Original Article

Clinical Relevance of Choroidal Thickness in Obese and Healthy Children: A Machine Learning Study

Erkan Bulut*, Sümeyra Köprübaşı**, Özlem Dayi***, Hatice Bulut****

*Gelişim University, Vocational School of Health Services, Department of Opticianry, İstanbul, Türkiye
** University of Health Sciences Türkiye, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Clinic of Ophthalmology, İstanbul, Türkiye
***Beylikdüzü State Hospital, Clinic of Ophthalmology, İstanbul, Türkiye
****İstanbul Gelişim University, Vocational School of Health Services, Department of Child Development, İstanbul, Türkiye

Abstract

Objectives: To analyze the effect of macular choroidal thickness (MCT) and peripapillary choroidal thickness (PPCT) on the classification of obese and healthy children by comparing the performance of the random forest (RF), support vector machine (SVM), and multilayer perceptrons (MLP) algorithms.

Materials and Methods: Fifty-nine obese children and 35 healthy children aged 6 to 15 years were studied in this prospective comparative study using optical coherence tomography. MCT and PPCT were measured at distances of 500 µm, 1,000 µm, and 1,500 µm from the fovea and optic disc. Three different feature selection algorithms were used to determine the most prominent features of all extracted features. The classification efficiency of the extracted features was analyzed using the RF, SVM, and MLP algorithms, demonstrating their efficacy for distinguishing obese from healthy children. The precision and reliability of measurements were assessed using kappa analysis.

Results: The correlation feature selection algorithm produced the most successful classification results among the different feature selection methods. The most prominent features for distinguishing the obese and healthy groups from each other were PPCT temporal 500 µm, PPCT temporal 1,500 µm, PPCT nasal 1,500 µm, PPCT inferior 1,500 µm, and subfoveal MCT. The classification rates for the RF, SVM, and MLP algorithms were 98.6%, 96.8%, and 89%, respectively.

Conclusion: Obesity has an effect on the choroidal thicknesses of children, particularly in the subfoveal region and the outer semi-circle at 1,500 µm from the optic disc head. Both the RF and SVM algorithms are effective and accurate at classifying obese and healthy children.

Keywords: Choroidal thickness, feature selection, machine learning, obese children, optical coherence tomography

Cite this article as: Bulut E, Köprübaşı S, Dayi Ö, Bulut H. Clinical Relevance of Choroidal Thickness in Obese and Healthy Children: A Machine Learning Study. Turk J Ophthalmol 2023;53:161-168

Address for Correspondence: Özlem Dayi, Beylikdüzü State Hospital, Clinic of Ophthalmology, İstanbul, Türkiye
E-mail: ozlemkuru_uutf@hotmail.com **ORCID-ID:** orcid.org/0000-0001-7008-077X
Received: 26.03.2022 **Accepted:** 21.09.2022

DOI: 10.4274/tjo.galenos.2022.36724

Introduction

Childhood obesity is an exceedingly prevalent health issue in the world. The World Health Organization (WHO) has declared obesity as an "escalating global epidemic."[1] Worldwide, 22 million children under the age of 5 years and 150 million school-age children have been reported to be severely overweight, with the prevalence of childhood obesity estimated to be 10%.[2] While there are several parameters to indicate a child's nutrition and growth status, the parameter recommended by WHO is the Z-score. The Z-score system displays a set of standard deviations (SD) from the reference median or mean. It allows more accurate assessments by standardizing measurements based on age and gender.[3] The Z-score system can be used to calculate a number of anthropometric values such as weight-for-age Z-scores, height-for-age Z-scores (HAZ), weight-for-height Z-scores, and body mass index-for-age Z-scores (BMIZ). Body mass index (BMI) is the most frequently used metric in ophthalmological research to define children's nutrition and development. However, BMIZ has been reported to be the most helpful technique for assessing obesity.[4]

Obesity has been associated with multiple ocular diseases, including cataract, glaucoma, dry eye, diabetic retinopathy, and age-related macular degeneration.[5,6,7] Although the reason for the relationship between obesity and eye diseases is unclear, it is thought to be related to obesity-related chronic oxidative stress, endothelial dysfunction, and vascular damage.[6] Changes in choroidal thickness are also observed in various systemic diseases, including diabetes, hypertension, and endocrine diseases.[8,9] There are a few studies on the effects of obesity on the eyes, but no detailed assessment of macular choroidal thickness (MCT) and peripapillary choroidal thickness (PPCT) has been conducted.[10,11,12]

Due to advances in computing technology, artificial intelligence has begun to replace conventional parametric tests in data analysis. Machine learning, the most important subset of artificial intelligence, makes it possible to interpret information, classify data, and make predictions for the future by analyzing the structures and texture patterns of a large number of computer data.[13,14] Machine learning algorithms have been found to be more efficient, effective, and accurate than conventional statistical methods in the analysis of a large number of complex data.[13,15,16]

