

Article

nmODE-Unet: A Novel Network for Semantic Segmentation of Medical Images

Shubin Wang, Yuanyuan Chen *  and Zhang Yi

Intelligent Interdisciplinary Research Center and College of Computer Science, Sichuan University, Chengdu 610065, China; wangshubin@stu.scu.edu.cn (S.W.); zhangyi@scu.edu.cn (Z.Y.)

* Correspondence: chenyuanyuan@scu.edu.cn

Abstract: Diabetic retinopathy is a prevalent eye disease that poses a potential risk of blindness. Nevertheless, due to the small size of diabetic retinopathy lesions and the high interclass similarity in terms of location, color, and shape among different lesions, the segmentation task is highly challenging. To address these issues, we proposed a novel framework named nmODE-Unet, which is based on the nmODE (neural memory Ordinary Differential Equation) block and U-net backbone. In nmODE-Unet, the shallow features serve as input to the nmODE block, and the output of the nmODE block is fused with the corresponding deep features. Extensive experiments were conducted on the IDRiD dataset, e_ophtha dataset, and the LGG segmentation dataset, and the results demonstrate that, in comparison to other competing models, nmODE-Unet showcases a superior performance.

Keywords: deep neural networks; deep learning; diabetic retinopathy; semantic segmentation

1. Introduction

Diabetic retinopathy (DR) is a common and potentially blinding eye disease and is one of the leading causes of blindness in adults in developed countries. Early diagnosis and treatment are crucial in preventing the progression of diabetic retinopathy. Many healthcare institutions have successfully reduced the visual impairment caused by DR and lowered its blindness rate through the promotion of early screening and timely medical interventions [1,2]. However, the rapid increase in the number of DR patients still poses new challenges to public health and healthcare systems.

The main lesions associated with DR include hard exudates (EXs), soft exudates (SEs), hemorrhages (HEs), and microaneurysms (MAs). EXs are lipid deposits that remain after the gradual absorption of extravasated plasma substances due to abnormal permeability of the retinal capillary walls. Typically, EXs manifest as yellow-white or white spots or patches with clear boundaries [3,4]. SEs are a result of occlusion and damage to retinal microvessels, causing severe ischemia and hypoxia in the nourished tissue. Typically, SEs manifest as faint yellow or off-white cotton-like patches with blurred edges [5–7]. HEs are mainly located in the deep layers of the retina and can appear as dot-like or patch-like formations, resulting from damaged small blood vessels. Typically, HEs manifest as red, dark-red, or deep-colored spots or patches, varying in size and shape [8,9]. MAs are a typical early sign of DR, characterized by localized dilations and bulges of the retinal small blood vessel walls. Typically, MAs manifest as small red or orange circular or elliptical structures, commonly distributed in the central area of the retina [10–13]. Based on the above-mentioned features, it becomes evident that the color and shape of these lesions in retinal images can easily lead to confusion. For instance, EXs and HEs share similar colors, and HEs and MAs display close resemblances to each other in both color and morphology. MAs, which appear on the vessel walls, also have a color resembling the vessels themselves. Presently, the mainstream method for diagnosing DR still relies on ophthalmologists manually examining fundus images of patients. However, due to the



Citation: Wang, S.; Chen, Y.; Yi, Z. nmODE-Unet: A Novel Network for Semantic Segmentation of Medical Images. *Appl. Sci.* **2024**, *14*, 411. <https://doi.org/10.3390/app14010411>

Academic Editor: João M. F. Rodrigues

Received: 30 November 2023

Revised: 22 December 2023

Accepted: 29 December 2023

Published: 2 January 2024



Copyright: © 2024 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/>).

complexity and diversity of DR in retinal images, several issues arise when relying solely on manual diagnosis by eye care professionals. Firstly, manual diagnosis is susceptible to the influence of subjective judgments by doctors, and the diagnostic outcomes are constrained by their experience and expertise. Secondly, manual diagnosis is prone to misdiagnosis and oversights, particularly concerning subtle lesions that might be overlooked or misjudged. Lastly, manual diagnosis struggles to keep up with the demands of large-scale image data analysis and diagnostic requirements. Therefore, it is urgent and necessary to develop an automated detection method that utilizes computer technology to assist ophthalmologists with clinical diagnosis.

To address the challenging task of segmenting DR in fundus images and to develop an accurate segmentation method, we proposed a model named nmODE-Unet, which is based on nmODE and U-net backbone. The contributions were summarized as follows:

- (1) We proposed an nmODE to address the difficulty of segmenting DR lesions in fundus images. The nmODE has a unique global attractor, which exhibits strong resistance to noise, facilitating the extraction of lesion features.
- (2) We have constructed a new segmentation framework named nmODE-Unet. Experimental results on two different datasets demonstrate that nmODE-Unet exhibits an excellent performance in DR lesion segmentation tasks and displays strong robustness.

2. Related Works

2.1. Deep Neural Networks in DR Lesion Segmentations

In the past decade, some studies focusing on the recognition of lesions in DR fundus images have shown promising results.

Xue et al. [14] proposed a CNN architecture called “Deep Membrane Systems”, which incorporates an encoder and a decoder. This system utilizes FPN [15] and ResNet101 [16] as the backbone for Mask R-CNN [17], and employs two distinct branches for detection and segmentation. It combines both semantic segmentation and edge detection tasks. “Deep Membrane Systems” combine the advantages of multiple network structures and can effectively identify lesions of various sizes and shapes in DR fundus images. Experimental results from three challenging DR image datasets have demonstrated the excellent performance and robustness of this method. Guo et al. [18] proposed L-Seg, which utilizes a multi-scale feature fusion approach. L-Seg employs VGG16 [19] for feature learning. In L-Seg, a side feature extraction module is connected after each group of convolution layers, and feature maps are combined through weighted fusion. Otherwise, they proposed an improved multi-channel loss function to address the imbalance between the background and lesions in DR fundus images. The experiments on three public datasets have demonstrated the effectiveness of L-Seg combined with the improved multi-channel loss in DR segmentation tasks. Liu et al. [20] proposed a feature reassembly method named M2MRF. M2MRF effectively mitigates the potential loss of information about small lesions during the downsampling process. They replaced the bilinear interpolation and repeated stride convolution layers in HRnetV2 [21] with M2MRF to create a variant of HRNetV2. Experimental results demonstrate that the HRnetV2 variant utilizing M2MRF exhibits superior performance and generalization capabilities. The experiments on three public datasets have demonstrated the effectiveness of L-Seg combined with the improved multi-channel loss in DR segmentation tasks.

