Comparison of Capsule Networks and Other Networks for Object Segmentation Tasks

Zhankun Luo, *Member, IEEE*

Abstract—Recently, a new architecture was introduced by Hinton, named Capsule Networks. It shows excellent capability for processing video and classification with derivative-free layers (Capsule layers). A capsule is a group of neurons whose activity vectors represent the parameters of an entity. Hinton uses the length of the activity vector to represent the probability of an entity's existence, and uses its direction to represent the parameters. The first-level active capsule predicts the higher-level capsule through the parameters of the transformation matrix pair. The objective of this study is to compare Capsule networks with conventional networks: UNet and DenseNet for Object Segmentation Tasks. Using TensorFlow or PyTorch library, a python program is implemented to achieve these goals. These models ware tested on Drive dataset: the Retina Vessel dataset. Consequently, it showed that the accuray of SegCaps networks could not compete with that of U-Net and DenseNet on Drive dataset. Further study is needed to apply the SegCaps networks to other datasets to verify the assertion.

Index Terms—Segmentation, U-Net, DenseNet, SegCaps, Capsule.

1 Introduction

In the 1980s, Hinton and LeCun proposed and popularized a back-propagation algorithm that can be used to train multilayer neural networks. But at that time, CNN laid the bane.

The first is that backpropagation algorithms are biologically difficult to establish, and it is difficult to believe that the neural system can automatically form backpropagation structures that correspond to forward propagation (This requires precise derivatives, matrix transposition, and use of chain rules. No anatomical evidence was found for the system). Secondly, the back-propagation algorithm needs to be optimized by SGD and other methods. This is a highly non-convex problem, its mathematical nature is worrying, and it depends on fine-tuning parameters.

An interesting fact noticed by Hinton is that there are a large number of columnar structures (cortical pillars) in the cerebral cortex, with hundreds inside Neurons and many layers inside. A layer in the human brain is not similar to the current NNs, instead has a complex internal structure. Hinton believes that the mini-column must have played a role. Therefore, Hinton also proposed a corresponding structure called Capsule [1].

2 Visualization Method

To investigate the ability of the capsule network to perform object segmentation tasks, we selected the SegCaps network [2] as the comparison object, and its performance in lung segmentation is promising. The architecture of SegCaps is demonstrated in Fig.3.

For comparison, we also applied U-Net[3], DenseNet[4] for the same segmentation tasks. U-Net is well-known for its efficiency in segmentation. In addition, DenseNet is proved to be a useful and effective architecture. These architecture can be seen in Fig.1 and Fig.2.

The differences between Capsule and conventional CNN are listed in the Fig.4.

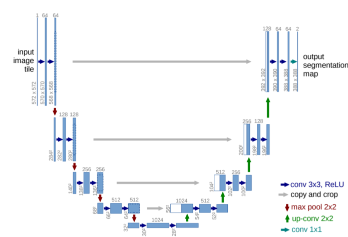


Fig. 1. U-Net Architecture.

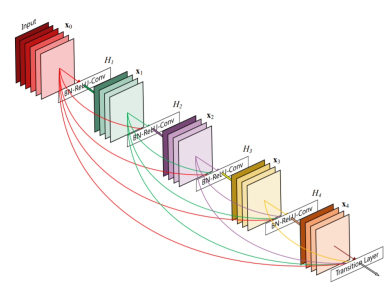


Fig. 2. DenseNet Architecture.

3 Dataset Description

To test the sensibility of networks for segmentation tasks, we chose a complicated segmentation dataset for training and test, named DRIVE: Digital Retinal Images for Vessel Extraction.

Photos from the DRIVE database are from the diabetic

• Z. Luo was with the Department of Electrical and Computer Engineering, Purdue University Northwest, Hammond, IN, 46323.
E-mail: dassein75325@gmail.com

Manuscript received Dec 11, 2019; revised Dec 11, 2019.

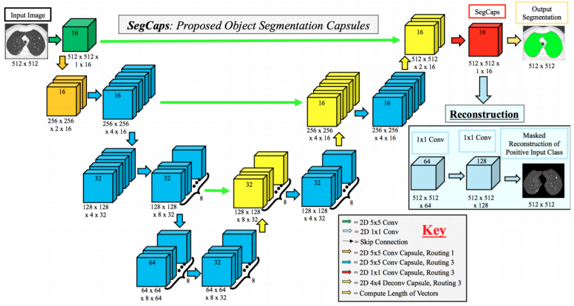


Fig. 3. SegCaps Architecture.