The random forest (RF) algorithm is a grouping, correlation, and other task-specific ensemble learning process.[17] Support vector machine (SVM) is a regulated classification algorithm with learning techniques for classification and correlation analysis. The SVM algorithm successfully allows multidimensional and nonlinear classifications.[18] Multilayer perceptrons (MLP) is a well-known correlation algorithm for determining the relationship between a continuous dependent variable and two or more independent variables.[19]

Several image classification studies have been conducted in the field of ophthalmology to classify different eye conditions. Dong et al.[20] conducted a study on eye state estimation with various feature sets using RF, random ferns, and SVM and reported high success with random forest/ferns. In another study, Agarwal et al.[21] demonstrated the feasibility of a multilayer-based methodology in detecting cataracts with a success rate of 94% and 75% with SVM and MLP, respectively. Improta et al.[22] studied the eye-tracking patterns of newborns acquired by electrooculography and infrared oculography to detect congenital nystagmus. They demonstrated the feasibility of a regression analysis performed through machine learning algorithms like RF, logistic regression tree, gradient boosted tree, K-nearest neighbor, MLP, and SVM to detect variables related to congenital nystagmus. Avilés-Rodríguez et al.[23] performed a quality assessment of eye fundus images acquired by digital fundoscopy with topological data analysis and machine learning methods like SVM, decision tree, k-NN, random forest, logistic regression (LoGit), and MLP. da Cruz et al.[24] studied dry eye syndrome classification using machine learning algorithms like SVM, RF, naive Bayes, MLP, random tree, and RBF Network and reported the highest performance using the RF classifier (97% accuracy).

In this study, we examined and compared the performance of RF, SVM, and MLP algorithms in the classification of obese and healthy children based on differences in MCT and PPCT. We aimed to examine the impact of childhood obesity on choroidal thickness and to recognize early clinical changes that could pose a risk for multiple ocular diseases by using machine learning algorithms, a modern method of analysis.

Materials and Methods

This research was reviewed by an independent ethical review board and conformed to the principles and applicable guidelines for the protection of human subjects in biomedical research.

In this prospective comparative study, healthy and obese children between 6 and 15 years of age who presented to the departments of pediatrics and ophthalmology for routine follow-up were recruited from 1 June 2020 to 1 December 2020. The exclusion criteria were as follows: presence of chronic diseases such as diabetes, hypertension, heart disease, and obstructive sleep apnea syndrome; history of any medication use; ocular diseases such as strabismus, cataracts, glaucoma, amblyopia, uveitis, optic disc anomaly, and retinal disease; history of prior eye surgery; more than 2 diopters of spherical or cylindrical refractive error; corneal, lens, or vitreous opacity which does not allow quality optical coherence tomography (OCT) imaging; and insufficient cooperation for OCT imaging.

Physical Examination

Height and weight measurements were taken using a digital scale and a wall-mounted Harpender stadiometer. Z-scores were determined using the WHO AnthroPlus software (www.who.int/tools/growth-reference-data-for-5to19-years/application-tools). Obesity was defined as greater than +2 SD, while normal weight was defined as between 1 and +1 SD for both BMIZ and HAZ.[3] After a resting period, blood pressure was measured using an automatic sphygmomanometer (Omron M2 HEM7121E, Omron Healthcare Co, Japan) at least three times within a 10-minute period. Blood pressure was measured as the average of a total of three consecutive measurements taken after the required resting time. Children with systolic and/or diastolic blood pressure levels greater than the 95th percentile were defined as hypertensive.[25]

Ophthalmological Examination

A detailed ophthalmological examination, including measures of best-corrected visual acuity, spherical equivalent, slit-lamp biomicroscopy, intraocular pressure (IOP), central corneal thickness (CCT), axial length (AXL), and anterior chamber depth (ACD), and OCT imaging were performed for each participant by an experienced ophthalmologist. Only the participants' right eyes were included in the study. Autokeratorefractometry (Topcon KR-800, Topcon Medical Systems, Inc., Fukuoka, Japan) was used for refractive measurements. IOP was measured using Goldmann applanation tonometry and CCT was measured using a non-contact tonopachymeter (NT-530P, Nidek Co., Gamagori, Japan). AXL and ACD were measured using optic biometry (Nidek Axial Length-Scan, Nidek Co., Gamagori, Japan). Retinal and choroidal thicknesses were assessed using Spectralis OCT (Cirrus HD OCT, Carl Zeiss Meditec, Dublin, CA, USA).

