3DCRNet: Deep learning-driven Classification of Intermediate Age-related Macular Degeneration Using Optical Coherence Tomography in Resource-Constrained Devices

Federico Carrara[a]

[a]*Politecnico di Milano,*

Abstract

Age-related macular degeneration (AMD) stands as the primary cause of irreversible blindness in the elderly population. In-vivo spectral-domain optical coherence tomography (SD-OCT) is a well-established method for obtaining high-quality cross-sectional scans of the retina, commonly used to identify AMD due to its detailed anatomical characterization. While distinguishing eyes with advanced stages of intermediate AMD is generally straightforward, manually reviewing numerous OCT scans proves to be excessively time-consuming. In this research, we introduce 3D Resource-Constrained Net (3DRCNet), a 3D Convolutional Neural Network designed to automatically identify the presence of AMD-related conditions in SD-OCT retinal images. The distinguishing feature of 3DRCNet compared to competitor models is its optimized architecture, enabling it to operate on devices with limited resources (≤6GB GPU). This characteristic makes it particularly valuable in medical environments where extensive computing resources may not be readily available. Despite these inherent constraints, the model demonstrated a 78% accuracy and a noteworthy 86% recall in classifying AMD in an SD-OCT test set. Consequently, we can assert that 3DRCNet strikes an excellent balance between computational demands and performance, making it a promising tool in the realm of AMD diagnosis.

Keywords: Optical Coherence Tomography, Age-related macular degeneration, Deep Learning, Convolutional Neural Networks

1. Introduction

Age-related Macular Degeneration (AMD) is a prevalent and debilitating eye condition, representing a leading cause of irreversible blindness among the elderly population. Early and accurate detection of AMD is crucial for timely intervention and effective management of the disease.

In recent years, Spectral-Domain Optical Coherence Tomography (SD-OCT) has emerged as a pivotal imaging modality for examining the retina in exquisite detail. SD-OCT provides high-resolution cross-sectional images of the retina, enabling clinicians to visualize structural changes associated with AMD, such as the presence of drusen of Geographic Atrophy (GA) (Fig. 1).

The detailed anatomical characterization offered by SD-OCT scans allows for the identification of subtle abnormalities, aiding in the early diagnosis and monitoring of AMD progression. However, the manual analysis of these scans is labor-intensive and time-consuming. Hence, the availability of efficient and accurate automated methods for the detection of AMD from SD-OCT retinal scans is certainly helpful. Deep Learning (DL), specifically Convolutional Neural Networks (CNNs), has demonstrated remarkable success in various medical image analysis tasks. In particular, the integration of Deep Learning techniques with SD-OCT scans has proven a promising avenue for automating AMD detection (Wang and Wang, 2019; Li et al., 2019; Cheng et al., 2019; Gómez-Valverde et al., 2019; Li et al., 2023).

Incorporating advanced technologies in certain medical settings with constrained resources proves challenging. For instance, underfunded healthcare facilities in developing countries frequently face a deficit in technological devices, particularly for resource-intensive applications such as DL-driven diagnostics. Consequently, defining a fair balance between computing power demand and the quality of outcomes becomes imperative.

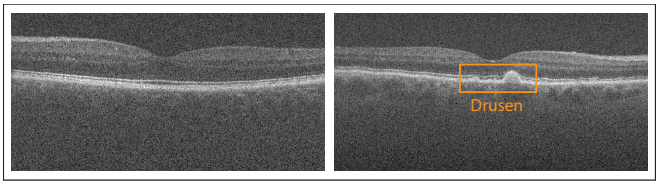

Figure 1: SD-OCT cross-sectional scan associated with a control patient (left) and presenting an AMD-related condition in the form of drusen (right).

In this paper, we present *3D Resource-Constrained Net* (3DRCNet), a 3D-CNN architecture designed to automatically identify the presence of AMD-related conditions in 3D SD-OCT retinal images. This model is specifically designed to optimally respond to the trade-off between computation requirements and performance. Indeed, on the one hand, it is capable of processing the rich 3D information inherent in SD-OCT scans, providing a comprehensive understanding of retinal structures. On the other hand, 3DRCNet has the advantage of being deployable on resource-constrained devices, making it practical for integration into diverse medical settings where extensive computational resources may be limited.

3DRCNet is an open-source project implemented in

Python. The code is available at `https://github.com/federico-carrara/3DRCNET`.

2. Methods

2.1. Dataset

For this study, we employed the dataset from the A2A SD-OCT Study (Leuschen et al., 2013). The data are open-source and available at `https://people.duke.edu/~sf59/RPEDC_Ophth_2013_dataset.htm` (Farsiu et al., 2014). The patients with AMD included in the study had to satisfy the following criteria: between 50 and 85 years of age, exhibiting intermediate AMD with large drusen ($> 125mm$) in both eyes or large drusen in 1 eligible eye and advanced AMD in the fellow eye, with no history of vitreoretinal surgery or ophthalmologic disease that might affect acuity in either eye. Age-appropriate control subjects were enrolled with the same inclusion criteria except that they must have had no evidence of macular drusen or AMD in either eye at the baseline visit or in the follow-up years. For this study, we limited the dataset to 160 SD-OCT retinal scans, of which 80 are associated with eyes affected by AMD and 80 control samples. Each SD-OCT scan is composed of 100 stacked cross-sectional scans of the retina centered at the fovea. In turn, each scan has size 512×1000. In the following, we will consider each SD-OCT scan as a 3D image of size $512 \times 1000 \times 100$ (resp. $H \times W \times D$).

