



OPEN

Machine Learning Based Non-Enhanced CT Radiomics for the Identification of Orbital Cavernous Venous Malformations: An Innovative Tool

Qinghe Han, MD,* Lianze Du, MD,* Yan Mo, MD,[†] Chencui Huang, MD,[†] and Qinghai Yuan, MD*

Purpose: To evaluate the capability of non-enhanced computed tomography (CT) images for distinguishing between orbital cavernous venous malformations (OCVM) and non-OCVM, and to identify the optimal model from radiomics-based machine learning (ML) algorithms.

Methods: A total of 215 cases of OCVM and 120 cases of non-OCVM were retrospectively analyzed in this study. A stratified random sample of 268 patients (80%) was used as the training set (172 OCVM and 96 non-OCVM); the remaining data were used as the testing set. Six feature selection techniques and thirteen ML models were evaluated to construct an optimal classification model.

Results: There were statistically significant differences between the OCVM and non-OCVM groups in the density and tumor location ($P < 0.05$), whereas other indicators were comparable (age, gender, shape, $P > 0.05$). Linear regression (area under the curve [AUC] = 0.9351; accuracy = 0.8657) and Stochastic Gradient Descent (AUC = 0.9448; accuracy = 0.8806) classifiers, both of which coupled with the f test and L1-based feature selection method, achieved optimal performance. The support vector machine (AUC = 0.9186; accuracy = 0.8806), Random Forest (AUC = 0.9288; accuracy = 0.8507) and eXtreme Gradient Boosting (AUC = 0.9147; accuracy = 0.8507) classifier combined with f test method showed excellent average performance among our study, respectively.

Conclusions: The effect of non-enhanced CT images in OCVM not only can help ophthalmologist to find and locate lesion, but also bring great help for the qualitative diagnosis value using radiomic-based ML algorithms.

Key Words: Machine learning, non-enhanced computed tomography, orbital cavernous venous malformation, radiomics

(*J Craniofac Surg* 2022;00: 00–00)

Orbital cavernous venous malformation (OCVM) is the most common orbital primary benign tumor. According to the latest International Society for the Study of Vascular Anomalies classification, OCVM accounts for 11% of all orbital primary masses.^{1,2} Orbital cavernous venous malformation present between 20 and 60 years of age and are more common in women, they are related to estrogen and progesterone according to previous research findings.³ Unilateral progressive exophthalmos in a gradient way is normally seen in OCVM patient, whereas usually, treatment strategy for whom is complete resection for the lesion.⁴ Treatment option as well as surgical paths are made due to precise preoperative diagnosis, for which computed tomography (CT) and magnetic resonance imaging (MRI) are most widely used.⁵ Previously, ultrasound was a noninvasive and inexpensive diagnostic method, but it had limited value in detecting the size, location, relationship of lesions, and relative to surrounding structures, which frequently led to missed diagnosis.⁶ Now non-contrast CT is the first choice of imaging examination method due to its low radiation and fast examination.⁷ Clinical studies have shown that most of OCVM located in the retrobulbar intraconal space, which is usually ovoid or round, lobulated when enlarged, well-circumscribed, and presented as smoothly margined lesions with homogeneous soft tissue density.⁸ By both MRI and CT scans, OCVM showed similar characteristics, whereas part of the lesion with point calcifications, and a few cases were irregularly lobulated.⁵

Differential diagnosis among OCVM and non-OCVM is of great importance, especially in the cases with similar images (schwannoma, fibrous histiocytoma, granulosa cell tumor, metastatic tumor, pleomorphic adenoma, etc).⁹ The similarities include size, location, morphology, density, and edge shape of the lesions, resulting in no differences on diagnosis accuracy between the junior and senior radiologist in differentiating OCVM from non-OCVM solely by non-enhanced CT or MRI.¹⁰ To facilitate better identification, enhanced CT or MRI scan were needed for qualitative diagnosis. According to the literature reports, most OCVM have the characteristics of progressive enhancement initiated from peripheral of lesions, which is the distinguishing feature from other similar space-occupying lesions. The accuracy of qualitative diagnosis of dynamic enhanced MRI can reach 90%, which is slightly better than CT enhancement.¹¹ Subject to dose of contrast agent, dynamic enhanced CT was seldom implemented, resulting in a relative lower rate of accuracy.¹² Beyond the condition stated above, some OCVM

From the *Radiology Department, The Second Hospital of Jilin University, Changchun; and [†]Deepwise AI Lab, Beijing Deepwise & League of PHD Technology Co, Ltd, Beijing, China.

Received August 16, 2021.

Accepted for publication November 27, 2021.

Address correspondence and reprint requests to Qinghai Yuan, MD, Radiology Department, The Second Hospital of Jilin University, Changchun 130041, China; E-mail: yuanqinghai123@sina.com

Qinghe Han and Lianze Du contributed equally to this work and co-first authors, the decision is agreed by all authors.

This work is supported by Jilin Province Education Department (JJKH20201062KJ), Department of Science and Technology of Jilin Province (20200404200YY).

The authors report no conflicts of interest.

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jcraniofacialsurgery.com).

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of Mutaz B. Habal, MD.

