Architecture
Image Studies

Vol.7, Issue 1, pp.2479-2488, 2026
DOI: https://doi.org/10.62754/ais.v7i1.1247
© by AP2 on Creative Commons 4.0
International License (CC BY-NC 4.0)
https://journals.ap2.pt/index.php/ais/index

Classification of Age-Related Macular Degeneration Disease with Deep Learning based on Optical Coherence Tomography images

Kiran Venneti[1], R. Murugan[2]

Abstract

Because Deep Learning can predict events accurately and quickly, it has garnered a lot of interest from the scientific community. In this study, age-related macular degeneration was categorized using deep learning. In persons over 50, age-related macular degeneration (AMD) is the main factor causing blindness. AMD first manifests as a black spot in the center of the field of vision and a loss of central vision. With an ageing population, the prevalence of this pathology is continuing to rise, creating a serious public health issue. AI and deep learning are already proving to be disruptive in industries like medicine. Even while in the current state of affairs, an AI's recommendations cannot completely replace a doctor's, such systems could nevertheless lessen the overall labour that the doctor has to do. We employ optical coherence tomography (OCT) images and divide them into four groups based on the already recognized retinal illnesses. These groups are called drusen, normal, diabetic macular edema (DME), and choroidal neovascularization (CNV). We have performed image augmentation to prevent the overfitting of the trained models. We train Xception with additional layer, Inception V3 with additional layer and VGG19 with additional layer, out of which VGG19 with additional layer, produced better testing accuracy of 90.93%, which is best when maintaining the train, validation and test split ratio as 0.7, 0.1 and 0.2.

Keywords: *Age related Macular Degeneration, VGG 19, Inception V3, Xception, Optical Coherence Tomography (OCT), Deep Learning.*

Introduction

The condition known as age-related macular degeneration (AMD) affects your retina. It takes place when the macule, a section of the retina, is harmed. You lose your centre vision if you have AMD. No matter how close or how far away you are from something, you cannot perceive small details. However, your side vision will continue to be normal. Consider looking at a clock with hands as an illustration. You might be able to see the clock's numerals with AMD but not its hands. People begin to lose their ability to drive, see faces, and read tiny type as the disease progresses. People may not be aware they have AMD in its early stages because there may be no symptoms or indicators. AMD is quite widespread. For those 50 years of age or older, it is the main cause of visual loss.

Dry AMD and Wet AMD are the two different kinds of AMD. Dry shape is rather typical. Approximately 80% (8 out of 10) of AMD patients have the dry form. Parts of the macula thin out with ageing and develop microscopic protein clumps called drusen, which is a condition known as dry AMD (also known as atrophic AMD). Early, intermediate, and late stages of dry AMD all occur together. Typically, it develops gradually over a number of years. Your central vision gradually fades. Late dry AMD has no known cure, although there are techniques to maximise your remaining vision. You can also take precautions to safeguard your other eye if you only have late dry AMD in one of your eyes. The Wet version is less frequent but significantly more dangerous. When new, abnormal blood vessels develop beneath the retina, it is called wet AMD (also known as advanced neo- vascular AMD). The macula may get scarred as a result of these arteries leaking blood or other fluids. Wet AMD causes visual loss more quickly than dry AMD. Wet AMD can develop at any stage of dry AMD, however wet AMD is always late stage. If you consume a diet heavy in saturated fat (found in foods like meat, butter, and cheese), are overweight, smoke cigarettes, are over 50, have hypertension (high blood pressure), or have a family history of AMD, you are more likely to develop the condition. The generation of cross-sectional retinal images with semi-histologic resolution is possible because to optical coherence tomography (OCT) technology [1]. It enables the definition of the location and kind of alterations in the

[1]Biomedical Imaging Lab (BIOMIL), Department of Electronics and Communication Engineering, National Institute of Technology, Silchar, Assam, India 788010.
[2] Biomedical Imaging Lab (BIOMIL), Department of Electronics and Communication Engineering, National Institute of Technology, Silchar, Assam, India 788010. Corresponding Author E-mail: murugan.rmn@ece.nits.ac.in.

retina and surrounding structures and assesses the retina's and those structures' thickness. The components of exudative macular degeneration, such as serous retinal detachments, haemorrhages, and subretinal neovascular membranes, are best understood via OCT. Early detection of these illnesses can improve the likelihood of successful patient treatment and visual recovery.

