applied sciences

MDPI

Article

Adaptive Deep Clustering Network for Retinal Blood Vessel and Foveal Avascular Zone Segmentation

Azaz Khan [1,2,3], Jinyi Hao [1,2,3], Zihao Dong [1,2,3,*] and Jinping Li [1,2,3,*]

[1] School of Information Science and Engineering, University of Jinan, Jinan 250022, China; engrazazkhan3132@gmail.com (A.K.); haojy2014@126.com (J.H.)
[2] Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China
[3] Shandong College and University Key Laboratory of Information Processing and Cognitive Computing in 13th Five-Year, Jinan 250022, China
* Correspondence: ise_dongzh@ujn.edu.cn (Z.D.); ise_lijp@ujn.edu.cn (J.L.)

Abstract: Optical coherence tomography angiography (OCTA) is a new non-invasive imaging technology that provides detailed visual information on retinal biomarkers, such as the retinal vessel (RV) and the foveal avascular zone (FAZ). Ophthalmologists use these biomarkers to detect various retinal diseases, including diabetic retinopathy (DR) and hypertensive retinopathy (HR). However, only limited study is available on the parallel segmentation of RV and FAZ, due to multi-scale vessel complexity, inhomogeneous image quality, and non-perfusion, leading to erroneous segmentation. In this paper, we proposed a new adaptive segmented deep clustering (ASDC) approach that reduces features and boosts clustering performance by combining a deep encoder–decoder network with K-means clustering. This approach involves segmenting the image into RV and FAZ parts using separate encoder–decoder models and then employing K-means clustering on each part separated by the encoder–decoder models to obtain the final refined segmentation. To deal with the inefficiency of the encoder–decoder network during the down-sampling phase, we used separate encoding and decoding for each task instead of combining them into a single task. In summary, our method can segment RV and FAZ in parallel by reducing computational complexity, obtaining more accurate interpretable results, and providing an adaptive approach for a wide range of OCTA biomarkers. Our approach achieved 96% accuracy and can adapt to other biomarkers, unlike current segmentation methods that rely on complex networks for a single biomarker.

Keywords: OCTA; RV; FAZ; segmentation; K-means

check for updates

Citation: Khan, A.; Hao, J.; Dong, Z.; Li, J. Adaptive Deep Clustering Network for Retinal Blood Vessel and Foveal Avascular Zone Segmentation. *Appl. Sci.* **2023**, *13*, 11259. https://doi.org/10.3390/app132011259

Academic Editor: Jan Egger

Received: 22 June 2023
Revised: 17 July 2023
Accepted: 21 July 2023
Published: 13 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Clear vision is essential for a fulfilling life, and retinal biomarkers are critical for identifying and diagnosing retinal disorders. These biomarkers offer valuable insights into human vision and can advance our understanding of retinal health [1]. RV and FAZ are crucial biomarkers that ophthalmologists use to diagnose and monitor retinal disease patients. Alterations in the RV and FAZ area are closely related to several diseases, including DR, HR, cardiovascular disease, brain stroke, acute renal injury, choroidal neovascularization (CNV), and arterial hypertension [2–4]. Therefore, early detection of changes in these regions is crucial for identifying these disorders. RV and FAZ segmentation allow ophthalmologists to discover retinal diseases early before they become severe. Hence, automatic segmentation of retinal biomarkers can aid in early disease detection.

OCTA is a state-of-the-art retinal imaging modality because it is non-invasive. It captures 3D volumetric retina data with dense scanning and high frequency, providing micron-level resolution of retinal blood vessels and the vascular system [5]. Visualizing retinal plexuses by projecting OCTA data from different retinal layers helps understand retinal vascular diseases. Recent studies have shown that OCTA scans may correctly

reveal pathologic alterations associated with retinal disease, leading many researchers to use OCTA image biomarkers to support clinical assessment in the early stages of retinal diseases. Some researchers have used OCTA to uncover pathologic alterations connected with retinal vein occlusion (RVO), including neovascularization, retinal nonperfusion, and FAZ abnormalities [6]. OCTA imaging visualizes blood flow in the retinal vasculature, which can indicate retinal health. Variations in the retinal vasculature detected through OCTA may indicate specific retinal disease [7], while precise segmentation helps the ophthalmologist diagnose RV-related diseases. Another essential structure for retinal health is the FAZ, a zone at the center of the fovea, also termed the fovea centralis, with the maximal cone concentration in the photoreceptor layer. Smaller FAZ individuals have excessive myopia because of poor foveal vascular circulation. At the same time, an enlarged FAZ may signify visual insight loss due to an ischemia situation in the fovea [8].

Consequently, RV and FAZ segmentation are crucial tasks in retinal disease investigation. These familiar structures in every retina exhibit spontaneous changes when disease occurs. However, segmentation is challenging due to limitations in current methods. These methods consist of deep learning (such as U-Net and encoder–decoder-based architecture) based approaches and traditional methods. Deep learning methods primarily consist of a single biomarker, such as FAZ segmentation [9–13] and RV segmentation [14,15]. However, the critical problem with RV segmentation is the thickness difference, particularly at the thin vessel ends as show in Figure 1, which have poor contrast, which remains the main barrier to RV segmentation. Another essential factor is similar structure interference, such as non-perfusion and FAZ, because their gray level values are similar and can cause difficulties in segmentation.

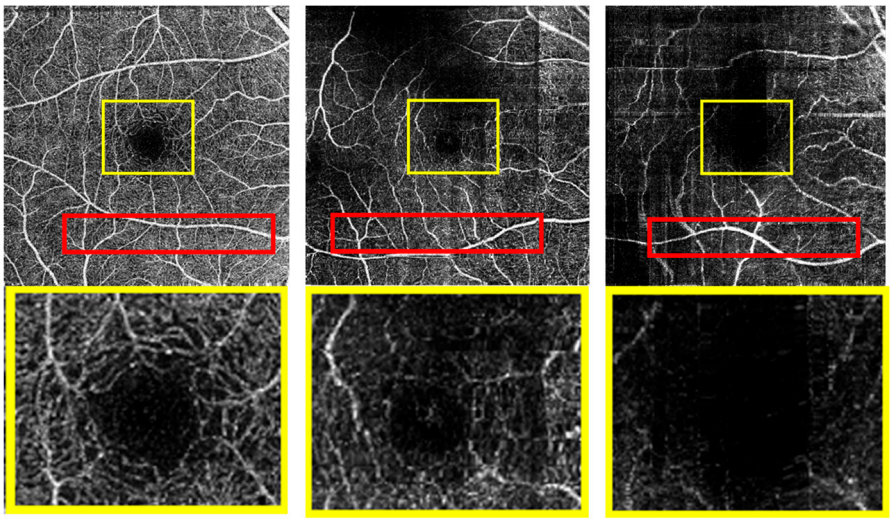


Figure 1. Variations in OCTA image quality: thickness difference, poor contrast, and inhomogeneity. The yellow square indicates thin vessels, while the red rectangle shows thick vessels. The images also demonstrate differences in contrast across varying scales of vessels, with thick vessels having good contrast and thin vessels exhibiting poor contrast.

Although the approach in [16] showed promising results, it suffered from non-uniform image quality, possible confusion between RV and FAZ regions, high computational complexity, and memory consumption. These challenges highlight the need for more efficient and accurate segmentation methods for OCTA images.

Segmentation based on traditional image processing methods such as RV [2,7] and FAZ segmentation [17,18] uses filter-based models, random field models, morphological operations (opening and closing), and locally adaptive thresholding. The poor segmentation performance illustrates the drawbacks of traditional techniques in feature extraction. These algorithms create assumptions about blood vessel shapes and flow signal distribution that

could not be accurate across all areas and disease states, which makes it difficult to segment multi-scale vessels and irregular FAZ precisely. The algorithm in [1] uses a frequency-based technique called a Gabor filter bank that only permits a specified frequency for vessel segmentation, which results in erroneous large vessel segmentation due to inhomogeneity in the OCTA image. Hence, simultaneous segmentation of RV and FAZ structures is a challenging task due to the intricate edges of the FAZ, non-uniform image quality, and the presence of multi-scale vessels and non-perfusion. While some studies have focused on extracting only one biomarker, multi-scale vessel segmentation is challenging when segmenting thin and thick vessels or terminal branches in noisy OCTA scans. The researchers' preference comes from the fact that OCTA image data are relatively hard to annotate, leading to a lack of standard available datasets with high-quality FAZ and RV annotations. Subsequently, various large OCTA image datasets were published, and the authors began investigating the extraction strategy of two biomarkers using the same method.

Determining boundaries is a significant challenge in segmenting large areas such as the FAZ. Improved boundary segmentation methods are needed to meet clinical requirements and address the complexities involved in this task. While encoder–decoder models can effectively capture unique features, they may produce unclear or inaccurate outputs, mainly when dealing with complex and variable features. The fundamental cause is that the encoder–decoder network is generally inefficient; that is, during the down-sampling process, the loss of image features such as boundary information steadily increases [19], and such data cannot be recovered during the up-sampling process in the decoding stage.

The extraction of biomarker features using traditional approaches is another issue. This study addresses this challenge by using separate encoding and decoding for each task instead of combining them into a single task. Combining the encoder–decoder models with K-means clustering can further refine and optimize the segmentation mask. By combining the RV and FAZ segmentation tasks into a single approach, we can achieve improved accuracy, fine-grained control over the training process, and scalability to other OCTA biomarkers, facilitating early retinal disease diagnosis by ophthalmologists.