ISSN: 1049-2275

DOI: 10.1097/SCS.00000000000008446

samples quit the typical mode of enhancement, embarrassing whose differential diagnosis from fibroma, solitary fibrous tumor or malignant pleomorphic adenoma.¹⁰ Besides, enhanced MRI and CT are both expensive and time consuming examinations with potential risk of adverse reactions such as contrast agent extravasation and allergy, and contrast agent allergy, renal insufficiency, claustrophobia and similar paramagnetic foreign bodies such as pacemakers and implanted defibrillators are absolutely contraindicated.¹³ Completely surgically removed OCVM brings satisfactory prognosis, while incompletely removal of tumor capsule in some other solid tumors (the non-OCVMs) is prone to relapse and the prognosis is poor.⁴ Due to the large variety in treatment and prognosis, it is necessary to diagnose and different OCVM and non-OCVM in clinical practice. We did not tend to perform enhanced CT or MRI scans, and applied radiomics methods based on non-enhanced CT as a substitute to explore whether it could perform differential diagnosis of OCVM and non-OCVM.

In recent years, radiomics and artificial intelligence (AI) have been used more and more widely in medical imaging. Radiomics is to obtain a large number of advanced quantitative imaging features from medical images with high throughput, extracting feature data to extend the traditional pictures from which for further analyze.^{14–16} These quantitative characteristics, which are different from the visual images we perceive, can provide potential value to predict tumor development and prognosis. In this field, machine learning (ML) algorithms is employed to select the best features, and develops and improves models, which has the potential to improve predictive capabilities.¹⁷ Recent reports on orbital diseases, such as Guo and his colleagues,¹⁸ showed comparable accuracy between models trained with radiomic features and specialist in ophthalmic imaging diagnosis in distinguish orbital lymphoma and lymphoproliferative diseases. A recent AI study focusing on differential diagnosis OCVM and schwannoma achieved 90% of precision in judging lesion location and lesion type, which reach the level of expert radiologist with MRI enhancement images, proving that the neural network models based on large samples can quickly identify, diagnose and predict.¹⁹

In this study, we aimed to use non-enhanced CT images of the orbit to construct an ML model to assist clinicians in distinguishing between OCVM and non-OCVM. The model with high qualitative diagnostic ability may provide more valuable clinical information, facilitating decreasing in number of enhanced CT and MRI for clinical suspected OCVM patients for fewer side effects and lower costs. Meanwhile, the completion of this study will be able to provide ideas for the further research of large-scale OCVM radiomics and AI, and new research identifying different types of orbital tumors based on non-enhanced CT and MRI radiomic methods would be applicable.

MATERIALS AND METHODS

Patients

This retrospective study included 2 cohorts of patient undergoing CT examination between March 2014 and May 2020 at the Second Hospital of Jilin University, which were confirmed to be OCVM or non-OCVM by pathological diagnosis. The ethical approval was obtained, and the informed consent requirement was waived by our institutional reviewing board. The inclusion and exclusion criteria are as follows: Inclusion criteria: (1) Patients who underwent CT examination and were judged by doctors to be suspected of OCVM (non-enhanced CT showed occupying lesion inside or outside from the muscle cone, in shape of round or oval or slightly lobulated with smooth edge, which is isodense to the extraocular muscle with uniform density or punctate calcification), and then confirmed by pathological diagnosis as OCVM or non-

OCVM (including ocular pleomorphic adenoma, lymphoma, schwannoma, meningioma, adenoid cystic carcinoma, melanoma, etc); (2) No history of eye or orbital surgery, trauma, radiation therapy, or other local treatment; (3) Qualified image: complete, clear, and normal posture. Exclusion criteria: (1) No pathological diagnosis result; (2) Imaging manifestation of surgery, trauma, radiation therapy, or other local treatment of eye or orbit; (3) Orbital mass causing damage to adjacent bone; (4) Unqualified image: defective, unclear, and abnormal posture.

As shown in Figure 1, a total of 335 patients were included in the study, including 215 patients with OCVM, 120 patients with non-OCVM. Those 335 patients who satisfied each inclusion or exclusion criterion were identified for the whole cohort and were further assigned to either the training cohort or testing cohort randomly in a ratio of 8:2.

Based on the inclusion and exclusion criteria, a total of 335 patients were enrolled in this study as the whole cohort and were further distributed randomly to the training or testing cohort to explore and verify the discrimination performance of the model between OCVM and non-OCVM.

CT Acquisition

The equipment for image acquisition was Philips Brilliance 256 slice intelligence computed tomography (Philips Healthcare, Cleveland, OH). Patients took supine position, scanning range was from orbital floor to orbital roof. The parameters of the scanning were 120 kV/300 mA, 5 mm for layer thickness and 1 mm for reconstruction thickness. Finally, the image was transmitted to the post-processing workstation in real time, and the multi planar reconstruction was used for post-processing, in accordance with the medical digital image and communication standard.

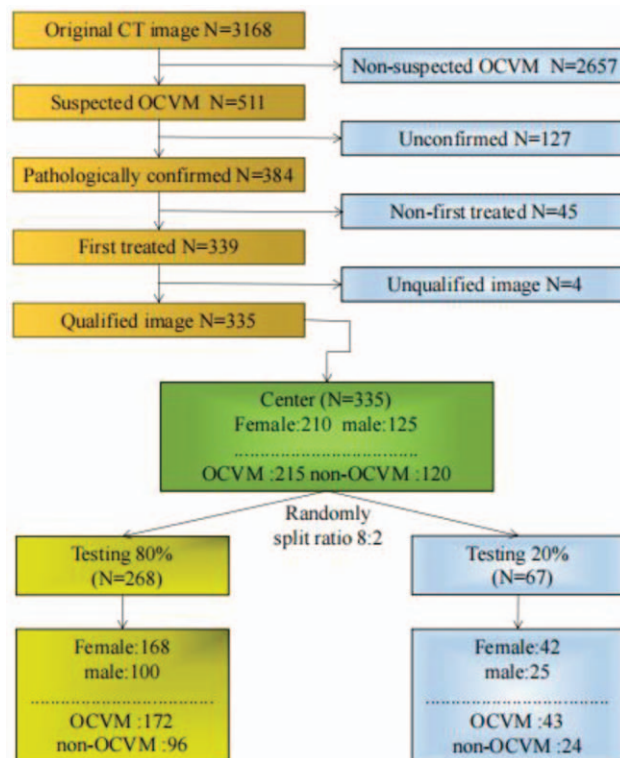


FIGURE 1. Flow chart of the selection of patients. Based on the inclusion and exclusion criteria, a total of 335 patients were enrolled in this study as the whole cohort and were further distributed randomly to the training or testing cohort to explore and verify the discrimination performance of the model between OCVM and non-OCVM. OCVM, orbital cavernous venous malformation.