In this study, diabetic macular edema (DME), drusen, and choroidal neovascularization (CNV) are the main types of eye diseases that are taken into consideration.

1. *Choroidal Neovascularization (CNV***):** It is a subtype of exudative age-related degeneration (AMD), which is characterized by an aberrant development of capillaries that pass through the Bruch's membrane from the choroidal vasculature to the neurosensory retina.
2. *Diabetic Macular Edema (DME):* It results from the retina's blood vessels being dam- aged. If untreated, a buildup of fluids in the macular would result in a swelling region on the retinal layer and ultimately irreversible blindness. In contrast to wet AMD, which is characterized by aberrant blood vessel growth, DME is brought on by damage to the retina's pre-existing blood vessels.
3. *Drusen:* It is one of AMD pathology's earliest symptoms. Ageing causes the macula's cells to begin to alter. Drusen are brought on by dry AMD 90% of the time and wet AMD 10% of the time.

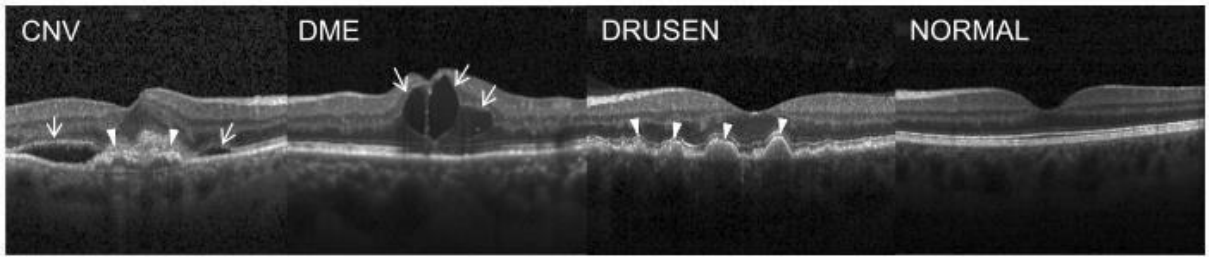

Figure 1: (Far left) choroidal neovascularization (CNV) with neovascular membrane (white arrowheads) and associated subretinal fluid (arrows). (Middle left) Diabetic macular edema (DME) with retinal-thickening-associated intraretinal fluid (arrows). (Middle right) Multiple drusen (arrowheads) present in early AMD. (Far right) Normal retina with preserved foveal contour and absence of any retinal fluid/edema [2]

The paper is organized as follows: Section II presents the related works. Section III dis- cusses the materials and methods used. Section IV explains the results and discussion. Section V discusses the conclusion of this study.

