Machine learning and novel ophthalmologic biomarkers for Alzheimer’s disease screening: Systematic Review


1 Biostatistics and Informatics Unit, Department of Epidemiology, Clinical Research and Community Health, Faculty of Medicine and Pharmacy of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco
2 Laboratory of Artificial Intelligence, Data Science and Emerging Systems, National School of Applied Sciences Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco
3 Department of Epidemiology, Clinical Research and Community Health, Faculty of Medicine and Pharmacy of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco

*Corresponding author. Email: hbourkhime@gmail.com, noura.qarmiche@usmbe.ac.ma, mohammed.omari@usmbe.ac.ma, nassibabahra@gmail.com, nabil.tachfouti@usmbe.ac.ma, samira.elfakir@usmbe.ac.ma, nada.otmani@usmbe.ac.ma

ABSTRACT

Background: Alzheimer’s disease (AD) is a progressive neurodegenerative disease that leads to dementia and eventual death, the reason why screening is so beneficial in its early stages. Recent evidence suggests that memory and vision impairments are closely linked to Alzheimer’s disease. Moreover, assessing vision disorders may improve early detection and treatment of dementia. Therefore, some research has been conducted on screening for AD disease using new machine learning (ML) techniques on novel ophthalmologic biomarkers data.

Objective: To summarize existing findings on machine learning models exploring eye changes data to predict cognitive decline in the context of AD.

Methods: Systematic review of original research between January 2016 and August 2021. A search covered two databases on (Scopus) and (PubMed).

Results: From 104 search results, 13 articles were selected after using the eligibility criteria: 5 machine learning models used retinal texture data, 5 models included eye movement data, 2 proposed models used iris change data, and 1 proposed model used corneal nerve loss data.

Conclusion: Promising results are reported in almost all 13 studies, but very few have been implemented in research or clinical practice. The principal constraints in this area are limited standardization and comparability of results.

Keywords: Alzheimer’s disease, Machine learning, Screening, Early detection, Ophthalmologic biomarkers.
1. INTRODUCTION

Alzheimer’s disease (AD) is a fatal progressive brain disease and the most common cause of dementia. It begins with brain changes that are unnoticeable and only after many years that symptoms arise, such as memory loss and language problems.[1]

Therefore, developing accurate tools to screen its onset presents a real need. However, old and current tools are either invasive (cerebrospinal fluid analysis), costly (neuroimaging), or time-consuming (neuropsychological assessment), which hinders their accessibility for first-line AD screening.[2]

In the way to resolve that challenge, recent studies are focusing more on finding new biomarkers of AD, such as ophthalmologic ones, and new innovative techniques such those using machine learning (ML), which could be more useful, noninvasive, and cost-effective to identify subjects in AD early stages.[2] Indeed, the eye is anatomically considered as an extension of the brain and its examination can provide a lot of information on cognitive disorders, notably by exploring the morphology of the retina, the eye movements, the quality of the iris responses to light, etc...[2]

On the other hand, ML is a computer science aimed at training computers to learn and act without being specifically programmed, it involves building algorithms that adjust their models to improve their predictive ability [3].

Recently, there have been increasing efforts to apply ML techniques for classifying population most at risk for some diseases [3].

In this paper, we aimed to summarize existing findings on machine learning models exploring eye changes data to predict cognitive decline in the context of AD.

2. METHODS

A systematic review adhering to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PRISMA) statement 2020 (Page et al., 2021) [4] was conducted to identify studies reporting the development or attempted validation of a ML model to screen for AD using novel ophthalmologic biomarkers.

2.1. Information sources

We performed an exhaustive electronic literature search of English-language studies from PubMed and Scopus from 01/01/2016 to 08/31/2021.

Selected studies for review were manually searched to determine additional potentially pertinent studies in the reference lists.

2.2 Search strategy

The search strategy was based on the following equation:

- "Machine learning" AND Alzheimer AND (ophthalmologic OR retina OR retinal OR vision OR visual OR eye OR ocular OR optic) AND (Screening OR "early detection").

2.3. Eligibility criteria

2.3.1. Inclusion criteria

Articles were included based on passing all the selection criteria:

- Published as a primary research paper in a peer-reviewed journal and Conference papers only;
- Full paper available in English;
- Described either the development and validation, or a proposal, of a ML model to screen AD using ophthalmologic biomarkers;
- Described a neurodegeneration in the context of AD.
- Described a ML method for AD screening, detection, and prediction.
- Described a specific ophthalmologic biomarker such as: Retina texture, eye movement, iris changes and cornea nerve loss.

