



# Differentiate cavernous hemangioma from schwannoma with artificial intelligence (AI)

Shaowei Bi<sup>1#</sup>, Rongxin Chen<sup>1#</sup>, Kai Zhang<sup>1,2#</sup>, Yifan Xiang<sup>1</sup>, Ruixin Wang<sup>1</sup>, Haotian Lin<sup>1,3</sup>, Huasheng Yang<sup>1</sup>

<sup>1</sup>State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China; <sup>2</sup>School of Computer Science and Technology, Xidian University, Xi'an, China; <sup>3</sup>Center for Precision Medicine, Sun Yat-sen University, Guangzhou, China

**Contributions:** (I) Conception and design: H Yang, H Lin; (II) Administrative support: H Lin; (III) Provision of study materials or patients: H Yang; (IV) Collection and assembly of data: S Bi, R Chen, Y Xiang; (V) Data analysis and interpretation: S Bi, K Zhang, R Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Professor Haotian Lin, MD, PhD; Professor Huasheng Yang, MD, PhD. Zhongshan Ophthalmic Center, Sun Yat-sen University, Xian Lie South Road 54#, Guangzhou, China. Email: haot.lin@hotmail.com; yanghs64@126.com.

**Background:** Cavernous hemangioma and schwannoma are tumors that both occur in the orbit. Because the treatment strategies of these two tumors are different, it is necessary to distinguish them at treatment initiation. Magnetic resonance imaging (MRI) is typically used to differentiate these two tumor types; however, they present similar features in MRI images which increases the difficulty of differential diagnosis. This study aims to devise and develop an artificial intelligence framework to improve the accuracy of clinicians' diagnoses and enable more effective treatment decisions by automatically distinguishing cavernous hemangioma from schwannoma.

**Methods:** **Material:** As the study materials, we chose MRI images as the study materials that represented patients from diverse areas in China who had been referred to our center from more than 45 different hospitals. All images were initially acquired on films, which we scanned into digital versions and recut. Finally, 11,489 images of cavernous hemangioma (from 33 different hospitals) and 3,478 images of schwannoma (from 16 different hospitals) were collected. **Labeling:** All images were labeled using standard anatomical knowledge and pathological diagnosis. **Training:** Three types of models were trained in sequence (a total of 96 models), with each model including a specific improvement. The first two model groups were eye- and tumor-positioning models designed to reduce the identification scope, while the third model group consisted of classification models trained to make the final diagnosis.

**Results:** First, internal four-fold cross-validation processes were conducted for all the models. During the validation of the first group, the 32 eye-positioning models were able to localize the position of the eyes with an average precision of 100%. In the second group, the 28 tumor-positioning models were able to reach an average precision above 90%. Subsequently, using the third group, the accuracy of all 32 tumor classification models reached nearly 90%. Next, external validation processes of 32 tumor classification models were conducted. The results showed that the accuracy of the transverse T1-weighted contrast-enhanced sequence reached 91.13%; the accuracy of the remaining models was significantly lower compared with the ground truth.

**Conclusions:** The findings of this retrospective study show that an artificial intelligence framework can achieve high accuracy, sensitivity, and specificity in automated differential diagnosis between cavernous hemangioma and schwannoma in a real-world setting, which can help doctors determine appropriate treatments.

**Keywords:** Artificial intelligence (AI); differential diagnosis; multicenter

Submitted Oct 23, 2019. Accepted for publication Feb 27, 2020.

doi: 10.21037/atm.2020.03.150

View this article at: <http://dx.doi.org/10.21037/atm.2020.03.150>

## Introduction

Cavernous hemangioma is one of the most common primary tumors that occur in the orbit, accounting for 3% of all orbital lesions (1-3), while schwannoma is a benign orbital tumor with a prevalence of less than 1% among all orbital lesions (1). It is necessary to distinguish these two tumors at treatment onset because they have different treatment strategies (4-6): complete removal is the treatment goal for cavernous hemangioma while for schwannoma, the goal is to ensure that no capsules remain. Moreover, clear differentiation provides useful information that fosters better vessel management (2). If the wrong surgical regimen is chosen, the tumor will recur, and the patient would need to undergo an additional operation.

