Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Computer-Aided Age-Related Macular Degeneration Diagnosis with the Fusion of Both Color Fundus and Fluorescein Angiography

Zekuan Yu
Peking University
Jianchen Hao
Peking University First Hospital
Zifeng Tian
Peking University
Bin Qiu
Peking University
Shujin Zhu
Nanjing University of Posts and Telecommunications
Wenbo Zhang
Peking University First Hospital
Hongping Nie
Peking University First Hospital
Shijie Zhang
Peking University First Hospital
Jun Li
Peking University First Hospital
Liu Yang
Peking University First Hospital
Qiushi Ren
Peking University
Yanye Lu (yanye.lu@pku.edu.cn)
Peking University https://orcid.org/0000-0002-3063-8051

Research

Keywords: age-related macular degeneration, deep learning, dual modality, computer aided diagnosis

Posted Date: August 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-53293/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Computer-aided Age-related Macular Degeneration Diagnosis with the Fusion of Both Color Fundus and Fluorescein Angiography

ZEKUAN YU[1,3,4*], JIANCHEN HAO[2*], ZIFENG TIAN[1,3,4*], BIN QIU[1,3,4], SHUJIN ZHU[6], WENBO ZHANG[2], HONGPING NIE[2], SHIJIE ZHANG[2], JUN LI[2], LIU YANG[2 §], QIUSHI REN[1,3,4], AND YANYE LU[1,3,4,5 §]

Abstract

Background: Age-related macular degeneration (AMD) is one of the most severe vision-threatening diseases, and yet Fundus Fluorescein Angiography (FFA) is the gold standard for AMD diagnosis. In recent years, many AMD computer-aided diagnosis (CAD) systems have been developed based on either color fundus images or OCT images. However, there is no CAD technique that integrates FFA with other ophthalmic imaging so far.

Methods: In order to improve the performance of AMD CAD system, we propose a pioneering CAD pipeline that combines color fundus and FFA photography. This novel pipeline is the first work that incorporates FFA with any other modality. Six deep neural networks (ResNet-18, ResNet-50, ResNet-101, Inception-V3, Inception-ResNetV2, and DenseNet-201) were utilized to extract feature vectors to facilitate five classifiers (Random Forest, K-Nearest Neighbor, and Support Vector Machine with Linear, Gaussian, and Quadratic functions) for AMD diagnosis. The pipeline was validated on 664 pairs of color fundus and FFA images using 10-fold cross-validation.

Results and conclusion: The accuracy and area under curve (AUC) value achieves 93.8% and 0.97, respectively. The results demonstrate that combining color fundus images and FFA images in CAD system is beneficial for AMD diagnosis, indicating promising potential to clinical practice in the future.

Key Words age-related macular degeneration; deep learning; dual modality; computer aided diagnosis

1. Background

The aging of society has become a global problem as the population ages above 60 will increase from a current number of 901 million to 1.4 billion by 2030 [1]. Elderly people are expecting a longer and better life nowadays [2], nevertheless, age-related diseases are still threatening the expectancies. The emerging problem of population aging and age-related diseases bring heavier burdens on healthcare providers. Among various age-related diseases, age-related macular degeneration (AMD) is one of the

most severe vision-threatening diseases in people age 50 and older in the developed countries[3-5]. In the United States, there are two million advanced AMD patients and more than eight million mid-term AMD patients. Moreover, this number will even increase by 50% by 2020 [6]. As society ages, the number of AMD patients would increase significantly in the next decades, which decreases the quality of elderly people's life, affects their relatives, and brings an increasing burden on economies [2].

AMD is a chronic eye disease that mainly affects the macular area in the retina. The macular area that are responsible for most visual functions and visual acuity the influences central vision [7]. Basically, the AMD is of two types: non-neovascular (i.e. dry), and neovascular (i.e. wet). Dry AMD can be further divided into early stage, middle stage and late stage (geographic atrophy). At the early stage, drusen, which are yellow sediments in different sizes, shapes, and distributions, can be observed under the retina [8]. As the disease progresses, the number and size of drusen will increase. At the late stage, geographic atrophy results in progressive atrophy of RPE, capillaries, and photoreceptor cells. In wet AMD, choroidal neovascularization breaks through the retinal pigment epithelium (RPE) layer and reaches the retina, causing leakage of liquid, lipid, and blood, then finally leads to the formation of a fibrous scar.

There was no effective treatment for AMD until the discovery of anti-vascular endothelial growth factor (anti-VEGF) therapies. With anti-VEGF therapies, the prognosis of AMD has undergone earth-shaking changes: nearly 95% of patients can avoid vision loss, and 40% of patients get improvements in their vision [9-11]. Since early detection and early treatment are critical to prevent severe visual impairment, the correct clinical diagnosis and classification of AMD is crucial, which directly affects the prognosis of patients.

In order to early diagnose AMD, techniques such as fundus color photography, optical coherent tomography (OCT), and fundus fluorescein angiography (FFA) are widely used to detect comprehensive information of the clinical manifestations in the posterior pole such as drusen, map-like atrophy, choroidal neovascularization, etc. [12, 13] Among all the retinal imaging methods used for AMD detection, color fundus photography is the most convenient imaging method. However, some lesions (e.g., leakages) cannot be clearly observed on color fundus images for the ophthalmologists to diagnosis AMD. In fact, FFA represents the golden standard for AMD diagnosis, since it can precisely detect the leakage of dye (hyper fluorescence), neovascular, drusens and other lesions. Nevertheless, differential diagnosis of AMD takes a lot of time, and the conclusion is often subjective [14]. Therefore, CAD

systems can help ophthalmologists reduce workload and improve diagnostic accuracy. In recent years, many AMD CAD systems based on color fundus images as well as some other modalities have been developed [15-23]. In their pipelines, various feature extracting approaches with different classifiers have been demonstrated to achieve promising results. Mookiah et al. [16] applied Local Configuration Pattern (LCP) features and Support Vector Machine (SVM) classifier to the AMD screening. In their research [17] and [18], discrete wavelet transform (DWT) and the empirical mode decomposition (EMD) were respectively examined as feature extractors for AMD diagnosis. They [19] also surveyed the performance of greyscale features including various entropies, Higher Order Spectra (HOS), bispectra features, Fractional Dimension (FD), and Gabor wavelet.

Classifiers consisting of Naive Bayes (NB), k-Nearest Neighbour (k-NN), Probabilistic Neural Network (PNN), Decision Tree (DT) and Support Vector Machine (SVM) were also examined and evaluated in the research [19]. Acharya et al.[20] proposed a AMD CAD tool with pyramid of histogram of oriented gradients (PHOG) features to diagnose normal, dry AMD, and wet AMD, which is the first study on three-class AMD color fundus images. In their related research [21], Radon Transform (RT), DWT, and Locality Sensitive Discriminant Analysis (LSDA) were utilized as feature extractor, and DT, SVM, PNN, and k-NN were applied as classifiers for AMD identification. They [22] also used Bi-dimensional Empirical Mode Decomposition (BEMD) technique and various entropy methods as feature extractor in their Posterior Segment Eye Diseases (PSED) diagnosis, which include diabetic retinopathy (DR), glaucoma, and AMD. Feature extraction is considered essential but intractability in applied machine learning, therefore extracting adequate features always bother researchers. Some researchers have applied convolutional neural networks (CNNs) to AMD diagnosis. Tan et al. [23] proposed a AMD CAD system with a 14 layer CNN model, achieving 95.45% accuracy with ten-fold cross-validation, which is the first work to apply deep learning method to perform AMD CAD. Serener et al. applied ResNet model to fast classify dry and wet AMD based on the OCT images [24]. Nevertheless, none of these researches used FFA imaging. In addition, some of their datasets are small, which limits the generalization of their results.

