Research Article Vol. 16, No. 11 / 1 Nov 2025 / Biomedical Optics Express 4380
Biomedical Optics EXPRESS

AI-aided segmentation of four types of drusen in volumetric OCT

YUKUN GUO,[1,2] TRISTAN T. HORMEL,[1] AN-LUN WU,[1,3] MIN GAO,[1,2] THOMAS S. HWANG,[1] STEVEN T. BAILEY,[1] AND YALI JIA[1,2,*]

[1]Casey Eye Institute, Oregon Health & Science University, Portland, OR 97239, USA
[2]Department of Biomedical Engineering, Oregon Health & Science University, Portland, OR 97239, USA
[3]Department of Ophthalmology, Mackay Memorial Hospital, Hsinchu 300044, Taiwan
[*]jiaya@ohsu.edu

Abstract: Drusen are a hallmark biomarker of age-related macular degeneration (AMD), with their size, number, and morphology (type) closely linked to disease severity and progression. Accurate segmentation and classification of drusen from optical coherence tomography (OCT) images are essential for objective AMD assessment and monitoring. In this work, we present a deep learning framework that combines a convolutional neural network for automated drusen segmentation with a dedicated classification module to distinguish four clinically relevant, distinct drusen types based on segmentation output. We evaluated our approach on a comprehensive dataset and achieved a mean Dice score of 0.74 ± 0.21 for voxel-wise segmentation accuracy and a critical success index of 0.69 ± 0.24 for drusen count accuracy. This method demonstrates substantial improvements in the quantitative drusen analysis and offers a promising tool for enhanced AMD diagnosis and tracking of disease progression.

© 2025 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Introduction

Age-related macular degeneration (AMD) is one of the most prevalent causes of irreversible vision loss in the aging population, currently affecting over 200 million individuals worldwide [1]. A defining feature of early and intermediate AMD is the presence of drusen - extracellular deposits that form between the retinal pigment epithelium (RPE) and Bruch's membrane. Traditionally, drusen have been identified and assessed using color fundus photography (CFP), which provides a two-dimensional view of the retina. Although CFP has been integral to clinical grading systems, its relatively low resolution and lack of depth information limit its ability to accurately characterize drusen morphology and quantity, or to reliably differentiate drusen from other retinal pathologies.

Optical coherence tomography (OCT), a non-invasive imaging modality that provides high-resolution cross-sectional images of the retina, addresses many of CFP's limitations. Recent studies by Kim et al. [2] reported that the drusen diameter measured on CFP is significantly correlated with those on OCT. Similarly, Cheung et al. [3] reported that measurement of drusen volume in OCT correlated closely with conventional CFP-derived AMD severity and individual AMD features. Compared to CFP, the research from Jain et al. [4] and Khanifar et al. [5] demonstrated that OCT offers better visualization of drusen morphology and can measure their anatomic distribution, enabling more accurate diagnosis and monitoring of AMD progression.

Segmenting drusen from OCT images remain challenging due to substantial variability in drusen type, size, and shape. Manual segmentation is not only time-consuming but also subjective. To overcome these limitations, various automated segmentation methods have been developed. Our group previously proposed an algorithm that leverages intensity and contrast features derived from *en face* OCT projections of slabs located above Bruch's membrane [6]. While effective, this approach relies on accurate segmentation of Bruch's membrane, which is not always reliable.

#566705 https://doi.org/10.1364/BOE.566705
Journal © 2025 Received 1 May 2025; revised 22 Jul 2025; accepted 17 Sep 2025; published 10 Oct 2025

Biomedical Optics EXPRESS

Other groups have explored other approaches. Chen et al. applied intensity thresholding and morphological operations on cross-sectional OCT [7]. Similarly, Khalid et al. segmented drusen using retinal layer boundaries from cross-sectional OCT [8]. However, these methods are vulnerable to segmentation errors and often require manual parameter tuning, limiting their robustness and scalability.

Contemporary deep-learning-based methods have demonstrated significant performance advantages in medical image segmentation tasks [9–15]. Applied to OCT images, deep learning has shown notable success in retinal layer segmentation [16–18], retinal fluid segmentation [19–21], and disease classification [22,23]. Several groups have also applied deep learning specifically for drusen segmentation in OCT images. For instance, Wang et al. proposed a deep learning model incorporating a multi-scale transformer and global attention mechanism [24], which achieved higher Dice score compared to prior methods. Zadeh et al. introduced a U-net like convolutional neural network (CNN) [25] to segment drusen on OCT and Asgari et al. proposed a U-net with spatial pyramid pooling [26]. For the 3D segmentation on OCT data, Goyanse et al. adopted nnU-net [27] to perform 3D segmentation of drusen on OCT data [28]. However, these methods treat drusen as a single class and do not differentiate between drusen types, even though drusen types as well as size and number are strongly associated with AMD progression [29]. Recent studies have explored multi-class segmentation in retinal OCT analysis. Rasti et al. proposed RetiFluidNet, a self-adaptive, multi-attention network for segmenting various fluid types in OCT scans [30]. Similarly, Ni et al. introduced a 3D segmentation framework using channel–spatial attention and multi-scale feature aggregation for retinal lesion segmentation [31]. While these methods demonstrate the potential of multi-class segmentation, they focus on fluid and lesion types rather than drusen subtypes. In contrast, our work addresses this gap by explicitly segmenting and differentiating drusen types, which are key indicators of AMD progression.

In this study, we develop a dual-decoder convolutional neural network to analyze drusen in AMD. Additionally, we introduce a classification algorithm to categorize four types of drusen-pseudodrusen, hard drusen, soft drusen, and drusenoid pigment epithelial detachment (drusenoid PED)- based on the segmentation outputs from the deep learning model. Our approach offers three key innovations: 1) a shared encoder and dual-decoder architecture which reduces computational cost; 2) a novel classification algorithm that leverages segmentation outputs to distinguish between four clinically relevant drusen types; and 3) a comprehensive hyperparameter search process that optimizes and validates the effectiveness of our model.

