation, thereby increasing the accuracy of extracted radiomic features. Table 3 lists the characteristics of the evaluated techniques. Our proposal has a hybrid structure by combining AI and ML through the sequential execution of AI-based FAZ area segmentation and ML-based AD diagnosis based on multiple radiomic features extracted from the FAZ.

*Extraction of additional multiple radiomic features.* For comparison with baselines 1 and 2, four radiomic features were added to our technique, as defined below and illustrated in Fig. 3. Let  $S \in \{0, 1\}^{h \times w}$  be the FAZ binary segmentation mask.

- **Solidity.** The solidity measures the degree of curvature of the FAZ interface as the ratio of the FAZ inner area to its convex hull region:

$$Solidity(S) = \frac{Area(S)}{Area_{cvh}(S)} \quad (4)$$

where  $Area(S)$  is the FAZ area (area in which  $S$  is 1) and  $Area_{cvh}(S)$  is the FAZ convex hull area.

- **Compactness.** The compactness measures the degree of curvature of the FAZ interface as the ratio of the FAZ inner area to its perimeter:

$$Compact(S) = \frac{4\pi \cdot Area(S)}{p(S)^2} \quad (5)$$

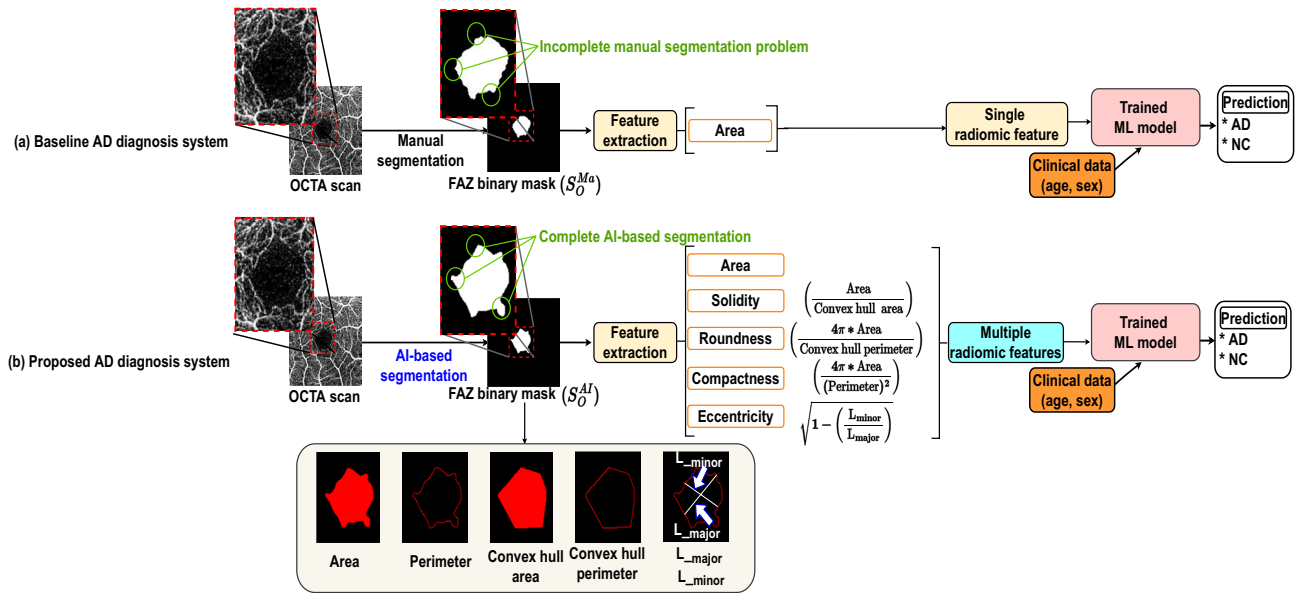
where  $p(S)$  is the FAZ perimeter.

- **Roundness.** The roundness is similar to the compactness but uses the perimeter of the convex hull rather than the perimeter of the FAZ:

$$Round(S) = \frac{4\pi \cdot Area(S)}{p_{cvh}(S)^2} \quad (6)$$

where  $p_{cvh}(S)$  is the perimeter of the FAZ convex hull.





**Figure 3.** Feature extraction from FAZ segmented on OCTA scan. Multiple radiomic features are used for training in the proposed technique.

- Eccentricity. The eccentricity is obtained as the ratio of the longest ( $a(S)$ ) to the shortest ( $b(S)$ ) straight-line length within the FAZ,  $S$ . It allows to measure the FAZ closeness to an ellipse as follows:

$$Eccent(S) = \sqrt{1 - \frac{b(S)^2}{a(S)^2}} \quad (a(S) \geq b(S)) \tag{7}$$

*AI-based FAZ segmentation.* Baseline 2 requires ophthalmologists to perform manual segmentation of the FAZ on OCTA scan  $O$  to obtain FAZ binary mask  $S_O^{ma} \in \{0, 1\}^{h \times w}$ . In contrast to baseline 2 with manual FAZ segmentation, the proposed technique uses an AI model to automatically segment the FAZ. Thus, the input is OCTA scan  $O$ , and the output is the extracted FAZ. For training, we used a public dataset, whereas our hospital data were used for evaluating FAZ segmentation and comparison with manual annotations in baseline 2. We denote the automatically segmented binary mask as  $S_O^{AI} \in \{0, 1\}^{h \times w}$ , with 0 and 1 indicating the outer and inner parts of the FAZ, respectively.

*Inference.* Using multiple radiomic features and automatic FAZ segmentation, the proposed technique performs inference as shown in Eq. (8), which is different from inference for the baselines given by Eqs. (1) and (3). In addition, the proposed technique learns to classify AD and NC through an ML model using the FAZ area, like in baseline 2, in addition to other four FAZ features.

$$Proposed(O, C) = ML(Area(S_O^{AI}), Solidity(S_O^{AI}), Compact(S_O^{AI}), Round(S_O^{AI}), Eccent(S_O^{AI}), C) \in \{AD, NC\} \tag{8}$$

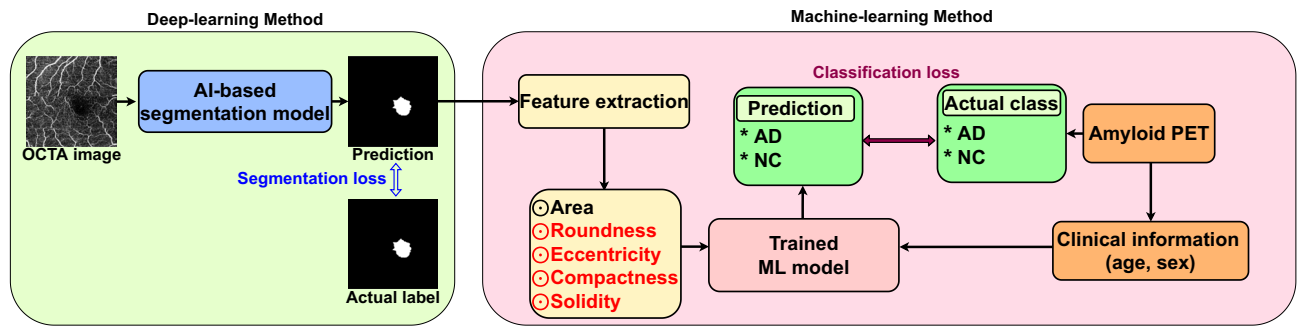
where  $ML(\cdot)$  represents the same ML model used in baseline 2,  $S_O^{AI}$  is the FAZ binary mask automatically extracted from OCTA scan  $O$ , and  $C$  represents the two clinical datapoints commonly used for all the techniques evaluated in this study.

