

ISICs 2018 Leison Segmentation using Mask R-CNN

This project uses [Mask R-CNN](#) [1] to predict lesion segmentation boundaries within dermoscopic images [ISIC 2018 Challenge - Task 1: Lesion Boundary Segmentation](#) [2] . The Mask R-CNN model will be trained with the goal of acheiving an Intersection over Union (IoU) score of 0.8 for the predicted image segmentation. Training is done using a Nvidia GeForce GTX 1050 Ti (4GB) GPU.

ISICs 2018 Challenge: Lesion Boundary Segmentation

The ISICs Lesion Boundary segmentation pre-processed dataset contains a total of 2,594 Lesion images (in .jpg format) with accompanying mask segmentations (in .png format). For the mask segmentations the pixel values 0 upto less than 255 represent the background, and 255 represents the target object. The Figure 1 shows the Lesion image with its accompanying mask.



Figure 1: Lesion Images + Accompanying Masks

Data Pre-processing

Before any pre-processing of the data, to ensure that there is no data leakage among the augmentations of the original image and the original image itself, data is split into Training, Validation, and Test at the beginning. First, data is split into Train (75%), and Test (25%), and second the Train set is further split into Train (75%), and Validation (25%). One of the biggest challenges in CNNs is the amount of computation required to process large resolution images as seen in Figure 2 where most of the images are above 1 megapixel. In order to overcome this challenge, the images were resized to 25% of original when it exceed 700*700 pixels (around 0.5 megapixels), otherwise, resolution is retained. Additionally, to account for varying positions of the Lesion in images, Images are flipped vertical, and horizontal separately. Images are normalized by dividing them by 255. Lastly, the target object is then labelled as "leison". The pre-processing sequence is shown in Figure 3, and the flipped images, and masks are shown in Figure 4. After augmentation, the total images per dataset are as follows: Train Set (4,374), Validation Set (1,461), Test Set (649).