Represented by CNNs, deep neural networks have shown significant advantages in automatic feature extraction. However, it is worth mentioning that fined extracted features can boost the performance of classifiers such as SVM, random forests, and even is superior than fully connected deep neural networks. Especially, the computational cost is saliently reduced. Therefore, we propose a new

AMD CAD pipeline that employs deep neural networks as representative learning to extract features along with binary classifiers. Both color fundus and FFA images were utilized to train the CAD systems for AMD diagnosis. In this study, six pre-trained CNN models (including ResNet-18, ResNet-50, ResNet-101, Inception-V3, Inception-ResNetV2, and DenseNet-201) were implemented to extract the features 664 pairs of color fundus and FFA images, while five binary classifiers (including Random Forest, K-Nearest Neighbor, and Support Vector Machine with Linear, Gaussian, and Quadratic functions) were investigated using 10-fold cross-validation as well.

2. Results

2.1. Evaluation Metrics

The AMD CAD task is evaluated on image classification of sensitivity (Sen), specificity (Spe) and accuracy (Acc). These evaluation indexes are computed from the following equations:

$$Sensitivity(Sen) = TP/(TP + FN) \tag{12}$$

$$Specificity(Spe) = TN/(TN + FP) \tag{13}$$

$$Accuracy(Acc) = (TP + TN)/(TP + TN + FP + FN) \tag{14}$$

TP (True Positive) represents the number of AMD disease patients correctly classified as AMD patients. TN (True Negative) is the number of normal people who are correctly classified as normal people. FP (False Positive) and FN (False Negative) are misclassified as the wrong predicted label. Furthermore, the receiver operating curve and the value of the area under curve (AUC) is used to evaluate the AMD 10-fold cross-validation classification results. Detailed results are listed in Table 2 and Table 3. The results in Table 2 demonstrate that for all CNN models, in terms of the accuracy, sensitivity, specificity, and AUC, using dual-modality images for diagnosis is the most competitive, i.e., it performs better than using only color fundus or FFA images. This result also reveals the potential advantage of combining feature vectors of different modalities which are extracted by deep neural networks. Among all the state-of-the-art models, Resnet-50 achieved the best performance, proving that this feature representation is an effective off-the shelf descriptor for AMD screening task. In addition, SVM with quadratic kernel function outperformed all the other examined classifiers such as K-NN, and Random Forest. The dual modality CAD system achieved 93.8% accuracy, 97% sensitivity, 74% specificity, which outperforms the single color fundus CAD system and single FFA system. We can observe in figure

1 that the AUC of single color fundus is 0.86, that of single FFA is 0.95, and that of dual modality is 0.97. Multiple modalities images include color fundus and FFA images contain complementary information which boost the classification accuracy.

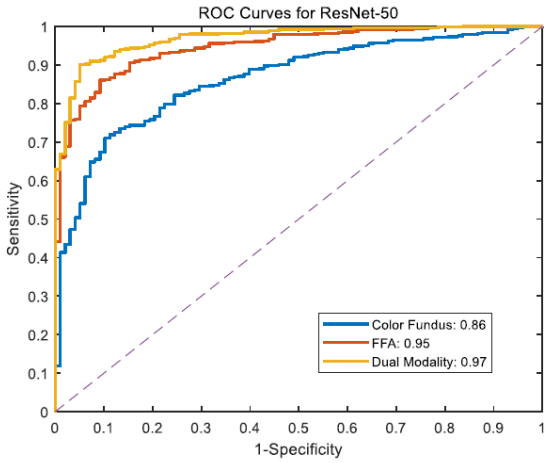


Figure 1: ROC curve of ResNet-50.

TABLE 2: The performance of AMD diagnosis from different CNN models

Color Fundus Image				FFA Image				Dual Modality			
Acc (%)	Sen (%)	Spe (%)	AUC	Acc (%)	Sen (%)	Spe (%)	AUC	Acc (%)	Sen (%)	Spe (%)	AUC
85.8	95	32	0.85	90.2	93	58	0.92	90.4	95	63	0.95
85.8	93	42	0.86	92	95	69	0.95	93.8	97	74	0.97
83.1	93	23	0.82	90.1	95	64	0.93	92.2	97	66	0.96
83.9	95	22	0.81	89.8	95	57	0.92	90.7	95	63	0.95
86	94	38	0.83	88.7	94	59	0.94	91.6	96	66	0.96
85.8	95	33	0.84	90.2	95	65	0.94	90.5	96	65	0.95

TABLE 3: The performance of AMD diagnosis from different classifiers

Classifiers	Acc (%)	Sen (%)	Spe (%)	AUC
SVM (Linear function)	92.5	97.0	67.0	0.96
SVM (Gaussian function)	92.0	98.0	59.0	0.96
SVM (Quadratic function)	93.8	97.0	74.0	0.97
K-NN	83.9	86.0	73.0	0.80
Random Forest	88.4	96.0	45.0	0.92

TABLE 4: A summary of automated identification of normal and AMD diseases

Author	Year	Number of Images	Methods	Performance
Zheng et al.[25]	2012	N:98, AMD:160	Weighted frequent sub-graph mining	Ten-fold, Acc:99.6%, Sen:99.4%, Spe:100%
Hijazi et al. [26]	2012	N:60, AMD:101	Hierarchical decomposition (1262 features)	Ten-fold, Acc:100%, Sen:100%, Spe:100%
Hijazi et al.[27]	2015	N:98, AMD:165	Tree-based approach	Ten-fold, Acc:99.9%, Sen:100.00%, Spe:99.00%
Acharya et al.[22]	2016	N:400, Abnormal:400	BEMD, SVM classifier	Ten-fold, Acc:88.63%, Sen:86.25%, Spe:91.00%
Acharya et al.[20]	2017	N: 404, Dry AMD: 517, PHOG, nonlinear features, Wet AMD: 24	PHOG, nonlinear features, SVM classifier	Ten-fold, Acc:83.30%, Sen:82.60%, Spe:84.80%
Tan et al.[23]	2018	N:402, AMD: 708	14-layer CNN	Ten-fold, Acc:95.45%, Sen:96.43%, Spe:93.75%
Our work	2020	Both Color and FFA images N:99 (pair), AMD:539 (pair)	ResNet-50, off-the-shelf	Ten-fold, Acc: 93.8%, Sen: 97%, Spe:74%, AUC:0.97

3. Discussion

To the best of our knowledge, this is the first study to combine color fundus and FFA images for AMD CAD with deep learning methods. Previous study has shown that the possibility of first combining with fundus color photography and OCT can improve the diagnostic accuracy of AMD diagnosis compared to the data alone [28]. In fact, as the gold standard modality for diagnosing AMD, FFA is effective for detecting the presence of choroidal neovascularization (CNV). In FFA screening, a contrast agent is given by intravenous injection in order to collect photos of vascular visualization, and then the existence and range of CNV can be determined by observing the presence or absence of contrast agent in the morphological memory of blood vessels [29, 30]. FFA is effective in determining the classification of the CNV (classic or occult), boundaries, composition, location of the neovascular complex, and can guide laser or anti-VEGF therapy [31]. However, our experiment results show that when applying proposed two modality CAD algorithms, single modality (i.e., only use FFA itself in diagnosis) performs worse than the one combining both FFA and color fundus images. Normally, patients who take FFA screening should have taken color fundus imaging already, so based on this check process, our pipeline does not cause extra cost.

In Tan's work [23], they did not give the training time of their 14 layer CNN models and did not compare with other state-of-the art CNN models. In our AMD CAD pipeline, ResNet-50 with Quadratic SVM have achieved the best performance compared with other five Convolutional Neural Networks (CNN) models and other four classifiers. Based on the off-the-shelf training strategy, our method have less training time than training from scratch strategy, ResNet-50 as a feature extractor, in our two modality images, the time of feature extraction is less than 11 minutes, and the training time of SVM with Quadratic is less than 10 seconds. Furthermore, it needs more data samples and more complicated CNN models to train dual modality images with training from scratch. Our AMD CAD pipeline have the good advantage of the time complexity with less training time and it is easier to be implemented into the clinical application.