2.3.2. Exclusion criteria

Articles were excluded if they:

- Included other neurodegenerative diseases such as Parkinson's disease (unless comorbid with AD);
- Described an ML method for a distinctive classification of AD, to support diagnosis, etc;
- Exclusively included central visual system and network changes, brain imaging, brain atrophy, and visual brain atrophy rating scale;
- Included exclusively electroencephalographic recordings without eye tracking;
- Included the exclusive use of Visual Neuropsychological tests [5] with measure functions as: visual-motor coordination, visuospatial function, visual memory and attentive functions.
2.4. Data extraction

Data extraction was based on the characteristics of each model.

Three tables regarding the general characteristics of the study, country, number of participants, year of publication, algorithms applied and their characteristics, model prediction parameters, model performance, including accuracy, discrimination, sensitivity, and specificity rates.

2.5. Data Synthesis and Analysis

We summarize the data in narrative form, following the structure provided by the features summarized in each table.

3. RESULTS

3.1. Search results

The literature search identified 104 studies. Of these, 17 duplicates were removed. Then 74 articles were excluded because they did not meet the criteria. Only 13 were included after full-text analysis.

The PRISMA flow chart for the systematic review process is shown in Figure 1.

5 papers discussed ML models using retinal texture data. Another 5 papers included eye movement data in the ML models. There were also 2 model proposals on iris changes and 1 on corneal nerve loss in AD.

![](image)

**Figure 1** Flow diagram of process of systematic literature search in accordance with PRISMA guidelines.

3.2 Model Development and Validation

The ML algorithms used to screen the AD including retina texture data were a pipeline of convolutional neural network (CNN)+ supporting vector machine (SVM) in one [6], SVM in two [7] [8], SVM + linear discriminant analysis + random forest in one [9], and a ML wasn’t identified in one [10]. Table 1 shows the details of the development of these models.

The ML algorithms used to screen the AD including eye movement data were: SVM + Random Forrest + Neural Network in one [11], SVM+ linear and nonlinear methods + logistic regression in one [12], transfer learning + Deep CNN + Support Vector Regression in one [13], Decision Tree + Logistic Regression + Random Forest + Gradient Boosting in one [4], and a ML model wasn’t identified in one [15]. Table 2 shows the details of the development of these models.

The ML algorithms used to screen the AD including iris and pupil changes data were supervised learning (Naïve Bayes+ ZeroR + Multilayer Perceptron) in one [16] and supervised learning + unsupervised learning in one [17]. Table 3 shows the details of the development of these models.

The ML algorithms used to screen the AD including corneal neuropathy data was CNN+ Adaptive Neuro Fuzzy Inference System (ANFIS) [18]. Table 3 shows the details of the development of these models.

3.3. Study populations

3 studies [6] [10] [7] on retinal texture data used a cohort design to train the models. 2 other studies [8] [9] on retinal texture and all included studies [11] [12] [13] [14] [15] on eye movement data used a case-control design to train the models.

Studies [16] [17] [18] on iris changes and corneal neuropathy data were proposals with no available study design.

3.4. Features selection

5 models essentially used retinal imaging data in addition to demographic data, or clinical data, or brain imaging data.

5 models used essentially eye tracking and neuropsychological data in addition to demographic data, or clinical data, or EEG data, or brain imaging data, or laboratory data.

2 papers involved iris imaging.

And one proposed model used essentially confocal microscopy imaging of the cornea and clinical data.

https://doi.org/10.1051/itmconf/20224301009
3.5. Models Discrimination

The accuracy was reported for 12 models. Accuracy values ranged from 0.61 [16] to 0.95 in three models [9] [15] [18]. 9 of these models and the left one reported an area under the receiver operating characteristic curve between 0.632 [16] and 0.98 [9].

3.6. Models sensitivity and specificity

Sensitivity was reported in six models with values ranging from 0.74 [16] to 0.98 [7]. Specificity was reported for the same models, and their values ranged from 0.47 [16] to 0.97 [9].

4. DISCUSSION

To the best of our knowledge, this is the first systematic review of the ML use to screen early AD stages by investigating novel ophthalmologic biomarkers data.

It shows that 13 screening models exist for predicting the Alzheimer disease in early stages. The best discrimination and specificity model (0.95; 0.97 respectively) were reported by Qiu Y, 2020.

Given the heterogeneity of the models, we opted for a qualitative analysis of the results. Indeed, ocular biomarkers of AD are still in the early stages of discovery, which further clarifies the limited results found.

Eye-tracking tools and sensitive imaging modalities (such as fundus cameras and optical coherence tomography) are already widely available in eye clinics. Therefore, ophthalmic biomarker data are growing steadily and their study in combination with ML models is a promising, noninvasive, and inexpensive novel tool. However, the development and adaptation of these models face many challenges such as biomedical data privacy, data volume and quality...