To address the challenges associated with OCTA image data analysis, such as the extensive data annotation required for accurate segmentation, the loss of information in the down-sampling process for RV and FAZ, the difficulties in feature extraction using traditional methods, and the extensive training and inference times associated with complex network structures, we propose a new ASDC approach that reduces features and boosts clustering performance by combining a deep encoder–decoder with K-means clustering to segment the two essential structures (RV and FAZ) in OCTA-500 scans. This approach allows for the simultaneous segmentation of two important biomarkers, RV and FAZ, using parallel processing to reduce the time required for analysis. Combining the RV and FAZ segmentation tasks into a single approach aims to improve accuracy and scalability while reducing computational cost and time in OCTA image analysis. Our method has the potential to help ophthalmologists diagnose diseases early and manage retinal diseases more efficiently. To improve K-means clustering performance, we use an encoder–decoder network to reduce the number of features and prepare the data for clustering. K-means struggles to evaluate high-dimensional spaces, which is common in OCTA images with large numbers of pixels and intensity values. By reducing the dimensionality of the data, we can enhance the clustering effectiveness and accurately identify meaningful groupings.

The importance of K-means clustering preserves the quality of small and large vessels. As a result, in non-uniform OCTA scans, we attempt to acquire simultaneous RV and FAZ segmentation and precise multi-scaled vessel extraction while avoiding interference from similar structures such as non-perfusion and FAZ. Our dual segmentation network based on an encoder–decoder network with K-means treats the segmentation problem independently instead of treating it as a single class. The framework treats each problem as a single task to segment RV and FAZ and learn their relationship. Feature extraction through the encoder and feature expansion using a decoder followed by the K-means algorithm obtains the candidate for segmentation. Then, it converges to the desired target segmentation based

on cluster pixel similarity, which results in multi-scaled RV segmentation with a smooth gradient by removing noisy dots. A similar model is used for the FAZ candidate, and the final segmented output is obtained by merging the dual model results. Our approach provides parallel segmentation of RV and FAZ, reducing computational complexity and producing accurate, interpretable results for distinct parts. Additionally, the additivity can be applied to a variety of OCTA biomarkers. Figure 1, demonstrate the OCTA-500 ILM_OPL scans, highlighting the variation in vessel thickness across different vessels. The yellow square indicates thin vessels, while the red rectangle shows thick vessels. The images also demonstrate differences in contrast across varying scales of vessels, with thick vessels having good contrast and thin vessels exhibiting poor contrast. Moreover, the image quality varies across the scans, making it challenging to accurately segment the vessels and FAZ. Inhomogeneity in image quality further adds to the complexity of the biomarker segmentation.

In summary, the encoder–decoder models extract relevant features from the input image and produce a compressed feature representation of the object of interest. K-means clustering refines and optimizes the segmentation mask to produce a more precise representation of the biomarkers.

The proposed framework comprises two main phases: (1) In phase 1, the first step is to divide the target image (ground truth image) into two parts, FAZ and RV. Then, we convert the input image into grayscale format. These three images will be fed to phase 2 for training the encoder–decoder model. (2) In phase 2, we train two models independently which is Model 1 and Model 2, for RV and FAZ segmentation; each model is comprised of an encoder–decoder network followed by a K-means clustering algorithm applied to each part to obtain the desired segmented output. Finally, the two segmented parts, predicated image 1 corresponding to the FAZ part and predicated image 2 corresponding to the RV part, are concatenated to obtain the final segmented output.

Our primary contributions are listed below:

(1) Introducing the ASDC approach for the simultaneous segmentation of RV and FAZ biomarkers in an automatic and scalable manner, thus reducing the computational cost and time required for OCTA image analysis. This approach can potentially aid in early retinal disease investigation by providing more accurate and efficient segmentation.
(2) Conducting a comparative analysis of both unsupervised and supervised learning methods for OCTA image segmentation, providing insights into the strengths and weaknesses of different approaches, and helping to inform future research in this area.
(3) Demonstrating a thorough understanding of the current state-of-the-art in OCTA image segmentation research and introducing a scalable approach for processing OCTA data. This research contributes to developing more efficient and accurate methods for analyzing OCTA images, which can aid in diagnosing and managing retinal diseases.

The visual and calculated segmentation results demonstrate that our framework can obtain the desired segmentation results. Compared to current methods, the ASDC approach improves accuracy while reducing the computing time for OCTA image analysis.

Our study examines both supervised and unsupervised learning approaches for OCTA image segmentation and demonstrates improved segmentation accuracy for RV biomarker segmentation compared to previous methodologies. We validated our model on the OCTA-500 dataset using OCTA scans with a maximal projection between the internal limiting membrane (ILM) and the outer plexiform layer (OPL). This projection is derived from the inner retina's highest projection and clearly shows the retina's blood vessels and the structure of the FAZ. The ILM and OPL are commonly used as reference points when segmenting the FAZ and RV biomarkers.

The remainder of the paper is organized as follows: Section 2 explains the related work. Section 3 explains the method. Section 4 explains the training process. Section 5 explains the experimental results and discussion. Section 6 presents the conclusion.

2. Related Work

OCTA is a non-invasive novel modality; thus, a limited number of RV and FAZ segmentation studies are available. Most researchers have worked on color fundus images [20] for retinal blood vessel segmentation [21]. The retinal vasculature precise segmentation is a challenging task because of the low contrast of the image data, interference of homogenous structure in the retinal vasculature, and variation in vessel size; thus, automatic RV and FAZ segmentation is a crucial issue in retinal vasculature segmentation. This paper discusses mostly RV segmentation-related work based on color fundus images because of the few studies available on OCTA.

Most RV segmentation was conducted on color fundus images, and several studies have been reported [22–30] etc. This is because OCTA is a relatively rare, available technology, and some RV segmentation techniques have been reported, such as Eladawi et al. [2] who used a noise reduction filter using the generalized Gauss–Markov random field (GGMRF) model. Then they used a joint Markov–Gibbs random field (MGRF) model for RV segmentation. Breger A et al. [1], proposed a new frequency-based method using Gabor filter banks for RV and FAZ segmentation in OCTA en face images and compared the result within a built device called Cirrus HD-OCT with a difference of $1.58 \pm 1.08\%$ compared to the device for FAZ where a mean dice similarity coefficient of 0.89 was obtained.

Li A et al. [31], proposed an angiogenesis tracking algorithm using top hat filtering and optimally oriented flux (OOF) algorithms to segment the retinal capillary plexus. Y. Ma et al. [14] presented the retinal OCTA segmentation (ROSE) data set. They proposed a new split-based coarse-to-fine RV segmentation network able to detect thick and thin vessels, and this study mainly focused on RV segmentation without segmenting the FAZ part. Lu et al. [17] proposed an OCTA FAZ segmentation method based on a generalized gradient vector flow (GGVF) snake model to identify the FAZ area. Díaz et al. [18] proposed an image processing method using morphological operators to enhance the input image and edge detection technique to extract the FAZ. Guo et al. [11] proposed a deep learning model which is called MEDnet, a multi-scaled encoder–decoder neural network to segment the retinal capillary non-perfusion area which also includes the FAZ region in the retinal superficial vascular complex while the highest accuracy of the proposed method was 0.89 ± 0.04. Mingchao Li et al. [16] proposed a 3D-to-2D image projection network (IPN), IPN-V2, and IPN-V2+, which gave promising results in RV and FAZ segmentation. The author also developed the OCTA-500 dataset.

Despite their advancements, these methods remain limited and need vast training data and model parameters. Such as is seen in the case of [14,16]; extensive parameter training and a required 3D volumetric OCTA scan for training, suffering from non-uniform image quality resulting in confusion between RV and FAZ. Another essential factor is similar structure interference, such as non-perfusion and FAZ, because their greyscale values are similar and can cause difficulties in segmentation. Traditional image processing methods such as RV and FAZ segmentation [17,18] use filter-based models, random field models, morphological operation (opening and closing), and locally adaptive thresholding methods, which produce rough vessel boundaries caused by the non-uniform image quality and excessive background noise caused by threshold-based and filter-based segmentation algorithms. Some of the approaches mentioned above [7] are limited to segmenting a single biomarker, due to issues such as ambiguous FAZ with intrusive structures, unavoidable outliers caused by erroneous layer projection, and the inability to segment low-contrast FAZ from its surrounding region.

In summary, only a limited number of studies are currently available for the simultaneous segmentation of RV and FAZ. Due to the complexity of small and large retinal vessels, non-uniform image quality, and confusion between non-perfusion and FAZ areas, it is challenging to simultaneously segment the RV and FAZ biomarkers. Hence, we performed efficient parallel segmentation of the two biomarkers to successfully implement automatic simultaneous segmentation of RV and FAZ in one system with sufficient accuracy using OCTA 2D en face images.

3. Method

The primary motivation behind the proposed methods was to develop an accurate automatic RV and FAZ segmentation system to detect the RV and FAZ simultaneously by using a parallel approach. The experiment was performed on OCTA maximum projection between the internal limiting membrane (ILM) layer and the outer plexiform layer (OPL), clearly showing inner retinal morphology and FAZ shape. The study's second objective was to compare unsupervised (K-means clustering) and supervised learning algorithms for RV and FAZ segmentation and illustrate their effectiveness.