2.2. Data Pre-processing

As previously mentioned, SD-OCT is a non-invasive imaging technique, highly effective in providing high-resolution images of retinal structures. However, SD-OCT scans are susceptible to noise due to various factors, such as optical imperfections, scattering of light within the eye, involuntary eye movements during the scan process, and variations in tissue reflectivity. As a result, SD-OCT scans commonly exhibit a salt-and-pepper noise pattern (Fig. 2). The presence of such noisy signals emphasizes the need for advanced image processing techniques to enhance the diagnostic utility of these images.
In our study, we employed the following image pre-processing steps:

1. *Total Variation* (TV) *Denoising* as proposed by *Chambolle* (Chambolle et al., 2009).
2. *Adaptive Histogram Equalization* (AHE) (Pizer et al., 1987).

TV denoising seeks to minimize the total variation norm, calculated as the L1 norm of the image gradient, while maintaining proximity to the original image. In this instance, we chose to utilize TV denoising to mitigate the impact of the salt-and-pepper noise pattern. Conversely, AHE functions as a contrast enhancement method with the specific aim of accentuating drusen structures in the deeper layers of the retina.

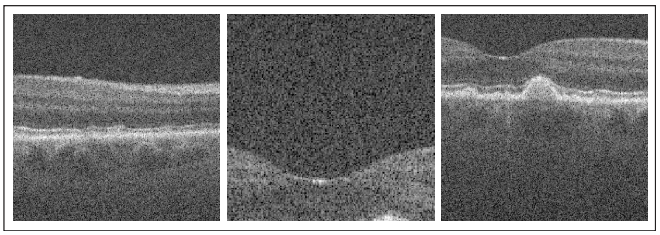

Figure 2: Salt-and-pepper noise pattern in SD-OCT cross-sectional scans.

2.3. Model

3DRCNet is a 3D CNN that operates on 3D SD-OCT images. The choice for 3D convolutions on the complete volume, as opposed to 2D convolutions on isolated slices, stems from the ability of 3D convolution to capture inter-voxel relationships throughout the entire volume. This approach facilitates the recognition of patterns extending across multiple slices, enhancing the model's ability to understand spatial dependencies. It is important to note, however, that the utilization of a 3D CNN comes with the trade-off of heightened GPU memory usage. The demand for 3D convolutional filters and the storage of 3D volumes poses greater challenges compared to the requirements for processing 2D slices. Therefore, to guarantee a fair compromise between computing resources and performance, a meticulous optimization of the 3D CNN model becomes essential. To minimize GPU memory usage during the training phase, we opted to segment the $512 \times 1000 \times 100$ SD-OCT scans into $512 \times 100 \times 100$ patches (Fig. 3). The decision to crop the volumes exclusively along the width dimension was made to avoid generating patches devoid of meaningful signals. Moreover, we designed for 3DRCNet a rather shallow architecture, with three convolutional blocks and two fully-connected (FC) layers. Each convolutional block comprehends 3D convolutional layers (3x3x3 filter, "same" padding) with an increasing number of kernels (8, 16, and 32), followed by batch normalization, leaky ReLU activation, and max pooling with $(4, 2, 2)$ kernel. The first FC layer is characterized by a hidden size of 32, while the second provides the input to a sigmoid function that outputs a single logit value. As a result, the 3DRCNet model contains less than 1.2M trainable parameters. To provide a term of comparison, the 3D-CNN versions of *VGG-16*, *ResNet-18* and *DenseNet-121* contain, respectively, 179.1M, 33.3M, and 11.4M trainable parameters.

3. Results

3.1. Training details

From the original dataset of 160 SD-OCT scans (80 per class) training, validation, and test sets were randomly selected with a 60/20/20 proportion and preserving the class balance.
The 3DRCNet model was trained for 20 epochs on one NVIDIA RTX2060 GPU (6 GB). Network parameters were learned by minimizing the BCE Loss using the Adam optimizer with an initial learning rate equal to 0.001. It is meaningful to notice that patches are created at the beginning of the training phase. To further reduce the computational requirements we

limited the batch size to 8 patches. Evaluation during training was performed by monitoring validation loss and accuracy.

Accuracy	*AUC*	*Precision*	*Recall*	*F1-score*
0.78	0.78	0.74	0.86	0.80

Table 1: Results on 3DRCNet evaluation on SD-OCT test set.