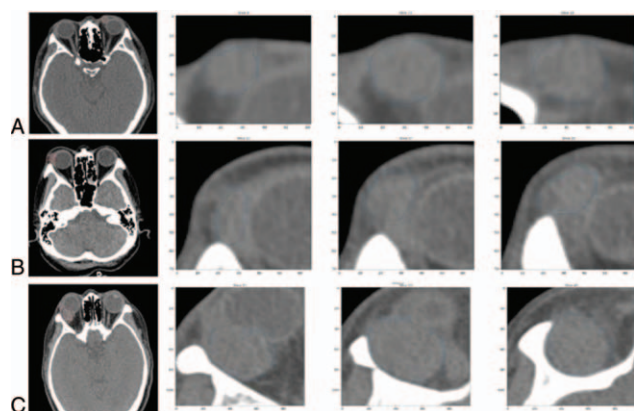


FIGURE 2. Some cases of lesions drawn by radiologists. (A) A 36-year-old male patient with orbital cavernous hemangioma (OCVM) outside the orbital muscle cone. (B) A 56-year-old woman with plasmacytoma (non-OCVM) outside the orbital muscle cone. (C) A 44-year-old male patient with pleomorphic adenocarcinoma (non-OCVM) located inside or outside the orbital muscle cone. OCV, orbital cavernous venous malformation.

Standard Reference

The gold standard for this study was the histopathology diagnosis. In order to facilitate doctors to draw CT slices and run ML models, all data were desensitized and uploaded to Dr. Wise Multimodal Research Platform (Beijing Deepwise & League of PHD Technology Co., Ltd, Beijing, China, <http://keyan.deepwise.com/>). Two 5-year-experienced radiologists visually evaluated OCVM tumors and drew region of interest (ROI) for each slice. Another 10-year-experienced radiologist confirmed the drawing effect (Fig. 2). All radiologists are blind to the histopathology diagnosis of each patient.

Radiomic Analysis

An overview of the study workflow including ROI segmentation, image preprocessing, feature extraction, feature selection, classification model construction, and model evaluation is illustrated in Figure 3. This study used the Dr. Wise Multimodal Research Platform (<https://keyan.deepwise.com>) (Beijing Deepwise & League of PHD Technology Co., Ltd, Beijing, China) for radiomics analysis, including image annotation, feature extraction and selection and modeling.

Image Preprocessing and Feature Extraction

We use B-spline interpolation sampling technology to resample images with different resolutions, making all images resampled to the same resolution. Based on the non-enhanced CT images, we carried out the pre-processing and image transformation. Ten image

preprocessing methods were used, including Wavelet, LOG, Square, Square Root, Logarithm, Exponential, Gradient, LBP2D, LBP3D, etc, and adjusted the optimal parameters.^{20,21} After image processing, a total of 1852 features were obtained from each ROI in our study, including 3 categories: first order feature, shape features, and texture features.²²

First order features (19 features) uses basic first-order statistics such as mean, variance, entropy, standard deviation, etc, to describe the pixel intensity and distribution within the ROI. Shape features (3D:16 features; 2D:10 features) describes the shape and size information of ROI in 2D or 3D, such as volume, diameter, roundness, etc. Texture features mainly include Gray Level Co-occurrence Matrix (24 features), Gray Level Size Zone Matrix (16 features), Gray Level Run Length Matrix (16 features), Gray Level Dependence Matrix (14 features), and Neighboring Gray Tone Difference Matrix (5 features), which describe the gray-scale relationship between a certain pixel and surrounding pixels. Our study named the features as 3 levels, the first is the image pre-processing method, the second is the feature type, and the third is the specific feature. For example, wavelet-LLH_glcml_lmc1, which means that the wavelet transform of LLH is used to process the image, and the lmc1 feature under Gray Level Co-occurrence Matrix in the texture feature is extracted.

Feature extraction was performed by a comprehensive open-source platform called PyRadiomics, which enables the processing and extraction of radiomic features from medical image data and is implemented in Python.²³

Inter-Class Correlation Coefficient (ICC) and Pearson Correlation Analysis

In order to eliminate the errors of 2 observers in plotting ROI at different times and obtain stable features, inter-observer consistency analysis was conducted. We calculated the ICC among all radiomic features of 40 patients. Stable features were retained with $ICC \geq 0.80$, which was in good agreement among observers. Then we perform Pearson correlation analysis to eliminate highly correlated features (Pearson correlation > 0.9) to reduce feature redundancy, and retain lower correlated features for further analysis.

Feature Selection

Feature selection simplifies the model by reducing the number of features, whose purpose is to improve learning efficiency, reduce over-fitting, enhance the generalization ability of the model, and improve the independence and reliability of features, whereas ensuring less information loss.

