



OPEN ACCESS

EDITED BY

Bingyao Tan,
Nanyang Technological University, Singapore

REVIEWED BY

Christophe Orssaud,
Georges Pompidou European Hospital, France
Yi-Ting Hsieh,
National Taiwan University Hospital, Taiwan

*CORRESPONDENCE

Givago Silva Souza
✉ givagosouza@ufpa.br

RECEIVED 09 August 2023

ACCEPTED 27 November 2023

PUBLISHED 14 December 2023

CITATION

Farias FM, Salomão RC, Rocha Santos EG,
Sousa Caires A, Sampaio GSA, Rosa AAM,
Costa MF and Silva Souza G (2023) Sex-related
difference in the retinal structure of young
adults: a machine learning approach.
Front. Med. 10:1275308.
doi: 10.3389/fmed.2023.1275308

COPYRIGHT

© 2023 Farias, Salomão, Rocha Santos, Sousa
Caires, Sampaio, Rosa, Costa and Silva Souza.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Sex-related difference in the retinal structure of young adults: a machine learning approach

Flávia Monteiro Farias^{1,2}, Railson Cruz Salomão²,
Enzo Gabriel Rocha Santos³, Andrew Sousa Caires¹,
Gabriela Santos Alvarez Sampaio¹,
Alexandre Antônio Marques Rosa⁴, Marcelo Fernandes Costa⁵
and Givago Silva Souza^{1,2*}

¹Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil, ²Núcleo de Medicina Tropical, Universidade Federal do Pará, Belém, Brazil, ³Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, Brazil, ⁴Instituto de Ciências da Saúde, Universidade Federal do Pará, Belém, Brazil, ⁵Departamento de Psicologia, Instituto de Psicologia, Universidade de São Paulo, São Paulo, Brazil

Purpose: To compare the accuracy of machine learning (ML) algorithms to classify the sex of the participant from retinal thickness datasets in different retinal layers.

Methods: This cross-sectional study involved 26 male and 38 female subjects. Data were acquired using HRA + OCT Spectralis, and the thickness and volume of 10 retinal layers were quantified. A total of 10 features were extracted from each retinal layer. The accuracy of various algorithms, including k-nearest-neighbor, support vector classifier, logistic regression, linear discriminant analysis, random forest, decision tree, and Gaussian Naïve Bayes, was quantified. A two-way ANOVA was conducted to assess the ML accuracy, considering both the classifier type and the retinal layer as factors.

Results: A comparison of the accuracies achieved by various algorithms in classifying participant sex revealed superior results in datasets related to total retinal thickness and the retinal nerve fiber layer. In these instances, no significant differences in algorithm performance were observed ($p > 0.05$). Conversely, in other layers, a decrease in classification accuracy was noted as the layer moved outward in the retina. Here, the random forest (RF) algorithm demonstrated superior performance compared to the others ($p < 0.05$).

Conclusion: The current research highlights the distinctive potential of various retinal layers in sex classification. Different layers and ML algorithms yield distinct accuracies. The RF algorithm's consistent superiority suggests its effectiveness in identifying sex-related features from a range of retinal layers.

KEYWORDS

retina, retinal thickness, macula, machine learning, sex-related differences

Introduction

Over the past 30 years, optical coherence tomography (OCT) has been used as a non-invasive method for image acquisition to evaluate the anterior and posterior segments of the eye in both diseased and healthy conditions of the human retinal structure (1–3). The development of eye diseases can occur throughout life due to the natural aging process, exposure to unhealthy lifestyle habits, systemic disorders, or genetic inheritance. In addition to these factors, sex-related factors, such as the concentrations of sex hormones that vary throughout an individual's life, can also influence the development of eye diseases (4–6).

The existence of sexual dimorphism of the retina in humans has been investigated using OCT. The first findings of retinal sexual dimorphism pointed to a larger total retinal thickness in male subjects than in female subjects (7–12). However, the debate regarding retinal layers remains open, as some studies have observed that some retinal layers are thicker in male subjects than in female subjects, while other investigations have found no or few sex-related differences (13–19).

Overall, investigating sex-related features in the human retina is an important area of research that could lead to new insights into the causes of retinal diseases, the development of sex-specific treatments, and the design of more effective medical devices for the eye and the possible impact of postmenopausal hormone replacement anti-estrogenic therapy (20).

Due to the large amount of data extracted from the retina during an OCT scan, the use of machine learning methods could be an alternative candidate for analyzing OCT data. Machine learning methods have been used due to their ability to capture complex relationships, work with high-dimensional data, generalize to new data, be flexible and adaptable, and perform automated learning of relevant features, reducing the need for human intervention (21–23). Compared to norms based on populational averages, which may not account for the significant individual variability that exists within each sex group, machine learning models can capture and leverage this variability, allowing for more precise and individualized assessments of retinal thickness. This individualized precision can be particularly valuable in clinical decision-making as it takes into account the uniqueness of each patient's condition. Additionally, retinal thickness datasets can exhibit complex patterns and subtle variations that may not be fully captured by simple norm-based criteria.

In the present study, we aimed to evaluate the performance of several machine learning algorithms to predict the sex of the participants based on information from retinal structure features. Our primary goal was to identify which retinal layers are best to correctly classify the sex of the participant and which machine learning algorithms are better for predicting the participant's sex in the different retinal layers.

Materials and methods

Ethical considerations

The present study was approved by the Ethical Committee for Research in Humans of the Universidade Federal do Pará (report number 3.285.557). All participants were informed

about the experimental procedures and gave written consent to participate in the study.

Participants

The sample consisted of 26 male participants (mean age \pm standard deviation: 26.19 ± 4.96 years) and 38 participants (mean age \pm standard deviation: 26.05 ± 4.68 years). All participants had normal visual acuity or were corrected to 20/20 visual acuity using a refractive lens. Only two participants (one male and one female) used optical corrections of -0.5 and -0.7 diopters and we considered that any imprecision of their OCT measurements had little or no influence on the results. Participants with neurological, systemic, eye, or retinal diseases that affected the structure or function of the visual system were excluded.