3.1. K-Means Clustering and Image Processing (Unsupervised Learning)

Segmentation of RV and FAZ in OCTA (ILM_OPL) 2D en face images using image processing and K-means (KM) clustering technique: K-means clustering (KM) is a well-known clustering approach for unsupervised learning applications. K-means can be applied to image segmentation via grouping the pixels by maximizing intersegment similarity. KM clustering is one of the widespread partitioning methods and is broadly used due to its ease, effectiveness, and comfort of implementation. Based on a distance measure, it partitions data samples into different clusters. Then it finds a partition such that the squared error between the centroid and the points in the cluster is minimized. Let $O = \{O_1, O_2, \ldots, O_n\}$, a set of n data specimens to be clustered into a set of K cluster $C = \{C_i, i = 1, \ldots, k\}$. The objective of KM is to reduce the sum of the squared error over all K clusters, defined as follows in Equation (1):

$$J(C) = \sum_{i=1}^{k} \sum_{O_l \in C_i} (O_l - \mathrm{Z_i})^2 \quad (1)$$

where i is the index of the cluster, ranging from 1 to k. l is the index of the data sample within a cluster, ranging from 1 to the number of data samples in the i-th cluster. C_i is the i-th cluster, $\mathrm{Z_i}$ the centroid for the i-th cluster, O_l is data samples from the i-th cluster, and k represents the total number of clusters [32].

Now we can provide the steps of implementation as follows:

(1) Convert the given image into a grayscale image and smooth the image with a 3×3 kernel Gaussian blur [33] filter. Gaussian blur reduces image noise by blurring it with a Gaussian function. It is a non-uniform low-pass filter that lowers image noise and noiseless details. Reshape the image array (3D) as a vector (1D) to apply to the cluster. The effectiveness of a Gaussian blur filter is to enhance vascular contrast and connectivity, which enhances the quality of the OCTA angiograms [34].

(2) We convert reshaping values to a floating point because each feature should be in a single column for K-means. Specify K-means convergence criteria, the point at which the algorithm stops iterating. Once the specified requirements are met, the algorithm terminates. The convergence criteria used in the segmentation function specifies that the K-means clustering algorithm should stop when either the desired accuracy or the maximum number of iterations is reached, whichever comes first. The maximum number of iterations is set to 10, and the desired accuracy is set to 1.0. This means that the algorithm will stop after 10 iterations or earlier if the desired accuracy is achieved. The convergence criterion is an important parameter in the K-means algorithm as it determines when the algorithm has converged to a stable solution and can be used to optimize the clustering process.

(3) Apply the K-means algorithm with K = 4 to divide the OCTA en face image into four clusters. After performing K-means clustering, we extract the compactness, labels, and centers to identify the regions. Finally, it returns a list of labels corresponding to each cluster and their centers (in this case, where they are located on the original image).

(4) We access the labels and convert the label array to a clustered image; every pixel represents its clustering label. Therefore, we can establish a list of the images in our clustered image, calculate the minimum and maximum values for each image,

then create an empty 3-dimensional array to record each pixel's color. Next, it loops through each image's pixels to assign color values.

(5) Specifically, the maximum pixel values are converted to 255, and the minimum pixel values are converted to 100. Other two-pixel values represent 0. Then, a smooth function is applied to remove noisy dots in the image. In summary, setting K = 4 in Step (5) allows the segmentation function to accurately capture the boundary between the foreground (RV and FAZ) and background, while considering only 3 different pixel values in the resulting image ensures that the RV and FAZ objects are not lost and that the segmentation is robust to any additional clusters that may have been created by the K-means clustering algorithm.

(6) Smooth function: The goal of this step is to remove noise in the image by averaging pixel values in a small window around each pixel and returning the image with a smooth gradient. By cycling through all pixels in the original image, check if each pixel's current value differs from the average of all values. If so, that pixel becomes black (0). After each pixel is processed, we get a smooth gradient image. Then we check how much a pixel differs from its average value over input pixels, which is set to 0.8. This ensures that no single pixel has more than 80% of its total area covered by black pixels, which would result in a very dark image, or only white pixels, which would result in an extremely bright image. Figure 2 shows the proposed K- means implementation.

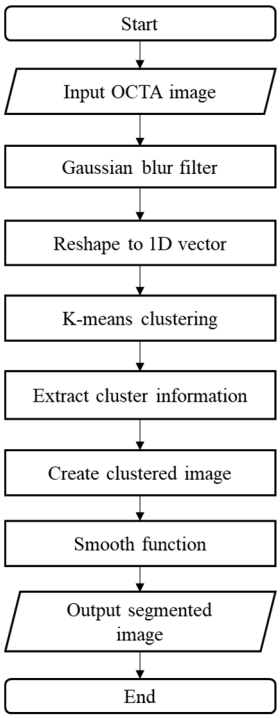


Figure 2. Flow chart of the proposed K-means implementation.

3.2. Encoder–Decoder Network Architecture (Supervised Learning)

We suggest a new deep convolutional neural network (DCNN) encoder–decoder architecture to segment the RV and FAZ area. The encoder is responsible for extracting a collection of features from the input image to produce encoded data, which are then used by the decoder to provide a segmented result. The encoder portion creates a low-resolution feature map of the input image using a generic convolutional neural network (CNN) architecture without completely connected layers. On the other hand, in the decoder part, a parallel architecture is used to replicate a high-resolution feature map [35]. Each encoder consists of various convolutional layers responsible for producing feature maps [36]. Then we used a batch normalization layer to enhance model generalization [37]. Then to intro-

duce nonlinearity, a rectified linear unit (ReLU) max(0,x) is applied to an efficient CNN activation function.

After that, a low-resolution feature map is obtained using several convolutional layers. Next we use a flattening transformation [38] to reduce the dimension of the feature vector obtained in the last layer and input it to the decoder part, which is then up-sampled using transposed convolution in the decoder part. The last layer in the encoder part is the flattening layer which reshapes the feature vector into a 1-dimensional vector to avoid loss of information [39].

The same structure is used in the decoder part, starting from an unflattened layer. The next layer is to perform transposed convolution [40] to up-sample the feature map, followed by ReLU activation function and batch normalization to obtain a high-dimensional feature representation. The final decoder output is fed to a sigmoid activation function which is the last layer to obtain the final segmented output. The model summary is shown in Tables 1 and 2.

Table 1. Encoder model detailed layer configuration.

Number	Layer (Type)	Input Shape	Parameters	Trainable Parameters
1	Conv2D	[1, 1, 400, 400]	60	60
2	ReLU	[1, 6, 200, 200]	0	0
3	Conv2D	[1, 6, 200, 200]	660	660
4	BatchNorm2D	[1, 12, 100, 100]	24	24
5	ReLU	[1, 12, 100, 100]	0	0
6	Conv2D	[1, 12, 100, 100]	2616	2616
7	ReLU	[1, 24, 98, 98]	0	0
8	Conv2D	[1, 24, 98, 98]	10,416	10,416
9	ReLU	[1, 48, 96, 96]	0	0
10	Conv2D	[1, 48, 96, 96]	41,568	41,568
11	ReLU	[1, 96, 94, 94]	0	0
12	Flatten	[1, 96, 94, 94]	0	0
Total parameters: 55,344; Trainable parameters: 55,344; Non-trainable parameters: 0				

Table 2. Decoder model detailed layer configuration.

Number	Layer (Type)	Input Shape	Parameters	Trainable Parameters
1	Unflatten	[1, 848256]	0	0
2	ConvTranspose2D	[1, 96, 94, 94]	41,520	41,520
3	ReLU	[1, 48, 96, 96]	0	0
4	ConvTranspose2D	[1, 48, 96, 96]	10,392	10,392
5	ReLU	[1, 24, 98, 98]	0	0
6	ConvTranspose2D	[1, 24, 98, 98]	2604	2604
7	BatchNorm2D	[1, 12, 100, 100]	24	24
8	ReLU	[1, 12, 100, 100]	0	0
9	ConvTranspose2D	[1, 12, 100, 100]	654	654
10	BatchNorm2D	[1, 6, 200, 200]	12	12
11	ReLU	[1, 6, 200, 200]	0	0
12	ConvTranspose2D	[1, 6, 200, 200]	55	55
Total parameters: 55,261; Trainable parameters: 55,261; Non-trainable parameters: 0				

3.3. Encoder–Decoder Architecture plus K-Means Clustering OCTA Segmentation Framework

Encoder–decoder plus K-means clustering-based OCTA segmentation framework illustrates the framework adopted in this study, which is comprised of two main phases. (1) In phase 1, the first step is to divide the target image (ground truth image) into two parts, namely (1) FAZ and (2) RV. Then we convert the input image into grayscale format; these three images will be fed to phase 2 for training the DCNN encoder–decoder model. (2) In phase 2, we trained two models independently: Model 1 and Model 2 for RV and FAZ segmentation. Each model is comprised of an encoder–decoder network followed

by a K-means clustering algorithm applied to each part to obtain the desired segmented output. Finally, the two segmented parts, predicated image 1 corresponding to the FAZ part and predicated image 2 corresponding to the RV part, are concatenated to obtain the final segmented output image of OCTA.

- Phase 1. In this phase, the target image is separated into RV and FAZ components as shown in Figure 3, and the test image is then fed into phase 2. In the process of preparing the data for training, we extract the regions of interest for the RV and FAZ classes by identifying the corresponding pixel values of 255 and 100, respectively, based on the histogram of the ground truth data.

