AI CHALLENGER 全球AI挑战赛 2018

眼底水肿病变区域自动分割

参赛队：DeepSeg

队员：王沈　　北京大学&深睿医疗

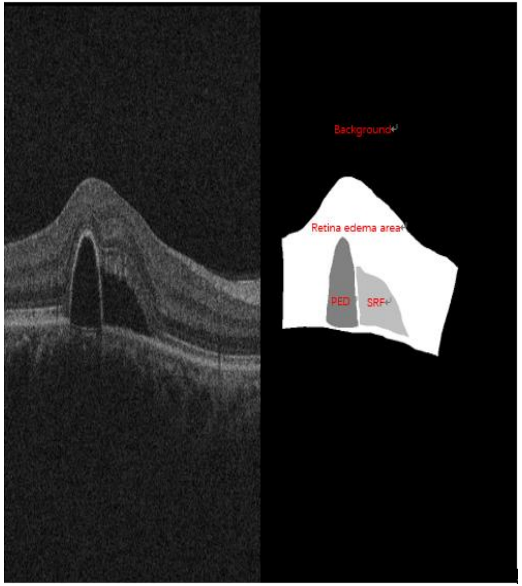
导师：王亦洲教授　　北京大学

CATALOGUE

AI CHALLENGER 全球AI挑战赛 2018

1.1 Background

- Retinal edema is an eye disease that can cause blurry vision and affect normal life.
- OCT (optical coherence tomography) can be used to assist doctors in the judgment of retinal edema.
- Early detection of edema symptoms can play a key role in the treatment of the disease.
- Our task is to design algorithms to automatically detect the type of retinal edema and segment the area of retinal edema based on OCT images.



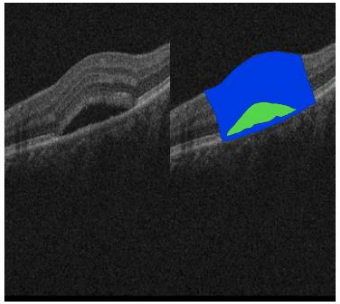
1.2 Data Overview

- Data statistical information

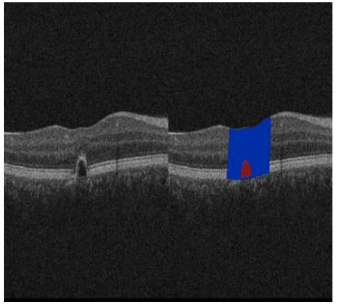
Stage	Mean	Deviation	Shape	Number
train	36.68	25.01	128*512*1024	70
val	36.47	24.71	128*512*1024	15
test	35.73	24.12	128*512*1024	15

Stage	Level	Class 1	Class 2	Class 3
train	patient	70	55	29
	slice	4664	2934	402
	pixel	2.91E+08	3.96E+07	1.29E+06
	size	62355	13502	3223
val	patient	15	8	6
	slice	834	360	41
	pixel	4.79E+07	4.36E+06	2.81E+04
	size	57399	12113	684

- Data visualization



Val : 0012



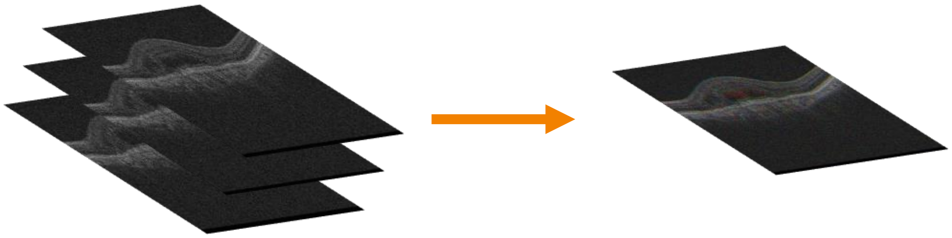
Val : 0014

REA SRF PED

- Edge bending of the retina
- Inclusion relationship between lesions

1.3 Data Preprocess

- Stack the upper and lower slices to form a three-channel input



- Normalization

$$x_i^* = \frac{x_i - \mu_i}{\sigma_i}$$

where μ_i is the mean of the image in channel i

σ_i is the deviation of the image in channel i

x_i is the pixel value of the image position x in channel i

- Data augmentation(Only Random Horizontal Flip)