We used 6 widely used feature selections namely analysis of variance (F test), least absolute shrinkage and selection operator (LASSO), Pearson Correlation Coefficient, mutual information, tree model and Recursive feature elimination. Feature selection based on analysis of variance (F value) compares the means between different features and retains the features with significant differences.²⁴ Feature selection based on LASSO (L1-based) establishes a linear model on the training data and obtains a sparse coefficient matrix, and retains features with larger absolute values of coefficients.²⁵ Feature selection based on Pearson Correlation Coefficient calculates the linear correlation coefficient between features and dependent variables, and retains features with higher correlation coefficients.²⁶ Feature selection based on mutual information calculates the mutual information index between the feature and the dependent variable, and retains the feature with higher mutual information index.¹⁷ Feature selection based on the tree model builds a tree model on the training data, and calculates the impurity-based feature importance based on information impurity to select important features.²⁷ Recursive feature elimination fits the

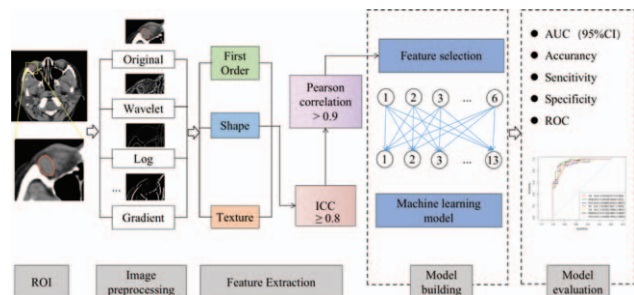


FIGURE 3. A flow chart of the study.

linear model iteratively on the training set, and removes some unimportant features based on the current model each time until the best performance was achieved on the training set.²⁸

The above method uses the open-source functions of “sklearn.-feature_selection” and “pyradiomics” of the python language.

Classification Algorithms

Classification is based on the label of training set to infer the outcome, which is a widely used supervised ML method.^{29,30} Our study tested 13 ML models, namely logistic regression (LR), support vector machine (SVM), linear SVC, decision tree, random forest (RF), AdaBoost, Gradient Boosting (GB), eXtreme Gradient Boosting, BernoulliNB, Gaussian Naive Bayesian (NB), KNear-estNeighbors (KNN), linear discriminant analysis, and Stochastic Gradient Descent (SGD).

All data is randomly divided into training and test cohort at a ratio of 8:2. In order to obtain a higher area under the receiver operating characteristic curve (AUC) of the test set, we conducted a global search to find the optimal parameters of the current model. The study uses indicators such as AUC, accuracy, sensitivity and specificity to comprehensively evaluate the ability of classification algorithms to distinguish between OCVM and non-OCVM.

Statistical Analysis

Clinical characteristics of the training set and testing set are statistically described using mean, standard deviation, frequency and percentage. To compare the statistical differences between the 2 cohorts, 2 independent-sample *t* tests were used for normally distributed continuous variables, Mann–Whitney *U* test was used for skewed distributions, and Pearson chi-square analysis was used for categorical variables. Delong test is used to compare the statistical differences of the ROC curves of different models.

Our study uses International Business Machines Corporation Statistical Product and Service Solutions (<https://spss.en.softonic.com/>), python 3.0 and R language (<https://cran.r-project.org/>) for data analysis. The 95% Confidence Interval is based on the double-precision exact test. *P* values <0.05 were deemed to be statistically significant.

RESULTS

Clinical Characteristics

A total of 335 cases were collected in this study, 215 OCVM patients and 120 non-OCVM patients, including 45 polymorphous adenoma, 13 lymphoma, 19 schwannoma, 6 solitary fibroma, 6 meningioma, 7 adenoid cystic carcinoma, 3 melanoma, 3 granular cell tumor, 2 molluscum simplex, 2 plasmacytoma, 2 fibroplastic proliferation, others are 1 for each (malignant pleomorphic fibroma, high-grade ductal adenocarcinoma, mesenchymal origin tumor, collagen fibroma, glioma, lymphoid hyperplasia, squamous cell carcinoma, chronic granuloma, sarcoma, invasive fibrous histiocytoma, fibroma, hemangioblastoma). The imaging findings of different tumors collected were as shown in Figure 4.

A stratified random sampling method was conducted to extract 80% (268 cases) of the total sample as the training set to train the model. The remaining 20% (67 cases) were used for testing the models' performance. In the training cohort, patients had an average age of 49.3 ± 10.2 years, 168 (62.7%) were female, 172 (64.2%) were OCVM. In the testing cohort, patients had an average age of 47.581 ± 12.309 years, 42 (62.7%) were female, 43 (64.2%) were OCVM. The details of the demographic and clinical characteristics of the training and testing cohorts as shown in Supplementary Digital Content, Table 1, <http://links.lww.com/SCS/D688>.

Two independent samples *t* test and chi-square analysis showed that there were statistically significant differences between the

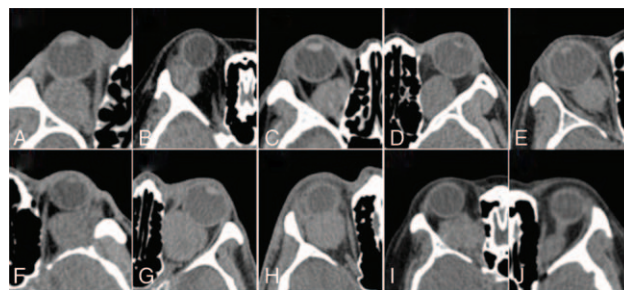


FIGURE 4. The imaging findings of different tumors collected. (A) Polymorphous adenoma; (B) lymphoma; (C) schwannoma; (D) solitary fibroma; (E) meningioma; (F) adenoid cystic carcinoma; (G) granular cell tumor; (H) molluscum simplex; (I) plasmacytoma; (J) fibroplastic proliferation.