OCT imaging

Retinal OCT imaging was performed using the Spectralis HRA + OCT system (Heidelberg Engineering GmbH, Heidelberg, Germany). Each session consisted of a 25-line horizontal raster scan in a $20^\circ \times 20^\circ$ area centered on the fovea, followed by 24 automated real-time repetitions. The Heidelberg Eye Explorer software (Heidelberg Engineering GmbH, Heidelberg, Germany) was used to segment retinal layers [total retina (TR), retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), and retinal pigmented epithelium (RPE)] and three combinations of retinal layers [overall retinal, outer retinal layers (ORL), which range from the external limiting membrane to Bruch's membrane, and inner retinal layers (IRL), which range from the inner limiting membrane to the external limiting membrane]. The thickness and volume of each layer were quantified. Visual inspection of the segmentation was performed to avoid possible errors. The outcome of the image segmentation of retinal layers was the mean thickness of nine macular subfields (central, nasal inner, temporal inner, superior inner, inferior inner, nasal outer, temporal outer, superior outer, and inferior outer), following the Early Treatment Diabetic Retinopathy Study (ETDRS) grid. The volume of each layer was also extracted.

For each participant, the examination was performed by the same operator following the manufacturer's guidelines. Two images were obtained in sequence for each eye on the same day. The first image was used as a reference to scan the same parts of the retina during the second image (device's follow-up mode). The thickness of both images was averaged for subsequent analysis. Data were acquired from 128 eyes with the Spectralis HRA + OCT system, and 64 eyes were randomly selected for analysis.

Machine learning algorithms

Prior to the application of ML algorithms, a bootstrap resampling method was employed, utilizing 200 replications for each feature derived from OCT readings. A total of 10 features were used for each retinal layer, comprising nine subfield thicknesses and the volume of the retinal layer. Python scripts were utilized for data

analysis and normalization, feature selection, and the execution of ML algorithms through the training and testing phases. The performance of the ML was subsequently evaluated.

We utilized the *StandardScaler* function from the *sklearn*. Preprocessing package to standardize the features into standard deviation units, as shown in Equation 1.

$$\text{Standardized_feature} = (\text{feature} - \text{mean}) / \text{standard_deviation} \quad (\text{Equation 1})$$

The standardized features were used to train and test seven supervised ML algorithms:

The *sklearn.neighbors.KNeighborsClassifier* function was employed to implement the k-nearest neighbors (kNN) algorithm, utilizing the Minkowski distance and a k-value within the range of 5–10. The optimal k-value, which yielded the highest accuracy, was determined using the *GridSearchCV* function.

The support vector classifier (SVC) utilizes *sklearn.svm.SVC* function with the radial basis function kernel. The gamma and C parameters are set to 1 and 10, respectively.

The *sklearn.linear_model.LogisticRegression* function is utilized for logistic regression (LR), with the parameters “*penalty*” and “*solver*” set to “*l1*” and “*liblinear*,” respectively.

The *sklearn.discriminant_analysis.LinearDiscriminantAnalysis* function is utilized for linear discriminant analysis. The parameters “*solver*” and “*store_covariance*” are set to “*svd*” and “*true*,” respectively.

The *sklearn.ensemble.RandomForestClassifier* function is utilized in the application of random forest (RF), with the parameters set as follows: “*criterion*” is set to “*gini impurity*,” “*n_estimators*” is set to 50, and “*max_depth*” is set to 6.

The decision tree (DT) employs the *sklearn.tree.DecisionTreeClassifier* function, maintaining identical parameter values for “*criterion*,” “*n_estimators*,” and “*max_depth*,” as utilized in the RF algorithm.

Gaussian Naïve Bayes (GNB) using the *sklearn.naive_bayes.GaussianNB* function.

The accuracy of ML algorithms in correctly classifying the data was evaluated (Equation 2).

$$\text{Accuracy} = (\text{true positives} + \text{true negatives}) / \text{total} \quad (\text{Equation 2})$$

True positives represent the data points correctly classified as male, while true negatives denote those accurately identified as female. The total refers to the overall number of data points.

The *ShuffleSplit* function from the *Scikit-learn* library (version 0.21.3) was utilized to divide the data, allocating 70% for model training and 30% for model testing.

Statistics

We used a *t*-test to compare the thickness of the different datasets obtained from both eyes of male and female subjects and to later carry out an intergroup comparison of retinal layer thickness. We conducted a one-way ANOVA to evaluate the influence of macular field in the retinal thickness as well as two-way ANOVA to evaluate the influence of the classifier type and retinal dataset factors on the accuracies (model training and model testing) of the classifier. For multiple comparisons, we employed the Tukey HSD *post-hoc* test. We compared the accuracies of the model training and model testing using a *t*-test for repeated measures. A confidence level of 5% was applied for the statistical comparisons.

Results

Inter-eye comparison of the retinal thickness for male and female subjects

To ensure that the selection of the eye did not introduce any bias, we conducted a comparison of the thickness of various retinal layers between the right and left eyes of participants of both sexes. Our analysis revealed that no significant differences were observed in any of the retinal layers between the eyes. Based on these findings, we opted to randomly select one eye from each participant for data extraction concerning retinal thickness. **Table 1** displays the comparison of retinal thickness in the various datasets obtained from both eyes within the sample.

We randomly select one eye to extract retinal thickness and compared this feature between male and female groups, as depicted in **Table 2**. Our findings indicated significant differences in the total

TABLE 1 Comparison of the retinal thickness obtained from both eyes of male and female groups.

Thickness (μm)	Male group		p-value	Female group		p-value
	Right eye	Left eye		Right eye	Left eye	
TR	323.6 ± 10.8	322.3 ± 10.8	0.65	311.1 ± 12.7	309.8 ± 13.4	0.72
RNFL	28 ± 2	27.3 ± 2.2	0.22	25.9 ± 2.3	25.3 ± 2	0.33
GCL	43.4 ± 2.4	43.3 ± 2.4	0.87	40.4 ± 3.6	40.3 ± 3.7	0.89
IPL	36.1 ± 1.6	36.1 ± 1.7	0.86	34 ± 2.7	34.2 ± 2.8	0.88
INL	35.6 ± 2.1	35.7 ± 2.1	0.91	33.5 ± 2.6	33.3 ± 2.3	0.79
OPL	28.4 ± 2.7	29.3 ± 2.8	0.23	28.8 ± 4.1	29.06 ± 4.1	0.81
ONL	69.2 ± 8.1	68.5 ± 8.3	0.76	66.6 ± 6.4	66.4 ± 7.1	0.90
RPE	15 ± 1.1	15 ± 1	0.93	14.9 ± 1.2	14.64 ± 1.4	0.45

TR, total retina; RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; RPE, retinal pigmented epithelium.

TABLE 2 Comparison of retinal layer thickness between male and female groups.

