

Deep Learning Analysis of Retinal Structures and Risk Factors of Alzheimer's Disease

Seowung Leem*

*J. Crayton. Pruitt Department of
Biomedical Engineering
University of Florida
Gainesville, USA
leem.s@ufl.edu*

Yunchao Yang

*University of Florida Research
Computing
University of Florida
Gainesville, USA
yunchaoyang@ufl.edu*

Adam J. Woods

*Department of Clinical and
Health Psychology
University of Florida
Gainesville, USA
ajwoods@phhp.ufl.edu*

Ruogu Fang*

*J. Crayton. Pruitt Department of
Biomedical Engineering
University of Florida
Gainesville, USA
ruogu.fang@bme.ufl.edu*

Abstract—The importance of early Alzheimer's Disease screening is becoming more apparent, given the fact that there is no way to revert the patient's status after the onset. However, the diagnostic procedure of Alzheimer's Disease involves a comprehensive analysis of cognitive tests, blood sampling, and imaging, which limits the screening of a large population in a short period. Preliminary works show that rich neurological and cardiovascular information is encoded in the patient's eye. Due to the relatively fast and easy procedure acquisition, early-stage screening of Alzheimer's Disease patients with eye images holds great promise. In this study, we employed a deep neural network as a framework to investigate the relationship between risk factors of Alzheimer's Disease and retinal structures. Our result shows that the model not only can predict several risk factors above the baseline but also can discover the relationship between the retinal structures and risk factors to provide insights into the retinal imaging biomarkers of Alzheimer's disease.

Keywords—Deep Learning, Alzheimer's Disease, Oculomics, Risk Stratification

I. INTRODUCTION

The importance of managing risk factors for Alzheimer's Disease (AD) is increasingly acknowledged for the early screening of patients. The absence of a cure post-onset of AD and the increasing number of AD-related deaths coupled with the decreasing death rate of other major diseases underline the significance of risk management [1], [2]. The AD screening & diagnosis involve a series of tests from diverse domains. The domains include cognitive tests [3], blood assays [4], cerebrospinal fluid (CSF) [5], and magnetic resonance imaging [6]. Although the thorough examination of these risk factors contributes to the accurate diagnosis, each test requires a substantial amount of time for acquisition and analysis.

Oculomics, which employs macro and microscopic image-driven features from the retinal structures to analyze a patient's health status, provides the means to overcome the limitations of the existing screening process [7]. To analyze health via oculomic features, there's a need for a feature extraction tool that proficiently captures the characteristics of the retina. Previously, risk factor management was performed by the hand-driven features from the ophthalmologists. While analysis through features measured by experts yielded clear results, it demanded supervision from ophthalmologists and a

meticulously designed hypothesis to reveal the associations [8], [9]. In contrast to statistical approaches, deep learning technique holds promise on various fronts. It does not require specific parameter settings, or supervision from the experts to extract the retinal features related to the risk factors.

In this study, we developed the deep learning model for risk factor prediction using only colored retinal fundus photography as an input. Our model was designed to capture the visual features from the retinal structures available in fundus photography and to effectively map the computed features for the prediction of risk factors. For evaluation, we not only performed a general model performance analysis but also analyzed how much the model utilized the retinal structures for correct prediction.

II. METHODS

A. Dataset

Table. 1 shows the general information of our dataset used for the model development. We acquired our dataset from the UK Biobank, the comprehensive biomedical dataset retaining diverse information from more than 500,000 subjects from the UK. The technical details of how the acquisition of each data was performed are explained in this website (<https://www.ukbiobank.ac.uk/>). From the dataset, we obtained the colored fundus image and the list of risk factors (6 categorical & 6 continuous). The risk factor list was based on the prior clinical studies which were found to be significantly correlated with the onset of Alzheimer's Disease.

