Automatic Drusen Segmentation for Age-Related Macular Degeneration in Fundus Images Using Deep Learning

Quang T. M. Pham, Sangil Ahn, Su Jeong Song, and Jitae Shin

1 College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, Korea; quangpham@skku.edu (Q.T.M.P.); il2s@skku.edu (S.A.)
2 Department of Ophthalmology, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul 03181, Korea
3 Biomedical Institute for Convergence (BICS), Sungkyunkwan University, Suwon 16419, Korea
* Correspondence: ssjeye@skku.edu (S.J.S.); jtshin@skku.edu (J.S.)

Received: 31 August 2020; Accepted: 24 September 2020; Published: 1 October 2020

Abstract: Drusen are the main aspect of detecting age-related macular degeneration (AMD). Ophthalmologists can evaluate the condition of AMD based on drusen in fundus images. However, in the early stage of AMD, the drusen areas are usually small and vague. This leads to challenges in the drusen segmentation task. Moreover, due to the high-resolution fundus images, it is hard to accurately predict the drusen areas with deep learning models. In this paper, we propose a multi-scale deep learning model for drusen segmentation. By exploiting both local and global information, we can improve the performance, especially in the early stages of AMD cases.

Keywords: deep learning; drusen segmentation; high-resolution

1. Introduction

Age-related macular degeneration (AMD) is one of the most common diseases that can lead to blindness worldwide. The pathognomonic feature of AMD is drusen, which are the bright/yellow regions of various sizes in color fundus images. Drusen and early AMD are defined in accordance with the international classification developed by the International Age-Related Maculopathy Epidemiological Study Group [1]. Detecting and monitoring drusen in the early stage of AMD is a crucial step to diagnosis and selecting patients with a high risk of progression for early AMD. Until now, drusen location, size, numbers, and associated features are known to be associated with increased risk of AMD progression [2,3]. Thus objective and automatic detection of drusen is an important process for both researchers and clinicians as well.

Nowadays deep learning has become a powerful tool for not only computer vision research but also in medical image analysis using convolutional neural networks (CNNs) with overwhelming results. However, we observe several challenges in the drusen segmentation task. Firstly, drusen can appear anywhere with various sizes and numbers. As shown in Figure 1, in the early stage, drusen can be small and vague, making them difficult to detect. In several cases, the small drusen can be confused with noise without a global context. Secondly, high-resolution fundus images can have different colors according to the machine. This may cause several problems for deep-learning based methods. If we resize and predict drusen on the whole low-resolution fundus image, the result might be not accurate since the model cannot recognize the small drusen. On the other hand, if we crop the full image into patches and train the deep learning model on the patch-level, we may miss important global information, leading to the wrong predictions. Therefore, most of the current works only use public datasets such as STARE [4] and ARIA [5] for evaluation. These datasets have a limited number of early
AMD images and most of them are in the intermediate and advanced stages, making the detection of drusen changes difficult (Figure 1). Thus it is important to develop a method that can detect drusen changes in the very early AMD stages.

![Image](image1.png) ![Image](image2.png)

**Figure 1.** The comparison of drusen between early AMD and intermediate AMD.

To solve these problems, we introduce a multi-scale deep learning model for drusen segmentation. Firstly, we make a segmentation model for detecting drusen throughout the whole image. This model learns to extract the global context on the image-level. Then, to refine the segmentation results, we introduce another segmentation model for detecting the drusen at the patch level. Moreover, we introduce an efficient way to exploit global information from the image-level model into the patch-level model. Therefore, the second model takes both local and global information (from the first model) to make the prediction. By this setting, we can make accurate predictions, even in high-resolution images. Overall, our contributions of this research are described below:

1. We proposed a multi-scale deep learning model for drusen segmentation, which can detect the drusen by using both global and local information. It helps to overcome the issue of small and vague drusen.
2. We apply a simple and efficient way to use global information in the patch-level model. Thus, it is possible to use the well-known backbone network with less computational complexity. This allows us to benefit from pre-trained models.
3. We evaluate the proposed method on the public dataset, STARE, and our dataset from Kangbuk Samsung Hospital. The experimental results show the effectiveness of our proposal compared to other state-of-the-art (SOTA) methods.