2. Methods

2.1. Data acquisition

The study was approved by the Institutional Review Board of the Oregon Health & Science University, Portland, and adhered to the tenets of the Declaration of Helsinki. Informed consent was obtained from all participants. The Volumetric OCT data were acquired over the central macular 6×6-mm region using a 120-kHz OCT commercial OCT system (SOLIX; Optovue/Visionix, Inc.). Two repeated B-scans were taken at each 512 raster positions, and each B-scan consists of 512 A-lines. The structural OCT was generated by averaging the two repeated B-scans.

2.2. Convolutional neural network design

In this study, we propose a CNN for drusen segmentation featuring a shared encoder and dual-decoder architecture to improve the segmentation of both the RPE-drusen complex (RPEDC) and drusenoid volume in OCT images (Fig. 1). We designed the shared encoder with convolutional layers and progressive down-sampling to capture multi-scale features, while adding short-cut connections between the encoder and decoder layers to prevent vanishing gradients. The dual

Biomedical Optics EXPRESS

decoders allow us to separately segment the RPEDC and drusenoid volume, utilizing residual blocks [32] to enhance feature representation. We incorporated down-sampling residual blocks to reduce the computational complexity, and identity residual blocks to retain spatial details. By concatenating feature maps from the encoder to the decoders and using up-sampling layers, we ensure accurate segmentation of the target regions. The input to the CNN is a series of cross-sectional OCT images, and the output is a single segmentation image for the cross-sectional OCT image positioned in the middle of the series.

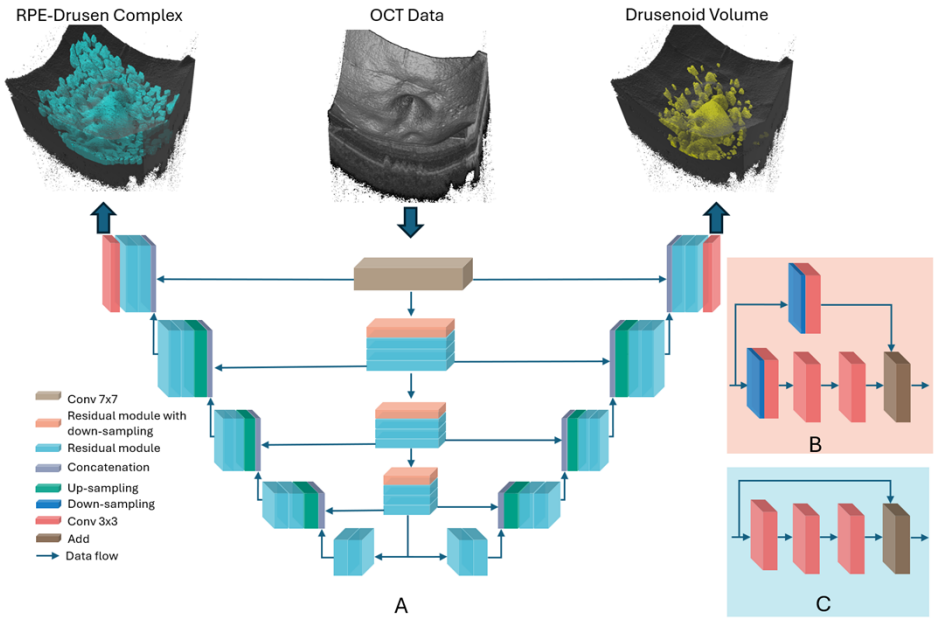


Fig. 1. The architecture of our proposed convolutional neural network for drusen segmentation in optical coherence tomography (OCT) images. A. The proposed dual-encoder CNN. The network consists of a shared encoder (central pathway) and two distinct decoders that specialize in segmenting the retinal pigment epithelium-drusen complex and drusenoid volume. Skip connections connected the encoder and decoders to preserve spatial resolution. The encoder processes the input OCT data and extracts hierarchical features using residual modules with down-sampling and passes these features through skip connections to the decoders. Each decoder up-samples the features and reconstructs the target segmentation. B. Residual module with down-sampling. C. Identity residual module.

2.3. Dataset preprocessing

To enhance the quality of the input images, we preprocess the OCT data by averaging adjacent pairs of images, a technique that effectively reduces noise and improves contrast. Following this, the images undergo normalization to achieve zero-mean and unit variance, ensuring consistent data representation for the CNN model. The ground truth for segmenting the RPEDC and drusenoid volume is generated through manual annotation of the OCT images using in-house developed software [33]. The annotation process was performed by two experienced graders (Y.G. and M.G.) and consisted of two components: the RPEDC region and the drusenoid volume. For RPEDC annotation, the graders first utilized a retinal layer segmentation algorithm [34] to identify the upper boundary of the ellipsoid zone and the Bruch's membrane. A thickness map was then computed as the point-wise difference between these two boundaries. To obtain an initial estimate of the RPEDC region, a threshold based on the mean thickness value was applied.

Biomedical Optics EXPRESS

Subsequently, the graders manually reviewed and refined the annotations on each B-scan. A similar procedure was followed for the drusenoid volume. In this case, the thickness map was generated using the lower boundary of the RPE and the Bruch's membrane to capture the relevant drusenoid regions. Any discrepancies between the two graders were resolved through discussion and consensus to ensure consistency and accuracy.

2.4. Training settings

To address the challenge of class imbalance, we used a composite loss function that combines categorical cross-entropy and the Dice coefficient, aiming to balance both pixel-wise classification accuracy and overlap-based segmentation quality. This combined loss function is defined as

$$L = -\frac{1}{N}\sum_{i=1}^{N}\{Y_i \cdot \log\hat{Y}_i + (1 - Y_i) \cdot \log(1 - \hat{Y}_i)\} + \left(1 - \sum_{i=1}^{N}\frac{2\hat{Y}_i \cdot Y_i}{\hat{Y}_i + Y_i + \varepsilon}\right) \tag{1}$$

where N denotes the number of pixels in the input image. Y_i represents the ground truth label for pixel i, $\hat{Y}_i$ is the corresponding prediction made by the model. ε is a small constant added to prevent division by zero. By combining these two loss components, the model is guided to focus not only on individual pixel classification but also on the overall structure of the segmented regions. The Adam optimizer with an initial learning rate of 0.001 was used to train the model, ensuring effective and stable convergence.

