

## Article

# Efficient Deep Learning-Based Automated Pathology Identification in Retinal Optical Coherence Tomography Images

Qingge Ji <sup>1,2</sup>, Wenjie He <sup>1,2</sup>, Jie Huang <sup>1,2</sup> and Yankui Sun <sup>2,3,\*</sup>

<sup>1</sup> School of Data and Computer Science, Sun Yat-sen University, 132 East Waihuan Road, Guangzhou Higher Education Mega Center, Guangzhou 510006, China; issjqg@mail.sysu.edu.cn (Q.J.); 13926101907@163.com (W.H.); huangj229@mail2.sysu.edu.cn (J.H.)

<sup>2</sup> Guangdong Key Laboratory of Big Data Analysis and Processing, Guangzhou 510006, China

<sup>3</sup> Department of Computer Science and Technology, Tsinghua University, 30 Shuangqing Road, Haidian District, Beijing 100084, China

\* Correspondence: syk@mail.tsinghua.edu.cn; Tel.: +86-10-6278-2609

Received: 4 May 2018; Accepted: 17 June 2018; Published: 20 June 2018



**Abstract:** We present an automatic method based on transfer learning for the identification of dry age-related macular degeneration (AMD) and diabetic macular edema (DME) from retinal optical coherence tomography (OCT) images. The algorithm aims to improve the classification performance of retinal OCT images and shorten the training time. Firstly, we remove the last several layers from the pre-trained Inception V3 model and regard the remaining part as a fixed feature extractor. Then, the features are used as input of a convolutional neural network (CNN) designed to learn the feature space shifts. The experimental results on two different retinal OCT images datasets demonstrate the effectiveness of the proposed method.

**Keywords:** optical coherence tomography; image analysis; age-related macular degeneration; diabetic macular edema; convolutional neural network

## 1. Introduction

The macula, which is located in the central part of the retina, is the most sensitive area of vision. Its healthiness can be affected by a number of pathologies such as AMD and DME. As technology advances, OCT has been extensively used to capture the textural and morphological variations in the retina [1–3]. The ability to provide cross-sectional images of tissue makes it possible to diagnose AMD and DME [4–6].

A majority of the previous works on the retina focused on the methods of retinal layer segmentation [7–22]. During the past decades, a multitude of classification algorithms were proposed to facilitate the automatic detection of retinal disorders (especially AMD and DME). Liu et al. [23] proposed a methodology for computing the Local Binary Pattern (LBP) to encode the texture and shape information in retinal OCT images, and represented images using a multiple-scale spatial pyramid followed by a principal component analysis for dimension reduction. Sugruk et al. [24] segmented the OCT images to find the retinal pigment epithelium (RPE) layer and used a binary classification to classify between AMD and DME. Srinivasan et al. [25] utilized multiple-scale histograms of oriented gradient (HOG) descriptors as feature vectors of a support vector machine (SVM)-based classifier. Hassan et al. [26] extracted five distinct features (three thickness profiles of the sub-retinal layers and two cyst fluids within the sub-retinal layers) from the images and used SVM for classification. Sun et al. [27] took advantage of sparse coding and dictionary learning based on scale-invariant feature transform (SIFT) descriptor for automated detection of AMD, DME, and normal (NOR).

Venhuizen et al. [28] proposed an unsupervised feature learning approach which extracted a set of small descriptive image patches from the training data and used the patches to create a patch occurrence histogram for each image. Wang et al. [29] systematically selected the linear configuration pattern (LCP) based features of the OCT images by the Correlation-based Feature Subset (CFS) selection algorithm. All the aforementioned approaches mainly consist of two components, namely extracting features and training classifier.

The emergence of deep learning brings new methods and ideas to the classification of retinal OCT images. Especially the development of the CNN has had an important impact on images classification [30]. The CNN makes it possible to process images in the form of pixels as input and to give the desired classification as output, which can avoid the complicated feature extraction in the traditional algorithm. Karri et al. [31] fine-tuned (fix the weights in the lower levels and retrain the weights of the upper levels with back propagation) the GoogLeNet [32] to identify retinal pathologies on spectral domain OCT (SD-OCT) images. In the same year, Rasti et al. [33] proposed a multiple-scale convolutional mixture of expert ensemble model which applies CNN on multiple-scale sub-images. Recently, Kermany et al. [34] proposed an image-based deep learning (IBDL) method, which fixed the weights of the network and was used as a feature extractor.

All of the methods mentioned above are used for classifying OCT B-scans, although the method in Ref. [25] could classify three-dimensional (3-D) retinal OCT images. Nevertheless, Albarrak et al. [35] first proposed a method based on decomposition and graph analysis which could directly deal with the 3-D OCT volume image. More recently, Fang et al. [36] presented a method utilizing the principal component analysis network (PCANet) to extract the features from each B-scan of the 3-D retinal OCT images and fusing multiple kernels as a composite kernel to exploit the strong correlations among features of the 3-D OCT images for classification.

CNN is widely used in computer vision (especially in image classification) since the ImageNet [37] competition in 2012. In recent years, variants of CNN architectures have been developed, such as AlexNet [30], GoogLeNet [32], ResNet [38]. However, annotating biomedical images is not only tedious and time consuming, but also demanding of costly, specialty-oriented knowledge and skills [39]. When deep learning is applied to the specific problems in the field of biomedicine, the problem of overfitting is often caused by the lack of training data, and a good result cannot be obtained. In the case of only a small number of annotated data, compared with initializing the weights of the deep CNN randomly, transfer learning is an effective way to accelerate model convergence and improve the classification accuracy [31,34,39,40]. Fine-tuning is a commonly used technique to learn the feature space shifts whose weights in the lower levels are fixed and those of the upper levels can be updated with back propagation. In each round of training, errors are calculated through forward propagation in the entire network and weights of the network are updated through backward propagation. The propagation is time-consuming because of too many parameters. To address the problem of insufficient training data and time-consuming propagation, we propose a method for automatic diagnosis of AMD, DME, and NOR in retinal OCT B-scans which can effectively identify different pathologies.