Similar to the diagnosis of many other tumors, imaging techniques are applied as the predominant methods to diagnose these two tumors. Magnetic resonance imaging (MRI) is the most commonly used approach because of its high resolution, which clearly reflects the tissues to determine the appropriate surgical approach (5-7). However, because it manifests similarly to cavernous hemangioma, especially in MRI images, schwannoma often evokes an improper diagnosis (2,7,8) even highly experienced ophthalmologists or radiologists can make inaccurate diagnoses (9).

In recent years, the application of artificial intelligence (AI) in medicine has achieved physician-equivalent classification accuracy in the diagnosis of many diseases, including diabetic retinopathy (10-13), lung diseases (14), cardiovascular disease (15), liver disease, skin cancer (16), and thyroid cancer (17) and others.

Therefore, the goal of this project was to develop an AI framework that uses MRI image sets from 45 hospitals in China as input to automate the differential diagnosis between cavernous hemangioma and schwannoma with high accuracy, sensitivity and specificity.

## Methods

### Overall architecture

Considering the current dominance of MRI in the

differential diagnosis of the two studied tumor types, we selected MRI images as the research materials in this study. The research framework included of three types of functional models. Each type consists of eight groups of models with different arrangements and combinations of slice orientations (coronal and transverse) as well as weighted sequences (T1-weighted, T1-weighted contrast-enhanced, T2-weighted and T2-weighted fat suppression). Each group was sorted into 4 models and trained according to the principle of four-fold cross validation. In summary, a total of 96 models were obtained ( $3 \times 8 \times 4 = 96$ ) (*Figure 1*).

As mentioned above, we established 3 types of functional models to achieve the goal of distinguishing cavernous hemangioma from schwannoma. First, to reduce interference from unnecessary information, eye-positioning models were designed to identify the eye area from the complete images. Then, to further narrow the recognition range, tumor-positioning models were created to locate tumors within the identified eye area. Finally, tumor classification models were trained to classify the tumors. As shown in *Figure 2*, when an MRI image is input, the framework first delineates the eye area from the whole image; then it localizes the tumor scope from the eye area; and finally, it specifically classifies the tumor. The eye-positioning and tumor-positioning models were trained using the Faster-RCNN algorithm, while the tumor classification models used the ResNet-101 algorithm.

### Data set

The data set consisted of digital data scanned from MRI films representing patients from all over the country (most were from Southern China) who came to Sun Yat-sen University Zhongshan Ophthalmic Center (one of the most famous ophthalmic hospitals in China) for treatment. For all these patients, the diagnostic conclusions were supported by pathology and examined by the members of our team.

First, the MRI films brought by the patients from 45 different hospitals were scanned into a digital format and then screened, rotated and cropped. After this step, we obtained 6,507 images of cavernous hemangioma (from 33 different hospitals, *Table 1*) and 2,993 images of

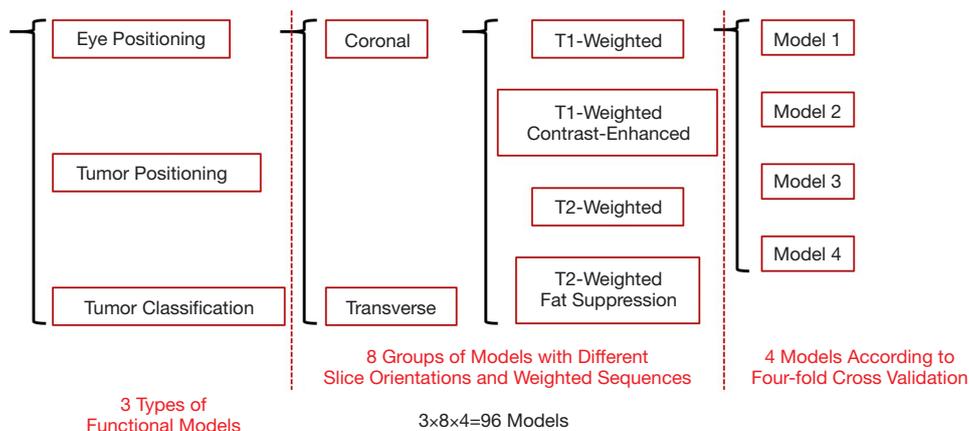


Figure 1 Branching diagram of all 96 models.