### 2. Related Works

The segmentation task in the medical domain has been studied for a long time. The superpixel method was used for 3D prostate MR images [6]. For improvement, a multi-atlas fusion framework was introduced in [7]. The watershed algorithm is commonly used for segmentation. Huang et al. [8] applied it for breast tumors segmentation. On the other hand, the combination of the watershed algorithm and neural networks was introduced in [9] for liver segmentation in MRI images. The active-contour based methods were also used in medical imaging [10]. More specifically, Zhao et al. [11] exploited an active contour model with hybrid region information for vessel segmentation in fundus images.

The methods of drusen segmentation can be divided into two main approaches. The first direction is based on traditional image processing techniques. Usually, several local features are extracted, and then, a machine learning algorithm (e.g., SVM [12]) is made for the classification task [13–15]. The main goal is to detect the drusen region directly [16,17], or to determine the boundary of the drusen [18]. Kim et al. [17] tried to apply multi-filters on the candidate regions for detecting drusen while the GrowCut segmentation method was exploited in [15]. To handle the various sizes of drusen, the multiscale local image descriptors were applied in [14,19]. On the other hand, a color normalization method was introduced in [20] to deal with the color variants problem in fundus images.
The gradient-based features were also used for drusen segmentation [18,21]. Since most of those methods are based on local features, they might not work well in the case of early AMD. To detect small vague drusen in high-resolution fundus images, it requires local features and global context information. Although there are several global features exploited for improvement [14,16], they are not good enough to detect drusen in an early stage.

Recently, many deep learning models are applied to segmentation tasks in medical image analysis. Many researchers tried to apply deep learning models in Optical Coherence Tomography (OCT) images [22,23]. Although we can easily detect drusen in OCT images, the OCT test is expensive. The fundus images are more popular and easier to access than the OCT test. For fundus image analysis, deep learning also shows many applications. For example, UNet [24] was successfully applied to vessel segmentation [25] and optic disc detection [26].

Many works tried to perform the segmentation task on diabetic retinopathy lesions. The L-seg [27] was introduced for multi-lesion segmentation of diabetic retinopathy. Khojasteh et al. [28] used a deep convolutional network to detect exudates, hemorrhages, and microaneurysms in patch-level. To improve performance, a combination of global-local U-Nets was exploited in [29]. However, there are not many works of drusen segmentation using the deep learning method. In 2018, Fang Yan et al. [30] proposed a deep learning model with a random walk module to detect drusen in the patch level.

3. Method

3.1. Model Architecture

As we can see in Figure 2, our proposed model includes two networks. The first network is the Image-Level network. It takes the whole fundus images as input and predicts the drusen segmentation mask. The second network is the Patch-Level network. This network refines the prediction result from the Image-Level Network. The input of the Patch-Level network is a combination of patch images and their corresponding probability maps of drusen segmentation from the Image-Level network. The Patch-Level network outputs the prediction on the patch-level. We use the results of this network as the final prediction. By using the prediction of drusen segmentation from the Image-Level network as the input of the Patch-Level network, it helps the Patch-Level network employ global information for making predictions.

Figure 2. The proposed multi-scale model for drusen segmentation.
The reason we choose the probability maps instead of the feature maps of the Image-Patch network as the extra input of the Patch-Level network is because of the computational cost. The feature maps of intermediate layers can bring more information than prediction maps. However, it is not computational efficient to use the large feature maps as the input of a network. The recent SOTA segmentation model DeepLabV3+ [31] has 256 output feature maps. To exploit those pre-trained models, we decided to use the probability maps. In Section 4.4, we show that the pre-trained models can improve the performance significantly.