2.5. Hyperparameter search

We conducted a comprehensive analysis to evaluate the impact of key hyperparameters on the segmentation performance of our CNN model. The hyperparameters assessed included the number of downsampling steps in encoder, the number of filter kernels in the convolutional layers, and the number of input data channels. (1) Downsampling in the encoder reduces the spatial resolution of feature maps, enhancing the CNN's robustness to noise, improving computational efficiency, and lowering the risk of overfitting. However, excessive downsampling can result in information loss, limiting the model's ability to extract fine-grained details. Therefore, identifying the optimal number of downsampling steps is crucial for maximizing model performance. (2) The number of filter kernels in the convolutional layers influences the model's feature extraction capability. While increasing the number of kernels enhances feature extraction, it also raises computational complexity. Thus, determining the optimal number of filter kernels requires balancing model performance and computational efficiency. (3) The number of input data channels plays a significant role in enabling the model to learn features from adjacent cross-sectional OCT images. However, too many channels might introduce irrelevant information, hindering the model's ability to capture useful features. Optimizing these three hyperparameters is essential to achieve the best performance from the model. Several hyperparameter search methods have been proposed for deep learning models [35–37].

In this study, since our goal is to find the optimal combination of just three hyperparameters, we used a simplified grid search approach, in which one hyperparameter varies at a time while keeping the others fixed. Model performance was rigorously evaluated using five-fold cross-validation across the entire dataset, with segmentation accuracy quantified by sensitivity [Eq. (2)], specificity [Eq. (3)], and the Dice coefficient score [Eq. (4)]:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{3}$$

$$\text{Dice - score} = \frac{2 \times \text{TP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{4}$$

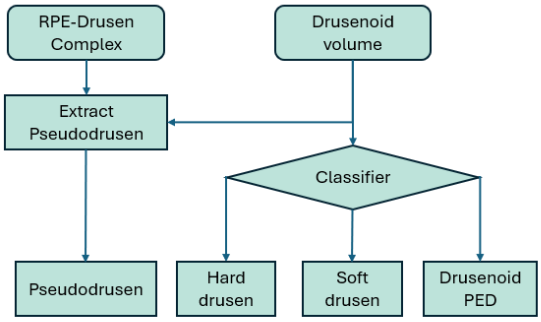
Biomedical Optics EXPRESS

In these equations, TP (True Positive) represents the number of pixels correctly classified as belonging to the drusen region. FN (False Negative) denotes pixels incorrectly classified as part of the background. FP (False Positive) refers to pixels incorrectly predicted as part of the drusen region when they belong to the background. TN (True Negative) indicates pixels correctly predicted as part of the background. This methodology enabled us to identify the optimal hyperparameter configurations for our model. We checked the range of downsampling steps with [1,2,3,4,5]. The searching range of filter kernel numbers in [8,16,24,32,48, 64, 96, 128, 192, 256], for each round, four continuous numbers were selected for evaluation. The number of input data channels is ranged in [1,3,5,7,9,11,13,15].

2.6. Classification algorithm for differentiating four types of drusen

The classification algorithm for differentiating four types of drusen is derived from the segmentation outputs of the deep learning model, leveraging the RPEDC and the drusenoid volume segmentations. The procedure begins by labeling regions within the RPEDC and creating an intersection volume between RPEDC and the drusenoid volume. Unique region IDs are extracted from the labeled regions, followed by extracting the drusenoid volume from RPEDC. The classification algorithm then labels the regions in the drusenoid volume, performing a maximum projection along the axial direction to facilitate the differentiation process. Major and minor axis lengths are calculated for each region, allowing the classification algorithm to classify drusen into three categories based on their volumetric properties: hard drusen, which have the smallest size; soft drusen; and drusenoid PED, which has the largest volume. Additionally, while the RPEDC spans the area between the Ellipsoid zone and Bruch's membrane, pseudodrusen are clinically defined as deposits located above the RPE. To align with this definition, we excluded the RPE slab when extracting pseudodrusen volumes from RPEDC.

A flowchart depicting the classification algorithm is shown in Fig. 2(A). The classification criterion for drusen types follows the definition of different types of drusen [38–40]. Pseudodrusen are determined as regions located within the RPEDC but outside the drusenoid volume. For the hard drusen, we defined the major axis length of the drusenoid volume smaller than 63-μm, and the drusenoid PED is defined as the major axis length larger than 350-μm. The soft drusen defined the major axis length between 63-μm and 350-μm. The detailed steps of the algorithm are outlined in Table 1.



A

Criterion for drusen classification

Drusen Types	Size Range
Pseudodrusen	-
Hard drusen	Major Axis < 63 um
Soft drusen	Major Axis ∈ [63 um ~ 350 um]
Drusenoid PED	Major Axis > 350 um

B

Fig. 2. Flow chart and classification criterion of the proposed classification algorithm for differentiating four types of drusen. A. The flow chart of the proposed classification algorithm. Firstly, the pseudodrusen is extracted from the RPE-Drusen Complex by removing the region that co-occurred with drusenoid volume. Then, the hard drusen, soft drusen, and drusenoid PED are classified from drusenoid volume based on the criterion in (Tab. 1). B. Criterion for drusen classification.

Biomedical Optics EXPRESS

3. Results

3.1. Study population

We collected data from a total of 120 eyes (80 with drusen and 40 healthy controls) from 120 subjects enrolled in a clinical AMD study.