To ensure the proper development of the model, preprocessing for both the image and the risk factor data was performed. For images, the resizing, and quality assessment were performed using an established pipeline [10]. We filtered out the low-quality images to acquire 62,874 images from 37,254 subjects for the model development. All right fundus images were flipped horizontally for consistency. The augmentations of the fundus images were not performed to keep the local features of retinal structures during the model development. For risk factors, imputations were not performed for the missing data to maintain the distribution of the data as closely as possible. Instead, we excluded subjects if any variables were missing, and variables that couldn't be utilized as predictors (e.g. -1: Do not know, -3: Prefer not to answer).

TABLE I. BASELINE CHARACTERISTICS OF SUBJECTS IN DEVELOPMENT AND VALIDATION SETS AFTER PROCESSING

Risk Factors	Dataset				Reference
	Development set		Validation set		
	Classification(n=21,435)	Regression(n=19,528)	Classification(n=5,359)	Regression(n=4,882)	
Gender: male %	47.47	n/a	47.70	n/a	[2], [11], [12]
Age: mean years (s.d.)	n/a	56.38 (8.18)	n/a	56.31 (8.29)	[2], [11], [13]
Education: mean years (s.d.)	n/a	16.99 (2.40)	n/a	16.94 (2.33)	[2], [11], [14]
Sleeplessness: %	28.10 Never/Rarely 47.00 Sometimes 24.90 Usually	n/a	28.40 Never/Rarely 45.46 Sometimes 26.14 Usually	n/a	[2], [15]
Current Smoker: %	8.83	n/a	9.40	n/a	[2], [16]
Alcohol Use: %	21.70, Daily/Almost daily 24.59, 3~4 times a week 24.88, 1~2 times a week 11.29, 1~3 times a month 10.69, Special Occasions 6.84, Never	n/a	21.18, Daily/Almost daily 23.89, 3~4 times a week 24.86, 1~2 times a week 11.48, 1~3 times a month 11.42, Special Occasions 7.18, Never	n/a	[2], [17]
Recurrent Depression: % positive	27.29	n/a	27.08	n/a	[2], [16], [18]
Economic Status: %	17.72, Less than 18k £ 23.18, 18k~31k £ 26.55, 31k~52k £ 24.28, 52k~100k £ 8.27, > 100k £	n/a	18.72, Less than 18k £ 23.68, 18k~31k £ 26.31, 31k~52k £ 23.18, 52k~100k £ 8.12, > 100k £	n/a	[11], [19]
BMI: mean (s.d.)	n/a	27.62 (4.74)	n/a	27.61 (4.82)	[2], [11], [14]
Diastolic BP: mean mmHg (s.d.)	n/a	82.12 (10.00)	n/a	82.01 (10.28)	[2], [12], [16]
Systolic BP: mean mmHg (s.d.)	n/a	137.60 (18.35)	n/a	137.56 (18.70)	[2], [12], [16]
HbA1C: mean mmol/mol (s.d.)	n/a	36.05 (6.64)	n/a	36.06 (6.81)	[2], [16]

This operation was performed separately for categorical and continuous risk factors. As a result of the operation, the data from 26,794 (Development Set: 21,435 & Validation Set: 5,439) subjects for classification and 24,410 (Development Set: 19,528 & Validation Set: 4,882) subjects for regression were determined for model development and validation. Thus, we acquired left and right fundus images for model development (classification=36,310 & regression=32,868) and validation (classification=9,042 & regression=8,211). After the finalization of the development and validation set, we applied additional transforms to some variables that required refinement. For continuous risk factors, the 2 measures of blood pressure values were averaged to a single value. In addition, we normalized the factors with mean and standard deviation, obtained from the development set to prevent data leakage. For categorical variables, the smoking status was modified for classification variables as a binary variable, representing a ‘current smoker’. In addition, the depression was also modified as a binary classification problem, representing a possibility of ‘recurrent major depression’.

B. Model Development

In this study, we used a Swin-Transformer [20], a well-recognized vision transformer model developed for computer vision, as a foundation model (Swin-L model). To efficiently train the model, we performed transfer learning with parameters pretrained with the ImageNet dataset. Since vision transformer-base models require a sizable dataset for desirable performance, transfer learning allowed us to build a model with decent

performance using a comparatively small dataset size. For model architecture, only the last fully connected layer of the network was replaced with the fully connected layers designed for either classification or regression tasks. Thus, we trained 2 different models for each task.