Nevertheless, this research is also limited by some shortcomings. Especially, in terms of the paired color fundus and FFA images of healthy people, this dataset is seriously unbalanced. Only 98 healthy color fundus images with paired FFA images can be used in our study. Since FFA is an invasive test that requires intravenous infusion of sodium fluorescein injection as a contrast agent, applying this test on

healthy people would cause meaningless safety risks and financial burden. Therefore, healthy people are commonly not encouraged to take FFA test, which results in the lack of healthy FFA images. The imbalance of dataset indeed affects the result: the highest specificity value is 74%, whereas the highest sensitivity value is 97%. Even so, the performance of combined modalities is still better than either single modality: the best specificity values of only using color fundus and FFA image are 42% and 69%, respectively. That is to say, the dataset imbalance does not affect the superiority of combined modalities. Furthermore, the image quality of fundus and FFA images is a very important factor, i.e., low image quality images with large noises and artifacts will affect the performance of feature extraction by CNN models. The extracted features from more complicated models (such as ResNet101, DenseNet-201 and Inception-ResnetV2) are not well matched for our data samples in our pipeline, but it cannot completely represent that the ResNet-50 always has better performance than the other models. If the paired training samples are larger, these state-of-the art CNN models are suggested to be examined and chosen to be the best CNN models. Since SVM has been well applicated in CAD systems, three kinds of SVM have been compared in this study. Among them, the Quadratic kernel function has better performance than Linear function and Gaussian (Radial Basis) function.

In addition, the lack of proper labels of dry and wet AMD images is also a limitation of our research. The identification of AMD types (i.e., dry-AMD and wet-AMD) is also an important research problem. In some cases, distinguishing between dry AMD and wet AMD is difficult even for a well-trained ophthalmologist. Our research is a binary classification which focusing on the identification of AMD. After the identification of AMD, distinguishing the dry and wet AMD can be treated as another binary classification problem. If we could have achieved a dataset with dry and wet AMD labeled, we would implement our pipeline on it for further classification of dry and wet AMD.

4. Conclusion

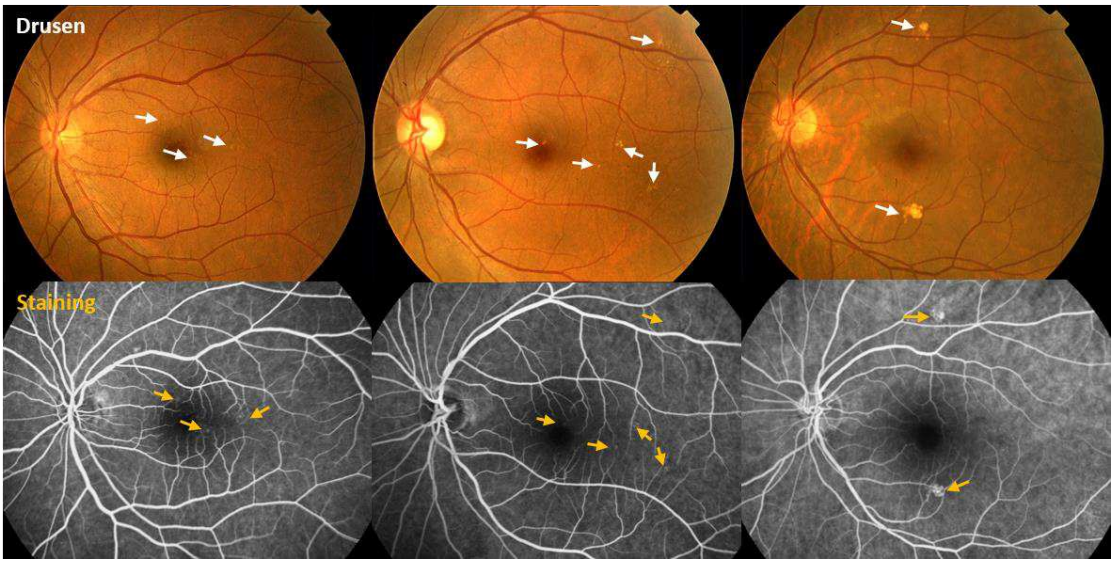
In this study, we proposed a new AMD computer-aided diagnosis system with both color fundus and FFA images. This system merges the feature vectors extracted from those two modalities using deep neural networks, and then accomplishes diagnosis with SVM classifier. The evaluation results show that the advantage of this dual-modality CAD system which outperforms both single fundus and single FFA

CAD system. The proposed method achieved 93.8% accuracy and 0.97 AUC value, which can help ophthalmologist reduce the workload and improve the accuracy of diagnosis. The results also indicate that ResNet-50 with Quadratic SVM are suggested to be used as feature extraction and classifier in this CAD pipeline which achieved the best performance in this study, which has feasibility of clinical application. In the future work, along with color fundus and fluorescein angiography imaging, optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA) are also suggested to extract multi-modality features. We could incorporate these two modalities into our pipeline in order to further improve the feature representation and classification result.

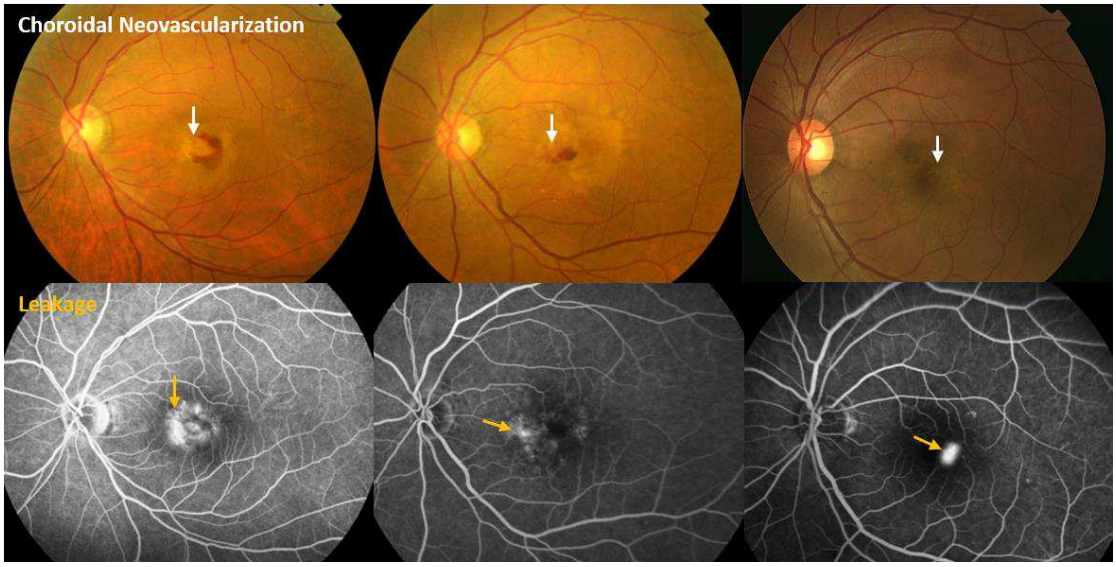
5. Method

5.1. Datasets

We collected the data of patients who had undergone colored fundus photographs and fluorescein angiography in Peking University First Hospital Department of Ophthalmology between 2008-2017. Fundus photographs were captured using a TOPCON TRC-50DX fundus camera (Topcon Corporation, IMAGE net 2000TM, Tokyo, Japan). The Heidelberg SPECTRALIS HRA was used to obtain the conventional fluorescein angiograms. Angiography was performed using an intravenous injection of 3 ml of 20% sodium fluorescein after mydriasis. High-speed mode FFA images are used in our experiment. The resolution of color fundus image is 576×768, and that of FFA image is 768×768. 664 paired color fundus and FFA images were used in this study, which include 98 normal eyes and 566 AMD eyes. Some typical paired color images and FFA images along with drusens and leakages in our study are shown in figure 1. As shown in this figure, drusens can be observed as yellow sediments of different sizes, shapes and distributions under the retina, and are usually stained with fluorescein in FFA. FFA can reveal drusens that bind fluorescein and hyperfluorescence in the middle and late stage of angiography. Meanwhile, CNV presents as a hyperfluorescence lesion (namely, leakage) that enlarges in size and intensity in late phases as the fluorescein leaks from the neovascular membranes.