The rest of this paper is organized as follows. Section 2 illustrates our classification method based on transfer learning in detail. Experimental results and analysis on two SD-OCT are presented in Section 3, and Section 4 outlines conclusions and suggests some future work.

## 2. Proposed Method

Our classification approach mainly consists of three parts, as illustrated in Figure 1: (1) preprocess the OCT B-scans to reduce the morphological variations; (2) apply a pre-trained Inception V3 architecture [41] on each OCT B-scan to extract its middle-level features automatically; (3) train the CNN to extract the semantic level features of OCT B-scans for image classification.

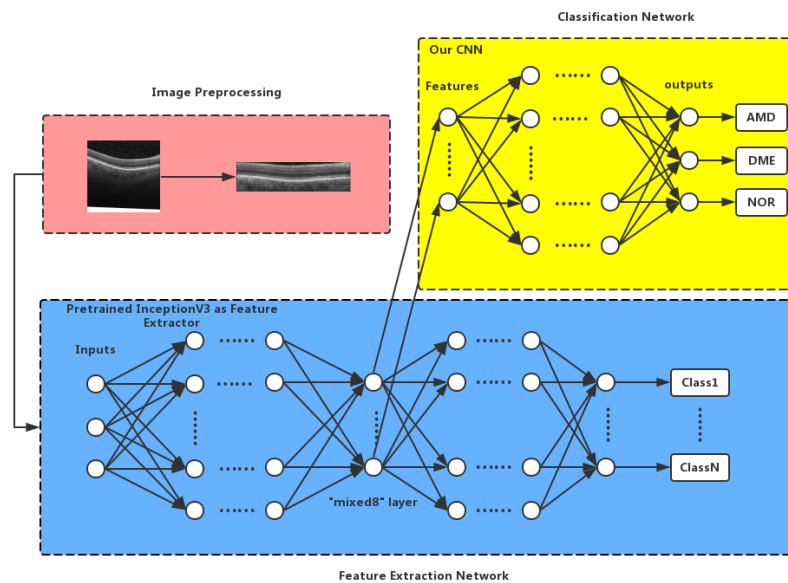
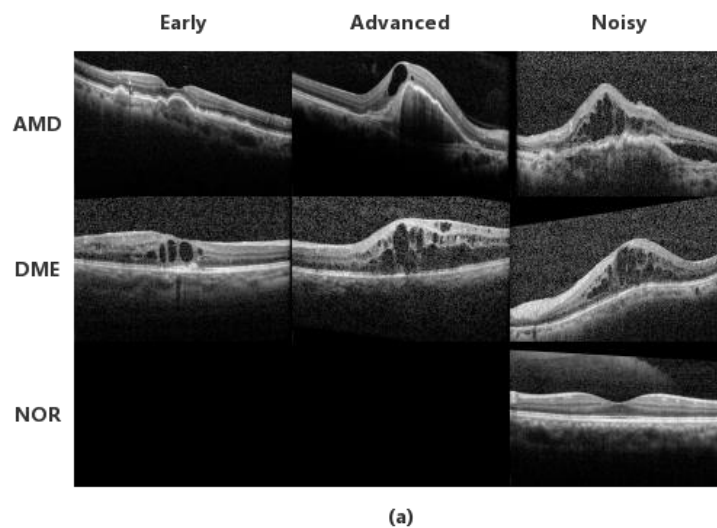


Figure 1. Outline of the proposed algorithm.

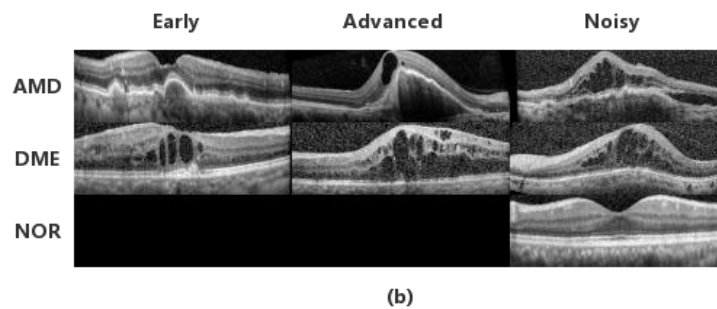
### 2.1. Image Preprocessing

We adopt the image preprocessing method in Ref. [27]. It is mainly divided into three stages: (1) In the perceiving stage, the method detects the overall morphology of a retina. Firstly, the sparsity-based block matching and 3-D-filtering (BM3D) denoising method [42] are used to reduce noises of the OCT image. Then, the binarization, median filtering, morphological closing and morphological opening methods are used to obtain the subject of the image. (2) In the fitting stage, the method automatically chooses the set of data points and a fitting method (linear fitting or second-order polynomial fitting). (3) In the normalizing stage, the method normalizes the retinas by aligning them to a relatively unified morphology and crops the images to trim out insignificant space. Examples can be found in Figure 2. The first column shows the early stage of AMD and DME, the second column shows the advanced stage of AMD and DME and the last column shows the noisy OCT B-scans of each category. Regardless of whether it's a deformed or a noisy OCT B-scan, the preprocessing method can detect the whole morphology of a retina and normalize it to a relatively unified morphology, as can be seen in Figure 2.



(a)

Figure 2. Cont.