All OCT imaging and evaluations were performed by the same experienced ophthalmologist without pupil dilatation. All examinations were performed between 9:00 and 11:00 a.m. to reduce diurnal variances. Retinal thickness and mean ganglion cell layer and inner plexiform layer (GCL + IPL) thickness were measured using automated segmentation values of the Spectralis OCT system with a macular cube position of 512x128. The OCT HD 1-line-EDI protocol's high-resolution scan through the fovea was used for MCT measurements. Choroidal thickness

was assessed manually from the outer edge of the hyperreflective line corresponding to the retinal pigment epithelium to the inner layer of the sclera. MCT measurements were performed at the foveal center and at distances of 500 µm, 1,000 µm, and 1,500 µm nasally and temporally from the foveal center. For PPCT assessment, scans were carried out in vertical and horizontal planes through the middle of the optic disc using the OCT HD 5-Line Raster-EDI protocol.[26] In this scan, the optic disc is divided into two equal sections in both the horizontal and vertical planes. Then, in each of the nasal, temporal, superior, and inferior regions, PPCT measurements were taken at distances of 500 µm, 1,000 µm, and 1,500 µm from the optic disc margin (Figure 1). Both MCT and PPCT measurements were performed at 100% magnification by two masked ophthalmologists (E.B., O.D.) during different sessions for inter-observer reproducibility. The OCT Disc Cube 200x200 protocol was used for retinal nerve fiber layer thickness (RNFLT) and cup-to-disc ratio analysis. Superior, inferior, nasal, temporal, and average RNFLT values were calculated automatically.

Data Analysis

Feature Extraction and Selection

We manually measured all the features considered significant and tested whether these parameters validated our hypothesis or not. All of the manually extracted features are given in Table 1.

Feature selection techniques are based on the procedure of selecting the most important parameters. Feature selection primarily focuses on removing non-informative or irrelevant predictors from the model to minimize the number of parameters. The classification efficiency of different systems is influenced by their capabilities in data classification. In order to produce an easier, faster, and efficient classification system, we used three feature selection algorithms: variable ranking (VR), correlation feature selection (CFS), and principal component analysis (PCA). All the extracted features were entered into the VR, CFS, and PCA algorithms, and the most prominent features were selected to form the feature vector. This feature vector is used as an input for the classification algorithms (Figure 2).

Classifiers for Machine Learning

After the feature selection process, we looked at how well RF, SVM, and MLP performed with the selected prominent features and compared them to see if they could differentiate between obese and healthy children. We analyzed and compared the efficiency of RF, SVM, and MLP based on selected features. The efficiency of the different algorithms can vary, since they are structured differently. RF works through building a large number of decision trees during training and then extracting the test.[17,18] The SVM algorithm uses a training dataset to assign characteristics to just one or another subclass, making it a binary and linear classifier that cannot be predicted.[18] MLP is often used to determine which variable has the largest influence on the expected output and which variables relate to each other.[19]

Artificial intelligence-based categorization systems may be measured using precision (positive predictive), recall (sensitivity), and F-measure. Unlike precision, which only looks at correct positive predictions, recall also looks at positive predictions that did not come true. The F-measure gives us the harmonic mean of the values of precision and recall. The primary purpose of utilizing the F-measure value is to avoid selecting an inappropriate model of non-uniformly distributed datasets. The F-measure is a method for combining precision and recall into a single measure that includes all qualities. We conducted kappa analysis to assess the reliability and accuracy of our measurements. The kappa value ranges from 0 to +1. System reliability improves as the kappa value approaches 1.[27]

Results

This study included 59 obese children (35 girls, 24 boys) as the study group and 35 healthy children (21 girls, 14 boys) as the control group.

The CFS algorithm produced the most successful classification results among the three different feature selection methods. The CFS algorithm determined that subfoveal choroidal thickness is the most distinguishing feature, along with PPCT measurement locations including temporal 500 µm, temporal 1,500 µm, nasal 1,500 µm, and inferior 1,500 µm. In addition to these