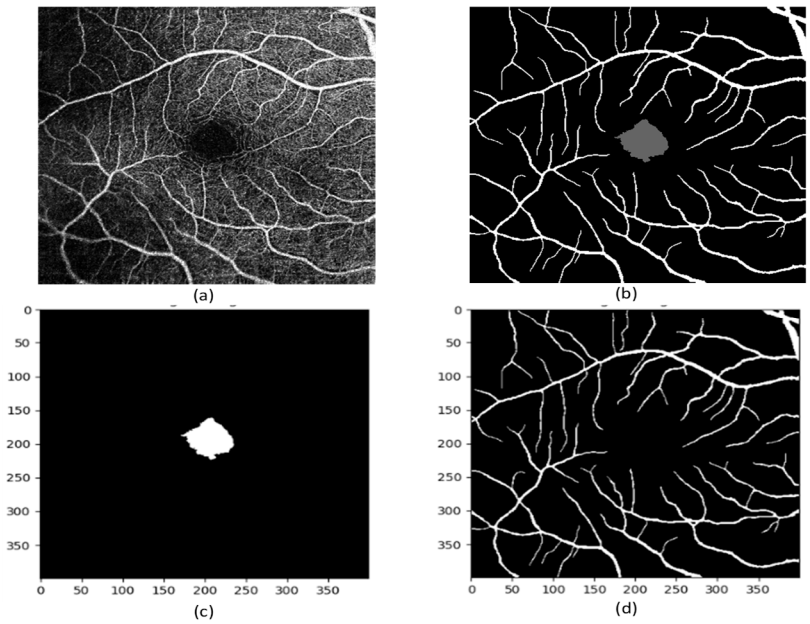


Figure 3. Phase 1. (**a**) OCTA-500 ILM_OPL input image; (**b**) OCTA-500 ILM_OPL ground truth (target image); (**c**) target image 1; (**d**) target image 2.

- Phase 2. This phase involves feeding Model 2 (the FAZ green part), the test input image, and the FAZ part from the previous phase to train the network and predict the FAZ part and RV, respectively. A concatenation operator is used to combine the two images after predicting the two images independently to produce the final segmented output. We present the encoder-decoder architecture for OCTA segmentation in Figure 4. This architecture is then used in combination with the K-means clustering algorithm in the Encoder-decoder plus K-means clustering-based OCTA segmentation framework, which is shown in Figure 5 and provides the final output. Figure 6 illustrates the internal structure of the two models, which is comprised of the encoder-decoder architecture presented in Figure 4, followed by the K-means clustering algorithm shown in Figure 2.The key importance of the K-means clustering algorithm is to refine the output and group the pixels which belong to the RV and FAZ areas in each task.

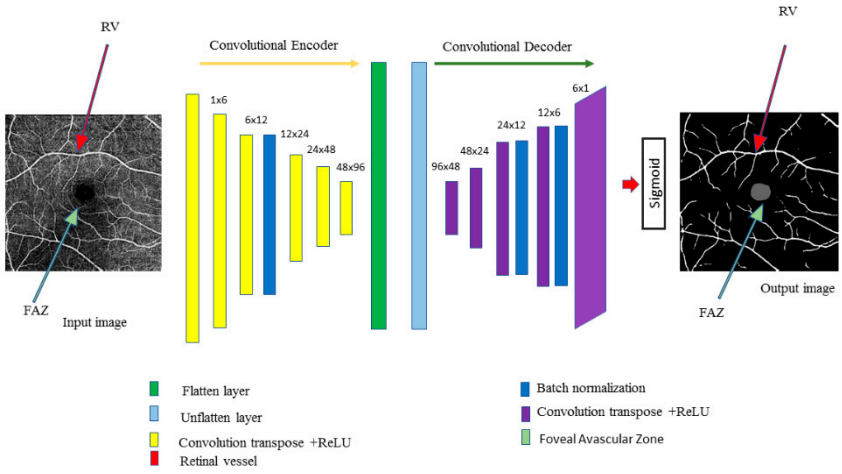


Figure 4. Overall model architecture for OCTA RV and FAZ segmentation.

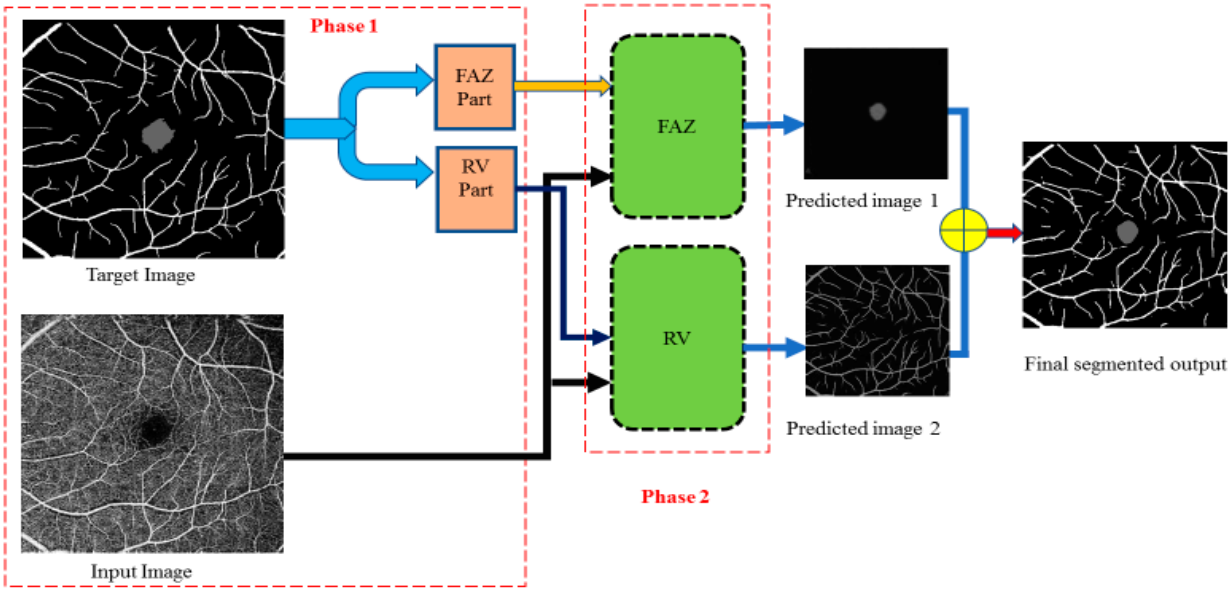


Figure 5. Encoder–decoder plus K-means clustering based OCTA segmentation framework.

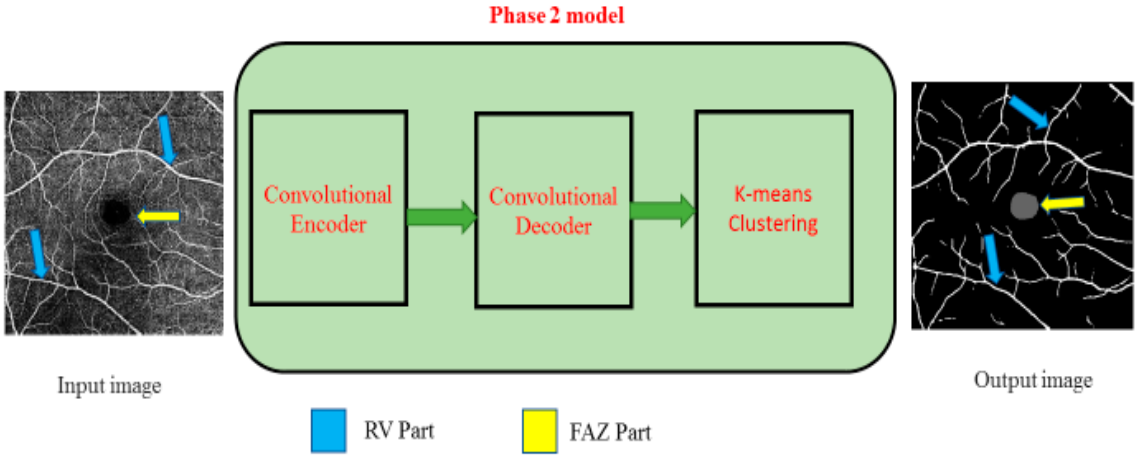
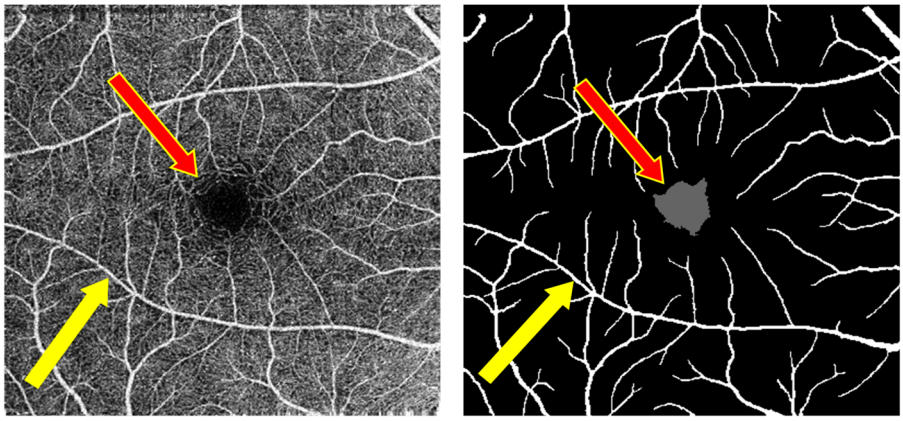


Figure 6. Phase 2, internal structure which include encoder–decoder part followed by K-Means clustering algorithm.

