Deep convolutional neural network-based automated lesion detection in wireless capsule endoscopy

Jeon, Yejin, Cho, Eunbyul, Moon, Sehwa, Chae, Seung-Hoon, Jo, Hae Young, et al.

Yejin Jeon, Eunbyul Cho, Sehwa Moon, Seung-Hoon Chae, Hae Young Jo, Tae Oh Kim, Chang Mo Moon, Jang-Hwan Choi, "Deep convolutional neural network-based automated lesion detection in wireless capsule endoscopy," Proc. SPIE 11050, International Forum on Medical Imaging in Asia 2019, 110501N (27 March 2019); doi: 10.1117/12.2522159

Event: 2019 Joint International Workshop on Advanced Image Technology (IWAIT) and International Forum on Medical Imaging in Asia (IFMIA), 2019, Singapore, Singapore

Deep Convolutional Neural Network-Based Automated Lesion Detection in Wireless Capsule Endoscopy

Yejin Jeon[a*], Eunbyul Cho[a*], Sehwa Moon[a], Seung-Hoon Chae[b], Hae Young Jo[c], Tae Oh Kim[c], Chang Mo Moon[c¶], Jang-Hwan Choi[a¶]

[a]Division of Mechanical and Biomedical Engineering, Ewha Womans University, Seoul, Korea; [b]Department of Bio-Medical IT Convergence Research, Electronics and Telecommunications Research Institute, Daejeon, Korea; [3]Department of Internal Medicine, Ewha Womans University College of Medicine, Seoul, Korea; *These authors contributed equally to this work; [¶]These corresponding authors also contributed equal to this work

ABSTRACT

Because most of the capsule-endoscopic images contain normal mucous membranes, physicians spend most of their reading time observing normal areas. Thus, a significant reduction in their reading time would be possible if only the portion of the image frame for which a particular lesion is suspected can be read intensively. This study aims to develop a deep convolutional neural-network-based model capable of automatically detecting lesions in the capsule-endoscopic images of a small bowel. The proposed model consists of two deep neural networks in parallel, each of which takes in images in RGB and CIELab color spaces, respectively. The neural-networks model is based on transfer-learned GoogLeNet architecture. Our proposed algorithm showed promising results in classifying endoscopic images where lesions exist (98.56% accuracy). If the proposed algorithm is used to screen abnormal images, it is expected to reduce a physician's reading time and to improve his/her reading accuracy.

Keywords: Deep neural networks; Convolutional neural Networks; Wireless capsule endoscopy; Lesion detection; Small bowel tumor, Small-bowel Wireless Capsule Endoscopy

INTRODUCTION

Development of wireless capsule endoscopy (WCE) provides superior endoscopic visualization of the small intestine by sending images taken from a wireless capsule to an external data-collecting device. However, the challenge now is that tens of thousands of data per patient are generated, leading to long clinical reading times. Long reading times also increase reader fatigue, which reduces reading accuracy. To solve this problem, many lesion classification and detection algorithms using image processing technology based on existing machine learning have been proposed [1, 2]. Recently, algorithms based on deep learning techniques have been published and reported superior performance over existing processing techniques [3]. However, most deep learning-based studies are focused on colonoscopies that have a relatively large amount of open data. Also, most of the studies deal mainly with polyps [3] which are relatively easy to recognize morphologically or Angiodysplasia [4, 5] lesions that are clearly bordered by color. In this study, we propose an algorithm that can detect the abnormality by applying the deep learning algorithm to various lesions in the WCE images.

International Forum on Medical Imaging in Asia 2019, edited by Feng Lin, Hiroshi Fujita, Jong Hyo Kim, Proc. of SPIE Vol. 11050, 110501N · © 2019 SPIE
CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2522159

METHODS AND MATERIALS

Dataset Preparation and Patch Generation

We used 431 WCE RGB images from 52 patients with lesions of different size and shape to generate patches. We transformed the colors of WCE images to CIELab representations using ImageJ, and resized both RGB and CIELab images to 576 × 576 for consistency. Afterwards, we created patches in the same manner for RGB and CIELab images.

