Preliminary results of red lesion segmentation in WCE images

Said Charfi^{1,*}, Mohamed El Ansari^{1,2,**}, Ayoub Ellahyani^{3,***}, and Ilyas El Jaafari^{3,****}

¹LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco.

²Informatics and Applications Laboratory, Department of Computer Science, Faculty of Science, My Ismail University Meknès, Morocco

³LabSIE, Department of Mathematics and Computer Science, multidisciplinary faculty, Ibn Zohr University, BP 638, 45000 Ouarzazate, Morocco.

Abstract. Wireless capsule endoscopy (WCE) is a novel imaging technique that can view the entire small bowel in human body. Thus, it is presented as an excellent diagnostic tool for evaluation of gastrointestinal diseases compared with traditional endoscopies. However, the diagnosis by the physicians is tedious since it requires reviewing the video extracted from the capsule and analysing all of its frames. This tedious task has encouraged the researchers to provide automated diagnostic techniques for WCE frameworks to detect symptoms of gastrointestinal illness. In this paper, we present the preliminary results of red lesion detection in WCE images using Dense-Unet deep learning segmentation model. To this end, we have used a dataset containing two subsets of anonymized video capsule endoscopy images with annotated red lesions. The first set, used in this work, has 3,295 non-sequential frames and their corresponding annotated masks. The results obtained by the proposed scheme are promising.

1 Introduction

Diseases of the digestive tract, such as esophagus, stomach and small intestine, colon, and other digestive organs cancers pose a serious threat to human health. Many types of endoscopy are employed to examine the patient's gastrointestinal tract. For example, gastro-copy, pressure enteroscopy and colonoscopy are used to examine the human digestive system. However, most of the above endoscopy tests are limited to examine the human small intestine. To overcome these limitations, the WCE examination developed by Given Imaging in 2000 was presented as an excellent diagnostic tool for evaluation of gastrointestinal bleeding, ulcers, Crohn's disease and others digestive abnormalities. Capsule endoscopy is mainly used to find the cause of unexplained bleeding in the digestive tract or in case of inflammation, red lesion or tumors in the small intestine. Nevertheless, this review or analysis of these videos by physicians can take several hours, which is tiring, so they may miss parts where abnormalities of the gastrointestinal tract are present, since these parts are often present only in a few frames of the video sequence. Furthermore, the size and distribution of the anomalies make their identification with the naked eye very difficult. Thus, computer-assisted diagnosis and offline post processing have been developed to facilitate the reading of the WCE video sequences, as well as to alleviate the heavy burden of doctors.

Many approaches have been proposed for automatic detection of GI tract diseases using handcrafted features in WCE images [1]. In [2, 3], a multi-scale analysis-based grey-level co-occurrence matrix (GLCM) is conducted for ulcer detection. In the second work, local features extractor is proposed for abnormalities detection in WCE images.

In this work, we mainly focus on those based on deep learning models. In [4], Zou et al. used deep convolutional neural network (DCNN) framework to classify WCE images digestive organs. Authors in [5] proposed a novel WCE classification system by a hybrid CNN with extreme learning machine (ELM). Instead of utilizing the conventional full-connection classifier in DCNN classification system, the cascaded ELM was used as a strong classifier in the previous work. Zhu et al. [6] introduced DCNN model to extract generic features of WCE images and applied SVM as a classifier to detect lesions. In [7, 8], systems for small intestine motility characterization and bleeding detection, based on Deep CNN were introduced. In [9] authors proposed a novel deep feature learning method, named stacked sparse autoencoder with image manifold constraint (SSAEIM). In [10], authors proposed to split the image into several patches and extract color features pertaining to each block using a CNN. A novel offline and online 3D deep learning integration framework by leveraging the 3D fully convolutional network (3D-FCN) was stated in [11]. A study in [12], proposed a novel method that is based on CNN to automatically recognize polyp in small bowel WCE image. They utilized the Alexnet architecture, one of the classical CNN, to ex-