Retinal layers	Retinal thickness (μm)		p -value
	Male group	Female group	
TR	322.9 \pm 11.3	310.6 \pm 12.9	0.0005
RNFL	27.8 \pm 2.1	25.3 \pm 1.8	<0.0001
GCL	43.2 \pm 2.2	40.4 \pm 3.5	0.0012
IPL	36 \pm 1.6	34.1 \pm 2.8	0.0039
INL	35.6 \pm 2.2	33.4 \pm 2.5	0.0016
OPL	29.1 \pm 2.6	29.3 \pm 4.2	0.78
ONL	68.7 \pm 8.3	66.4 \pm 6.8	0.28
RPE	15.1 \pm 1	14.7 \pm 1.2	0.22

TR, total retina; RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; RPE, retinal pigmented epithelium.

retina and layers comprising information from the inner retina (RNFL, GCL, IPL, INL), with the male group exhibiting greater thickness compared to the female group ($p < 0.01$). Conversely, no significant differences were discerned in the layers within the outer retina (OPL, ONL, RPE; $p > 0.01$).

In the intergroup comparison, considering the thickness of different macular fields (Table 3), we observed that in datasets representing the total retina and data from the inner retina, the male group had thicker tissues across all fields than the female group ($p < 0.01$). However, in the datasets from the outer retina, we observed a predominance of non-significant differences.

Machine learning accuracies during model training

Table 4 presents the mean accuracies (\pm standard deviation) derived from model training across various classifiers and retinal datasets. The results of a two-way ANOVA revealed significant effects attributed to both the algorithm factor, the retinal dataset factor, and the interaction between these two factors, as summarized in Table 5. Notably, *post hoc* multiple comparisons demonstrated that the accuracies achieved by all algorithms were markedly superior when utilizing the total retina dataset and datasets originating from the inner retina (RNFL, GCL, IPL, INL), as compared to datasets from the outer retina (OPL, ONL, and RPE).

In evaluating the accuracies of different algorithms across the diverse retinal datasets, multiple comparisons indicated a notable absence of significant differences in algorithm performance within the total retina dataset and the inner retina datasets ($p > 0.05$). Conversely, in the OPL dataset, it was evident that random forest (RF), support vector classifier (SVC), and decision tree (DT) exhibited significantly higher accuracies when contrasted with other algorithms. Similarly, in the ONL dataset, random forest and decision tree outperformed their counterparts. Notably, in the RPE dataset, random forest demonstrated the highest accuracy among all algorithms.

Machine learning accuracies during model testing

Table 6 displays the mean accuracies (\pm standard deviation) derived from model testing across various classifiers and retinal datasets. Once again, the results of a two-way ANOVA revealed significant effects associated with the algorithm factor, the retinal dataset factor, and their interaction (as summarized in Table 7). *Post hoc* multiple comparisons further substantiated that, much like the training model, all algorithms achieved significantly higher accuracy levels when employing the total retina dataset and datasets from the inner retina, in comparison to the datasets from the outer retina. Consistent with the training model, the results of multiple comparisons within the total retina dataset and datasets from the inner retina indicated an absence of significant differences in algorithm accuracies ($p > 0.05$). In contrast, concerning the outer retina, random forest (RF) exhibited notably higher accuracy compared to other algorithms ($p < 0.05$).

Comparison of the accuracies estimated for the models in the training and testing stages

The comparison of the accuracies calculated for the models in the training and testing showed that 10.8% of the comparisons had significant differences, and all of them showed higher accuracy of the model in the training (Figure 1).

After finding that the random forest classifier outperformed other methods in classifying the datasets, we examined feature importance scores, which indicate the extent to which each feature influences the model's predictions. Random forest employs the Gini impurity, which reveals how frequently a feature is used to split the data in its decision trees. Figure 2 displays the feature importance scores for macular thickness in different fields. We conducted one-way ANOVA to assess the impact of the macular field on the feature importance score for each dataset. We found that in all datasets, there were significant differences ($p < 0.01$), with one or more fields having a greater importance than others in the classification decision.

Discussion

This study's findings reveal significant patterns in the classification accuracy of sex-specific data, utilizing various retinal layers and ML algorithms. The most reliable accuracies for accurately distinguishing between male and female participants were observed when analyzing data from the total retinal structure and the retinal nerve fiber layer. These results suggest that these retinal layers possess unique sex-related characteristics that were effectively identified by the employed ML techniques. Interestingly, the highest classification accuracies were consistently achieved using these retinal layers, yet no statistically significant differences were detected among the accuracies derived from the various ML algorithms used in this study. This suggests that the algorithms consistently performed when tasked with sex classification based

TABLE 3 Comparison of retinal dataset thickness in the different macular fields from measurements obtained from both groups.

