Polyp Classification Using Multiple CNN-SVM Classifiers from Endoscope Images

Masataka Murata Hiroyasu Usami Yuji Iwahori

Department of Computer Science
Chubu University
Kasugai, 487-8501 Japan
Email: {mmurata|usami}@cvl.cs.chubu.ac.jp
Email: iwahori@cs.chubu.ac.jp

Wang Aili

Higher Education Key Lab
Harbin University of Science and Technology
Harbin, 150080 China
Email: aili925@hrbust.edu.cn

Naotaka Ogasawara Kunio Kasugai

Department of Gastroenterology
Aichi Medical University
Nagakute, 480-1195 Japan
Email: {nogasa|kuku3487}@aichi-med-u.ac.jp

Abstract—**This paper proposes a classification approach of a malignant or bening polyp type by multiple CNN-SVM classifiers using Convolutional Neural Networks (CNN) as feature extractor and Support Vector Machine (SVM) as classifier from three kinds of endoscope images as white light image, dye image and Narrow Band Image (NBI). First, the polyp feature is extracted using CNN as feature extractor from three kinds of endoscope images using each datasets. Second, classifiers are generated as many as three kinds of combinations using SVM and each image is classified. Finally, the final classification result is judged by voting processing from the result obtained by each classifier. The effectiveness of the proposed method was confirmed through experiments in which both validity and accuracy of multiple CNN-SVM voting results were evaluated using actual malignant or benign polyp images.**

Keywords–Polyp Classification; Endoscope Image; Voting Processing; Pre-Trained Network; Convolutional Neural Network; Support Vector Machine.

I. Introduction

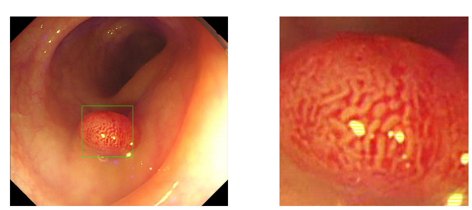
The polyp diagnosis is conducted using the endoscope in the medical scene, according to the prevalence rate of colorectal cancer has been increasing. There are various forms of polyps, such as protuberance type, surface flat type, surface recessed type and so on. These shapes are used as a reference when judging the malignancy/benignity of polyps. However, it is difficult to judge if a polyp is benign/malignant only by its shape, in some cases, and the diagnostic result of polyp using endoscope depends on the experience of the medical doctor. There are many cases where correct diagnosis is obtained by the medical doctor as the pathological diagnosis judges correctly. Therefore, it is necessary to develop a computer-aided system with computer vision technology to eliminate the difference in the diagnosis results from the experience of the doctor and to reduce the burden of the medical provider.

As a method to judge the malignant/benign polyp from endoscope images, some methods [1][2] have been proposed. In these methods [1][2], a ultra-high magnification endoscope is used for the polyp diagnosis with high precision. The ultra-high magnification endoscope has higher magnification than regular endoscope and it enables the diagnosis at the cell level. However, it requires a lot of diagnosis time when ultrahigh magnification is used, and this would put additional burden on the patient.

Therefore, this paper proposes a method to classify malignant or benign polyp using regular endoscope images.

Actually, there are many non-polyp scenes in endoscope video of the regular endoscope, which makes it difficult to classify the malignant or benign polyp. Therefore, for our proposed method, a necessary condition is that only the polyp images be used as the target. Paper [3] and [4] were proposed for polyp detection. These papers detect polyps with the rectangles (as shown in Figure 1). There are three types of images which are taken by the regular endoscope: with white light, dye and narrow band image (NBI) in general. These three kinds of images have different characteristics and the difficulty of classification level of malignant or benign polyp depends on the condition of each image. In this paper, the polyp region is extracted with the rectangle by methods [3][4] and three types of images taken by the regular endoscope are used for the classification. Accurate classification of malignant or benign polyp are tried from each image features for supporting the medical diagnosis.

Section II introduces the proposed method. Section III gives the result of our experiment. Finally, Section IV concludes the proposed method.



(a)Result of Detection   (b)Region of Rectangle

Figure 1. Detected Polyp

II. PROPOSED METHOD

Our proposed method uses features [5][6] obtained by pre-trained network for malignant or benign polyp classification. Specifically, each feature is extracted from Convolutional Neural Network (CNN) [7] using each of three kinds of images with white light, dye and NBI, respectively. Multiple Support Vector Machine (SVM) [8] is used for the classification of diagnosis using extracted CNN features.

