CONVOLUTIONAL NEURAL NETWORKS FOR INTESTINAL HEMORRHAGE DETECTION IN WIRELESS CAPSULE ENDOSCOPY IMAGES

Panpeng Li[1], *Ziyun Li*[2], *Fei Gao*[1], *Li Wan*[3], *Jun Yu*[1]

[1] Key Laboratory of Complex Systems Modeling and Simulation,
School of Computer Science and Technology, Hangzhou Dianzi University
[2] School of Communication Engineering, Hangzhou Dianzi University
[3] College of Computer Science, Chongqing University
lipanpeng@yeah.net, liziyun2014@gmail.com, {gaofei, yujun}@hdu.edu.cn, wanli@cqu.edu.cn

ABSTRACT

Wireless capsule endoscopy (WCE) can painlessly capture a large number of images inside the intestine. However, only a small portion of these WCE images contain hemorrhage. It is thus critical to develop automated hemorrhage detection method to facilitate the diagnosis of intestinal diseases. However, automated hemorrhage detection is complicated by 1) the extreme imbalance between the amount of hemorrhage images and that of normal images; and 2) the variety of the appearance, texture, and luminance inside the intestine. In this paper, we proposed to learn a robust intestinal hemorrhage detection model via Convolutional Neural Networks (CNNs), because of CNNs' extraordinary performance in solving various image understanding tasks. Specially, we explored different CNN architectures and data augmentation methods. Besides, we investigated the correlation between hemorrhage detection accuracy and image quality. Across about $1.3k$ hemorrhage images and $40k$ normal images, the learned CNN model achieves an F-measure of 98.87%.

Index Terms— Automated hemorrhage detection, convolutional neural networks, deep learning, gastrointestinal, wireless capsule endoscopy

1. INTRODUCTION

Wireless capsule endoscopy (WCE) enjoys increasing popularity now, because it allows painless endoscopy imaging of the whole intestine. The capsule can be easily swallowed and cause no discomfort [1]. Besides, the wireless endoscope takes thousands of color images during its journey in the intestine. On one hand, the WCE can thus provide sufficient details about the intestine, such as hemorrhage, which are essential symbols of various intestinal diseases. On the other hand, it is time-consuming for physicians to look through all the WCE images. In addition, sometimes hemorrhage is present in only a few images and may be missed by the physicians because of oversight. As a result, it is critical to develop automated hemorrhage detection algorithms to reduce the burden of physicians and benefit the early diagnosis of intestinal diseases [2].

This work was supported in part by the Zhejiang Provincial Science Foundation under Grants LQ16F030004 and LR15F020002, by the National Natural Science Foundation of China under Grants 6147210, 61501349, 61501349, 61601158, and 61602136.

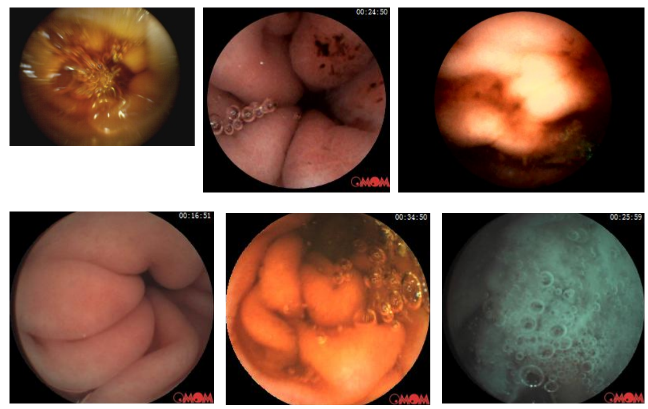

Fig. 1. Illustration of hemorrhage intestinal images (top) and normal ones (bottom), captured by WCEs.

Fig. 1 illustrates several intestinal WCE images. The left image is with hemorrhage; the right one is normal. There are great varieties in the appearance, texture, and illumination in these images. Besides, the hemorrhage regions vary greatly in term of size and pattern. Furthermore, the amount of hemorrhage images is much less than that of normal images. This imbalance makes it difficult to learn a robust model that can work well for both types of images. Thus, automated hemorrhage detection is complicated and challenging.

To date, several attempts have been made to design automated intestinal hemorrhage detection methods for the WCE images. These methods typically focus on extracting low-level features, such as raw pixel values [3], statistical features [4] or texture features [5, 6], to discriminate the hemorrhage regions and normal regions. Afterwards, traditional machine learning techniques, e.g. the Support Vector Machines

978-1-5090-6067-2/17/$31.00 ©2017 IEEE

(SVMs) [3, 7], are adopted to perform a binary classification of normal or hemorrhage images. However, these handcrafted features have been demonstrated limited for complicated image understanding tasks [8]. It is necessary to learn robust models that work well in limiting conditions.