While CNN-based methods have made significant advancements in image segmentation, the task of segmenting DR lesions in fundus images still remains challenging.

2.2. Neural Ordinary Differential Equation

In 2018, Chen et al. [22] introduced the concept of the Neural Ordinary Differential Equation (neuralODE), treating neural networks as approximations of ODEs and enabling continuous modeling of neural network. Since then, research on neuralODE has been consistently expanding. Poli et al. [23] introduced an innovative approach that combines Graph Neural Networks (GNNs) with ODEs, enabling the modeling of dynamic evolution

in graph data within continuous time. By incorporating ODEs into graph data modeling, Graph Neural ODEs provide a more flexible and powerful modeling tool for adapting to the dynamic nature of graph data. It offers a completely new approach to handling dynamic graph data. Li et al. [24] introduced an innovative approach that is based on discretization techniques within stochastic differential equations (SDEs) and employs gradient estimation to enhance computational efficiency. This method significantly improves the gradient computation efficiency of SDEs, thereby accelerating their application in machine learning and other fields. Yi [25] proposed a novel neuralODE model named nmODE. Regardless of external input, nmODE possesses a unique global attractor, thus embedding a memory mechanism within it. nmODE demonstrates universality in image classification or segmentation tasks, laying the foundation for practical applications of neuralODE.

In this paper, we proposed a network structure named nmODE-Unet for the segmentation of DR lesions in fundus images.

3. Methodology

3.1. nmODE

For a single-layer neural network, it is sufficient to multiply its input by the weight and add the bias term to obtain the neuron's output. It is typically expressed in the following functional form:

$$y = f[Wx + b]. \quad (1)$$

In general, such a single-layer neural network can only address linearly separable problems, and it is unable to capture nonlinear relationships, making it unsuitable for handling complex data patterns or tasks. In general, the computational capacity of neural networks tends to increase with the improvement of their nonlinearity. The nonlinearity of Equation (1) can be enhanced by constructing an implicit mapping equation, which takes the following form:

$$y = f[y + Wx + b]. \quad (2)$$

Yi [25] has already demonstrated that using $\sin^2(\cdot)$ as the activation function in a single-layer neural network effectively enhances the network's expressive power, and addressing nonlinear problems such as the XOR problem. When employing $\sin^2(\cdot)$ as the activation function in Equation (2), the following expression is obtained:

$$y = \sin^2[y + Wx + b]. \quad (3)$$

Yi [25] has already demonstrated that Equation (3) has one and only one global attractor. Let $\dot{y} = 0$, Equation (3) can be transformed into an ODE, referred to as the "nmODE", which we incorporate as a block in the neural network. Its expression is as follows:

$$\dot{y} = -y + \sin^2[y + Wx + b]. \quad (4)$$

Clearly, the nmODE exhibits a unique global attractor, a property intimately associated with memory within dynamic systems. Throughout the training process, the nmODE effectively decouples the initial values and external inputs, where x represents the external input, and the initial value $y(0)$ is set to 0. This separation of initial values and external inputs endows the neural network with two distinct types of neurons: learning neurons and memory neurons. The primary function of learning neurons is to acquire learnable parameters, while memory neurons obtain features through ODEs. Any neuron representing the network's state at any time t , referred to as $y(t)$, is recognized as a memory neuron. Figure 1 provides an intuitive illustration of the nmODE.

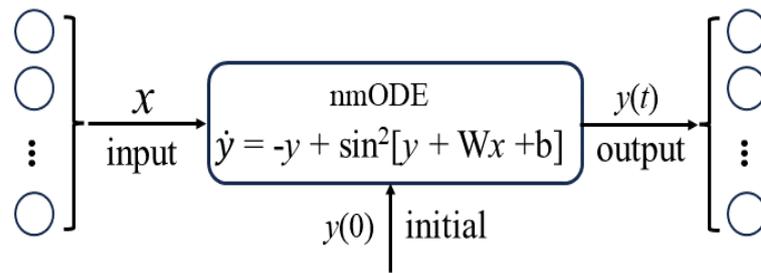


Figure 1. Intuitive diagram of the nmODE.

3.2. Network Architecture

As is widely known, U-Net [26] has demonstrated an outstanding performance in medical image segmentation. It is commonly utilized as a benchmark test model for semantic segmentation tasks.

We proposed a network architecture named nmODE-Unet for segmenting DR lesions, as shown in Figure 2. Within this proposed network architecture, we have made minor modifications to the U-Net. Specifically, we inserted an nmODE block into each skip connection of the U-Net backbone. The shallow-level features serve as inputs to the nmODE blocks, and the outputs of the nmODE blocks replace the shallow-level features, engaging in feature fusion with corresponding deep-level features.