*e-mail: charfisaid@gmail.com

**e-mail: melansari@gmail.com

***e-mail: a.ellahyani@uiz.ac.ma

****e-mail: eljaafari.ilyas@gmail.com

tract the features of WCE images and classify polyp images from normal ones. He et al. [13] proposed a novel deep hookworm detection framework (DHDF) for WCE images. This new framework simultaneously models visual appearances and tubular patterns of hookworms. In [14], authors proposed a method for WCE videos summarization based on a hybrid unsupervised method using long short-term memory (LSTM), variational autoencoder (VAE), pointer network (Ptr-Net), generative adversarial network (GAN), and de-redundancy mechanism (DM) techniques. In [15], authors used deep CNN and metric learning with Triplet Loss function for polyp detection. Mohebbian et al. [16] exposed an approach based on learning single models using samples from only one class, and ensemble all models for multiclass classification is proposed. In the first step, deep features are extracted based on an autoencoder architecture from the pre-processed images. Then, these features are oversampled using Synthetic Minority Over-sampling Technique and clustered using Ordering Points to Identify the Clustering Structure. To create one-class classification model, the Support Vector Data Descriptions are trained on each cluster with the help of Ant Colony Optimization, which is also used for tuning clustering parameters for improving F1-score. This process is applied on each class and ensemble of final models used for multiclass classification. Training a CNN model, from scratch, for WCE images classification was introduced by researchers in [17]. Their method aims to multi diseases detection i.e., normal mucosa, bile predominant, air bubbles, debris, inflamed mucosa, a typical vascularity, and bleeding. Coelho et al. [18] presented a method for red lesion detection Using U-net architecture. An approach for intestinal hemorrhage detection was stated in [19]. The scheme is based on CNN architecture after reducing the color palette using minimum variance quantization. A framework for best features selection was introduced in [20], the authors presented a fully automated system for stomach infection recognition based on deep learning features fusion and selection. In this design, ulcer images are assigned manually and support to a saliency-based method for ulcer detection. Later, pre-trained deep learning model named VGG16 is employing and re-trained using transfer learning. Features of re-trained model are extracted from two consecutive fully connected layers and fused by array-based approach. Besides, the best individuals are selected through the meta-heuristic approach name PSO along mean value-based fitness function. The selected individuals are finally recognized through Cubic SVM. Ellahyani et al. [21] proposed a method using ELM. They extract extracted HOG features from images in HSV colour space. In [20], a novel semi-supervised learning method with Adaptive Aggregated Attention (AAA) module for automatic WCE image classification is proposed. Firstly, a novel deformation field based image preprocessing strategy is proposed to remove the black background and circular boundaries in WCE images. Then they proposed a synergic network to learn discriminative image features, consisting of two branches The first branch consists of an estimator of abnormal regions and a distiller of abnormal information.

The first branch employs the proposed AAA module to extract global dependencies and incorporate context information to highlight the most meaningful regions. The second branch focuses on these calculated attention regions for accurate and robust abnormality classification. Finally, these two branches are jointly optimized by minimizing the proposed discriminative angular (DA) loss and Jensen-Shannon divergence (JS) loss with labeled data as well as unlabeled data. Hajabdollahi et al. [22], proposed a bifurcated structure with one branch performing classification, and the other performs the segmentation. Initially, separate network structures are trained for each abnormality separately and then primary parts of these networks are merged. A research in PubMed for all original publications on the subject of deep learning applications in WCE was conducted in [23]. A survey on deep learning models for WCE image analysis was carried out in [24]. In [25], a parametric rectified nonlinear unit activation function was stated.

The rest of the paper is organized as follows. Section 2 details the new proposed method for red lesion detection through the WCE images. Experimental results are illustrated in section 3. At the end, the paper is concluded in section 4.

2 Proposed method

In this section, the method used for red lesion detection is presented. The model used in this study is the Dense-Unet architecture, proposed in [26] to segment images. The Dense-UNet, which is based on U-net structure, employed the dense concatenation to deepen the depth of the network architecture and achieve feature reuse. This model included four expansion modules (each module consisted of four down-sampling layers) to extract features. Figure 1 depicts the model architecture. It can be seen as a two symmetrical paths i.e., upsampling and downsampling paths. These two paths are concatenated at the end. The Dense-UNet contains 10 dense_blocks. These blocks can be split into 5 dense_blocks in the dense upsampling path and 5 dense_blocks in the dense downsampling path. Each dense_block composes of 4 densely layers connected and having the same feature size [26].