### Evaluation metrics

To evaluate the proposed FAZ multiple radiomic features of Alzheimer’s diagnosis (binary classification of NC and AD), we used the area under the curve (AUC), accuracy, sensitivity, and specificity of the receiver operating characteristic (ROC) curve. For the ROC curve, we chose the most commonly used decision threshold of 0.5 and calculated the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) rates based on this threshold. We then calculated the accuracy, sensitivity, and specificity values as follows:

$$Accuracy = \frac{(TN + TP)}{(TN + TP + FN + FP)}, \tag{9}$$

$$Sensitivity = \frac{(TP)}{(TP + FN)}, \tag{10}$$



**Figure 4.** Training overview of proposed diagnosis technique. The proposed technique comprises AI-based FAZ segmentation and ML-based AD diagnosis using multiple radiomic FAZ features. Segmentation and classification loss functions are used as training losses for the AI and ML models, respectively. Data from our hospital are used for training and evaluation with fivefold cross-validation.

$$\text{Specificity} = \frac{(TN)}{(TN + FP)}, \quad (11)$$

## Results

### Fivefold cross-validation results on the training set

We compared the diagnostic performance of the proposed technique and the baseline technique for AD diagnosis on the training set. We divided the training set into 85 OCTA scans, which were divided into five sets to apply fivefold cross-validation. Each diagnosis technique was trained five times, and the mean validation performance was considered as the final diagnostic performance. The training details for each technique are detailed below.

### Training details

#### Training for baseline 1

As the CNN backbone used in baseline 1, we tested four representative models: ResNet<sup>31</sup>, DenseNet<sup>32</sup>, EfficientNet<sup>33</sup>, and Inception<sup>34</sup>. Each model was trained with fivefold cross-validation using the pretraining parameters on the ImageNet dataset for initialization. Each training procedure proceeded for 50 epochs by applying the cross-entropy loss<sup>41</sup> to a two-dimensional output probability vector for binary classification of positive (AD) and negative (NC) samples. The optimal learning rates were  $1e^{-2}$  for EfficientNet<sup>33</sup>,  $1e^{-2}$  for ResNet<sup>31</sup>,  $1e^{-5}$  for Inception<sup>34</sup>, and  $1e^{-2}$  for DenseNet<sup>32</sup>.

#### Training for baseline 2

Baseline 2 required manual extraction of FAZ binary mask  $S_O^{ma}$  from an OCTA scan. Thus, ophthalmologists extracted the FAZ binary masks from the 85 OCTA scans used in this study. In baseline 2, given the FAZ binary mask  $S_O^{ma}$  provided by the ophthalmologists, the area was calculated, and the ML model was applied to learn and evaluate the AD diagnosis by fivefold cross-validation. We used the LGBM<sup>39</sup> as the ML model.

#### Training for proposed technique

Unlike baseline 2, the proposed technique performs AI-based segmentation. It receives an OCTA scan as input and predicts the FAZ binary mask as the output. To train the segmentation model, we used 2000 OCTA scan–FAZ mask pairs<sup>42</sup> based on nnUNet<sup>43</sup>. As the learning objective function, the conventional pixel-based cross-entropy loss was used for training over 100 epochs under Adam optimization<sup>44</sup> with a learning rate of 0.01. Thus, 85 FAZ mask prediction results for the 85 evaluation OCTA scans were obtained from the learned segmentation model, and multiple radiomic features were extracted. Then, the ML model (LGBM<sup>39</sup> for a fair comparison with baseline 2) was applied for AD diagnosis in fivefold cross-validation. Training of the proposed technique is illustrated in Fig. 4.

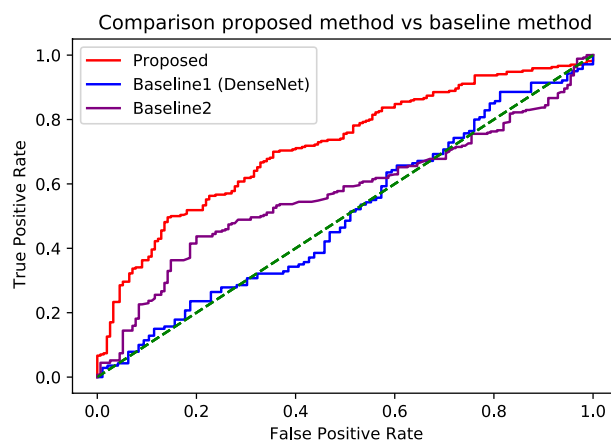
### Diagnostic performance

The diagnostic performance results (AUC) per fold and across folds of the evaluated techniques are listed in Table 4. The AUC of the proposed technique was at least 13% higher than that of the baselines. The baselines did not provide clinically meaningful results because all the AUC values were below 60%. In contrast, the proposed technique could achieve clinical significance with AUC values above 70%. Furthermore, the proposed method also demonstrates statistical significance with very low p-values when compared to the baselines, providing evidence of its statistical significance ( $p < 0.05$ ). Hence, this is the first technique demonstrating that multiple radiomic FAZ features are meaningful biomarkers for AD diagnosis.

Figure 5 shows the receiver operating characteristic curves of each technique for the aggregate AUC values derived from the cross-fold mean. In all the areas, regardless of the threshold, the proposed technique demonstrated higher sensitivity than the baselines, confirming its superiority.

| Technique                               | Fold 1    | Fold 2    | Fold 3    | Fold 4    | Fold 5    | Mean (%)          | <i>p</i> -value |
|---|-----------|-----------|-----------|-----------|-----------|-------------------|-----------------|
| Baseline 1 (ResNet) <sup>31</sup>       | 50        | 47        | 58        | 45        | 55        | 50.6 ± 4.4        | 0.0001          |
| Baseline 1 (DenseNet) <sup>32</sup>     | 51        | 49        | 60        | 52        | 50        | 52.4 ± 3.9        | 0.0001          |
| Baseline 1 (EfficientNet) <sup>33</sup> | 36        | 51        | 37        | 57        | 59        | 48 ± 9.7          | 0.0020          |
| Baseline 1 (Inception) <sup>34</sup>    | 51        | 50        | 38        | 56        | 61        | 51.2 ± 7.6        | 0.0014          |
| Baseline 2                              | 51        | 64        | 49        | 60        | 68        | 59.1 ± 7.1        | 0.0126          |
| <b>Proposed</b>                         | <b>76</b> | <b>76</b> | <b>65</b> | <b>70</b> | <b>72</b> | <b>72.2 ± 4.2</b> | –               |

**Table 4.** AUC of differential diagnosis between AD and NC. The mean and standard deviation are calculated from fivefold cross-validation.