TR	Retinal thickness (μm)						
	Male	Female	p-value	RNFL	Male	Female	p-value
C	274.9 ± 15.9	258.6 ± 18.9	<0.01*	C	12 ± 1.5	10.3 ± 2.1	<0.01*
NI	353.1 ± 12.8	338.4 ± 16.9	<0.01*	NI	21.1 ± 2.1	18.8 ± 1.4	<0.01*
NO	328.8 ± 13.3	318 ± 14.0	0.04*	NO	48.6 ± 5.6	44.8 ± 4.5	<0.01*
TI	336.5 ± 10.3	320.6 ± 14.7	<0.01*	TI	16.9 ± 1.4	16 ± 0.9	<0.01*
TO	297.6 ± 13.9	283.3 ± 11.4	<0.01*	TO	18.8 ± 1.2	17.7 ± 1	<0.01*
SI	353 ± 12.8	340 ± 16.2	<0.01*	SI	24.5 ± 2.9	21.9 ± 2.4	<0.01*
SO	313.8 ± 13.3	305.7 ± 13.2	0.02*	SO	39.5 ± 5	36.6 ± 4.1	0.02*
II	351.1 ± 11.7	336.4 ± 16.8	<0.01*	II	26.1 ± 2.6	23.8 ± 2.2	<0.01*
IO	300.8 ± 10.8	291.9 ± 12.5	0.02*	IO	41.8 ± 4.8	39 ± 4.3	0.02*
GCL	Male	Female	p-value	IPL	Male	Female	p-value
C	15.4 ± 2.8	12.9 ± 4.1	<0.01*	C	21.2 ± 2.6	19.18 ± 3.1	<0.01*
NI	54.7 ± 3.9	50.7 ± 5.9	<0.01*	NI	45.2 ± 3	41.6 ± 4.7	<0.01*
NO	41.7 ± 3.5	40.7 ± 3.3	0.45	NO	31.9 ± 2.5	31.5 ± 3	0.51
TI	51.1 ± 3.4	46.3 ± 5.7	<0.01*	TI	43.7 ± 2.3	41.9 ± 4.3	0.02*
TO	41.2 ± 3.8	36.9 ± 3.5	<0.01*	TO	34.4 ± 2.6	32.3 ± 2.3	<0.01*
SI	55.4 ± 3.2	52.3 ± 5.7	0.01*	SI	44 ± 2.2	41.4 ± 3.4	<0.01*
SO	38.1 ± 2.8	37.3 ± 3.2	0.168	SO	31.1 ± 2.3	30.3 ± 2.5	0.15
II	55.5 ± 2.6	53.3 ± 5.3	0.02*	II	43.6 ± 1.9	41.2 ± 4.1	0.04*
IO	36.2 ± 3.2	35.3 ± 3.5	0.34	IO	29.5 ± 2	28.6 ± 2.5	0.21
INL	Male	Female	p-value	OPL	Male	Female	p-value
C	17.8 ± 4.1	13.9 ± 4.4	<0.01*	C	21.9 ± 4	22.1 ± 5.8	0.92
NI	41.59 ± 4	38.3 ± 4	<0.01*	NI	33 ± 6.5	32.1 ± 6.5	0.66
NO	37 ± 1.7	36.8 ± 2.8	0.4	NO	29.8 ± 3.1	29 ± 4.4	0.42
TI	37.7 ± 3.3	34.9 ± 3.4	0.01*	TI	29.3 ± 2.4	31.5 ± 7.1	0.20
TO	36.7 ± 2.2	34.8 ± 2.5	<0.01*	TO	27.3 ± 2	27.24 ± 3.8	0.95
SI	40.7 ± 3.2	39.3 ± 3.7	0.07	SI	32.4 ± 5.2	36.4 ± 9.5	0.11
SO	33.8 ± 2	32.9 ± 2.4	0.39	SO	27 ± 1.7	28.3 ± 4.8	0.23
II	42.2 ± 3.5	39.2 ± 4.3	<0.01*	II	32 ± 5.1	31.3 ± 5.4	0.46
IO	33.4 ± 2.3	31.8 ± 2.4	0.09	IO	26.8 ± 2.3	26.3 ± 2.8	0.23
ONL	Male	Female	p-value	RPE	Male	Female	p-value
C	93.6 ± 9	91 ± 11.63	0.43	C	18.3 ± 1.5	17.3 ± 1.7	0.21
NI	72.3 ± 13	71.1 ± 11.7	0.99	NI	16.2 ± 1.2	16 ± 1.6	0.92
NO	57.4 ± 8.8	58.7 ± 6.7	0.88	NO	14 ± 1.3	13.8 ± 1.4	0.77
TI	75.3 ± 7.8	71.1 ± 8	0.04*	TI	15.1 ± 1.3	14.4 ± 1.3	0.07
TO	60.8 ± 7.2	57.4 ± 6	0.09	TO	13.3 ± 1.2	13.2 ± 1.2	0.27
SI	71.3 ± 10.7	67.1 ± 11	0.07	SI	16 ± 1.5	15.7 ± 1.6	0.50
SO	64.1 ± 8.2	60.5 ± 6.8	0.15	SO	14.1 ± 1.2	13.6 ± 1.5	0.24
II	68.5 ± 10.3	68 ± 8.6	0.88	II	15.4 ± 1	15 ± 1.2	0.24
IO	54 ± 6.3	53 ± 5.5	0.68	IO	13.6 ± 1.2	13.1 ± 1.1	0.21

TR, total retina; RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; RPE, retinal pigmented epithelium. C, central retina; NI, nasal inner; NO, nasal outer; TI, temporal inner; TO, temporal outer; SI, superior inner; SO, superior outer; II, inferior inner; IO, inferior outer. *p < 0.05.

on retinal data, regardless of their inherent methodologies. Moreover, a fascinating trend was observed where classification accuracies showed a decreasing trend as the analysis moved toward the outer retinal layers. Additionally, some algorithms demonstrated statistically significant deviations from others in terms of classification accuracy. Notably, the RF algorithms displayed higher accuracies compared to the others in this context.

While the sex of a patient is typically known during a consultation, it is not always evident whether the retinal thickness of that patient aligns with the sex-based patterns expected. Comparing a patient's retinal thickness to sex-based populational norms can be a valuable tool in evaluating the patient. However, alternative approaches, such as machine learning, can complement conventional statistical methods. For instance, our study revealed

TABLE 4 Comparison of mean accuracies (\pm standard deviation) obtained from the machine learning algorithms to classify the sex-related differences in the retinal layers (and total retina) for model training.

Layers	Machine learning algorithm accuracies (%)						
	SVC	GNB	RF	KNN	LR	LD	DT
RPE	72.5 \pm 9.6	71.8 \pm 5.1	83 \pm 4.5	64.3 \pm 5.5	68.5 \pm 6.2	69.3 \pm 7.6	71.3 \pm 9.9
ONL	69.3 \pm 6.5	67.8 \pm 7.1	90.5 \pm 5.5	68.5 \pm 9.4	67.5 \pm 8.7	65.3 \pm 6.3	82.8 \pm 5.2
OPL	79.8 \pm 7.5	70.5 \pm 4.1	89.5 \pm 5.8	71.8 \pm 5	65.5 \pm 7.5	66.8 \pm 7.8	79.8 \pm 6.8
INL	84.5 \pm 4.5	89 \pm 3.4	92.3 \pm 3.8	81.5 \pm 6.5	85 \pm 6.2	85.8 \pm 6.7	81.5 \pm 6
IPL	85.8 \pm 3.3	88 \pm 3.3	92 \pm 5	83.8 \pm 5.4	84.5 \pm 6.1	82.3 \pm 5.3	82.3 \pm 7.2
GCL	89.5 \pm 3.7	90 \pm 4.3	94.5 \pm 2	89.8 \pm 4.3	85.5 \pm 4.7	87.5 \pm 4.3	88 \pm 3.9
RNFL	92.3 \pm 2.8	91.5 \pm 3.8	94.5 \pm 2.8	91 \pm 4.6	90.5 \pm 3.9	91 \pm 3.2	90.8 \pm 4.7
TR	92.5 \pm 2.3	96 \pm 2.4	96.5 \pm 3.9	91 \pm 4.6	93.5 \pm 5.3	93.3 \pm 3.9	89 \pm 4.3

TR, total retina; RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; RPE, retinal pigmented epithelium; SVC, support vector classification; GNB, Gaussian Naïve Bayes; RF, random forest; kNN, k-nearest neighbors; LR, logistic regression; LD, linear discriminant; DT, decision tree.