3 Experimental Results

In this section, the obtained results of the proposed approach are presented. Accuracy and loss have been employed to assess the choices of the Dense-Unet model. First, the dataset exploited in this work is introduced. Then, the results of the red lesion detection using Dense-Unet are given. Finally, comparison between Dense-Unet and Unet is carried out.

3.1 Dataset

The dataset used in this work was first presented in [18]. It contains 3,295 images with red lesion and 3,295 annotated masks, each corresponding to an abnormal image. The size

Figure 1. Schematic diagram of Dense-UNet architecture [26].

of the images is 320×320 pixels. The dataset was split into training and testing sets. For training we have used 2,636 images and the remaining ones for testing. Fig. 2 shows some of the images of the dataset.

3.2 Implementation Details

The Dense-UNet network was trained from scratch with red lesion dataset, which was split randomly in 80% for training and 20% for validation, to detect and segment red lesions. The training was made using binary cross entropy as cost function, in 1 cycle of 20 epochs with the Adam optimizer. The learning rate was $1E-4$. The model evaluation was performed by comparing its predictions with the annotated masks, used as ground truth, based on accuracy metric. The network was implemented in Python 3.6 and all experiments were performed on a machine provided by Google Colab with GPU and 12GB RAM. The Dense-UNet was implemented using Keras with TensorFlow as backend.

3.3 Results

Before exposing the results obtained by the proposed method, definition of the measures employed to assess the performance is given below.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total samples}} \quad (1)$$

$$Loss = -\frac{1}{N} \times \sum_{i=1}^N y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i) \quad (2)$$

As mentioned before, the proposed approach to detect red lesion in WCE images, is carried out using Dense-UNet. This model achieved 96.10% in accuracy and a loss of 0.1761 in the training set. In the validation set, the model reached 0.3118 and 92.79% in the loss and accuracy, respectively. Figure 3, shows some test images in the segmented red lesion region.

3.4 Comparison with state-of-the-art methods

In order to assess the obtained results, we compared the proposed approach with state-of-the-art one. For this purpose, we compared the proposed method with the one presented in [18] since the authors used the same dataset in their work. The method presented in this work used U-net architecture for red lesion detection. Table 1 shows the results obtained for both approaches. As can be seen in the table the proposed method outperformed the method in [18] with 0.22%.

Table 1. Comparison between the proposed method and the published method using red lesion dataset.

Method	Accuracy	Loss
Method in [18]	95.88%	-
The proposed method	96.10%	0.1761

4 Conclusion and perspectives

In this paper, we presented the preliminary results of red lesion detection using Dense-UNet architecture in WCE images. We compared the results obtained with a literature approach in the same dataset and the performance is

Figure 2. Samples from the dataset and their corresponding annotated masks.

Figure 3. Samples from the dataset and their corresponding detected red lesion region.

satisfactory. Hence, we can conclude that the proposed scheme shows promising results in terms of accuracy and loss. It achieved 0.1761, 96.10% in terms of loss and accuracy, respectively. In the future work, we are planning to improve the robustness of the method by applying transfer learning. Moreover, fine tuning various parameters will lead to further improvements.

References

- [1] T. Rahim, M.A. Usman, S.Y. Shin, Computerized Medical Imaging and Graphics p. 101767 (2020)
- [2] M. Souaidi, M. El Ansari, IET Image Processing **13**, 2233 (2019)
- [3] S. Charfi, M. El Ansari, Soft Computing **24**, 4469 (2020)
- [4] Y. Zou, L. Li, Y. Wang, J. Yu, Y. Li, W. Deng, *Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network*, in *Digital Signal Processing (DSP), 2015 IEEE International Conference on* (IEEE, 2015), pp. 1274–1278
- [5] J.s. Yu, J. Chen, Z. Xiang, Y.X. Zou, *A hybrid convolutional neural networks with extreme learning machine for WCE image classification*, in *Robotics and Biomimetics (ROBIO), 2015 IEEE International Conference on* (IEEE, 2015), pp. 1822–1827

