

Enhancing Fundus Image Quality for Improved Age-Related Macular Degeneration Detection using Deep Learning

Ceren Yilmaz, Sinem Yesil, Elif Ilkay Ozkan, Erdem Okur
Izmir University of Economics, Izmir, Turkiye

Abstract—Age-related Macular Degeneration (AMD) is a leading cause of vision loss in the elderly, where early detection plays a critical role in slowing disease progression. Deep learning approaches have shown strong potential for automated AMD diagnosis from retinal images; however, their performance can be hindered by image quality variations, illumination inconsistencies, and artifacts. In this study, we propose a custom contrast enhancement mask to improve lesion visibility in fundus images prior to classification. Using the publicly available ADAM challenge dataset, three deep learning architectures—YOLOv8ncls, InceptionV3, and a modified U-Net encoder—were trained and evaluated on both the original and enhanced datasets. Experimental results demonstrate that the enhancement method substantially improves classification performance across all models, with YOLOv8ncls achieving the highest accuracy with 91.70%, and specificity with 96.76%. These findings highlight the importance of preprocessing in medical image analysis and suggest that lightweight models, when combined with effective enhancement techniques, can achieve high accuracy suitable for clinical deployment.

Keywords—AMD, Age-related Macular Degeneration, deep learning, image processing, neural networks

I. INTRODUCTION

Age-related Macular Degeneration (AMD) is one of the leading causes of irreversible vision loss among elderly people globally, with a significant impact on quality of life. This neurodegenerative macular disorder of the retina, which is responsible for sharp and detailed vision, has a tendency to advance insidiously in the early phase. Early and accurate diagnosis of AMD is essential for early treatment, which can slow down the progression of the disease and preserve residual vision [1]. Traditional diagnostic methods have a significant reliance on ophthalmic expertise and manual examination of fundus images, which can be subjective, labor-intensive, and prone to inter-observer variability. The increasing prevalence of AMD due to a globally aging population further strains healthcare resources, echoing an urgent need for efficient, objective, and scalable diagnostic solutions.

In the last several years, the emergence of deep learning (DL) and computer vision has presented promising prospects for the automation of detection and categorization of various retinal diseases from fundus images [2]. Convolutional Neural Networks (CNNs), in particular, have demonstrated superior performance in feature extraction and pattern recognition,

making them highly suitable for medical image analysis [3]. A number of studies have explored the application of CNNs to AMD detection, with different architectures and methods being employed to classify fundus images according to characteristic features [4][5][6].

Other than fundus images there are different datasets available with other type of images for instance, OCT images, as well as various approaches from several aspects. Serener and Serte aimed to classify dry and wet AMD types using OCT images [6]. The dataset they have used consisted of 83484 images categorized into four classes; healthy, diabetic macular edema (DME), dry AMD (drusen), and wet AMD (CNV). AlexNet and ResNet-18 architectures were compared, and ResNet demonstrated particularly high performance in dry AMD detection [7][8]. This study highlighted the effectiveness of transfer learning-based deep learning models, especially in the classification of dry AMD.

Ahnaf et al. conducted classification on 723 AMD and 1407 normal images from the Duke SD-OCT dataset using ResNet50 and VGG16 architectures [9][10][11][12]. The novelty of this study lied in its use of the Grad-CAM method to provide explainability in the classification process [13]. It was observed that the ONL-ISM (Outer Nuclear Layer to Inner Segment Myeloid) layer played a decisive role in both AMD and normal class decisions. This approach enhances the transparency and reliability of model decisions for clinical applications.

Lastly, in the study conducted by Singh et. al., the goal was to automatically classify ocular diseases such as diabetic retinopathy, glaucoma, AMD, and cataract [14]. Working on the ODIR dataset [15], which consists of images from 5000 patients, the researchers compared VGG19 and Inception V3 architectures. Interestingly, VGG19 outperformed InceptionV3 [16][17]. This success was attributed to VGG19's ability to capture fine details through its narrower receptive fields and its generalization capability across different demographics.

Despite the progress, however, there are still challenges to achieving robust and highly accurate automatic AMD detection from fundus images. Fundus images typically possess variable qualities, non-uniform illumination, and artifacts, which could potentially impair the performance of even state-of-the-art deep learning models. To address these limitations, our study proposes an enhanced framework for automated AMD detection with a particular focus on the critical role

played by image preprocessing. We hypothesize that by improving the quality and consistency of the fundus images prior to classification, we can boost the performance of off-the-shelf deep learning models.