TABLE 5 Two-way ANOVA results for model training.

	SS	DF	MS	F	p-value
Algorithms	7,937	6	1,322.9	42.87	< 0.001
Retinal datasets	37,140	7	5,305.7	171.93	< 0.001
Algorithms*retinal datasets	6,703	42	159.6	5.17	< 0.001
Residuals	15,554	504	30.9		

SS, sum of squares; DF, degrees of freedom; MS, mean squares.

TABLE 6 Comparison of mean accuracies (\pm standard deviation) obtained from the machine learning algorithms to classify the sex-related differences in the retinal layers (and total retina) for model testing.

Layers	Machine learning algorithms (%)						
	SVC	GNB	RF	KNN	LR	LD	DT
RPE	68.8 \pm 9.5	70.8 \pm 5.8	76.8 \pm 3.7	59.8 \pm 5.2	66 \pm 5.6	66.5 \pm 7.3	69.5 \pm 6
ONL	70.5 \pm 8.8	63.5 \pm 6.7	89.3 \pm 5.4	68 \pm 6.5	64.3 \pm 5.9	66.8 \pm 10	75.5 \pm 6
OPL	76.3 \pm 4.3	74.5 \pm 7.9	88.5 \pm 3.8	72 \pm 5.2	64.5 \pm 9	62.8 \pm 7.2	83 \pm 6.2
INL	82.8 \pm 6.3	85.8 \pm 5.5	88.8 \pm 4	80.5 \pm 6.3	77.5 \pm 6.6	80.8 \pm 5.3	85.3 \pm 7
IPL	85.3 \pm 6.6	83.3 \pm 5.5	93.8 \pm 2.4	80.5 \pm 5.1	82.3 \pm 5.8	84.3 \pm 4.3	85.8 \pm 7.1
GCL	89.3 \pm 4.4	82 \pm 4.1	92.3 \pm 4.2	84 \pm 7.4	86.3 \pm 5.3	88.8 \pm 5	89.8 \pm 4.4
RNFL	91.5 \pm 3.8	91 \pm 3.9	94.5 \pm 4.7	92.3 \pm 4.6	90.3 \pm 5.1	91.3 \pm 4.6	90 \pm 5.3
TR	91.8 \pm 2.7	93.3 \pm 3.7	94 \pm 4.1	89.5 \pm 5.1	91.3 \pm 3.6	91.5 \pm 5.7	91.3 \pm 4.3

TR, total retina; RNFL, retinal nerve fiber layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; RPE, retinal pigmented epithelium; SVC, support vector classification; GNB, Gaussian Naïve Bayes; RF, random forest; kNN, k-nearest neighbors; LR, logistic regression; LD, linear discriminant; DT, decision tree.

that, even in retinal layers where there were no significant differences in thickness between the male and female groups, such as the datasets from the outer retina, we achieved a sex classification accuracy exceeding 75%. What would it signify if a male patient were classified as female based on retinal thickness patterns, or vice versa? It is crucial to emphasize that this classification does not pertain to the patient’s actual sex but rather reflects the retinal thickness patterns expected for each sex. The clinical implications of a disparity between a patient’s actual sex and a different sex classification based on retinal structure remain unclear, but further investigations may shed light on this question.

An investigation has previously been conducted using a deep learning method to predict sex through macular OCT images (24). It showed that the differences between male and female subjects might not be uniform throughout the macula. The best accuracy in separating data from male and female subjects occurred in the central fovea (around 75%) and lower accuracy was found in the external limit of the fovea (around 70%). They also fed models considering different macular sectors and found non-uniformity in the accuracies (ranging between 52 and 62%). The data they used are comparable to the total retina dataset of the present study. We interpreted that our accuracies were higher because we had fed our models with thickness information of all the macular sectors, and they used information from each sector for their classification.

TABLE 7 Two-way ANOVA results for model testing.

	SS	DF	MS	F	p-value
Algorithms	8,299	6	1,383.1	41.91	<0.001
Retinal datasets	41,178	7	5,882.6	178.25	<0.001
Algorithms*retinal datasets	6,276	42	149.4	4.53	<0.001
Residuals	16,633	504	33		

SS, sum of squares; DF, degrees of freedom; MS, mean squares.

Taking into account the significance of macular field thickness, our results align with the findings achieved using deep learning approaches for the total retinal dataset, wherein the temporal fields were identified as the most crucial for classifying sex. The current study also revealed that in other retinal layers, the field of greatest importance varied.