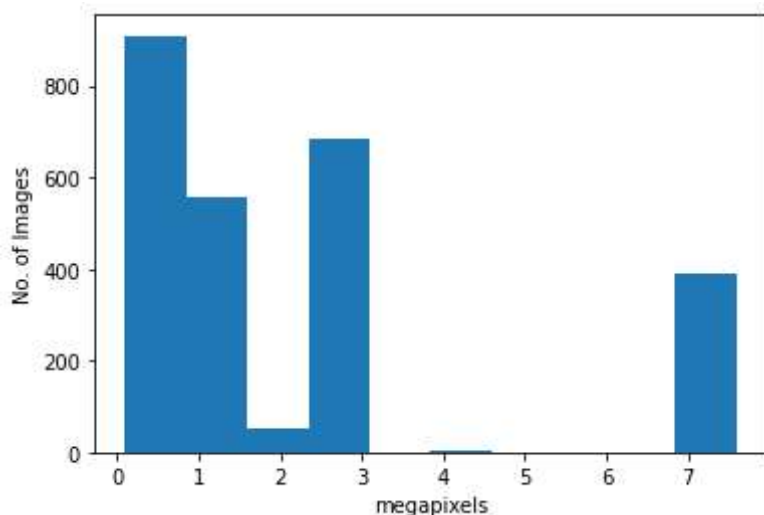


Figure 2: Pixel Resolution of Original Images



Figure 3: Image Pre-processing Steps

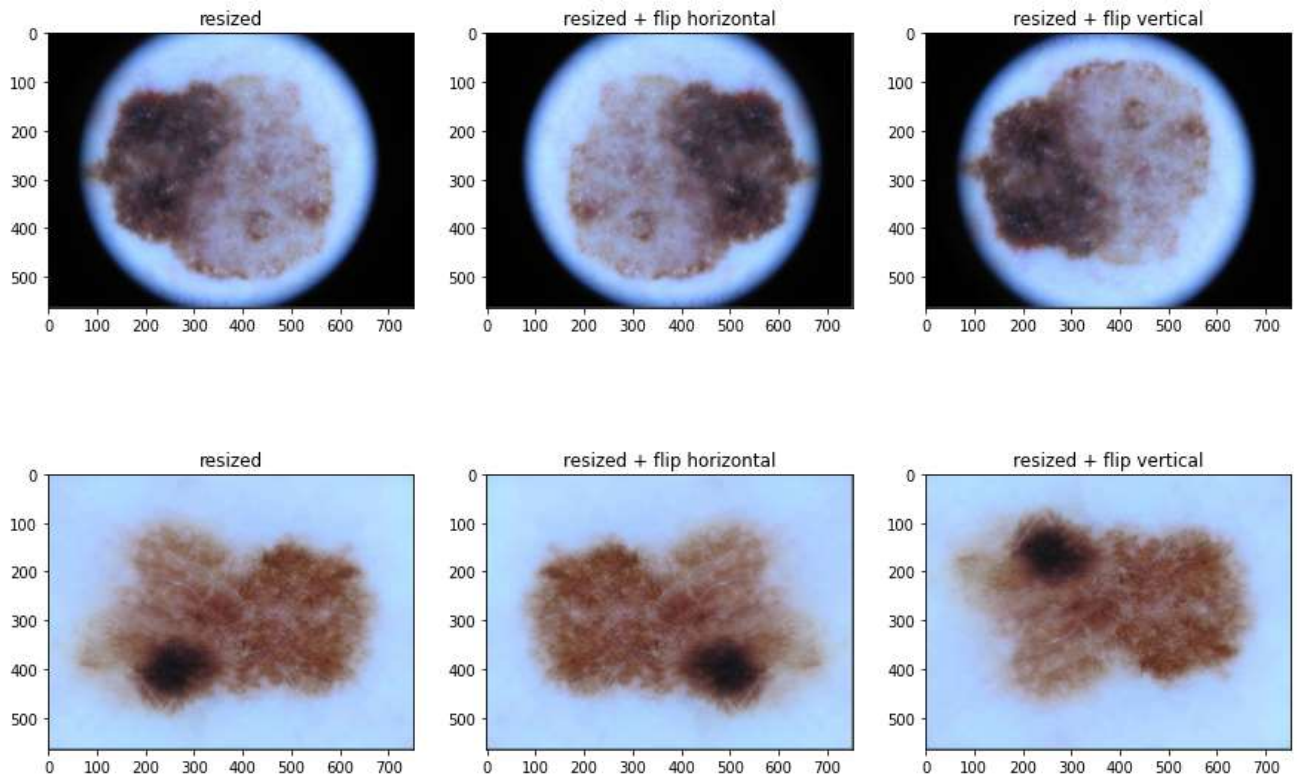


Figure 4: Image and Mask After Pre-processing

Mask R-CNN Architecture

Mask-RCNN is an extension of [Faster R-CNN](#) [3] where it also uses a *Region Proposal Network (RPN)* that proposes candidate bounding boxes during the first stage, and then modifies the second stage where the same *RoIPool* is used to extract features then adds a parallel generates an output for each *Region of Interest (RoI)*. Mask R-CNN optimizes the three (3) loss functions for each *RoI* namely, classification loss (L_{cls}), bounding-box class loss (L_{box}), and mask loss (L_{mask}). The total loss (L) is defined as $L = L_{cls} + L_{box} + L_{mask}$ [1]. The Mask RCNN architecture is show in Figure 5.