1.4 Problems and challenges

- How can the two tasks(segmentation and detection) promote each other?
  - Multi-task learning framework
- Multi-scale of the retinal edema lesions
  - UNet and UNet++
- Sample imbalance among three type of retinal edema
  - Exponential logarithmic loss
- How to enlarge receptive field to detect edge bending?
  - Dilated module

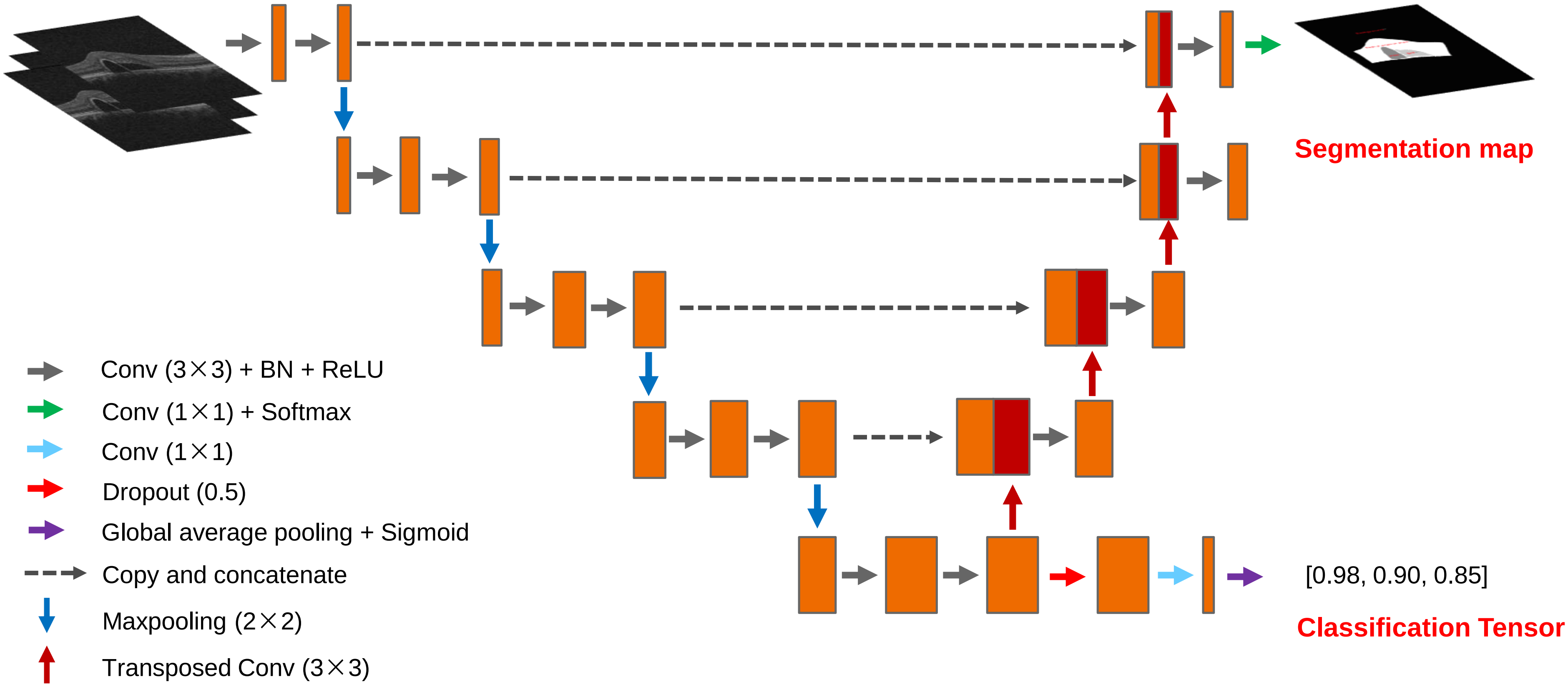
2.1 Baseline

- Segmentation – UNet
  - Downsample 16×
  - Channel : [16,32,64,128,256]
  - Input : original image
- Detection – ResNet18
  - Resize 224*224

Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.9036
ResNet18(nopre*)	No	11.18 M	BCE	-	0.964	-	-
UNet	No	2.47 M	WCE	0.753	-	-	-
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-

** pre : pretrained model on ImageNet   nopre : train model from scratch*

2.2 Multi-task Framework



2.2 Multi-task Framework

AI CHALLENGER 全球AI挑战赛 2018

- Jointly learn the segmentation and detection
- Reduce time and computational cost
- Improve the effects of both tasks

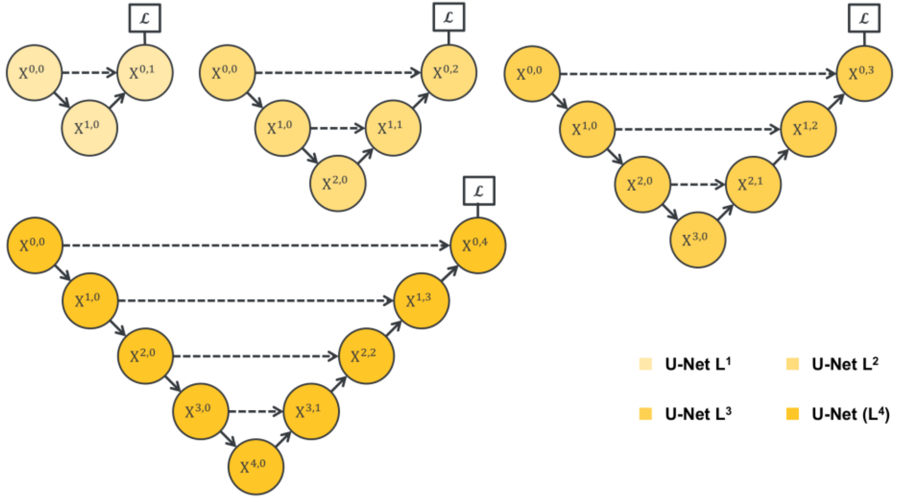
Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.904
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-
UNet	Yes	2.47 M	WCE+Dice +BCE	**0.785 (+0.013)**	**0.985 (+0.014)**	**0.701 (+0.018)**	**0.968 (+0.064)**

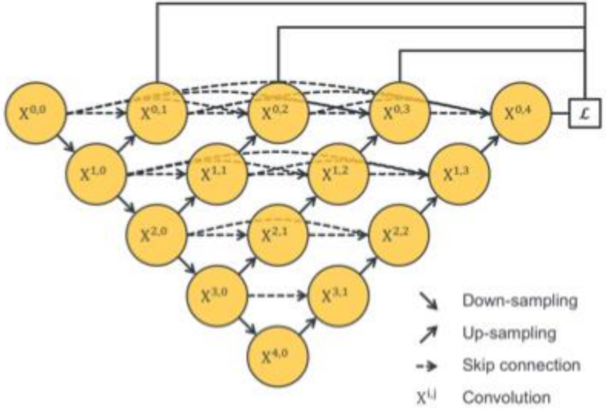
** pre : pretrained model on ImageNet*

2.3 UNet++ & DeepSupervision

- Dense future connection
- Deep Supervision

- Fuse low-level and high-level feature more effectively





Reference : *Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018: 3-11.*

2.3 UNet++ & DeepSupervision

AI CHALLENGER 全球AI挑战赛 2018

Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.904
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-
UNet	Yes	2.47 M	WCE+Dice +BCE	**0.785 (+0.013)**	0.985 (+0.014)	**0.701 (+0.018)**	**0.968 (+0.064)**
UNet++	Yes	2.95 M	WCE+Dice +BCE	0.784 (+0.012)	**0.986 (+0.015)**	-	-