To find the best parameters for each task, we used an AdamW optimizer [21] with a learning rate of $1e-4$. The training objective of the model was to minimize the loss determined for each task, which is categorical cross-entropy loss for classification, and mean squared error for the regression task. In specific, the loss of each risk factor was computed separately and aggregated as a single loss term for the backpropagation. During 100 training epochs, the parameters with the best result in internal validation were chosen to test on the clinical validation set. We chose the maximum batch size that our hardware could withhold, which was 16 in our experiment. The model was trained using HiPerGator AI, with 32 NVIDIA DGX A100 GPUs.

C. Model Evaluation & Analysis

We computed the different evaluation metrics for each task. For classification, accuracy, and area under the receiver operating characteristics curve (AUROC) was computed. In the regression task, the coefficient of determination (R^2 Score) was computed to assess the model’s ability to accurately predict continuous values. To provide the confidence intervals for the metric, we performed bootstrapping; images were sampled equal to the number in the validation set with replacement and

TABLE II. MODEL PERFORMANCE ON PREDICTING RISK FACTORS OF ALZHEIMER'S DISEASE

Risk Factors	Prediction Result		
	Metric	Performance	Random
Gender	Accuracy	0.8584 (0.8556, 0.8612)	0.5
	AUROC	0.9353 (0.9334, 0.9371)	0.5
Age	R ²	0.6701 (0.6653, 0.6747)	0
	R ²	0.1034 (0.0897, 0.1170)	0
Education	Accuracy	0.4503 (0.4462, 0.4544)	0.3333
	AUROC	0.5468 (0.5433, 0.5502)	0.5
Sleeplessness	Accuracy	0.9078 (0.9054, 0.9101)	0.5
	AUROC	0.6503 (0.6427, 0.6582)	0.5
Current Smoker	Accuracy	0.2624 (0.2589, 0.2611)	0.1667
	AUROC	0.5888 (0.5857, 0.5917)	0.5
Alcohol Use	Accuracy	0.7157 (0.7120, 0.7194)	0.5
	AUROC	0.5646 (0.5593, 0.5699)	0.5
Depression	Accuracy	0.3076 (0.3038, 0.3113)	0.2
	AUROC	0.6281 (0.6253, 0.6311)	0.5
Economic Status	R ²	0.2442 (0.2324, 0.2564)	0
	R ²	0.3159 (0.3064, 0.3258)	0
Diastolic BP	R ²	0.2025 (0.1916, 0.2131)	0
	R ²	0.3301 (0.3139, 0.3468)	0
Systolic BP	R ²	0.3301 (0.3139, 0.3468)	0
	R ²	0.3301 (0.3139, 0.3468)	0
HbA1C	R ²	0.3301 (0.3139, 0.3468)	0
	R ²	0.3301 (0.3139, 0.3468)	0

repeated 2000 times. To determine the retinal structures employed by the DL model for prediction, we acquired the saliency map through class activation mapping (CAM) [22]. Specifically, we selected the first normalization layer in the last Swin Transformer block and acquired 12 different saliency maps for each risk factor. Moreover, we computed the extent to which retinal structures contributed to the model's prediction (r). Equation (1) shows how we performed this computation. We initially obtained the binary segmentation map ($M_{i,j}$) of the artery, vein, optic cup, and disc (σ) from the existing segmentation algorithm [10]. Subsequently, we applied a threshold of 0.5 to binarize the saliency map ($S_{i,j}$), considering that the maps were generated as a normalized value ranging from 0 to 1. After computing the overlap region between the segmentations of the retinal structures and the binarized saliency map, we obtained the ratio of retinal structures contributing to the model prediction.

