2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Precise Bleeding and Red lesions localization from Capsule Endoscopy using Compact U-Net

Aparna Kanakatte[1] and Avik Ghose[2]

Abstract— **Wireless capsule endoscopy is a non-invasive and painless procedure to detect anomalies from the gastrointestinal tract. Single examination results in up to 8 hrs of video and requires between 45 - 180 mins for diagnosis depending on the complexity. Image and video computational methods are needed to increase both efficiency and accuracy of the diagnosis. In this paper, a compact U-Net with lesser encoder-decoder pairs is presented, to detect and precisely segment bleeding and red lesions from endoscopy data. The proposed compact U-Net is compared with the original U-Net and also with other methods reported in the literature. The results show the proposed compact network performs on par with the original network but with faster training and lesser memory consumption. Also, the proposed model provided a dice score of 91% outperforming other methods reported on a blind tested WCE dataset with no images from this set used for training.**

I. Introduction

Gastrointestinal (GI) hemorrhage results in approximately 300,000 hospital admissions in the US annually. Upper GI bleeding occurs at a rate of 40-150 episodes per 100,000 persons per year, with a mortality rate of 6% - 10%[1]. The small bowel is one of the major organs where frequent bleeding occurs, known as Obscure Gastrointestinal Bleeding (OGIB). The small bowel organ length and its morphological diversity limit full and direct visualization through endoscopy and colonoscopy. This limitation is overcome by using Wireless Capsule endoscopy (WCE) which is a non-invasive and painless procedure that uses a tiny wireless camera to take pictures of the gastrointestinal tract for diagnosis of anomalies. This capsule is swallowed by the patient and is passively propelled by peristalsis. WCE imaging is extremely useful in detecting bleeding and other anomalies like Crohn's, Polyps, ulcers, tumors, Barrett's esophagus, small bowel syndrome, etc in GI.

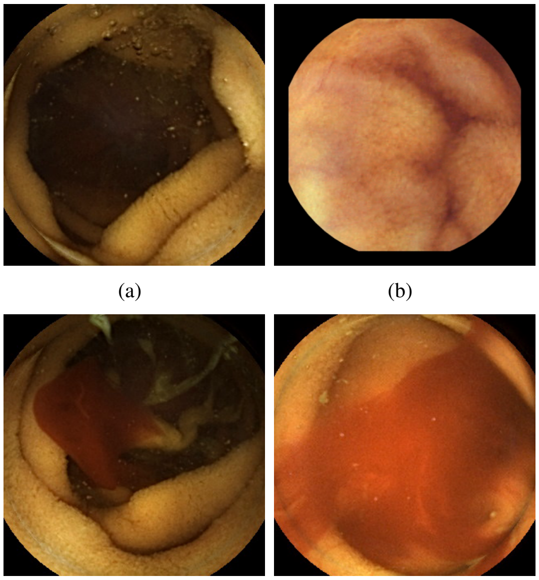
The sensitivity of capsule endoscopy (CE) for detection of lesions relies on the quality of images and number of images the capsule takes in 1 second. The quality of the image is improving progressively with depth of view and luminosity control improving from one generation to another and for every brand of capsule. The number of images is very important in assessing sensitivity, but an increasing number of images leads to an increase in diagnosis time for physicians as they have to go through all the images to identify the anomaly. With an acquisition rate of 2-14 frames per second (according to its model and producer), the capsule acquires and transmits around 72000 frames during its battery lifetime of 7-8 hrs among which only 1% may be of clinical interest [3]. A specialized doctor needs to view all these images, annotate the relevant ones supporting diagnosis and create a report. This diagnosis and annotation is intense and time-consuming which may vary between 45 and 180 minutes [19] depending on the complexity. Some image processing techniques have been incorporated by the manufacturers to speed up the process, such as, Given Imaging's Rapid Reader Software which includes a blood detection tool called Suspected Bleeding Indicator (SBI). However, studies have shown that the performance of SBI is suboptimal and inconsistent [4]. Automatic detection of anomalies like bleeding spots, ulcers, polyps with high accuracies will help physicians in reducing this diagnosis time as they need not have to go through all the images but look for the ones flagged by automated detection systems. In this paper, a faster and memory-efficient automated method to detect and localize the bleeding spots and red lesions from capsule endoscopy data has been attempted. Few examples of normal and bleeding capsule endoscopy images from the red lesion dataset are shown in Fig 1. This paper is organized as follows. Literature review is provided in the section II. The proposed compact U-Net is explained in the section III. Implementation details, results and discussions are explained in the section IV. Conclusion and future work are provided in the section V.