- **Implementation of CIELab Color Spaces**: We transformed the colors of WCE images to CIELab representations to obtain more perceptual representations by separating intensity from chromatic image components [1]. We combined the softmax output units of RGB and CIELab images for classification into normal and abnormal, by a support vector machine learning algorithm, to improve performance of lesion detection model.

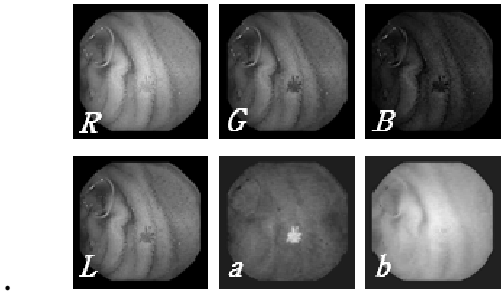

Figure 1. Original red, green, blue (RGB) wireless capsule endoscopy image in R, G, B (first row). Color transformation of RGB image components to CIE Lab image components in L, a, b (second row).

- **Normal patch**: We created normal patches that does not contain any lesioned parts in the ground truth label images. The resulting normal patches were the same size as 64 × 64.

Figure 2. Representative normal patches of RGB and CIELab images in (A) and (B), respectively.

- **Abnormal patch**: We first obtained the four vertex coordinates of the bounding box containing the lesion through the label images to create abnormal patches. Using the coordinate values, abnormal patches are cropped in 8 ways: a tight box, boxes with 3 pixel margin, 5 pixel margin, 10 pixel margin and boxes with 3 pixel shift to the upper left, 3 pixel shift to the down right.

- **Model Regularization**: 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 [2]. Furthermore, prior studies demonstrated that data augmentation improves deep learning performance, a process of synthetically generating additional training examples by using image transformations. Doing so forces the learned model to can deal with degraded images and improve performance. We used horizontal and vertical flipping, and a 90 degrees rotation on the left and right. The other technique we used to reduce over-fitting is 'early stopping', where a validation set is reserved exclusively for monitoring the CNN's accuracy during training, and the weights of the network at the point of best performance are saved, as opposed to the weights obtained at the end of training [3]. Afterwards, we restored the saved graph and weights for testing.

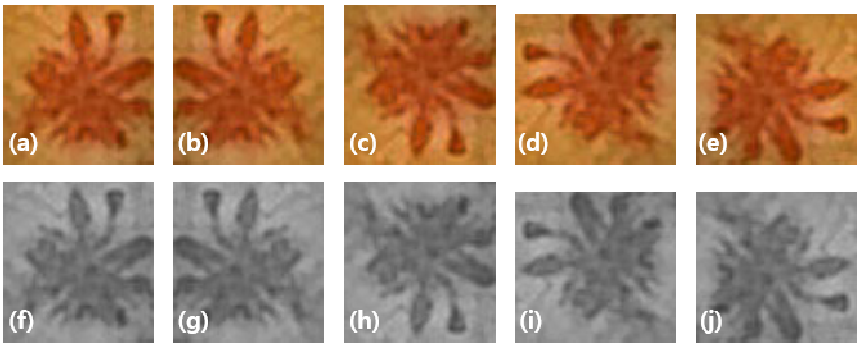

Figure 3. (a)-(e) are RGB patches, original in (a), horizontal flip in (b), vertical flip in (c), 90 degrees left rotation in (d), 90 degrees right rotation in (e). And (f)-(j) are CIELab patches, original in (f), horizontal flip in (g), vertical flip in (h), 90 degrees left rotation in (i), 90 degrees right rotation in (j).

CNN Model Architecture and Training

We trained the GoogLeNet model, which has been pre-trained on the ImageNet data, before refining the weights on our data set. We added the final layer with three FC (fully-connected) layers, a softmax layer, and cross-entropy loss. We used the Adam optimizer to update the parameters in this CNN model. All experiments were implemented using Tensorflow software libraries.