Table 1. Pseudo code of the Algorithm for Differentiating Four Types of Drusen[a]

```
Algorithm 1. Differentiating Drusen Types
Procedure Diff-DRUSEN( RPEDC, DV )
    RPEDC_L ← Label Regions( RPEDC )
    I ← RPEDC & DV
    I_L ← RPEDC_L
    I_L (not I) = 0
    IDs ← Unique( I_L )
    Pseudodrusen ← RPEDC_L
    for i in IDs
        Pseudodrusen ( Pseudodrusen equals i ) = 0
    end for
    DV_L ← Label Regions (DV)
    DV_L2D ← Maximum Projection along z axis( DV_L )
    IDs ← Unique( DV_L2D )
    MajAxisLens← Get Region Proprieties( DV_L2D )
    PED ← 0, SoftDrusen ← 0, HardDrusen ← 0,
    for i, l in ( IDs.length, MajAxisLens)
        if l less or equals HardDrusen _Size
            HardDrusen ← HardDrusen + ( DV_L equals i )
        elseif l > DrusenoidPED_Size
            Drusenoid_PED ← Drusenoid_PED + ( DV_L equals i )
        else
            SoftDrusen ← SoftDrusen + ( DV_L equals i )
        end if
    end for
End Procedure
```

RPEDC: The segmentation result of RPE-Drusen complex from proposed model.

RPEDC_L: Labeled volumetric data of RPE-Drusen complex, each component in RPE-Drusen complex has a unique value for distinguishing with others.

DV: The segmentation result of Drusenoid Volume from proposed model.

DV _L: Labeled volumetric data of Drusenoid Volume, each component Drusenoid Volume has a unique value for distinguishing with others.

DV_L2D: Two-dimensional image obtained by projecting the 3D Drusenoid Volume along the depth axis.

MajAxisLens: The major axis length of each region in DV_L2D.

[a]RPEDC: The segmentation result of RPE-Drusen complex from proposed model.
[a]RPEDC_L: Labeled volumetric data of RPE-Drusen complex, each component in RPE-Drusen complex has a unique value for distinguishing with others.
[a]DV: The segmentation result of Drusenoid Volume from proposed model.
[a]DV _L: Labeled volumetric data of Drusenoid Volume, each component Drusenoid Volume has a unique value for distinguishing with others.
[a]DV_L2D: Two-dimensional image obtained by projecting the 3D Drusenoid Volume along the depth axis.
[a]MajAxisLens: The major axis length of each region in DV_L2D.

3.2. Hyperparameter search results

The results of the hyperparameter search are summarized in Table 2. Based on these results, we selected the optimal combination of three hyperparameters: downsampling step number of 3, kernel sizes of [24,32,48, 64], and an input data channel number of 11. This combination

Biomedical Optics EXPRESS

was chosen for the subsequent performance evaluations. As shown in the "Number of Kernels" row in Table 2, the residual modules at each downsampling level use a consistent number of convolutional filters corresponding to their respective stages. Each residual module follows the standard ResNet architecture [32], comprising two convolutional layers, each followed by batch normalization and ReLU activation. We have expanded our evaluation by including a comprehensive comparison with several state-of-the-art segmentation models including ResUNet [41], UNet++ [42], Swin Unet [43], Attention Unet [44], TransUNet [45]. The results are shown in Table S1. We also conducted an experiment to investigate the effect of varying weight combinations on the final performance of our proposed model. Different weight ratios between Dice and cross-entropy losses ranging from [1, 0] to [0, 1]. The results, reported in Table S2, show the Dice coefficients (mean ± standard deviation) for both RPEDC and drusenoid volume segmentation.

Table 2. Agreement (in Voxels, Mean ± Standard Deviation) Between Automated Detection and Manual Delineation of Drusen (5-fold cross-validation)

Network Settings		RPE-Drusen Complex			Dusenoid volume		
		Sensitivity	Specificity	Dice core	Sensitivity	Specificity	Dice core
Number of down sampling steps[a]	**1**	0.70 ± 0.02	1.00 ± 0.00	0.71 ± 0.03	0.71 ± 0.02	1.00 ± 0.00	0.63 ± 0.06
	2	0.77 ± 0.03	1.00 ± 0.00	0.77 ± 0.03	0.78 ± 0.02	1.00 ± 0.00	0.75 ± 0.01
	3	**0.79 ± 0.02**	**1.00 ± 0.00**	**0.78 ± 0.02**	**0.80 ± 0.02**	**1.00 ± 0.00**	**0.77 ± 0.02**
	4	0.77 ± 0.03	1.00 ± 0.00	0.77 ± 0.04	0.78 ± 0.02	1.00 ± 0.00	0.76 ± 0.04
	5	0.76 ± 0.02	1.00 ± 0.00	0.77 ± 0.01	0.76 ± 0.01	1.00 ± 0.00	0.75 ± 0.02
Number of kernel[b]	**[8,16,2,32]**	0.70 ± 0.00	1.00 ± 0.00	0.69 ± 0.01	0.63 ± 0.01	1.00 ± 0.00	0.57 ± 0.01
	[16,24,32,48]	0.72 ± 0.03	1.00 ± 0.00	0.71 ± 0.03	0.67 ± 0.03	1.00 ± 0.00	**0.64 ± 0.02**
	[24,32,48, 64]	0.74 ± 0.01	**1.00 ± 0.00**	**0.73 ± 0.02**	**0.70 ± 0.01**	**1.00 ± 0.00**	0.63 ± 0.07
	[32,48, 64, 96]	**0.75 ± 0.01**	1.00 ± 0.00	0.72 ± 0.01	0.62 ± 0.04	1.00 ± 0.00	0.58 ± 0.03
	[48, 64, 96, 128]	0.70 ± 0.00	1.00 ± 0.00	0.68 ± 0.02	0.67 ± 0.02	1.00 ± 0.00	0.62 ± 0.00
	[64, 96, 128, 192]	0.71 ± 0.03	1.00 ± 0.00	0.69 ± 0.03	0.63 ± 0.02	1.00 ± 0.00	0.58 ± 0.03
	[96, 128, 192, 256]	0.68 ± 0.06	1.00 ± 0.00	0.68 ± 0.06	0.64 ± 0.04	1.00 ± 0.00	0.61 ± 0.03
Number of input data channel[c]	**1**	0.76 ± 0.01	1.00 ± 0.00	0.76 ± 0.01	0.80 ± 0.00	1.00 ± 0.00	0.76 ± 0.01
	3	0.76 ± 0.02	1.00 ± 0.00	0.76 ± 0.02	0.79 ± 0.01	1.00 ± 0.00	0.77 ± 0.04
	5	0.77 ± 0.03	1.00 ± 0.00	0.76 ± 0.04	0.78 ± 0.04	1.00 ± 0.00	0.75 ± 0.02
	7	0.78 ± 0.01	1.00 ± 0.00	0.77 ± 0.02	0.76 ± 0.01	1.00 ± 0.00	0.74 ± 0.02
	9	0.81 ± 0.02	1.00 ± 0.00	0.80 ± 0.02	0.82 ± 0.01	1.00 ± 0.00	0.79 ± 0.03
	11	**0.81 ± 0.02**	**1.00 ± 0.00**	**0.81 ± 0.02**	**0.83 ± 0.02**	**1.00 ± 0.00**	**0.83 ± 0.02**
	13	0.77 ± 0.04	1.00 ± 0.00	0.78 ± 0.02	0.82 ± 0.02	1.00 ± 0.00	0.78 ± 0.00
	15	0.78 ± 0.02	1.00 ± 0.00	0.78 ± 0.01	0.78 ± 0.03	1.00 ± 0.00	0.75 ± 0.04