(a) Drusens and Stainings



(b) Choroidal Neovascularizations and leakages

Figure 2 Examples of collected fundus images and FFA images with **(a)** drusens on color fundus images and the staining of them on FFA images and **(b)** suspected Choroidal Neovascularizations on color fundus images and their corresponding leakages on FFA images.

5.2. Pre-processing

AMD screening is mainly based on the lesions in the macular area. In addition, the high brightness of the optic disc can affect the feature extraction of macular area. Therefore, we crop the color fundus and FFA images and only keep the macular area, so that we could maintain and focus on the information that is closely related to AMD. We calculate the centroid of optic disk by implementing region growing

method using optic disk centroid and papilla diameter. The centroid of macular can be located indirectly. Square area of macular with radius of 2.5 papilla diameter is utilized as input to extract universal features. According to the corresponding coordinate positions, paired macular area of FFA images can also be cropped. Since the distance between the center of the macular area and the center of the optic disc is about 2.5 times of papilla diameter, the macular areas of the color fundus and FFA images were located and cropped by batch processing according to the corresponding coordinate positions. Since some macular areas were difficult to locate (for example, some images were collected from patients with serious diseases, whose macular areas were nearly all covered by lesions), manual cropping was needed to extract the exact macular area patches. The resolution of macular area patches was from 300×300 to 400×400, and all these macular images were subsequently normalized 224×224 and 299×299 in order to load into different Convolutional Neural Network (CNN).

5.3. Methods

Convolutional Neural Network (CNN) is a powerful machine learning technique from the field of deep learning. CNNs are trained using large collections of diverse images. From these large collections, CNNs can learn rich feature representations for a wide range of images. These feature representations often outperform hand-crafted features such as HOG, LBP, or SURF [32]. An easy way to leverage the power of CNNs. In order to fuse two modality's features, pre-trained CNNs with off the shelf strategy were used to extract feature. ResNet-18, ResNet-50, ResNet-101 [33] Inception-V3 [34], Inception-ResnetV2 [35], and DenseNet-201 [36] CNN models which achieved good performance in the ImageNet Competitions were implemented in this study. The pipeline of our study is shown in figure 3. We also used several fully connected network (FCN) layers of these pre-trained CNN models as feature extractors. The extracted features are therefore used to train different classifiers. The names of the six examined pretrained CNN models and their corresponding FCN layers are listed in Table 1.

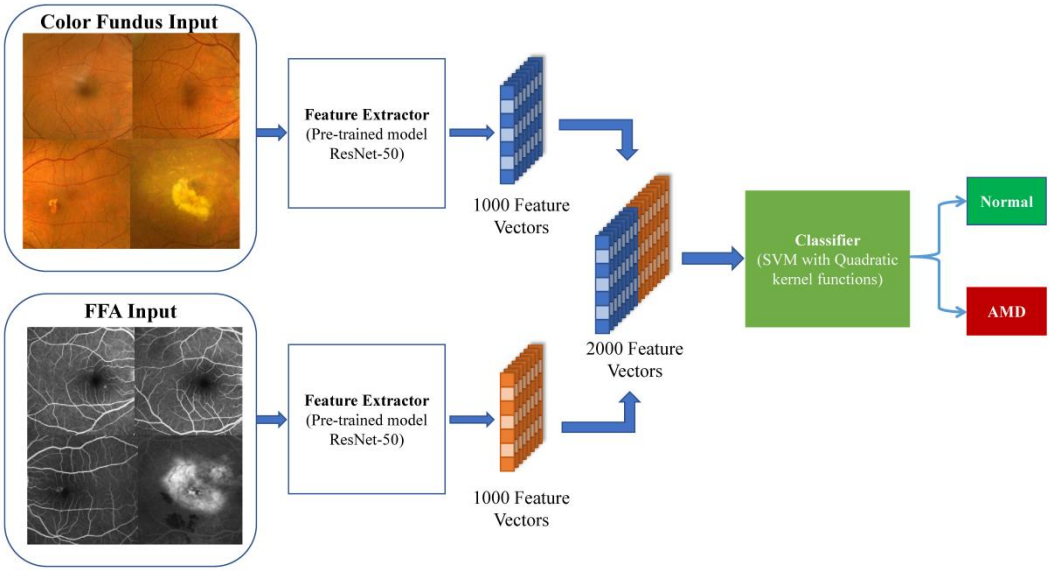


Figure 3 The pipeline of the Dual Modality AMD computer-aided diagnosis

Table 1 CNN models and layers

Models	Layer Names	Feature Vectors
ResNet-18	fc1000	1000
ResNet-50	fc1000	1000
ResNet-101	fc1000	1000
Inception-V3	predictions	1000
Inception-ResNetV2	predictions	1000
DenseNet-201	fc1000	1000

5.3.1 Deep Feature Extraction

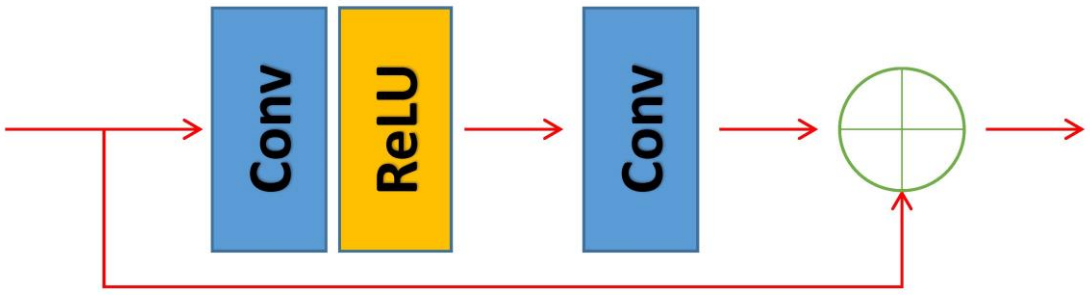
The pre-trained models trained on the ImageNet weights were employed in our study. Parameters in the convolutional and fully connected layers are fixed and are used as image extractors. During the feature extraction process, only the weights of the convolutional part were retained for extracting deep features from the last fully connected layer of each pre-trained CNN model. All the experiments were implemented on MATLAB 2019b, Intel (R) Core (TM) i5-3300 CPU (3.0GHZ).

The extracted features by deep learning neural networks would be sent to the classifiers to categorize as "normal" and "AMD". The feature vector of color fundus images is $F^c = [F_c^1, F_c^2, \dots, F_c^n]$,

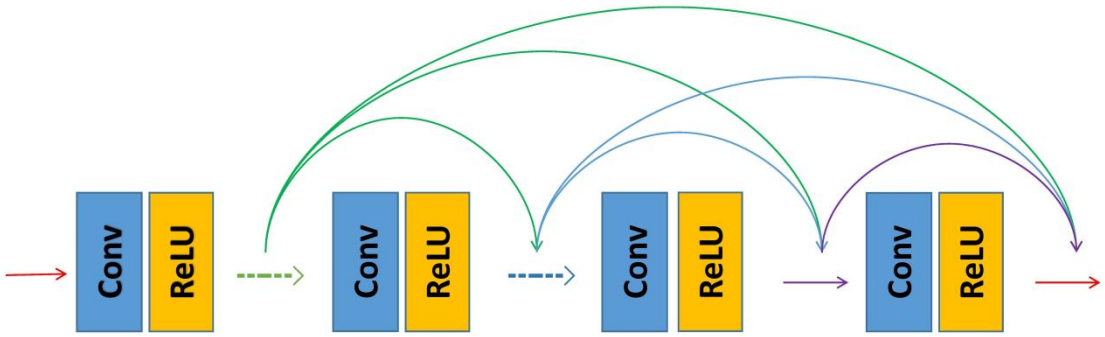
and the feature vector of FFA images is $F^f = [F_f^1, F_f^2, \ldots, F_f^n]$. Thus, the merged dual-modality feature vector is $F^d = [F_c^n, F_f^n]$. The corresponding label(or classification) $Y = [Y_1, Y_2, \cdots, Y_n] \in [0,1], n \in R$, where 0 and 1 mean that the fundus and FFA images was taken from normal people (without AMD) and AMD patients, respectively. Here we briefly describe the pre-trained CNN models that were examined in our research.