3.2. Image-Level Network

The goal of this network is to determine the candidate regions that might contain drusen. Firstly, due to the various sizes of raw data, all the fundus images are resized to 1200 × 1200. The input of the Image-Level network is the 512 × 512 fundus images, and the output is the probability map of the drusen. We use the pre-trained DeepLabV3+ [31] for the Image-Level network.

Since the drusen can be vague in the early stage of AMD, it is important to make sure that the Image-Level network does not miss those drusen. The drusen area is much smaller than the background area. This class imbalance problem leads the network to predict the confused region as non-drusen. Therefore, we train the Image-Level network with the weighted binary cross-entropy (weighted BCE) loss as below:

\[
\text{loss}_{\text{weightedBCE}} = - \sum_{i=0}^{N} w_i y_i \log(x_i) + (1 - y_i) \log(1 - x_i)
\]

where \( y_i \) and \( x_i \) are the ground truth and the prediction of the \( i \)-th pixel in the total \( N \) pixels. The weight of the positive samples is \( w \).

Although the other loss functions (e.g., Focal loss or Dice loss) can have better Dice scores, the network ignores confused regions which leads to low sensitivity if we use those loss functions. The weighted BCE loss can increase sensitivity and help our model focus on small drusen. This is the reason we choose pre-trained DeepLabV3+ as the Image-Level model since it obtains the best sensitivity on detecting early drusen (refer to Table 3).

3.3. Patch-Level Network

The Patch-Level Network takes both global and local information to make the final prediction. Firstly, we crop 1200 × 1200 fundus images and their corresponding probability maps from the Image-Level network into 128 × 128 patches. Then, the input of the Patch-Level network is comprised of the patch images and patch probability maps. The detailed architecture is described in Figure 3.

![Figure 3. The Patch-Level Network architecture. The parameters of the convolution layer are: k is the kernel size, n is the number of channels and s is the stride.](image)

The probability maps show the Patch-Level network where it should focus. Since the patch images do not contain the global context, the drusen area can be confused as a noise area in the case of
small drusen or low contrast. By employing such global information, the Patch-Level network now focuses on the boundary of drusen and makes a more accurate boundary. That is the reason why we should increase the sensitivity instead of the Dice score when training the Image-Level network.

We denote true positive, true negative, false positive, and false negative as \( TP \), \( TN \), \( FP \), and \( FN \), respectively. For training the Patch-Level network, we use Focal Tversky loss as below:

\[
\text{loss}_{\text{FocalTversky}} = (1 - TI)^\gamma
\]

where \( TI \) is the Tversky Index:

\[
TI = \frac{TP}{TP + \alpha FN + \beta FP}.
\]

In Equations (2) and (3), \( \gamma \), \( \alpha \), and \( \beta \) are the parameters of the Focal Tversky loss.

The Focal Tversky loss can work better than weighted BCE loss or Focal loss for dealing with class imbalance data. Especially, it is suitable for detecting small drusen. Moreover, we witness a faster coverage by using Focal Tversky loss.

4. Experiments

4.1. Dataset

This study adhered to the tenets of the Declaration of Helsinki, and the study protocol was reviewed and approved by the Institutional Review Board of Kangbuk Samsung Hospital (No. KBSMC 2019-01-014). The requirement for written informed consent was waived because the study used retrospective and anonymized retinal images.

Our dataset includes 775 high-resolution color fundus photographs from early AMD patients from Kangbuk Samsung Hospital Department of Ophthalmology from 2007–2018. Fundus photographs were taken with nonmydriatic fundus cameras of various manufacturers, including TRC NW300, TRC-510X, NW200, and NW8 (Topcon, Japan); CR6-45NM and CR-415NM (Canon, Japan); and VISUCAM 224 (Zeiss, USA). Digital images of the fundus photographs were analyzed with a picture archiving communication system (INFINITT, Republic of Korea). We used 697 images for training and the rest of the 78 images for testing. Moreover, we also tested our model on the public dataset STARE [4]. This dataset contains 400 fundus images with different diseases. We selected 28 AMD images from STARE for testing. All the ground truths are annotated by the expert. Notice that we do not apply any specific pre-processing techniques except cropping the center image to remove redundant background regions.