**Figure 5.** Receiver operating characteristic curves of differential diagnosis between AD and NC. The AUC values are  $72.2 \pm 4.2$  (95% confidence interval, 66.7 – 75.5),  $52.4 \pm 3.9$  (95% confidence interval, 49 – 55.8), and  $59.1 \pm 6.2$  (95% confidence interval, 52.9 – 65.3) for the proposed technique, baseline 1 with DenseNet, and baseline 2, respectively.

| Model                       | Fold 1    | Fold 2    | Fold 3    | Fold 4    | Fold 5    | Mean (%)          | <i>p</i> -value |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|-------------------|-----------------|
| XGBoost <sup>37</sup>       | 65        | 72        | 62        | 68        | 59        | 65.5 ± 6.0        | 0.0346          |
| Random forest <sup>38</sup> | 64        | 72        | 53        | 64        | 56        | 62.3 ± 4.0        | 0.0632          |
| <b>LGBM<sup>39</sup></b>    | <b>76</b> | <b>76</b> | <b>65</b> | <b>70</b> | <b>72</b> | <b>72.2 ± 4.2</b> | –               |

**Table 5.** Performance of differential diagnosis between AD and NC using different ML models in the proposed technique. The mean and standard deviation are obtained from fivefold cross-validation.

## Analysis of proposed technique and its elements

### Performance of different ML models

The proposed technique diagnosed AD by feeding multiple radiomic FAZ features into an ML classifier. The results using LGBM<sup>39</sup> as a representative ML model are listed in Table 4, and the performance comparison of other ML models (i.e., XGBoost<sup>37</sup> and random forest<sup>38</sup>) is presented in Table 5. LGBM showed the highest performance, thus being selected as the ML model for the proposed technique. For models other than LGBM, the mean diagnostic performance was at least 60%. Thus, the proposed technique showed higher performance than the baselines (AUC of 60% or less), as shown in Table 4. This demonstrates that the proposed method is superior to the baselines regardless of the underlying ML model.

### Ablation study for proposed technique

The proposed technique uses multiple radiomic FAZ features (i.e., area, compactness, eccentricity, roundness, and solidity) instead of only one feature (i.e., area used in baseline 2). Table 6 shows that the diagnostic performance gradually improved when each of these features was added to the proposed technique. Hence, the diagnostic performance was improved by adding the four features to the area, justifying their inclusion in the technique.



| Features                      | Fold 1    | Fold 2    | Fold 3    | Fold 4    | Fold 5    | Mean (%)                 |
|-------------------------------|-----------|-----------|-----------|-----------|-----------|--------------------------|
| *Area                         | 54        | 69        | 62        | 60        | 69        | 63.2 ± 5.0               |
| Area + A                      | 57        | 71        | 60        | 61        | 69        | 64.0 ± 5.4 (+ 0.8)       |
| Area + A + B                  | 71        | 72        | 63        | 64        | 64        | 67.2 ± 3.7 (+ 4.0)       |
| Area + A + B + C              | 67        | 74        | 63        | 69        | 68        | 68.8 ± 3.3 (+ 5.6)       |
| <b>**Area + A + B + C + D</b> | <b>76</b> | <b>76</b> | <b>65</b> | <b>70</b> | <b>72</b> | <b>72.2 ± 4.2 (+9.0)</b> |

**Table 6.** Comparison of AD diagnostic performance by including radiomic features. A, compactness; B, eccentricity; C, roundness; D, solidity. The mean and standard deviation are obtained from fivefold cross-validation. \*Baseline 2 method (i.e., Using only area feature). \*\*Proposed method (i.e., Using five multiple radiomic features).

| Segmentation     | Fold 1    | Fold 2    | Fold 3    | Fold 4    | Fold 5    | Mean (%)                  |
|------------------|-----------|-----------|-----------|-----------|-----------|---------------------------|
| Manual           | 65        | 67        | 55        | 68        | 52        | 61.9 ± 6.5                |
| <b>Automatic</b> | <b>76</b> | <b>76</b> | <b>65</b> | <b>70</b> | <b>72</b> | <b>72.2±4.2* (+ 10.3)</b> |

**Table 7.** AD diagnostic performance for manual and automatic segmentation. The mean and standard deviation are obtained from fivefold cross-validation. \* $p < 0.05$ .

| Method          | Accuracy (%)              | Sensitivity (%)           | Specificity (%)          | AUC (%)                 |
|-----------------|---------------------------|---------------------------|--------------------------|-------------------------|
| Baseline        | 50.7±5.5                  | 35.2±15.6                 | 78.7±13.5                | 58.0±0.9                |
| <b>Proposed</b> | <b>64.8±4.3* (+ 14.1)</b> | <b>54.4±5.0* (+ 19.2)</b> | <b>83.7±6.3* (+ 5.0)</b> | <b>72.0±4.8* (+ 14)</b> |

**Table 8.** Diagnostic performance comparison between proposed and baseline method. Accuracy, sensitivity, specificity, and AUC score obtained from the holdout test set. Values in parentheses represent improvements in performance between the proposed and baseline method diagnoses. The mean and standard deviation are obtained from the holdout test. \* $p < 0.05$ .

#### *Validity of AI-based segmentation in proposed technique*

The proposed approach automatically extracts the FAZ by AI-based segmentation. Table 7 lists the diagnostic performance when the proposed technique uses the binary masks manually annotated by an ophthalmologist, as in baseline 2, instead of the automatically segmented FAZ. The AUC of AD diagnosis was enhanced by 10.3% when AI-based automatic segmentation was used compared with manual segmentation. This performance improvement followed from the more accurate and precise AI-based FAZ extraction compared with manual annotations. We explain this performance improvement in section Discussion.

#### **Comparison of diagnostic performance of proposed methods on holdout test set**

We conducted a holdout test using a set of 45 OCTA scan holdout datasets. The holdout test was compared and verified by baseline<sup>35</sup>(i.e., baseline 2 method only area feature) and ophthalmologists, respectively. accuracy, sensitivity, specificity, and AUC measured the evaluation of each diagnostic technique.

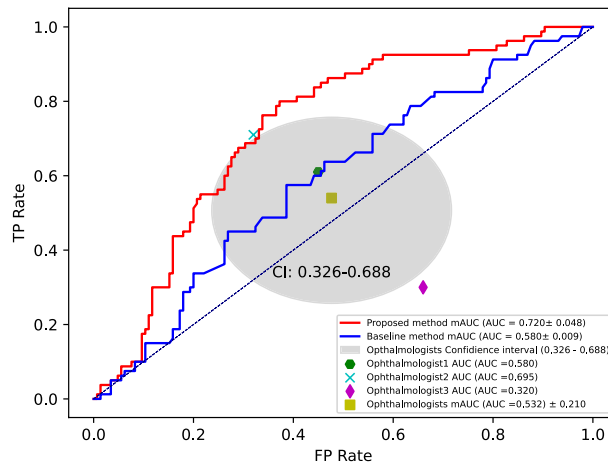
#### *Comparison diagnostic performance between baseline and proposed method*

The diagnostic performance results for the baseline<sup>35</sup> method (i.e., the baseline 2 method uses only a single area feature) and the proposed method (multiple radiomic features) are detailed in Table 8.