OCVM and non-OCVM groups in tumor lesion location ($P < 0.05$), whereas there were no statistically significant differences in other indicators (age, gender, lesion density, lesion edge) ($P > 0.05$). There were no statistically significant differences between training cohort and testing cohort in clinical indicators such as age, gender, lesion density, lesion edge, lesion location, and pathological grouping ($P > 0.05$).

Radiomics Features

We extracted 1852 radiomic features from each ROI, of which 360 features were eliminated according to ICC analysis, and 1085 strongly correlated features were eliminated through Pearson correlation coefficient analysis. The Pearson correlation heat map of 50 features is shown in Figure 5. Several highly correlated features with a correlation coefficient close to 1 will eventually retain only 1 feature, and the remaining 407 features enter further analysis.

Performance of Various Machine Learning Models

In order to obtain a better classification performance, our study combined 6 feature selections and 13 ML models to form 78 classifiers and adjusted the parameter configuration of each model to be optimal through global super-parameter search. Afterward, we

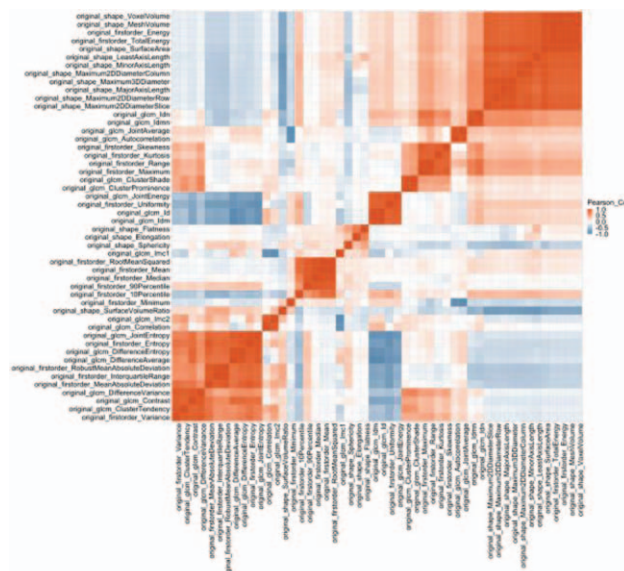


FIGURE 5. Pearson correlation heat map of 50 features.

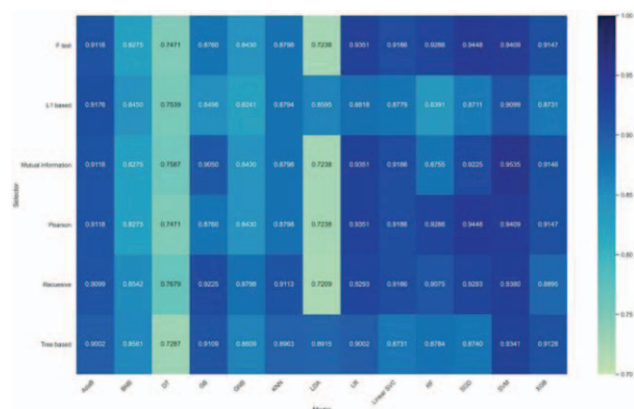


FIGURE 6. AUC values of test cohort for model combinations of different feature selection and different machine learning. AUC, area under the receiver operating characteristic curve.

select the classifier with the highest AUC under each training model combination to test the test set, and get the AUC values of different ML models on the test set as shown in Figure 6. We also got the AUC heat map on the test set as shown in Figure 6. The darker the blue was, the closer the AUC was to 1, which meant the better the performance of the current model in the test set was.

As shown in Supplementary Digital Content, Table 2, <http://links.lww.com/SCS/D689> the highest AUC (0.9448 and 0.9351), accuracy (0.8806 and 0.8657), sensitivity (0.907 and 0.8837), and specificity (0.8333 and 0.875) were shown separately in SGD and LR models on the test cohort. The SVM, RF, liner SVC, GB, and KNN models also show excellent AUC performance at 0.9186, 0.9288, 0.9186, 0.9225, and 0.9113, respectively. The performance of the other models in the test cohort was general. Area under the receiver operating characteristic curve and other values were not as good as those of the above 7 models. Therefore, in this study, the performance ability of the model was $\text{SGD} > \text{LR} > \text{SVC} = \text{RF} > \text{GB} > \text{SVM} > \text{KNN} > \text{others}$.

Detailed metrics on the test cohort for all models as shown in Supplementary Digital Content, Table 2, <http://links.lww.com/SCS/D689>.

The top 20 features and coefficients of the 6 ML models with higher AUC in Supplementary Digital Content, Table 2, <http://links.lww.com/SCS/D689> were presented in Figure 7. Among the top 20 features in the LR model, there are 14 texture features (70%), 4 first-order features (20%) and 2 shape features (10%). In SVM and SVC, there were 15 texture features (75%), 5 first-order features

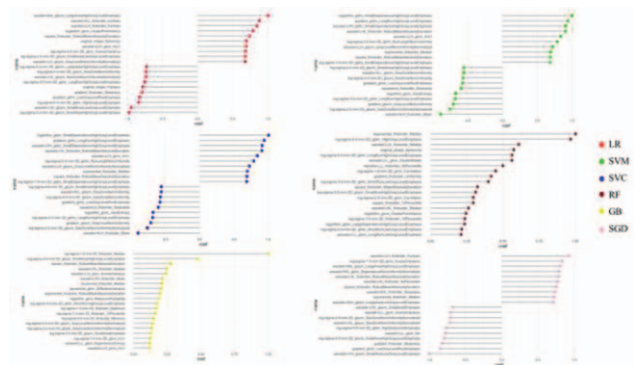


FIGURE 7. The top 20 features of different models.