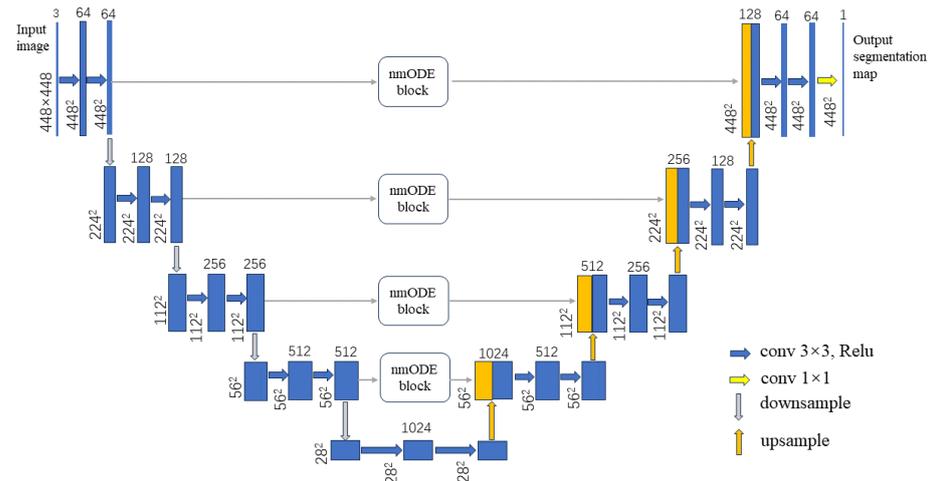


Figure 2. The architecture of the network.

4. Experiments and Results

4.1. Dataset

The experiments were conducted using the IDRiD dataset [27], the e-optha dataset [28], and the LGG segmentation dataset [29,30], all of which contain ground truth annotations by medical professionals. In the IDRiD dataset, 81 fundus images were meticulously annotated at the pixel level. Following the data split outlined in the IDRiD competition, we used 54 images for training and 27 images for validation. The e-optha dataset comprises two types of lesions: EX and MA, with 47 images of EX lesions and 148 images of MA lesions. We utilized 31 images of EX lesions and 98 images of MA lesions for training, and 16 images of EX lesions and 50 images of MA lesions for validation. The LGG segmentation dataset includes 1319 brain MRI images. Out of these, 1000 images were used for training, and the remaining 319 images were used for testing.

To mitigate the risk of overfitting caused by the small dataset size, we employed diverse data augmentation techniques. In practice, we applied horizontal flips, vertical flips, and rotations of 90°, 180°, and 270° to augment the data. Furthermore, to reduce

computational overhead, all images were resized to a standardized size of 448×448 , which served as the input to the network.

4.2. Experimental Setup

This method was implemented using PyTorch 1.3.0 on an NVIDIA Tesla k40m GPU. In the experiments, all models were trained using the Adam optimizer for 2×10^3 iterations, with an initial learning rate of 10^{-3} . The Dice loss was used as the loss function. Each fundus image was resized to $448 \times 448 \times 3$ and was used as input for the models. Various evaluation metrics were calculated, including the Dice, IoU, ROC-AUC, and AUPR.

4.3. Results

4.3.1. Comparison with Baseline Models

We conducted experiments on several high-performing segmentation models, including FCN [31], DeepLab v3+ [32], U-net [26], and Unet++ [33], and used the results of these models as baselines.

The dice score, IoU score, ROC-AUC score, and AUPR score of nmODE-Unet and the baseline models on the IDRiD dataset are summarized in Tables 1 and 2. The ROC curves and the PR curves are shown in Figures 3 and 4, respectively. In addition, the binary segmentation images on the IDRiD dataset are visualized in Figure 5.

As observed from Tables 1 and 2 and Figure 5, FCN excels in recognizing EX and SE lesions, but its segmentation results for HEs and MAs are not satisfactory. DeepLab v3+ excels in recognizing EXs and SEs, but performs poorly in its segmentation results for MAs. Unet++ demonstrates a strong overall performance in recognizing SE, HE, and MA lesions. Unet performs well in recognizing EX and HE lesions, but its performance in recognizing SE is notably much lower compared to other baseline models. Overall, the baseline models are effective in identifying larger lesions, but encounter difficulties in segmenting very small lesions like MAs.

Compared to the baseline models, the nmODE-Unet not only performs remarkably in identifying larger lesions, but also significantly improves the segmentation performance for smaller lesions like MAs. In terms of evaluation metrics, nmODE-Unet achieved higher dice scores by 7.45%, 11.20%, 1.41%, and 1.75%, higher IoU scores by 5.79%, 3.95%, 2.33%, and 1.52%, and higher AUPR scores by 6.67%, 6.02%, 5.23%, and 8.10% in the segmentation of EXs, SEs, HEs, and MAs, respectively, when compared to the second-place model. Regarding ROC-AUC scores, it obtained higher scores by 1.51%, 0.06%, and 4.47% in the segmentation of EXs, SEs, and MAs, respectively. Although the ROC-AUC score of the U-net model in HE segmentation is higher than that of nmODE-Unet by 0.55%, it is evident that its performance is notably insufficient in comparison to nmODE-Unet across other indicators.

Table 1. The segmentation results on the IDRiD dataset: Dice and IoU.

Model	Dice				IoU					
	EX	SE	HE	MA	mDice	EX	SE	HE	MA	mIoU
FCN [31]	0.6304	0.5835	0.4952	0.3805	0.5224	0.5055	0.4623	0.3522	0.2379	0.3898
DeepLab v3+ [32]	0.6741	0.5709	0.5394	0.4155	0.5500	0.5845	0.4977	0.3882	0.2813	0.4379
U-net [26]	0.6770	0.4645	0.5620	0.4376	0.5353	0.5654	0.3696	0.4205	0.2845	0.4100
Unet++ [33]	0.6552	0.5861	0.5896	0.5087	0.5849	0.5520	0.4624	0.4240	0.3465	0.4462
nmODE-Unet	0.7515	0.6981	0.6037	0.5262	0.6449	0.6424	0.5372	0.4473	0.3617	0.4972

Table 2. The segmentation results on the IDRiD dataset: ROC-AUC and AUPR.