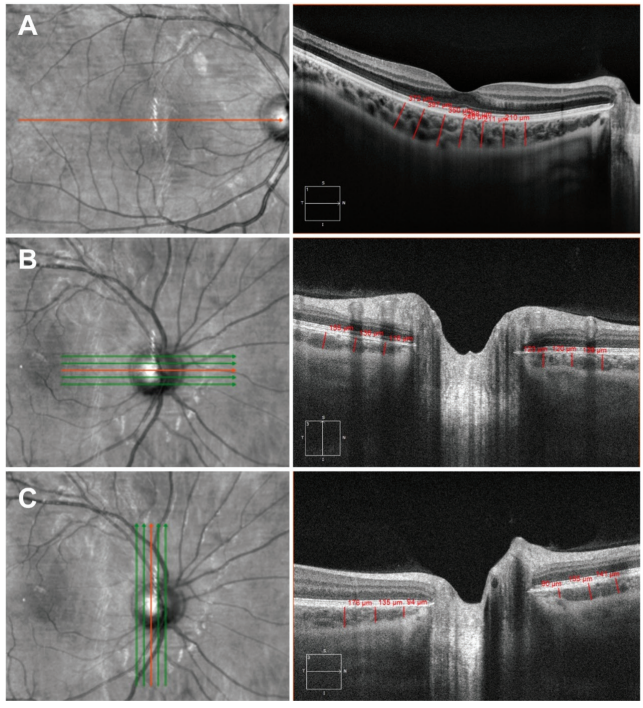

Figure 1. Example of macular and peripapillary choroidal thickness measurements (right eye). A) Macular choroidal thickness was measured at the central fovea (left panel: line denotes where the scan was taken relative to the fundus; right panel: lines show the measurement sites in the nasal (left) and temporal (right) quadrants. B) Peripapillary choroidal thickness measurements on the horizontal plane through the center of the optic disc (left panel: lines denote where the scan was taken relative to the fundus; right panel: lines show the measurement sites in the nasal (left) and temporal (right) quadrants. C) Peripapillary choroidal thickness measurements in the vertical plane through the center of the optic disc (left panel: lines denote where scan was taken relative to the fundus; right panel: lines show the measurement sites in the superior (right) and inferior (left) quadrants)

Table 1. All extracted features

Physical examination-based features	Ocular examination-based features	OCT imaging-based PPCT features	OCT imaging-based MCT features	OCT imaging-based other features
Age	Spherical equivalent	PPCT temporal 500	MCT fovea	GCL + IPL complex thickness
Sex	AXL	PPCT temporal 1000	MCT temporal 500	MT
Height	ACD	PPCT temporal 1500	MCT temporal 1000	Average c/d ratio
Weight	IOP	PPCT nasal 500	MCT temporal 1500	Vertical c/d ratio
BMI	Pachymetry	PPCT nasal 1000	MCT nasal 500	RNFLT temporal
BMIZ		PPCT nasal 1500	MCT nasal 1000	RNFLT nasal
HAZ		PPCT superior 500	MCT nasal 1500	RNFLT superior
Systolic BP		PPCT superior 1000		RNFLT inferior
Diastolic BP		PPCT superior 1500		RNFLT average
		PPCT inferior 500		
		PPCT inferior 1000		
		PPCT inferior 1500		

ACD: Anterior chamber depth, AXL: Axial length, BP: Blood pressure, BMI: Body mass index, BMIZ: BMI-for-age Z-score, c/d: Cup-to-disc, GCL + IPL: Ganglion cell layer + inner plexiform layer, HAZ: Height-for-age Z-score, IOP: Intraocular pressure, MCT: Macular choroidal thickness, MT: Macular thickness, OCT: Optical coherence tomography, PPCT: Peripapillary choroidal thickness, RNFLT: Retinal nerve fiber layer thickness

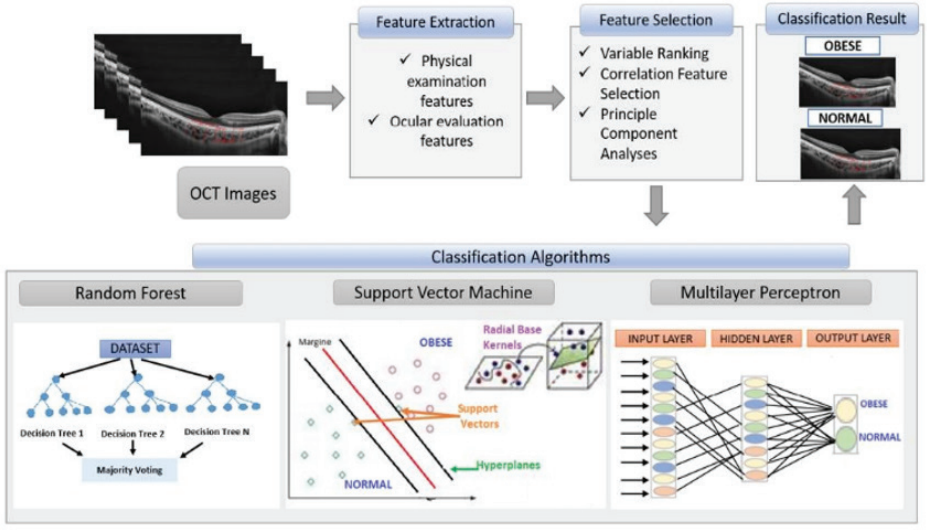

Figure 2. Flow chart of the proposed recognition system

features, the PCA algorithm selected the spherical equivalent value feature. However, when the spherical equivalent feature was absent, the classification results showed a higher success rate.