The procedure of the proposed method is as follows (as shown in Figure 2).

Step 0 Assign labels to endoscope images.
Step 1 Extract CNN features obtained from each input image of three kinds of images.
Step 2 Construct multiple SVM classifiers using CNN features.
Step 3 Extract features for evaluation with CNN as Step 1.
Step 4 Classify malignant or benign polyp using multiple SVM classifiers constructed in Step 2 using features obtained in Step 3.
Step 5 Determine the final result by a voting process using the classified result of multiple SVM classifiers.

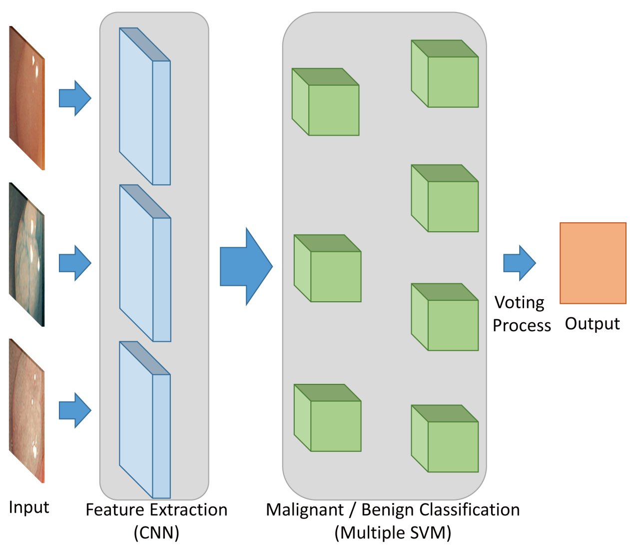


Figure 2. Flow of The Proposed Method

A. Assign Label to Endoscope Images

There are White Light (Figure 3(a)(d)), Dye (Figure 3(b)(e)) and NBI (Figure 3(c)(f)) that can be taken by the regular endoscope. These endoscope images have different characteristics and they have the following features.

White Light: Taken in normal condition.
Dye: Taken with indigo carmine stain solution or crystal violet stain solution sprayed on the polyp, and the irregularities of the lesion are emphasized.
NBI: Taken in the state irradiated with light which is easily absorbed by hemoglobin in the blood different from normal light and its blood vessels and patterns are emphasized around the lesion.

Assign labels to these image as malignant polyp (Figure 3(d)(e)(f)) or benign (Figure 3(a)(b)(c)) polyp and also assign labels on the types of the above endoscope images. Six kinds of labels are attached, as shown in the Figure 3.

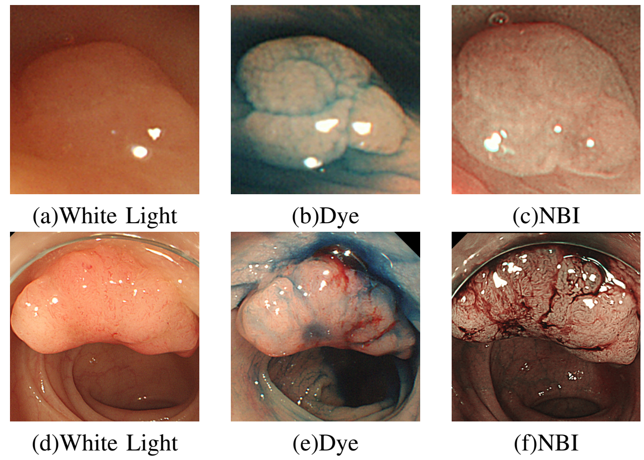


Figure 3. Endoscope Image

B. Feature Extraction Using CNN

Differences in polyp features are necessary to classify the malignancy/benignity of a polyp. However, it is difficult in general to use the empirical feature, such as Scale invariant feature transform (SIFT) [9] to classify malignancy/benignity polyp. CNN is highly evaluated as a feature extractor in recent years and the CNN feature is used for feature extraction in case of the polyp images. AlexNet [10] is used as a model of CNN for feature extraction and corresponding 4096-dimensional polyp features are extracted from each of the seventh layers among totally connected layers with input of each of three kinds of endoscopic images: white light, dye, and NBI.