FAZ binary masks were obtained manually and automatically (i.e., AI-based segmentation) for a holdout dataset for both the baseline and the proposed method. The FAZ binary masks were then used to test pre-trained models through a fivefold cross-validation process. In the holdout test, the proposed method showed a significant improvement in AUC compared to the baseline, with an improvement of 14% (baseline AUC 58.0% vs proposed AUC 72.0%). Furthermore, when compared to the baseline, the proposed method showed a 14.1% increase in accuracy (baseline accuracy 50.7% vs proposed accuracy 64.8%), a 5.0% increase in specificity (baseline specificity 78.7% vs proposed specificity 83.7%), and a 19.2% increase in sensitivity (baseline sensitivity 35.2% vs proposed sensitivity 54.4%). These results demonstrate the robustness of the proposed method in the holdout test, demonstrating superior performance in all evaluation metrics compared to the baseline ( $p < 0.05$ ). Therefore, the proposed method demonstrates excellent diagnostic performance for the diagnosis of Alzheimer's disease based on FAZ.

| Method          | Accuracy (%)              | Sensitivity (%)          | Specificity (%)           | AUC (%)                   |
|-----------------|---------------------------|--------------------------|---------------------------|---------------------------|
| Humans          | 53.1±15                   | 52.8±14.1                | 53.7±18.8                 | 53.2±21.0                 |
| <b>Proposed</b> | <b>64.8±4.3* (+ 11.7)</b> | <b>54.4±5.0* (+ 1.6)</b> | <b>83.7±6.3* (+ 30.0)</b> | <b>72.0±4.8* (+ 18.8)</b> |

**Table 9.** Diagnostic performance comparison between proposed method and ophthalmologist. Accuracy, sensitivity, specificity, and AUC score obtained from the holdout test set. Values in parentheses represent improvements in performance between the proposed method and ophthalmologist diagnoses. The mean and standard deviation are obtained from the holdout test. \*  $p < 0.05$ .



**Figure 6.** Comparison diagnostic performance of the proposed method and average of AUCs for three ophthalmologists on holdout test. The AUC values are  $72.0 \pm 4.8$  (95% confidence interval, 67.7 – 76.2),  $58.0 \pm 0.009$  (95% confidence interval, 57.9 – 58.0), and  $53.2 \pm 21.0\%$  (95% confidence interval, 32.0 – 69.5) for the proposed technique, baseline, and average of AUCs for three ophthalmologists.

#### Comparison diagnostic performance between the ophthalmologist and proposed method

diagnosis of humans was conducted with the evaluation of three experienced ophthalmologists who did not participate in the collection of the OCTA holdout test dataset. For the diagnosis of humans, training was performed using the FAZ binary mask of 85 labeled training sets, and subsequently, an evaluation was performed using the FAZ binary mask of 45 unlabeled holdout datasets. The results of comparing the proposed method with the ophthalmologists are presented in Table 9.

The proposed method showed superior performance in all metrics (sensitivity, specificity, accuracy, and AUC) compared to ophthalmologists, particularly showing a significant improvement of over 30% in specificity ( $p < 0.05$ ). This suggests that the proposed method is more sensitive in reducing false positives compared to ophthalmologists (humans' specificity 53.7% vs proposed specificity 83.7%). In other words, it can significantly reduce the rate of false positive predictions for normal patients, which is cost-effective by saving on additional testing expenses (humans' specificity 53.7% vs proposed specificity 83.7%). Furthermore, the proposed technique demonstrated higher sensitivity compared to ophthalmologists (humans' sensitivity 52.8% vs proposed sensitivity 54.4%) and showed strong discriminative power for false negatives (humans' AUC 53.2% vs proposed AUC 72.0%). Consequently, the proposed method shows potential utility as a clinical support tool for Alzheimer's diagnoses based on FAZ in the future.

Figure 6 shows the AUC results for the binary classification of AD and NC using the proposed method and three ophthalmologists. The proposed method yields results that are 18.8% higher than the average AUC of the three ophthalmologists (humans' AUC 53.2% vs proposed AUC 72.0%). In addition, the AUC of the proposed method showed a 14% improvement compared to the baseline (baseline AUC 58.0% vs proposed AUC 72.0%). This confirms that the proposed method (i.e., using multiple radiomics features including area) exhibited a significant performance improvement by considering multiple radiomics features, in contrast to the baseline method that relied on a single feature (i.e., using only area) for Alzheimer's diagnosis (baseline AUC 58.0% vs proposed AUC 72.0%). Notably, a significant performance improvement was achieved even when compared to ophthalmologists (humans' AUC 53.2% vs proposed AUC 72.0%). This indicates the potential of multiple radiomics features as a novel biomarker in FAZ-based Alzheimer's diagnosis.

#### Discussion

We showed that multiple radiomic FAZ features can be extracted by an AI model to support AD diagnosis. To the best of our knowledge, this is the first report using multiple radiomic FAZ features for diagnosis in patients

with AD. We developed an automatic AD diagnosis technique comprising AI-based FAZ segmentation and ML-based AD diagnosis using the automatically extracted FAZ features.

#### *Clinical implications of multiple radiomic FAZ features*

Early detection of AD is of paramount importance, as it allows intervention prior to the onset of irreversible brain degeneration. Nevertheless, the current gold-standard diagnostic methods for AD, such as amyloid PET scans or CSF analysis, are insufficient as early screening tools. The retina, due to its embryological similarities with the brain and its easily and safely examined anatomical features, presents a promising avenue for the early detection of AD. The FAZ is a potential retinal biomarker for AD. The FAZ can be extracted from OCTA, which is a noninvasive retinal imaging modality. A recent meta-analysis revealed an enlargement of the FAZ in AD<sup>14</sup>. Another meta-analysis reported an enlarged FAZ in patients with mild cognitive impairment but no significant enlargement in AD<sup>45</sup>, while another meta-analysis showed no significant enlargement of the FAZ in AD<sup>46</sup>. Although limitations included the heterogeneity of OCTA equipment, diverse scanning protocols, and unmeasured confounders, previous studies only investigated the FAZ area, neglecting the FAZ shape, which may be a reliable indicator of retinal disorders<sup>15,17,47</sup>.

While area is frequently utilized as a primary metric for characterizing the FAZ, it is essential to recognize the substantial normal variation in FAZ size<sup>48</sup>. This variability may potentially constrain its utility as a pathological indicator in cross-sectional screening applications<sup>49</sup>. Evaluating the regularity of the FAZ's overall shape, measured in terms of roundness or circularity, may offer a more precise indication of disease due to reduced variability within the healthy population<sup>50</sup>. Consequently, it is imperative to investigate whether biomarkers related to the shape of the FAZ possess diagnostic capabilities in individuals with AD. Recently, not only have there been reports of studies using ophthalmic imaging and AI for the diagnosis of ophthalmic diseases, but there have also been reports on their use for diagnosing AD<sup>29,51–53</sup>. In this study, we first revealed that multiple radiomic FAZ features, including roundness, eccentricity, compactness, and solidity, can improve the AD diagnostic performance compared with the FAZ area alone. Therefore, multiple radiomic FAZ features are useful for diagnosis and should be considered when evaluating the FAZ as new biomarkers for AD.