For the patch images, all full-size fundus images are resized to 1200 × 1200 and then cropped to 128 × 128 patches. To avoid all background samples, we only select the patches including at least one pixel as drusen. Overall, we have about 11,000 patches.

4.2. Implementation Details

For training the DeepLabV3+ as the Image-Level network, we use the input size 512 × 512. The Adam optimizer is applied with a learning rate of \( 10^{-4} \). The number of epochs is 10 and the batch size is 4. For training the UNet as the Patch-Level network, we use the input size 128 × 128 with 4 channels. The number of epochs is 25 and the batch size is 32. The high learning rate of 0.01 is used with the Adam optimizer for the first 15 epochs and then the learning rate is reduced to 0.001 for the rest of the training process. To avoid overfitting, we also apply several data augmentation methods: shift, scale, rotate, and horizontal flip. The input images are normalized to \([-1,1]\) so that it has the same range with the probability maps (range [0,1] from the sigmoid activation of the last layer of Image-Level network). The implementation can be found in the repository https://github.com/QuangBK/drusen_seg.
4.3. The Metrics

The common metrics for drusen segmentation are sensitivity, specificity, accuracy, and Dice score:

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (4)
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (5)
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)
\]

\[
\text{Dice score} = \frac{2 \times TP}{2 \times TP + FP + FN}. \quad (7)
\]

Since the drusen segmentation task leads to the imbalanced class problem, the Dice score is the best metric compared to the others. Due to the domination of the background class, the true negative should be much larger than the false positive. Thus, specificity and accuracy are always high. On the other hand, by sacrificing specificity a bit, we can increase the sensitivity significantly. In that case, all confused regions would be classified as drusen (increasing false positive) and the prediction mask is not accurate. Meanwhile, the Dice score is sensitive to both true positive and false positive even if the dataset is imbalanced. The comparison between those metrics is illustrated in Figure 4.

4.4. Ablation Study

We first examine the benefit of global and local information on drusen segmentation. The Image-Level network and the Patch-Level network are trained independently to check the impact on performance. In this experiment, for fair comparison, we used the same UNet architecture for both networks and used the same loss function of Focal Tversky loss. As we can see from Table 1, the Image-Level network provides a better Dice score than the Patch Level network. Drusen are small and not obvious in the early AMD stage. Thus it is understandable that the network cannot detect drusen well in the patch-level since drusen may look like a noise area. For example, we can see that the Patch-Level model predicts several noise areas and the bright area near the optic disc as drusen in Figure 5. By getting a global context, the Image-Level network can find the drusen regions better. However, predicting drusen on low-resolution images (512 × 512) does not give detailed segmentation masks, especially on small drusen (Figure 5). By combining both global and local information with our method, we can generate accurate drusen segmentation masks with high resolution (1200 × 1200). We witness a significant improvement of our model compared to the other two models with a Dice score of 0.508. Our model also achieves the best specificity, sensitivity, and accuracy.
Table 1. The impact of global, local, and combination of both information for drusen segmentation. We evaluate on our dataset from Kangbuk Samsung Hospital.

| Model                      | Specificity | Sensitivity | Accuracy | Dice Score |
|----------------------------|-------------|-------------|----------|------------|
| Only Image-Level network (UNet) | 0.991       | 0.510       | 0.986    | 0.372      |
| Only Patch-Level network (UNet) | 0.992       | 0.558       | 0.989    | 0.356      |
| Proposed method            | 0.995       | 0.674       | 0.993    | 0.508      |

Figure 5. The segmentation results of Image-Level model, Patch-Level model, and combined model.

