Automating Melanoma Segmentation

Ning (Louis) Li,[1] Grace Wang,[1] Makena Low[2]
{louisli, gracenol, makenal}@stanford.edu

[1]Department of Computer Science, Stanford University
[2]Department of Electrical Engineering, Stanford University

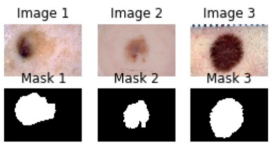
Stanford
Computer Science

Abstract

Manual lesion segmentation for diagnosis is time and labor intensive, and segmentation results can vary between providers. Thus, a robust and timely automated segmentation model could greatly improve diagnosis and detection. Our team introduces two novel U-Net based models that takes in a dermoscopic lesion image from the ISIC 2018 competition dataset and outputs a binary prediction mask. We use the Jaccard Index to measure performance. Our best model scored a Jaccard Index of 0.826 and a Threshold Jaccard Index of 0.785 on the test set, placing our model in the top 10 of the 2018 competition. [1]

Data

Our data was extracted from the "ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection" grand challenge datasets, with 2594 samples in total. Each sample consisted of a dermoscopic lesion RGB image with a labeled ground truth image (binary black and white masks, where 1 is the presence of a lesion and 0 is the absence). [2][3]



Features

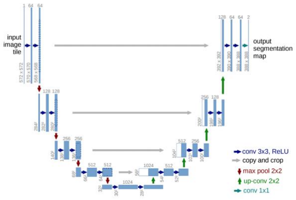
Original image sizes ranged from 600 x 600 to 5000 x 4000 pixels. For consistency among images, the lesion and ground truth images were reshaped to 256 x 256 resolution, then the RGB lesion images were then converted to grayscale and normalized. Keras data augmentation was used to flip, rotate, zoom, shear and shift the images, as augmentation is an essential component of U-Net.

U-Net

We used the pixel-wise Jaccard index to measure the performance of a prediction :

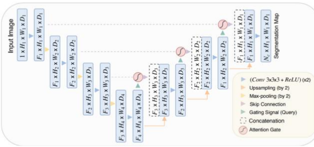
$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

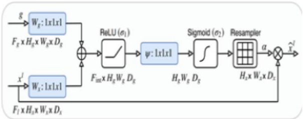
- The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. [4]



Attention U-Net

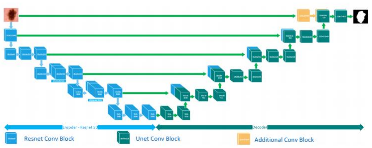
- The attention model traditionally has been used heavily in RNN, in which it helps neural networks make choices about which features need more attention [5]. It was an innovative idea to apply the attention model in a CNN architecture by adding the Attention Gate architecture to the existing U-Net. [5]





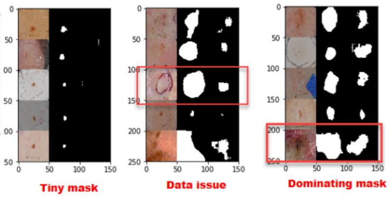
Extensions to Models

- Original U-Net + LeakyReLU + one more layer on the both encoder side and the decoder side.
- BigLeakyU-Net + one more layer on each side.
- BiggerLeakyU-Net with Attention U-Net
- Transfer learning with VGG16 as the encoder [7]
- Transfer learning with ResNet50 as the encoder [8]



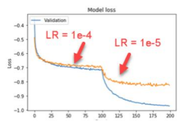
Error Analysis

The 10 lowest scores were due to tiny masks. To improve, we explored a higher resolution of 512 x 512 images as well as RGB (3 color channels instead of one). The result was greatly improved - the tiny mask issue disappeared after using RGB images. A higher resolution of images did not significantly improve performance.

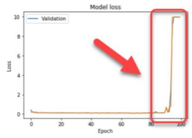


Experiments

- Hyperparameter Tuning
  - The best learning rate was 1e^-5



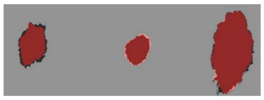
- Focal Loss was experimented with , but was unstable



Results

Due to the small dataset size, the data was split into 60/20/20 (1555/519/520) for training, validation and testing, respectively.

Model	Training	Validation
ResNet50	0.97	0.819
Attention UNet	0.818	0.815
FullUnet	0.813	0.812
BiggerLeaky	0.818	0.809
BiggerLeaky	0.779	0.778
VGG16	0.858	0.775



Discussion

- With a smaller amount of trainable parameters and the same level of performance, ResNet+U-Net was able to achieve the highest validation score. This model had a learning rate of 1e-5, batch size of 2, and a input resolution of 256x256.
- The fact that the ResNet encoder did better than the VGG16 makes sense because of the ResNet's depth and use of residual blocks.
- The Attention U-Net, built on top of the BiggerLeakyU-Net, also achieved comparable results, as expected.
- This is because the Attention U-Net is merely an extension of the U-Net, but increases model sensitivity by suppressing irrelevant regions and highlighting important ones.
- In the end, our results were comparable to the top 10 models in the ISIC 2018 competition.

Future

With more time, we would explore stabilizing the Focal Loss function and fine tuning Resnet + U-Net model. We also started testing with another popular image segmentation model - Mask RCNN at the end of this project, and it is showing promising results as we finalize this report.

References

[1] Cancer.org. (2019). *Skin Cancer | Skin Cancer Facts | Common Skin Cancer Types*. [online] Available at: https://www.cancer.org/cancer/skin-cancer.html [Accessed 10 Jun. 2019].
[2] Tschandl, P., Rosendahl, C. and Kittler, H. (2019). *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*.
[3] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "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)," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018.
[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. CoRR, abs/1505.04597, 2015. URL http://arxiv.org/abs/1505.04597.
[5] O. Ronneberger, "Invited Talk: U-Net Convolutional Networks for Biomedical Image Segmentation," *Informatik aktuell Bildverarbeitung für die Medizin 2017*, pp. 3–3, 2017.
[6] Simonyan, Karen, Zisserman, and Andrew, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv.org*, 10-Apr-2015. [Online]. Available: https://arxiv.org/abs/1409.1556. [Accessed: 10-Jun-2019].
[7] Zhang, Ren, Sun, and Jian, "Deep Residual Learning for Image Recognition," *arXiv.org*, 10-Dec-2015. [Online]. Available: https://arxiv.org/abs/1512.03385. [Accessed: 10-Jun-2019].