1) ResNet

ResNet is a robust deep learning network series [33]. It uses residual framework to decrease the difficulty of the training of "very deep" networks. Let a single image x_0 go through a L layers convolutional network. Each layer of the network corresponds to a non-linear transformation $H_l(\cdot)$, where l represents the index of the layer. Let x_1 be the output of the lth layer. In traditional convolutional networks, the input of the $l+1$th layer is the output of the lth layer, which can be expressed as:$x_{l+1} = H_{l+1}(x_l)$. On the other hand, unlike traditional CNNs, ResNet implements a residual block (shown in figure 4) that sums up the identity mapping of the input to the output of a layer, where the output can be depicted as: $x_{l+1} = H_{l+1}(x_l) + x_l$. This operation eases the convergence during training. The structure of residual block is shown in figure 4(a). We chose 3 typical depth of ResNet, i.e., ResNet-18, ResNet-50, ResNet-101, as the models to be examined and compared in our research.



(a) Residual block



(b) Dense block

Figure 4 (a) Residual and **(b)** dense block structure.

2) Inception-V3

Since the success in ILSVRC2014, GoogLeNet has been widely applied in image classification tasks. Nevertheless, some disadvantages of the network, such as the complexity of its "Inception" structure, limited the progress and development of GoogLeNet. Thus, the researchers, Szegedy et al., reconsidered the Inception module in GoogLeNet and presented some general principles so that the Inception network can be modified and optimized [34]. According to these principles, they proposed a new network architecture named as Inception-V3. Inception-V3 has a depth of 48 layers, consisting of six convolutional layers, two pooling layers, three Inception structures, one linear layer, and one softmax classifier. The total parameters of this network is 23.9 million. In the off-the-shelf pre-trained Inception-V3 model, the image input size is 299×299.

3) Inception-ResnetV2

Inception-ResnetV2 is a combined model of Inception-V3 and residual network [35]. This model was proved to have similar performance of the next generation Inception model Inception-V4 in terms of top-1 error and top-5 error in certain evaluations. The depth of Inception-ResnetV2 is 164, and the total number of parameters in the network is about 55.9 million. The pre-trained Inception-ResnetV2 model constrains the input size as 299×299.

4) DenseNet-201

Unlike the models above, who focused more in extending the depth and width of networks, DenseNet "densely" connected each block to achieve the re-usage of features (shown in figure 4(b)), which made the model relatively easy to fit and convergence [36]. In addition, bottleneck was introduced to further reduce the computing burden. More specifically, The feature maps of all previous layers are treated as separate inputs by connecting them to a single tensor $[x_0, x_1, \dots, x_l]$, while their own feature maps are passed as input to all subsequent layers. Layer l + 1 receives the feature maps of all previous layers and can be expressed as: $x_{l+1} = H_{l+1}([x_0, x_1, \dots, x_l])$. The total number of parameters in this network is 20.0 million, and the input size of image is 224×224.

5.3.2 Classifiers

For the purpose of quantitatively assessing the performance of classifiers, we tested and compared Support Vector Machine (SVM) with different kernel functions, K-Nearest Neighbor (K-NN), and Random Forest. We also applied tenfold cross-validation to better reflect data distribution. In the SVM methods, we examined three kernel functions (linear function, Gaussian function, and Quadratic function).

1) Random Forest

Random forest (RF), also known as random decision forest, is a learning based method that have been widely applied in multiple tasks such as classifications and regressions. This algorithm was firstly proposed by Kam in 1995 [37], and subsequently developed by Breiman in 2001 [38]. The random forest, basically, uses multiple non-correlated decision trees to optimize the accuracy. While the decision tree is a widely welcomed learning-based method, the accuracy of it is limited by the overfitting issue. A sampling and averaging algorithm called bootstrap aggregating (or bagging) is introduced to reduce the variance and further resolve the overfitting. In our implementation of the random forest, the number of trees (ntree) was set to 200.

2) K-Nearest Neighbor

The K-NN approach is a famous non-parametric machine learning method. The theory is that in the dataset, it finds a group of k samples that are nearest to unknown samples. For this classifier in our study, the parameter k was set as 2 due to that we were performing a binary classification task.

3) Support Vector Machine

SVM is a classifier which finds the optimal separating hyperplane (OSH) to categorize data into, normally, two classifications. The prototype of SVM was originally proposed in 1963, and was continuously improved until 1995 when it became current well-known pattern [39, 40]. The gap between the OSH and the nearest point in the data is called "support vectors". Assume that each point in the dataset can be defined as$x_i \in R^n, i = 1,2, \dots, N$, and belongs to a class $y_i \in \{-1,1\}$. For linear classification,

$$w \cdot x_i + b \geq 1, yi = 1$$

$$w \cdot x_i + b \leq 1, yi = -1 \quad (1)$$

can classify into two categories. This can be expanded as

$$y_i \cdot [(w \cdot x_i)] \geq 1, i = 1,2, \dots, l \quad (2)$$

Thus, the distance between support vectors is defined as

$$d = \frac{2}{\|w\|} \quad (3)$$

When $\|w\|$ is minimized, distance d gets maximized, and the separating hyperplane is optimized, i.e., the OSH. This problem can be represented as a constrained quadratic programming convex optimization problem. Using the Lagrangian function:

$$L(w, b, a) = \frac{\|w\|^2}{2} - \sum_{i=1}^{1} a_i \cdot \{y_i \cdot [(w \cdot x_i) + b - 1]\} \quad (4)$$

convert into a dual form of quadratic programming convex optimization:

$$max \sum_{i=1}^{l} a_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} a_i \cdot a_j \cdot y_i \cdot y_j (x_i^T x_j)$$

$$s.t., a_i \geq 0, i = 1,2, \dots, l \quad (5)$$

$$\sum_{i=1}^{l} a_i \cdot y_i = 0$$

This problem has a single optimized solution. Let a_i be the optimized solution, then

$$w^* = \Sigma_{i=1}^{l} a_i^* y_i x_i \tag{6}$$

The support vector consists of all the a_i that is not 0. On the other hand, b can be solved with the constraints. Thus, the optimized classifying function is:

$$f(x) = sgn((w^*)^T x + b^*) \tag{7}$$

SVM uses non-linear transforms to map the input space into a high-dimension space. In the mapped new space, solving the OSH could be easier. The non-linear transform is performed by kernel functions (inner products). Let $K(x_i, x_j) =< \Phi(x_i)\Phi(x_j) >$replace $x_i^T x_j$ According to the Hilbert-Schmidt's theorem, if a operation suits Mercer's theorem, it can be used as inner product. The common kernel functions includes:

a: Linear kernel function

$$K(x_i, x_j) = x_i^T x_j \tag{8}$$

b: Polynomial kernel function

$$K(x_i, x_j) = (x_i^T x_j + 1)^q \tag{9}$$

c: Gaussian kernel function

$$\mathrm{K}(x_i, x_j) = exp\left\{-\frac{\|x_i - x_j\|^2}{\sigma^2}\right\} \tag{10}$$

d: Quadratic kernel function

$$K(x_i, x_j) = (1 + x_i^T x_j)^2 \tag{11}$$

In our implementation, parameter of penalty coefficient (C) and kernel width parameter (γ) are optimized by grid search method.