While our advanced AI-based methodology demonstrated a successful diagnosis of AD with a favorable diagnostic accuracy of “72.2 ± 4.2%,” it is important to acknowledge that this figure falls short of direct comparison with current gold standard diagnostic methods. Notably, our diagnosis was solely based on retinal imaging, without the utilization of traditionally established diagnostic tools for AD, such as amyloid PET scans, CSF tapping, brain imaging, and even the Mini-Mental State Examination. Nevertheless, the findings from our study have significant clinical implications. They bridge a well-recognized diagnostic gap by providing a non-invasive and cost-effective means for screening AD, circumventing the need for invasive and expensive tests like PET, CSF tapping, and brain MRI. This innovative approach not only offers potential clinical utility but also signals a promising avenue for further refinement. Moreover, our study's results indicate the potential for further refinement of AI-based diagnostic techniques, which holds promise for future research endeavors focused on enhancing the early detection of AD. This work not only contributes to the field's knowledge but also paves the way for continued exploration and development in the realm of AD diagnosis.

#### *Possible mechanisms for FAZ changes in AD*

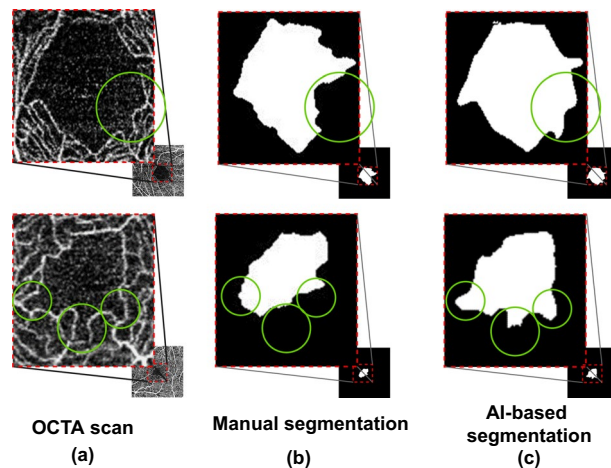
Vascular dysfunction in patients with AD likely leads to cerebral hypoperfusion during AD development<sup>54–58</sup>. In vivo and autopsy data have revealed that AD is associated with the deposition of amyloid and collagen within the cerebral capillaries, which can result in cellular apoptosis and vessel dropout<sup>59–62</sup>. In addition, various studies have found the accumulation of beta-amyloid plaques in the inner retina of postmortem tissue extracted from patients with AD<sup>63–66</sup>. Therefore, FAZ changes in patients with AD may be secondary to retinal degeneration owing to beta-amyloid accumulation within the retina.

#### *Performance improvement by AI-based FAZ segmentation*

To evaluate the effectiveness of AI-based FAZ segmentation integrated into the proposed technique, we compared it with manual FAZ segmentation, obtaining the results listed in Table 7. Manual segmentation was the same as that in baseline 2. Compared with manual segmentation, AI-based segmentation improved the diagnostic performance in terms of AUC from 61.9 to 72.2 (improvement of 10.3%). As shown in Fig. 7, the performance improvement was due to AI-based FAZ segmentation overcoming problems and errors in manual annotation, which showed some inaccurate or mistaken results. Nguyen et al.<sup>67</sup> reported the high performance of AI-based FAZ segmentation. We observed that the AI-based FAZ segmentation extracted the FAZ more precisely. Thus, the multiple radiomic FAZ features were more precisely determined, thereby improving AD diagnosis.

#### *Performance improvement by multiple radiomic features*

Different from previous studies<sup>35</sup>, we considered multiple FAZ features (i.e., area, roundness, eccentricity, compactness, and solidity) to diagnose AD. Existing techniques relied only on the FAZ area, and their AD diagnostic performance was not high. We demonstrated that various FAZ features contributed to further improving AD diagnosis, as indicated in Table 10. Every feature considered in this study (i.e., roundness, eccentricity, compactness, and solidity) contributed to the diagnosis, individually leading to comparable performance to that of the area. This individual validation may indicate the diagnostic performance improvement achieved by feature combination, as shown in Table 5, with the performance gradually improving as more features were added.



**Figure 7.** Comparison between FAZ segmentation methods. (a) Original OCTA scan and results from (b) manual and (c) AI-based segmentation.

| Feature             | Fold 1    | Fold 2    | Fold 3    | Fold 4    | Fold 5    | Mean (%)          |
|---------------------|-----------|-----------|-----------|-----------|-----------|-------------------|
| Area                | 54        | 69        | 62        | 60        | 69        | 63.2 ± 5.0        |
| Solidity            | 59        | 75        | 50        | 66        | 71        | 64.6 ± 8.0        |
| Roundness           | 54        | 68        | 47        | 62        | 66        | 59.9 ± 7.0        |
| Compactness         | 60        | 67        | 58        | 72        | 71        | 66.0 ± 5.0        |
| Eccentricity        | 63        | 78        | 61        | 53        | 60        | 63.5 ± 8.0        |
| <b>All features</b> | <b>76</b> | <b>76</b> | <b>65</b> | <b>70</b> | <b>72</b> | <b>72.2 ± 4.2</b> |

**Table 10.** AD diagnostic performance obtained from every radiomic feature. The mean and standard deviation are obtained from fivefold cross-validation.

#### Technical implications

Our hybrid technique achieved an AUC of 72.2%, thus improving the AD diagnostic performance using FAZ biomarkers by 13.1% compared with existing techniques. This result holds notable clinical significance because it confirms that the FAZ is a suitable biomarker for AD, even though it was previously overlooked due to its low diagnostic performance in AD diagnosis. Hence, high AD diagnostic performance may be achieved by using FAZ biomarkers along with well-known biomarkers (e.g., global retinal nerve fiber layer, retinal thickness, vascular density, and FAZ area) that have been used for noninvasive AD diagnosis.

#### Promising performance of additional features in isolated feature analysis

We conducted comparative experiments by individually diagnosing isolated features, including the previously reported area<sup>35</sup> feature and the four additional multiple features introduced in this study (i.e., solidity, compactness, eccentricity, roundness). The results for each of these isolated single features are detailed in Table 10.

Compared to the area, which was previously reported in FAZ-based Alzheimer's diagnosis, the additional multiple features demonstrated their diagnostic potential, with solidity at 64.6% (+ 1.4%), roundness at 59.9% (− 3.3%), compactness at 66.0% (+ 2.8%) and eccentricity at 63.5% (+ 0.3%) in the single feature comparison experiments. This confirms the excellent performance of the majority of these features. Furthermore, it suggests that these multiple features (i.e., solidity, compactness, eccentricity, roundness) have significant correlations with structural changes in FAZ caused by Alzheimer's disease, beyond only area. This research not only contributes to the significant impact of FAZ-based Alzheimer's diagnosis but also provides the first study presenting meaningful biomarkers for detecting structural changes due to other ocular diseases.