**Figure 2.** Retinal dataset examples. (a) Original B-scans. (b) Corresponding preprocessed B-scans.

## 2.2. Inception V3 Feature Extraction

Using the Keras and Tensorflow we can easily adapt an Inception V3 architecture pretrained on the ImageNet. The Inception V3 architecture is represented in Figure 3. Kermany et al. [34] fixed the convolutional layer and used a feature extractor to obtain the bottleneck features. The bottleneck features described in Ref. [34] is the output of InceptionE2, as shown in Figure 3. However, the features learned by deep neurons are too specialized (such as eyes, nose) which do not describe the B-scans well, while using the features learned from low-level neurons does not fully exploit the advantages of deep learning. Different from the methods in Ref. [34], we extract the middle-level feature from the “mixed8” layer (the last layer of the InceptionD as shown in Figure 3) instead of the bottleneck (the last layer of the InceptionE2 as shown in Figure 3). The features learned by middle-level neurons can better describe the characteristics of the B-scans because of the difference between the natural pictures and retinal OCT images. On the other hand, retraining an Inception V3 is not only time consuming, but is also likely to have problems with overfitting because of insufficient training set. Therefore, we use a strategy to use a modified Inception V3 architecture pretrained on the ImageNet as a feature extractor, whose weights are fixed and cannot be updated. The middle-level features of OCT B-scans can be initially calculated and stored in order to remove redundant training processes and reduce the training time of the method.

## 2.3. Convolutional Neural Network

CNN is a special neural network whose hidden units are only connected to a local receptive field. We trained a simple CNN with the features extracted from the previous step as input to improve the classification accuracy. The weights of the CNN will be randomly initialized to learn the feature space shifts between ImageNet and retinal OCT images. As a result, the features extracted from the CNN can better describe the high-level characteristics of retinal OCT images. The CNN architecture are illustrated in Table 1. The convolutional layer captures the salient features in the local area, and the batch normalization layer can avoid overfitting. More importantly, the max-pooling layer reduces the number of parameters and makes the model easier to train. After the last batch normalization layer, the flatten layer can change the dimensions of the data for the transition of the batch normalization layer and the fully connected layer. The three outputs of the fully connected layer with an activation function soft-max shows the probability that the input B-scan belongs to each category (AMD, DME, and NOR). The CNN can extract information and identify the different characteristics from the training set. Therefore, it can improve the identification accuracy of the proposed method.

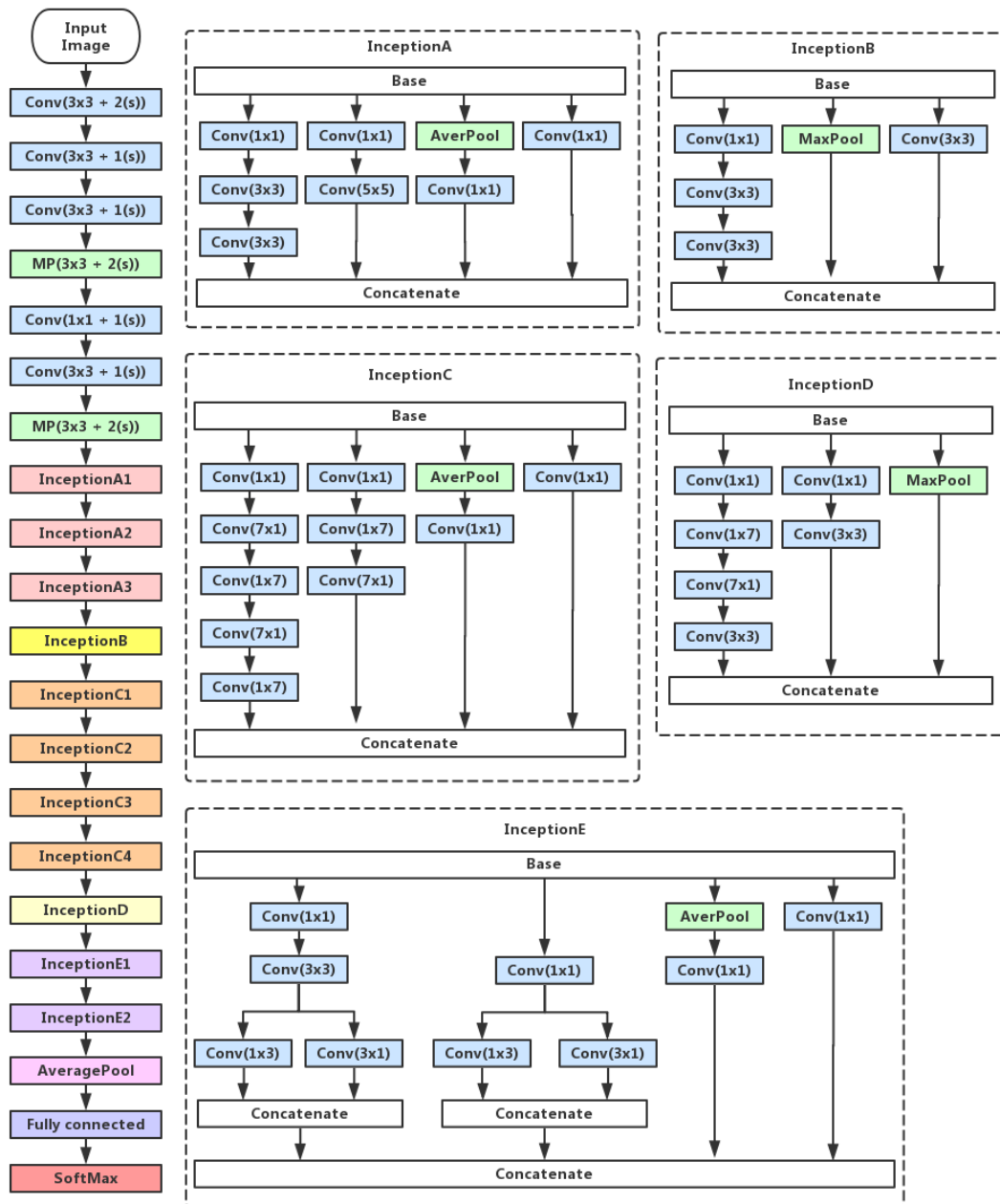


Figure 3. Inception V3 with filter dimensions and illustration of the inception layer.