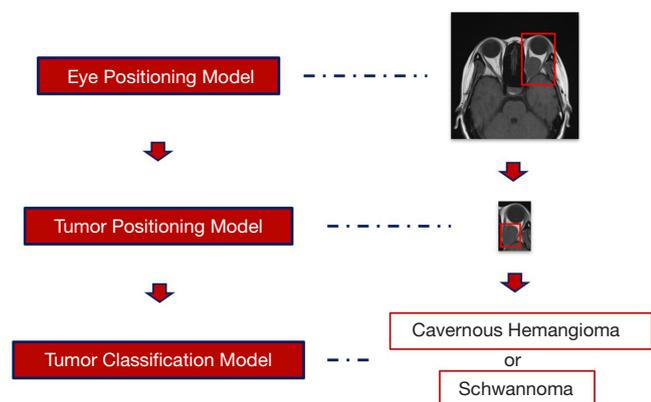


Figure 2 Work flow of the AI framework. AI, artificial intelligence.

schwannoma (from 16 different hospitals, Table 2). Then, to form training sets and validation sets, we used the image processing software named LabelImg [Tzutalin. LabelImg. Git code (2015). <https://github.com/tzutalin/labelImg>] to interpret and manually label all the images. The purpose of interpretation is to generate coordinates that delineate the extent of the ranges of eyes and tumors according to anatomical knowledge. The labels include eye, cavernous hemangioma and schwannoma supported by pathological diagnosis. Next, all these processed data were randomly divided into two parts: a training set and a validation set. The training set included 6,669 images for the eye-positioning model, 3,367 images for the tumor-positioning model and 3,131 images (2,059 images for cavernous hemangioma and 1,072 images for schwannoma) for the classification model. The validation set included 468 images for cavernous hemangioma and 217 images for schwannoma

(Table 3).

### Experimental settings

The settings of this study were based on Caffe (18), the Berkeley Vision and Learning Center deep-learning framework (BVLC), and TensorFlow (19). All the models were trained in parallel on three NVIDIA Tesla P40 GPUs.

In terms of the classification problem, the key performance evaluation metrics were estimated as follows (20):

$$Accuracy = (\sum_{i=1}^k TP_i) / N \tag{1}$$

$$Sensitivity_i (TPR, Recall) = TP_i / (TP_i + FN_i) \tag{2}$$

$$Specificity_i = TN_i / (TN_i + FP_i) \tag{3}$$

where N represented the quantity of samples;  $P_i$  represented the number of correctly classified samples within the  $i$ th class;  $k$  denoted the number of classes in this specific classification problem;  $TP_i$  indicated the number of correctly classified samples within the  $i$ th class;  $FP_i$  denoted the number of wrongly recognized samples within the  $i$ th class;  $FN_i$  denoted the number of wrongly classified samples within the  $j$ th class,  $j \in [1, c] / i$ ; and  $TN_i$  denoted the number of samples that were correctly recognized as not belonging to the  $j$ th class,  $j \in [1, c] / i$ . All these parameters can be integrated into a confusion matrix. Additionally, the receiver operating characteristic (ROC) curves (21), which indicated how many samples of the  $i$ th class were recognized conditioned on a specific number of  $j$ th class ( $j \in [1, c] / i$ ) samples classified as the  $i$ th class, together with the area under the curve (AUC), were adopted to assess the performance. The performance evaluation parameters (accuracy, sensitivity, specificity, and