4. Training

The OCTA-500 data set was used in this study, which has 500 subjects and consists of two fields of view (FOV) with six types of projection maps, two types of volumes, four types of text labels, and two types of pixel-level labels. There are more than 360 K images and 80 GB of data. Out of two FOVs, we used OCTA_6M ILM_OPL projection images as shown in Figure 7, which comprises 300 subjects with a 6 mm × 6 mm FOV. This data was collected from Jiangsu Province Hospital between March 2018 and September 2019. OCTA_6 M, whose volume size is 400 px × 400 px × 640 px. The imaging range is 6 mm × 6 mm × 2 mm positioned on the fovea.



(a) (b)

Figure 7. OCTA_6M ILM_OPL projection images: (**a**) input image, (**b**) corresponding ground truth: The red arrows represent RV and the yellow represents FAZ.

Out of six projection images, we chose OCT B-scan-5 maximum projection between the inner retina's internal limiting membrane (ILM) and outer plexiform layer (OPL). This projection map is often used for RV and FAZ segmentation: 80% of the data were used for training, 10% for testing, and 10% for validation. The network was trained using ILM_OPL input images and ground truth maps.

4.1. Hyperparameter Setting

The proposed method was implemented using a cloud-based machine learning platform called Google Colab environment with pre-configured CUDA [41].

4.2. Loss Function

The mean square error (MSE) loss function was used during the training phase. This is shown in Equation (2):

$$E = \frac{1}{N}\sum_{i=1}^{N}\left(Yi - \hat{Yi}\right)^2, \quad (2)$$

where, E is the mean square error, N is the number of observations in a training batch, Y is the label, $\hat{Y}$ is the predicted value [11].

4.3. Optimizer

Adam optimization is easy to implement, computationally efficient, and requires less memory. The Adam optimization method [42] was applied to minimize the loss of the network and ensure an effective calculation. The learning rate was 0.0005, and the number of epochs was set to 20 in this experiment.

5. Results and Discussion

This section describes the results of the unsupervised and supervised learning methods and shows their comparison based on the results from both methods. Then we compare the

results of our proposed encoder–decoder model with other state-of-the-art methods [15]. All the method's performance was evaluated using three well-known metrics, namely, precision, recall, and accuracy, as shown in Equations (3)–(5).

$$Accuracy = \frac{TP + TN}{TP + FP + TP + TN}, \tag{3}$$

$$Precison = \frac{TP}{TP + FP}, \tag{4}$$

$$Recall = \frac{TP}{TP + FN}, \tag{5}$$

where TP represents true positives (correctly predicted RV and FAZ pixels), FN are false negatives, FP is false positives, and TN is true negatives. The proposed methods were evaluated using the performance metrics namely accuracy, precision, and recall. Our deep learning method performs best over unsupervised learning, as shown in Table 3, on RV and FAZ segmentation using OCTA-500 ILM_OPL 2D en face images. We used the trained model on the test set to evaluate ASDC performance. In the context of image segmentation, accuracy is a measure of the overall correctness of the segmentation result, which is typically defined as the ratio of the number of correctly classified pixels to the total number of pixels in the image. In other words, accuracy represents the percentage of pixels that are correctly classified by the segmentation algorithm. The function we used consists of two arguments: the flattened ground truth image and the flattened segmented image. It returns a scalar value between 0 and 1, where 1 indicates perfect accuracy, and 0 indicates complete disagreement between the ground truth and segmented images.

Table 3. Quantitative comparison of the proposed methods on OCTA-500 data.

S/N	Issue	Learning Method	Proposed Method	Projection Map	Data Set: OCTA-500		
					Accuracy (Acc)	Precision (Pre)	Recall (Rec)
1	RV and FAZ	Unsupervised	K-Means clustering	OCTA_6M ILM_OPL	0.89	0.67	0.79
2	segmentation	Supervised	ASDC		**0.96**	**0.91**	**0.83**
3	RV segmentation	Supervised	ASDC		**0.97**	**0.92**	**0.91**

In Figure 8, the results from the two proposed methods are shown. The importance of large superficial retinal blood vessels is of great significance in quantifying retinal disease analysis, and the best projection map is the ILM_OPL map which clearly shows the morphology of the large vessel. The irregular white lines represent the large vessel, and the center point represents the FAZ region. The area highlighted in the red circle shows the ability of our proposed method to segment the final RV and FAZ accurately, even with inhomogeneous images and using one projection map. The result from the deep encoder–decoder + K-means clustering outperforms and gains the best outcomes, which are close to the ground truth of the OCTA-500 data set. Figure 8 compares our proposed ASDC approach with the K-means algorithm. Our proposed simple architecture achieves the same comparable results mentioned in the related work. One advantage over other approaches is that our proposed method simultaneously segments RV and FAZ and achieves comparable results. In Figure 8, the input OCTA images are shown in the first column, the K-means clustering algorithm result is shown in the second column, the proposed deep encoder–decoder + K-means clustering method result is shown in the third column, and the ground truth images from the OCTA-500 data set are shown in the last column for all projection maps. In Table 4, we compared the segmentation results of RV and FAZ using various approaches as listed. The table shows that most studies only focused on a single biomarker, whereas our approach simultaneously segmented both biomarkers, resulting in better accuracy metrics.

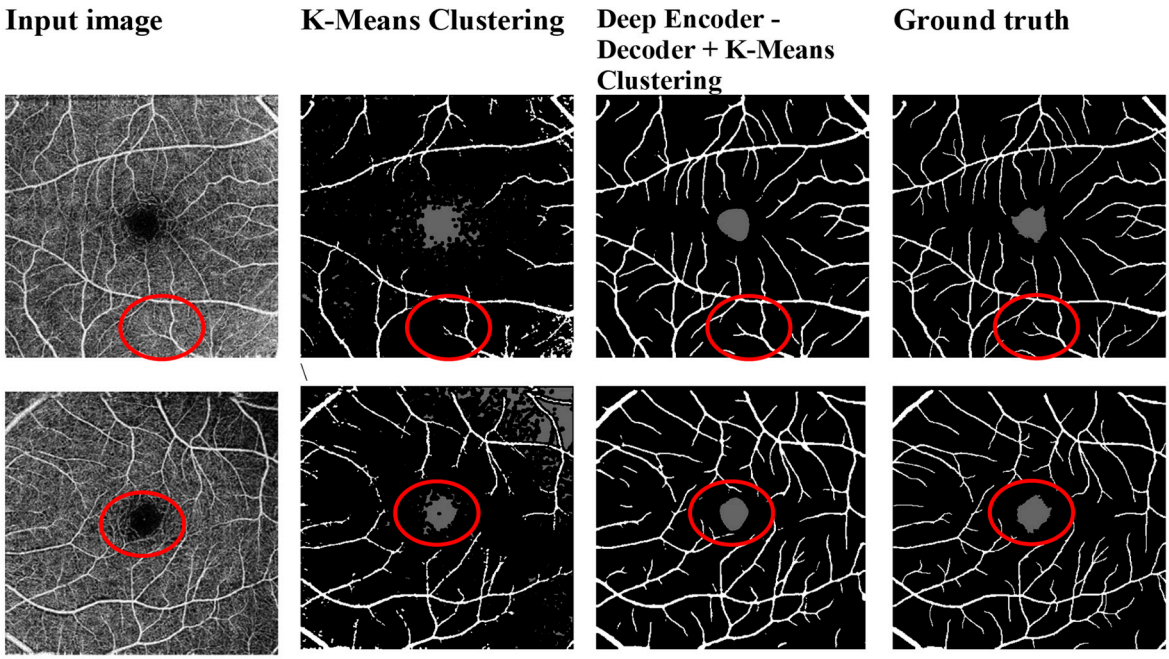


Figure 8. Segmentation results of K-means and ASDC approach on OCTA_6M projection map (ILM_OPL).

Table 4. Comparison of different methods for retinal vasculature segmentation using OCTA images.

Number	Target	Methods	Accuracy
1 [2]	RV segmentation	Soares et al.	0.84
		Nguyen et al.	0.70
		Azzopardi et al.	0.80
		Eladawi et al.	0.93
2 [11]	FAZ segmentation	Guo et al.	0.89
3	RV and FAZ	**Ours**	**0.96**

In Figure 9**,** we compared the latest advanced algorithm for segmenting RV biomarkers. Our study extracts RV from the OCTA scan and segments the FAZ area, which is crucial for detecting retinal diseases such as DR. We have included additional results in Figure 10 to further highlight the effectiveness of our proposed method.

5.1. RV Segmentation Results

In this section, we compare the performance of our proposed method for RV segmentation on the OCTA-500 dataset with several state-of-the-art segmentation methods. Since most previous studies have focused on RV segmentation using complex models, our objective was to implement a lightweight model that would achieve good results while being computationally efficient. We performed our performance calculations on 60 OCTA-500 images, extracting the RV ground truth labels from the OCTA-500 ground truth. Using our ASDC approach, we segmented the input test images and extracted only the RV part, which we saved in a directory. To evaluate the performance of our segmentation results, we used the dice similarity coefficient (DSC) and Jaccard index (JAC) as evaluation indexes, which are defined in Equations (6) and (7) and are commonly used to measure the similarity between predicted values and labels:

$$DSC = \frac{2 \times TP}{FP + FN + 2 \times TP'} \quad (6)$$

$$JAC = \frac{TP}{FP + FN + TP}, \tag{7}$$

with, *TP* (true positive), *FP* (false positive), and *FN* (false negative). *DSC* measures the overlap between the predicted segmentation and the ground truth, with values ranging from 0 to 1. *JAC* is similar to *DSC* but places more weight on the true positives and less weight on the false positives and false negatives.