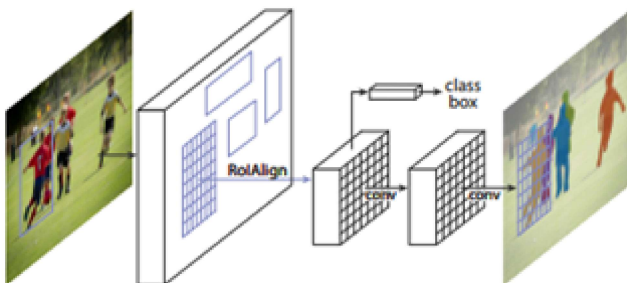


Figure 5: Mask R-CNN Architecture [1]

Training Configuration and Results

Configuration

The model was trained using pre-trained model weights using [COCO](#) dataset. The training follows a two-step procedure that takes advantage of selecting the depth of the model to decrease/increase amount of feature learning. Two layer configurations were used, namely, the "heads" (The RPN, classifier and mask heads of the network) and "3+" (*heads* + Train Resnet upto stage 3) layers. The "heads" layers on the first 5 epochs, and then "3+" layers upto the 30th epoch. The dataset is shuffled every epoch, and 1500 steps were taken for each. Non-geometric augmentations are also applied during training where none upto all of the augmentations are used randomly (Figure 6).

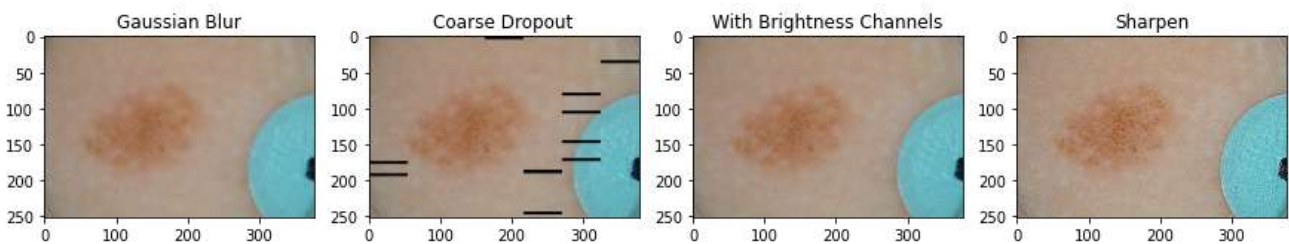


Figure 6: Train Augmentations

Training Loss

Figure 7 shows the overall training loss after each epoch, other loss metrics are shown in the model_loss folder. Note that step 0 is the 1st epoch, and step 29 is the 30th.