$$r^\sigma = M^\sigma \cap S / M^\sigma \quad (1)$$

III. RESULTS

A. Model Performance of Risk Factor Prediction

We assessed our model's performance in predicting AD risk factors solely based on colored fundus images. Table. 2 shows the model performance for each risk factor variable. In the regression task, the DL model demonstrated high confidence predictions for age. Furthermore, the performance of the other continuous variables including diastolic BP, systolic BP, BMI, and HbA1C was better than the baseline. However, the model faced challenges in accurately predicting the age of completing full-time education and performing less effectively compared to other risk factors. In the classification task, the model excelled in predicting the gender of the subjects. For the alcohol use, and economic status, the model achieved accuracy above the baseline, but the low AUROC suggests that the model had difficulties predicting minority classes for each risk factor. While the classification model faced challenges in predicting positive samples for current smokers and depression due to the limited number of such samples, high accuracy gives a promise that DL models hold the potential for

TABLE III. MODEL'S INFERENCE OF THE RELATIONSHIP BETWEEN RETINAL STRUCTURES AND AD RISK FACTORS

Risk Factors	Model's Inference (%)			
	Artery	Vein	Optic cup	Optic disc
Gender	4.26	6.47	79.28	100
Age	35.66	30.01	100	100
Education	11.39	11.04	0	0
Sleeplessness	2.92	6.1	0	0
Current Smoker	19.86	17.63	100	100
Alcohol Use	8.37	9.76	0	0
Depression	1.44	1.28	0	0
Economic Status	0.29	0.72	24.37	2.75
BMI	29.43	28.2	58.56	27.19
Diastolic BP	15.99	22.37	0	0
Systolic BP	9.10	10.00	0	0
HbA1C	14.74	16.7	93.08	100

accurate prediction of the smoking and depression status when enough data is available.

B. Model's Inference of Correlation Between the Retinal Structures and Alzheimer's Disease Risk Factors

Fig. 1 illustrates the saliency map generated by the DL during the prediction of each risk factor. The visualization of the saliency indicates that the model makes predictions not by randomly selecting features, but by extracting the information from the retinal structures in the input. Table. 3 presents the model's inference of the relationship between retinal structure and AD risk factor calculated using the saliency map. The inference percentage suggests that retinal arteries and veins have an akin relationship with age, BMI, and blood pressure. The percentage for optic cups and discs indicates these structures encode information related to gender, age, smoking status BMI, and HbA1C. For risk factors without correlation to artery, vein, optic cup, and disc (e.g. sleeplessness, economic status, alcohol use), the model's saliency map indicates the focus was made in the macula region.

IV. DISCUSSION

The model's performance in both classification and regression tasks demonstrated the robust prediction ability of DL models related to gender and age. While the model could predict other risk factors more than the baseline, the improvement was not significant. While the model's performance may not have been notable for certain risk factors, assessing the capability of DL models only based on a metric like AUROC and R² underestimates the capability as a tool to study the relationship between the AD and fundus images.

The visualization of the model's class activation map and subsequent analysis of the model's inference regarding the contribution of retinal structures to prediction demonstrates that the model could identify the relationship between retinal structures and risk factors without any prior background information or supervision. This outcome suggests that DL models can serve as a tool for uncovering the underlying relationship between the retinal images and risk factors that may not be easily correlated with existing rationale and hypothesis. Hence, in future research, we aim to extend this analysis to other risk factors such as CSF biomarkers and brain measures. This will explicitly show the effectiveness of the DL model in identifying correlations between risk factors and associated retinal structures.