II. Related work

Researchers have proposed automatic or semi-automatic methods for the detection of polyps, ulcers, bleeding, Crohn's and other abnormalities from WCE images. Jia *et al* [2] proposed bleeding detection using histograms of K-Means clustering as features and classification by CNN. Fu *et al* [5] used superpixel color features and SVM for bleeding detection. Tuba *et al* [6] used local binary patterns extracted from channel I of HIS color space and SVM for extracting bleeding areas. Sainju *et al* [7] used statistical features and region growing method for automated bleeding detection. Konstantin *et al* [8] has compared the performance of color versus texture features in bleeding detection from capsule endoscopy videos. Iakovidis *et al* [9] used salient superpixels to detect bleeding regions.

Advancement in deep learning has presented significantly better results than previous methods by automatically extracting characteristics from data and thus supporting new developments in computer-aided diagnosis systems. Anjany *et al* [10] used CNN for various abnormality detection.

[1]Aparna Kanakatte is with TCS Research and Innovation, Bengaluru, India aparna.kg@tcs.com
[2]Avik Ghose is with TCS Research and Innovation, Kolkata, India avik.ghose@tcs.com



(a) (b) (c) (d)

Fig. 1: Normal (a, b) and bleeding (c, d) images

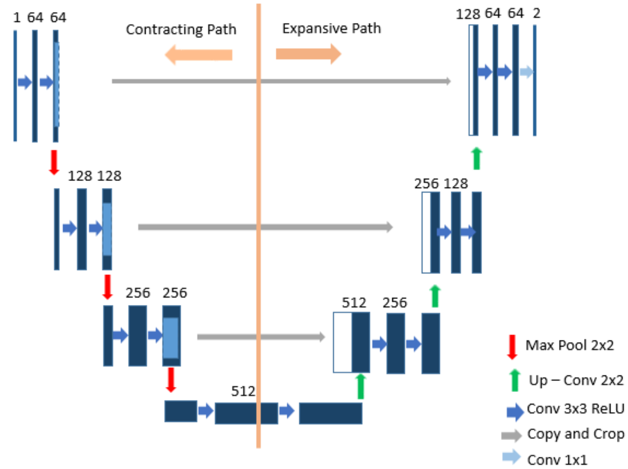


Fig. 2: Architecture of compact U-Net

Coelho *et al* [11] used U-Net to detect bleeding areas and also compared the performance of this with commercial tools. In this paper, we have proposed a simplified compact U-Net with reduced filters and compared the performance of this with standard U-Net in bleeding detection and localization using an open-source dataset. We have also shown an improved performance of the proposed model when blind tested on another dataset that has not been used for training.

III. Proposed Compact U-Net

U-Net is a fully convolutional neural network developed by Ronneberger *et. al.* [12] for fast and precise segmentation of medical images. This presents a U shape due to the symmetrical form presented by the contracting and expansive path following encoder-decoder architecture which captures context, thus enabling precise localization. The network can be trained end to end with lesser images and the results outperforms the sliding window convolution network. The basic idea of U-Net is that it takes multiple patches overlapping within an image and performs pixel-wise reconstruction of the segmentation mask. As a result of this process, every pixel results in more than one probability score, which is then averaged to make the final prediction. Due to the nature of training and skip connections, U-Net results in low validation loss and excellent segmentation.

The proposed architecture of compact U-Net shown in fig 2 is a three-level U-Net with lesser encoder-decoder pairs than the original U-Net. The contracting path (left) follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. At each downsampling step, we double the number of feature channels. Every step in the expansive path (right) consists of an upsampling of the feature map followed by a 2x2 convolution (up-convolution) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer, a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.

The original architecture of U-Net consists of four levels of a contracting path and an expansive path. The patch size of the regular U-Net is 32x32 while the compact patch is 16x16. This reduction helps in faster training and lesser memory consumption.

IV. Results and Discussion

To validate the proposed algorithm red lesion endoscopy dataset [11] and WCE dataset [17] were used. We have used Set 1 of the red lesion endoscopy dataset for both training and testing as the annotations were available for this set. Set 1 contains 3295 images of which 1131 containing bleeding or red lesions and 2164 with no lesions or blood. This set contains images from different cameras such as MicroCam, PillCam SB1, SB2 and SB3 with different red lesions such as angioectasias, angiodysplasias, bleeding, and others. All lesions were annotated manually by experts. The images have 320×320 or 512×512 resolutions depending on the camera settings.