Nevertheless, based on this process, we discovered that there is a need to standardize and validate the various ML model algorithms for forecasting the preclinical stage of AD, as well as to explore the importance of applying the predictive model to clinical practice.

We also found a lack of African-developed models, given the unique properties of the continent, building an African model is highly recommended, especially since individuals spend varying lengths of time in each part of the AD continuum [20].

Indeed, the duration of each phase of the continuum is influenced by age, genetics, gender, and other factors [20], making it more important to study these AD screening models, especially to select the perfect AD individuals for clinical trials.

4. CONCLUSION

Alzheimer's disease is a significant, increasing, and expensive problem. As more effective medication become available to treat this disease, better early detection tests will be needed. The use of ML models on new ophthalmologic biomarkers data can play a vital role in providing primary prevention measures.

In this systematic review, many models showed improved discrimination and accuracy, but further research is needed before these models can be incorporated into routine clinical practice. Advances in medical science and informatics should lead to further exploration for better prediction of AD.
<table>
<thead>
<tr>
<th>First author, Year, Reference</th>
<th>Country</th>
<th>Source of data</th>
<th>Study design</th>
<th>Main Variables</th>
<th>ML algorithms</th>
<th>Validation method, feature</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jianqiao Tian, 2021 [9]</td>
<td>USA</td>
<td>DRIVE database: UK Biobank database</td>
<td>Cohort: N=97 (122 fundus images of 67 AD)</td>
<td>Retinal vasculature, fundus images by OCT</td>
<td>Pipeline CNN++ SVM+ +saliency maps</td>
<td>five-fold cross validations, FS</td>
<td>Sensitivity=84.18% Specificity=86.64% Accuracy=82.44%</td>
</tr>
<tr>
<td>Sophie Lemmens, 2020 [10]</td>
<td>Belgium</td>
<td>not available</td>
<td>Cohort: N=39 (10 AD, 7 amyloid proven AD and 22 HC*)</td>
<td>Amyloid accumulation and retinal nerve fiber layer thickness (RNFL) (retinal imaging hyperspectral snapshot camera+OCT)</td>
<td>ML model not identified</td>
<td>leave-one-out cross validation, FS, DP</td>
<td>AUROC of 0.74</td>
</tr>
<tr>
<td>Ana Nunes, 2019 [9]</td>
<td>Portugal</td>
<td>Dementia and the Movement Disorders Units of the Neurological Department of the Centro Hospitalar e Universitario de Coimbra</td>
<td>Case-control: N=75 (20 AD, 28 PD*** and age-matched HC)</td>
<td>Retinal texture, Retinal thickness (fundus images by OCT), Demographic data</td>
<td>SVM</td>
<td>k-fold cross-validation, FS, DP</td>
<td>Sensitivity=88.7% Specificity=84.9% Accuracy=82.2%</td>
</tr>
<tr>
<td>Yunyi Cui, 2020 [9]</td>
<td>Canada</td>
<td>University of Waterloo, Waterloo, Ontario Canada</td>
<td>Case-control: N = 28 donors with AD high deposits retina imaging, cumulative score of AD with neuropathologic change, 4 fluorescence and with moderate score, 1 polarimetric microscopy with low score; all had (Mueller Matrix evidence of brain amyloid) [Polarimetry]</td>
<td>Retinal Amyloid deposits, Retinal imaging thickenings (fundus images with OCT), Retinal imaging by AD with moderate score, Retinal imaging by AD with low score, Retinal imaging by AD with high deposits</td>
<td>Linear discriminant analysis, and random forest, SVM</td>
<td>10-fold cross-validation, FS, DP</td>
<td>The random forest performance was the highest: Sensitivity= 94% Specificity= 97% Accuracy= 95% AUROC of 0.986</td>
</tr>
<tr>
<td>Benny Zee, 2021 [7]</td>
<td>China</td>
<td>The Chinese University of Hong Kong</td>
<td>Cohort: N=65 (a community-based cohort with participants, age 65 or above)</td>
<td>Retinal vasculature, retinal imaging by Canon non-mydriatic fundus camera and Topcon non-mydriatic retinal camera, Brain MRI, Vascular risk factors, Cardiovascular conditions</td>
<td>SVM</td>
<td>10-fold cross-validation, FS, DP</td>
<td>Sensitivity=98.8% Specificity=92.0% Accuracy=93.3% AUROC of 0.955</td>
</tr>
</tbody>
</table>

* Healthy controls ** Optical Coherence Tomography *** Parkinson Disease **** Magnetic resonance imaging
Table 2. Development details of ML models using eye movement data.