Related Works

Recently, several classification techniques have been proposed by various researchers in the biomedical image processing domain that is used for the classification of AMD disease [3– 7]. The authors in [8] proposed two deep convolutional neural networks (CNN) which were evaluated using 5-fold cross validation on the OCT data at baseline to attempt to predict which eyes would convert to advanced AMD. First VGG16 model was used which is a popular CNN for image recognition was fine-tuned, and a novel, simplified CNN architecture was trained from scratch achieving AUC of 0.87 at the volume level. The authors in [9] used 52,690 normal macular OCT images and 48,312 AMD macular OCT images. A deep neural network was trained to categorize images as either normal or AMD. At the image level, they achieved an area under the ROC of 92.78% with an accuracy of 87.63%. The authors proposed a deep CNN architecture in [10] and compared the obtained results with pretrained models such as Inception V3 and VGG-16, their proposed CNN architecture gave an accuracy of 98.5 % and Inception V3 model gave an accuracy up to 99.27 % on the test set while VGG-16 gave only 53%. They later modified the VGG-16 architecture by adding more convolution layers and regularization terms to obtain a result up to 93.5%. The authors in [11] attained training accuracy and validation accuracies nearly 100%, and the cross-entropy was 0.005. The authors in [2] obtained 108,309 OCT images (37,206 with choroidal neovascularization, 11,349 with diabetic macular edema, 8,617 with drusen, and 51,140 normal) from 4,686 patients and used to train the AI system. The model was tested with 1,000 images (250 from each category) from 633 patients and achieved an accuracy of 96.6%, with a sensitivity of 97.8%, a specificity of 97.4%, and a weighted error of 6.6%. The authors in [12] trained

several CNNs and implemented an ensemble learning model based on CNNs to further improve the performance. Among the CNNs, ResNet152 shows the best results with 0.9810 accuracy, 0.9810 sensitivity, and 0.9937 specificity, and the ensemble learning based on three ResNet152 shows 0.989 accuracy, 0.989 sensitivity, and 0.996 specificity.

Materials and Methods

This section presents the database used in this study, the proposed ensembled model architecture, method to train our proposed model.

Materials

In this paper, the database used is Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images [13]. The database has 109309 OCT images, and its size is 8 GB. A training set is created from the images (108309 images, 37205 images from CNV, 11348 images from DME, 8616 images from Drusen and 51140 images from Normal class) and a testing set (1000 images, 250 images from each class). The four image directories are CNV, DME, DRUSEN, and NORMAL.

We then performed data augmentation techniques such as left-right and top-bottom flip, rotation, skew and zoom operations resulting in 164309 images. We then split these augmented images into a new train, validation and test folders in the ratio of 0.7, 0.1 and 0.2. The pro- posed ensemble model with transfer learning was trained on the new 115015 training images and 16429 validation images and tested on 32865 testing images. In this way, AMD disease classification by the proposed model was performed.

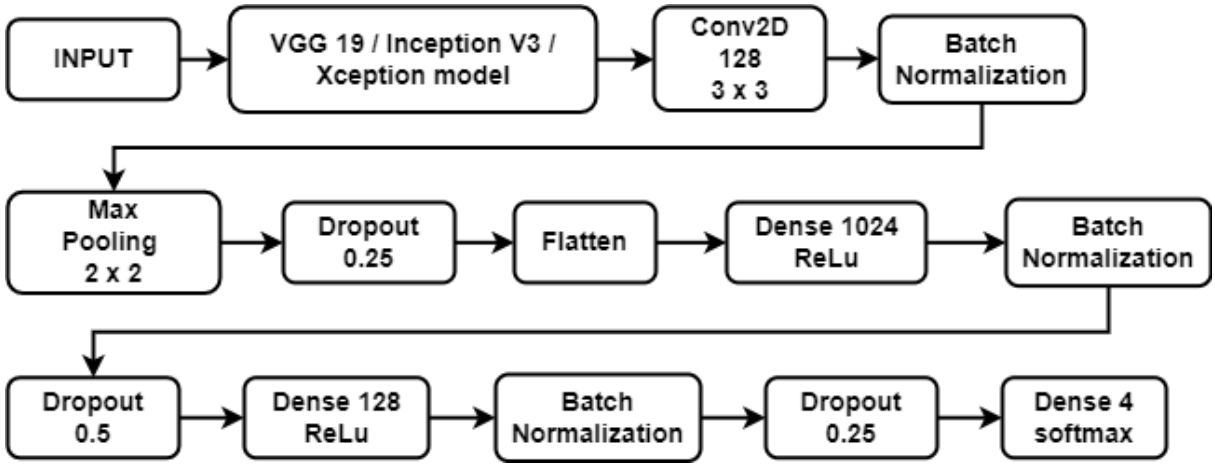

Figure 2: The pretrained VGG 19, Inception and Xception model with additional layers architecture

The Proposed Model Architecture

The model is Xception model with additional layer, Inception V3 with additional layer and VGG19 with the developed models were non trainable keeping only the additional layers trainable. The models with additional layers were individually trained by training set.