Model	ROC-AUC					AUPR				
	EX	SE	HE	MA	mROC-AUC	EX	SE	HE	MA	mAUPR
FCN [31]	0.8954	0.9584	0.8150	0.7571	0.8561	0.6624	0.6658	0.4885	0.3255	0.5356
DeepLab v3+ [32]	0.9415	0.9030	0.8724	0.8291	0.8865	0.7694	0.6779	0.5417	0.3200	0.5773
U-net [26]	0.9374	0.8005	0.8964	0.8948	0.8823	0.7620	0.4350	0.5739	0.3803	0.5334
Unet++ [33]	0.9473	0.9089	0.8638	0.8506	0.8927	0.7610	0.6103	0.5670	0.4367	0.5938
nmODE-Unet	0.9624	0.9590	0.8909	0.9395	0.9380	0.8361	0.7381	0.6262	0.5177	0.6795

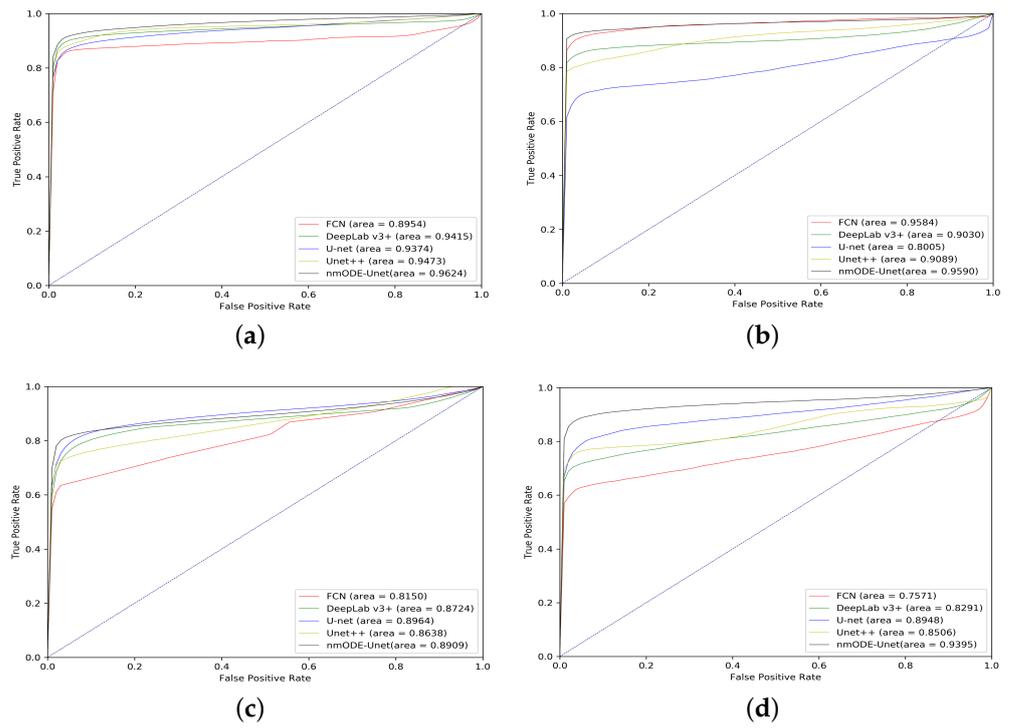


Figure 3. (a–d): The ROC curves for EXs, SEs, HEs, and MAs on the IDRiD dataset.

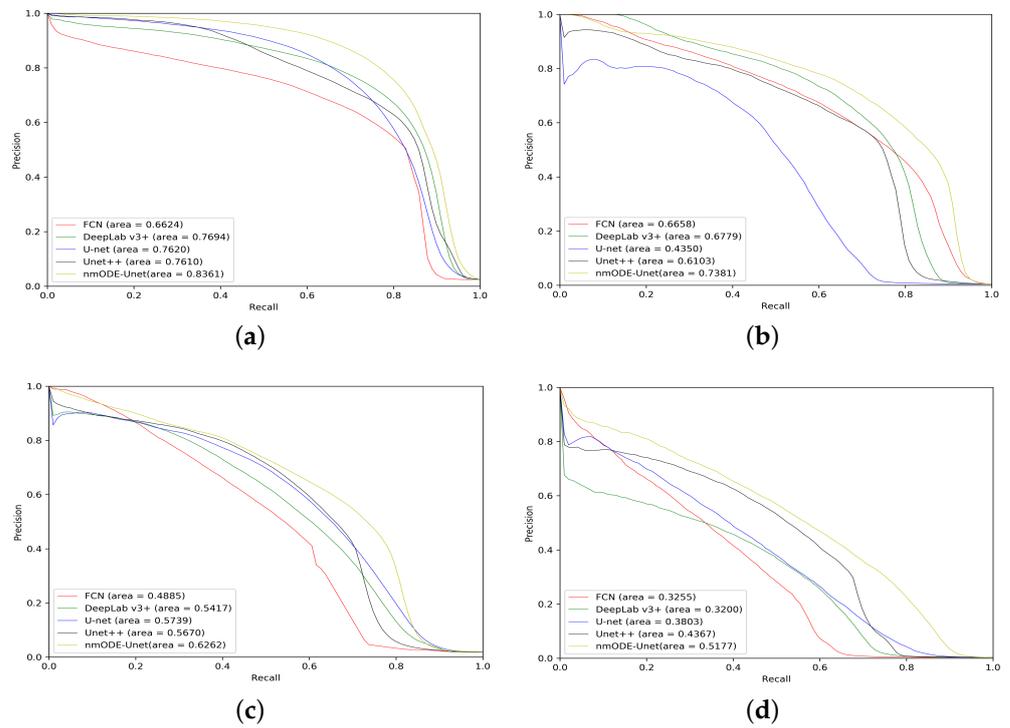


Figure 4. (a–d): The PR curves for EXs, SEs, HEs, and MAs on the IDRiD dataset.