[a]Kernel number = [24,32,48, 64], Input data channel number = 5
[b]Down sampling step number = 3, Input data channel number = 5
[c]Down sampling step number = 3, Kernel number = [24,32,48, 64]
[c]Bold: the best performance in each parameter setting.

3.3. *Four types of drusen in a single eye*

Using the hyperparameters identified in Section 3.1, we retrained our deep learning model and evaluated its performance on the test dataset. Figure 3 illustrates the segmentation and classification results on an AMD eye. The RPEDC (Fig. 3(A)) and Drusenoid volume (Fig. 3(B)) were segmented using our deep learning model, while four types of drusen were classified

Biomedical Optics EXPRESS

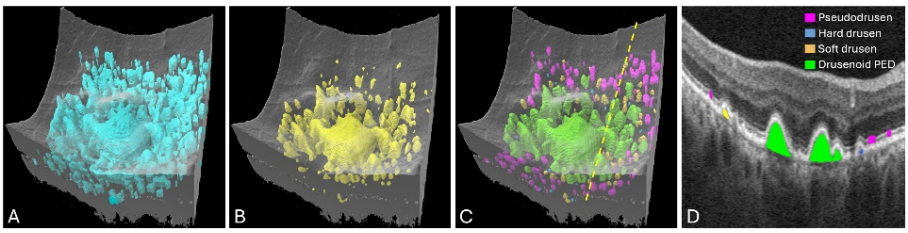


Fig. 3. Illustration of segmentation of four types of drusen in an AMD eye. A. Segmentation of the RPE-Drusen-complex. B. Segmentation of drusenoid volume. C. Classification of four drusen types (pseudodrusen, hard drusen, soft drusen, and drusenoid PED) using the proposed algorithm. D. Cross-sectional OCT image overlaid with the classified drusen types. The yellow dotted line in C indicates the location of the cross-section shown in D.

through the proposed classification algorithm (Fig. 3(C)). The cross-sectional OCT image (Fig. 3(D)) overlaid with the four classified types demonstrates high consistency between the segmented drusen regions and their actual anatomical locations, confirming the model's accuracy in capturing diverse drusen types.

3.4. Segmentation accuracy quantification

The segmentation of the RPEDC and drusenoid volume demonstrates a high level of consistency with the manually delineated ground truth [Fig. 4(A), (B)]. Additionally, the differentiating algorithm effectively distinguished the four drusen types, with segmentation results closely matching the ground truth [Fig. 4(C)-(F)].

We further evaluated the performance of the proposed deep learning model in segmenting four distinct drusen types. Given that both the size and the count of drusen are critical factors in AMD severity grading (Table 3), the model's performance was assessed at the voxel level and by counting the number of segmented drusen. These metrics provide a comprehensive evaluation of the model's ability to accurately quantify drusen, which is essential for determining AMD severity and tracking disease progression.

Table 3. Detection accuracy across different drusen types (in voxels and counts, Mean ± Standard Deviation)

	Pseudodrusen	Hard Drusen	Soft Drusen	Drusenoid PED	Overall
Drusen Volume detection Accuracy					
Sensitivity	0.68 ± 0.07	0.59 ± 0.37	0.60 ± 0.14	0.83 ± 0.12	0.71 ± 0.21
Specificity	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
Dice score	0.76 ± 0.06	0.57 ± 0.38	0.65 ± 0.12	0.85 ± 0.09	0.74 ± 0.21
Drusen Counts detection Accuracy					
Sensitivity	0.99 ± 0.06	0.68 ± 0.22	0.69 ± 0.14	0.94 ± 0.14	0.81 ± 0.22
Precision[a]	0.69 ± 0.12	0.66 ± 0.13	0.85 ± 0.10	0.94 ± 0.05	0.81 ± 0.21
CSI[b]	0.69 ± 0.13	0.49 ± 0.12	0.62 ± 0.16	0.89 ± 0.15	0.69 ± 0.24

[a]Precision = TP/(TP + FP)
[b]Critical success index, CSI = TP/(TP + FP + FN)