**Table 1** Data-set sources: MRI images of cavernous hemangioma

Serial number	Hospitals	MRI images of cavernous hemangioma
1	Renai Hospital of Guangzhou	2,796
2	Guang Kong Hou Qin Hospital	930
3	The First Affiliated Hospital, Sun Yat-sen University	351
4	Guangzhou Panyu Central Hospital	186
5	Jiangmen Central Hospital	153
6	Unknown	147
7	Guangzhou General Hospital of PLA	146
8	Foshan second People's Hospital	117
9	The 458 PLA Hospital	129
10	Zhongshanyi Town Health Centre	120
11	Tianjin Huaxing Hospital	106
12	The Fifth Affiliated Hospital, Sun Yat-sen University	98
13	Shenzhen People's Hospital	93
14	Jiangxi Ji'an Central Hospital	89
15	Huizhou City People's Hospital	85
16	Anhui Yijishan Hospital of Wannan Medical College	77
17	Hainan General Hospital	72
18	Meizhou People's Hospital	72
19	Jiangmen Wuyi TCM Hospital	71
20	Hengyang Central Hospital	63
21	Liupanshui Mineral Bureau Hospital	62
22	Guangzhou Huaxing Kangfu Hospital	56
23	Affiliated Hospital of Xiangnan University	54
24	The Second Affiliated Hospital of Guangzhou Medical University	50
25	Guangzhou TCM No. 1 Hospital	50
26	Hainan Province Nongken Sanya Hospital	48
27	Jinshazhou Hospital of Guangzhou University of Chinese Medicine	47
28	Dongguan SDBRM Hospital	41
29	Maoming TCM Hospital	39
30	Jiangxi TCM Hospital	36
31	Maoming Nongken Hospital	35
32	Liuzhou City Worker Hospital	34
33	Beijing Boren Hospital	32
34	Armed Police Chengdu Hospital	22
Total		6,507

MRI, magnetic resonance imaging.

**Table 2** Data-set sources: MRI images of schwannoma

Serial number	Hospitals	MRI images of schwannoma
1	Renai Hospital of Guangzhou	1,609
2	Guang Kong Hou Qin Hospital	225
3	Unknown	148
4	Jiangxi People's Hospital	144
5	Guangdong Hospital of TCM	99
6	Shenzhen Hengsheng Hospital	95
7	Xinhui People's Hospital	87
8	Shenzhen Longgang Central Hospital	83
9	Guangdong Second TCM Hospital	80
10	Jiangsu Subei People's Hospital	79
11	Shenzhen Shekou Hospital	73
12	Foshan Hospital of TCM	72
13	Sanya City People Hospital	56
14	Huizhou Boluo People's Hospital	53
15	Guangzhou Huaxing Kangfu Hospital	41
16	Hunan Chenzhou First Hospital	30
17	Hainan Province Nongken Sanya Hospital	19
Total		2,993

MRI, magnetic resonance imaging.

**Table 3** Components of the training and validation sets

Slice orientation	Sequence	Training sets of eye positioning models	Training sets of tumor positioning models	Training sets of tumor classification models		Validation sets of tumor classification models	
				Cavernous hemangioma	Schwannoma	Cavernous hemangioma	Schwannoma
Coronal	T1-weighted	1,224	544	341	176	52	30
	T1-weighted contrast-enhanced	511	256	129	112	59	30
	T2-weighted	238	135	93	41	7	0
	T2-weighted fat suppression	185	108	57	45	22	0
Transverse	T1-weighted	1,276	623	368	203	81	43
	T1-weighted contrast-enhanced	1,211	612	397	171	86	38
	T2-weighted	1,016	530	326	150	79	39
	T2-weighted fat suppression	1,008	559	348	174	82	37
Total		6,669	3,367	2,059	1,072	468	217

**Table 4** AP of the eye-positioning models and tumor-positioning models

Sequence	AP of eye positioning models (%)	AP of tumor positioning models (%)
T1-weighted	100	100
T1-weighted contrast-enhanced	100	0
T2-weighted	100	100
T2-weighted fat suppression	100	100
T1-weighted	100	100
T1-weighted contrast-enhanced	100	100
T2-weighted	100	91
T2-weighted fat suppression	100	100

AP, average precision.

**Table 5** Performances of the tumor classification models

Slice orientation	Sequence	Performance of internal validation (%)			Performance of external validation (%)		
		Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Coronal	T1-weighted	89.76	80.49	94.19	69.51	66.67	71.15
	T1-weighted contrast-enhanced	92.74	91.67	93.75	60.67	93.33	44.07
	T2-weighted	94.44	90.91	96.00	–	–	–
	T2-weighted fat suppression	96.00	90.91	100.00	–	–	–
Transverse	T1-weighted	93.01	90.20	94.57	69.57	67.44	71.43
	T1-weighted contrast-enhanced	95.07	88.37	97.98	91.13	86.84	93.02
	T2-weighted	94.07	89.19	96.30	77.12	53.85	88.61
	T2-weighted fat suppression	93.02	79.07	100.00	64.71	86.49	54.88

ROC curve with AUC) were applicable only for binary classification problems. The accuracy and confusion matrix were applied to evaluate multiclass classification problems.