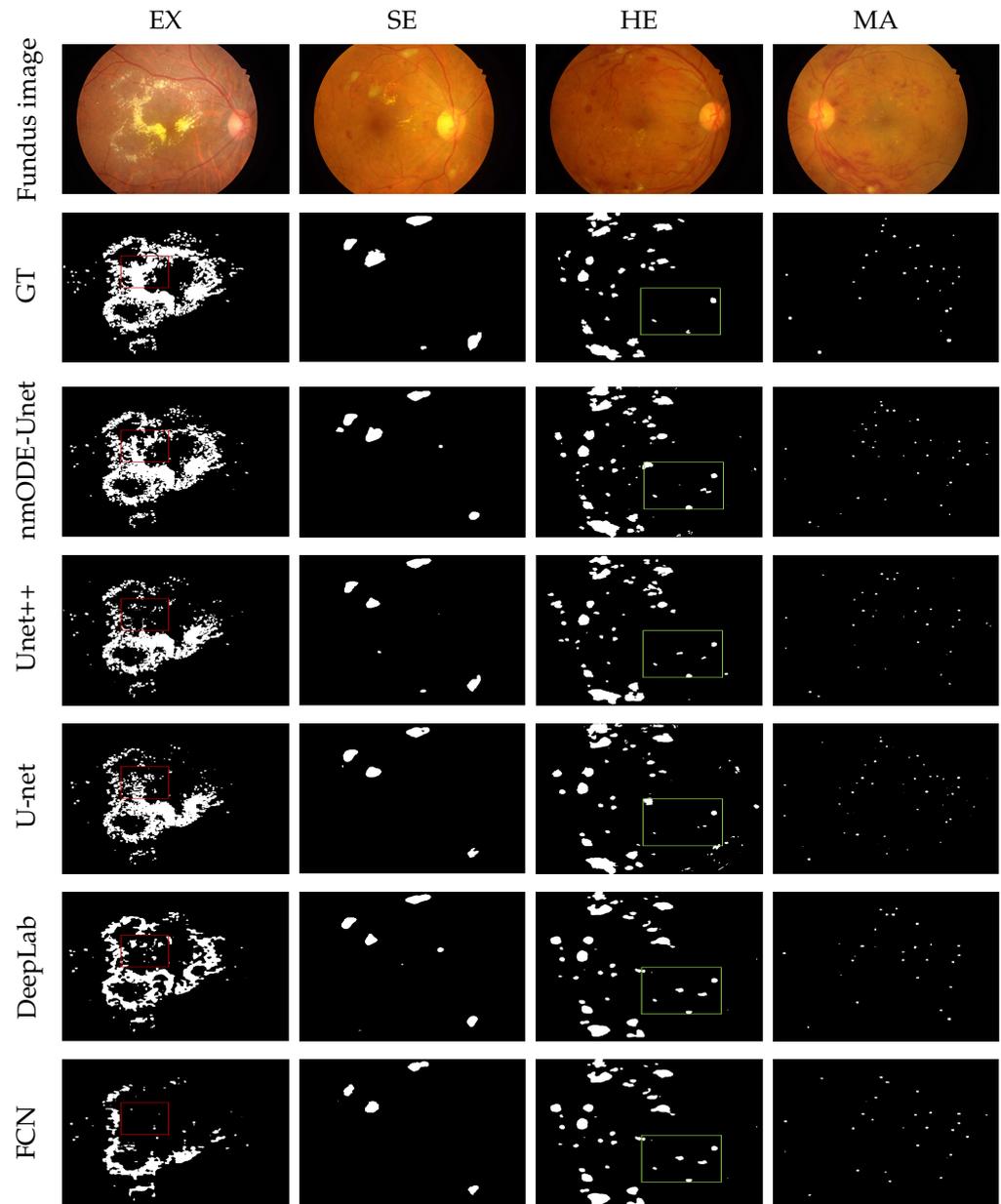


Figure 5. The binary segmentation results of four kinds of lesions. The first row shows the color fundus images, the second row shows the ground truth. The text on the left side from the third to the seventh row indicates the models used, and the corresponding segmentation results of each model are shown on the right side.

From Figure 5, it can be observed that, for the segmentation of EXs, the lesions within the red box in the ground truth have a large area and appear as blocks. The baseline model exhibits poor segmentation in this area, identifying lesions as dots and showing a higher rate of false negatives. For the segmentation of HEs, the lesions within the green box in the ground truth consist of only three patches. However, all models segment more lesions in this area, indicating a noticeable presence of false positives. Overall, the segmentation performance of nmODE-Unet is significantly superior to that of the baseline model.

We also conducted experiments on the e-ophtha dataset. The Dice score, IoU score, ROC-AUC score, and AUPR score of nmODE-Unet and the baseline models on the e-ophtha dataset are summarized in Tables 3 and 4. The ROC curves and the PR curves on the e-ophtha dataset are shown in Figures 6 and 7, respectively. From Tables 3 and 4, as well as Figures 6 and 7, it can be observed that nmODE-Unet outperforms the baseline models in

all evaluation metrics. Furthermore, when comparing the segmentation results with those from the IDRiD dataset, it becomes evident that all methods show a significant decrease in segmentation performance on the e-ophtha dataset. However, the extent of the decline in segmentation results for nmODE-Unet is notably lower than that of the baseline models, particularly in the case of MA lesions. The baseline models exhibited reductions in Dice score and AUPR score ranging from 0.1225 to 0.2573 and from 0.1823 to 0.2934, respectively, while nmODE-Unet demonstrated decreases of 0.0930 and 0.1282, respectively.

Table 3. The segmentation results on the e-ophtha dataset: Dice and IoU.

Model	Dice			IoU		
	EX	MA	mDice	EX	MA	mIoU
FCN [31]	0.3841	0.2580	0.3211	0.2451	0.1480	0.1930
DeepLab v3+ [32]	0.5402	0.2033	0.3718	0.3633	0.1148	0.2391
U-net [26]	0.5157	0.2062	0.3610	0.3412	0.1182	0.2438
Unet++ [33]	0.5626	0.2514	0.4070	0.3764	0.1464	0.2614
nmODE-Unet	0.6659	0.4332	0.5496	0.5064	0.2774	0.3919

Table 4. The segmentation results on the e-ophtha dataset: ROC-AUC and AUPR.