Fig.2 depicts the additional layers added to the pre-trained models consisting of one 3 × 3 convolutional layer followed by a rectified linear unit (ReLu), batch normalization, a 2 × 2 max pooling operation, dropout with a rate of 0.25, flatten layer, dense layer of 1024 units with an activation function as ReLu, batch normalization, dropout with a rate of 0.5, dense layer of 128 units with an activation function as ReLu, batch normalization, dropout with a rate of 0.25 and dense layer of 4 units with an activation function as softmax.

Method to Train the proposed Model

We have trained the OCT images over various models. At the beginning we increased the initial data set consisting of 109309 images to 164309 images by performing image augmentation and split

the dataset into train, test and validation set in the ratio of 0.7, 0.2 and 0.1. We used the training images to train the VGG19 with additional layer, Inception V3 with additional layer and Xception with additional layer as depicted in the Fig.2 and validated them during the training period using the validation set. Number of steps per epoch is 230 and validation steps is 1026. We trained them for 25 epochs. We also save the trained models which will be used in our proposed ensemble model. The proposed ensemble model takes an average of the individual predictions given from VGG19 with additional layer, Inception V3 with additional layer and Xception with additional layer and predict output. We use the unseen test data set to test the accuracy of the models. For training all the models we have used adam optimizer and categorical cross-entropy for the loss function.

Result and Discussion

This section presents the environment used to run the proposed model, performance metrics used to evaluate the model, the results of the proposed ensemble architecture. Later the obtained results are compared with the other models.

Environment

We have performed image augmentation on Jupyter Notebook with NVIDIA RX 6600 GPU. Python 3.7, Tensorflow, and Keras 2.11.0 are used to run the model in Kaggle, together with an NVIDIA Tesla P100-PCIE GPU and 25.46 GB of RAM.

Performance metrics

This study uses the performance metrics Accuracy, Precision, Recall, F1-Score, and AUC- ROC to assess the trained model. The numerical measurement of the classification outcome is called accuracy. Greater accuracy suggests a more accurate classification model. Equation (1) determines the accuracy.

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

Where TP (True Positives) is the proportion of accurately predicted instances that belong to the same class as the actual instances. The number of accurately predicted cases that do not fall under the actual class is known as TN (True Negatives). The number of inaccurately predicted examples that are regarded as belonging to the actual class is known as FP (False Positives). False Negatives (FN) are examples that are once included in the actual class but are now dis- regarded as not being a part of it. The proportion of real positives to all expected positives is known as precision. The ratio of real positives to all the positives in the ground truth is essentially what makes up a recall. Equations (2) and (3), respectively, yield it.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Precision and recall are taken together in the F1-score measure. In reality, the harmonic mean of the two is the F1 score. It is specified in Equation (4).

$$F1 - score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (4)$$

AUC-ROC is area under receiver operating characteristics curve. It makes use of true positive rates (TPR) and false positive rates (FPR). It is given by given by Equation (5) and Equation (6).