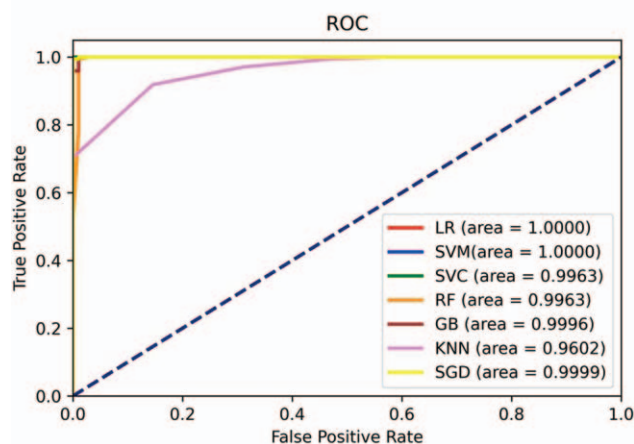
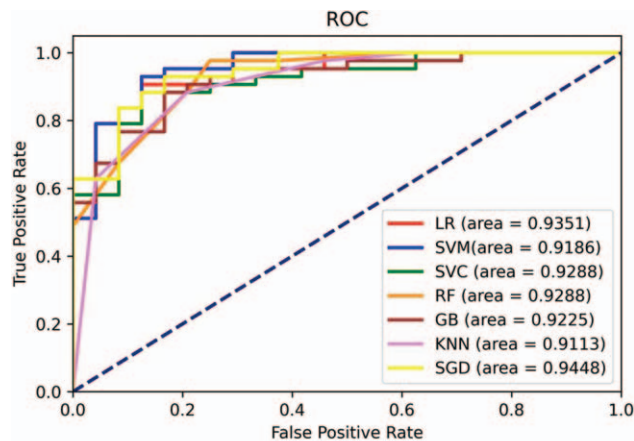


FIGURE 8. The ROC curve of 7 models in the test set and training set.

(25%), and 0 shape features (0%). In RF, there were 11 (55%), 8 (40%), and 1 (5%) respectively, whereas in GB, they were 11 (55%), 9 (45%), 0 (40%). In SGD, they were 13 (65%), 7 (35%), and 0 (40%). In our study, texture and first-order features contributed more to the classification model, which is related to the density and edge characteristics used by radiologists.

Finally, we summarized the ROC of the 7 types of ML models with higher AUC on the training set and test set, and obtained the results in Figure 8. The Delong test was used to test the ROC curves of the 7 models in pairs, and it was found that there were no significant differences between LR and SGD ($P = 0.8272$), RF and GB ($P = 0.56$), GB and KNN ($P = 0.7004$), whereas there were significant differences between the ROC curves of the other models in pairs ($P < 0.05$) on the test set. On the training set, there was no significant difference between ROC curves of all models ($P > 0.05$).

DISCUSSION

The incidence rate of OCRM is very high in orbital mass. It is necessary to identify OCRM early and quickly, which is helpful to slow down the progression of lesions and reduce the pressure of patients. Through the optimized clinical path, we can also quickly screen whether patients have orbital masses. In terms of quantification and reproducibility, ultrasound is highly dependent on the operator, which is not conducive to automation and tool implementation.⁶ Although MRI clearly show tissues and structures, the cost of MRI is high, the appointment and examination time is long, some

special populations are also absolutely contraindicated. Enhanced scan has the potential risk of adverse reactions such as contrast agent extravasation and allergy. Non-contrast CT is widely used, low radiation, fast examination, easy to be accepted by patients and clinical.

Orbital cavernous venous malformation is the most common disease among orbital masses, and it is also the first diagnosis to be ruled out when clinically diagnosing orbital space-occupying lesions. Non-enhanced CT is 1 of the first choice for OCVM, whereas differentiating OCVM from non-OCVM based on non-contrast CT images is challenging.⁵ The results of this study showed that gender, age, lesion location, shape, density, edge, etc could not make a rapid and accurate qualitative diagnosis of OCVM on non-enhanced CT. As reported, most of OCVM was located in the orbital muscle cone, but there were still a large proportion of cases located inside and outside the muscle vertebrae, or outside the muscle vertebrae.⁷ In this research, there were statistical differences between the test group and the verification group in the proportion of lesions inside the muscle cone, which made the diagnosis more difficult.

These patients collected previously did not treated with medication such as propranolol. Definite radiological diagnosis, especially the preoperative judgment of benign and malignant, plays a guiding role in the formulation of surgical plan. The mainstay options for management of OCVM include observation and surgery.³¹ If there are no clinical symptoms, the patient and treatment team can consider a period of observation or medication, otherwise, surgical treatment shall be adopted, the choice of surgical treatment approach depends on the location of the lesion within the orbit and its relationship to critical neurovascular structures.³² Non-contrast CT has the ability to meet this requirement.