A 10-fold cross-validation process was used to test the stability and reliability of the RF, SVM, and MLP algorithms. The dataset was divided into two sections, with 70% of the data used for training and 30% for testing. To reduce selection bias, random sampling was conducted ten times to generate separate training and testing sets from the dataset.

The confusion matrix and classification rates of the RF, SVM, and MLP algorithms to classify children as normal or obese according to choroidal thickness are shown in Table 2. The overall accuracy rate of our system was 98.9% based on RF, 96.8% based on SVM, and 89.4% based on MLP.

Although the RF and SVM algorithms were equally successful at classifying the healthy group, RF was more successful in tagging the obese group. While the RF algorithm identified all obese data sets correctly, the SVM algorithm incorrectly classified two obese datasets as healthy. The BMIZ values of misclassified children were respectively 2.01 and 2.02. The thickness of the choroidal layer differed between obese and healthy children, and this difference was crucial in classifying groups using both the RF and SVM algorithms.

Despite using different learning rates and architecture, success with the MLP algorithm only increased from 85.83% to 89.36%. The reason for this small change is most probably because the dataset is limited, falls into the local extremum, and lacks spatial information.

The overall precision rate was high for RF (98.9%) and SVM (96.8%) but was relatively unsatisfactory for the MLP system (89.4%). Similarly, the overall F-measurement results of RF and SVM were both high (98.9% and 96.8%, respectively), whereas the result of MLP was low (89%). The overall recall rates for the RF and SVM systems were also 98.9% and 96.8%, respectively.

However, recall values for the obese group for the RF and SVM systems were 100% and 96.6%, respectively, which confirms the power of the proposed system's capability to recognize choroidal thickness measurements (Table 2). The average recall rate for the MLP system was 89.4%. However, the recall values of the obese and healthy groups were 98.3% and 74.3%, respectively (Table 2).

Reliability analysis yielded kappa coefficients of 0.9771, 0.9305, and 0.7600 for RF, SVM, and MLP, respectively.

Discussion

According to the findings of the current study, obesity had an effect on choroidal thickness at specific measurement regions but not at all measurement sites. The results suggest that obesity-related metabolic alterations affect choroidal thickness, particularly in the subfoveal region and the outer semi-circle at 1,500 µm from the optic disc head. This study is noteworthy because it not only comprehensively assessed choroidal thickness in obese children, but also utilized machine learning techniques in its analysis.

There are a few studies in the literature that assess the impact of childhood obesity on ocular structures. Baran et al.[10] found that obese children had higher IOP and lower RNFLT than healthy children and reported that childhood obesity may contribute to the development of glaucoma. They assessed choroidal thickness in the central subfoveal region alone and discovered no statistically significant differences. However, they did not conduct a comprehensive evaluation of MCT and PPCT. Bulus et al.[11] determined that obese children had thicker MCT than healthy children, but they did not evaluate PPCT. Additionally, they also used the BMI SD score, which is equal to the BMIZ for childhood nutrition and growth classification reported by the WHO in 2006. Bulus et al.[11] reported a strong positive correlation between BMI SD score and subfoveal MCT. Consistent with this study, we found that subfoveal MCT is affected by obesity and is a distinguishing feature between the obese and control groups.

While there are several literature studies assessing MCT in various diseases, there are few studies evaluating PPCT. Read et al.[28] identified normal PPCT values and variations in healthy children and confirmed that myopic refractive errors cause a reduction in PPCT. Ozcimen et al.[29] documented thinning in both PPCT and MCT in chronic obstructive pulmonary diseases. They attributed the choroidal thinning to vascular resistance resulting from hypoxia. Komma et al.[30] evaluated PPCT and subfoveal choroidal thickness in healthy subjects and glaucoma patients using spectral domain OCT and swept-source OCT. They discovered that choroidal thickness was significantly thicker in glaucoma subjects than controls in the peripapillary region, but not in the macular region on swept-source OCT.

This is the first research that we are aware of that evaluates PPCT in childhood obesity. Furthermore, conventional statistical methods have been employed in previous studies, including choroidal evaluation in various disorders. There is no prior study in the current literature that evaluates both MCT and PPCT using machine learning algorithms.