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

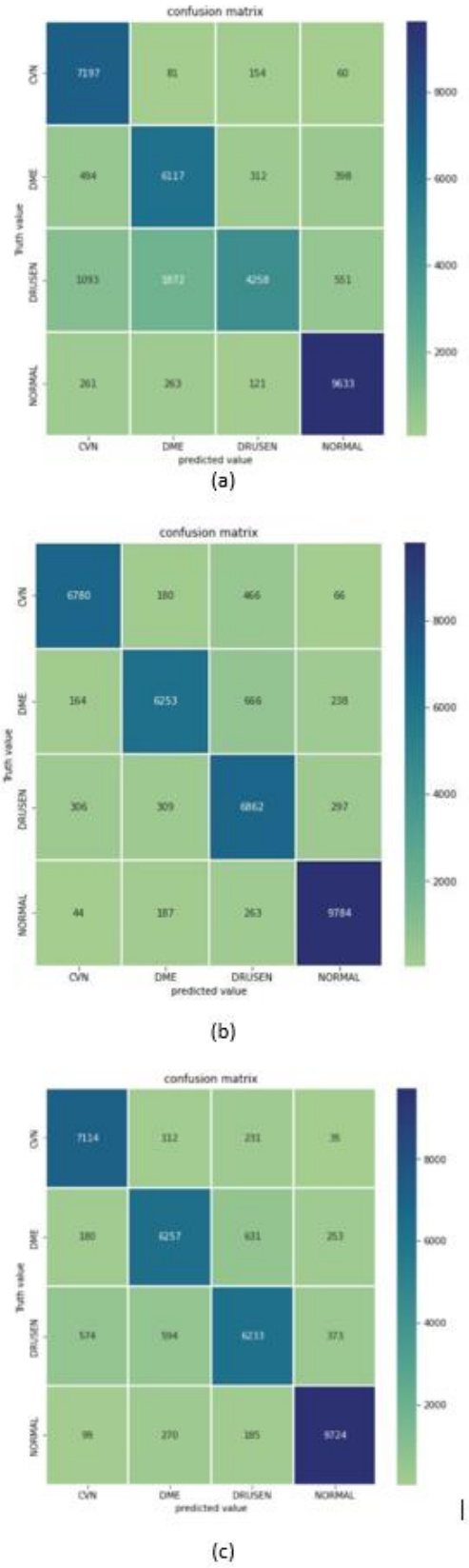

Figure 3: Confusion matrix obtained after training (a) Xception with added layers after augmentation (b) Inception V3 with added layers after augmentation (c) VGG 19 with added layers after augmentation

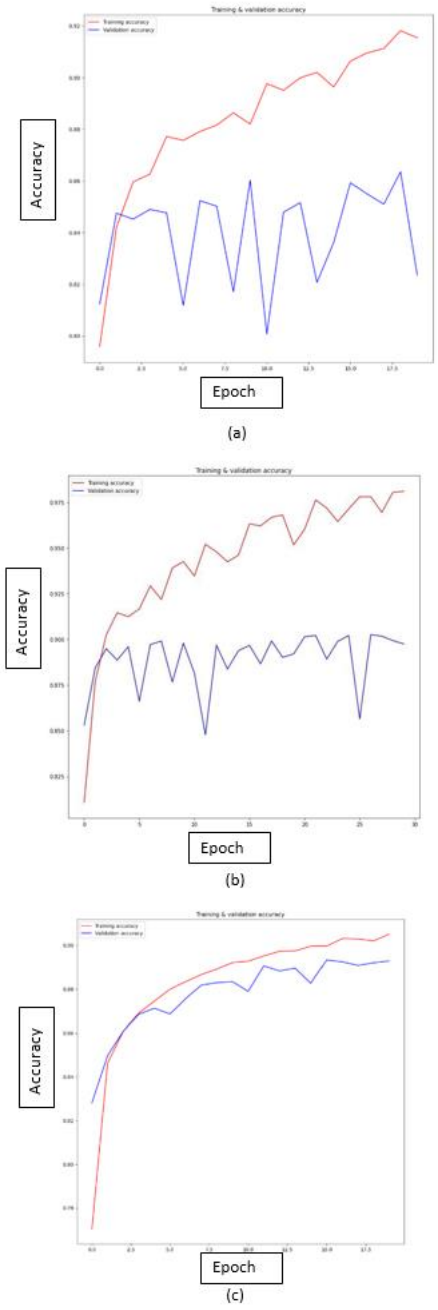

Figure 4: Training and Validation accuracy curve obtained for (a) Xception with added layers after augmentation (b) Inception V3 with added layers after augmentation (c) VGG 19 with added layers after augmentation