Model	ROC-AUC			AUPR		
	EX	MA	mROC-AUC	EX	MA	mAUPR
FCN [31]	0.8648	0.5772	0.7120	0.4382	0.1137	0.2760
DeepLab v3+ [32]	0.7321	0.7383	0.7352	0.4963	0.1399	0.3181
U-net [26]	0.8163	0.8710	0.8437	0.4909	0.1246	0.3078
Unet++ [33]	0.8402	0.8205	0.8304	0.5021	0.1433	0.3227
nmODE-Unet	0.9251	0.9153	0.9202	0.7121	0.3895	0.5508

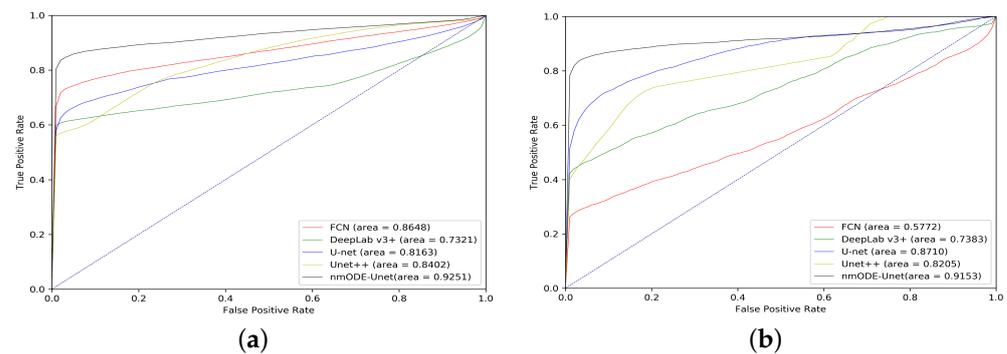


Figure 6. (a,b): The ROC curves for EXs and MAs on the e-ophtha dataset.

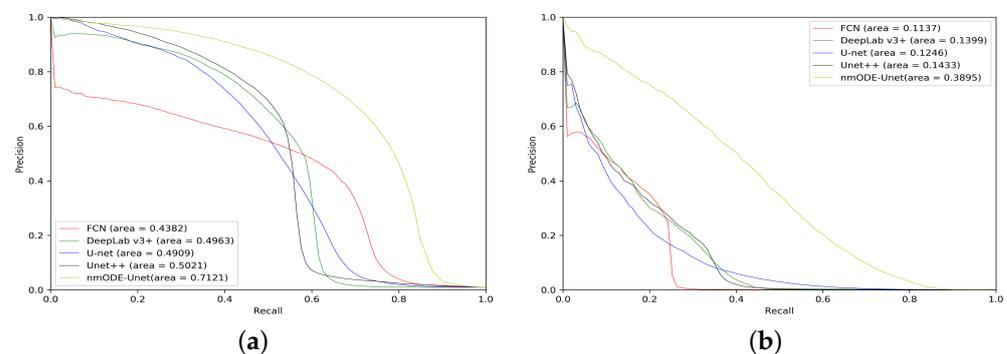


Figure 7. (a,b): The PR curves for EXs and MAs on the e-ophtha dataset.

To assess the versatility and robustness of nmODE-Unet, experiments were conducted using the LGG segmentation dataset. The results on the LGG segmentation dataset are summarized in Table 5. From Table 5, it can be observed that nmODE-Unet outperforms the baseline models. In terms of evaluation metrics, nmODE-Unet outperformed the second-place model with higher Dice scores, IoU scores, ROC-AUC scores, and AUPR

scores by 1.18%, 2.56%, 0.49%, and 2.37%, respectively. Clearly, nmODE-Unet is applicable to different types of medical images and demonstrates excellent robustness.

Table 5. The segmentation results on the LGG segmentation dataset.

Model	Dice	IoU	ROC-AUC	AUPR
FCN [31]	0.7591	0.6066	0.8536	0.6977
DeepLab v3+ [32]	0.8228	0.7005	0.9534	0.8615
U-net [26]	0.8057	0.6748	0.8988	0.8024
Unet++ [33]	0.8352	0.7185	0.9188	0.8028
nmODE-Unet	0.8470	0.7441	0.9583	0.8825

4.3.2. Comparison with IDRiD Challenge Teams

The comparison of AUPR scores between nmODE-Unet and the top ten teams in the IDRiD competition is shown in Table 6. It can be observed that there is a significant variation in the results among these competing teams. The second-ranked team, PATech, did not complete the SE segmentation task, but they achieved the highest score in EX segmentation and demonstrated an excellent performance in HE and MA segmentation. Compared to these competing teams, nmODE-Unet ranks No. 4 in EX segmentation, No. 1 in SE segmentation, No. 3 in HE segmentation, No. 1 in MA segmentation, and achieves the highest average AUPR score among the four kinds of lesions. Clearly, nmODE-Unet exhibits significant competitiveness when compared to these competing teams.

Table 6. Comparison with top ten teams in IDRiD challenge.

Team	EX	SE	HE	MA	mAUPR
VRT (1st)	0.7127	0.6995	0.6804	0.4951	0.6469
PATech (2nd)	0.8850	-	0.6490	0.4740	-
iFLYTEK-MIG (3rd)	0.8471	0.6588	0.5588	0.5017	0.6484
SOONER (4th)	0.7390	0.5396	0.5395	0.4003	0.5539
SAIHST (5th)	0.8582	-	-	-	-
lzyuncc_fusion (6th)	0.8202	0.6259	-	-	-
SDNU (7th)	0.5018	0.5374	0.4572	0.4111	0.4769
CIL (8th)	0.7554	0.5024	0.4886	0.3920	0.5346
MedLabs (9th)	0.7863	0.2637	0.3705	0.3397	0.4401
AIMIA (10th)	0.7662	0.2733	0.3283	0.4627	0.4367
nmODE-Unet	0.8361	0.7381	0.6262	0.5177	0.6795

4.3.3. Comparison with State-of-the-Art Segmentation Methods

Using AUPR as the evaluation metric, the comparison of the performance for four kinds of lesions between nmODE-Unet and state-of-the-art methods is shown in Table 6.