#### Instrumental applicability of the proposed method in real clinical settings

The proposed method showed a high-specificity model, but it was possible to derive a model with high sensitivity by adjusting the thresholds. Threshold adjustment resulted in a sensitivity of 90% and a specificity of 33%. This means that it can identify 90% of Alzheimer's patients while detecting around 30% of the normal control group. Specifically, it excels at accurately detecting 90% of Alzheimer's patients, enabling them to be referred for secondary testing such as amyloid PET scans and CSF analysis. At the same time, it provides a basis for reducing the cost of secondary testing in around 30% of normal patients. This is because, unlike ophthalmologists, the proposed technique uses AI technology to achieve a model with high sensitivity through various threshold

| Study                         | OCTA scans as input | Target disease                                      | AI-based segmentation? | Multiple FAZ features | Diagnosis considering multiple FAZ features |
|-------------------------------|---------------------|---|------------------------|-----------------------|---|
| Werner et al. <sup>71</sup>   | O                   | Retinal vein occlusion.                             | X                      | X                     | X   |
| Shiihara et al. <sup>69</sup> | O                   | NC  | X                      | O                     | X   |
| Philip et al. <sup>70</sup>   | O                   | Exfoliation glaucoma<br>Primary open-angle glaucoma | X                      | O                     | X   |
| Chan et al. <sup>68</sup>     | O                   | Diabetic retinopathy                                | O                      | X                     | X   |
| Mirshahi et al. <sup>50</sup> | O                   | Diabetic retinopathy                                | O                      | X                     | X   |
| O'Bryhim et al. <sup>35</sup> | O                   | AD  | X                      | X                     | X   |
| <b>This study</b>             | <b>O</b>            | <b>AD</b>   | <b>O</b>               | <b>O</b>              | <b>O</b>                                    |

**Table 11.** Characteristic of proposed and existing techniques. Unlike previous studies, our study covers all the listed aspects.

adjustments. As a result, we can provide a model that, for the first time, detects up to 90% of actual Alzheimer's patients while providing a false positive rate of less than 10%.

#### *Comparison between proposed and existing techniques*

We compared and analyzed the differences between the proposed and existing techniques regarding various aspects, as summarized in Table 11.

OCTA provides scans in a short time, enabling efficient noninvasive FAZ analysis. In only one other study, OCTA was used for AD diagnosis (third column of Table 11)<sup>35</sup>. However, that study used only the FAZ area, discarding other radiomic features. We demonstrated the importance of using multiple FAZ features for AD diagnosis by improving the diagnostic performance when using the proposed technique compared with conventional techniques that use a single feature (i.e., baseline 2).

Among existing studies using OCTA, Chan et al.<sup>68</sup> and Mirshahi et al.<sup>50</sup> used AI-based segmentation to extract the FAZ (fourth column of Table 11). They reported that using AI enabled the extraction of FAZ boundaries with better accuracy than existing signal processing methods, thereby validating the use of AI-based FAZ segmentation in our technique. However, the contribution of the extracted FAZ to diagnosis was not confirmed in those studies. Our study has both technical and clinical significance because we showed that AI-based FAZ segmentation can improve the diagnostic performance for AD.

Shiihara et al.<sup>69</sup> and Philip et al.<sup>70</sup> extracted multiple FAZ features like in our study (fifth column of Table 11). However, they did not develop an ML model for diagnosing a specific disease using multiple FAZ features (sixth column of Table 11). Shiihara et al.<sup>69</sup> found a small difference between individuals in other FAZ features in addition to the area for healthy subjects, thereby suggesting their potential as biomarkers. However, their study was limited to healthy subjects, without confirming the possibility of using multiple biomarkers for diagnosing specific diseases. In contrast, we demonstrated that features other than the FAZ area are useful biomarkers for AD and developed an ML model for diagnosis. Philip et al.<sup>70</sup> analyzed whether multiple FAZ features were individually correlated with primary open-angle glaucoma or exfoliation glaucoma, but they did not observe a correlation with a specific disease by combining multiple features. Therefore, they did not validate feature combinations. In addition, they did not implement a technique for disease diagnosis taking those features as inputs. Our study provides clinical and technical significance by overcoming existing limitations in AD diagnosis by implementing an ML model that receives multiple radiomic FAZ features as inputs and provides the AD diagnosis result as output.

#### *Limitations*

A major limitation of our study was the small sample size, which consisted solely of Asian individuals. Nevertheless, to overcome this limitation, a fivefold cross-validation and a holdout test were applied in the paper. The limitation of this holdout test is that it relies on data from a single institution and lacks external validation. However, in our study, we tried to collect data for the fivefold cross-validation experiment and the holdout test at different periods, attempting to separate the data as effectively as possible. Furthermore, the exclusion of patients with known vascular disease from our study was another limitation. We could not evaluate whether these results are applicable to individuals who may have retinal microvascular alterations by other causes. In addition, the inclusion of participants with cognitive changes and positive biomarkers for AD limited comparisons with subjects with preclinical and positive biomarkers, such as mild cognitive impairment. Nevertheless, we demonstrated the diagnostic ability of FAZ with AI for AD and all individuals in this study, including those in the AD and NC groups, were screened by amyloid PET. In future work, a comparison between patients with mild cognitive impairment and NCs and longitudinal changes in FAZ in AD will be considered.

## **Conclusion**

Employing an advanced AI-based methodology, we successfully automated FAZ segmentation and extracted a comprehensive array of radiomic FAZ features. The integration of FAZ area with these additional features presents a promising avenue for the development of robust and potentially transformative biomarkers for AD. Furthermore, our AI-driven FAZ analysis, encompassing automatic segmentation and multi-feature extraction,



not only holds substantial promise for AD diagnosis but also extends its utility to the broader spectrum of retinal disorders, underlining its pivotal role in advancing clinical ophthalmology and neurology.

### Data availability

The main data supporting the results of this study are available within the paper. The raw datasets from Samsung Medical Center are protected to preserve patient privacy, but they can be made available upon reasonable request if approval is obtained from the corresponding Institutional Review Board. For the request, please contact Don-Il Ham (di.ham@samsung.com).

### Code availability

The code used in this study is available at <https://github.com/kskim-phd/Hybrid-FAZ>, which will be disclosed upon the publication or second/revised submission.