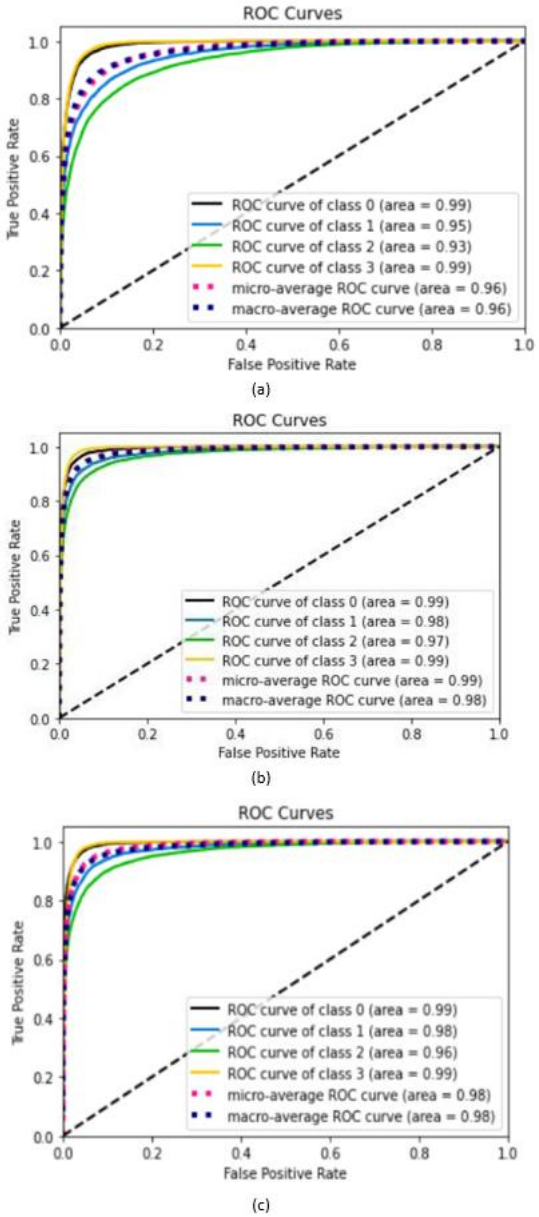

Figure 5: ROC curve obtained for (a) Xception with added layers after augmentation (b) Inception V3 with added layers after augmentation (c) VGG 19 with added layers after augmentation.

Result of proposed ensemble model

The model's training and validation accuracy are 93.31% and 91.22%, respectively, as shown in Fig.4. We obtained a testing accuracy of 90.93 %. Fig.3 and Fig.5 display the confusion matrix and ROC curve for this model, respectively.

Comparison of results

All the output metrics obtained from models used for training and testing of the OCT images as mentioned in the method section are noted in the Table I for the study of comparison of different models.

Table 1: Comparison with other models

	model	accuracy			class	precision	recall	f1-score	support	ROC curve area
		train	val	test						