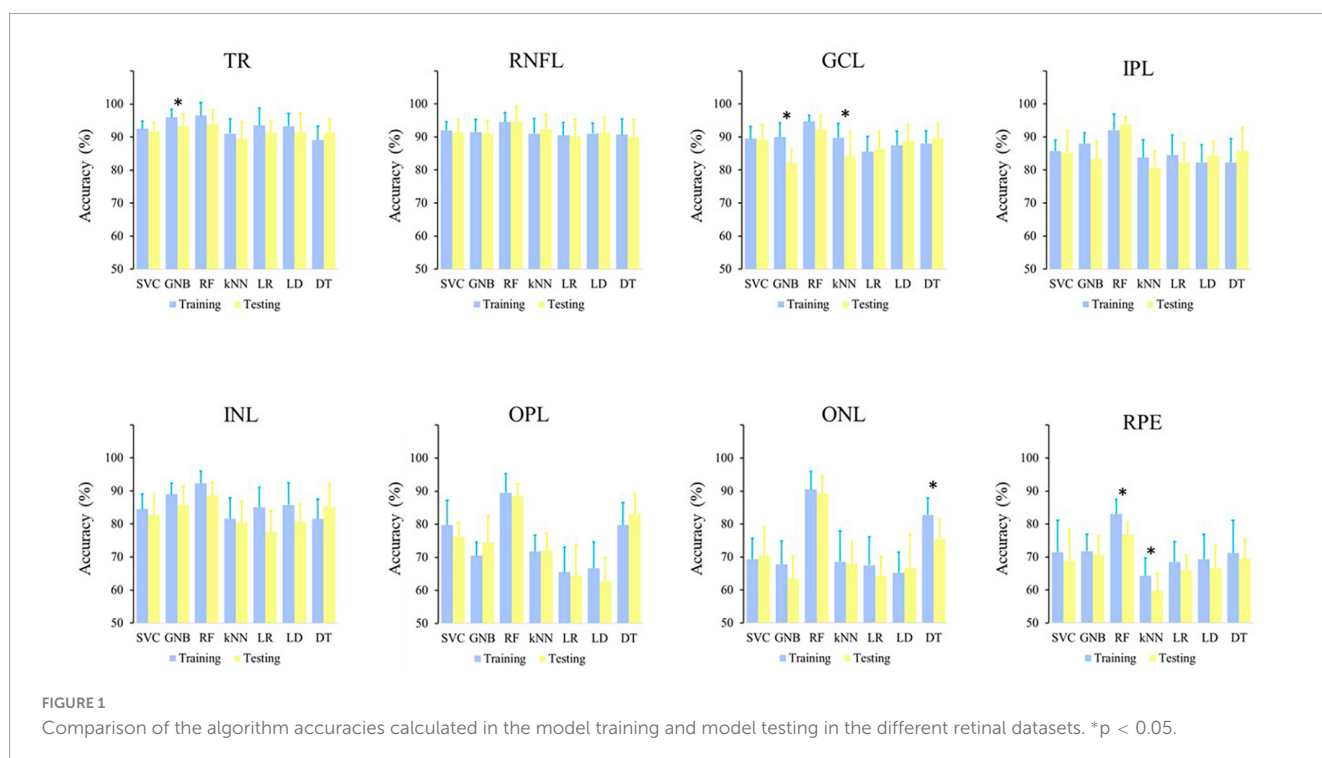
The difference between the accuracies of the training and testing models is a crucial aspect in the evaluation of machine learning models. This difference can provide insights into how well the model is generalizing to unseen data, which is essential for determining the model’s robustness. In the current study, the vast majority of comparisons showed no significant discrepancy between the training and testing accuracies, which is a positive indication. It suggests that the model, which fits the training data well, also exhibits good generalization to new data. This alignment between training and testing accuracies indicates that the model is not overfitting the training data and has the potential for reliable performance on new, unseen data.

The superior performance of random forest in achieving higher accuracies compared to alternative machine learning algorithms in our study can be attributed to several key advantages of this ensemble learning technique. random forest

harnesses the power of multiple decision trees, where each tree is trained on a different subset of the data and with feature randomness (25). This inherent diversity and randomness help mitigate overfitting, a common challenge in machine learning, by reducing the model’s sensitivity to noise and outliers (26). Moreover, random forest’s ability to handle both classification and regression tasks, its capacity to capture complex non-linear relationships in the data, and its robustness to multicollinearity make it particularly well-suited for a wide range of datasets (27). Additionally, the ensemble nature of random forest allows it to aggregate the predictions from multiple trees, reducing the risk of bias that can be associated with individual models. Consequently, the comprehensive nature of random forest, combining predictive power and robustness, positions it as an attractive choice for achieving high accuracy in diverse machine learning tasks.

Prior research has suggested that male participants typically display a greater retinal thickness compared to female participants (7–12). The impact of sex on retinal layers is still a topic of ongoing debate. Some studies (13–17) have reported thicker retinal layers in male subjects (GCL, IPL, INL, OPL, and ONL), while others have observed minimal or no sex-related differences (18, 19). Some studies have shown that female subjects had a thicker peripapillary RNFL than male subjects (28, 29). The present study uncovers a greater thickness in the inner retinal layers of male subjects compared to female subjects. Sexual hormones interacting with receptors such as estrogen and androgen receptors can affect ocular tissue. However, despite their influence on various ocular structures, the effect of these hormones on retinal layer thickness remains largely uninvestigated (30–35).

Neglecting to account for sex differences in comparisons of retinal thickness between healthy individuals and patients could result in erroneous diagnoses, particularly for inner retinal