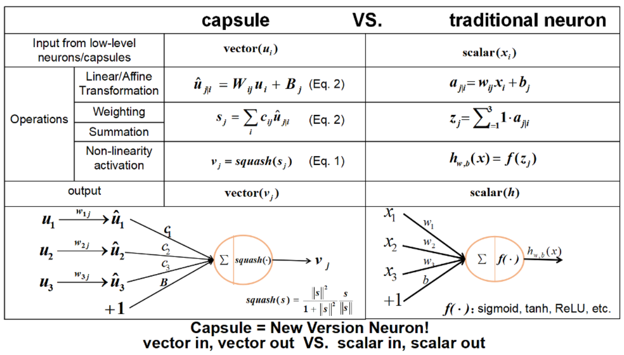


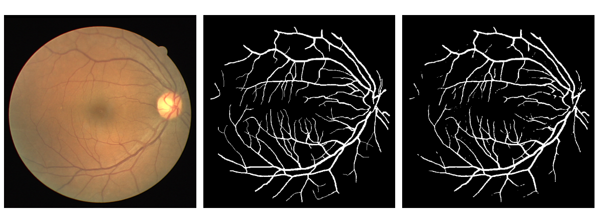
Fig. 4. Differences of Capsule and CNN.

retinopathy screening program in the Netherlands. 40 photos were randomly selected 400 diabetic patients aged 25-90 years.
DRIVE dataset are separated into 2 sub folders, one is Training folder, the other is Test folder, both contain 20 images. For the training images, manual annotation for each image is available. For the test cases, manual segmentations are lacked.

4 Results and Discussion

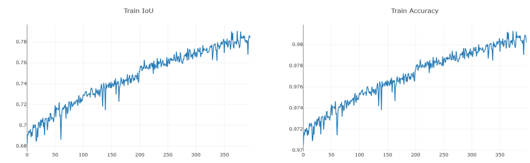
4.1 The U-Net networks:

After training for 350 epochs, the Accuracy reached 98%, IoU(Intersection over Union) reached 80%.



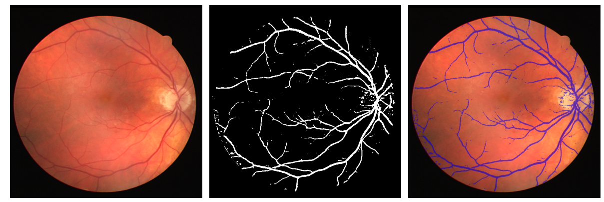
(a) Original Image (b) Target Image (c) U-Net Prediction

Fig. 5. Training, target, predict images in Training data set for U-Net



(a) Training IoU (b) Training Accuracy

Fig. 6. Training IoU and accuracy of U-Net

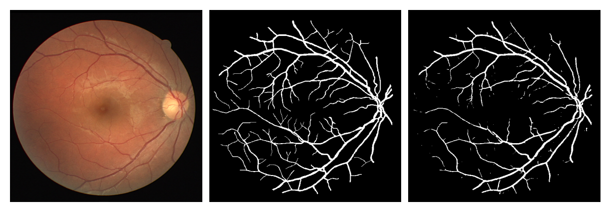


(a) Test Image (b) Test Prediction (c) Masked Image

Fig. 7. Test, predicted, masked images for U-Net

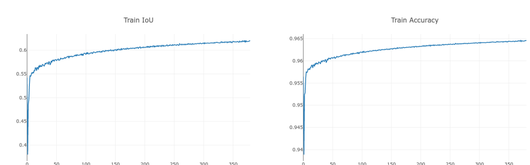
4.2 The DenseNet networks:

After training for 350 epochs, the Accuracy reached 96.5%, IoU reached 62%.