Biomedical Optics EXPRESS

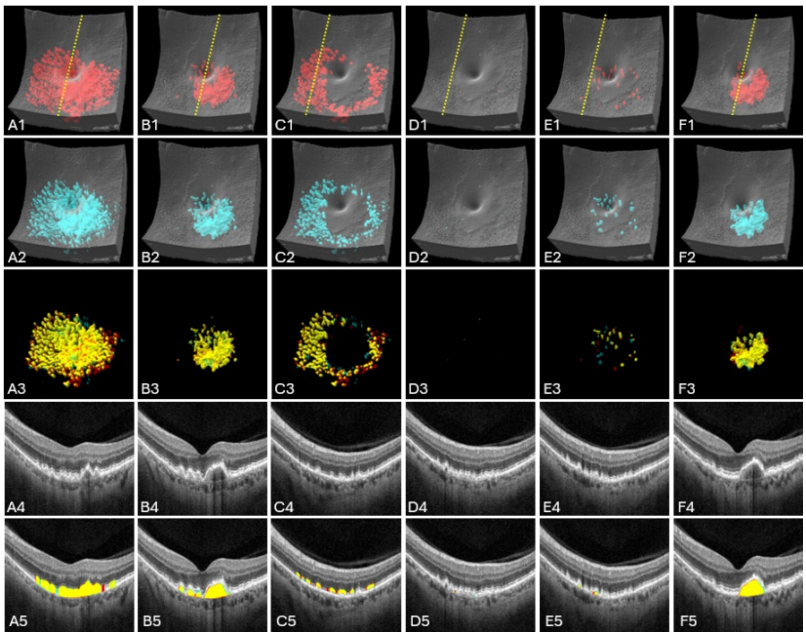


Fig. 4. Segmentation results of the proposed deep learning model for four subtypes of drusen on OCT data. Columns show the RPE-Drusen complex (A), the drusenoid volume (B), pseudodrusen (C), hard drusen (D), soft drusen (E), and drusenoid pigment epithelial detachment (PED; F). Row 1: Ground truth drusen segmentation. The yellow line indicates the position of the cross-sectional OCT slice shown in the fourth and fifth rows. Row 2: Segmentation outputs generated by the proposed deep learning model. Row 3: Visualization of the overlap between the ground truth and the model's predictions. Yellow regions indicate areas where the model's predictions coincide with the ground truth. Red regions are ground truth and blue regions are model's predictions. Row 4: Cross-sectional OCT images taken at the positions indicated by the arrows in A1-F1. Row 5: Overlays of the cross-sectional OCT images and the corresponding drusen segmentation (blue: false positive, red: false negative, yellow: true positive).

4. Discussion

We developed a deep learning model integrated with a newly designed classification algorithm to segment and identify four distinct types of drusen in volumetric OCT data. Our results demonstrate that this model, in conjunction with the classification algorithm, can accurately differentiate between four drusen types based on the segmentation output.

In our study we classified four types of drusen—pseudodrusen, hard drusen, soft drusen, and drusenoid PED—covering the majority of drusen types found in AMD eyes. To enable our deep learning model to accurately detect and differentiate between these types, we designed two distinct decoders: one for segmenting the RPEDC and another for drusenoid volume. A shared encoder was implemented to reduce computational complexity, with skipping connections between the encoder and decoders to preserve spatial resolution. Additionally, a residual module was embedded in both the encoder and decoders to enhance extraction of features and mitigate the vanishing gradient problem. A comparative experiment was conducted, and the results demonstrated the effectiveness of the proposed dual-decoder architecture in accurately segmenting both RPEDC and drusenoid volume (Table S1).

The hyperparameter search experiments (Table 2) indicate that the choice of hyperparameters is crucial for the performance of a deep learning model. Notably, the input image channel number

Biomedical Optics EXPRESS

plays a vital role in enabling the network to capture adjacent features of the drusen region, resulting in more accurate and smoother drusen boundaries. When comparing the best-performing model, which utilized 11 input channels, to a model with a single input channel, the 11-channel model demonstrated a significant improvement in Dice score. This is particularly important because the adjacent cross-sectional OCT images provide three-dimensional information that enhances the model's ability to accurately analyze and segment drusen structures. Additionally, we conducted an experiment to assess the impact of different weight combinations in the combined loss function. As shown in Table S2, balanced weightings between Dice and cross-entropy losses resulted in more stable and higher segmentation performance compared to other combinations.

The proposed classification algorithm for differentiating the four types of drusen also shows promising performance (Table 3). As shown in Table 3, our method has slightly lower sensitivity in detecting hard drusen, primarily due to their small size. Consequently, our deep learning model struggles to fully segment all hard drusen. For soft drusen, the model's performance is also slightly lower than for pseudodrusen and drusenoid PED. This is likely due to the small drusenoid volume within soft drusen and the limited spatial resolution of OCT, which can cause the model to miss them. The metrics for drusenoid PED were higher than other drusen types. This is likely because the drusenoid PED is the most prominent among all types, making it easier to segment, and its larger size facilitates classification by the proposed algorithm.

Our approach provides several opportunities for future research. With its ability to precisely quantify the volume size and locate different types of drusen, we can measure the distribution of each type in AMD eyes. Additionally, It also supports longitudinal tracking of each type of drusen, offering valuable insight into disease progression. Furthermore, it opens possibilities to explore relationships between various drusen types, paving the way for additional valuable insights.

While our method demonstrates high accuracy in segmenting and differentiating four types of drusen, there are some limitations to this study. Given the complexity of drusen nomenclature, we focused on summarizing only four types, which may not capture all variations observed in AMD eyes. This limitation could be addressed in future research by establishing clear naming conventions and categorical definitions in clinical practice [46–48]. Another limitation of our classification algorithm is its reliance on the length of the semi-major axis from the projected volumetric drusen segmentation as the primary criterion for classification, while overlooking other potentially relevant features. This choice was made to align with the traditional definition of drusen size in color fundus photography. This restriction arises from the inherent variability among drusen types. To overcome this limitation, further investigation into the characteristics of different drusen types is necessary, such as analyzing OCT signal intensity and the internal patterns within the drusenoid volume.

5. Conclusions

In summary, we developed a deep learning model coupled with a classification algorithm specifically for segmenting and distinguishing four types of drusen in volumetric OCT scans. Through hyperparameter tuning and five-fold cross-validation, our model achieved strong performance in drusen segmentation. With its capability to differentiate among the four types of drusen, our approach offers a more comprehensive quantification of drusen and may improve AMD severity grading.