Table 1. Overview of the convolutional neural network (CNN) architecture.

Name	Patch Size/Stride	Padding	Output Size
Convolution1	$3 \times 3/1$	same	$128 \times 8 \times 8$
Max Pooling1	$2 \times 2/2$	valid	$128 \times 4 \times 4$
BatchNormalization1			$128 \times 4 \times 4$
Convolution2	$3 \times 3/1$	same	$128 \times 4 \times 4$
Max Pooling2	$2 \times 2/2$	valid	$128 \times 2 \times 2$
BatchNormalization2			$128 \times 2 \times 2$
Convolution3	$3 \times 3/1$	same	$128 \times 2 \times 2$
BatchNormalization3			$128 \times 2 \times 2$
Flatten			512
Dense			3

### 3. Experiments and Results

In this section, we introduce two different datasets applied in our experiments and detail several comparative experiments. Our method mainly compares with the following methods: HOG-SVM [25], ScSPM [27], and deep CNN [31]. In addition to the results from the proposed method, we also quote some results directly from the literature.

The effectiveness of the proposed method is evaluated from the following experiments. In the following experiments 1 and 2, we apply our method to the Duke dataset to classify the volume data and compare it with the other method. In experiment 3, we apply our method on Beijing clinical datasets and compare with the good-performing ScSPM [27] in traditional methods and the transfer learning-based method, IBDL [34], to evaluate the classification performance of our method on retinal OCT B-scans. In experiment 4, we discuss the efficiency of our method. In experiment 5, we focus on the model structure itself and compare the results of different architectures.

In our experiments, we utilize cross validation to evaluate the classification performance of the proposed method. The cross validations are repeated with different random seeds to avoid dataset splitting bias. The accuracy, sensitivity, and specificity for each category (AMD, DME, and NOR) are defined by following formula [36]:

$$accuracy = \frac{TP + TN}{TP + FN + FP + FN} \quad (1)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$specificity = \frac{TN}{TN + FP} \quad (3)$$

where  $TP$  is true positive,  $FN$  is false negative,  $TN$  is true negative,  $FP$  is false positive.

Our classification algorithm is implemented using Keras and it is conducted on a desktop computer with an Intel (R) Core (TM) i5-4590 central processing unit (CPU), 8 GB of RAM and a Nvidia GeForce GTX 1070 graphics processing unit (GPU). The parameters selected for the proposed method are listed below:  $299 \times 299$  for the image size and the pixel value of the images is divided by 255 for normalization; “mixed8” for the middle layer name (the output of InceptionD discussed in Section 2.3); 128 for the filters amount,  $3 \times 3$  for the kernel size and “same” for the padding; the optimizer for training model is adadelta [43] optimizer.

#### 3.1. Datasets

We apply two different datasets in our experiments and present the experimental results over them. One of the datasets applied for experiments is the publicly available OCT dataset provided by the joint efforts from Duke University, Harvard University, and University of Michigan [25]. The Duke OCT dataset consists of retinal SD-OCT images from 45 subjects (15 AMD, 15 DME, and 15 NOR). The number of OCT B-scans in each subject varies from 36 to 97.

The other dataset is obtained from clinics in Beijing, using CIRRUS TM (Heidelberg Engineering Inc., Heidelberg, Germany) SD-OCT device. The dataset consists of 1680 OCT B-scan images (560 AMD, 560 DME and 560 NOR).

#### 3.2. Result Comparisons

##### Experiment 1: Comparison to the traditional method [25,27]

We first conduct our test on the Duke datasets. To compare our method with that in Ref. [25] and Ref. [27] fairly, we also use leave-three-out cross-validation for 45 times as done in Refs. [25,27]. In each experiment, 42 volumes are chosen as training set and the other 3 volumes (one volume from each class) as testing set so that each of the 45 SD-OCT volumes can be classified once. Since a volume contains many OCT scans of a specific person, majority voting of all predicted labels for each subject

is treated as class (AMD, DME, or NOR) for the corresponding subject [25]. We do not artificially exclude the normal tissue slices of AMD or DME as mentioned in the Ref. [25]. The cross-validation results can be seen in Table 2. As can be seen from Table 2, 100% of 45 OCT volumes are correctly classified with our method which invariably performs better than the HOG-SVM [25] (95.56%) and ScSPM [27] (97.78%).

**Table 2.** Fraction of volumes correctly classified with the three methods on the Duke dataset.

	HOG-SVM [25]	ScSPM [27]	Ours
AMD	15/15 = 100.00%	15/15 = 100.00%	15/15 = 100.00%
DME	15/15 = 100.00%	15/15 = 100.00%	15/15 = 100.00%
NOR	13/15 = 86.67%	14/15 = 93.33%	15/15 = 100.00%
Overall	43/45 = 95.56%	44/45 = 97.78%	45/45 = 100.00%

### Experiment 2: Comparison to transfer learning-based method [31]

Recently, some methods based on transfer learning have been proposed to classify the retinal OCT B-scans. According to Ref. [31], The dataset is divided into 15 folds, with each fold containing three subjects (one from each class). However, different from the setting in experiment 1, each validation involves eight folds (24 volumes) for training and the rest seven folds (21 volumes) for testing. Folds are sequentially chosen instead of randomly. The mean of decision pooling across all validations for AMD, DME, and NOR are reported in Table 3. The sensitivity of the HOG-SVM and deep CNN are quoted from Ref. [36]. As can be seen, the proposed method delivers better performance than HOG-SVM and deep CNN in total.