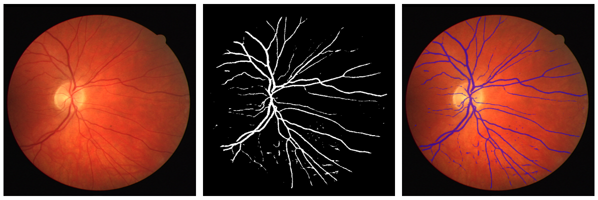
(a) Original Image (b) Target Image (c) U-Net Prediction

Fig. 8. Training, target, predicted images for DenseNet



(a) Training IoU (b) Training Accuracy

Fig. 9. Trainig IoU and accuracy of DenseNet

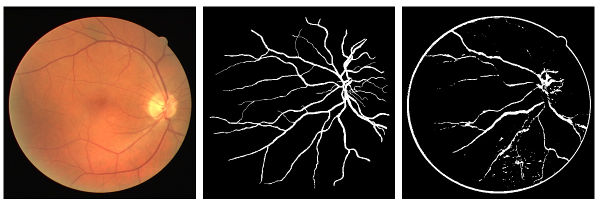


(a) Test Image (b) Test Prediction (c) Masked Image

Fig. 10. Test, predicted, masked images for DenseNet

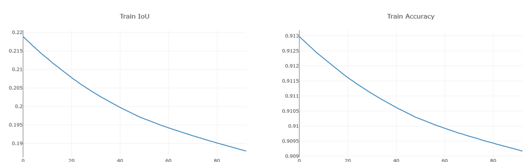
4.3 The SegCaps networks:

After training 350 epochs, the loss function did not converge. IoU and accuracy don't improve. IoU is less than 0.2.



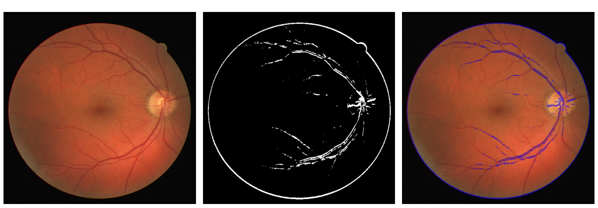
(a) Original Image (b) Target Image (c) U-Net Prediction

Fig. 11. Training, target, predicted image for SegCaps



(a) Training IoU (b) Training Accuracy

Fig. 12. Training IoU and accuracy of SegCaps



(a) Test Image (b) Test Prediction (c) Masked Image

Fig. 13. Test, predicted, masked images for SegCaps

5 Conclusion

The U-Net architecture showed the best performance on the DRIVE dataset. Good results are also obtained byThe DenseNet. The SegCaps networks could not converge on the DRIVE dataset. It may be proved that SegCaps is unable to handle complicated segmentation tasks. There are still great rooms for researchers to improve Capsules networks for segmentation tasks.

6 Contributions

Thanks to Dr. Sidike for giving me useful instructions and offering inspiring suggestions. The work is mainly conducted by the author-Zhankun Luo.

Appendix A
Architecture of U-Net Networks

Appendix one text goes here.

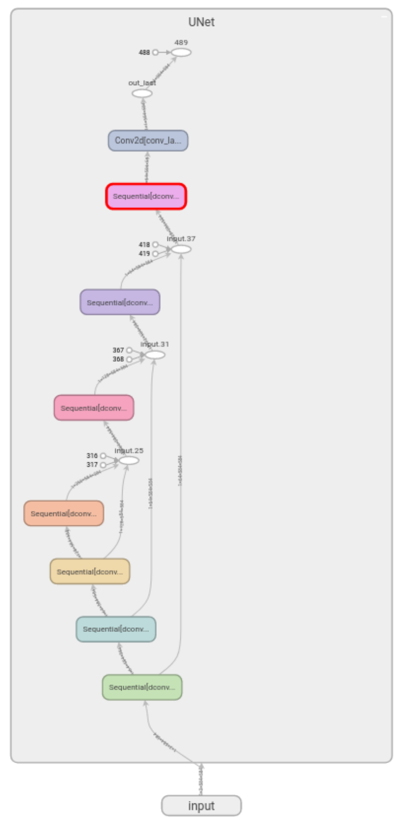


Fig. 14. Overview of U-Net Architecture.

Appendix B
Architecture of DenseNet Networks

Acknowledgments

Kind help and patient guidance of Dr. Sidike would be gratefully acknowledged. Besides, the author extended gratitude to IEEE and Overleaf for providing the IEEE Journal template.