** pre : pretrained model on ImageNet*

2.4 Exponential Logarithmic Loss

$$Loss = L_{Seg} + L_{cls} = L_{ELDice} + L_{CE} + L_{BCE}$$

$$L_{ELDice} = E(f_{\gamma}(Dice_i)) = E((-\ln(Dice_i))^{\gamma}) \quad (\gamma = 0.5)$$

$$Dice_i = \frac{2 \cdot \sum_x \delta_{il}(x) p_i(x) + \varepsilon}{\sum_x (\delta_{il}(x) + p_i(x)) + \varepsilon}$$

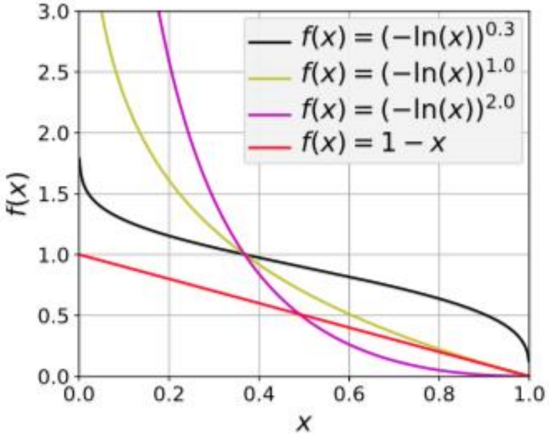


Figure . Loss functions with different nonlinearities, where x can be $Dice_i$

x : pixel position

i : label

l : ground truth label at x

d_{il} : Kronecker delta

e : pseudocount for additive smoothing

$p_i(x)$: Softmax probability which acts as the portion of pixel x owned by label i

Reference : Wong K C L, Moradi M, Tang H, et al. 3d segmentation with exponential logarithmic loss for highly unbalanced object sizes[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2018: 612-619.

2.4 Exponential Logarithmic Loss

AI CHALLENGER 全球AI挑战赛 2018

Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.9036
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-
UNet	Yes	2.47 M	WCE+Dice +BCE	0.785 (+0.013)	0.985 (+0.014)	0.701 (+0.018)	**0.968 (+0.064)**
UNet++	Yes	2.95 M	WCE+Dice +BCE	0.784 (+0.012)	0.986 (+0.015)	-	-
UNet++	Yes	2.95 M	WCE+ ELDice+BCE	**0.799 (+0.027)**	**0.989 (+0.018)**	**0.736 (+0.053)**	-

** pre : pretrained model on ImageNet*

How about small object segmentation?(class 2,3)

Model	Loss	Dice	Dice_1	Dice_2	Dice_3
UNet++	WCE+ Dice+BCE	0.784	**0.761**	0.832	0.760
UNet++	WCE+ ELDice+BCE	**0.799**	0.755	**0.859**	**0.784**

2.5 Dilated Module

AI CHALLENGER 全球AI挑战赛 2018

- Receptive field calculation formula

$$RF_l = RF_{l-1} + (k_l - 1) \times \prod_{i=0}^{l-1} s_i \times d_l,\ l \geq 1,\ RF_0 = 1$$