**Table 3.** Fraction of volumes correctly classified with the three methods on the Duke dataset.

	HOG-SVM [25]	Deep CNN [31]	Ours
AMD	89	89	89
DME	83	86	92
NOR	90	99	100

### Experiment 3: Comparisons on classification performance of retinal OCT B-scans

To evaluate the classification performance of our method on retinal OCT B-scans, we conduct more experiments on the Beijing clinic dataset with different proportion. To obtain reliable results, we repeat the experimental process 10 times with different randomly selected images in the clinic dataset for training the rest for testing. Firstly, we choose a quarter of the images of each class (140 AMD, 140 DME, and 140 NOR) for training and the rest (420 AMD, 420 DME, and 420 NOR) for testing. Then we choose half of the images of each class (280 AMD, 280 DME, and 280 NOR) and the rest (280 AMD, 280 DME, and 280 NOR) for testing. Tables 4 and 5 detail the classification results of each category. The overall classification results are tabulated in Table 6 (the best results among different methods are labeled in bold). From the results we can see that the proposed method performs better than ScSPM [27] and IBDL [34]. In particular, IBDL uses more than 100,000 OCT B-scans as training set with accuracy, sensitivity, and specificity above 95% according to Ref. [34]. However, the classification performance of IBDL method running on our dataset is not very good, indicating that it is not suitable for small-scale dataset. In conclusion, the proposed method has better classification performance on small-scale datasets.

**Table 4.** Classification results (%) on 1/4 clinic dataset.

Methods	Classes	Accuracy	Sensitivity	Specificity
ScSPM	AMD	97.35 $\pm$ 0.58	96.19 $\pm$ 0.85	97.94 $\pm$ 0.45
	DME	97.17 $\pm$ 0.44	93.81 $\pm$ 0.51	98.87 $\pm$ 0.46
	NOR	97.87 $\pm$ 0.15	98.73 $\pm$ 0.49	97.44 $\pm$ 0.08
IBDL	AMD	91.23 $\pm$ 0.38	85.40 $\pm$ 0.81	94.25 $\pm$ 0.58
	DME	94.77 $\pm$ 0.29	96.83 $\pm$ 0.22	93.65 $\pm$ 0.57
	NOR	91.63 $\pm$ 0.40	85.32 $\pm$ 0.40	94.92 $\pm$ 0.66
Ours	AMD	98.51 $\pm$ 0.19	98.14 $\pm$ 0.49	98.69 $\pm$ 0.44
	DME	97.80 $\pm$ 0.35	94.57 $\pm$ 0.76	99.43 $\pm$ 0.31
	NOR	98.35 $\pm$ 0.20	99.33 $\pm$ 0.41	97.85 $\pm$ 0.38

**Table 5.** Classification results (%) on 1/2 clinic dataset.

Methods	Classes	Accuracy	Sensitivity	Specificity
ScSPM	AMD	97.75 $\pm$ 0.21	96.43 $\pm$ 0.58	98.43 $\pm$ 0.08
	DME	97.60 $\pm$ 0.29	95.48 $\pm$ 0.89	98.67 $\pm$ 0.09
	NOR	97.91 $\pm$ 0.31	98.10 $\pm$ 0.84	97.81 $\pm$ 0.40
IBDL	AMD	93.36 $\pm$ 0.32	88.84 $\pm$ 2.78	95.66 $\pm$ 1.02
	DME	96.96 $\pm$ 0.14	98.13 $\pm$ 0.46	96.33 $\pm$ 0.24
	NOR	93.39 $\pm$ 0.25	89.11 $\pm$ 2.21	95.57 $\pm$ 0.99
Ours	AMD	99.01 $\pm$ 0.30	99.02 $\pm$ 0.39	99.01 $\pm$ 0.37
	DME	98.51 $\pm$ 0.27	96.34 $\pm$ 1.08	99.60 $\pm$ 0.20
	NOR	99.07 $\pm$ 0.21	99.55 $\pm$ 0.46	98.83 $\pm$ 0.32

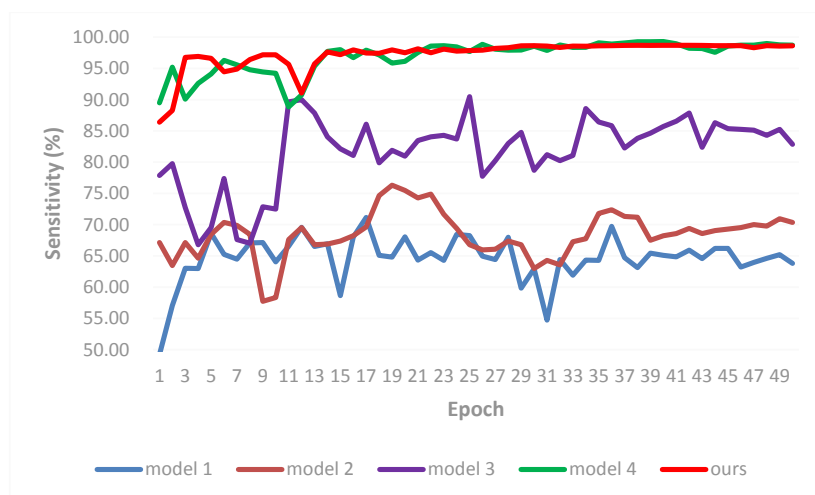
**Table 6.** Overall classification results (%) of different methods on clinic dataset.