Next, the impact of different loss functions and pre-trained models are considered. All the models in this experiment are trained on full images with a $512 \times 512$ input size. We trained the UNet (with random initialization) with different loss functions: weighted BCE, Focal loss, and Focal Tversky loss. It is easy to see that the Focal Tversky loss achieves the best performance among three loss functions with a Dice score of 0.372 in Table 2. Moreover, there is a remarkable improvement when using pre-trained models instead of the Unet without pre-training. We applied the pre-trained Squeeze-and-Excitation (SE) Network [32] with a ResNet50 [33] as the backbone (SE-ResNet50). The SE-ResNet50 was trained on ImageNet and use as the encoder of the UNet. It can be seen from Table 2 that the UNet with pre-trained SE-ResNet50 achieves 0.561 for the Dice score. In addition, the DeepLabV3+ also obtains a decent Dice score of 0.517. The other metrics (specificity, sensitivity, and accuracy) also increase by applying pre-trained models.

Table 2. The impact of different loss functions and pre-trained models on the Image-Level models. We evaluate on our dataset from Kangbuk Samsung Hospital.

| Model                                              | Specificity | Sensitivity | Accuracy | Dice Score |
|----------------------------------------------------|-------------|-------------|----------|------------|
| UNet + Weighted BCE                                | 0.984       | 0.684       | 0.981    | 0.327      |
| UNet + Focal loss                                  | 0.999       | 0.172       | 0.992    | 0.227      |
| UNet + Focal Tversky loss                          | 0.991       | 0.510       | 0.986    | 0.372      |
| UNet (with SE-Resnet50) + Focal Tversky loss *     | 0.994       | 0.716       | 0.993    | 0.561      |
| DeepLabV3+ + Focal Tversky loss *                  | 0.992       | 0.725       | 0.990    | 0.517      |

* Note: two last models use pre-trained models.

4.5. Comparison with The SOTA Methods

Finally, we compare our final results with the other methods. Our final model has DeepLabv3+ as the Image-Level network. The Patch-Level network is the UNet with the pre-trained of SE-ResNet50 as the encoder. The model is trained with Focal Tversky loss. We compare those drusen segmentation methods on both our validation dataset (Table 3) and the public dataset STARE (Table 4). The results
show that our model outperforms the other methods in both datasets. As we can see in Figure 6, our method provides an accurate and detailed prediction for both large and small drusen.

Table 3. The comparison of drusen segmentation methods on Kangbuk Samsung Hospital dataset.

| Model                        | Specificity | Sensitivity | Accuracy | Dice Score |
|------------------------------|-------------|-------------|----------|------------|
| Our method                   | 0.997       | 0.662       | 0.995    | 0.625      |
| Kim et al. [17]              | 0.988       | 0.058       | 0.980    | 0.040      |
| Yan et al. [29]              | 0.997       | 0.606       | 0.994    | 0.539      |
| Unet (with SE-ResNet50)      | 0.994       | 0.716       | 0.993    | 0.561      |
| DeepLabV3+                   | 0.992       | 0.725       | 0.990    | 0.517      |

Table 4. The comparison of drusen segmentation methods on STARE dataset.

| Model                        | Specificity | Sensitivity | Accuracy | Dice Score |
|------------------------------|-------------|-------------|----------|------------|
| Our method                   | 0.991       | 0.588       | 0.981    | 0.542      |
| Kim et al. [17]              | 0.980       | 0.478       | 0.968    | 0.328      |
| Yan et al. [29]              | 0.990       | 0.491       | 0.979    | 0.446      |
| Unet (with SE-ResNet50)      | 0.955       | 0.847       | 0.951    | 0.404      |
| DeepLabV3+                   | 0.992       | 0.317       | 0.975    | 0.322      |

Figure 6. The segmentation results of our method and compared methods.

We first compare our method with several segmentation models trained on full images. Our method outperformed the DeepLabV3+ and Unet (with pre-trained SE-ResNet50) with a Dice score of 0.624 on our validation dataset and 0.534 on the STARE dataset. This indicates that our method, which uses both global and local information, can have a better result than the previous models trained on only Image-Level.