Through a detailed analysis of the data in Table 7, we can obtain the following results: (1) compared to [18], the AUPR score of nmODE-Unet is 8.66%, 2.68%, and 5.50% higher in the segmentation of EXs, SEs, and MAs, respectively, but 1.12% lower in HE segmentation; (2) compared to [20], the AUPR score of nmODE-Unet is 1.45%, 4.49%, and 2.97% higher in the segmentation of EXs, SEs, and MAs, respectively, but 6.07% lower in HE segmentation; (3) compared to [34], the AUPR score of nmODE-Unet is 4.51% and 4.42% lower in the segmentation of EXs and HEs, but 1.00% and 10.25% higher in the segmentation of SEs and MAs, respectively; (4) compared to [35], the AUPR score of nmODE-Unet is 3.14% and 1.27% lower in the segmentation of EXs and HEs, but 2.56% and 0.29% higher in the segmentation of SEs and MAs, respectively; (5) the method in [20] achieved the highest value in HE segmentation, the method in [34] achieved the highest value in EX segmentation, the method in [35] achieved the highest value in mAUPR, and nmODE-Unet achieved the highest value in SE and MA segmentation, respectively.

Table 7. Comparison with other state-of-the-art methods.

Model	EX	SE	HE	MA	mAUPR
L-seg [18]	0.7495	0.7113	0.6374	0.4627	0.6515
M2MRF [20]	0.8216	0.6932	0.6869	0.4880	0.6724
SAA [34]	0.8812	0.7281	0.6704	0.4152	0.6738
CARNet [35]	0.8675	0.7125	0.6389	0.5148	0.6834
nmODE-Unet	0.8361	0.7381	0.6262	0.5177	0.6795

4.4. Discussion

Due to the small size of DR lesions and the high interclass similarity in terms of location, color, and shape among different lesions, the segmentation task is highly challenging. To further enhance the accuracy of DR lesion segmentation, we proposed a novel framework, nmODE-Unet, and conducted comparative experiments with commonly used image segmentation networks and other state-of-the-art methods. The experimental results indicate that our approach is competitive. Particularly, the segmentation of SE and MA lesions on the IDRiD dataset has achieved the highest values we are aware of.

Due to the poor image quality and noise interference in the e-optha dataset, all methods experienced a significant decrease in segmentation accuracy, particularly in the case of MA lesions. MA lesions are extremely small in size, share similarities with retinal blood vessels, and are susceptible to information loss during the downsampling process, exacerbated by noise, leading to mis-segmentation issues. However, while our method also exhibited reduced accuracy, it remained significantly superior to other approaches. As evident from Tables 3 and 4, the segmentation performance of nmODE-Unet surpasses that of other methods by a wide margin, primarily because nmODE possesses a unique global attractor. The presence of this global attractor enhances the network's resistance to noise.

We have measured the FLOPs, parameters, memory usage, and inference time for nmODE-Unet and the baseline model. The results are presented in Table 8. Unet++ and DeepLab v3+ exhibited the highest computational complexity. Unet++ recorded the highest FLOPs, while DeepLab v3+ had the highest parameter count and memory usage. In comparison to the U-Net backbone, nmODE-Unet does not increase FLOPs and parameters. However, both inference time and memory usage are increased due to the iterative process required by the nmODE block for solving ordinary differential equations. Despite the absence of increased FLOPs and parameters, nmODE-Unet exhibits an improved segmentation performance. Nonetheless, the drawback lies in the longer inference time, indicating the need for further improvement in this regard.

Table 8. Computational efficiency and resource requirements.

Model	FLOPs (G)	Parameters (M)	(s)/Batch	Memory (M)
FCN [31]	156.18	18.64	0.01	298.07
DeepLab v3+ [32]	136.21	59.34	0.05	953.50
U-net [26]	181.35	19.20	0.01	304.74
Unet++ [33]	213.78	9.16	0.02	153.98
nmODE-Unet	181.35	19.20	1.13	485.96

5. Conclusions

In this paper, we proposed a novel framework named nmODE-Unet, which is based on the nmODE block and U-net backbone for the segmentation of DR lesions. The performance of nmODE-Unet was evaluated on the IDRiD dataset, the e-optha dataset, and the LGG segmentation dataset. nmODE-Unet exhibited a superior performance over the baseline model across all datasets. On the IDRiD dataset, nmODE-Unet demonstrated a competitive performance compared to state-of-the-art methods and the top ten teams in the IDRiD challenge.

While the nmODE block has demonstrated an excellent performance, its adaptability to a wider range of medical image types remains to be verified. MedSAM [36] exhibits a robust performance across a diverse array of datasets, proving that a single foundation model can handle various segmentation tasks, eliminating the need for task-specific models.

Integrating the nmODE block into more deep learning networks to build foundation models, similar to MedSAM and capable of handling diverse segmentation tasks, is a topic worthy of further research.

Our research goal is to apply nmODE-Unet to clinical diagnosis; we believe that, by training nmODE-Unet with a more diverse set of medical images representing various types, it has tremendous potential to become a valuable tool for clinical diagnostics.

Author Contributions: Conceptualization, S.W., Y.C. and Z.Y.; Methodology, S.W., Y.C. and Z.Y.; Software, S.W., Y.C. and Z.Y.; Validation, S.W.; Formal analysis, S.W., Y.C. and Z.Y.; Investigation, S.W., Y.C. and Z.Y.; Resources, Y.C. and Z.Y.; Data curation, S.W.; Writing—original draft, S.W.; Writing—review & editing, Y.C. and Z.Y.; Visualization, S.W.; Supervision, Y.C. and Z.Y.; Project administration, Y.C. and Z.Y.; Funding acquisition, Y.C. and Z.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 61702349) and National Major Science and Technology Projects of China (Grant No. 2018AAA0100201).