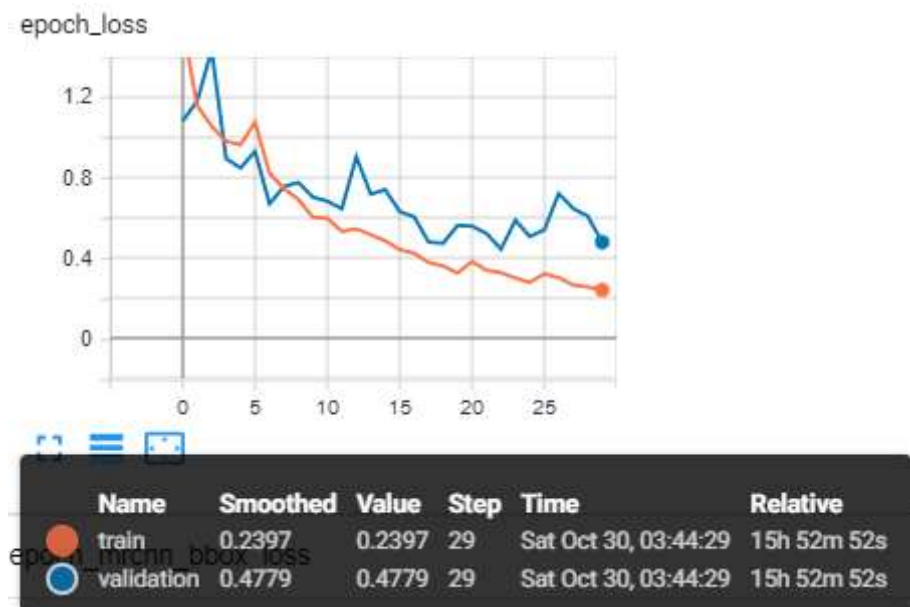


Figure 7: Overall Training Loss

Prediction Results

The goal of the project is to reach an IoU score of ≥ 0.80 , and this project has reached 0.85 on the test set where Figure 8 shows a boxplot of the IoU score distribution per image. It is known that there is only 1 object per image, and to account for the possibility of having detected multiple instances, the maximum IoU score per image was taken. One of the prediction results are shown in Figure 9 (where the number shown in the image is the model's prediction confidence), and its accompanying original image, and original mask in Figure 10 (see predictions folder for other results).