where RF_l is the receptive field of layer l

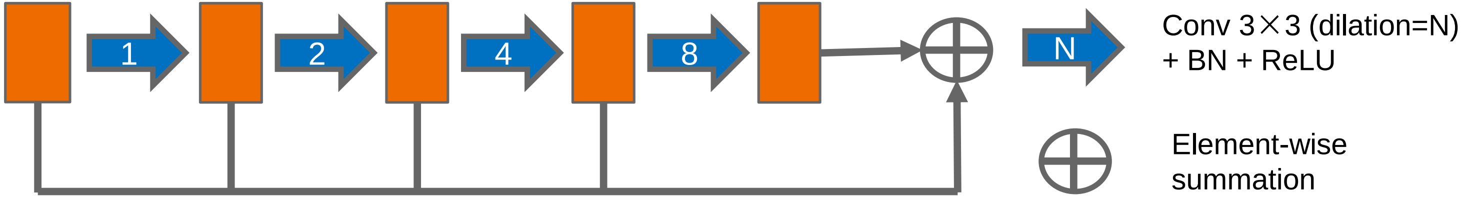
k_l is the kernel size of layer l

s_i is the strider of layer i

d_l is the dilation of layer l

- Receptive field of UNet encoder

Dilated rate	Stage	Receptive Field
-	1	5
-	2	14
-	3	32
-	4	68
-	5	140
1,2,4	**5**	**364**
1,2,4,8	5	**620**



2.5 Dilated Module

AI CHALLENGER 全球AI挑战赛 2018

Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.904
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-
UNet	Yes	2.47 M	WCE+Dice +BCE	0.785 (+0.013)	0.985 (+0.014)	0.701 (+0.018)	0.968 (+0.064)
UNet++	Yes	2.95 M	WCE+Dice +BCE	0.784 (+0.012)	0.986 (+0.015)	-	-
UNet++	Yes	2.95 M	WCE+ ELDice+BCE	0.799 (+0.027)	0.989 (+0.018)	0.736 (+0.053)	-
Dilated Unet++	Yes	5.32 M	WCE+ ELDice+BCE	**0.807 (+0.035)**	0.978 (+0.006)	-	-

** pre : pretrained model on ImageNet*

How about big object segmentation?(class 1)

Model	Loss	Dice	Dice_1	Dice_2	Dice_3
UNet++	WCE+ ELDice+BCE	0.799	0.755	0.859	0.784
Dilated UNet++	WCE+ ELDice+BCE	**0.807**	**0.769**	**0.863**	**0.789**

3.1 Experiment Summary

AI CHALLENGER 全球AI挑战赛 2018

Model	Multi-Task	Params	Loss	Val_Dice	Val_Auc	Test_Dice	Test_Auc
ResNet18(pre*)	No	11.18 M	BCE	-	0.971	-	0.904
UNet	No	2.47 M	WCE+Dice	0.772	-	0.683	-
UNet	Yes	2.47 M	WCE+Dice +BCE	0.785 (+0.013)	0.985 (+0.014)	0.701 (+0.018)	0.968 (+0.064)
UNet++	Yes	2.95 M	WCE+Dice +BCE	0.784 (+0.012)	0.986 (+0.015)	-	-
UNet++	Yes	2.95 M	WCE+ ELDice+BCE	0.799 (+0.027)	0.989 (+0.018)	0.736 (+0.053)	-
Dialted Unet++	Yes	5.32 M	WCE+ ELDice+BCE	**0.807 (+0.035)**	0.978 (+0.006)	-	-
Unet & Unet++ Fusion	-	-	-	0.805 (+0.033)	**0.991 (+0.020)**	**0.744 (+0.061)**	**0.986 (+0.082)**