In this work, we present a novel approach to the automatic AMD detection task, which integrates a preprocessing method for quality improvement of fundus images with their subsequent classification by well-known convolutional neural network models. We would like to demonstrate the impact of careful preprocessing on the accuracy and reliability of AMD detection, leading to a more stable and practical solution for clinical application. The remainder of this paper is organized as follows: Section II explains our new preprocessing approach and the deep learning architectures employed as well as the reasons why they are chosen. Section III presents the experimental setup and results. Finally, Section IV concludes the paper and addresses future work.

II. METHODOLOGY AND PROPOSED ENHANCEMENT

The methodology to test whether the contrast mask proposed here affects the classification is simple. Three deep learning architectures —YOLOv8n-cls, InceptionV3, and a modified U-Net encoder classifier— were trained and evaluated for binary classification of retinal fundus images into Age-related Macular Degeneration (AMD) and Normal (NO) categories [18][17][19]. All models were trained on a uniformly preprocessed dataset and images resized to appropriate size for the input layer of models augmented with horizontal flips and random rotations to overcome class imbalance. They were also trained and evaluated with another dataset which is essentially the same but preprocessed with proposed enhancement mask which constitutes the novel step of this work. It is a custom contrast enhancement mask to improve lesion visibility for the network models.

The network models are chosen for their well-known structures and versions readily available to use due to popularity. Moreover, they were selected over other architectures due to certain unique properties. For instance, YOLOv8n-cls prioritizes computational efficiency and real-time capability. The InceptionV3 leverages a custom classification head for deep semantic learning. Lastly, U-Net retains its encoder for spatially rich feature extraction. In the end, they are trained, tested for both datasets and compared against each other.

Following subsections will explain the dataset used in the study and preprocessing step which includes the contrast enhancement mask generation in detail.

A. The Dataset

The Automatic Detection challenge on Age-related Macular Degeneration (ADAM), introduced at ISBI 2020, is the first international competition designed to benchmark algorithms for AMD detection and characterization [20]. Addressing the lack of large, standardized datasets and comparable evaluation protocols in medical image analysis, ADAM builds upon the successes of the REFUGE and AGE challenges. The dataset provided by the challenge now consists of over 1500 retinal

images categorized as AMD or NO (normal). All images were obtained from CT Angiography (BTA), not traditional frontal fundus photography. This distinction is crucial because BTAdetected images are often easier to acquire in clinical workflows and are less dependent on ophthalmology-specific imaging hardware. Their standardization across patient records also enhances model transferability. Specifically two devices -a Zeiss Visucam 500, producing at a resolution of 2124×2056 pixels, and a Canon CR-2, yielding at 1444×1444 pixels- are used to collect the images. All images were taken in strict adherence to the standard operating procedures for each camera.

B. Preprocessing

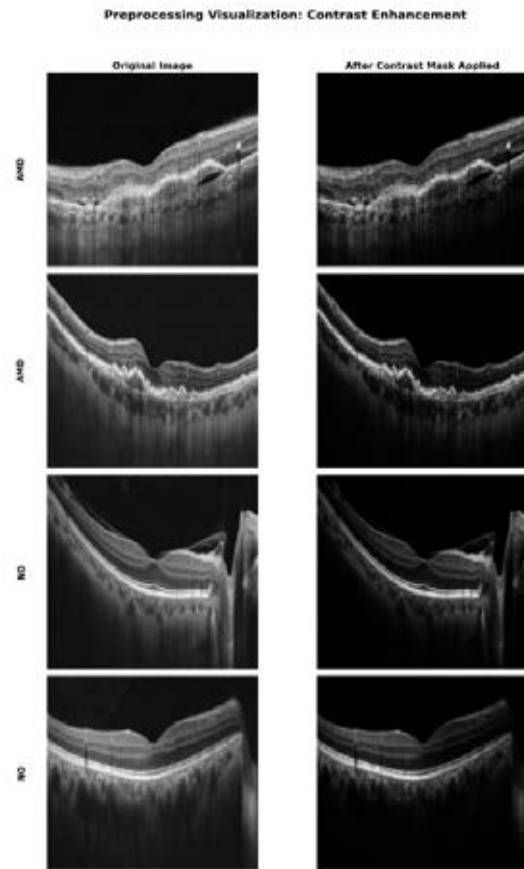


Fig. 1: After contrast enhancement preprocessing step the images as they are shown on right side. They look darker to the eye but the pixel value variation between tissue layers are higher now for detection.