Funding. National Institutes of Health (R01 EY036429, R01 EY035410, R01 EY024544, R01 EY027833, R01 EY031394, R43EY036781, P30 EY010572, T32 EY023211, UL1TR002369); Jennie P. Weeks Endowed Fund; Malcolm M. Marquis, MD Endowed Fund for Innovation; Edward N. and Della L. Thome Memorial Foundation; Unrestricted Departmental Funding Grant; Research to Prevent Blindness (Dr. H. James and Carole Free Catalyst Award); Bright Focus Foundation (G2020168, M20230081).

Biomedical Optics EXPRESS

Disclosures. Yukun Guo: Optovue/Visionix (P), Genentech/Roche (P, R); Ifocus Imaging (P), Tristan T. Hormel: Ifocus Imaging (I), Steven T. Baily: Optovue/Visionix (F), Yali Jia: Optovue/Visionix (P, R), Genentech/Roche (P, R, F), Ifocus Imaging (I, P), Optos (P), Boeringer Ingelheim (C), Kugler (R)

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request. See Supplement 1 for supporting content.

Supplemental document. See Supplement 1 for supporting content.

References

1. A. Abdelsalam, L. Del Priore, and M. A. Zarbin, "Drusen in Age-Related Macular Degeneration," Surv. Ophthalmol. **44**(1), 1–29 (1999).
2. D. Y. Kim, J. Loo, S. Farsiu, *et al.*, "Comparison of Single Drusen Size on Color Fundus Photography and Spectral-Domain Optical Coherence Tomography," Retina **41**(8), 1715–1722 (2021).
3. C. M. G. Cheung, Y. Shi, Y. C. Tham, *et al.*, "Correlation of Color Fundus Photograph Grading with Risks of Early Age-related Macular Degeneration by using Automated OCT-derived Drusen Measurements," Sci. Rep. **8**(1), 12937 (2018).
4. N. Jain, S. Farsiu, A. A. Khanifar, *et al.*, "Quantitative Comparison of Drusen Segmented on SD-OCT versus Drusen Delineated on Color Fundus Photographs," Invest. Ophthalmol. Vis. Sci. **51**(10), 4875 (2010).
5. A. A. Khanifar, A. F. Koreishi, J. A. Izatt, *et al.*, "Drusen Ultrastructure Imaging with Spectral Domain Optical Coherence Tomography in Age-related Macular Degeneration," Ophthalmology **115**(11), 1883–1890.e1 (2008).
6. R. Zhao, A. Camino, J. Wang, *et al.*, "Automated drusen detection in dry age-related macular degeneration by multiple-depth, en face optical coherence tomography," Biomed. Opt. Express **8**(11), 5049 (2017).
7. Q. Chen, T. Leng, L. Zheng, *et al.*, "Automated drusen segmentation and quantification in SD-OCT images," Med. Image Anal. **17**(8), 1058–1072 (2013).
8. S. Khalid, M. U. Akram, T. Hassan, *et al.*, "Automated Segmentation and Quantification of Drusen in Fundus and Optical Coherence Tomography Images for Detection of ARMD," J Digit Imaging **31**(4), 464–476 (2018).
9. L. Cai, J. Gao, and D. Zhao, "A review of the application of deep learning in medical image classification and segmentation," Ann Transl Med **8**(11), 713 (2020).
10. K. Fritscher, P. Raudaschl, P. Zaffino, *et al.*, "Deep Neural Networks for Fast Segmentation of 3D Medical Images," in *SpringerK Fritscher, P Raudaschl, P Zaffino, MF Spadea, GC Sharp, R SchubertMedical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th, 2016•Springer* (Springer Verlag, 2016), 9901 LNCS, pp. 158–165.
11. S. M. Waldstein, W. D. Vogl, H. Bogunovic, *et al.*, "Characterization of Drusen and Hyperreflective Foci as Biomarkers for Disease Progression in Age-Related Macular Degeneration Using Artificial Intelligence in Optical Coherence Tomography," JAMA Ophthalmol **138**(7), 740–747 (2020).
12. Y. Guo, T. T. Hormel, S. Pi, *et al.*, "An end-to-end network for segmenting the vasculature of three retinal capillary plexuses from OCT angiographic volumes," Biomed. Opt. Express **12**(8), 4889 (2021).
13. Y. Guo, T. T. Hormel, H. Xiong, *et al.*, "Development and validation of a deep learning algorithm for distinguishing the nonperfusion area from signal reduction artifacts on OCT angiography," Biomed. Opt. Express **10**(7), 3257–3268 (2019).
14. Y. Guo, A. Camino, J. Wang, *et al.*, "MEDnet, a neural network for automated detection of avascular area in OCT angiography," Biomed. Opt. Express **9**(11), 5147 (2018).
15. M. Gao, Y. Guo, T. T. Hormel, *et al.*, "A Deep Learning Network for Classifying Arteries and Veins in Montaged Widefield OCT Angiograms," Ophthalmology Science **2**(2), 100149 (2022).
16. Q. Li, S. Li, Z. He, *et al.*, "DeepRetina: Layer Segmentation of Retina in OCT Images Using Deep Learning," Trans. Vis. Sci. Tech. **9**(2), 61 (2020).
17. Z. Mishra, A. Ganegoda, J. Selicha, *et al.*, "Automated Retinal Layer Segmentation Using Graph-based Algorithm Incorporating Deep-learning-derived Information," Sci. Rep. **10**(1), 9541 (2020).
18. P. Zang, S. S. Gao, T. S. Hwang, *et al.*, "Automated boundary detection of the optic disc and layer segmentation of the peripapillary retina in volumetric structural and angiographic optical coherence tomography," Biomed. Opt. Express **8**(3), 1306 (2017).
19. D. Lu, M. Heisler, S. Lee, *et al.*, "Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network," Med. Image Anal. **54**, 100–110 (2019).
20. M. X. Li, S. Q. Yu, W. Zhang, *et al.*, "Segmentation of retinal fluid based on deep learning: application of three-dimensional fully convolutional neural networks in optical coherence tomography images," Int. J. Ophthalmol., Engl. Ed. **12**(6), 1012–1020 (2019).
21. Y. Guo, T. T. Hormel, H. Xiong, *et al.*, "Automated Segmentation of Retinal Fluid Volumes From Structural and Angiographic Optical Coherence Tomography Using Deep Learning," Trans. Vis. Sci. Tech. **9**(2), 54 (2020).
22. C. S. Lee, D. M. Baughman, and A. Y. Lee, "Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images," Ophthalmol Retina **1**(4), 322–327 (2017).
23. P. Zang, T. T. Hormel, T. S. Hwang, *et al.*, "Deep-LearningeAided Diagnosis of Diabetic Retinopathy, Age-Related Macular Degeneration, and Glaucoma Based on Structural and Angiographic OCT," Ophthalmology Science **3**(1), 100245 (2023).