** pre : pretrained model on ImageNet*

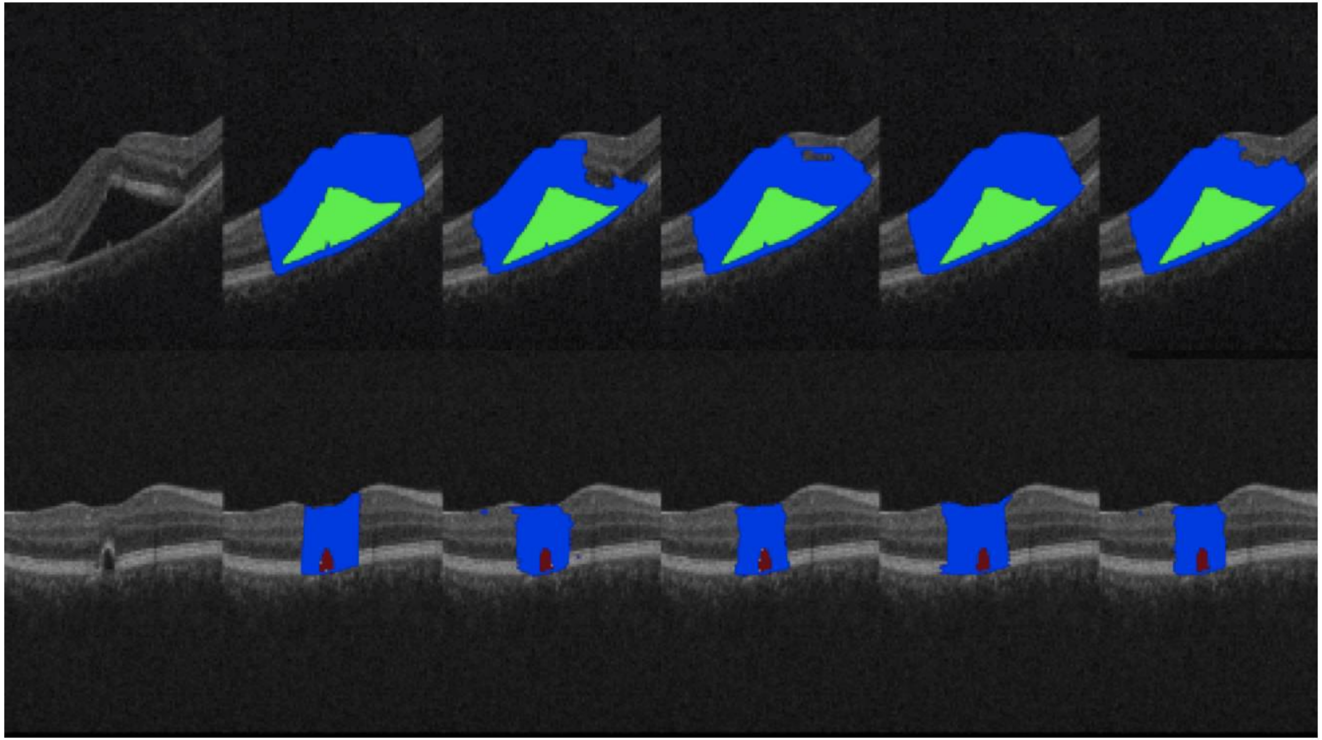
- Memory : 7.3 G(batch=8), Inference time : 9.5 s/patient

3.2 Visualization

Image	Ground Truth	UNet	Unet++	Dilated UNet++	Fusion

Val : 0010

Val : 0014



REA

SRF

PED

3.3 Future Work

- Detection framework (Mask R-CNN)
- 3D semantic segmentation model
- Use Res-block or Dense-block in backbone
- Consider the relation of lesions

3.4 Conclusion

- Build an end-to-end multi-task framework that can simultaneously detect and segment retinal edema lesions.
- Use the latest UNet++ model to better integrate high-level and low-level features.
- Use a novel exponential logarithmic loss function to enhance the segmentation of two types of small lesions.
- Introduce the dilated convolution module, which significantly increases the receptive field of the model.
- Only random horizontal flip data augmentation and no post-processing.
- The dice of single model on the test set is 0.736. The dice of fusion model on the test set is 0.744 and the detection AUC is 0.986. In addition, the memory of inference stage is 7.3G when we set batch is 8 and the inference time is 9.5s per patient.

3.5 Acknowledge

- GPU support of Deepwise
- Mentor, Prof. Wang’s guidance

AI CHALLENGER 全球AI挑战赛 2018

Thank You!

Welcome the questions and advice.

Code:
https://github.com/ShawnBIT/AI-challenger-Retinal-Edema-Segmentation

队名：DeepSeg

队员：王沈　　北京大学&深睿医疗
导师：王亦洲教授　北京大学