In machine learning, feature selection helps boost classification efficiency by avoiding over-fitting, creating a time-saving model, and making the designed model more human-friendly. There are

Table 2. Classification results of obese and healthy children based on choroidal thickness by algorithm

	TP rate	FP rate	Precision	Recall	F-measure	Confusion matrix	
Random forest algorithm							
Obese	1	0.029	0.983	1	0.992	59	0
Normal	0.971	0.000	1	0.971	0.986	1	34
Weighted average	0.989	0.018	0.990	0.989	0.989		
Support vector machine algorithm							
Obese	0.966	0.029	0.983	0.966	0.974	57	2
Normal	0.971	0.034	0.944	0.971	0.958	1	34
Weighted average	0.968	0.031	0.968	0.968	0.968		
Multilayer perceptrons algorithm							
Obese	0.983	0.257	0.866	0.983	0.921	58	1
Healthy	0.743	0.017	0.963	0.743	0.839	9	26
Weighted average	0.894	0.168	0.902	0.894	0.890		

TP: True positive, FP: False positive

several feature selection approaches in the literature to minimize the number of features for classification purposes. Different subsets can be created with each feature selection method. We ran all of the data through a feature selection process using three different algorithms: VR, CFS, and PCA. None of the parameters associated with MCT and PPCT were excluded in any of the three analyses, and they were found to be distinctive in all of them. According to the results, obese and healthy children have significantly different choroidal thicknesses at specific measurement regions. These measurement regions were PPCT temporal 500 µm, PPCT temporal 1,500 µm, PPCT nasal 1,500 µm, PPCT inferior 1,500 µm, and the subfoveal region. In the PCA algorithm, spherical equivalent value was chosen in addition to the distinguishing features chosen in the CFS algorithm. There was no statistically significant difference between the two groups' spherical equivalent values. The CFS algorithm outperforms PCA in classification because the spherical equivalent value was not a distinguishing feature for these groups. While machine learning algorithms identify distinct features in classification for the two groups, they do not show the relative value of these features in each group. As machine learning algorithms reveal the importance of features, classification is performed on all of the selected features.

In this study, we compared the results of three different classification algorithms (RF, SVM, and MLP) because it is difficult to predict which machine learning algorithm will perform better in classification. We selected RF because it is a good comparison and classification technique and can detect outliers very well. SVM is a very robust technique for solving high-dimensional problems and creating accurate classifications. MLP is an accessible technique with the ability to create a simple architecture, easily build it, and quickly calculate the model. The risk of falling into the local extremum, weak overfitting skills, a lack of theoretically-based rigid design programs, and difficulty managing the training program are disadvantages of the MLP algorithm. SVM may be more determinant in some cases, even though the RF algorithm is generally more successful in classification. We had several difficulties applying the SVM and MLP algorithms because of the limited and unbalanced datasets used in this study. To overcome this challenge, we focused on kernel selection, which had an effect on the kernel's success in implementing the SVM algorithm. We used polynomial and radial base kernels to improve classification efficiency by reducing our margin of error. Additionally, the success of the MLP algorithm was influenced by the network structure. The more complicated the network's structure, the more successful it will be. However, we did not increase the number of layers in order to reduce the margin of error.

While RF produces better results against outliers and noise than SVM, it is not as successful in handling the dataset imbalance problem. Although our dataset was slightly unbalanced, the results with RF were quite successful. MLP was found to be less successful than SVM and RF in the classification according to choroidal thickness.

The MLP algorithm had the highest rate of misclassification of all of the classification techniques. The MLP algorithm misclassified ten children, three of whom were also misclassified by the SVM algorithm. We found no similarities in terms of features such as height or weight in cases misclassified by the MLP algorithm. In terms of group classification, we discovered that the SVM algorithm outperformed the MLP algorithm. The main reason for misclassification based on the SVM algorithm may be that the children were at the threshold of obesity according to their BMIZ values. As a result, the classification success of the SVM algorithm is higher in obese cases with high BMIZ values.

The performance of machine learning algorithms, as well as the complexity of the models used, are influenced by the quality and quantity of data. To the best of our knowledge, there is no open dataset in the literature that is comparable to our dataset. The drawback of our analysis is the limited size of the dataset. However, the majority of medical research faces difficulty in achieving a sufficient number of cases. Obtaining large quantities of high-quality data for medical research is a time-consuming and difficult task. There is medical research in the literature that uses machine learning algorithms with small datasets. Ruiz Hidalgo et al.[31] used machine learning algorithms to classify keratoconus using five Pentacam-derived parameters of 131 eyes. An et al.[32] developed classification criteria that could aid in the clinical management of glaucoma by using machine learning algorithms to classify 163 glaucomatous optic discs. Cartes et al.[33] evaluated the variability of tear osmolarity in 20 patients with dry eye using machine learning techniques. It has been demonstrated that machine learning algorithms can conduct self-diagnosis and classification analyses of OCT images with high accuracy, speed, and consistency.[34] However, in the classification tests, we measured kappa values to ensure that the small dataset did not affect the reliability of our results and to maximize success. The kappa value is a measure that contrasts the observed precision with the predicted precision (random chance). This is a far more reflective indicator of model efficiency. Kappa values were measured as 0.9771, 0.9305, and 0.7600 for the RF, SVM, and MLP analyses, respectively. According to the kappa statistics, RF is the most accurate test, but the reliability of SVM is also very similar to RF. Despite the limited number of datasets, kappa analyses showed that both RF and SVM were very successful and reliable in the classification of obese and healthy children.