- [6] R. Zhu, R. Zhang, D. Xue, *Lesion detection of endoscopy images based on convolutional neural network features*, in *Image and Signal Processing (CISP), 2015 8th International Congress on* (IEEE, 2015), pp. 372–376
- [7] S. Seguí, M. Drozdal, G. Pascual, P. Radeva, C. Malagelada, F. Azpiroz, J. Vitrià, *Computers in biology and medicine* **79**, 163 (2016)
- [8] X. Jia, M.Q.H. Meng, *A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images*, in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the* (IEEE, 2016), pp. 639–642
- [9] Y. Yuan, M.Q.H. Meng, *Medical physics* **44**, 1379 (2017)
- [10] A.K. Sekuboyina, S.T. Devarakonda, C.S. Seelamantula, *A convolutional neural network approach for abnormality detection in Wireless Capsule Endoscopy*, in *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on* (IEEE, 2017), pp. 1057–1060
- [11] L. Yu, H. Chen, Q. Dou, J. Qin, P.A. Heng, *IEEE journal of biomedical and health informatics* **21**, 65 (2017)
- [12] L. Xu, S. Fan, Y. Fan, L. Li, *Automatic detection of small bowel polyp in wireless capsule endoscopy images*, in *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications* (International Society for Optics and Photonics, 2018), Vol. 10579, p. 1057919
- [13] J.Y. He, X. Wu, Y.G. Jiang, Q. Peng, R. Jain, *IEEE Transactions on Image Processing* **27**, 2379 (2018)
- [14] L. Lan, C. Ye, *Knowledge-Based Systems* **222**, 106971 (2021)
- [15] P. Laiz, J. Vitrià, H. Wenzek, C. Malagelada, F. Azpiroz, S. Seguí, *Computerized Medical Imaging and Graphics* **86**, 101794 (2020)
- [16] M.R. Mohebbian, S.S. Vedaiei, K.A. Wahid, P. Babyn, *arXiv preprint arXiv:2103.08508* (2021)
- [17] S.H. Kim, Y. Hwang, D.J. Oh, J.H. Nam, K.B. Kim, J. Park, H.J. Song, Y.J. Lim (2020)
- [18] P. Coelho, A. Pereira, M. Salgado, A. Cunha, *A deep learning approach for red lesions detection in video capsule endoscopies*, in *International Conference Image Analysis and Recognition* (Springer, 2018), pp. 553–561
- [19] H.S. Pannu, S. Ahuja, N. Dang, S. Soni, A.K. Malhi, *Multimedia Tools and Applications* **79**, 21941 (2020)
- [20] M.A. Khan, S. Kadry, M. Alhaisoni, Y. Nam, Y. Zhang, V. Rajinikanth, M.S. Sarfraz, *IEEE Access* **8**, 132850 (2020)
- [21] A. Ellahyani, I. El Jaafari, S. Charfi, M. El Ansari, *Signal, Image and Video Processing* pp. 1–8 (2020)
- [22] M. Hajabdollahi, R. Esfandiarpour, E. Sabeti, N. Karimi, S.R. Soroushmehr, S. Samavi, *Biomedical Signal Processing and Control* **57**, 101792 (2020)
- [23] S. Soffer, E. Klang, O. Shimon, N. Nachmias, R. Eliakim, S. Ben-Horin, U. Kopylov, Y. Barash, *Gastrointestinal endoscopy* (2020)
- [24] P. Muruganantham, S.M. Balakrishnan, *International Journal of Cognitive Computing in Engineering* **2**, 83 (2021)
- [25] I. El Jaafari, A. Ellahyani, S. Charfi, *Signal, Image and Video Processing* **15**, 241 (2021)
- [26] S. Cai, Y. Tian, H. Lui, H. Zeng, Y. Wu, G. Chen, *Quantitative imaging in medicine and surgery* **10**, 1275 (2020)