Received: 18 February 2023; Accepted: 7 January 2024

Published online: 22 January 2024

### References

- Association, A. *et al.* 2016 Alzheimer's disease facts and figures. *Alzheimer's Dement.* **12**, 459–509 (2016).
- Hebert, L. E., Weuve, J., Scherr, P. A. & Evans, D. A. Alzheimer disease in the united states (2010–2050) estimated using the 2010 census. *Neurology* **80**, 1778–1783 (2013).
- Grossman, I. *et al.* Alzheimer's disease: Diagnostics, prognostics and the road to prevention. *EPMA J.* **1**, 293–303 (2010).
- Gauthier, S., Scheltens, P. & Cummings, J. *Alzheimer's Disease and Related Disorders* (CRC Press, 2005).
- Reitz, C. & Mayeux, R. Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem. Pharmacol.* **88**, 640–651 (2014).
- Arvanitakis, Z., Capuano, A. W., Leurgans, S. E., Bennett, D. A. & Schneider, J. A. Relation of cerebral vessel disease to Alzheimer's disease dementia and cognitive function in elderly people: A cross-sectional study. *Lancet Neurol.* **15**, 934–943 (2016).
- Brown, W. R. & Thore, C. R. Cerebral microvascular pathology in ageing and neurodegeneration. *Neuropathol. Appl. Neurobiol.* **37**, 56–74 (2011).
- Smith, E. E. & Greenberg, S. M.  $\beta$ -amyloid, blood vessels, and brain function. *Stroke* **40**, 2601–2606 (2009).
- Thal, D. R., Attems, J. & Ewers, M. Spreading of amyloid, tau, and microvascular pathology in Alzheimer's disease: Findings from neuropathological and neuroimaging studies. *J. Alzheimer's Dis.* **42**, S421–S429 (2014).
- London, A., Benhar, I. & Schwartz, M. The retina as a window to the brain—from eye research to CNS disorders. *Nat. Rev. Neurol.* **9**, 44–53 (2013).
- Thomson, K. L., Yeo, J. M., Waddell, B., Cameron, J. R. & Pal, S. A systematic review and meta-analysis of retinal nerve fiber layer change in dementia, using optical coherence tomography. *Alzheimer's Dement. Diagn. Assess. Disease Monit.* **1**, 136–143 (2015).
- McGrory, S. *et al.* The application of retinal fundus camera imaging in dementia: A systematic review. *Alzheimer's Dement. Diagn. Assess. Disease Monit.* **6**, 91–107 (2017).
- Spaide, R. F., Fujimoto, J. G., Waheed, N. K., Sadda, S. R. & Staurengi, G. Optical coherence tomography angiography. *Prog. Retin. Eye Res.* **64**, 1–55 (2018).
- Rifai, O. M. *et al.* The application of optical coherence tomography angiography in Alzheimer's disease: A systematic review. *Alzheimer's Dement. Diagn. Assess. Disease Monit.* **13**, e12149 (2021).
- Krawitz, B. D. *et al.* Acircularity index and axis ratio of the foveal avascular zone in diabetic eyes and healthy controls measured by optical coherence tomography angiography. *Vis. Res.* **139**, 177–186 (2017).
- Tang, F. Y. *et al.* Determinants of quantitative optical coherence tomography angiography metrics in patients with diabetes. *Sci. Rep.* **7**, 1–10 (2017).
- Kwon, J., Choi, J., Shin, J. W., Lee, J. & Kook, M. S. Alterations of the foveal avascular zone measured by optical coherence tomography in glaucoma patients with central visual field defects. *Investig. Ophthalmol. Vis. Sci.* **58**, 1637–1645 (2017).
- Kang, S. H. *et al.* Amyloid positivity in the Alzheimer/subcortical-vascular spectrum. *Neurology* **96**, e2201–e2211 (2021).
- Fazekas, F. *et al.* The relation of cerebral magnetic resonance signal hyperintensities to Alzheimer's disease. *J. Neurol. Sci.* **142**, 121–125 (1996).
- Kang, Y., Jang, S. & Na, D. *Seoul Neuropsychological Screening Battery (SNSB)* (Human Brain Research & Consulting Co., 2012).
- Petersen, R. C. *et al.* Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* **56**, 303–308 (1999).
- Guze, S. B. Diagnostic and statistical manual of mental disorders, (dsm-iv). *Am. J. Psychiatry* **152**, 1228–1228 (1995).
- Tombaugh, T. N. & McIntyre, N. J. The mini-mental state examination: A comprehensive review. *J. Am. Geriatr. Soc.* **40**, 922–935 (1992).
- O'Bryant, S. E. *et al.* Staging dementia using clinical dementia rating scale sum of boxes scores: A Texas Alzheimer's research consortium study. *Arch. Neurol.* **65**, 1091–1095 (2008).
- Huang, D. *et al.* Optical coherence tomography. *Science* **254**, 1178–1181 (1991).
- Grewal, D. S., Polascik, B. W., Hoffmeyer, G. C. & Fekrat, S. Assessment of differences in retinal microvasculature using OCT angiography in Alzheimer's disease: A twin discordance report. *Ophthalmic Surg. Lasers Imaging Retin.* **49**, 440–444 (2018).
- Bulut, M. *et al.* Evaluation of optical coherence tomography angiographic findings in Alzheimer's type dementia. *Br. J. Ophthalmol.* **102**, 233–237 (2018).
- Enders, C. *et al.* Quantity and quality of image artifacts in optical coherence tomography angiography. *PLoS ONE* **14**, e0210505 (2019).
- Wisely, C. E. *et al.* Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. *Br. J. Ophthalmol.* **106**, 388–395 (2022).
- O'Shea, K. & Nash, R. An introduction to convolutional neural networks. [arXiv:1511.08458](https://arxiv.org/abs/1511.08458) (2015).
- 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*, 770–778 (2016).
- Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708 (2017).
- Tan, M. & Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*, 6105–6114 (PMLR, 2019).
- Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI Conference on Artificial Intelligence* (2017).
- O'Bryhim, B. E., Apte, R. S., Kung, N., Coble, D. & Van Stavern, G. P. Association of preclinical Alzheimer disease with optical coherence tomographic angiography findings. *JAMA Ophthalmol.* **136**, 1242–1248 (2018).