Recently, deep learning technologies, especially Convolutional Neural Networks (CNNs), have shown extraordinary performance in various image understanding tasks [9]. Researchers therefore explore CNNs in the task of medical image analysis [10]. As expected, CNNs lead to distinct improvement in performance. Thus, we proposed to exploit CNNs for learning intestinal hemorrhage detection models.

In this paper, we explored different CNN architectures, including LeNet [11], AlexNet [9], GoogLeNet [12], and VGGNet [8]. These architectures are different in the input format, depth, and modules. In addition, to improve the robustness of the learned model, we augmented the data with various transformations, including rotation, blurring, luminance change, and Poison noise. Besides, we investigated the correlation between hemorrhage detection accuracy and image quality. To our knowledge, this is the first usage of CNNs for intestinal hemorrhage detection in WCE images. Across about $1.3k$ hemorrhage images and $40k$ normal images, the proposed method achieves an F-measure of 98.87%.

The rest of this paper is organized as below. Section 2 briefly reviews the existing works about intestinal hemorrhage detection and deep learning based medical image analysis. Section 3 details the proposed intestinal hemorrhage detection method via CNNs. Experimental results are presented and analyzed in Section 4. Section 5 concludes this paper.

2. RELATED WORK

In this section we briefly review the related works about 1) intestinal hemorrhage detection in capsule endoscopy images; and 2) the usage of deep learning in medical image analysis.

Intestinal hemorrhage detection. Traditionally, researchers adopt the raw pixel values or their statistics [3] [13] to discriminate the hemorrhage regions from the normal regions. For example, Liu and Yuan [3] explore three types of features, i.e. the raw pixel values, color histogram, and the pixel-wise ratio different color components. Afterwards, they use a SVM to predict whether a WCE image is normal or bleeding based on these features. Additionally, researchers [5] [6] combine the color texture feature and uniform local binary patterns (LBPs) to represent WCE images, and use the multilayer perceptron neural network or SVM to detect the bleeding regions. For example, Hwang [14] combines color features and Gabor features, and then takes the idea of bag-of-visual-words for feature learning. More recently, Hassan and Haque [7] extract features from the discrete Fourier transform (DFT) coefficients of the WCE images, and then use a SVM to formulate a classifier.

Medical image analysis via deep learning. Recently, researchers begin to explore deep learning technologies in the tasks about computer-aided medical image analysis [15, 16]. For example, Grinsven et al. [10] utilize a 9-layer CNN for bleeding detection in fundus images, and speed up the CNN training by dynamically selecting misclassified negative samples during training. More recently, Tajbakhsh et al. [15] explore the usage of CNNs in four distinct tasks, e.g. polyp detection in colonoscopy images. They find that it is promising to fine-tune a CNN model (pre-trained in natural image processing tasks) for medical image analysis.

Encouraged by the success of deep learning techniques, especially CNNs, in medical image analysis, we proposed to learn robust intestinal hemorrhage detection models via CNNs. Details are introduced in the following section.

3. THE PROPOSED METHOD

In this paper, to mitigate the overfitting risk, we first augmented the data with various transformations. Afterwards, we explored a number of classic CNN architectures, and trained them in the intestinal hemorrhage detection task. Details are sequentially represented in the following subsections.

3.1. Data augmentation

During the journey inside the intestine, a wireless capsule endoscope might take photos of the local intestine regions from random and arbitrary angles and distances. As illustrated in Fig. 1, the WCE images might be blurring due to poor focusing, or contain bubbles. Besides, the illuminance condition varies greatly, the resulted images might be dark, low-contrast, and contain Poisson noise. The learned model should be capable of dealing with such degraded images. In this paper, we formulated these degradations by using the following transformations. Fig. 2 shows one original image and its transformed images.

Rotation: We rotated the image by an angle of α° with the nearest-neighbor interpolation. Here, α is uniformly chosen from 60 to 300 with a step of 60. As a result, we obtained 5 rotated versions of each original image.

Luminance change: We changed the luminance of each image by a random factor. Specially, we first converted the image from the RGB space to the YCbCr space. Afterwards, we uniformly and randomly chose one scalar λ from the interval of $[0, 1]$, and then multiplied λ with the Y channel. Finally, we converted the modified image to the RGB space and saved it into an image file. We repeated this process five times, and thus obtained five versions of the original image. Here, we chose random factors instead of fixed ones to make the luminance of the augmented images spread across the whole possible value space. As shown in Fig. 2, the luminance of the modified images might be very low.