Kim’s method [17] using the traditional image processing technique has the lowest Dice score. On the STARE dataset, the drusen areas are bigger and more obvious, and the Dice score is 0.328. However, by using only local features and handcrafted filters, Kim’s method cannot work well in the case of early AMD. As we expected, it presents a poor performance on our Kangbuk Samsung Hospital dataset with the Dice score of 0.040.

Yan’s method [29] also applies two UNets to exploit both global and local information. However, they use the feature maps as global information and fuse them in the last layer of their Patch-Level network (called LocalNet). This setting leads to the problem in which we should have feature maps with a limited number of channels to avoid large computations. Because of that, this method cannot take advantage of the pre-trained models. Note that training our model is 5 times faster than that for Yan’s model. Furthermore, in Yan’s method, they fuse the global information in the last layer of the Patch-Level network, so the Patch-Level network may not benefit much while we use global information as the input of the network. Compared to Yan’s method, our method has a noticeable improvement.
5. Discussion and Conclusions

In this paper, we propose a multi-scale deep learning model for drusen segmentation. While most of the previous work on medical segmentation used UNet-based architecture on the whole image [24,29,34], our method applies a multi-scale learning method to make a fine segmentation prediction. It is suitable for drusen segmentation with high-resolution fundus images. To solve the high-resolution image problem, other researches tried to analyze on cropped image [28,35], which may lose the global information. By combining global and local information, our model is able to predict more accurate drusen segmentation masks. Moreover, by exploiting the pre-trained model and the combination of different loss functions, we can improve the performance of detecting drusen in the early stage of AMD. The experimental results show that our method obtains the best Dice score on both our validation dataset and the STARE dataset.

Since this segmentation task has an imbalanced class problem where the negative class (non-drusen pixel) is dominated over the positive class (drusen pixel), the metric of sensitivity is quite sensitive. Due to the vague drusen area and small amount of drusen pixel (Positive class), we cannot increase both the specificity and sensitivity at the same time. As shown in Figure 4, although we can sacrifice the specificity a bit to boost the sensitivity (by increasing FP and decreasing FN), the prediction is not accurate visually. Therefore, we decided to focus on the Dice index for evaluation. The Dice score is the common metric for image segmentation, especially in medical images [36]. However, drusen segmentation is a challenging task, and we fully acknowledge this. The final goal for our research is to improve the Dice score.

However, our model still has several limitations. The current setting is suitable for binary segmentation. For multi-class segmentation, the number of input channels of Patch-Level Network increases according to the number of classes. It can cause scalability and computational complexity problems. In addition, to generate the drusen prediction, our model needs to predict many patches for each high-resolution fundus image. Therefore, the inference time of our model can be longer than the other models.

Our study provides a fundamental step for further applications of deep learning methods in the screening and diagnosis of AMD. Further studies are needed to validate and expand the clinical implications of our algorithm. For improving performance, there are several possible directions. The loss function can be changed to increase sensitivity. Furthermore, we could introduce better network architecture or additional modules to refine the predictions and deal with multi-class segmentation. Moreover, it is also important to reduce the inference time by presenting a more efficient way to combine global and local information.

Author Contributions: Conceptualization, Q.T.M.P., S.J.S. and J.S.; methodology, Q.T.M.P. and S.A.; formal analysis, Q.T.M.P., S.A. and J.S.; software, validation, Writing—Original draft preparation, Q.T.M.P.; Writing—Review and editing, Q.T.M.P., S.A., S.J.S. and J.S.; supervision, project administration, and funding acquisition, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2020R1F1A1065626) and was partly supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01798) supervised by the IITP (Institute for Information & communications Technology Promotion). It was also partly supported by the research fund from Biomedical Institute for Convergence (BICS), Sungkyunkwan University.