36. Mayerhoefer, M. E. *et al.* Introduction to radiomics. *J. Nucl. Med* **61**, 488–495 (2020).
37. Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 (2016).
38. Liaw, A. *et al.* Classification and regression by randomforest. *R News* **2**, 18–22 (2002).
39. Ke, G. *et al.* Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems* **30** (2017).
40. Wirth, M. A. Shape analysis and measurement. University of Guelph. CIS **6320** (2001).
41. De Boer, P.-T., Kroese, D. P., Mannor, S. & Rubinstein, R. Y. A tutorial on the cross-entropy method. *Ann. Oper. Res.* **134**, 19–67 (2005).
42. Wang, Y. *et al.* OCTA image dataset with pixel-level mask annotation for FAZ segmentation (2021).
43. Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
44. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
45. Hui, J. *et al.* Detection of retinal changes with optical coherence tomography angiography in mild cognitive impairment and Alzheimer's disease patients: A meta-analysis. *PLoS ONE* **16**, e0255362 (2021).
46. Katsimpris, A. *et al.* Optical coherence tomography angiography in Alzheimer's disease: A systematic review and meta-analysis. *Eye* **36**, 1419–1426 (2022).
47. Musik, I. & Biliński, S. Reaction of 1-acetyl-4 (o-tolyl)-thiosemicarbazide with alpha-halogenketones. i. cyclization with chloroacetone. In *Annales Universitatis Mariae Curie-Skłodowska. Sectio D: Medicina*, vol. 43, 83–90 (1988).
48. Linderman, R. E. *et al.* Variability of foveal avascular zone metrics derived from optical coherence tomography angiography images. *Transl. Vis. Sci. Technol.* **7**, 20–20 (2018).
49. Kostic, M. *et al.* Investigating the fractal dimension of the foveal microvasculature in relation to the morphology of the foveal avascular zone and to the macular circulation in patients with type 2 diabetes mellitus. *Front. Physiol.* **9**, 1233 (2018).
50. Mirshahi, R. *et al.* Foveal avascular zone segmentation in optical coherence tomography angiography images using a deep learning approach. *Sci. Rep.* **11**, 1–8 (2021).
51. Singh, L. K., Pooja, Garg, H. & Khanna, M. An IoT based predictive modeling for glaucoma detection in optical coherence tomography images using hybrid genetic algorithm. *Multimed. Tools Appl.* **81**, 37203–37242 (2022).
52. Singh, L. K., Khanna, M. & Thawkar, S. A novel hybrid robust architecture for automatic screening of glaucoma using fundus photos, built on feature selection and machine learning-nature driven computing. *Expert Syst.* **39**, e13069 (2022).
53. Singh, L. K., Khanna, M., Thawkar, S. & Singh, R. Collaboration of features optimization techniques for the effective diagnosis of glaucoma in retinal fundus images. *Adv. Eng. Softw.* **173**, 103283 (2022).
54. Johnson, K. A. *et al.* Florbetapir (F18-AV-45) PET to assess amyloid burden in Alzheimer's disease dementia, mild cognitive impairment, and normal aging. *Alzheimer's Dement.* **9**, S72–S83 (2013).
55. Di Marco, L. Y. *et al.* Vascular dysfunction in the pathogenesis of Alzheimer's disease—a review of endothelium-mediated mechanisms and ensuing vicious circles. *Neurobiol. Dis.* **82**, 593–606 (2015).
56. Park, K. W. *et al.* Regional cerebral blood flow differences in patients with mild cognitive impairment between those who did and did not develop Alzheimer's disease. *Psychiatry Res. Neuroimaging* **203**, 201–206 (2012).
57. Caroli, A. *et al.* Cerebral perfusion correlates of conversion to Alzheimer's disease in amnesic mild cognitive impairment. *J. Neurol.* **254**, 1698–1707 (2007).
58. Nobili, F. *et al.* Brain SPECT in subtypes of mild cognitive impairment. Findings from the DESCRIPA multicenter study. *Eur. J. Nucl. Med. Mol. Imaging* **34**, S133–S133 (2007).
59. Kalaria, R. N. & Pax, A. B. Increased collagen content of cerebral microvessels in Alzheimer's disease. *Brain Res.* **705**, 349–352 (1995).
60. De Jong, G., De Vos, R., Steur, E. J. & Luiten, P. Cerebrovascular hypoperfusion: A risk factor for Alzheimer's disease? animal model and postmortem human studies. *Ann. N. Y. Acad. Sci.* **826**, 56–74 (1997).
61. Christov, A., Ottman, J., Hamdheydari, L. & Grammas, P. Structural changes in Alzheimer's disease brain microvessels. *Curr. Alzheimer Res.* **5**, 392–395 (2008).
62. Farkas, E., De Jong, G. I., de Vos, R. A., Jansen Steur, E. N. & Luiten, P. G. Pathological features of cerebral cortical capillaries are doubled in Alzheimer's disease and Parkinson's disease. *Acta Neuropathol.* **100**, 395–402 (2000).
63. Koronyo-Hamaoui, M. *et al.* Identification of amyloid plaques in retinas from Alzheimer's patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model. *Neuroimage* **54**, S204–S217 (2011).
64. Koronyo, Y. *et al.* Retinal amyloid pathology and proof-of-concept imaging trial in Alzheimer's disease. *JCI Insight* **2**, e93621 (2017).
65. Tsai, Y. *et al.* Ocular changes in TgF344-AD rat model of Alzheimer's disease. *Investig. Ophthalmol. Vis. Sci.* **55**, 523–534 (2014).
66. La Morgia, C. *et al.* Melanopsin retinal ganglion cell loss in Alzheimer disease. *Ann. Neurol.* **79**, 90–109 (2016).
67. Nguyen, Q. N., Nguyen, V. T.-Q., Hsu, T., Vajzovic, L. & Ngo, H. T. Automatic foveal avascular zone segmentation using hessian-based filter and U-Net deep learning network. In *International Conference on the Development of Biomedical Engineering in Vietnam*, 835–842 (Springer, 2020).
68. Chan, F. *et al.* Machine learning-assisted automated quantification of foveal avascular zone parameters and perifoveal capillary density of prototype and commercial optical coherence tomography angiography (OCT-A) platforms in healthy and diabetic eyes. *Investig. Ophthalmol. Vis. Sci.* **58**, 827–827 (2017).
69. Shihara, H. *et al.* Objective evaluation of size and shape of superficial foveal avascular zone in normal subjects by optical coherence tomography angiography. *Sci. Rep.* **8**, 1–9 (2018).
70. Philip, S. *et al.* Macula vessel density and foveal avascular zone parameters in exfoliation glaucoma compared to primary open-angle glaucoma. *Investig. Ophthalmol. Vis. Sci.* **60**, 1244–1253 (2019).
71. Werner, J. U. *et al.* Comparison of foveal avascular zone between optical coherence tomography angiography and fluorescein angiography in patients with retinal vein occlusion. *PLoS ONE* **14**, e0217849 (2019).

## Acknowledgements

This study was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (2021R1A2C1007824). This work was also supported by the Future Medicine 20\*30 Project of the Samsung Medical Center (SMX1210791). The funding organizations had no role in the design or execution of this study.

## Author contributions

(1) Research Project: A. Conception and design, B. Acquisition of data, and C. Analysis and interpretation of data. (2) Manuscript: A. Writing of the first draft, B. Review and critique. (3) Others: A. Statistical analysis, B. Obtaining funding, C. Technical support, D. Study supervision, E. Supervision of data collection. J.M.Y.: 1A, 1B, 1C, 2A, 2B, C.Y.L.: 1C, 2A, 2B, 3A, 3C, H.N.: 1B, 2B, S.W.N.: 1B, 2B, S.Y.J.: 1B, 2B, M.J.K.: 1B, 2B, M.Y.S.:

1B, 2B, H.J.: 1B, 2B, H.J.K.: 1B, 2B, S.W.S.: 1B, 2B, D.L.N.: 1B, 2B, M.J.C.: 3C, D.I.H.: 1B, 2B, 3E, K.K.: 1A, 1C, 2A, 2B, 3A, 3C, 3D.

### Competing interests

The authors have no conflicts of interest or financial disclosures to declare.

### Additional information

**Correspondence** and requests for materials should be addressed to D.-I.H. or K.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits 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/4.0/>.

© The Author(s) 2024