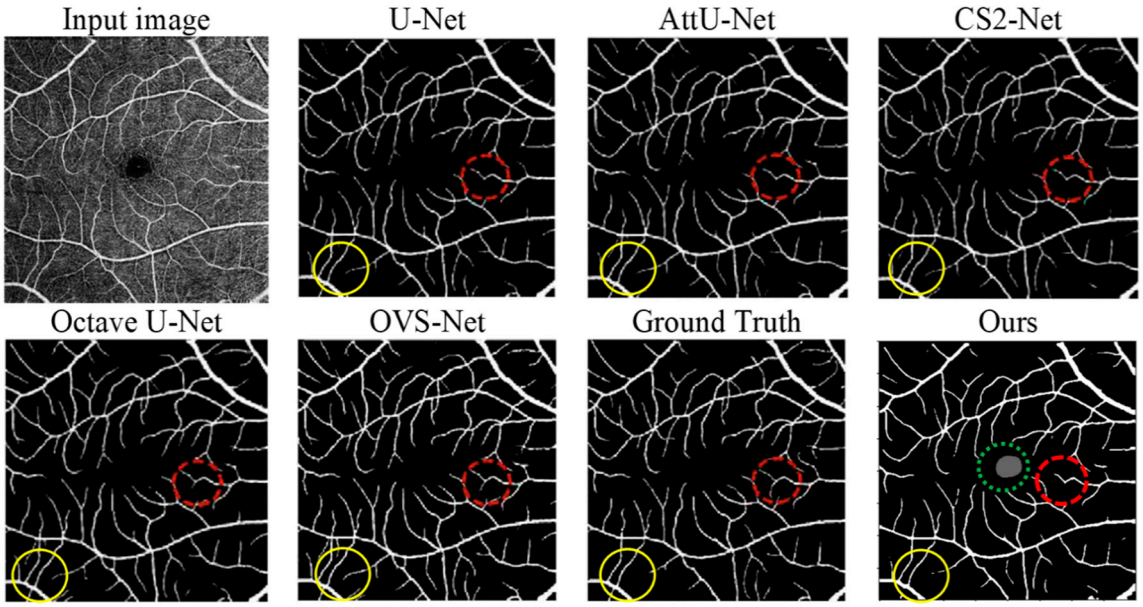


Figure 9. Comparison of RV segmentation results on OCTA-500 with OVS-net. Our proposed method achieves impressive segmentation outcomes for small and large vessels, as shown in row 2, column 4. The green circle represents the FAZ, an added benefit of our proposed method. The results obtained by our proposed method are compared with those reported by Zhu et al. [15].

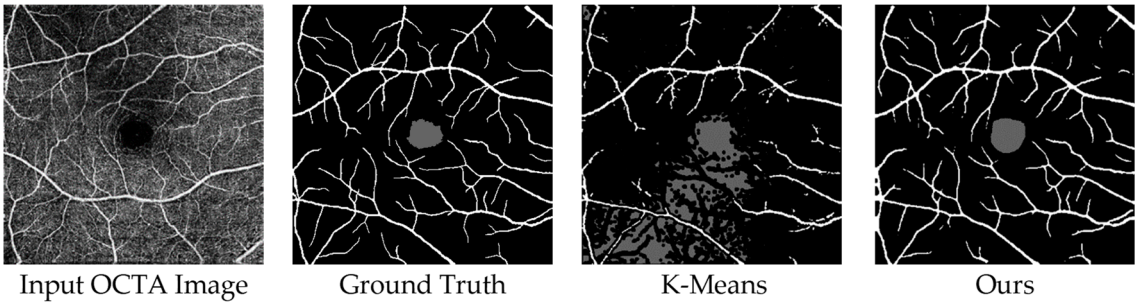


Figure 10. More results of our proposed methods on OCTA-500 data set.

Our results show that our proposed method achieved good performance while utilizing minimal resources. The results are shown in Figure 11 and Table 5. Despite differences in experimental setup, including the number of testing samples and hyperparameters, we were still able to make a fair comparison with other methods by using the same OCTA-500 dataset. Our visual and quantitative results were acceptable in comparison with other methods, suggesting that our proposed method may be a promising approach for RV segmentation in OCTA images. These results highlight the potential of our lightweight model to achieve good results while being computationally efficient, with potential applications in the diagnosis and monitoring of retinal diseases. In Table 5, the DICE and JAC values for RV segmentation and RV + FAZ segmentation are different because the RV + FAZ segmentation includes both the RV and the FAZ biomarkers, while the RV segmentation

only includes the RV biomarker. The inclusion of the FAZ in the RV + FAZ segmentation makes the segmentation task more challenging because the FAZ is a small and complex structure that can be difficult to accurately segment.

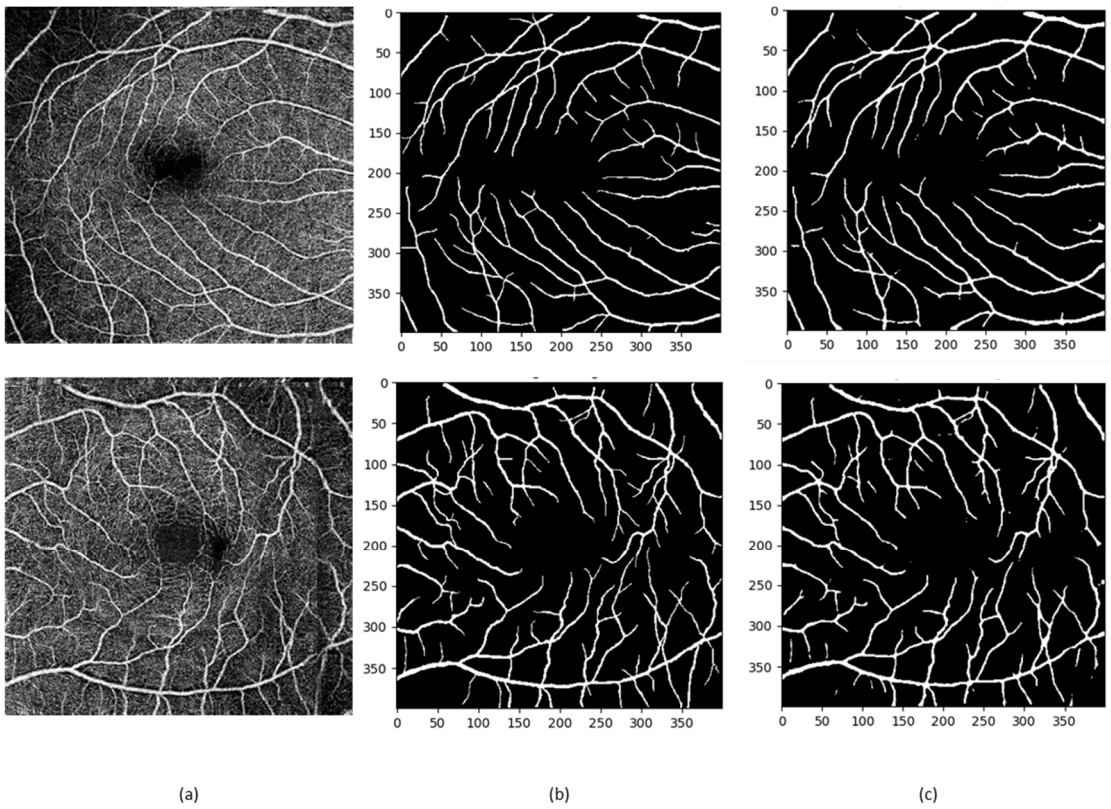


Figure 11. RV segmentation results on OCTA-500 scans: (**a**) input image, (**b**) ground truth, (**c**) ASDC segmented image.

Table 5. Comparison of RV segmentation results on the OCTA 500 dataset.

Data Set	Target	Methods		Dice (%)	Jac (%)
OCTA_6M	RV segmentation	U-Net	[15]	86.52	77.32
		R2U-Net4		81.22	69.29
		AttU Net5		86.47	77.23
		CE-Net6		83.04	72.04
		Cs2-Net4		86.13	76.71
		FARGO		89.15	80.50
		Ours		**91.60**	**85.20**
	RV and FAZ	**Ours**		**86.64**	**79.04**

In general, the inclusion of more structures in a segmentation task can lead to lower dice and JAC values because the segmentation becomes more difficult. This is reflected in the lower dice and JAC values for the RV + FAZ segmentation compared to the RV segmentation. Figure 12 presents an analysis of the performance metrics for RV segmentation of OCTA-500 scans.

5.2. Efficiency Analysis

- FLOPs and Trainable Parameters

To demonstrate the effectiveness of the ASDC method in improving the efficiency of biomarker extraction, the computational complexity of the proposed model was evaluated using FLOPs (floating point operations per second) and number of trainable parameters.

The number of FLOPs required by a neural network model is an important metric for evaluating its computational efficiency. FLOPs correspond to the number of arithmetic operations (such as additions and multiplications) performed by the model per input sample. We obtained the number of FLOPs required for a single forward pass on a 400 × 400 input image is approximately 17.70 billion or (17.70 G). This makes it relatively lightweight and computationally efficient, which can be beneficial for limited computational resources.