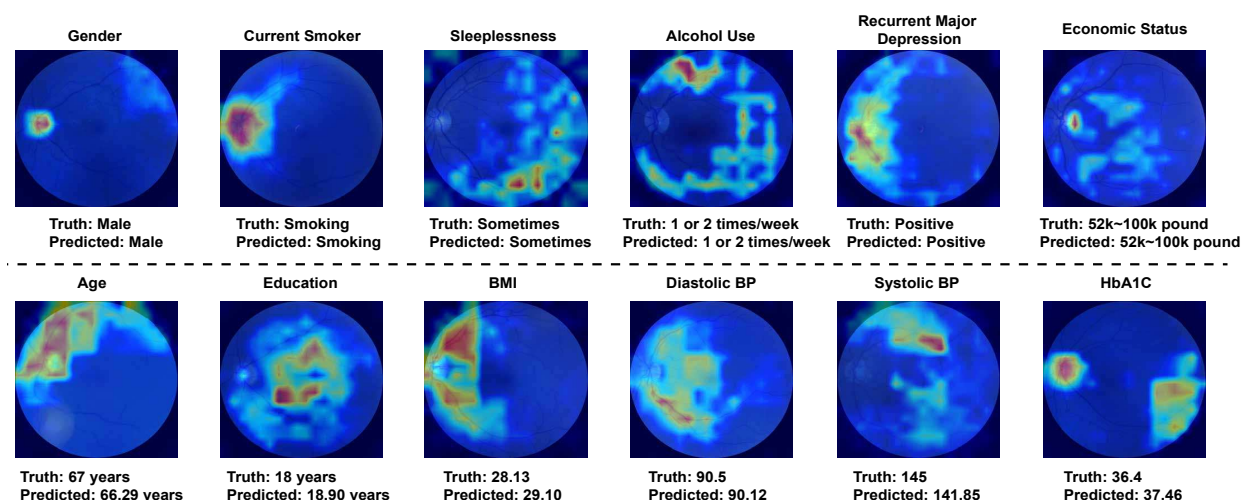


Fig. 1. Visualization of the saliency map for each risk factor. The results of the classification tasks are presented in the first row and regression tasks are presented in the second. For each risk factor, the subject's ground truth and the model's prediction are provided.