Partition	Methods	Overall-Acc	Overall-Se	Overall-Sp
1/4 dataset	ScSPM	97.46	96.24	98.08
	IBDL	92.54	89.18	94.27
	Ours	98.22	97.35	98.66
1/2 dataset	ScSPM	97.75	96.67	98.30
	IBDL	94.57	92.03	95.85
	Ours	98.86	98.30	99.15

#### Experiment 4: Comparison with the efficiency of the fine-tuning

In this experiment, we focus on comparing the proposed method with fine-tuning a pre-trained network, including performance, training time, etc. Fine-tuning a pre-trained network is adopted by Ref. [31]. We conduct an experiment on the Beijing clinic dataset. Half of the clinic dataset are chosen as training sets and the rest as testing images. The experiment is repeated 10 times with 50 epochs for each time. We use Inception V3 as the pre-trained network (pre-trained on ImageNet dataset) instead of GoogLeNet in Ref. [31] for fairness. In this experiment, we fine-tune different layers for comparison: (1) The weights of fully connected layers are randomly initialized and the other weights of the network are fixed (model 1). (2) The weights of the InceptionE2 layer and the fully connected layer can be updated while the others are fixed (model 2). (3) The weights of the InceptionE1, InceptionE2 and the fully connected layer can be updated while the others are fixed (model 3). (4) All of the weights in the network can be updated (model 4). The mean sensitivity of each epoch is shown in Figure 4. After 50 epochs for training, the sensitivity of each model (from model 1 to model 4) are 63.81%, 70.36%, 82.86%, and 98.73%, respectively, and ours is 98.61%. In terms of training efficiency, the average time of the fine-tuning method to train 50 epochs is 474 s while the proposed method is 95 s. To prove the computational efficiency advantage of our method, we try to use only the CPU to perform calculation,

the average time of the fine-tuning method to train 50 epochs is 10,106 s while that of our method is only 151 s. In general, our method demonstrates competitive performance in classification accuracy and reduces the training time in comparison to the fine-tuning methods.

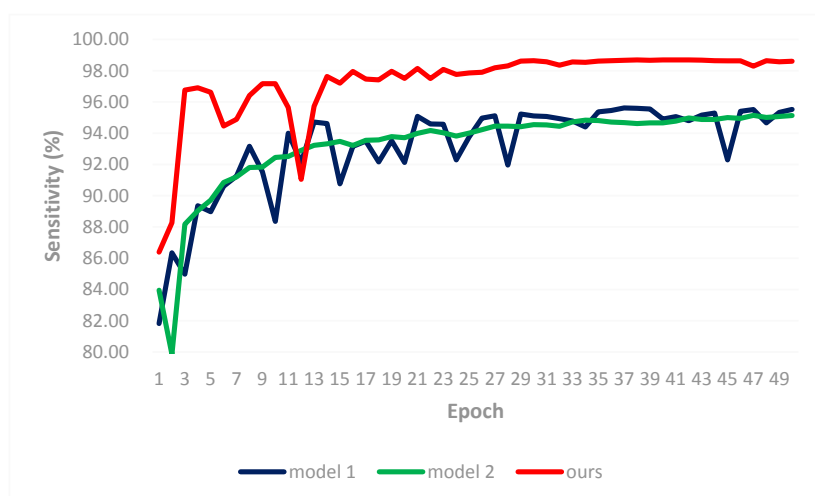


**Figure 4.** The classification sensitivity of different methods.

#### Experiment 5: Effectiveness of different architectures

Here, we conduct an experiment on the Beijing clinic dataset to prove the effectiveness of different architectures. The dataset is divided as in experiment 4. The experiment is repeated 10 times with 50 epochs for each time. The accuracy of each epoch is recorded for calculating the mean accuracy of the 10 experiments.

We build two classification models for comparison. The first model (model 1) extracts features from the bottleneck instead of middle layer and the features are sent to the CNN to extract advanced semantic features. Another model (model 2) extracts features from the “mixed8” layer (the output of InceptionD) but the features are used directly as input to the fully connected layer. The mean sensitivity of each model was shown in Figure 5. After 50 epochs for training, the sensitivity for model 1, model 2, and the proposed method are 95.52%, 95.13%, and 98.61%, respectively. From the result, we can conclude that: (1) Extracting the features from the middle layer is better than from the bottleneck in our case. (2) The CNN can effectively improve the classification accuracy.



**Figure 5.** The mean accuracy of different architectures.

#### 4. Conclusions

In recent years, many deep learning architectures have been proposed and used for natural pictures classification, especially for the classification in ImageNet database. On the other hand, biomedical images are not easily accessible because annotating them is time consuming and requires specialty-oriented knowledge. For small-scale biomedical training data, transfer learning is one of the best methods applied to biomedical image classification. A common method is to fine-tune the weights of the last several layers of the pretrained network and fix the others. In this paper, we present a fully automatic method to identify AMD, DME, and NOR from retinal OCT B-scans. The proposed method extracts the middle-level features from the middle layer of a pretrained network, then trains a CNN to learn the abstract semantic features of the retinal OCT B-scans. The proposed method is successfully tested in two different datasets for the detection of AMD and DME. The experimental results show that the proposed method demonstrates competitive performance in classification accuracy and has an advantage in computational efficiency compared with the fine-tuning method and shows better classification performance than that of traditional methods. To summarize, our algorithm is an effective method for retinal OCT detection and is a potentially impactful tool for computer-aided diagnosis and screening of ophthalmic diseases.

To help both researchers and practitioners further investigate retinal OCT image classification, we provide a Python implementation of the proposed method. The code is available at [https://github.com/HeWenjie/retina\\_OCT\\_images\\_classification](https://github.com/HeWenjie/retina_OCT_images_classification).