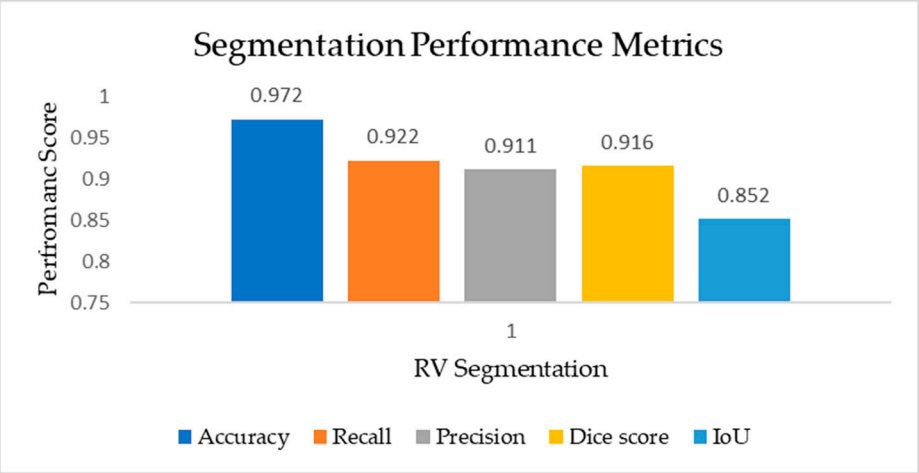


Figure 12. RV segmentation performance metrics on OCTA-500 scans.

The number of trainable parameters in our model is approximately 0.11 million (or 0.11 M) which is a relatively small number compared to other deep learning models. This small number of parameters means that the model requires less memory to store its parameters and is less prone to overfitting to the training data. The U-Net model has a high number of trainable parameters (up to 31 million) and requires a significant number of FLOPs to process input images, resulting in a large model size and high computational complexity. This can make real-time implementation of U-Net models challenging, especially on resource-constrained hardware [43]. Overall, the combination of low FLOPs and a smaller number of trainable parameters makes the ASDC method an efficient and compact alternative to other more complex neural network models, while still being able to extract useful biomarkers from the input images. This can be particularly advantageous in medical applications where computational efficiency and model size are critical factors for real-time analysis and diagnosis.

- Consumed Time

To evaluate the performance of our ASDC using the time consumed for a single image to be segmented, we performed execution time analysis as shown in Figure 13, which is the time consumed for a program or algorithm to complete its task, and it is essential for evaluating the performance of an automatic retinal image processing algorithm. We used the Google Colab environment, and our study found that our ASDC approach is faster than traditional K-means clustering for processing OCTA images of 400 × 400 resolution. This is because deep learning models can automatically learn the relevant features and patterns from large datasets, making them more efficient for handling complex and high-dimensional data. Therefore, our ASDC approach may be a more suitable and efficient method for ophthalmologists to process OCTA images and detect relevant biomarkers with less computational time. The ASDC method took less than a second to segment the OCTA images, with the shortest time being 0.415 s.

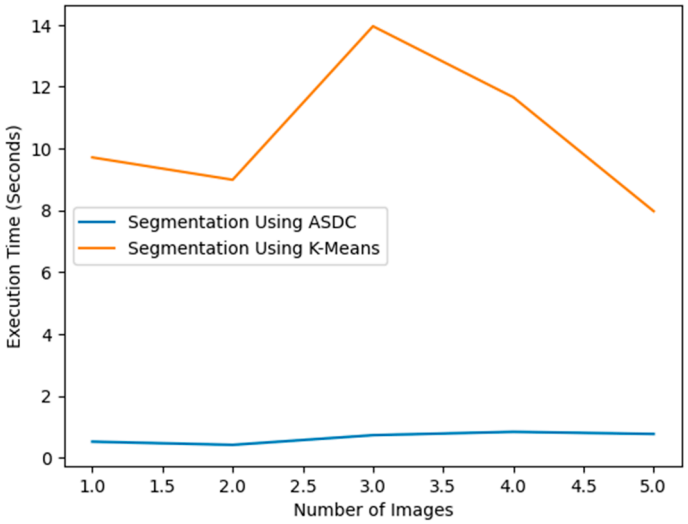


Figure 13. Comparison of computational complexity.

6. Conclusions

This research introduced a new lightweight ASDC approach for the automatic segmentation of RV and FAZ biomarkers in OCTA-500 scans using 2D en face images of the ILM_OPL projection map. The proposed method outperforms existing techniques both visually and quantitatively, achieving comparable results with less computational complexity. Additionally, the proposed method can segment RV and FAZ simultaneously, whereas current approaches only focus on RV segmentation. To ensure fair comparison with other methods, we individually extracted the RV part from our results. These findings suggest that the proposed method offers a promising approach for retinal image analysis and clinical applications.

Recent approaches used 3D OCTA volume data; ours uses 2D en face images. Currently, FAZ and RV extractions are typically done as individual tasks. However, it would be beneficial to understand the connections between the biomarkers in order to extract the necessary information to meet the clinical needs more effectively. The primary issue in segmenting large areas, such as FAZ, is in accurately determining the boundaries and in highlighting the complexities and relevance of accurate boundary segmentation.

The fundamental cause is that the encoder–decoder network is generally inefficient; during the down-sampling process; the loss of image features, such as boundary information, steadily increases, and such data cannot be recovered during the up-sampling process in the decoding stage. After segmenting these important biomarkers, we compared supervised learning and K-means clustering techniques to show the effectiveness of each method. We compared the results with other methods and highlighted the benefit of simultaneous segmentation of retinal biomarkers. Hence, simultaneously segmenting the retinal biomarkers RV and FAZ can enhance the investigation of retinal vascular diseases associated with these biomarkers and further advance our understanding of the retinal vascular system.

We need to point out that our study still has shortcomings, e.g., we acquire relatively smooth FAZ boundaries, which are rough and irregular in most cases, so our further study will be the improvement of the accuracy of FAZ boundaries. In the future, we will create a segmentation and classification system to examine retinal biomarkers used by ophthalmologists in DR early stages.

Author Contributions: Conceptualization, A.K., Z.D. and J.L.; Data curation, A.K., J.H. and J.L.; Formal analysis, A.K., J.H., Z.D. and J.L.; Funding acquisition, J.L.; Investigation, A.K. and J.L.; Methodology, A.K. and J.L.; Project administration, J.L.; Software, A.K.; Supervision, Z.D. and J.L.; Validation, A.K., J.H., Z.D. and J.L.; Visualization, A.K. and J.H.; Writing—original draft, A.K.;

Writing—review and editing, A.K., J.H., Z.D. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Department of Science & Technology of Shandong. Province (2017CXGC0810) and Shandong Natural Science Foundation (ZR2021QF043).

Data Availability Statement: The ASDC experiment utilized the comprehensive public dataset OCTA-500 to facilitate the analysis and interpretation of the research findings.

