Achieving Comparable Performance With Less Parameters in Segmentation of Melanoma Using Dense U-Nets

Charles Huang Gustavo Chau

Department of Bioengineering, Stanford University

Introduction

- Skin cancer is one of the most common types of cancer. Melanoma presents a high mortality rate.
- Growing interest in developing automatic methods for the diagnosis of Melanoma. Usual first step: **segmentation**.
- This project centers on the segmentation of melanoma lesions from dermatoscopic images of the ISIC challenge [1].

Dataset

- Task 1 of ISIC challenge: dermatoscopic image and the corresponding binary mask groundtruth [1].
- **2000 training examples, 150 validation examples, and 600 test examples**.
- Images resized to **192 x 256 pixels** and normalized by subtracting the mean and dividing by its standard deviation.
- Data augmentation did not help performance.

Acknowledgements

The authors would like to acknowledge and thank Yuan Xie, Yannan Yu, Enhao Gong, and Greg Zaharchuk for their help and advice.

References

[1] N. C. F. C. et al., "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (ISIC)," *CoRR*, vol. abs/1710.05006, 2017. [Online]. Available: http://arxiv.org/abs/1710.05006

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015. [Online]. Available: http://arxiv.org/abs/1505.04597

[3] Y. Yuan, M. Chao, and Y. Lo, "Automatic skin lesion segmentation with fully convolutional-deconvolutional networks," *CoRR*, vol. abs/1703.05165, 2017. [Online]. Available: http://arxiv.org/abs/1703.05165

[4] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *CoRR*, vol. abs/1608.06993, 2016. [Online]. Available: http://arxiv.org/abs/1608.06993

Methods

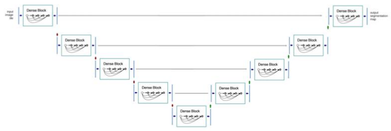


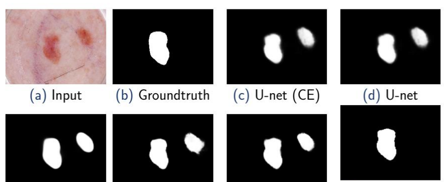
Figure: Generic dense U-net.

- We tested 4 architectures for solving the current problem: U-net [2], the 2017 winner which was a encoder-decoder [3], and dense U-nets (based on [4]).
- Learning rate=0.0001, trained until convergence, mini-batch size=16.

Results

Model	# param	Train Dice	Val Dice	Test Dice	Test(Post)
U-net(CE)	7.8M	0.85	0.90	0.77	0.80
U-net	7.8M	0.89	0.91	0.83	0.84
2017 winner	5.0M	0.85	0.92	0.81	0.83
small dense	0.7M	0.96	0.93	0.79	0.81
dense	2.7M	0.97	0.93	0.82	0.83

Table: Dice coefficient obtained for the different models.



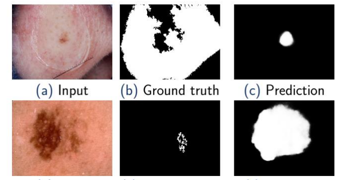
(a) Input (b) Groundtruth (c) U-net (CE) (d) U-net

(e) 2017 winner (f) Dense (g) Small Dense (h) Small Dense (postprocessing)

Figure: Example of Predicted images with the different methods.

Discussion

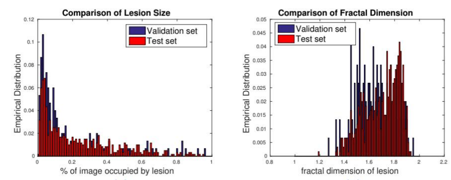
- Unpaired Mann–Whitney U test over percentage (PA) and the fractal dimension (FD) of validation and test obtained p-values ~ 0.0001. It is very likely that the validation and test sets are indeed coming from different distributions.



(a) Input (b) Ground truth (c) Prediction

(d) Input (e) Ground truth (f) Prediction

Figure: Example of discrepancy in the testing set of images caused by the significantly higher lesion border complexity or coloring of skin being mostly not melanoma tissue (bottom row).



(a) (b)

Figure: Distributions of PA and FD of validation and test set samples

Conclusion

- Dense U-net obtained **comparable validation and test dice scores**. The marked drop between validation and test dice in all trials was probably due to different distributions.
- Reduced the amount of parameters to as **small as 9% of the number of parameters** in a regular U-net.