1	Xception	94.99	88.15	88.19	CNV	0.88	0.94	0.91	7492	0.99
	with added layers				DME	0.91	0.79	0.84	7321	0.97
					DRUSEN	0.84	0.81	0.83	7774	0.96
					NORMAL	0.9	0.96	0.93	10278	0.99
					accuracy			0.88	32865	micro avg ROC
					macro avg	0.88	0.87	0.88	32865	macro avg ROC
					weighted avg	0.88	0.88	0.88	32865	
2	Inception V3	95.3	89.01	89.15	CNV	0.89	0.93	0.91	7492	0.99
	with added layers				DME	0.88	0.84	0.86	7321	0.97
					DRUSEN	0.85	0.82	0.84	7774	0.96
					NORMAL	0.93	0.95	0.94	10278	0.99
					accuracy			0.89	32865	micro avg ROC
					macro avg	0.89	0.89	0.89	32865	macro avg ROC
					weighted avg	0.89	0.89	0.89	32865	
3	VGG 19	93.31	91.22	90.93	CNV	0.92	0.94	0.93	7492	0.99
	with added layers				DME	0.86	0.91	0.88	7321	0.98
					DRUSEN	0.9	0.8	0.85	7774	0.97
					NORMAL	0.94	0.97	0.95	10278	0.99
					accuracy			0.91	32865	micro avg ROC
					macro avg	0.91	0.91	0.9	32865	macro avg ROC
					weighted avg	0.91	0.91	0.91	32865	

Comparison with the state-of-the-art

Table 2 presents a comparative analysis of the proposed model with existing state-of-the-art approaches for AMD disease classification. Hammou et al. [24] employed transformer-based architectures, including Vision Transformer (ViT) and Swin Transformer, achieving an accuracy of 82.76%. Although these models are effective in capturing global contextual features, their performance is limited by high computational complexity and the need for large-scale training data.

Moradi et al. [35] utilized a 3D CNN framework, which improved classification performance to 89.16% by leveraging volumetric feature representations. However, such architectures typically require extensive computational resources and longer training times.

Wang Q et al. [33] adopted a GAN-based approach, attaining an accuracy of 84.9%. While GANs enhance feature learning through data augmentation and adversarial training, they are often challenging to train and may suffer from instability issues.

In contrast, the proposed **Modified VGG19** model achieves the highest accuracy of **90.93%** for AMD disease classification. The improved performance can be attributed to the optimized network architecture, effective feature extraction, and better generalization capability, while maintaining relatively lower computational complexity compared to transformer- and GAN-based models. These

results demonstrate that the proposed approach provides a robust and efficient solution for accurate AMD disease detection.

Table 2: Comparison with the state-of-the-art

Author	Method / Architecture	Accuracy
Hammou et al.[14]	ViT, Swintransformer	82.76%
Moradi et al.[15]	3D CNN	89.16%
Wang Q et al. [16]	GAN	84.9%
Proposed Model	Modified VGG19	90.93%

Conclusion

In this paper, Xception with added layers, Inception V3 with added layers and VGG 19 with added layers, are developed for the classification of AMD disease. The proposed method has the edge over the individual VGG 19 with additional layer, performs exceptionally well even when the train, validation and test ratio is 70%, 10% and 20%. The proposed model obtained 90.93% testing accuracy.

By altering the CNN, we hope to increase the testing precision of the Drusen class and broaden the scope of the current investigation. In the future, we can develop a multi-task convolutional neural network model which can be used for the classification of the AMD disease.

References

[1] J. G. Fujimoto, C. Pitris, S. A. Boppart, and M. E. Brezinski, “Optical coherence tomog- raphy: an emerging technology for biomedical imaging and optical biopsy,” *Neoplasia*, vol. 2, no. 1-2, pp. 9–25, Jan. 2000.

[2] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, M. K. Prasadha, J. Pei, M. Y. Ting, J. Zhu, C. Li, S. Hewett, J. Dong, I. Ziyar, A. Shi, R. Zhang, L. Zheng, R. Hou, W. Shi, X. Fu, Y. Duan, V. A. Huu, C. Wen, E. D. Zhang, C. L. Zhang, O. Li, X. Wang, M. A. Singer, X. Sun, J. Xu, A. Tafreshi, M. A. Lewis, H. Xia, and K. Zhang, “Identifying medical diagnoses and treatable diseases by image-based deep learning,” *Cell*, vol. 172, no. 5, pp. 1122–1131.e9, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867418301545