In this study, the radiomics method based on non-enhanced CT was used to qualitatively diagnose the suspected OCVM, through multiple model optimization and feature screening. Among all models, SGD got the most satisfactory result of 0.9448 in AUC and LR was the second. Numerically, compared with imaging experts using plain scan plus dynamic enhanced MRI, this diagnostic performance in accuracy, sensitivity, and specificity was slightly lower.¹⁰ Since we only used CT plain scan image data, and it is not reasonable to make accurate qualitative diagnosis based on current medical knowledge in the clinic under such conditions, so we did not conduct a direct comparative analysis between imaging scientists and AI discussed in relative studies.^{33,34} In Bi's research, based T1-weighted contrast-enhanced sequence model, the diagnosis between OCVM and schwannoma reached an accuracy of 91.13%, a sensitivity of 86.84%, a specificity of 93.02%, and an AUC of 0.9535, which carried out similar diagnostic level.¹⁹ In another research, the result of MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation reached an accuracy of 73.02%, a sensitivity of 72.97%, a specificity of 73.08%, and an AUC of 0.73.¹⁸ The results of this study were significantly higher than that diagnostic performance in the lymphoma study. It may be due to differences in disease types, or differences in radiomics models or algorithms, but it basically reaches the diagnostic performance of imaging scientists, showing the potential application value and prospects of radiomic models in orbital diseases.³⁵

In this study, among the features of the classification model, the texture account for most of which. The regional characteristics of the heterogeneity between the focus and the peripheral region can more accurately distinguish OCVM and non-OCVM, revealing consistent trends with other studies.^{14–16} Texture feature refers to calculate the difference of regional heterogeneity according to gray level co-occurrence matrix, gray run matrix and gray area matrix.^{17,22} In previous reasonable studies in both radiomics and differentiating OCVM from schwannoma by AI model, different

heterogeneity was presented to reflect the differences of tumor metabolism, cell growth pattern and abnormal vascularization degree.^{18,19} However, in clinical applications, orbital hemangioma is not only to be differentiated from schwannoma, but also to include a variety of other diseases, including pleomorphic adenoma, fibrous histiocytoma, lymphoma, etc.⁵ Therefore, the purpose of this study was set to be true in clinical application, to distinguish among those suspected OCVM cases, which are diagnosed as OCVM, rather than choosing between 2 isolated diseases. Our results showed that radiomics models predicted OCVM well, providing evidence for further clinical interpretation.

In addition, there were a few false negative cases in the ML model on the test set. Technically, most of OCVM are single lesions. For radiologist, small lesion around the main part may disturb the judgement (Fig. 9A-B). This may also affected the model for predict. Orbital cavernous venous malformation mostly occurs in the muscle cone, with a clear boundary with the surrounding ocular muscles, and there is a small amount of fat between the lesions and the orbital apex. When the focus is located at the apex of the orbit, and the boundary between the lesion and the surrounding ocular muscles is not clear, the differential performance of the model may also be affected (Fig. 9C). For some patients with anemia or hypoproteinemia, the CT value of blood will reduce, inducing a declining in density value of lesion, which may potentially affect the result of model discrimination (Fig. 9D-E). Due to the relatively small number of such cases and the lack of clinical diagnosis experience, multicenter, large sample radiomics research and model optimization for such cases are our next research directions.

There were some limitations in this study. First, the cases we included were selected only when radiologists suspected a patient on OCVM by non-enhanced CT, which limited the sample size. Second, the radiomics model we used only conducted training and verification analysis on non-enhanced CT images. The CT enhanced sequence and non-enhanced MRI and multiphase enhanced sequence images have not yet been studied, which will be carried out in follow-up studies. Moreover, dynamic enhancement sequence plays a great role in the diagnosis and differential diagnosis of OCVM.³⁶ Establishment an AI model based on this dynamic sequence of human brain-eye recognition needs further research.

CONCLUSIONS

The radiomics model with ML method based on non-enhanced CT has achieved satisfactory results in the qualitative diagnosis of OCVM, which basically reaching the diagnostic efficiency of radiologist depending on images from CT and MRI enhanced scanning. Using this method, we can make a qualitative diagnosis

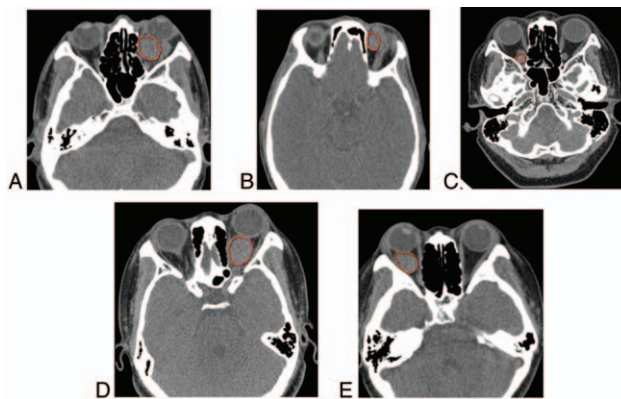


FIGURE 9. Pathologically confirmed OCVM cases were machine-identified as non-OCVM cases in the test set. OCVM, orbital cavernous venous malformation.

of OCVM in clinic without MRI scan and enhanced examination. This method can also reduce the use of contrast agents and the occurrence of side effects and save the time for appointments for examination and additional expenditure. Afterwards, we will explore the further large-scale verification of this research result and package it into a software, and experimentally embed it on the CT scanner to play a better role. In the future, our research team will optimize the our model, and implement further research on other diseases of the orbit based on which.