<table>
<thead>
<tr>
<th>First author, Year, Reference</th>
<th>Country</th>
<th>Source of data</th>
<th>Study design</th>
<th>Main Variables</th>
<th>ML algorithms</th>
<th>Validation method: feature detection (FS), data pre-processing (DP)</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marta Luísa Gonçalves de Freitas Pereira, 2020 [11]</td>
<td>Brazil</td>
<td>University hospital in São Paulo, Brazil</td>
<td>Case-control</td>
<td>Eye movement, demographic, neuropsychological data (esp. visuospatial abilities, visual attention, working memory)</td>
<td>SVM, Random Forest, Neural Network</td>
<td>10 k-fold cross validations, FS</td>
<td>Accuracy=70%, AUC=75%</td>
</tr>
<tr>
<td>Juanjuan Jiang, 2019 [12]</td>
<td>China</td>
<td>Shanghai University of Traditional Chinese Medicine</td>
<td>Case-control</td>
<td>Eye movement data, clinical data, brain imaging data, laboratory data</td>
<td>SVM, linear and nonlinear methods (logistic regression)</td>
<td>randomized cross-validation, FS, DP</td>
<td>Sensitivity=78.67%, Specificity=84.10%, Accuracy=81.51%, AUROC=0.9216</td>
</tr>
<tr>
<td>Rafi U Haque, 2020 [13]</td>
<td>USA</td>
<td>Emory Healthy Brain Study (EHSB) and the Goizueta Alzheimer Disease Research Center (ADRC) at Emory</td>
<td>Case-control</td>
<td>The best model: Eye movement (Eyetracking variables), clinical data, EEG data</td>
<td>transfer learning+CNN, SVR</td>
<td>5-fold cross-validation, FS</td>
<td>AUC=76%</td>
</tr>
<tr>
<td>Ivanna M Pavsic, 2017 [15]</td>
<td>United Kingdom</td>
<td>United Kingdom</td>
<td>Case-control</td>
<td>Eye movement, demographic data</td>
<td>ML model not identified</td>
<td>Cross-validation</td>
<td>Accuracy=95%</td>
</tr>
<tr>
<td>Benny Zee, 2021 [14]</td>
<td>Slovenia</td>
<td>several centres in Ljubljana in Slovenia</td>
<td>Case-control</td>
<td>Eye movement</td>
<td>Decision Tree; Logistic Regression; Random Forest; Gradient Boosting</td>
<td>10-fold cross-validation</td>
<td>The random forest performance was the higher Accuracy=80%, AUROC=0.85</td>
</tr>
</tbody>
</table>

* Healthy controls  **Mild cognitive impairment  ***Young Onset Alzheimer’s Disease
<table>
<thead>
<tr>
<th>First author, Year, Reference</th>
<th>Country</th>
<th>Source of data</th>
<th>Study design</th>
<th>Main Variables</th>
<th>ML algorithms</th>
<th>Validation method, feature detection (FS), data preprocessing (DP)</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tooba Salauuddin, 2019 [19]</td>
<td>Qatar</td>
<td>A local hospital in Qatar</td>
<td>not available</td>
<td>Neuropathy or nerve loss (Corneal microscopic confocal clinical data)</td>
<td>CNN+ Adaptive Neuro Fuzzy Inference System (ANFIS)</td>
<td>FS</td>
<td>Accuracy= 95%</td>
</tr>
<tr>
<td>Fernando Hernández, 2018 [16]</td>
<td>Ecuador</td>
<td>Foundation in the city of Quito</td>
<td>not available</td>
<td>Iris and pupil changes (Iris imaging)</td>
<td>supervised learning (Naive Bayes+ZeroR+Multilayer Perceptron)</td>
<td>cross-validation, DP</td>
<td>The Naive Bayes performance was the higher: Sensitivity= 74 % Specificity= 47.6191 % Accuracy= 61.9555 % AUC of 0.632</td>
</tr>
<tr>
<td>Fernando Hernández, 2019 [17]</td>
<td>Ecuador</td>
<td>Foundation in the city of Quito</td>
<td>not available</td>
<td>Iris and pupil changes (Iris imaging)</td>
<td>supervised learning (Naive Bayes+ZeroR+Multilayer Perceptron)</td>
<td>cross-validation, DP</td>
<td>The Naive Bayes performance was the higher: Sensitivity= 74 % Specificity= 47.6191 % Accuracy= 61.9555 % AUC of 0.632</td>
</tr>
</tbody>
</table>
REFERENCES