A. Implementation details

Ubuntu workstation with Intel Core i7, 32GB RAM, and NVIDIA GeForce 1080i GPU card is used for training and testing. The proposed approach is implemented using PyTorch [13] and OpenCV. Training set consists of 1400 (700 bleeding and 700 non-bleeding) frames, the validation set consists of 400 (200 bleeding and 200 non-bleeding) frames, and the test set consists of 431 (231 bleeding and 200 non-bleeding) frames from the red lesion dataset. All the images are resized to 512×512. Both U-Net and compact

TABLE I: Precise segmentation performance comparison using the Dice score of the proposed with other methods reported on red lesion dataset

	Dice
Coelho *et al* [11]	87.08%
Tuba *et al* [6]	84.00%
U-Net	96.55%
Proposed Compact U-Net	96.39%

U-Net were trained for 5000 epochs with a learning rate of 0.001 using Adam Optimizer with mean squared error loss function.

B. Performance comparison

The performance comparison of the algorithm in segmenting precise bleeding or red lesion regions is compared using Dice scores. The Dice coefficient is a relative metric that provides a similarity index between predicted and groundtruth segmentation. This is calculated by

$$D(X,Y) = \frac{2\,|X \cap Y|}{|X| + |Y|} \tag{1}$$

where $|X|$ and $|Y|$ are the cardinalities of the two sets (*i.e.* the number of elements in each set).

The detected bleeding regions from the proposed compact U-Net and the original U-Net on test images from the red lesion dataset are shown in figure 3. It can be seen that the compact U-Net has detected bleeding regions similar to U-Net. The dice score comparison with other methods reported in the literature on the red lesion dataset is provided in table I. It can be seen that the proposed compact U-Net provides a dice score of 96.39% and standard U-Net provides 96.55% on the red lesion dataset. The difference in dice score between the proposed compact and standard U-Net is very minimal. We have also compared the proposed model in terms of accuracy, True positive rate (TPR, also called sensitivity), and True negative rate (TNR, also called specificity) with other methods in table II. It can be seen that the proposed provides better accuracy of 95.90% and TPR of 99.57% as compared to other methods.

Eliminating an encoder-decoder from the original architecture has lesser memory consumption. It was also observed that the proposed compact U-Net was able to reduce the training time by nearly 30% when trained on standard U-Net with a learning rate of 0.001 for 5000 epochs when trained on Ubuntu workstation with Intel Core i7, 32GB RAM, and NVIDIA GeForce 1080i GPU card. The results have shown that the effect on dice score is minimal with the compact and standard U-Net but with a significant reduction in training time.

C. Blind Testing on WCE dataset

We have blind tested the proposed compact model on an entirely different WCE dataset which is free and available online [17]. This set contains 50 capsule endoscopy bleeding images of varying image resolution with groundtruth [17]. No images from this set have been used for training the proposed network. Few results from the proposed method on this dataset are shown in fig 4. It can be seen that the proposed segmentation output is very similar to groundtruth. The Dice score of 91% from the proposed method outperforms other methods reported in literature on this dataset as shown in table III.

TABLE II: Comparison of accuracy, true positive rate (TPR) and true negative rate (TNR) of the proposed with the other methods reported on Set 1 red lesion dataset

Authors	Accuracy (%)	TPR (%)	TNR (%)
Sainju *et al* [7]	93.00	96.00	90.00
Figueriredo *et al* [14]	92.70	92.90	> 90.00
Usman *et al* [15]	92.00	94.00	91.00
Xiong *et al* [16]	94.10	91.69	94.59
Coelho *et al* [11]	95.88	99.56	93.93
Proposed	**95.90**	**99.57**	91.00

TABLE III: Comparison of the Dice score of the proposed with other methods reported on WCE dataset. Note: None of the images from this dataset was used for training

	Dice
Tuba *et al* [6]	84%
Deeba *et al* [18]	81%
Proposed	**91%**

V. Conclusion and Future Work

In this paper, we have shown a compact U-Net network with lesser encoder-decoder pairs for detecting red lesions and bleeding spots from the capsule endoscopy dataset. The proposed method provided a better accuracy of 95.90% and a TPR of 99.57% in detecting bleeding and red lesions. The proposed method outperformed other reported methods on a blind tested WCE dataset with a dice score of 91% with none of the images from this dataset used for training the network. We have shown the results of the proposed compact network performs on par with the original U-Net network but with faster training and lesser memory consumption. The compact network trained nearly 30% faster than the standard U-Net. Our future work includes an extension of the network to detect other anomalies from the endoscopy dataset.