Declarations

Acknowledgements

Authors' contributions

Zekuan Yu and Zifeng Tian designed and conducted the experiments, as well as wrote the draft. Jianchen Hao, Bin Qiu, Wenbo Zhang, Hongping Nie, Shijie Zhang, and Jun Li collected and labeled the medical images, also checked and validated the experiment data. Shujin Zhu, Liu Yang, Qiushi Ren, and Yanye Lu reviewed and edited the manuscript.

Funding

This work was funded by the National Natural Science Foundation of China (61875123, 81421004); the National Key Instrumentation Development Project of China (2013YQ030651); the Natural Science Foundation of Hebei Province (H2019201378).

Availability of data and materials

Since the data used in this study include data collected in a clinical trial with patients, the data will not be shared.

Ethics approval and consent to participate

All participants gave their informed consent to participate in the study.

Consent for publication

All participants signed an informed consent to the use of all coded data collected during the study in scientific publications.

Competing interests

The authors declare no conflict of interest.

Authors' information

[1]Department of Biomedical Engineering, College of Engineering, Peking University, Beijing, 100871 China

[2]Department of Ophthalmology, Peking University First Hospital, Beijing, 100034 China

[3]Institute of Biomedical Engineering, Peking University Shenzhen Graduate School, Shenzhen, 518055 China

[4]Institute of Biomedical Engineering, Shenzhen Bay Laboratory 5F, Shenzhen, 518071 China

[5]Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen 91054 Germany

[6]School of Geography and Biological Information, Nanjing University of Posts and Telecommunications, Nanjing, 210023 China

*These authors contributed equally to this study and share the first authorship.

§Corresponding author: Yanye Lu (e-mail: yanye.lu@pku.edu.cn), Liu Yang (e-mail: liu_yang@bjmu.edu.cn)

Footnotes

None

Reference

1. Dey A: **World report on ageing and health.** *The Indian Journal of Medical Research* 2017, **145**(1):150.
2. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, Yusuf S: **The burden of disease in older people and implications for health policy and practice.** *Lancet* 2015, **385**(9967):549-562.
3. Friedman DS, O'Colmain BJ, Muñoz B, Tomany SC, McCarty C, de Jong PT, Nemesure B, Mitchell P, Kempen J: **Prevalence of age-related macular degeneration in the United States.** *Arch Ophthalmol* 2004, **122**(4):564-572.
4. Koh JE, Ng EY, Bhandary SV, Laude A, Acharya UR: **Automated detection of retinal health using PHOG and SURF features extracted from fundus images.** *Applied Intelligence* 2018, **48**(5):1379–1393.
5. Koh JEW, Ng EYK, Bhandary SV, Hagiwara Y, Laude A, Acharya UR: **Automated retinal health diagnosis using pyramid histogram of visual words and Fisher vector techniques.** *Comput Biol Med* 2018, **92**:204-209.
6. Chou R, Dana T, Bougatsos C, Grusing S, Blazina I: **Screening for Impaired Visual Acuity in Older**

Adults: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *Jama* 2016, **315**(9):915-933.

7. Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY: **Age-related macular degeneration.** *Lancet* 2012, **379**(9827):1728-1738.
8. Barondes M, Pauleikhoff D, Chisholm IC, Minassian D, Bird AC: **Bilaterality of drusen.** *Br J Ophthalmol* 1990, **74**(3):180-182.
9. Gragoudas ES, Adamis AP, Cunningham ET, Jr., Feinsod M, Guyer DR: **Pegaptanib for neovascular age-related macular degeneration.** *N Engl J Med* 2004, **351**(27):2805-2816.
10. Brown DM, Kaiser PK, Michels M, Soubrane G, Heier JS, Kim RY, Sy JP, Schneider S: **Ranibizumab versus verteporfin for neovascular age-related macular degeneration.** *N Engl J Med* 2006, **355**(14):1432-1444.
11. Rosenfeld PJ, Brown DM, Heier JS, Boyer DS, Kaiser PK, Chung CY, Kim RY: **Ranibizumab for neovascular age-related macular degeneration.** *N Engl J Med* 2006, **355**(14):1419-1431.
12. Ryan SJ, Sadda SR, Hinton D, Schachat AP, Wilkinson CP, Wiedemann P: *Retina Fifth Edition.* 2012.
13. Klein R, Davis M, Magli YL, Segal P, Klein BEK, Hubbard L: **The Wisconsin Age-Related Maculopathy Grading System.** *Ophthalmology* 1991, **98**:1128-1134.
14. Koh JEW, Acharya UR, Hagiwara Y, U R, Tan JH, Subbhuraam VS, Bhandary S, Rao A, Sivaprasad S, Chua K, et al: **Diagnosis of Retinal Health in Digital Fundus Images Using Continuous Wavelet Transform (CWT) and Entropies.** *Computers in Biology and Medicine* 2017, **84**.
15. Pead E, Megaw R, Cameron J, Fleming A, Dhillon B, Trucco E, MacGillivray T: **Automated detection of age-related macular degeneration in color fundus photography: a systematic review.** *Surv Ophthalmol* 2019, **64**(4):498-511.
16. Mookiah MR, Acharya UR, Fujita H, Koh JE, Tan JH, Noronha K, Bhandary SV, Chua CK, Lim CM, Laude A, Tong L: **Local configuration pattern features for age-related macular degeneration characterization and classification.** *Comput Biol Med* 2015, **63**:208-218.
17. Mookiah MR, Acharya UR, Koh JE, Chua CK, Tan JH, Chandran V, Lim CM, Noronha K, Laude A, Tong L: **Decision support system for age-related macular degeneration using discrete wavelet transform.** *Med Biol Eng Comput* 2014, **52**(9):781-796.
18. Mookiah MRK, Acharya UR, Fujita H, Koh JEW, Tan JH, Chua CK, Bhandary SV, Noronha K, Laude A, Tong L: **Automated detection of age-related macular degeneration using empirical mode decomposition.** *Knowledge-Based Systems* 2015, **89**:654-668.
19. Mookiah MRK, Acharya UR, Koh JEW, Chandran V, Chua CK, Tan JH, Lim CM, Ng EYK, Noronha K, Tong L, Laude A: **Automated diagnosis of Age-related Macular Degeneration using greyscale features from digital fundus images.** *Computers in Biology and Medicine* 2014, **53**:55-64.
20. Acharya UR, Hagiwara Y, Koh JEW, Tan JH, Bhandary SV, Rao AK, Raghavendra U: **Automated screening tool for dry and wet age-related macular degeneration (ARMD) using pyramid of histogram of oriented gradients (PHOG) and nonlinear features.** *Journal of Computational Science* 2017, **20**(May):41-51.
21. Acharya UR, Mookiah MR, Koh JE, Tan JH, Noronha K, Bhandary SV, Rao AK, Hagiwara Y, Chua CK, Laude A: **Novel risk index for the identification of age-related macular degeneration using radon transform and DWT features.** *Comput Biol Med* 2016, **73**:131-140.
22. Acharya UR, Mookiah MR, Koh JE, Tan JH, Bhandary SV, Rao AK, Fujita H, Hagiwara Y, Chua CK,

Laude A: **Automated screening system for retinal health using bi-dimensional empirical mode decomposition and integrated index.** *Comput Biol Med* 2016, **75**:54-62.