For the object positioning problem, interpolated average precision (AP) was adopted for the performance evaluation (22). The interpolated AP is computed from the precision recall (PR) curve as shown in Eq. [4].

$$AP = \frac{1}{11} \sum_{\gamma \in \{0, 0.1, \dots, 1\}} P_{interp}(\gamma) \quad [4]$$

$$P_{interp}(\gamma) = \max_{\tilde{\gamma} \geq \gamma} (p(\tilde{\gamma}))$$

where  $p(\tilde{\gamma})$  represents the measured precision at a specific recall value  $\tilde{\gamma}$ .

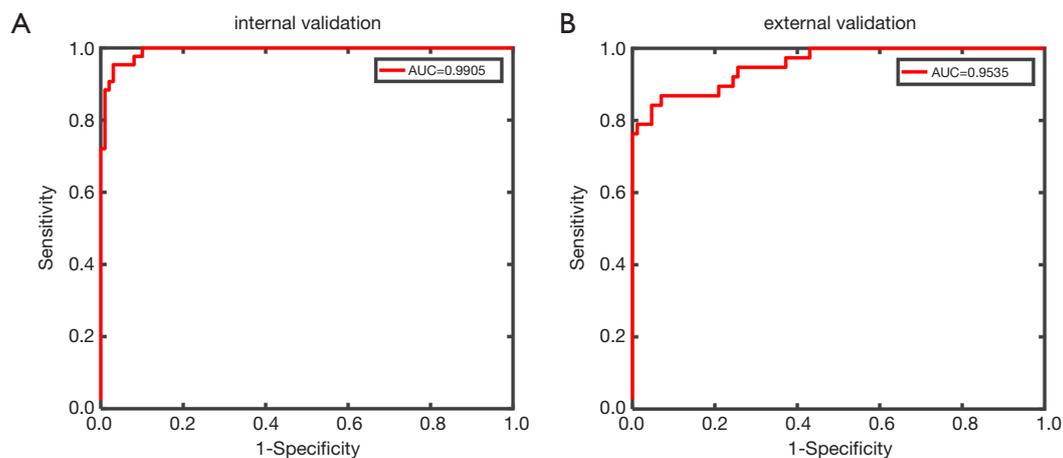
We adopted four-fold cross validation for the

performance evaluation to assess all the classification and positioning problems.

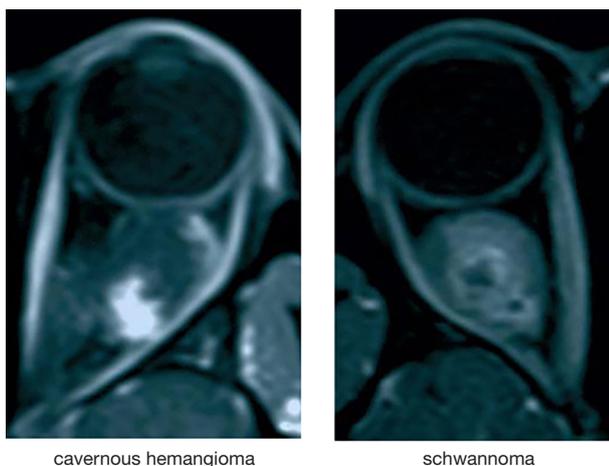
## Results

First, we conducted an internal four-fold cross-validation. The results showed that all the eye-positioning models achieved an AP of 100% and that the AP of the 28 tumor-positioning models exceeded 90% (Table 4). Similarly, the accuracy, sensitivity and specificity of almost all 32 tumor classification models were exceeded 90%, as shown in Table 5.

Next, we used the validation set for external validation. Considering that the tumor classification model results were mostly related to the differential diagnosis of cavernous



**Figure 3** Performance of the tumor classification model trained by the transverse T1-weighted contrast-enhanced sequence images.



**Figure 4** Manifestations of cavernous hemangioma and schwannoma in T1-weighted contrast-enhanced sequences.

hemangioma and schwannoma, the external verification of the tumor classification model predominantly represented the significance. The results showed that the transverse T1-weighted contrast-enhanced sequence model reached an accuracy of 91.13%, a sensitivity of 86.84%, a specificity of 93.02%, and an AUC of 0.9535. In contrast, the remaining models had significantly reduced performances compared with the internal verification results (see *Table 5* and *Figure 3*).

