SEMANTIC SEGMENTATION, DETECTION AND LOCALIZATION OF MUCOSAL LESIONS FROM GASTROINTESTINAL ENDOSCOPIC IMAGES USING SUMNET

Velmurugan Balasubramanian[1] *Rajiv Kumar*[2] *Sarasa Jyothsna Kamireddi*[3]
Rachana Sathish[4] *Debdoot Sheet*[5]

[1] Velmurugan Balasubramanian, School of medical science and Technology, IIT, Kharagpur, India
[2]Rajiv Kumar, Department of Chemical Engineering, IIT, Kharagpur, India
[3]Sarasa Jyothsna Kamireddi, Department of Electrical Engineering, IIT, Kharagpur, India
[4]Rachana Sathish, Department of Electrical Engineering, IIT, Kharagpur, India
[5]Debdoot Sheet, Centre of excellence in AI, Department of Electrical Engineering, IIT, Kharagpur, India

CEUR-WS.org/Vol-2595/endoCV2020_paper_id_s18.pdf

1. METHOD

We trained a fully convolutional neural network based on SUMNet [1] architecture, described in Fig.1, using Pytorch, for segmentation, detection and localization of lesions in Gastrointestinal endoscopic images using 386 images from EDD 2020 dataset [2]. An 80:20 training-validation split was followed with additional weights given to the under-represented classes depending upon their overall frequency of occurrence. We augmented the dataset with rotation, affine, scaling, projective and multi-crop transformations to accommodate for the variations caused due to scope positioning and augmented with variable brightness and HSV values to accommodate for images enhanced with narrow-band imaging and variable lighting conditions [3] [4]. We used the ADAM learning rate optimizer and binary cross-entropy loss function for training. SUMNet features (i) an encoder-decoder architecture with the pooling indices of encoder being passed to the corresponding decoder upsampling layers, (ii) encoder having a VGG11 like architecture pre-initialized with ImageNet pre-trained weights and (iii) concatenation of activations of the encoder with that of the decoder, combining the features of segmentation networks for natural and biomedical images [1].

2. RESULT

Our model was able to obtain dice coefficients of 0.977, 0.974, 0.986, 0.987, 0.961 and 0.545, 0.219, 0.172, 0.339, 0.573 on the training and validation sets for Barretts oesophagus, suspicious, high-grade dysplasia, cancer and polyp classes respectively. A class-wise distribution of the abnormalities detected in the test dataset is shown in Table 1 and

Copyright ©2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

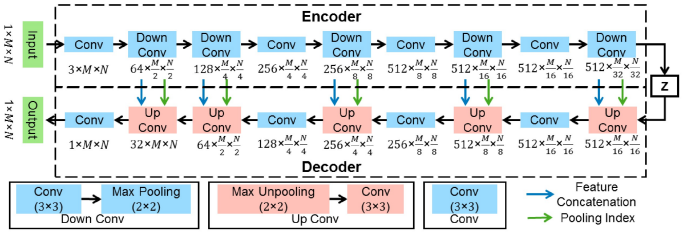

Fig. 1: SUMNet architecture.

Class Name	No of instances
Barrett's Esophagus	21
Suspicious	6
High grade dysplasia	9
Cancer	1
Polyp	30

Table 1: Distribution of the detected classes in the test dataset

Fig. 2 shows examples of the semantic masks and the bounding boxes we obtained for each of the five classes. We were able to obtain a semantic segmentation score of 0.538 with a standard deviation of 0.35 in our test submission and a mean detection score of 0.16 with a standard deviation of 0.074.

3. REFERENCES

[1] Sumanth Nandamuri, Debarghya China, Pabitra Mitra, and Debdoot Sheet. Sumnet: Fully convolutional model for fast segmentation of anatomical structures in ultrasound volumes. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Apr 2019.

[2] Sharib Ali, Noha Ghatwary, Barbara Braden, Dominique Lamarque, Adam Bailey, Stefano Realdon, Renato Can-

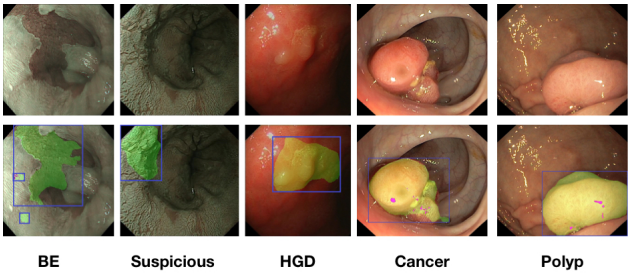

Fig. 2: Illustration of the original images (top) and their corresponding semantic masks (bottom).

nizzaro, Jens Rittscher, Christian Daul, and James East. Endoscopy disease detection challenge 2020. *arXiv preprint arXiv:2003.03376*, February 2020.

[3] Georg Wimmer, Andreas Uhl, and Andreas Vecsei. Evaluation of domain specific data augmentation techniques for the classification of celiac disease using endoscopic imagery. In *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–6. IEEE, 2017.

[4] Andrea Asperti and Claudio Mastronardo. The effectiveness of data augmentation for detection of gastrointestinal diseases from endoscopical images. *arXiv preprint arXiv:1712.03689*, 2017.