**Author Contributions:** Q.J., Y.S. conceived and designed the experiments; W.H. and J.H. performed the experiments and wrote the paper.

**Funding:** This research received no external funding.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China under Grant No. 61671272 and by the Opening Project of Guangdong Province Key Laboratory of Big Data Analysis and Processing under Grant No. 201803.

**Conflicts of Interest:** The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose. The research data were acquired and processed from patients by coauthors unaffiliated with any commercial entity.

#### References

1. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Grogory, K.; Puliafito, C.A.; et al. Optical coherence tomography. *Science* **1991**, *254*, 1178–1181. [[CrossRef](#)] [[PubMed](#)]
2. Podoleanu, A.G.; Rosen, R.B. Combinations of techniques in imaging the retina with high resolution. *Prog. Retinal Eye Res.* **2008**, *27*, 464–499. [[CrossRef](#)] [[PubMed](#)]
3. Cogliati, A.; Canavesi, C.; Hayes, A.; Tankam, P.; Duma, V.-F.; Santhanam, A.; Thompson, K.P.; Rolland, J.P. MEMS-based handheld scanning probe with pre-shaped input signals for distortion-free images in Gabor-Domain Optical Coherence Microscopy. *Opt. Express* **2016**, *24*, 13365–13374. [[CrossRef](#)] [[PubMed](#)]
4. Choma, M.A.; Sarunic, M.V.; Yang, C.; Izatt, J.A. Sensitivity advantage of swept-source and Fourier-domain optical coherence tomography. *Opt. Express* **2003**, *11*, 2183–2189. [[CrossRef](#)] [[PubMed](#)]
5. Virgili, G.; Menchini, F.; Casazza, G.; Hogg, R.; Das, R.R.; Wang, X.; Michelessi, M. Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy. *Cochrane Database Syst.* **2015**, *1*, CD008081. [[CrossRef](#)] [[PubMed](#)]
6. Keane, P.A.; Patel, P.J.; Liakopoulos, S.; Heussen, F.M.; Sadda, S.R.; Tufail, A. Evaluation of age-related macular degeneration with optical coherence tomography. *Surv. Ophthalmol.* **2012**, *57*, 389–414. [[CrossRef](#)] [[PubMed](#)]
7. Antony, B.J.; Abramoff, M.D.; Harper, M.M.; Jeong, W.; Sohn, E.H.; Kwon, Y.H.; Kardon, R.; Garvin, M.K. A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in SD-OCT volumes. *Biomed. Opt. Express* **2013**, *4*, 2712–2728. [[CrossRef](#)] [[PubMed](#)]
8. Carass, A.; Lang, A.; Hauser, M.; Calabresi, P.A.; Ying, H.S.; Prince, J.L. Multiple-object geometric deformable model for segmentation of macular OCT. *Biomed. Opt. Express* **2014**, *5*, 1062–1074. [[CrossRef](#)] [[PubMed](#)]