In this study a similar image enhancement mask that is described by Okur and Turkan is developed but applied on regular BT images [21]. The mentioned study utilizes three different masks and their combinations to detect best result yielding enhancement mask. This study creates a mask similar to brightness mask of theirs and utilizes only this one due to the characteristics of the dataset images. The developed mask first transforms the raw images from the dataset to an RGB

image. Then calculates the mask as the absolute difference between its value and the mean of the color channels at the same spatial location. A feature map is generated for each color channel and linearly normalized to the $[0, 1]$ range. When combined into an image, this process produces an RGB contrast-enhancement mask, where each pixel value functions as a weight. These weights derived from explained calculation then, applied on images by direct per-pixel weighting method. This mask ensures that the pixel value difference between each tissue that can be seen in the image is high and more separable. It is hypothesized that the applied mask will make chosen network models understand AMD related features better. Fig. 1 shows original images and images after the mask application.

The preprocessing step ends with resizing of dataset images to the appropriate resolution for the network model to be used. The resolutions are 320×320 for UNet, 299×299 for YOLO and Inception.

III. EXPERIMENTAL SETUP AND RESULTS

All experiments were conducted on an MSI Alpha 15 A3DDK-034 laptop equipped with an AMD Ryzen 7 3750H processor and an AMD Radeon RX 5500M GPU. The system was configured with 16 GB DDR4 memory and the operating environment consisted of Python 3.10 with all experiments executed under consistent software and hardware conditions to ensure reproducibility.

TABLE I: Performance comparison of models on regular and enhanced datasets. The U-Net result on the regular dataset is omitted as it performed below chance level.

Model	Dataset	Accuracy	Recall (Sensitivity)	Specificity	NPV	Precision	F1-score
U-Net	Enhanced	86.60%	79.81%	91.62%	91.37%	80.03%	79.92%
YOLOv8	Enhanced	91.70%	84.88%	96.76%	92.99%	89.43%	86.88%
InceptionV3	Enhanced	89.36%	83.03%	94.05%	92.55%	84.57%	83.76%
YOLOv8	Regular	66.38%	78.28%	57.57%	99.53%	69.10%	64.28%
InceptionV3	Regular	78.72%	50.00%	100.00%	78.72%	39.36%	44.05%

Table I presents the classification performance of U-Net, YOLOv8, and InceptionV3 on both the enhanced and regular retinal image datasets. The results clearly show that the proposed enhancement method significantly improved performance across all models, suggesting that better lesion visibility and contrast contribute to more discriminative feature learning. Among the tested architectures, YOLOv8 trained on the enhanced dataset achieved the best overall metrics, with an accuracy of 91.70%, specificity of 96.76%, and F1-score of 86.88%. This combination of high specificity and strong F1-score indicates that YOLOv8 not only excels at correctly identifying non-AMD cases but also maintains strong detection capability for AMD, avoiding the common pitfall of biasing toward the majority class. InceptionV3 with the enhanced dataset followed closely, achieving 89.36% accuracy and a well-balanced sensitivity–specificity profile, reflecting its capacity to capture subtle pathological features despite its larger model size and longer training requirements.

In stark contrast, the models trained on the regular dataset displayed lower overall accuracy and substantially reduced F1-scores. YOLOv8's performance dropped to 66.38% accuracy, primarily due to a steep decline in specificity with 57.57%, suggesting difficulty in rejecting non-AMD cases without enhanced image features. InceptionV3 trained on the regular dataset exhibited perfect specificity but very low sensitivity, indicating severe bias toward predicting the majority non-AMD class—a clear sign of imbalance exploitation rather than true discriminative learning. The omission of U-Net's regular dataset result was warranted, as it performed below chance level, further reinforcing that raw, unenhanced fundus images provide insufficient discriminative signal for reliable AMD detection.

Overall, these results confirm that the proposed enhancement method improves both generalization and class balance, particularly benefiting lightweight, real-time models like YOLOv8, which demonstrated the most favorable trade-off between sensitivity and specificity in this work.

IV. CONCLUSION

This work introduced a contrast-based enhancement mask designed to improve lesion visibility in retinal fundus images for AMD detection. The method was evaluated using three distinct deep learning architectures on the ADAM challenge dataset in both regular and enhanced forms. Results consistently showed that the enhancement step improved classification metrics, with YOLOv8n-cl delivering the best overall performance in terms of accuracy, specificity, and F1-score. InceptionV3 and U-Net also benefited from the preprocessing, demonstrating improved generalization and reduced class imbalance effects. These outcomes confirm that targeted image preprocessing can significantly boost the performance of deep learning models in medical imaging tasks. This work serves as a preliminary step, confirming that the enhancement method can successfully boost the performance of well-known architectures. Resonating on that, future work will explore integrating this enhancement method into end-to-end trainable pipelines and validating the approach on additional datasets to further assess its robustness and clinical applicability. If successful, it is considered to move on the multi-disease retinal datasets to eventually boost lightweight systems to be used as diagnostic aid tools.