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Ting, D.S.W.; Cheung, G.C.M.; Wong, T.Y. Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review. *Clin. Exp. Ophthalmol.* **2016**, *44*, 260–277. [[CrossRef](#)] [[PubMed](#)]
2. Ciulla, T.A.; Amador, A.G.; Zinman, B. Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies. *Diabetes Care* **2003**, *26*, 2653–2664. [[CrossRef](#)] [[PubMed](#)]
3. Avula Benzamin and Chandan Chakraborty. Detection of hard exudates in retinal fundus images using deep learning. In Proceedings of the 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 25–29 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 465–469.
4. Huang, C.; Zong, Y.; Ding, Y.; Luo, X.; Clawson, K.; Peng, Y. A new deep learning approach for the retinal hard exudates detection based on superpixel multi-feature extraction and patch-based cnn. *Neurocomputing* **2021**, *452*, 521–533. [[CrossRef](#)]
5. Joshi, S.; Karule, P.T. A review on exudates detection methods for diabetic retinopathy. *Biomed. Pharmacother.* **2018**, *97*, 1454–1460. [[CrossRef](#)] [[PubMed](#)]
6. Si, Z.; Fu, D.; Liu, Y.; Huang, Z. Hard exudate segmentation in retinal image with attention mechanism. *IET Image Process.* **2021**, *15*, 587–597. [[CrossRef](#)]
7. Kaur, J.; Kaur, P. Uniconv: An enhanced u-net based inceptionv3 convolutional model for dr semantic segmentation in retinal fundus images. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e7138. [[CrossRef](#)]
8. Sambyal, N.; Saini, P.; Syal, R.; Gupta, V. Modified u-net architecture for semantic segmentation of diabetic retinopathy images. *Biocybern. Biomed. Eng.* **2020**, *40*, 1094–1109. [[CrossRef](#)]
9. Gupta, A.; Chhikara, R. Diabetic retinopathy: Present and past. *Procedia Comput. Sci.* **2018**, *132*, 1432–1440. [[CrossRef](#)]
10. Dai, L.; Fang, R.; Li, H.; Hou, X.; Sheng, B.; Wu, Q.; Jia, W. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. *IEEE Trans. Med. Imaging* **2018**, *37*, 1149–1161. [[CrossRef](#)]
11. Chudzik, P.; Majumdar, S.; Caliva, F.; Al-Diri, B.; Hunter, A. Microaneurysm detection using deep learning and interleaved freezing. In *Medical Imaging 2018: Image Processing*; SPIE: Bellingham, WA, USA, 2018; Volume 10574, pp. 379–387.
12. Sarhan, M.H.; Albarqouni, S.; Yigitsoy, M.; Navab, N.; Eslami, A. Multi-scale microaneurysms segmentation using embedding triplet loss. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 174–182.
13. Perumal, T.S.R.; Jayachandran, A.; Kumar, S.R. Microaneurysms detection in fundus images using local fourier transform and neighbourhood analysis. *Knowl. Inf. Syst.* **2023**, 1–21. [[CrossRef](#)]
14. Xue, J.; Yan, S.; Qu, J.; Qi, F.; Qiu, C.; Zhang, H.; Chen, M.; Liu, T.; Li, D.; Liu, X. Deep membrane systems for multitask segmentation in diabetic retinopathy. *Knowl.-Based Syst.* **2019**, *183*, 104887. [[CrossRef](#)]
15. Lin, T.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
16. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
17. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.

18. 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](#)]
19. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
20. Liu, Q.; Liu, H.; Ke, W.; Liang, Y. Automated lesion segmentation in fundus images with many-to-many reassembly of features. *Pattern Recognit.* **2023**, *136*, 109191. [[CrossRef](#)]
21. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 3349–3364. [[CrossRef](#)] [[PubMed](#)]
22. Chen, R.T.Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D.K. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2018; Volume 31.
23. Poli, M.; Massaroli, S.; Park, J.; Yamashita, A.; Asama, H.; Park, J. Graph neural ordinary differential equations. *arXiv* **2019**, arXiv:1911.07532.
24. Li, X.; Wong, T.L.; Chen, R.T.Q.; Duvenaud, D. Scalable gradients for stochastic differential equations. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Palermo, Italy, 26–28 August 2020; pp. 3870–3882.
25. Yi, Z. nmode: Neural memory ordinary differential equation. *Artif. Intell. Rev.* **2023**, *56*, 14403–14438. [[CrossRef](#)]
26. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18. Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
27. Porwal, P.; Pachade, S.; Kamble, R.; Kokare, M.; Deshmukh, G.; Sahasrabudhe, V.; Meriaudeau, F. Indian diabetic retinopathy image dataset (idrid): A database for diabetic retinopathy screening research. *Data* **2018**, *3*, 25. [[CrossRef](#)]
28. Decencière, E.; Cazuguel, G.; Zhang, X.; Thibault, G.; Klein, J.-C.; Meyer, F.; Marcotegui, B.; Quéllec, G.; Lamard, M.; Danno, R.; et al. Teleophta: Machine learning and image processing methods for teleophthalmology. *IRBM* **2013**, *34*, 196–203. [[CrossRef](#)]
29. Mazurowski, M.A.; Clark, K.; Czarnek, N.M.; Shamsesfandabadi, P.; Peters, K.B.; Saha, A. Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with the cancer genome atlas data. *J. Neuro-Oncol.* **2017**, *133*, 27–35. [[CrossRef](#)]
30. Buda, M.; Saha, A.; Mazurowski, M.A. Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. *Comput. Biol. Med.* **2019**, *109*, 218–225. [[CrossRef](#)] [[PubMed](#)]
31. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
32. Chen, Li.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
33. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 20 September 2018; Proceedings 4; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–11.
34. Bo, W.; Li, T.; Liu, X.; Wang, K. Saa: Scale-aware attention block for multi-lesion segmentation of fundus images. In Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), Kolkata, India, 28–31 March 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–5.
35. Guo, Y.; Peng, Y. Carnet: Cascade attentive refinenet for multi-lesion segmentation of diabetic retinopathy images. *Complex Intell. Syst.* **2022**, *8*, 1681–1701. [[CrossRef](#)]
36. Ma, J.; Wang, B. Segment anything in medical images. *arXiv* **2023**, arXiv:2304.12306.

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.