3.2. Model performance

The performance of 3DRCNet for the task of AMD recognition in SD-OCT scans of the retina was evaluated on the hold-out test set using the following metrics: *Accuracy*, *AUC*, *Precision*, *Recall*, and *F1-score*. The results are reported in Table 1. It is worth noticing the high Recall (0.86) attained by the model. Recall holds particular significance as a metric in the realm of diagnostic devices, representing the percentage of correctly identified positive instances within all the ones present in the dataset. In our framework, a high Recall value signifies the 3DRCNet model's effectiveness in identifying a substantial proportion of retinal scans indicating AMD. This capability is particularly valuable for our diagnostic tool, as it prioritizes efficiently identifying positive cases, crucial in a context where early disease detection is paramount—even if it means retrieving some additional false positives.

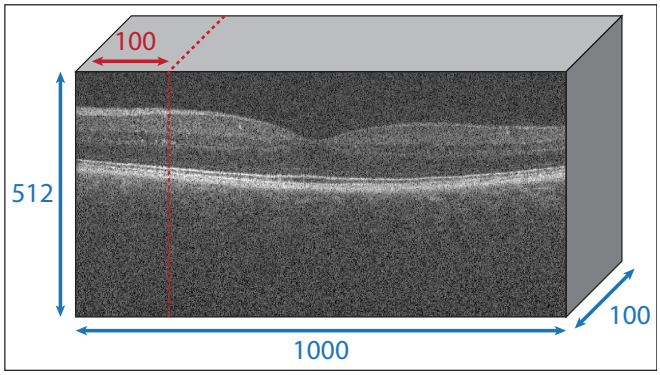

Figure 3: Schematic representation of the patch extraction process from full-size SD-OCT scans. This tiling approach ensure the inclusion of meaningful information in the resulting patches.

4. Discussion

In this paper, we introduced the 3D Resource-Constrained Net (3DRCNet), a 3D Convolutional Neural Network designed to automatically detect conditions associated with Age-related Macular Degeneration in spectral-domain optical coherence tomography images of the retina. The unique contribution of this study lies in the remarkable balance achieved by 3DRCNet between computational efficiency and performance. By leveraging 3D convolution to extract meaningful volumetric information from SD-OCT scans, the model consistently delivers accurate classification results. Notably, its high Recall makes 3DRCNet a reliable tool for detecting early-stage AMD conditions. Simultaneously, the adoption of smaller patches for training and the relatively shallow architecture significantly reduces the overall computational burden.

It's crucial to emphasize that the current iteration of 3DRCNet should be viewed as a prototype paving the way for a more refined model. Its existing performance falls short of justifying its standalone use as a diagnostic tool. Nonetheless, owing to its commendably high Recall, the current version of 3DRCNet can be seen as a facilitator, offering preliminary suggestions to aid medical decision-making.
iterations of this study, our objective is to enhance the performance of 3DRCNet through a meticulous refinement of the training routine and a comprehensive exploration of hyperparameter tuning. Additionally, to enhance its applicability in medical settings, we aspire to enhance interpretability by incorporating the HiResCAM approach (Draelos and Carin, 2020). This technique allows for the generation of heatmaps, highlighting regions with higher weights that played a more significant role in the final decision-making process.

References

Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T., 2009. An introduction to total variation for image analysis.

Cheng, S., Wang, L., Du, A., 2019. Histopathological image retrieval based on asymmetric residual hash and dna coding. IEEE Access 7, 101388–101400. doi:10.1109/ACCESS.2019.2930177.

Draelos, R.L., Carin, L., 2020. Use hirescam instead of grad-cam for faithful explanations of convolutional neural networks URL: http://arxiv.org/abs/2011.08891.

Farsiu, S., Chiu, S.J., O'Connell, R.V., Folgar, F.A., Yuan, E., Izatt, J.A., Toth, C.A., 2014. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography, pp. 162–172. doi:10.1016/j.ophtha.2013.07.013.

Gómez-Valverde, J.J., Antón, A., Fatti, G., Liefers, B., Herranz, A., Santos, A., Sánchez, C.I., Ledesma-Carbayo, M.J., 2019. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. Biomedical Optics Express 10, 892. doi:10.1364/BOE.10.000892.

Leuschen, J.N., Schuman, S.G., Winter, K.P., McCall, M.N., Wong, W.T., Chew, E.Y., Hwang, T., Srivastava, S., Sarin, N., Clemons, T., Harrington, M., Toth, C.A., 2013. Spectral-domain optical coherence tomography characteristics of intermediate age-related macular degeneration. Ophthalmology 120, 140–150. doi:10.1016/j.ophtha.2012.07.004.

Li, A., Winebrake, J.P., Kovacs, K., 2023. Facilitating deep learning through preprocessing of optical coherence tomography images. BMC Ophthalmology 23, 158. doi:10.1186/s12886-023-02916-2.

Li, F., Chen, H., Liu, Z., dian Zhang, X., shan Jiang, M., zheng Wu, Z., qian Zhou, K., 2019. Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomed. Opt. Express 10, 6204–6226. URL: https://opg.optica.org/boe/abstract.cfm?URI=boe-10-12-6204, doi:10.1364/BOE.10.006204.

Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K., 1987. Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 39, 355–368. doi:10.1016/S0734-189X(87)80186-X.

Wang, D., Wang, L., 2019. On oct image classification via deep learning. IEEE Photonics Journal 11, 1–14. doi:10.1109/JPHOT.2019.2934484.