## Discussion

### *Good performance in a real-world setting*

Based on clinical experience, T1-weighted contrast-

enhanced sequences can highlight the blood vessels. Progressive filling from center to periphery on enhancement is typical of cavernous hemangioma, while the enhancement pattern of schwannoma is partial and uneven (5,6) (see *Figure 4*). Therefore, these sequences are considered the most significant reference among all types of slices in the differential diagnosis of the two studied tumor types (23,24). The tumor classification model trained by the transverse T1-weighted contrast-enhanced sequence images and tested on the external validation sets achieved high accuracy, sensitivity, and specificity in automated cavernous hemangioma and schwannoma differential diagnosis in a real-world setting that is completely consistent with the clinical environment.

Our results showed that the performance of the tumor classification model trained by transverse T1-weighted contrast-enhanced sequence images reached an accuracy of 91.13%, a sensitivity of 86.84%, a specificity of 93.02% and an AUC of 0.9535. These results suggested that this model's performance quality meets the primary need for clinical application and that the goal of distinguishing cavernous hemangioma from schwannoma is achievable using this type of model.

### *A multicenter data-set*

Thanks to the popularity of our ophthalmology center in China, patients from all over the country come here for treatment; thus, we were able to obtain these valuable images. In this study, we included data from more than 45 different hospitals in China to reach the current data

amount. Moreover, due to the variety of equipment and operators among the different source hospitals, the data collection techniques were diverse, which enhances the wide generalizability of our diagnostic model.

### *Applying scanned versions rather than using DICOM*

In previous AI studies, researchers have typically preferred raw data (11-13,15-17,25,26), such as DICOM format, generated directly from the imaging equipment, because the DICOM format both preserves all the original data and allows convenient collection. However, the scanned format was chosen for this study because the resultant AI framework needs to be useful for doctors in remote areas. The information technology level of hospitals in remote areas was limited, and they often lack comprehensive medical record management systems (27,28). Because most clinicians rely on film images instead of computerized interfaces, it made sense that models trained from a film version would be more suitable in this type of situation.

### *Three steps to reach the final goal*

In previous studies, researchers commonly input entire MRI images for training (25,26). Here, we progressively designed three different types of models to achieve the goal of distinguishing cavernous hemangioma from schwannoma. First, because the eye area occupies only a small proportion of the entire MRI image, inputting the entire MRI image into the model directly would introduce considerable irrelevant information. To reduce the interference from such unnecessary information, we constructed an eye-positioning model that identifies the eye range within the full image; then, subsequent process can focus only on this range. Second, we established a tumor-positioning model to further narrow the scope for the final classifier and improve its precision. Third, we built a classification model to differentiate the located tumors to achieve the goal of automatically differentiating cavernous hemangioma from schwannoma.

### *Further subdividing the training sets instead of combining them*

According to the traditional wisdom, having sufficient data volume is the foundation of training the current AI techniques (11-17). The most fundamental and effective way to improve the accuracy, sensitivity and specificity of the

model is to augment the data in the training set. However, the MRI images for training had remarkable variations in different weighted sequences and slice orientations. If these images were blindly combined while ignoring these variations, the resultant incompatibilities would inevitably confuse the system, and its performance would deviate from the original intention. Therefore, we divided all the images into eight groups for training based on their different weighted sequences and slice orientations. The final result supported our conjecture: the performance of the transverse T1-weighted contrast-enhanced sequence was outstanding compared to that of other models. If all the training sets were combined, the accuracy of this model would be well below 91.13%.

### *Web-based automatic diagnostic system*

Early in our research, our team built a cloud platform for congenital cataract diagnosis (29); we will implement the models in this study on that platform at the appropriate time. In China, an objective technological gap exists between urban and rural areas, and this imbalance is particularly evident with regard to medical resources (30-32). The establishment of this AI cloud platform for disease diagnosis is an economical and practical approach to alleviate the problem of the uneven distribution of medical resources.