9. Chiu, S.J.; Izatt, J.A.; O'Connell, R.V.; Winter, K.P.; Toth, C.A.; Farsiu, S. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Investig. Ophthalmol. Vis. Sci.* **2012**, *53*, 53–61. [[CrossRef](#)] [[PubMed](#)]
10. Chiu, S.J.; Li, X.T.; Nicholas, P.; Toth, C.A.; Izatt, J.A.; Farsiu, S. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. *Opt. Express* **2010**, *18*, 19413–19428. [[CrossRef](#)] [[PubMed](#)]
11. DeBuc, D.C.; Somfai, G.M.; Ranganathan, S.; Tátrai, E.; Ferencz, M.; Puliafito, C.A. Reliability and reproducibility of macular segmentation using a custom-built optical coherence tomography retinal image analysis software. *J. Biomed. Opt.* **2009**, *14*, 064023. [[CrossRef](#)] [[PubMed](#)]
12. Fernández, D.C.; Salinas, H.M.; Puliafito, C.A. Automated detection of retinal layer structures on optical coherence tomography images. *Opt. Express* **2005**, *13*, 10200–10216. [[CrossRef](#)]
13. Ishikawa, H.; Stein, D.M.; Wollstein, G.; Beaton, S.; Fujimoto, J.G.; Schuman, J.S. Macular segmentation with optical coherence tomography. *Investig. Ophthalmol. Vis. Sci.* **2005**, *46*, 2012–2017. [[CrossRef](#)] [[PubMed](#)]
14. Lang, A.; Carass, A.; Hauser, M.; Sotirchos, E.S.; Calabresi, P.A.; Ying, H.S.; Prince, J.L. Retinal layer segmentation of macular OCT images using boundary classification. *Biomed. Opt. Express* **2013**, *4*, 1133–1152. [[CrossRef](#)] [[PubMed](#)]
15. Mayer, M.A.; Hornegger, J.; Mardin, C.Y.; Tornow, R.P. Retinal nerve fiber layer segmentation on FD-OCT scans of normal subjects and glaucoma patients. *Biomed. Opt. Express* **2010**, *1*, 1358–1383. [[CrossRef](#)] [[PubMed](#)]
16. Mishra, A.; Wong, A.; Bizheva, K.; Clausi, D.A. Intra-retinal layer segmentation in optical coherence tomography images. *Opt. Express* **2009**, *17*, 23719–23728. [[CrossRef](#)] [[PubMed](#)]
17. Mujat, M.; Chan, R.; Cense, B.; Park, B.; Joo, C.; Akkin, T.; Chen, T.; de Boer, J. Retinal nerve fiber layer thickness map determined from optical coherence tomography images. *Opt. Express* **2005**, *13*, 9480–9491. [[CrossRef](#)] [[PubMed](#)]
18. Paunescu, L.A.; Schuman, J.S.; Price, L.L.; Stark, P.C.; Beaton, S.; Ishikawa, H.; Wollstein, G.; Fujimoto, J.G. Reproducibility of nerve fiber thickness, macular thickness, and optic nerve head measurements using StratusOCT. *Investig. Ophthalmol. Vis. Sci.* **2004**, *45*, 1716–1724. [[CrossRef](#)]
19. Shahidi, M.; Wang, Z.; Zelkha, R. Quantitative thickness measurement of retinal layers imaged by optical coherence tomography. *Am. J. Ophthalmol.* **2005**, *139*, 1056–1061. [[CrossRef](#)] [[PubMed](#)]
20. Sun, Y.; Zhang, T.; Zhao, Y.; He, Y. 3D automatic segmentation method for retinal optical coherence tomography volume data using boundary surface enhancement. *J. Innov. Opt. Health Sci.* **2016**, *9*, 1650008. [[CrossRef](#)]
21. Vermeer, K.A.; van der Schoot, J.; Lemij, H.G.; de Boer, J.F. Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images. *Biomed. Opt. Express* **2011**, *2*, 1743–1756. [[CrossRef](#)] [[PubMed](#)]
22. Reisman, C.A.; Chan, K.; Ramachandran, R.; Raza, A.; Hood, D.C. Automated segmentation of outer retinal layers in macular OCT images of patients with retinitis pigmentosa. *Biomed. Opt. Express* **2011**, *2*, 2493–2503. [[CrossRef](#)]
23. Liu, Y.Y.; Chen, M.; Ishikawa, H.; Wollstein, G.; Schuman, J.S.; Rehg, J.M. Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. *Med. Image Anal.* **2011**, *15*, 748–759. [[CrossRef](#)] [[PubMed](#)]
24. Sugruk, J.; Kiattisin, S.; Leelasantham, A. Automated classification between age-related macular degeneration and diabetic macular edema in OCT image using image segmentation. In Proceedings of the 7th Biomedical Engineering International Conference, Fukuoka, Japan, 26–28 November 2014; pp. 1–4.
25. Srinivasan, P.P.; Kim, L.A.; Mettu, P.S.; Cousins, S.W.; Comer, G.M.; Izatt, J.A.; Farsiu, S. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed. Opt. Express* **2014**, *5*, 3568–3577. [[CrossRef](#)] [[PubMed](#)]
26. Hassan, B.; Raja, G.; Hassan, T.; Usman Akram, M. Structure tensor based automated detection of macular edema and central serous retinopathy using optical coherence tomography images. *J. Opt. Soc. Am. A* **2016**, *33*, 455–463. [[CrossRef](#)] [[PubMed](#)]
27. Sun, Y.K.; Li, S.; Sun, Z.Y. Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *J. Biomed. Opt.* **2017**, *22*, 16012. [[CrossRef](#)] [[PubMed](#)]

28. Venhuizen, F.G.; van Ginneken, B.; Bloemen, B.; van Grinsven, M.J.J.P.; Philipsen, R.; Hoyng, C.; Theelen, T.; Sánchez, C.I. Automated age-related macular degeneration classification in OCT using unsupervised feature learning. *Med. Imaging Comput.-Aided Diagn.* **2015**, *9414*. [[CrossRef](#)]
29. Wang, Y.; Zhang, Y.; Yao, Z.; Zhao, R.; Zhou, F. Machine learning based detection of age-related macular degeneration (AMD) and diabetic macular edema (DME) from optical coherence tomography (OCT) images. *Biomed. Opt. Express* **2016**, *7*, 4928–4940. [[CrossRef](#)] [[PubMed](#)]
30. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [[CrossRef](#)]
31. Karri, S.; Chakraborty, D.; Chatterjee, J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomed. Opt. Express* **2017**, *8*, 579–592. [[CrossRef](#)] [[PubMed](#)]
32. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015; pp. 1–9.
33. Rasti, R.; Mehrdehnavi, A. Macular OCT Classification using a Multi-Sacle Convolutional Neural Network Ensemble. *IEEE Trans. Med. Imaging* **2018**. [[CrossRef](#)] [[PubMed](#)]
34. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* **2018**, *172*, 1122–1131. [[CrossRef](#)] [[PubMed](#)]
35. Albarrak, A.; Coenen, F.; Zheng, Y.; Yu, W. Volumetric image mining based on decomposition and graph analysis: An application to retinal optical coherence tomography. *Comput. Intell. Inform.* **2012**, 263–268. [[CrossRef](#)]
36. Fang, L.; Wang, C.; Li, S.; Yan, J.; Chen, X.; Rabbani, H. Automatic classification of retinal three-dimensional optical coherence tomography images using principal component analysis network with composite kernels. *J. Biomed. Opt.* **2017**, *22*, 1–10. [[CrossRef](#)] [[PubMed](#)]
37. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.-F. Imagenet: A large-scale hierarchical image database. In Proceedings of the Computer Vision and Pattern Recognition, Miami Beach, FL, USA, 22–25 June 2009; pp. 248–255.
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 27–30 June 2016.
39. Zhou, Z.; Shin, J.; Zhang, L.; Gurudu, S.; Gotway, M.; Liang, J. Fine-tuning convolutional neural networks for biomedical image analysis: Actively and incrementally. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 21–26 July 2017; pp. 7340–7349.
40. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imaging* **2016**, *35*, 1299–1312. [[CrossRef](#)] [[PubMed](#)]
41. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 27–30 June 2016.
42. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]
43. Matthew, D.Z. ADADELTA: An adaptive learning rate method. *Tech. Rep.* **2012**, arXiv:1212.5701.