Conclusion

The current study indicates that MCT and PPCT differ in obese and healthy children and are effective in the categorization of these two groups using machine learning algorithms, especially when the RF or SVM algorithms were used. Additionally, obesity was shown to impact choroidal thickness in certain regions when compared to healthy children. The current study emphasizes the importance of subfoveal MCT as well as PPCT measurements in some regions (including temporal 500 µm, temporal 1,500 µm,

nasal 1,500 µm, and inferior 1,500 µm) in classifying children as obese or healthy. To improve classification performance, further deep learning studies with larger datasets are needed.

Acknowledgements

Special thanks to Aysun Sezer, who is a research engineer in Université Paris-Saclay, CEA LIST, Laboratory of Research on Software-intensive Technologies (LIST), Atomic Energy and Alternative Energies Commission. She worked as a data scientist and provided support with all statistical tests and machine learning algorithms for our research.

Ethics

Ethics Committee Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Ethics Committee of Biruni University (document number: 2020/40-06).

Informed Consent: Informed consent and oral consent was obtained from all individual participants and/or their legal guardians.

Peer-review: Externally peer reviewed.

Authorship Contributions

Concept: E.B., Ö.D., Design: E.B., S.K., Data Collection or Processing: E.B., Ö.D., Analysis or Interpretation: E.B., S.K., H.B., Literature Search: E.B., S.K., H.B., Writing: S.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Kosti RI, Panagiotakos DB. The epidemic of obesity in children and adolescents in the world. Cent Eur J Public Health. 2006;14:151-159.
2. Lobstein T, Baur L, Uauy R; IASO International Obesity TaskForce. Obesity in children and young people: a crisis in public health. Obes Rev. 2004;5(Suppl 1):4-104.
3. Mei Z, Grummer-Strawn LM. Standard deviation of anthropometric Z-scores as a data quality assessment tool using the 2006 WHO growth standards: a cross country analysis. Bull World Health Organ. 2007;85:441-448.
4. Cole TJ, Faith MS, Pietrobelli A, Heo M. What is the best measure of adiposity change in growing children: BMI, BMI %, BMI z-score or BMI centile? Eur J of Clin Nutr. 2005;59:419-425.
5. Age-Related Eye Disease Study Research Group. Risk factors associated with age-related nuclear and cortical cataract : a case-control study in the Age-Related Eye Disease Study, AREDS Report No. 5. Ophthalmology. 2001;108:1400-1408.
6. Abramson N, Abramson S. Hypercoagulability: clinical assessment and treatment. Sout Med J. 2001;94:1013-1020.
7. Karti O, Nalbantoglu O, Abali S, Tunc S, Ozkan B. The assessment of peripapillary retinal nerve fiber layer and macular ganglion cell layer changes in obese children: a cross-sectional study using optical coherence tomography. Int Ophthalmol. 2017;37:1031-1038.
8. Aiello LP, Avery RL, Arrigg PG, Keyt BA, Jampel HD, Shah ST, Pasquale LR, Thieme H, Iwamoto MA, Park JE, et al. Vascular endothelial growth factor in ocular fluid of patients with diabetic retinopathy and other retinal disorders. N Engl J Med. 1994;331:1480-1487.
9. Alam AA, Mitwalli AH, Al-Wakeel JS, Chaudhary AR, Zebaid MA. Plasma fibrinogen and its correlates in adult Saudi population. Saudi Med J. 2004;25:1593-1602.
10. Baran RT, Baran SO, Toraman NF, Filiz S, Demirbilek H. Evaluation of intraocular pressure and retinal nerve fiber layer, retinal ganglion cell, central macular thickness, and choroidal thickness using optical coherence tomography in obese children and healthy controls. Niger J Clin Pract. 2019;22:539-545.
11. Bulus AD, Can ME, Baytaroglu A, Can GD, Cakmak HB, Andiran N. Choroidal Thickness in Childhood Obesity. Ophthalmic Surg Lasers Imaging Retina. 2017;48:10-17.
12. Topcu-Yilmaz P, Akyurek N, Erdogan E. The effect of obesity and insulin resistance on macular choroidal thickness in a pediatric population as assessed by enhanced depth imaging optical coherence tomography. J Pediatr Endocrinol Metab. 2018;31:855-860.
13. Hansen M, Dubayah R, Defries R. Classification trees: an alternative to traditional land cover classifiers. Int J Remote Sens. 1996;17:1075-1081.
14. Huang C, Davis LS, Townshend JRG. An assessment of support vector machines for land cover classification. Int J Remote Sens. 2002;23:725-749.
15. Foody GM. Sample size determination for image classification accuracy assessment and comparison. Int J Remote Sens. 2009;30:5273-5291.
16. Friedl MA, Brodley CE, Strahler AH. Maximizing land cover classification accuracies produced by decision trees at continental to global scale. IEEE Trans Geosci Remote Sens.1999;37:969-977.
17. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci. 2003;43:1947-1958.
18. Jayadeva, Khemchandani R, Chandra S. Twin Support Vector Machines for pattern classification. IEEE Trans Pattern Anal Mach Intell. 2007;29:905-910.
19. Liu M, Wang M, Wang J, Li D. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. Sens Actuators B Chem. 2013;177:970-980.
20. Dong Y, Zhang Y, Yue J, Hu Z. Comparison of random forest, random ferns and support vector machine for eye state classification. Multimed Tools Appl. 2016;75:11763-11783.
21. Agarwal S, Kumar M, Jangir SK, Sharma C. Computer-Aided Cataract Detection Using MLP and SVM. In Artificial Intelligence and Global Society. 2021:103-114.
22. Improta G, Ricciardi C, Cesarelli G, D'Addio G, Bifulco P, Cesarelli M. Machine learning models for the prediction of acuity and variability of eye-positioning using features extracted from oculography. Health and Technology. 2020;10:961-968.
23. Avilés-Rodríguez GJ, Nieto-Hipólito JI, Cosío-León MLÁ, Romo-Cárdenas GS, Sánchez-López JD, Radilla-Chávez P, Vázquez-Briseño M. Topological Data Analysis for Eye Fundus Image Quality Assessment. Diagnostics (Basel). 2021;11:1322.
24. da Cruz LB, Souza JC, de Sousa JA, Santos AM, de Paiva AC, de Almeida JDS, Silva AC, Junior GB, Gattass M. Interferometer eye image classification for dry eye categorization using phylogenetic diversity indexes for texture analysis. Comput Methods Programs Biomed. 2020;188:105269.
25. Lurbe E, Agabiti-Rosei E, Cruickshank JK, Dominiczak A, Erdine S, Hirth A, Invitti C, Litwin M, Mancia G, Pall D, Rascher W, Redon J, Schaefer F, Seeman T, Sinha M, Stabouli S, Webb NJ, Wühl E, Zanchetti A. 2016 European Society of Hypertension guidelines for the management of high blood pressure in children and adolescents. J Hypertens. 2016;34:1887-1920.
26. Ho J, Branchini L, Regatieri C, Krishnan C, Fujimoto JG, Duker JS. Analysis of normal peripapillary choroidal thickness via spectral domain optical coherence tomography. Ophthalmology. 2011;118: 2001-2007.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.
28. Read SA, Alonso-Caneiro D, Vincent SJ, Collins MJ. Peripapillary choroidal thickness in childhood. Exp Eye Res. 2015;135:164-173.