References

[1] Kim. B.S, Li. B. T and Engel. A, *et.al.*: Diagnosis of gastrointestinal bleeding: a practical guide for clinicians, World J Gastrointest Pathophysiol. ,2014, vol. 5, pp 467-78

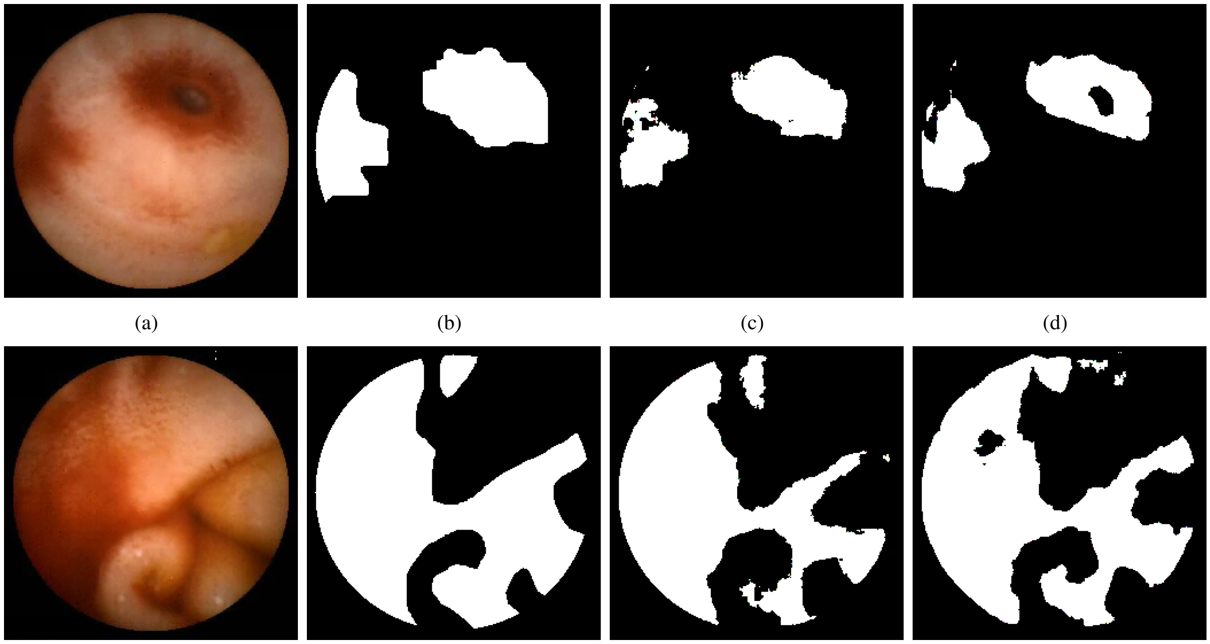
[2] Jia. X and Meng M.Q.H: Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features, IEEE EMBC, pp. 3154-3157, 2017

[3] Farah .D, Monzurul. I, Francis. M. B, Khan. A.W: Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection, 2018, Biomedical Signal Process. and Control, vol (40), pp 415-424

[4] Anastasios K, Emanuele R, Alexandros K: A ten-point contemporary review, World J Gastroenterol., 2013; 19(24): 3726-3746

[5] Fu. Y, Zhang. W, Mandal. M and Meng M.Q.H : Computer-aided bleeding detection in WCE video, IEEE journal of biomedical and health informatics, vol 18(2), pp. 636-642, 2014.