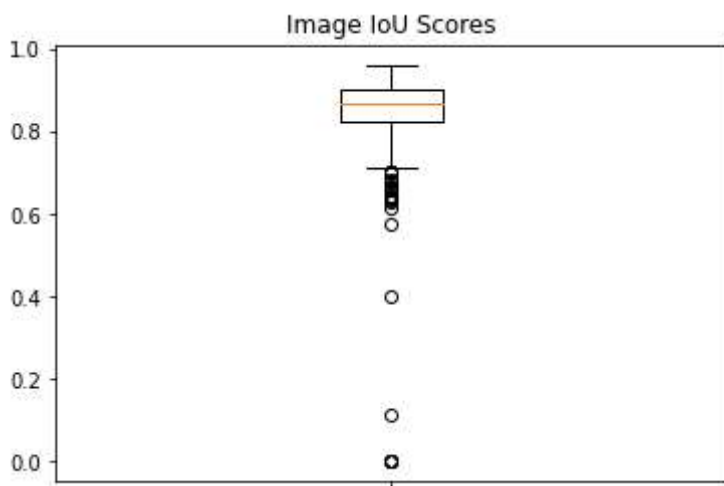


Figure 8: IoU Scores Boxplot

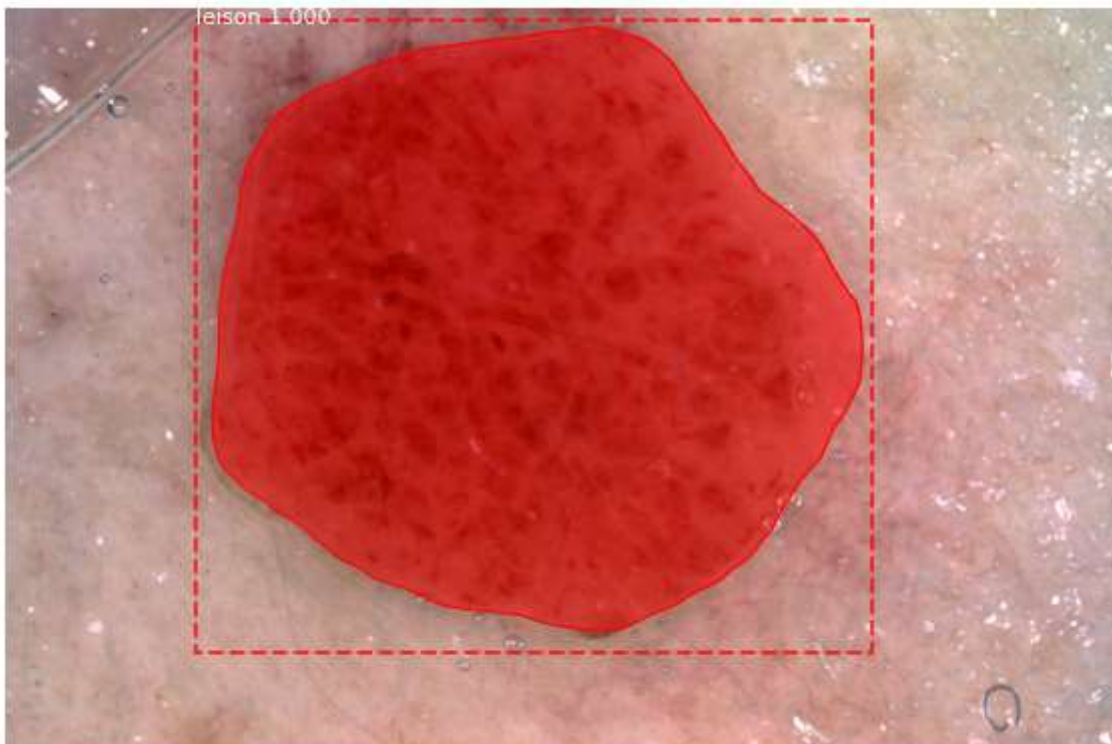


Figure 9: Prediction Result on Test Set

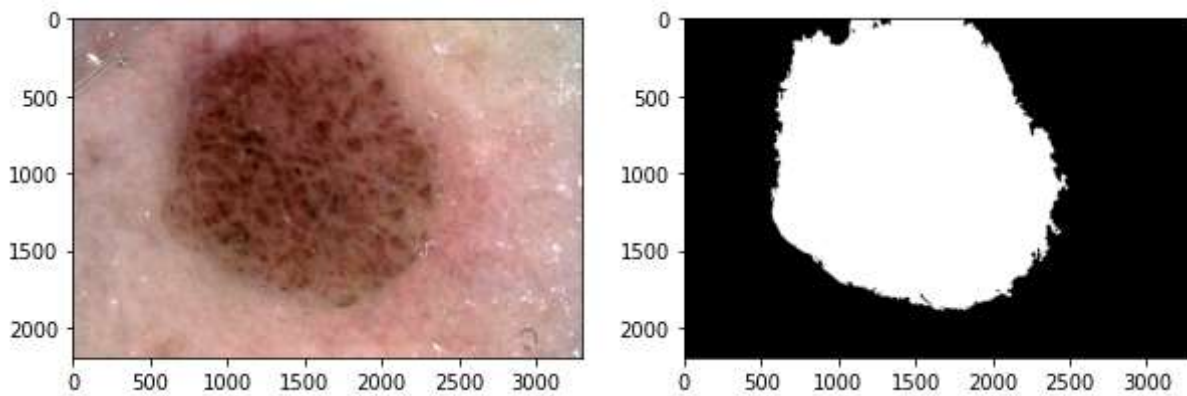


Figure 10: Original Image and Mask from Prediction

Pre-requisites

This project uses the Mask R-CNN implementation from [akTwelve](#) which uses Tensorflow 2 implementation of the original code from [matterport](#). A detailed step-by-step installation guide to setup akTwelve's Mask R-CNN implementation using Tensorflow 2 can be seen [here](#). Please ensure to check that you have the correct CUDA, and cuDNN installed for your Tensorflow version as described in the installation guide.

Libraries

- numpy
- scipy
- Pillow
- cython
- matplotlib
- scikit-image
- tensorflow>=2.0.0
- opencv-python
- cv2
- h5py
- imgaug
- os
- glob
- sys
- time
- gc

- IPython[all]

References

- [1] He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick, "Mask R-CNN," in 2017 IEEE International Conference on Computer Vision (ICCV), October 2017, pp. 2980–2988. [Online] Available: <http://arxiv.org/abs/1703.06870>.
- [2] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, A. Halpern, Dec 20, 2018, "ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection," distributed by International Skin Imaging Collaboration, [Online] Available: <https://challenge2018.isic-archive.com/>.
- [3] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. January, 2016, 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks', [Online] Available: <http://arxiv.org/abs/1506.01497>.
- Acknowledgements for the ISIC 2018 dataset:
- [4] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, Allan Halpern: "Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)", 2018, [Online] Available: <https://arxiv.org/abs/1902.03368>
- [5] Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161 doi:10.1038/sdata.2018.161 (2018).