Biomedical Optics EXPRESS

24. M. Wang, W. Zhu, F. Shi, *et al.*, "MsTGANet: Automatic Drusen Segmentation From Retinal OCT Images," IEEE Trans. Med. Imaging **41**(2), 394–406 (2022).
25. S. Gorgi Zadeh, M. W. M. Wintergerst, V. Wiens, *et al.*, "CNNs Enable Accurate and Fast Segmentation of Drusen in Optical Coherence Tomography," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer Verlag, 2017), **10553 LNCS**, pp. 65–73.
26. R. Asgari, S. Waldstein, F. Schlanitz, *et al.*, "U-Net with Spatial Pyramid Pooling for Drusen Segmentation in Optical Coherence Tomography," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer, 2019), **11855 LNCS**, pp. 77–85.
27. F. Isensee, P. F. Jaeger, S. A. A. Kohl, *et al.*, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," Nat. Methods **18**(2), 203–211 (2021).
28. E. Goyanes, S. Leyva, P. Herrero, *et al.*, "Fully-automatic end-to-end approaches for 3D drusen segmentation in Optical Coherence Tomography images," Procedia Comput Sci **246**(C), 1100–1109 (2024).
29. K. N. Khan, O. A. Mahroo, R. S. Khan, *et al.*, "Differentiating drusen: Drusen and drusen-like appearances associated with ageing, age-related macular degeneration, inherited eye disease and other pathological processes," Prog. Retinal Eye Res. **53**, 70–106 (2016).
30. R. Rasti, A. Biglari, M. Rezapourian, *et al.*, "RetiFluidNet: A Self-Adaptive and Multi-Attention Deep Convolutional Network for Retinal OCT Fluid Segmentation," IEEE Trans. Med. Imaging **42**(5), 1413–1423 (2023).
31. G. Ni, K. Cao, X. Qin, *et al.*, "Advanced 3D retinal lesion segmentation using channel-spatial attention-guided multi-scale feature aggregation," Biomed. Opt. Express **16**(5), 2093–2110 (2025).
32. K. He, X. Zhang, S. Ren, *et al.*, "Deep Residual Learning for Image Recognition," openaccess.thecvf.com (2015).
33. M. Zhang, J. Wang, A. D. Pechauer, *et al.*, "Advanced image processing for optical coherence tomographic angiography of macular diseases," Biomed. Opt. Express **6**(12), 4661 (2015).
34. Y. Guo, A. Camino, M. Zhang, *et al.*, "Automated segmentation of retinal layer boundaries and capillary plexuses in wide-field optical coherence tomographic angiography," Biomed. Opt. Express **9**(9), 4429 (2018).
35. D. M. Belete and M. D. Huchaiah, "Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results," International Journal of Computers and Applications **44**(9), 875–886 (2022).
36. G. Luo, "A review of automatic selection methods for machine learning algorithms and hyper-parameter values," Network Modeling Analysis in Health Informatics and Bioinformatics **5**(1), 18 (2016).
37. J. Wu, X. Y. Chen, H. Zhang, *et al.*, "Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization," J. Electron. Sci. Technol. **17**(1), 26–40 (2019).
38. N. Manafi, A. Mahmoudi, M. Emamverdi, *et al.*, "Topographic analysis of local OCT biomarkers which predict progression to atrophy in age-related macular degeneration," Graefe's Arch. Clin. Exp. Ophthalmol. **262**(7), 2083–2091 (2024).
39. M. A. Gamulescu, H. Helbig, and J. Wachtlin, *Retinal Pigment Epithelial Detachment* (2017).
40. C. Cukras, E. Agrón, M. L. Klein, *et al.*, "Natural History of Drusenoid Pigment Epithelial Detachment in Age-Related Macular Degeneration: Age-Related Eye Disease Study Report No. 28," Ophthalmology **117**(3), 489–499 (2010).
41. F. I. Diakogiannis, F. Waldner, P. Caccetta, *et al.*, "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data," ISPRS Journal of Photogrammetry and Remote Sensing **162**, 94–114 (2020).
42. Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, *et al.*, "Unet++: A nested u-net architecture for medical image segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11045 LNCS, 3–11 (2018).
43. H. Cao, Y. Wang, J. Chen, *et al.*, "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation," *Lecture Notes in Computer Science 13803 LNCS*, 205–218 (2023).
44. O. Oktay, J. Schlemper, L. Le Folgoc, *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," (2018).
45. J. Chen, J. Mei, X. Li, *et al.*, "TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers," Med. Image Anal. **97**, 103280 (2024).
46. J. Lu, Y. Cheng, J. Li, *et al.*, "Automated segmentation and quantification of calcified drusen in 3D swept source OCT imaging," Biomed. Opt. Express **14**(3), 1292 (2023).
47. Z. Yehoshua, F. Wang, P. J. Rosenfeld, *et al.*, "Natural History of Drusen Morphology in Age-Related Macular Degeneration Using Spectral Domain Optical Coherence Tomography," Ophthalmology **118**(12), 2434–2441 (2011).
48. X. Zhang and S. Sivaprasad, "Drusen and pachydrusen: the definition, pathogenesis, and clinical significance," Eye **35**(1), 121–133 (2021).