DEEP LEARNING BASED APPROACH FOR DETECTING DISEASES IN ENDOSCOPY

Vishnusai Y[1,*], *Prithvi Prakash*[1,*], *Nithin Shivashankar*[1]

[1] Mimyk Medical Simulations Pvt. Ltd,[*] Authors contributed equally

ABSTRACT

In this paper, we discuss our submissions for the Endoscopic Disease Detection Challenge (EDD2020) [1], which had two sub-challenges. The first task involved a bounding box based multi-class detection of diseases, namely Polyp, Barrett's Esophagus (BE), Cancer, Suspicious and High-Grade Dysplasia (HGD). The second task involved creating semantic masks of the images for the aforementioned class of diseases. For the disease detection task we submitted the predictions of a Faster R-CNN with a ResNeXt-101 backbone and achieved a dscore of 0.1335±0.0936. For the semantic segmentation task, we employed a U-NET with a ResNeXt-50 backbone that achieved an sscore of 0.5031.

1. METHOD

1.1. Disease Detection Task

For the disease detection task we made use of a Faster R-CNN [2] object detector with a ResNeXt-101 serving as the backbone. Prior to feeding the data into our Neural Network model we applied augmentation techniques based on RandAugment [3] to improve the generalization capability of the neural network. From a choice of 16 augmentation techniques, two augmentation transformations were selected at random. We observed that magnitudes of 4, 5, 6 gave out the most effective augmentations and hence, this was chosen. The Faster R-CNN model was trained for 10 epochs and the learning rate was set to 0.01. The images were resized to 1300x800 pixels.

1.2. Semantic Segmentation Task

The U-NET [4] Architecture was used for the semantic segmentation task. Five separate U-NET models were created to train individual models to segment out different diseases. Prior to feeding our data to each U-NET, the images and masks were scaled to 256x256 pixels. It was then split to ensure that a proportionate sample of the true classes was present in both the sets. This was done by the K-Means clustering algorithm and sampling an 80-20 split from each bucket. The number buckets was decided using the Elbow-Method. We then applied augmentations on the train images namely: flip, zoom, and rotate and then trained them on a U-NET with a ResNeXt-50 backbone for 150 epochs.

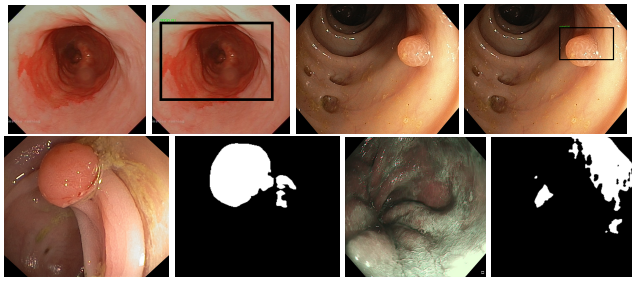

Fig. 1. Sample results on the test dataset for the disease detection and semantic segmentation tasks

2. RESULTS AND CONCLUSION

Disease Detection Task		
Sl No	**Model**	**mAP**
1	ResNet-101	0.1724
2	ResNeXt-101	0.2235

Semantic Segmentation Task			
Sl No	**Model**	**Train IoU**	**Val IoU**
1	Single Model	0.381	0.121
2	BE Model	0.871	0.542
3	Cancer Model	0.782	0.217
4	HGD Model	0.814	0.313
5	Polyp Model	0.932	0.571
6	Suspicious Model	0.434	0.115
7	Aggregate Model (2-6)	0.766	0.351

Table 1. Mean Average Precision of Test Data. Intersection over Union of Training and Validation Data.

The results of the disease detection and segmentation tasks are summarised in Table 1. From the disease detection section, we see that the ResNeXt-101 outperformed the ResNet-101. On submission we obtained a dscore of 0.1335±0.0936. From the semantic segmentation task section, we observe that the individual disease models performed better than a single model trained for all diseases. This prompted us to adopt an aggregate model that aggregated the results of the individual disease models. On submitting the predictions of this aggregate model on the test dataset, an sscore of 0.5031 was obtained.

3. REFERENCES

[1] Sharib Ali, Noha Ghatwary, Barbara Braden, Dominique Lamarque, Adam Bailey, Stefano Realdon, Renato Cannizzaro, Jens Rittscher, Christian Daul, and James East. Endoscopy disease detection challenge 2020. *arXiv preprint arXiv:2003.03376*, 2020.

[2] Ren et. al. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS15, page 9199, Cambridge, MA, USA, 2015. MIT Press.

[3] Ekin D. Cubuk et. al. Randaugment: Practical automated data augmentation with a reduced search space, 2019.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of *LNCS*, pages 234–241, 2015.