29. Ozcimen M, Sakarya Y, Kurtipek E, Bekci TT, Goktas S, Sakarya R, Yener HI, Demir LS, Erdogan E, Ivacik IS, Alpfidan I, Bukus A. Peripapillary choroidal thickness in patients with chronic obstructive pulmonary disease. Cutan Ocul Toxicol. 2016;35:26-30.
30. Komma S, Chhablani J, Ali MH, Garudadri CS, Senthil S. Comparison of peripapillary and subfoveal choroidal thickness in normal versus primary open-angle glaucoma (POAG) subjects using spectral domain optical coherence tomography (SD-OCT) and swept source optical coherence tomography (SS-OCT). BMJ Open Ophthalmol. 2019;4:e000258.
31. Ruiz Hidalgo I, Rozema JJ, Saad A, Gatinel D, Rodriguez P, Zakaria N, Koppen C. Validation of an Objective Keratoconus Detection System Implemented in a Scheimpflug Tomographer and Comparison With Other Methods. Cornea. 2017;36:689-695.
32. An G, Omodaka K, Tsuda S, Shiga Y, Takada N, Kikawa T, Nakazawa T, Yokota H, Akiba M. Comparison of Machine-Learning Classification Models for Glaucoma Management. J Healthc Eng. 2018;2018:6874765.
33. Cartes C, López D, Salinas D, Segovia C, Ahumada C, Pérez N, Valenzuela F, Lanza N, López Solís RO, Perez VL, Zegers P, Fuentes A, Alarcón C, Traipe L. Dry eye is matched by increased intrasubject variability in tear osmolarity as confirmed by machine learning approach. Arch Soc Esp Oftalmol (Engl Ed). 2019;94:337-342.
34. Tan Z, Scheetz J, He M. Artificial Intelligence in Ophthalmology: Accuracy, Challenges, and Clinical Application. Asia Pac J Ophthalmol (Phila) 2019;8:197-199.