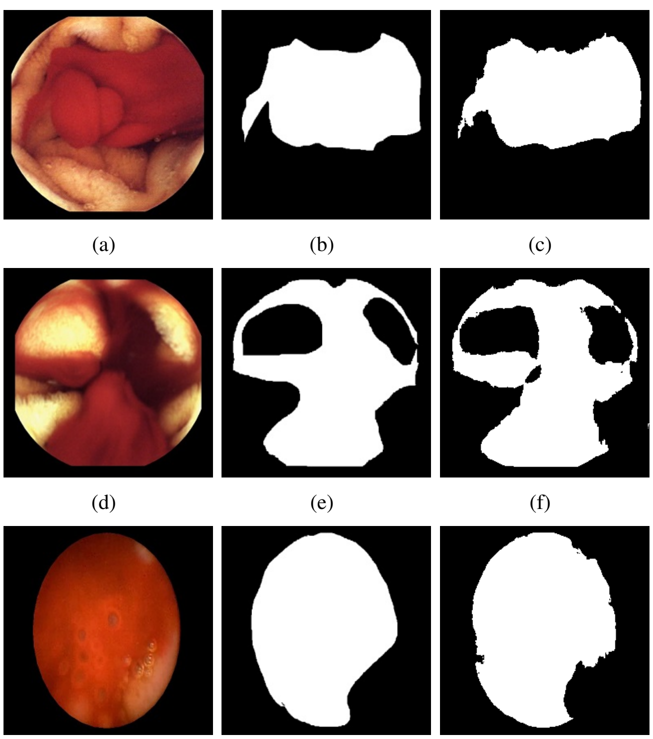
[6] Tuba. E, Tuba M and Jovanovic. R : An algorithm for automated segmentation for bleeding detection in endoscopic images, International Joint Conference on Neural Networks (IJCNN), pp.4579-4586, 2017.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 3: Results of Bleeding and red lesion detection on Red lesion dataset. Image a, e, are original images. b, f are ground-truth images. c, g are the U-Net output images. d, h are proposed compact U-Net output images



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 4: Results of Bleeding detection on blind tested WCE dataset. Image a, d, g are original images. b, e, h are ground-truth images. c, f, i are proposed compact U-Net output images. Note: None of the images from this dataset was used for training.

[7] Sainju. S, Bui. F.M and Wahid. K.A: Automated bleeding detection in capsule endoscopy videos using statistical features and region growing. J. Med. Syst. 38(4), 25 2014.

[8] Konstantin. P, Shipra. S, Fawnizu. A.H, Aamir. S.M, Olga. O, Michael. R, Pal. H, Shiaw. H.H and Khean-Lee. G.:Bleeding detection in wireless capsule endoscopy videos - Color versus texture features, J Appl Clin Med Phys, 2019, pp 141-154, vol 20(8)

[9] Iakovidis. D. K, Chatzis. D and Chrysanthopoulos. P: Blood detection in wireless capsule endoscope images based on salient superpixels, EMBC, pp 731 - 734, 2015

[10] Anjany. K.S, Surya. T. D and Chandra. S.S:,A convolutional neural network approach for abnormality detection in Wireless Capsule Endoscopy, ISBI 2017, pp 1057-1060

[11] Coelho P., Pereira A., Leite A., Salgado M. and Cunha A: A Deep Learning Approach for Red Lesions Detection in Video Capsule Endoscopies, Image Analysis and Recognition. ICIAR 2018. vol 10882, pp. 553–561

[12] Ronneberger. O, Fischer. P and Broz. T, U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI - 2015, pp.234-241.

[13] Paszake. A, Gross. S, Massa. F, et. al: PyTorch: An Imperative Style, High-Performance deep learning library, Advances in Neural Information Processing System 32 , 2019, pp 8024-8035.

[14] Figueiredo. I.N, Kumar. S, Leal. C, Figueiredo. P.N: Computer-assisted bleeding detection in wireless capsule endoscopy images, Biomed. Eng. Imaging Vis. 1(4), 198–210 (2013)

[15] Usman. M.A, Satrya. G, Usman. M.R, Shin. S.Y: Detection of small colon bleeding in wireless capsule endoscopy videos, Comput. Med. Imaging Graph. 54, pp 16–26 (2016)

[16] Xiong. Y, Zhu. Y, Pang. Z, Ma. Y, Chen. D, Wang. X: Bleeding detection in wireless capsule endoscopy based on MST clustering and SVM, SiPS, vol. 35, pp. 1–4,2015.

[17] Deeba. F: Bleeding images and corresponding groundtruth of CE images, 2016. Available:https://sites.google.com/site/farahdeeba073/Research/resources

[18] Deeba. F, Bui. F.M and Wahid K. A: Automated growcut for segmentation of endoscopic images, IJCNN, pp 4650-4657, 2016

[19] Barbara. P, Tammam. T, Marco. G, Enrico. M and Gabriella.O: A technique for blood detection in wireless capsule endoscopy images, EURASIP, pp 1864 - 1868, 2009