23. Tan JH, Bhandary S, Sivaprasad S, Hagiwara Y, Bagchi A, U R, Rao A, Raju B, Shetty N, Gertych A, et al: **Age-related Macular Degeneration detection using deep convolutional neural network.** *Future Generation Computer Systems* 2018, **87**.
24. Serener A, Serte S: **Dry and Wet Age-Related Macular Degeneration Classification Using OCT Images and Deep Learning.** In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*; *24-26 April 2019*. 2019:1-4.
25. Zheng Y, Hijazi MHA, Coenen F: **Automated "Disease/No Disease" Grading of Age-Related Macular Degeneration by an Image Mining Approach.** *Invest Ophthalmol Vis Sci* 2012, **53**(13):8310.
26. Hijazi MHA, Coenen F, Zheng Y: **Data mining techniques for the screening of age-related macular degeneration.** *Knowledge-Based Systems* 2012, **29**:p.83-92.
27. Hijazi MHA, Coenen F, Zheng Y: **Data mining for AMD screening: A classification based approach.** *International Journal of Simulation: Systems, Science and Technology* 2015, **15**(2):57-69.
28. Yoo TK, Choi JY, Seo J, Bhoopalan R, Selvaperumal S, Kim DW: **The possibility of the combination of OCT and fundus images for improving the diagnostic accuracy of deep learning for age-related macular degeneration: a preliminary experiment.** *Medical & Biological Engineering & Computing* 2018, **57**.
29. Kanagasingam Y, Bhuiyan A, Abramoff M, Smith R, Goldschmidt L, Wong TY: **Progress on Retinal Image Analysis for Age Related Macular Degeneration.** *Progress in retinal and eye research* 2013, **38**.
30. Al-Zamil W, Yassin S: **Recent developments in age-related macular degeneration: A review.** *Clinical Interventions in Aging* 2017, **Volume 12**:1313-1330.
31. Bressler NM, Bressler SB: **Chapter 66 - Neovascular (Exudative or "Wet") Age-Related Macular Degeneration.** In *Retina (Fifth Edition).* Edited by Ryan SJ, Sadda SR, Hinton DR, Schachat AP, Sadda SR, Wilkinson CP, Wiedemann P, Schachat AP. London: W.B. Saunders; 2013: 1183-1212.
32. Islam KT, Raj R, Al-Murad A: *Performance of SVM, CNN, and ANN with BoW, HOG, and Image Pixels in Face Recognition.* 2017.
33. He K, Zhang X, Ren S, Sun J: *Deep Residual Learning for Image Recognition.* 2016.
34. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna ZB: *Rethinking the Inception Architecture for Computer Vision.* 2016.
35. Szegedy C, Ioffe S, Vanhoucke V, Alemi A: **Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.** *AAAI Conference on Artificial Intelligence* 2016.
36. Huang G, Liu Z, van der Maaten L, Weinberger K: *Densely Connected Convolutional Networks.* 2017.
37. Ho TK: **Random decision forests.** In *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1.* IEEE Computer Society; 1995: 278.
38. Breiman L: **Random Forests.** *Machine Learning* 2001, **45**(1):5-32.
39. Cortes C, Vapnik V: **Support-vector networks.** *Machine Learning* 1995, **20**(3):273-297.
40. Boser BE, Guyon IM, Vapnik VN: **A training algorithm for optimal margin classifiers.** In *Proceedings of the fifth annual workshop on Computational learning theory.* Pittsburgh, Pennsylvania, USA: Association for Computing Machinery; 1992: 144–152.

Figures

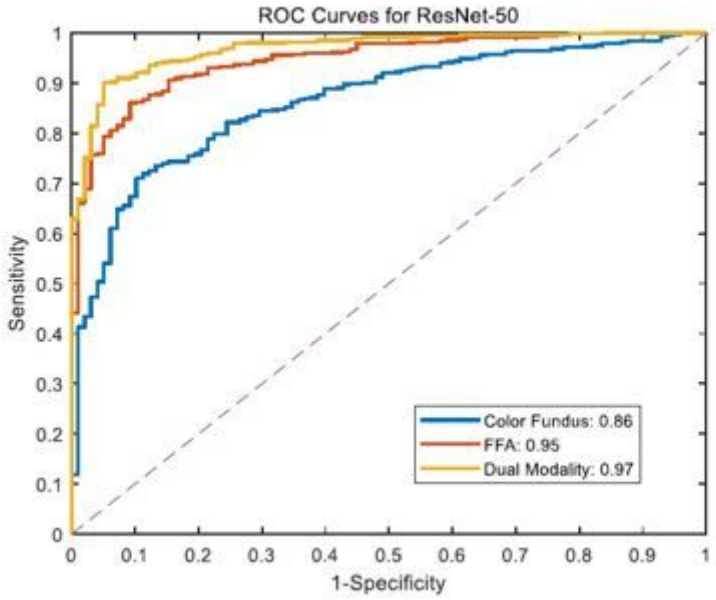
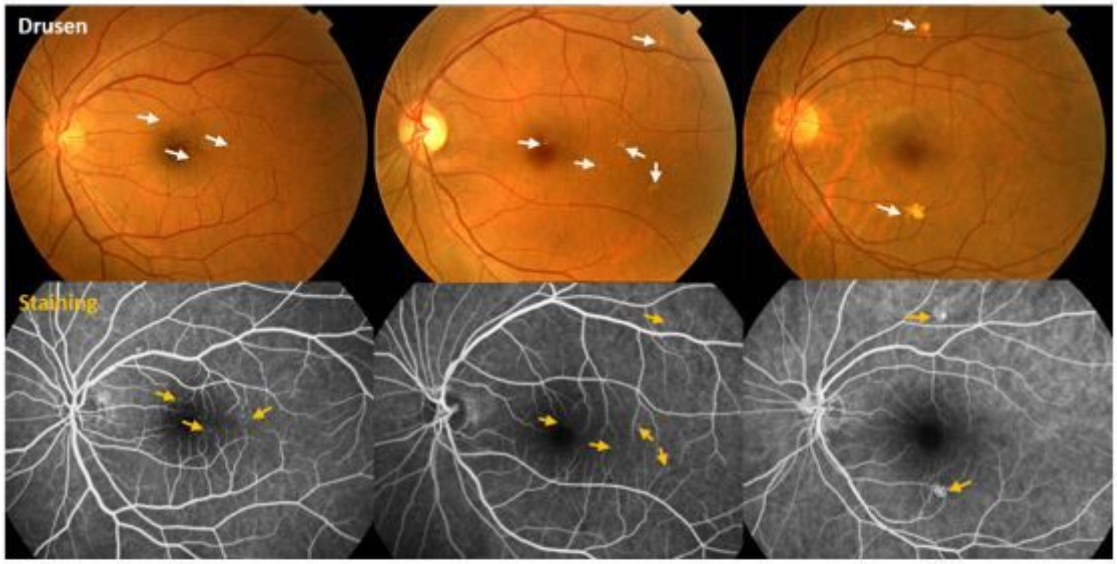
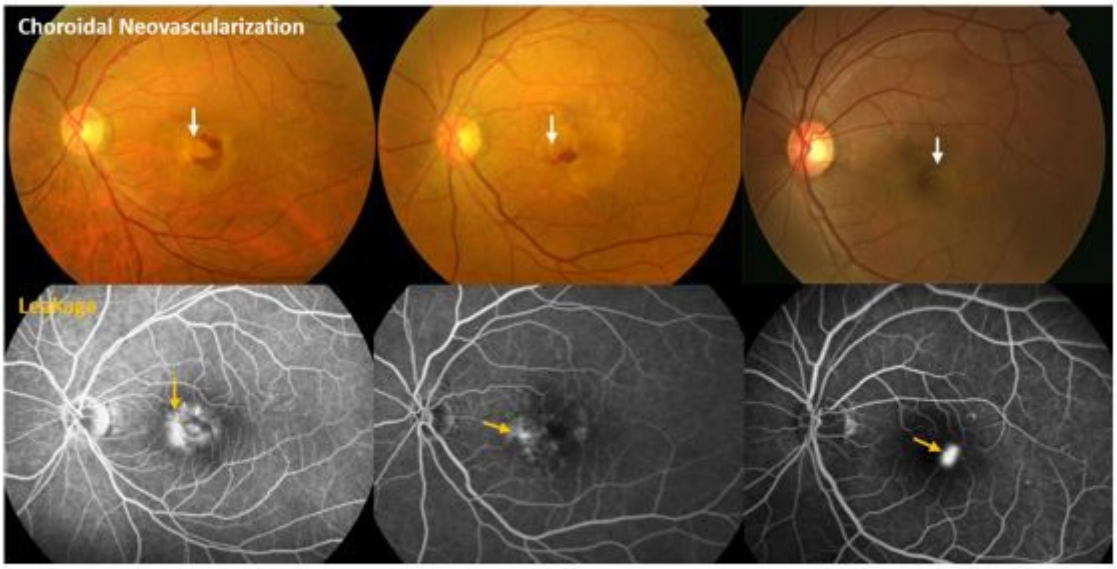


Figure 1

ROC curve of ResNet-50.