REFERENCES

1. Bonavolonta G, Strianese D, Grassi P, et al. An analysis of 2,480 space-occupying lesions of the orbit from 1976 to 2011. *Ophthalmic Plast Reconstr Surg* 2013;29:79–86
2. Shields JA, Shields CL, Scartozzi R. Survey of 1264 patients with orbital tumors and simulating lesions-The 2002 Montgomery Lecture, Part 1. *Ophthalmology* 2004;111:997–1008
3. Low CM, Stokken JK. Typical orbital pathologies: hemangioma. *J Neurol Surg B Skull Base* 2021;82:20–26
4. Strianese D, Bonavolontà G, Iuliano A, et al. Risks and benefits of surgical excision of orbital cavernous venous malformations (so-called cavernous hemangioma): factors influencing the outcome. *Ophthalmic Plast Reconstr Surg* 2021;37:248–254
5. Khan SN, Sepahdari AR. Orbital masses: CT and MRI of common vascular lesions, benign tumors, and malignancies. *Saudi J Ophthalmol* 2012;26:373–383
6. Zhang L, Li X, Tang F, et al. Diagnostic imaging methods and comparative analysis of orbital cavernous hemangioma. *Front Oncol* 2020;10:577452
7. Ansari SA, Mafee MF. Orbital cavernous hemangioma: role of imaging. *Neuroimaging Clin N Am* 2005;15:137–158
8. Harris GJ. Cavernous hemangioma of the orbital apex: pathogenetic considerations in surgical management. *Am J Ophthalmol* 2010;150:764–773
9. Mafee MF, Putterman A, Valvassori GE, et al. Orbital space-occupying lesions: role of computed tomography and magnetic resonance imaging. An analysis of 145 cases. *Radiol Clin N Am* 1987;25:529–559
10. Calandriello L, Grimaldi G, Petrone G, et al. Cavernous venous malformation (cavernous hemangioma) of the orbit: current concepts and a review of the literature. *Surv Ophthalmol* 2017;62:393–403
11. Tanaka A, Mihara F, Yoshiura T, et al. Differentiation of cavernous hemangioma from schwannoma of the orbit: a dynamic MRI study. *Am J Roentgenol* 2004;183:1799–1804
12. Unsold R, Hoyt WF, Newton TH. Ct-criteria of orbital hemangiomas and their importance in differential diagnosis of intraconal tumors (author's transl). *Klin Monbl Augenheilkd* 1979;175:773–785
13. Tonolini M, Campari A, Bianco R. Extravasation of radiographic contrast media: prevention, diagnosis, and treatment. *Curr Probl Diagn Radiol* 2012;41:52–55
14. Jiang C, Luo Y, Yuan J, et al. CT-based radiomics and machine learning to predict spread through air space in lung adenocarcinoma. *Eur Radiol* 2020;30:4050–4057
15. Gao X, Ma T, Cui J, et al. A CT-based radiomics model for prediction of lymph node metastasis in early stage gastric cancer. *Acad Radiol* 2020;28:e155–e164
16. Tobaly D, Santinha J, Sartoris R, et al. CT-based radiomics analysis to predict malignancy in patients with intraductal papillary mucinous neoplasm (IPMN) of the pancreas. *Cancers (Basel)* 2020;12:3089
17. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 2005;27:1226–1238
18. Guo J, Liu Z, Shen C, et al. MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation. *Eur Radiol* 2018;28:3872–3881
19. Bi S, Chen R, Zhang K, et al. Differentiate cavernous hemangioma from schwannoma with artificial intelligence (AI). *Ann Transl Med* 2020;8:710
20. Svoboda M, Matiu-Iovan L, Frigura-Iliasa F M, et al. B-spline interpolation technique for digital signal processing. International Conference on Information & Digital Technologies. IEEE. 2015
21. Pyradiomics. Community. <https://pyradiomics.readthedocs.io/en/latest/customization.html#image-types>
22. Pyradiomics. Community. <https://pyradiomics.readthedocs.io/en/latest/features.html>
23. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77:e104–e107
24. Villamil-Cubillos LF, Leon-Medina JX, Anaya M, et al. Evaluation of feature selection techniques in a multifrequency large amplitude pulse voltammetric electronic tongue. *Eng Proc* 2020;2:62
25. Wang S, Ji B, Zhao J, et al. Predicting ship fuel consumption based on LASSO regression. *Transp Res D Transp Environ* 2018;65:817–824
26. Patrik Waldmann. On the use of the Pearson correlation coefficient for model evaluation in genome-wide prediction. *Front Genet* 2019;10:899–1899
27. Han SH, Lu SX, Leung SC. Segmentation of telecom customers based on customer value by decision tree model. *Exp Syst Appl* 2012;39:3964–3973
28. Ke Yan, David Zhang. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sens Actuat B Chem* 2015;212:353–363
29. Bishop CM. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY: Springer-Verlag, Inc; 2006
30. Mitchell TM. *Machine Learning*. McGraw-Hill Science/Engineering/Math; 1997
31. Cheng JW, Wei RL, Cai JP. Transconjunctival orbitotomy for orbital cavernous hemangiomas. *Can J Ophthalmol* 2008;43:234–238
32. Ozturk O, Polat S, Inanli S, et al. Giant orbital cavernous hemangioma: a case report. *Otolaryngol Head Neck Surg* 2005;133:153–155
33. Song X, Liu Z, Li L, et al. Artificial intelligence CT screening model for thyroid-associated ophthalmopathy and tests under clinical conditions. *Int J Comput Assist Radiol Surg* 2021;16:323–330
34. Lin C, Song X, Li L, et al. Detection of active and inactive phases of thyroid-associated ophthalmopathy using deep convolutional neural network. *BMC Ophthalmol* 2021;21:39
35. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–762
36. Hou Y, Xie X, Chen J, et al. Bag-of-features-based radiomics for differentiation of ocular adnexal lymphoma and idiopathic orbital inflammation from contrast-enhanced MRI. *Eur Radiol* 2020;31:24–33