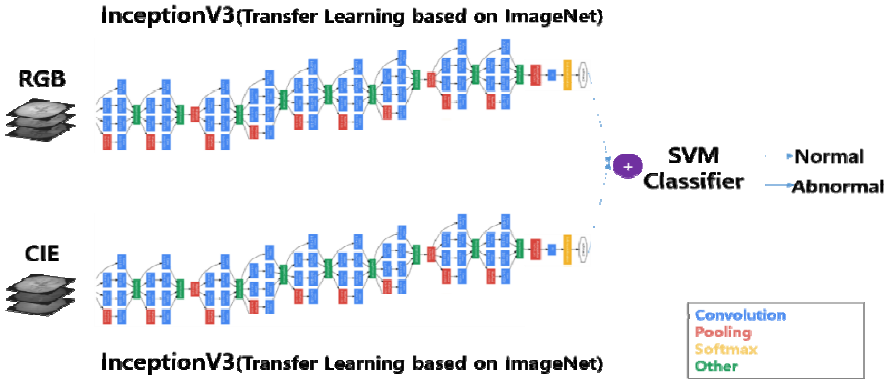

Figure 4. Deep neural network architecture.

The normal patches are randomly extracted from the dataset of normal patches so that they are 1:1 with the number of abnormal patches. We randomly split the augmented dataset into three parts: 50% for training, 30% for validation, and the rest 20% for testing. There is no overlap in content between the three datasets. For early stopping and hyperparameter optimization, we set validation dataset for monitoring the model's performance. We trained and evaluated the lesion-detection model separately for RGB and CIELab images. In the training process, we set the batch size to 100, the learning rate to 0.01, and the iteration times to 1600 to avoid overfitting. Afterwards, the Support Vector Machine (SVM) is adopted to concatenate the Softmax results of RGB and CIELab images and to perform a binary classification of normal or abnormal images.

RESULTS

We summarize lesion detection results in Table 1. To evaluate the lesion detection performance, we compare the predicted labels with the ground truth labels and compute three measures, which are accuracy, true positive rate (TPR), false-positive rate (FPR). Greater values of the accuracy and TPR and lower FPR indicate better performance [2]. The accuracy and TPR obtained with concatenation of RGB and CIELab testing results are higher than those of each model which is trained with either RGB or CIELab images. However, the best model (RGB+CIE) has the highest false positive rate (FPR) of 2.57%. The improved performance is due to the fact that CIELab is a perceptually uniform color space that has proved to perform better than RGB for color texture analysis [6]. This property of CIELab improves the model performance of detecting bleeding regions in WCE images which show different extents of differences in color texture compared to its surrounding regions [5].

Table 1. Quantitative comparison results of different patches and neural networks.

Patch	Accuracy (%)	TPR (%)	FPR (%)
RGB	98.27	97.16	**0.61**
CIE	96.39	94.83	**2.06**
RGB+CIE	**98.56**	**99.70**	**2.57**

CONCLUSION

We have built the endoscopic lesion detection and classification model using a convolutional neural network that was first trained on the ImageNet corpus of natural images and then re-trained on our WCE images. Furthermore, the lesion detection ability of model can be enhanced by concatenating the Softmax results of RGB and CIELab using SVM. We can detect and classify endoscopic images where lesions exist with an accuracy of 98.56%.

ACKNOWLEDGEMENTS

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government, MSIP (grant no: NRF-2017R1C1B5018287, NRF-2015M3A9A7029725, and NRF-2017M2A2A6A02070522, URL: http://nrf.re.kr). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

[1] D. K. Iakovidis, and A. Koulaouzidis, "Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software," Gastrointest Endosc, 80(5), 877-83 (2014).

[2] J. Bernal, J. Sanchez, and F. Vilarino, "Towards automatic polyp detection with a polyp appearance model," Pattern Recognition, 45(9), 3166-3182 (2012).

[3] G. Urban, P. Tripathi, T. Alkayali *et al.*, "Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy," Gastroenterology, 155(4), 1069-+ (2018).

[4] A. Shvets, V. I. Iglovikov, A. Rakhlin *et al.*, "Angiodysplasia Detection and Localization Using Deep Convolutional Neural Networks," CoRR, abs/1804.08024, (2018).

[5] B. Li, and M. Q. Meng, "Computer-aided detection of bleeding regions for capsule endoscopy images," IEEE Trans Biomed Eng, 56(4), 1032-9 (2009).

[6] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis *et al.*, "Computer-aided tumor detection in endoscopic video using color wavelet features," IEEE Trans Inf Technol Biomed, 7(3), 141-52 (2003).