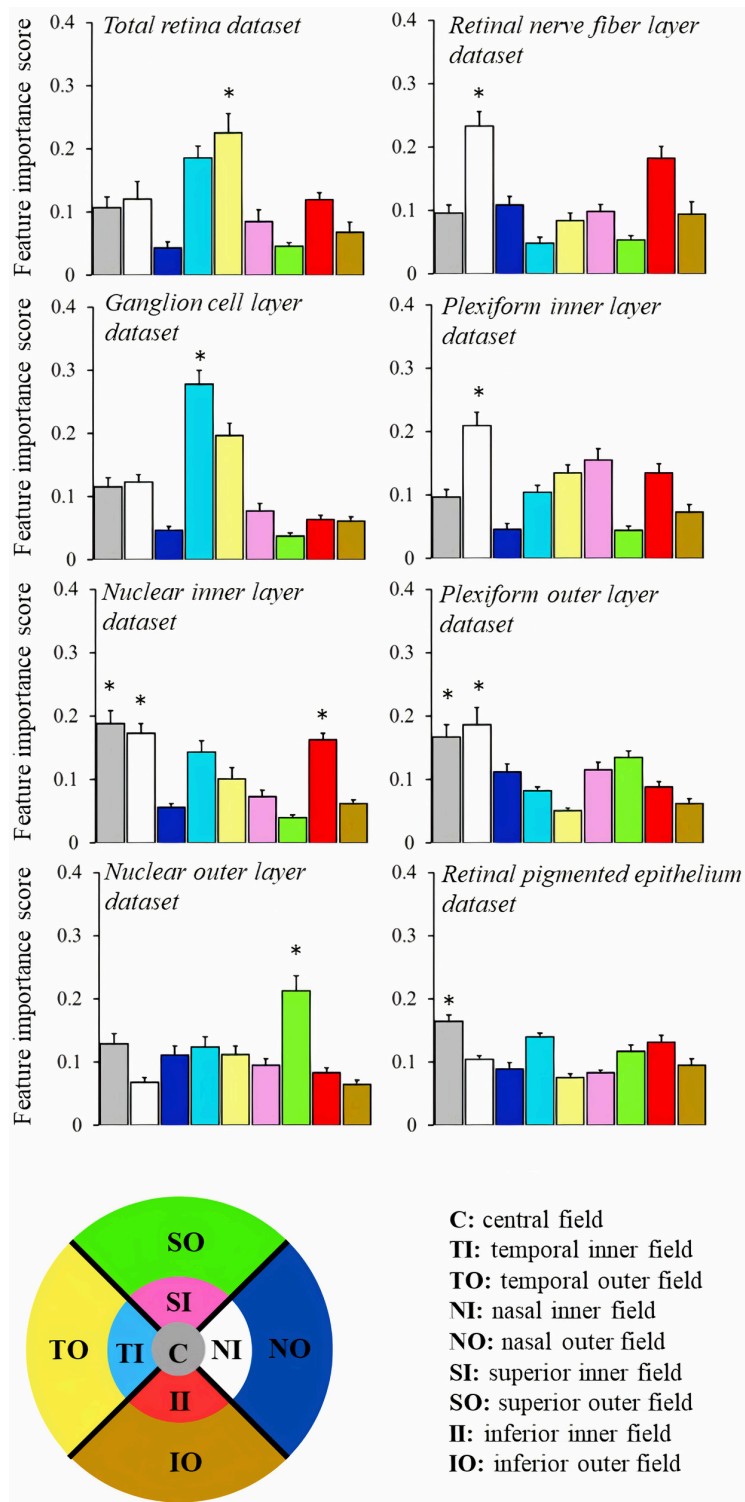


FIGURE 2 Comparison of the feature importance score obtained from random forest algorithm to classify the sex of the participant based on retinal thickness from different datasets. The color code is indicated at the bottom of the figure. * $p < 0.05$.

diseases that display substantial sex-related disparities. Conditions like glaucoma, macular holes, diabetic retinopathy, and age-related macular degeneration demonstrate varying prevalence rates between male and female subjects. This is likely attributable to changes in sex hormone concentrations after the age of 50 (36, 37).

The current investigation focuses on recruiting predominantly young adult participants, and as a result, the applicability of our findings may be limited to this specific age group. This demographic constraint represents a notable limitation of our study. To enhance the generalizability and robustness of our

conclusions, it is imperative for future research endeavors to encompass a broader spectrum of cases, incorporating individuals from various age ranges. In the present study, our primary aim was to demonstrate that various models can learn pertinent sex-related patterns within diverse retinal datasets. While the current sample size has proven adequate for this initial validation, it remains a limitation of the study and should be expanded in future research endeavors.

In conclusion, this research highlights the discriminative capacity of different retinal layers in sex classification, achieving varying levels of accuracy across distinct layers and ML algorithms. The consistently superior performance of the RF algorithm indicates its effectiveness in identifying sex-related characteristics in various retinal layers. Furthermore, the identified patterns of accuracy fluctuations across retinal layers offer invaluable insights for subsequent research and algorithmic advancement in the field of retinal data analysis.

Data availability statement

The original contributions presented in this study are available after inquiries directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethical Committee for Research in Humans of the Universidade Federal do Pará (report number 3.285.557). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

FF: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing.

RS: Conceptualization, Supervision, Writing – review & editing. ER: Formal analysis, Software, Writing – review & editing. AS: Investigation, Writing – review & editing. GSAS: Investigation, Writing – review & editing. AR: Investigation, Writing – review & editing. MC: Formal analysis, Funding acquisition, Writing – review & editing. GSS: Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by research grants from the Brazilian funding agencies: CNPq Edital Universal (#431748/2016-0). FF was a CAPES fellow for graduate students. MC and GSS are CNPq Fellows, Productivity Grants 302552/2017-0, and 309936/2022-5, respectively. The funders had no role in the study design.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Huang D, Swanson E, Lin C, Schuman J, Stinson W, Chang W, et al. Optical coherence tomography. *Science*. (1991) 254:1178–81. doi: 10.1126/science.1957169
- Swanson E, Izatt J, Hee M, Huang D, Lin C, Schuman J, et al. In vivo retinal imaging by optical coherence tomography. *Opt Lett*. (1993) 18:1864–6. doi: 10.1364/ol.18.001864
- Izatt J, Hee M, Swanson E, Lin C, Huang D, Schuman J, et al. Micrometer-scale resolution imaging of the anterior eye in vivo with optical coherence tomography. *Arch Ophthalmol*. (1994) 112:1584–9. doi: 10.1001/archoph.1994.01090240090031
- Pagon R. Retinitis pigmentosa. *Surv Ophthalmol*. (1988) 33:137–77. doi: 10.1016/0039-6257(88)90085-9
- Gardner T, Antonetti D, Barber A, LaNoue K, Levison S. Diabetic retinopathy: more than meets the eye. *Surv Ophthalmol*. (2002) 47 Suppl 2:S253–62. doi: 10.1016/s0039-6257(02)00387-9
- Lim L, Mitchell P, Seddon J, Holz F, Wong T. Age-related macular degeneration. *Lancet*. (2012) 379:1728–38. doi: 10.1016/S0140-6736(12)60282-7
- Wong A, Chan C, Hui S. Relationship of gender, body mass index, and axial length with central retinal thickness using optical coherence tomography. *Eye*. (2005) 19:292–7. doi: 10.1038/sj.eye.6701466
- Ooto S, Hangai M, Sakamoto A, Tomidokoro A, Araie M, Otani T, et al. Three-dimensional profile of macular retinal thickness in normal Japanese eyes. *Invest Ophthalmol Vis Sci*. (2010) 51:465–73. doi: 10.1167/iovs.09-4047
- Adhi M, Aziz S, Muhammad K, Adhi M. Macular thickness by age and gender in healthy eyes using spectral domain optical coherence tomography. *PLoS One*. (2012) 7:e37638. doi: 10.1371/journal.pone.0037638
- Çubuk M, Kasım B, Koçluk Y, Sukgen E. Effects of age and gender on macular thickness in healthy subjects using spectral optical coherence tomography/scanning laser ophthalmoscopy. *Int Ophthalmol*. (2018) 38:127–31. doi: 10.1007/s10792-016-0432-z
- Kelty P, Payne J, Trivedi R, Kelty J, Bowie E, Burger B. Macular thickness assessment in healthy eyes based on ethnicity using Stratus OCT optical coherence tomography. *Invest Ophthalmol Vis Sci*. (2008) 49:2668–72. doi: 10.1167/iovs.07-1000