[3] Y.-M. Chen, W.-T. Huang, W.-H. Ho, and J.-T. Tsai, “Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning,” *BMC Bioinformatics*, vol. 22, no. 5, p. 99, Nov 2021. [Online]. Available: https://doi.org/10.1186/s12859-021-04001-1

[4] T. Tsuji, Y. Hirose, K. Fujimori, T. Hirose, A. Oyama, Y. Saikawa, T. Mimura, K. Shiraishi, T. Kobayashi, A. Mizota, and J. Kotoku, “Classification of optical coherence tomography images using a capsule network,” *BMC Ophthalmology*, vol. 20, no. 1, p. 114, Mar 2020. [Online]. Available: https://doi.org/10.1186/s12886-020-01382-4

[5] A. Serener and S. Serte, “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)*, 2019, pp. 1–4.

[6] S. Saha, M. Nassisi, M. Wang, S. Lindenberg, Y. kanagasingam, S. Sadda, and Z. J. Hu, “Automated detection and classification of early amd biomarkers using deep learning,” *Scientific Reports*, vol. 9, no. 1, p. 10990, Jul 2019. [Online]. Available: https://doi.org/10.1038/s41598-019-47390-3

[7] S. Kuwayama, Y. Ayatsuka, D. Yanagisono, T. Uta, H. Usui, A. Kato, N. Takase, Y. Ogura, and T. Yasukawa, “Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images,” *Journal of Ophthalmology*, vol. 2019, p. 6319581, Apr 2019. [Online]. Available: https://doi.org/10.1155/2019/6319581

[8] D. B. Russakoff, A. Lamin, J. D. Oakley, A. M. Dubis, and S. Sivaprasad, “Deep Learning for Prediction of AMD Progression: A Pilot Study,” *Investigative Ophthalmology Visual Science*, vol. 60, no. 2, pp. 712–722, 02 2019. [Online]. Available: https://doi.org/10.1167/iovs.18-25325

[9] C. S. Lee, D. M. Baughman, and A. Y. Lee, “Deep learning is effective for the classifi- cation of OCT images of normal versus age-related macular degeneration,” *Ophthalmol Retina*, vol. 1, no. 4, pp. 322–327, Feb. 2017.

[10] M. Berrimi and A. Moussaoui, “Deep learning for identifying and classifying retinal dis- eases,” in *2020 2nd International Conference on Computer and Information Sciences (IC- CIS)*, 2020, pp. 1–6.

[11] M. Treder, “Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning.” graefe’s archive for clinical and experimental ophthalmology = albrecht von,” *Graefes Archiv fur klinische und exper- imentelle Ophthalmologie*, vol. 256, pp. 259–265, 2018.

[12] J. Kim and L. Tran, “Ensemble learning based on convolutional neural networks for the classification of retinal diseases from optical coherence tomography images,” in *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 2020,pp. 532–537.

[13] D. Kermany, “Large dataset of labeled optical coherence tomography (OCT) and chest X-Ray images,” 2018.

[14] Badr Ait Hammou, Fares Antaki, Marie-Carole Boucher, and Renaud Duval. Mbt: Model-based transformer for retinal optical coherence tomography image and video multi-classification. *International Journal of Medical Informatics*, 178:105178, 2023.

[15] Mousa Moradi, Yu Chen, Xian Du, and Johanna M Seddon. Deep ensemble learning for automated non-advanced amd classification using optimized retinal layer segmentation and sd-oct scans. *Computers in Biology and Medicine*, 154:106512, 2023.

[16] Chong Wang, Yuanhong Chen, Fengbei Liu, Michael Elliott, Chun Fung Kwok, Carlos Pena-Solorzano, Helen Frazer, Davis James McCarthy, and Gustavo Carneiro. An interpretable and accurate deep-learning diagnosis framework modeled with fully and semi-supervised reciprocal learning. *IEEE transactions on medical imaging*, 43(1):392–404, 2023.