(a) Drusens and Stainings



(b) Choroidal Neovascularizations and leakages

Figure 2

Examples of collected fundus images and FFA images with (a) drusens on color fundus images and the staining of them on FFA images and (b) suspected Choroidal Neovascularizations on color fundus images and their corresponding leakages on FFA images.

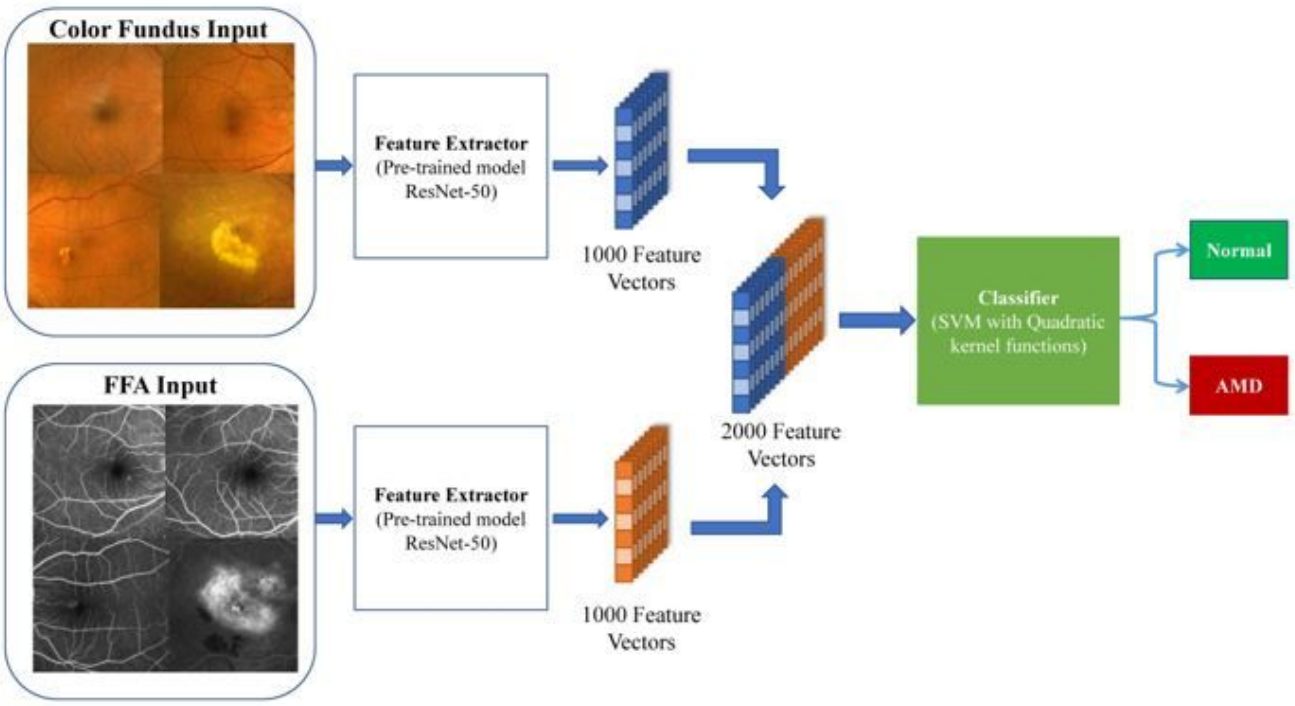


Figure 3

The pipeline of the Dual Modality AMD computer-aided diagnosis

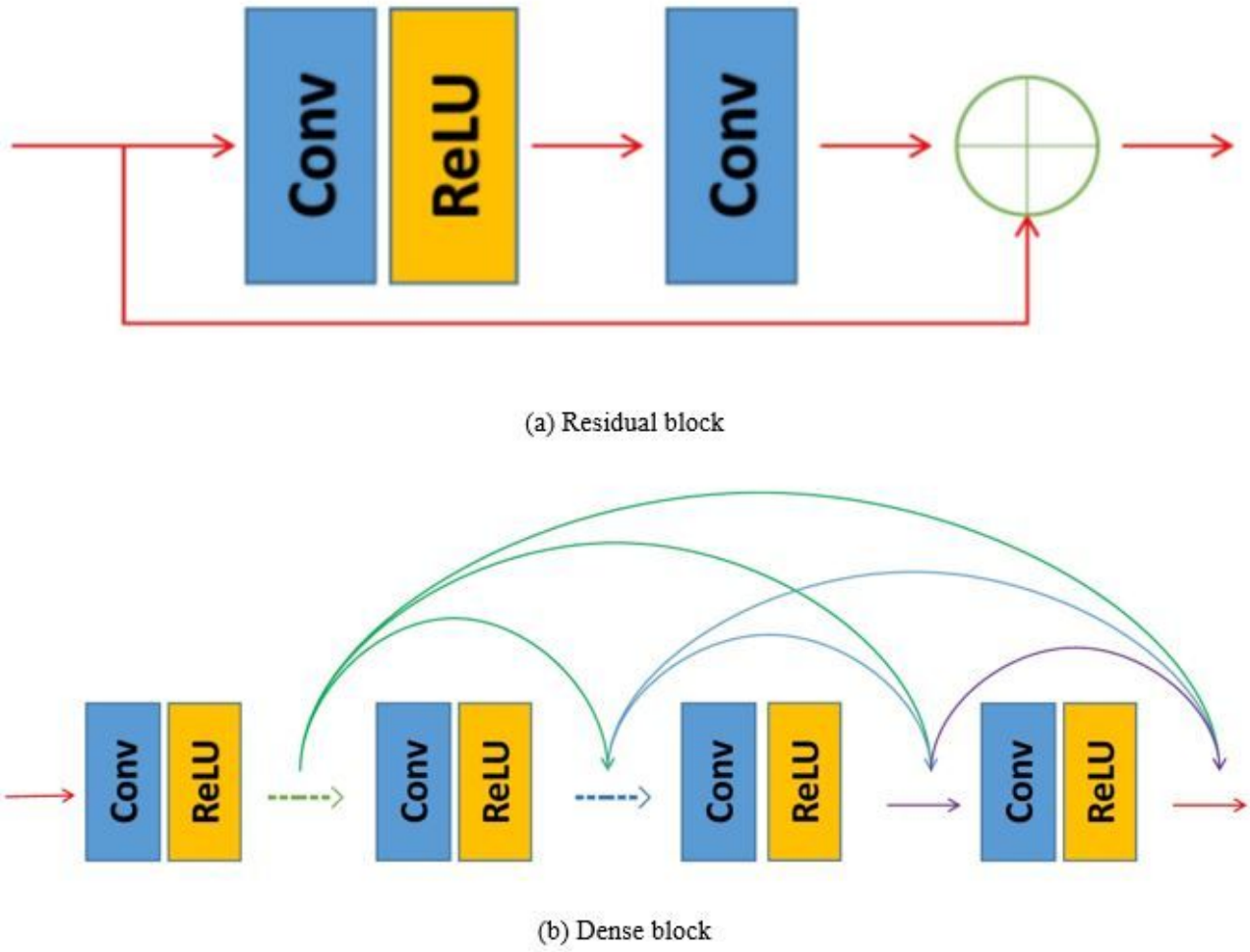


Figure 4

(a) Residual and (b) dense block structure.