### *Proper algorithms*

#### **Localization method**

Faster-RCNN is a widely used algorithm used to address positioning problems because of its practicability and efficiency. Evolving from RCNN and Fast-RCNN (33), Faster-RCNN generates region proposals quickly by using an anchor mechanism rather than by applying a superpixel segmentation algorithm. After adopting two-stage training, transformations of the bounding box regressor and classifier were achieved. In the first stage, Faster-RCNN generated region proposals. Then, it judged the authenticity of the proposals, and the topmost coordinates of each object were regressed. In the second stage, the class of each object was evaluated and each object was eventually regressed to obtain its coordinates. We adopted a pretrained Zeiler and Fergus (ZF) network (34) to reduce the training time.

#### **Convolutional neural network (CNN)**

The CNN is the most popular AI model used in medicine. In this study, we adopted ResNet, which has a thin CNN

architecture that includes numerous cross-layer connections and is suitable for rough classification tasks. To fit the residual function, the objective function was transformed, which resulted in a significant increase in efficacy, and we adopted a LogSoftMax loss function with class weights. The ResNet selected for this study has 101 layers, which is a sufficient depth to address the classification problems (20).

### *Limitations of our study*

The most important deficiency in the study is that we simply chose a model that achieved good efficacy rather than also considering other models. Although the model trained on the group containing the transverse T1-weighted contrast-enhanced sequence images achieved particularly remarkable performance and is already sufficiently robust to help doctors in clinical work, the other seven groups may also contain useful information for feature extraction. Thus, the diagnostic efficiency of the model should be improvable to some extent if we were to make rational use of the other seven groups of data. Such an approach requires involving multimodal machine learning (35-37), because MRI images with different weighted sequences should be processed as separate modes. Upon alignment, the models could be integrated under the joint representation principle. Our team will continue to investigate this aspect of the problem in future studies.

### **Conclusions**

The findings of our retrospective study show that the designed AI framework tested on external validation sets can achieve high accuracy, sensitivity, and specificity for differential diagnosis of automated cavernous hemangioma and schwannoma in real-world settings, which will contribute to the selection of appropriate treatments.

Although a partial accuracy rate of over 90% was achieved with the current data volume, AI algorithms can never have too much data. Thus, we plan to continue collecting additional cases to optimize the model by cooperating with hospitals in Shanghai to collect data in the eastern part of China, thereby supplementing our training set and enhancing model generalizability. Furthermore, at the appropriate time, we will design a web-based automatic diagnostic system to help solve the problem of obtaining advanced medical care in remote areas. In terms of algorithms, we will first investigate multimodal machine learning to take full advantage of these invaluable data.

Overall, the results show that further investigation of AI approaches are clearly a worthwhile effort that should be tested in prospective clinical trials.

### **Acknowledgments**

*Funding:* This study was funded by the National Key R&D Program of China (2018YFC0116500), the National Natural Science Foundation of China (81670887, 81870689 and 81800866), the Science and Technology Planning Projects of Guangdong Province (2018B010109008), and the Ph.D. Start-up Fund of the Natural Science Foundation of Guangdong Province of China (2017A030310549).

### **Footnote**

*Provenance and Peer Review:* This article was commissioned by the Guest Editors (Haotian Lin and Limin Yu) for the series “Medical Artificial Intelligent Research” published in *Annals of Translational Medicine*. The article was sent for external peer review organized by the Guest Editors and the editorial office.

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm.2020.03.150>). The series “Medical Artificial Intelligent Research” was commissioned by the editorial office without any funding or sponsorship. HL served as the unpaid Guest Editor of the series. The other authors have no other conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Data collection, analysis, and publication of this study were approved by the Ethics Committee of the Zhongshan Ophthalmic Center (No. 2016KYPJ028) according to the principles of the Declaration of Helsinki.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the