Acknowledgments: We want to express our sincere gratitude to Yukun Guo, Biomedical Engineering (BME) student at Oregon Health & Science University USA, who provided invaluable guidance and support throughout this research project. He offered insightful feedback and suggestions that significantly improved the quality of our work, and his expertise and dedication were instrumental in the success of this project. We would also like to thank Tiyu Fang for his collaboration in managing the data and providing guidance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Breger, A.; Goldbach, F.; Gerendas, B.S.; Schmidt-Erfurth, U.; Ehler, M. Blood vessel segmentation in en-face octa images: A frequency based method. In Proceedings of the Medical Imaging 2022: Computer-Aided Diagnosis, San Diego, CA, USA, 4 April 2022; Volume 12033, pp. 520–530.
2. Eladawi, N.; Elmogy, M.; Helmy, O.; Aboelfetouh, A.; Riad, A.; Sandhu, H.; Schaal, S.; El-Baz, A. Automatic blood vessels segmentation based on different retinal maps from octa scans. *Comput. Biol. Med.* **2017**, *89*, 150–161. [CrossRef]
3. Noronha, K.; Navya, K.T.; Nayak, K.P. Support system for the automated detection of hypertensive retinopathy using fundus images. In Proceedings of the International Conference on Electronic Design and Signal Processing ICEDSP, Manipal, India, 20–22 December 2012; pp. 1–5.
4. Pascual-Prieto, J.; Burgos-Blasco, B.; Avila Sanchez-Torija, M.; Fernández-Vigo, J.I.; Arriola-Villalobos, P.; Barbero Pedraz, M.A.; García-Feijoo, J.; Martínez-de-la-Casa, J.M. Utility of optical coherence tomography angiography in detecting vascular retinal damage caused by arterial hypertension. *Eur. J. Ophthalmol.* **2020**, *30*, 579–585. [CrossRef] [PubMed]
5. Kashani, A.H.; Chen, C.-L.; Gahm, J.K.; Zheng, F.; Richter, G.M.; Rosenfeld, P.J.; Shi, Y.; Wang, R.K. Optical coherence tomography angiography: A comprehensive review of current methods and clinical applications. *Prog. Retin. Eye Res.* **2017**, *60*, 66–100. [CrossRef] [PubMed]
6. Cardoso, J.N.; Keane, P.A.; Sim, D.A.; Bradley, P.; Agrawal, R.; Addison, P.K.; Egan, C.; Tufail, A. Systematic evaluation of optical coherence tomography angiography in retinal vein occlusion. *Am. J. Ophthalmol.* **2016**, *163*, 93–107.e106. [CrossRef] [PubMed]
7. Sheng, B.; Li, P.; Mo, S.; Li, H.; Hou, X.; Wu, Q.; Qin, J.; Fang, R.; Feng, D.D. Retinal vessel segmentation using minimum spanning superpixel tree detector. *IEEE Trans. Cybern.* **2018**, *49*, 2707–2719. [CrossRef] [PubMed]
8. Guo, M.; Zhao, M.; Cheong, A.M.; Corvi, F.; Chen, X.; Chen, S.; Zhou, Y.; Lam, A.K. Can deep learning improve the automatic segmentation of deep foveal avascular zone in optical coherence tomography angiography? *Biomed. Signal Process. Control* **2021**, *66*, 102456. [CrossRef]
9. Liang, Z.; Zhang, J.; An, C. Foveal avascular zone segmentation of octa images using deep learning approach with unsupervised vessel segmentation. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 1200–1204.
10. Lin, L.; Wang, Z.; Wu, J.; Huang, Y.; Lyu, J.; Cheng, P.; Wu, J.; Tang, X. BSDA-Net: A boundary shape and distance aware joint learning framework for segmenting and classifying octa images. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Strasbourg, France, 27 September–1 October 2021; pp. 65–75.
11. Guo, Y.; Camino, A.; Wang, J.; Huang, D.; Hwang, T.; Jia, Y. MEDnet, a neural network for automated detection of avascular area in OCT angiography. *Biomed. Opt. Express* **2018**, *9*, 5147–5158. [CrossRef]
12. Peng, L.; Lin, L.; Cheng, P.; Wang, Z.; Tang, X. Fargo: A joint framework for faz and rv segmentation from octa images. In Proceedings of the Ophthalmic Medical Image Analysis: 8th International Workshop, OMIA 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, 27 September 2021; Volume 8, pp. 42–51.
13. Meiburger, K.M.; Salvi, M.; Rotunno, G.; Drexler, W.; Liu, M. Automatic segmentation and classification methods using optical coherence tomography angiography (octa): A Review and Handbook. *Appl. Sci.* **2021**, *11*, 9734. [CrossRef]
14. Ma, Y.; Hao, H.; Xie, J.; Fu, H.; Zhang, J.; Yang, J.; Wang, Z.; Liu, J.; Zheng, Y.; Zhao, Y. ROSE: A retinal OCT-angiography vessel segmentation dataset and new model. *IEEE Trans. Med. Imaging* **2021**, *40*, 928–939. [CrossRef]
15. Zhu, C.; Wang, H.; Xiao, Y.; Dai, Y.; Liu, Z.; Zou, B. OVS-Net: An effective feature extraction network for optical coherence tomography angiography vessel segmentation. *Comput. Animat. Virtual Worlds* **2022**, *33*, e2096. [CrossRef]
16. Li, M.; Zhang, Y.; Ji, Z.; Xie, K.; Yuan, S.; Liu, Q.; Chen, Q. IPN-V2 and OCTA-500: Methodology and Dataset for Retinal Image Segmentation. *arXiv* **2020**, arXiv:2012.07261.

17. Lu, Y.; Simonett, J.M.; Wang, J.; Zhang, M.; Hwang, T.S.; Hagag, A.M.; Huang, D.; Li, D.; Jia, Y. Evaluation of automatically quantified foveal avascular zone metrics for diagnosis of diabetic retinopathy using optical coherence tomography angiography. *Investig. Opthalmol. Vis. Sci.* **2018**, *59*, 2212–2221. [CrossRef] [PubMed]
18. Díaz, M.; Novo, J.; Cutrín, P.; Gómez-Ulla, F.; Penedo, M.G.; Ortega, M. Automatic segmentation of the foveal avascular zone in ophthalmological OCT-A images. *PLoS ONE* **2019**, *14*, e0212364. [CrossRef] [PubMed]
19. Qin, X.; Xu, M.; Zheng, C.; He, C.; Zhang, X. Multi-scale feedback feature refinement u-net for medical image segmentation. In Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME), Shenzhen, China, 5–9 July 2021; pp. 1–6.
20. Saha, S.K.; Xiao, D.; Bhuiyan, A.; Wong, T.Y.; Kanagasingam, Y. Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: A review. *Biomed. Signal Process. Control* **2018**, *47*, 288–302. [CrossRef]
21. Yasser, I.; Khalifa, F.; Abdeltawab, H.; Ghazal, M.; Sandhu, H.S.; El-Baz, A. Automated diagnosis of optical coherence tomography angiography (octa) based on machine learning techniques. *Sensors* **2022**, *22*, 2342. [CrossRef] [PubMed]
22. Guo, Y.; Budak, Ü.; Şengür, A. A novel retinal vessel detection approach based on multiple deep convolution neural networks. *Comput. Methods Programs Biomed.* **2018**, *167*, 43–48. [CrossRef] [PubMed]
23. Hu, K.; Zhang, Z.Z.; Niu, X.R.; Zhang, Y.; Cao, C.H.; Xiao, F.; Gao, X.P. Retinal vessel segmentation of color fundus images using multi-scale convolutional neural network with an improved cross-entropy loss function. *Neurocomputing* **2018**, *309*, 179–191. [CrossRef]
24. Odstrcilik, J.; Kolar, R.; Budai, A.; Hornegger, J.; Jan, J.; Gazarek, J.; Kubena, T.; Cernosek, P.; Svoboda, O.; Angelopoulou, E. Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database. *IET Image Process.* **2013**, *7*, 373–383. [CrossRef]
25. Bernardes, R.; Serranho, P.; Lobo, C. Digital ocular fundus imaging: A review. *Ophthalmologica* **2011**, *226*, 161–181. [CrossRef]
26. Shah, S.A.A.; Tang, T.B.; Faye, I.; Laude, A. Blood vessel segmentation in color fundus images based on regional and Hessian features. *Graefe's Arch. Clin. Exp. Ophthalmol.* **2017**, *255*, 1525–1533. [CrossRef]
27. Christodoulidis, A.; Hurtut, T.; Tahar, H.B.; Cheriet, F. A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images. *Comput. Med. Imaging Graph.* **2016**, *52*, 28–43. [CrossRef] [PubMed]
28. Wang, X.; Jiang, X.; Ren, J. Blood vessel segmentation from fundus image by a cascade classification framework. *Pattern Recognit.* **2019**, *88*, 331–341. [CrossRef]
29. Li, X.; Ding, J.; Tang, J.; Guo, F. Res2Unet: A multi-scale channel attention network for retinal vessel segmentation. *Neural Comput. Appl.* **2022**, *34*, 12001–12015. [CrossRef]
30. Zhang, Y.; Lian, J.; Rong, L.; Jia, W.; Li, C.; Zheng, Y. Even faster retinal vessel segmentation via accelerated singular value decomposition. *Neural Comput. Appl.* **2020**, *32*, 1893–1902. [CrossRef]
31. Li, A.; You, J.; Du, C.; Pan, Y. Automated segmentation and quantification of oct angiography for tracking angiogenesis progression. *Biomed. Opt. Express* **2017**, *8*, 5604–5616. [CrossRef] [PubMed]
32. Xie, H.; Zhang, L.; Lim, C.P.; Yu, Y.; Liu, C.; Liu, H.; Walters, J. Improving k-means clustering with enhanced firefly algorithms. *Appl. Soft Comput.* **2019**, *84*, 105763. [CrossRef]
33. Kim, Y.W.; Krishna, A.V. A study on the effect of canny edge detection on downscaled images. *Pattern Recognit. Image Anal.* **2020**, *30*, 372–381. [CrossRef]
34. Chlebiej, M.; Gorczynska, I.; Rutkowski, A.; Kluczewski, J.; Grzona, T.; Pijewska, E.; Sikorski, B.L.; Szkulmowska, A.; Szkulmowski, M. Quality improvement of oct angiograms with elliptical directional filtering. *Biomed. Opt. Express* **2019**, *10*, 1013–1031. [CrossRef]
35. Karimpouli, S.; Tahmasebi, P. Segmentation of digital rock images using deep convolutional autoencoder networks. *Comput. Geosci.* **2019**, *126*, 142–150. [CrossRef]
36. Dumoulin, V.; Visin, F. A guide to convolution arithmetic for deep learning. *arXiv* **2016**, arXiv:1603.07285.
37. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; Volume 37, pp. 448–456.
38. Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Singh, V. Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet. *Chaos Solitons Fractals* **2020**, *138*, 109944. [CrossRef]
39. Gong, Q.; Wang, P.; Cheng, Z. An encoder-decoder model based on deep learning for state of health estimation of lithium-ion battery. *J. Energy Storage* **2022**, *46*, 103804. [CrossRef]
40. Zhou, S.; Song, W. Concrete roadway crack segmentation using encoder-decoder networks with range images. *Autom. Constr.* **2020**, *120*, 103403. [CrossRef]
41. Carneiro, T.; Da Nóbrega, R.V.M.; Nepomuceno, T.; Bian, G.B.; De Albuquerque, V.H.C.; Reboucas Filho, P.P. Performance analysis of google colaboratory as a tool for accelerating deep learning applications. *IEEE Access* **2018**, *6*, 61677–61685. [CrossRef]
42. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
43. Shah, K.D.; Patel, D.K.; Patel, H.A.; Nagrani, H.R. EMED-UNet: An efficient multi-encoder-decoder based unet for chest X-ray segmentation. In Proceedings of the 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, 1–3 July 2022; pp. 1–6.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.