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

References

1. Bird, A.; Bressler, N.; Bressler, S.; Chisholm, I.; Coscas, G.; Davis, M.; de Jong, P.; Klaver, C.; Klein, B.; Klein, R.; et al. An international classification and grading system for age-related maculopathy and age-related macular degeneration. *Surv. Ophthalmol.* **1995**, *39*, 367–374. [CrossRef]
2. Lee, J.; Choi, S.; Lee, C.S.; Kim, M.; Kim, S.S.; Koh, H.J.; Lee, S.C.; Byeon, S.H. Neovascularization in Fellow Eye of Unilateral Neovascular Age-related Macular Degeneration According to Different Drusen Types. *Am. J. Ophthalmol.* 2019, 208, 103–110. [CrossRef] [PubMed]
3. Joachim, N.D.L.; Mitchell, P.; Kifley, A.; Wang, J.J. Incidence, Progression, and Associated Risk Factors of Medium Drusen in Age-related Macular Degeneration: Findings From the 15-Year Follow-up of an Australian Cohort. *JAMA Ophthalmol.* 2015, 133, 698–705. [CrossRef] [PubMed]
4. STARE Dataset. Available online: https://cecas.clemson.edu/~ahoover/stare/ (accessed on 18 August 2020).
5. ARIA Dataset. Available online: https://www.researchgate.net/post/How_can_I_find_the_ARIA_Automatic_Retinal_Image_Analysis_Dataset (accessed on 18 August 2020).
6. Tian, Z.; Liu, L.; Zhang, Z.; Fei, B. Superpixel-Based Segmentation for 3D Prostate MR Images. *IEEE Trans. Med. Imaging* 2016, 35, 791–801. [CrossRef] [PubMed]
7. Nguyen, D.; Benameur, S.; Mignotte, M.; Lavoie, F. Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images. *Med. Image Anal.* 2018, 48, 58–74. [CrossRef] [PubMed]
8. Huang, Y.L.; Chen, D.R. Watershed segmentation for breast tumor in 2-D sonography. *Ultrasound Med. Biol.* 2004, 30, 625–632. [CrossRef] [PubMed]
9. Masoumi, H.; Behrad, A.; Pourmina, M.A.; Roosta, A. Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network. *Biomed. Signal Process. Control.* 2012, 7, 429–437. [CrossRef]
10. Ciecholewski, M.; Spodnik, J. Semi–Automatic Corpus Callosum Segmentation and 3D Visualization Using Active Contour Methods. *Symmetry* 2018, 10, 589. [CrossRef]
11. Zhao, Y.; Rada, L.; Chen, K.; Harding, S.P.; Zheng, Y. Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images. *IEEE Trans. Med. Imaging* 2015, 34, 1797–1807. [CrossRef]
12. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* 1995, 20, 273–297. [CrossRef]
13. Ren, X.; Zheng, Y.; Zhao, Y.; Luo, C.; Wang, H.; Lian, J.; He, Y. Drusen Segmentation From Retinal Images via Supervised Feature Learning. *IEEE Access* 2018, 6, 2952–2961. [CrossRef]
14. Zheng, Y.; Vanderbeek, B.; Daniel, E.; Stambolian, D.; Maguire, M.; Brainard, D.; Gee, J. An automated drusen detection system for classifying age-related macular degeneration with color fundus photographs. In Proceedings of the 2013 IEEE 10th International Symposium on Biomedical Imaging, San Francisco, CA, USA, 7–11 April 2013; pp. 1448–1451.
15. Liu, H.; Xu, Y.; Wong, D.W.K.; Liu, J. Growcut-based drusen segmentation for age-related macular degeneration detection. In Proceedings of the 2014 IEEE Visual Communications and Image Processing Conference, Valletta, Malta, 7–10 December 2014; pp. 161–164.
16. Raza, G.; Rafique, M.; Tariq, A.; Akram, M.U. Hybrid classifier based drusen detection in colored fundus images. In Proceedings of the 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEEC), Amman, Jordan, 3–5 December 2013; pp. 1–5.
17. Kim, Y.J.; Kim, K. Automated Segmentation Methods of Drusen to Diagnose Age-Related Macular Degeneration Screening in Retinal Images. *Comput. Math. Methods Med.* 2018, 2018, 1–8. [CrossRef] [PubMed]
18. Mittal, D.; Kumari, K. Automated detection and segmentation of drusen in retinal fundus images. *Comput. Electr. Eng.* 2015, 47, 82–95. [CrossRef]
19. Zheng, Y.; Wang, H.; Wu, J.; Gao, J.; Goo, J.C. Multiscale analysis revisited: Detection of drusen and vessel in digital retinal images. In Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, 30 March–2 April 2011; pp. 689–692.
20. Mohaimin, S.M.; Saha, S.K.; Khan, A.M.; Mohammad Arif, A.S.; Kanagasingam, Y. Automated method for the detection and segmentation of drusen in colour fundus image for the diagnosis of age-related macular degeneration. *Iet Image Process.* 2018, 12, 919–927. [CrossRef]
21. Mora, A.; Vieira, P.; Manivannan, A.; Fonseca, J. Automated drusen detection in retinal images using analytical modelling algorithms. *Biomed. Eng. Online* 2011, 10, 59. [CrossRef] [PubMed]
22. Huang, L.; He, X.; Fang, L.; Rabbani, H.; Chen, X. Automatic Classification of Retinal Optical Coherence Tomography Images With Layer Guided Convolutional Neural Network. *IEEE Signal Process. Lett.* 2019, 26, 1026–1030. [CrossRef]
23. Asgari, R.; Orlando, J.I.; Waldstein, S.; Schlanitz, F.; Baratsits, M.; Schmidt-Erfurth, U.; Bogunović, H. Multiclass Segmentation as Multitask Learning for Drusen Segmentation in Retinal Optical Coherence Tomography. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*; Shen, D., Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.T., Khan, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 192–200.

24. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

25. Mishra, S.; Chen, D.Z.; Hu, X.S. A Data-Aware Deep Supervised Method for Retinal Vessel Segmentation. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020; pp. 1254–1257.

26. Meyer, M.I.; Galdran, A.; Mendonça, A.M.; Campilho, A. A Pixel-Wise Distance Regression Approach for Joint Retinal Optical Disc and Fovea Detection. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*; Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 39–47.

27. Guo, S.; Li, T.; Kang, H.; Li, N.; Zhang, Y.; Wang, K. L-SEG: An end-to-end unified framework for multi-lesion segmentation of fundus images. *Neurocomputing* 2019, 349, 52–63. [CrossRef]

28. Khojasteh, P.; Aliahmad, B.; Kumar, D. Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms. *BMC Ophthalmol.* 2018, 18. [CrossRef] [PubMed]

29. Yan, Z.; Han, X.; Wang, C.; Qiu, Y.; Xiong, Z.; Cui, S. Learning Mutually Local-Global U-Nets For High-Resolution Retinal Lesion Segmentation In Fundus Images. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 597–600.

30. Yan, F.; Cui, J.; Wang, Y.; Liu, H.; Liu, H.; Wei, B.; Yin, Y.; Zheng, Y. Deep Random Walk for Drusen Segmentation from Fundus Images. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*; Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 48–55.

31. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 833–851.

32. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.

33. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

34. Jiang, Y.; Wang, F.; Gao, J.; Cao, S. Multi-Path Recurrent U-Net Segmentation of Retinal Fundus Image. *Appl. Sci.* 2020, 10, 3777. [CrossRef]

35. Tan, J.H.; Fujita, H.; Sivaprasad, S.; Bhandary, S.V.; Rao, A.K.; Chua, K.C.; Acharya, U.R. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Inf. Sci.* 2017, 420, 66–76. [CrossRef]

36. Eelbode, T.; Bertels, J.; Berman, M.; Vandermeulen, D.; Maes, F.; Bisschops, R.; Blaschko, M.B. Optimization for Medical Image Segmentation: Theory and Practice When Evaluating with Dice Score or Jaccard Index; IEEE: Piscataway, NJ, USA, 2020.