license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

- Shields JA, Bakewell B, Augsburger JJ, et al. Classification and incidence of space-occupying lesions of the orbit. A survey of 645 biopsies. *Arch Ophthalmol* 1984;102:1606-11.
- Tanaka A, Mihara F, Yoshiura T, et al. Differentiation of cavernous hemangioma from schwannoma of the orbit: a dynamic MRI study. *AJR Am J Roentgenol* 2004;183:1799-804.
- 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.
- Scheuerle AF, Steiner HH, Kolling G, et al. Treatment and long-term outcome of patients with orbital cavernomas. *Am J Ophthalmol* 2004;138:237-44.
- Kapur R, Mafee MF, Lamba R, et al. Orbital schwannoma and neurofibroma: role of imaging. *Neuroimaging Clin N Am* 2005;15:159-74.
- Ansari SA, Mafee MF. Orbital cavernous hemangioma: role of imaging. *Neuroimaging Clin N Am* 2005;15:137-58.
- 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.
- Young SM, Kim YD, Lee JH, et al. Radiological Analysis of Orbital Cavernous Hemangiomas: A Review and Comparison Between Computed Tomography and Magnetic Resonance Imaging. *J Craniofac Surg* 2018;29:712-6.
- Savignac A, Lecler A. Optic Nerve Meningioma Mimicking Cavernous Hemangioma. *World Neurosurg* 2018;110:301-2.
- Li Z, Keel S, Liu C, et al. An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. *Diabetes Care* 2018;41:2509-16.
- Ting DS, Cheung CY, Lim G, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* 2017;318:2211-23.
- Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016;316:2402-10.
- Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. *Ophthalmology* 2017;124:962-9.
- Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018;24:1559-67.
- Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2:158-64.
- Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-8.
- Li X, Zhang S, Zhang Q, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *Lancet Oncol* 2019;20:193-201.
- Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional Architecture for Fast Feature Embedding. *arXiv:1408.5093 [cs.CV]* 2014:675-8.
- Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. *arXiv:1603.04467 [cs.DC]* 2016:19.
- Zhang K, Liu X, Liu F, et al. An Interpretable and Expandable Deep Learning Diagnostic System for Multiple Ocular Diseases: Qualitative Study. *J Med Internet Res* 2018;20:e11144.
- Shi C, Pun C. Superpixel-based 3D deep neural networks for hyperspectral image classification. *Pattern Recogn* 2018;74:600-16.
- Everingham M, Zisserman A, Williams CKI, et al. The 2005 PASCAL Visual Object Classes Challenge. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006:117-76.
- Jinhu Y, Jianping D, Xin L, et al. Dynamic enhancement features of cavernous sinus cavernous hemangiomas on conventional contrast-enhanced MR imaging. *Am J Neuroradiol* 2008;29:577-81.
- He K, Chen L, Zhu W, et al. Magnetic resonance standard for cavernous sinus hemangiomas: proposal for a diagnostic test. *Eur Neurol* 2014;72:116-24.
- Yasaka K, Akai H, Abe O, et al. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology* 2018;286:887-96.
- Lu Y, Yu Q, Gao Y, et al. Identification of Metastatic Lymph Nodes in MR Imaging with Faster Region-

- Based Convolutional Neural Networks. *Cancer Res* 2018;78:5135-43.
27. Williams C, Asi Y, Raffenaud A, et al. The effect of information technology on hospital performance. *Health Care Manag Sci* 2016;19:338-46.
  28. Li H, Ni M, Wang P, et al. A Survey of the Current Situation of Clinical Biobanks in China. *Biopreserv Biobank* 2017;15:248-52.
  29. Long E, Lin H, Liu Z, et al. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nat Biomed Eng* 2017;1:0024.
  30. Zhang T, Xu Y, Ren J, et al. Inequality in the distribution of health resources and health services in China: hospitals versus primary care institutions. *Int J Equity Health* 2017;16:42.
  31. Anand S, Fan VY, Zhang J, et al. China's human resources for health: quantity, quality, and distribution. *Lancet* 2008;372:1774-81.
  32. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. *Lancet Oncol* 2014;15:e279-89.
  33. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mach Intell* 2017;39:1137-49.
  34. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. Cham: Springer International Publishing, 2014:818-33.
  35. Baltrusaitis T, Ahuja C, Morency LP. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Trans Pattern Anal Mach Intell* 2019;41:423-43.
  36. Atrey PK, Hossain MA, El Saddik A, et al. Multimodal fusion for multimedia analysis: a survey. *Multimedia Syst* 2010;16:345-79.
  37. Ramachandram D, Taylor GW. Deep Multimodal Learning: A Survey on Recent Advances and Trends. *IEEE Signal Proc Mag* 2017;34:96-108.

**Cite this article as:** Bi S, Chen R, Zhang K, Xiang Y, Wang R, Lin H, Yang H. Differentiate cavernous hemangioma from schwannoma with artificial intelligence (AI). *Ann Transl Med* 2020;8(11):710. doi: 10.21037/atm.2020.03.150