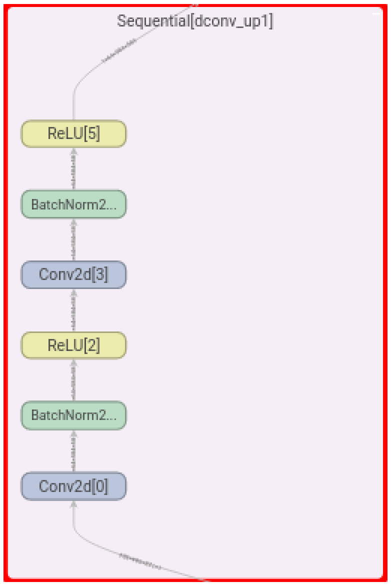


Fig. 15. Sequential Block in U-Net.

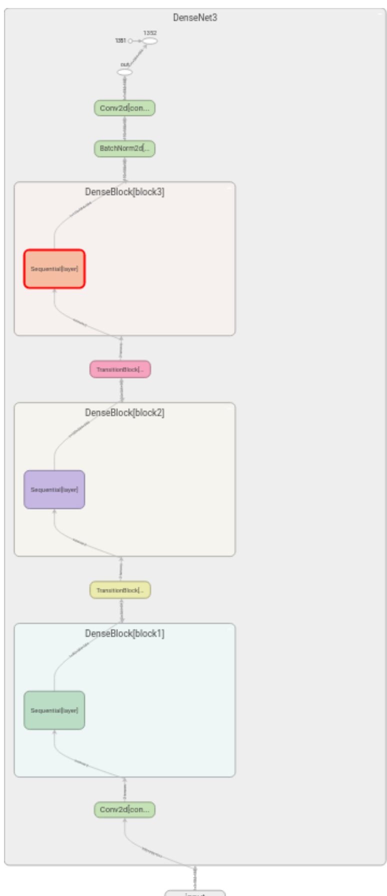


Fig. 16. Overview of DenseNet Architecture.

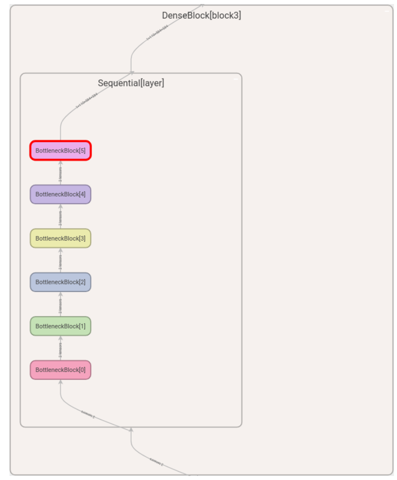


Fig. 17. Sequential Block in U-Net.

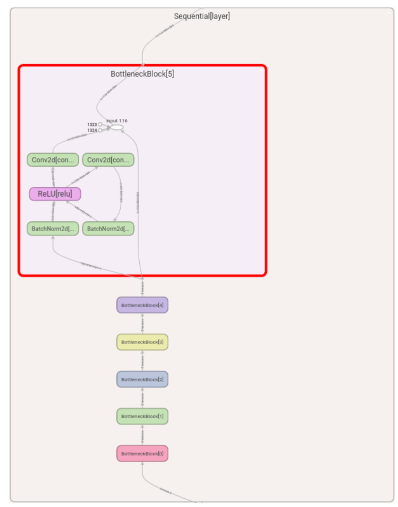


Fig. 18. BottleNeckBlock in Block of U-Net.

REFERENCES

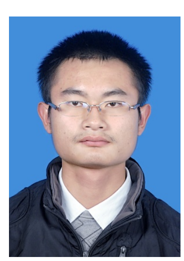
[1] S. Sabour and N. Frosst and G. E. Hinton, "Dynamic Routing Between Capsules," arxiv:1710.09829 [cs.CV], Nov. 2017.

[2] R. LaLonde and U. Bagci, "Capsules for Object Segmentation," arxiv:1804.04241 [stat.ML], April. 2018.

[3] O. Ronneberger and P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," arxiv:1505.04597 [cs.CV], May. 2015.

[4] G. Huang and Z. Liu and L. v. Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," arxiv:1608.06993 [cs.CV], Jan. 2018.

[5] H. Liao, "naturomics/CapsNet-Tensorflow", *Github*, 2019. Available: https://github.com/naturomics/CapsNet-Tensorflow. [Accessed: 11-Dec-2019].



Zhankun Luo received the BS degree in Telecommunication Engineering from Beijing Institute of Technology in 2019. He is a Master student with the Department of Electrical and Computer Engineering. His research interest includes deep learning, structured light and image processing.