REFERENCES

- [1] "Age-related macular degeneration (amd)," <https://www.hopkinsmedicine.org/health/conditions-and-diseases/agerelated-macular-degeneration-amd>, Accessed: 2025-07-22.
- [2] Y. Jeong, Y.-J. Hong, and J.-H. Han, "Review of machine learning applications using retinal fundus images," *Diagnostics*, vol. 12, no. 1, 2022.
- [3] P. Celard, E.L. Iglesias, J.M. Sorribes-Fdez, R. Romero, A.S. Vieira, and L. Borraro, "A survey on deep learning applied to medical images: from simple artificial neural networks to generative models," *Neural. Comput. & Applic.*, vol. 35, pp. 2291–2323, 2023.
- [4] A. Bhuiyan, T.Y. Wong, D.S.W. Ting, A. Govindaiah, E.H. Souied, and R.T. Smith, "Artificial intelligence to stratify severity of age-related macular degeneration (amd) and predict risk of progression to late amd," *Trans. Vis. Sci. & Tech.*, vol. 9, no. 2, pp. 25–25, 2020.



- [5] M.A. Zapata, D. Royo-Fibla, O. Font, J.I. Vela, I. Marcantonio, E.U. Moya-Sánchez, A. Sánchez-Pérez, D. Garcia-Gasulla, U. Cortés, E. Ayguadé, and J. Labarta, "Artificial intelligence to identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma," *Clin. Ophthalmol.*, pp. 419–429, 2020.
- [6] A. Serener and S. Serte, "Dry and wet age-related macular degeneration classification using oct images and deep learning," in *2019 EBBT*, 2019, pp. 1–4.
- [7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*. 2012, vol. 25, Curran Associates, Inc.
- [8] Goldy Verma, "Efficient retinal disease detection with resnet-18: A deep learning approach," in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, 2024, pp. 812–816.
- [9] S. M. Azoad Ahnaf, G. M. Atiqur Rahaman, and Sajib Saha, "Understanding cnn's decision making on oct-based amd detection," in *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)*, 2021, pp. 1–4.
- [10] Stephanie J. Chiu, Joseph A. Izatt, Rachelle V. O'Connell, Katrina P. Winter, Cynthia A. Toth, and Sina Farsiu, "Validated automatic segmentation of amd pathology including drusen and geographic atrophy in sd-oct images," *Investigative Ophthalmology I& Visual Science*, vol. 53, no. 1, pp. 53–61, 01 2012.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.
- [12] Shuying Liu and Weihong Deng, "Very deep convolutional neural network based image classification using small training sample size," in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 2015, pp. 730–734.
- [13] Ramprasaath R. Selvaraju, Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, and Dhruv Batra, "Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization," *CoRR*, vol. abs/1610.02391, 2016.
- [14] Preet Singh, Taniya Hasija, and K. R. Ramkumar, "Optimizing ocular disease detection: A deep learning approach with vgg19 and inceptionv3," in *2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, 2024, pp. 1–6.
- [15] Yin jianing, "Odir - dataset," 2022.
- [16] Ramyea Ravindaran, Kasthuri N, Preethi S, Adithya B, Gunaranjan Sp, Dharanidharan K, and Aravinth S, "Performance analysis of a vgg based deep learning model for classification of weeds and crops," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, pp. 1–5.
- [17] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the inception architecture for computer vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.
- [18] Athraa Kh. Al-Zihairy and A.E. Abdelkareem, "Optimizing yolov8-clis: A step towards smarter edge environments," in *2024 1st International Conference on Emerging Technologies for Dependable Internet of Things (ICETI)*, 2024, pp. 1–6.
- [19] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015.
- [20] Huihui Fang, Fei Li, Huazhu Fu, Xu Sun, Xingxing Cao, Fengbin Lin, Jaemin Son, Sunho Kim, Gwenole Quelled, Sarah Matta, Sharath M. Shankaranarayana, Yi-Ting Chen, Chuen-Heng Wang, Nisarg A. Shah, Chia-Yen Lee, Chih-Chung Hsu, Hai Xie, Baiying Lei, Ujjwal Baid, Shubham Innani, Kang Dang, Wenxiu Shi, Ravi Kamble, Nitin Singhal, Ching-Wei Wang, Shih-Chang Lo, José Ignacio Orlando, Hrvoje Bogunović, Xiulan Zhang, and Yanwu Xu, "Adam challenge: Detecting age-related macular degeneration from fundus images," *IEEE Trans. Med. Imag.*, vol. 41, no. 10, pp. 2828–2847, 2022.
- [21] Erdem Okur and Mehmet Turkan, "Patch enhancement for melanoma detection with bag of visual words," in *2022 Medical Technologies Congress (TIPTEKNO)*, 2022, pp. 1–4.