REFERENCES

- [1] K. B. Rajan, J. Weuve, L. L. Barnes, E. A. McAninch, R. S. Wilson, and D. A. Evans, "Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020-2060)," *Alzheimer's Dement. J. Alzheimer's Assoc.*, vol. 17, no. 12, pp. 1966–1975, Dec. 2021, doi: 10.1002/alz.12362
- [2] "2023 Alzheimer's disease facts and figures," *Alzheimer's Dement. J. Alzheimer's Assoc.*, vol. 19, no. 4, pp. 1598–1695, Apr. 2023, doi: 10.1002/alz.13016
- [3] I. Arevalo-Rodriguez *et al.*, "Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI)," *Cochrane Database Syst. Rev.*, vol. 2015, no. 3, p. CD010783, Mar. 2015, doi: 10.1002/14651858.CD010783.pub2
- [4] S. J. Kiddle, N. Voyle, and R. J. B. Dobson, "A Blood Test for Alzheimer's Disease: Progress, Challenges, and Recommendations," *J. Alzheimer's Dis. JAD*, vol. 64, no. s1, pp. S289–S297, 2018, doi: 10.3233/JAD-179904
- [5] A. Anoop, P. K. Singh, R. S. Jacob, and S. K. Maji, "CSF Biomarkers for Alzheimer's Disease Diagnosis," *Int. J. Alzheimer's Dis.*, vol. 2010, p. 606802, Jun. 2010, doi: 10.4061/2010/606802
- [6] K. A. Johnson, N. C. Fox, R. A. Sperling, and W. E. Klunk, "Brain Imaging in Alzheimer Disease," *Cold Spring Harb. Perspect. Med.*, vol. 2, no. 4, p. a006213, Apr. 2012, doi: 10.1101/cshperspect.a006213
- [7] S. G. Honavar, "Oculomics – The eyes talk a great deal," *Indian J. Ophthalmol.*, vol. 70, no. 3, p. 713, Mar. 2022, doi: 10.4103/ijo.IJO_474_22
- [8] C. G. Owen *et al.*, "Retinal Vasculometry Associations with Cardiometabolic Risk Factors in the European Prospective Investigation of Cancer-Norfolk Study," *Ophthalmology*, vol. 126, no. 1, pp. 96–106, Jan. 2019, doi: 10.1016/j.ophtha.2018.07.022
- [9] R. Klein, B. E. K. Klein, S. E. Moss, T. Y. Wong, and A. R. Sharrett, "Retinal vascular caliber in persons with type 2 diabetes: the Wisconsin Epidemiological Study of Diabetic Retinopathy: XX," *Ophthalmology*, vol. 113, no. 9, pp. 1488–1498, Sep. 2006, doi: 10.1016/j.ophtha.2006.03.028
- [10] Y. Zhou *et al.*, "AutoMorph: Automated Retinal Vascular Morphology Quantification Via a Deep Learning Pipeline," *Transl. Vis. Sci. Technol.*, vol. 11, no. 7, p. 12, Jul. 2022, doi: 10.1167/tvst.11.7.12
- [11] L. Wang, P. Li, M. Hou, X. Zhang, X. Cao, and H. Li, "Construction of a risk prediction model for Alzheimer's disease in the elderly population," *BMC Neurol.*, vol. 21, no. 1, p. 271, Jul. 2021, doi: 10.1186/s12883-021-02276-8
- [12] Z. He, S. Tian, A. Erdengasileng, N. Charness, and J. Bian, "Temporal Subtyping of Alzheimer's Disease Using Medical Conditions Preceding Alzheimer's Disease Onset in Electronic Health Records," *AMIA Jt. Summits Transl. Sci. Proc. AMIA Jt. Summits Transl. Sci.*, vol. 2022, pp. 226–235, 2022.
- [13] D. Tjandra, R. Q. Migrino, B. Giordani, and J. Wiens, "Cohort discovery and risk stratification for Alzheimer's disease: an electronic health record-based approach," *Alzheimer's Dement. N. Y. N.*, vol. 6, no. 1, p. e12035, 2020, doi: 10.1002/trc2.12035
- [14] W. Xu *et al.*, "Meta-analysis of modifiable risk factors for Alzheimer's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 86, no. 12, pp. 1299–1306, Dec. 2015, doi: 10.1136/jnnp-2015-310548
- [15] L. Peter-Derex, P. Yammine, H. Bastuji, and B. Croisile, "Sleep and Alzheimer's disease," *Sleep Med. Rev.*, vol. 19, pp. 29–38, Feb. 2015, doi: 10.1016/j.smrv.2014.03.007
- [16] M. V. F. Silva, C. de M. G. Loures, L. C. V. Alves, L. C. de Souza, K. B. G. Borges, and M. das G. Carvalho, "Alzheimer's disease: risk factors and potentially protective measures," *J. Biomed. Sci.*, vol. 26, no. 1, p. 33, May 2019, doi: 10.1186/s12929-019-0524-y
- [17] W. Xu *et al.*, "Alcohol consumption and dementia risk: a dose-response meta-analysis of prospective studies," *Eur. J. Epidemiol.*, vol. 32, no. 1, pp. 31–42, Jan. 2017, doi: 10.1007/s10654-017-0225-3
- [18] B. S. Diniz, M. A. Butters, S. M. Albert, M. A. Dew, and C. F. Reynolds, "Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies," *Br. J. Psychiatry J. Ment. Sci.*, vol. 202, no. 5, pp. 329–335, May 2013, doi: 10.1192/bjp.bp.112.118307
- [19] K. Y. Lai, C. Webster, S. Kumari, J. E. J. Gallacher, and C. Sarkar, "The associations of socioeconomic status with incident dementia and Alzheimer's disease are modified by leucocyte telomere length: a population-based cohort study," *Sci. Rep.*, vol. 13, no. 1, Art. no. 1, Apr. 2023, doi: 10.1038/s41598-023-32974-x
- [20] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows." arXiv, Aug. 17, 2021, doi: 10.48550/arXiv.2103.14030. Available: <http://arxiv.org/abs/2103.14030>. [Accessed: Nov. 20, 2023]
- [21] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," *arXiv.org*, Nov. 14, 2017. Available: <https://arxiv.org/abs/1711.05101v3>. [Accessed: Jan. 14, 2024]
- [22] H. Wang *et al.*, "Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks." arXiv, Apr. 13, 2020, doi: 10.48550/arXiv.1910.01279. Available: <http://arxiv.org/abs/1910.01279>. [Accessed: Jan. 15, 2024]