12. Song W, Lee S, Lee E, Kim C, Kim S. Macular thickness variations with sex, age, and axial length in healthy subjects: a spectral domain-optical coherence tomography study. *Invest Ophthalmol Vis Sci.* (2010) 51:3913–8. doi: 10.1167/iovs.09-4189
13. Ooto S, Hangai M, Tomidokoro A, Saito H, Araie M, Otani T, et al. Effects of age, sex, and axial length on the three-dimensional profile of normal macular layer structures. *Invest Ophthalmol Vis Sci.* (2011) 52:8769–79. doi: 10.1167/iovs.11-8388
14. Won J, Kim S, Park Y. Effect of age and sex on retinal layer thickness and volume in normal eyes. *Medicine.* (2016) 95:e5441. doi: 10.1097/MD.0000000000005441
15. Nieves-Moreno M, Martínez-de-la-Casa J, Morales-Fernández L, Sánchez-Jean R, Sáenz-Francés F, García-Feijó J. Impacts of age and sex on retinal layer thicknesses measured by spectral domain optical coherence tomography with Spectralis. *PLoS One.* (2018) 13:e0194169. doi: 10.1371/journal.pone.0194169
16. Invernizzi A, Pellegrini M, Acquistapace A, Benatti E, Erba S, Cozzi M, et al. Normative data for retinal-layer thickness maps generated by spectral-domain OCT in a white population. *Ophthalmol Retina.* (2018) 2:808.e–15.e. doi: 10.1016/j.oret.2017.12.012
17. Palazon-Cabanes A, Palazon-Cabanes B, Rubio-Velazquez E, Lopez-Bernal M, Garcia-Medina J, Villegas-Perez M. Normative Database for All Retinal Layer Thicknesses Using SD-OCT Posterior Pole Algorithm and the Effects of Age, Gender and Axial Length. *J Clin Med.* (2020) 9:3317. doi: 10.3390/jcm9103317
18. Grover S, Murthy R, Brar V, Chalam K. Normative data for macular thickness by high-definition spectral-domain optical coherence tomography (spectralis). *Am J Ophthalmol.* (2009) 148:266–71. doi: 10.1016/j.ajo.2009.03.006
19. Appukuttan B, Giridhar A, Gopalakrishnan M, Sivaprasad S. Normative spectral domain optical coherence tomography data on macular and retinal nerve fiber layer thickness in Indians. *Indian J Ophthalmol.* (2014) 62:316–21. doi: 10.4103/0301-4738.116466
20. Nuzzi R, Scalabrin S, Becco A, Panzica G. Gonadal Hormones and Retinal Disorders: A Review. *Front Endocrinol.* (2018) 9:66. doi: 10.3389/fendo.2018.00066
21. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, et al. Artificial intelligence: A powerful paradigm for scientific research. *Innovation.* (2021) 2:100179. doi: 10.1016/j.xinn.2021.100179
22. Ali O, Abdelbaki W, Shrestha A, Elbasi E, Alryalat M, Dwivedi YK. A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *J Innov Knowledge.* (2023) 8:100333.
23. Qin F, Lv Z, Wang D, Hu B, Wu C. Health status prediction for the elderly based on machine learning. *Arch Gerontol Geriatr.* (2020) 90:104121. doi: 10.1016/j.archger.2020.104121
24. Chueh K, Hsieh Y, Chen H, Ma I, Huang S. Identification of sex and age from macular optical coherence tomography and feature analysis using deep learning. *Am J Ophthalmol.* (2022) 235:221–8. doi: 10.1016/j.ajo.2021.09.015
25. Steffens M, Lamina C, Illig T, Bettecken T, Vogler R, Entz P, et al. SNP-based analysis of genetic substructure in the German population. *Hum Hered.* (2006) 62:20–9. doi: 10.1159/000095850
26. Abellán J, Mantas C, Castellano J, Moral-García S. Increasing diversity in random forest algorithm via imprecise probabilities. *Expert Syst Applic.* (2018) 97:228–43.
27. Svetnik V, Liaw A, Tong C, Culberson J, Sheridan R, Feuston B. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci.* (2003) 43:1947–58. doi: 10.1021/ci034160g
28. Li D, Rauscher F, Choi E, Wang M, Baniyasi N, Wirkner K, et al. Sex-specific differences in circumpapillary retinal nerve fiber layer thickness. *Ophthalmology.* (2020) 127:357–68. doi: 10.1016/j.ophtha.2019.09.019
29. Rougier M, Korobelnik J, Malet F, Schweitzer C, Delyfer M, Dartigues J, et al. Retinal nerve fibre layer thickness measured with SD-OCT in a population-based study of French elderly subjects: the Alienor study. *Acta Ophthalmol.* (2015) 93:539–45. doi: 10.1111/aos.12658
30. Ogueta S, Schwartz S, Yamashita C, Farber D. Estrogen receptor in the human eye: influence of gender and age on gene expression. *Invest Ophthalmol Vis Sci.* (1999) 40:1906–11.
31. Rocha E, Wickham L, da Silveira L, Krenzer K, Yu F, Toda I, et al. Identification of androgen receptor protein and 5alpha-reductase mRNA in human ocular tissues. *Br J Ophthalmol.* (2000) 84:76–84. doi: 10.1136/bjo.84.1.76
32. Wickham L, Gao J, Toda I, Rocha E, Ono M, Sullivan D. Identification of androgen, estrogen and progesterone receptor mRNAs in the eye. *Acta Ophthalmol Scand.* (2000) 78:146–53. doi: 10.1034/j.1600-0420.2000.078002146.x
33. Munaut C, Lambert V, Noël A, Frankenne F, Deprez M, Foidart J, et al. Presence of oestrogen receptor type beta in human retina. *Br J Ophthalmol.* (2001) 85:877–82. doi: 10.1136/bjo.85.7.877
34. Gupta P, Johar K, Nagpal K, Vasavada A. Sex hormone receptors in the human eye. *Surv Ophthalmol.* (2005) 50:274–84. doi: 10.1016/j.survophthal.2005.02.005
35. Cascio C, Deidda I, Russo D, Guarneri P. The estrogenic retina: The potential contribution to healthy aging and age-related neurodegenerative diseases of the retina. *Steroids.* (2015) 103:31–41. doi: 10.1016/j.steroids.2015.08.002
36. Evans J, Schwartz S, McHugh J, Thamby-Rajah Y, Hodgson S, Wormald R, et al. Systemic risk factors for idiopathic macular holes: a case-control study. *Eye.* (1998) 12:256–9. doi: 10.1038/eye.1998.60
37. Rudnicka A, Jarrar Z, Wormald R, Cook D, Fletcher A, Owen C. Age and gender variations in age-related macular degeneration prevalence in populations of European ancestry: a meta-analysis. *Ophthalmology.* (2012) 119:571–80. doi: 10.1016/j.ophtha.2011.09.027