1) Convolutional Neural Network: CNN is a network consisting of convolution layers that perform local feature extraction of images and pooling layers that collect extracted features where feature extraction and classification are performed in a network. Recently, it has been treated as a feature extractor by using only the feature extraction location, and it has been proved to have highly general versatility as a feature extractor.

2) AlexNet: AlexNet is a model learned for image classification using the classification task of ILSVRC 2012 and it is CNN consisting of 8 layers (as shown in Figure 4). This CNN model extracts features of 4096-dimensions for each input image and performs classification of 1000 classes. In this paper, feature extraction is obtained from the seventh layer as all connected layers of AlexNet.

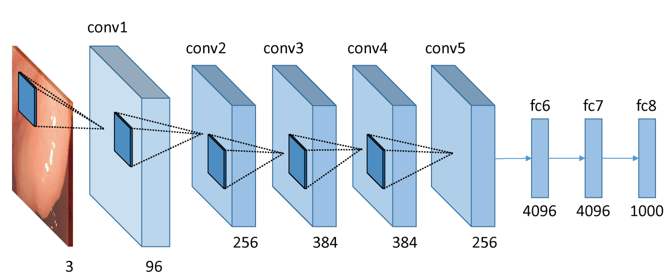


Figure 4. Alexnet layers

C. Construction of Classifiers Using Extracted Features

Classifiers of malignant or benign polyps are constructed using the extracted features described in Section II-B. SVM is used as classifier and it is constructed for three kinds of features consisting of white light, dye and NBI extracted from CNN, but the condition changes based on which image type is easy to be classified as malignant or benign polyp. Classifiers are constructed for the maximum number of combinations consisting of three kinds of features, and each classifier corresponds to each kind of image. Each classifier easily classifies malignant or benign polyp or not depending on polyp. Here, the input of each classifier is corresponding image features which were used when constructing each one. The output of each classifier is each diagnosis result of input images. Table I shows the combination type of features and the number of classifications.

TABLE I. COMBINATION

Combination of Features	*Number of Classifications*
White Light	1
Dye	1
NBI	1
White Light + Dye	2
White Light + NBI	2
Dye + NBI	2
White Light + Dye + NBI	3

D. Classification Result with Voting Processing

The result of each classifier constructed with the method from Section II-C may be different even for the same polyp depending on the kind of image. Therefore, the final result is determined by combining the results from each classifier. In the voting processing, classification score as the classification result obtained from each SVM is added to the evaluation score so that the reliability of the final score is improved rather than only handling one classification as one vote. Here, the approach handles the classification score as a weight of one vote. The calculation formula of the voting process is shown in Equation (1).

Here, ”Label” represents the classification score derived from Equation (2), ”Score” represents the classification score of the result classified by SVM, n represents the number of classification classes, ”Decision” represens the classification result of SVM, ”Benign” indicates probability of a benign polyp, ”Malignant” indicates probability of a malignant polyp.

$$Label = \sum_{n=0}^{12} Score_n \tag{1}$$

$$Label = \begin{cases} Benign & (if\ Decision = Benign) \\ Malignant & (otherwise) \end{cases} \tag{2}$$

Based on the probabilities of a benign polyp and the probability of a malignant polyp calculated by Equation (1), the final result is determined by the larger value as shown in Equation (3).

Here, ”result” represents the final result.

$$result = \begin{cases} Benign & (if\ Benign > Malignant) \\ Malignant & (otherwise) \end{cases} \tag{3}$$

As described above, voting processing is performed using classification scores from the results classified from seven classifiers. This solves the difficulties of classification derived from the difference of polyps. Simultaneously, the accuracy of classification becomes higher than classification by each classifier.

III. EXPERIMENT

Experiments were performed to validate the proposed method. The datasets used in the experiment were polyp images obtained as the rectangle detected by methods [3][4]. In order to increase the dataset, images were added with three types of rotation processing to the original image. In addition, since the label of the dataset of the learning image is unbalanced, undersampling on malignant/benign labels was performed in this experiment. Tables II and III show the number of the learning images and the test images, respectively.

TABLE II. TRAINIMAGE

	Malignant	*Benign*
White Light	188	380
Dye	112	408
NBI	32	140

TABLE III. TESTIMAGE

	Malignant	*Benign*
White Light	180	180
Dye	180	180
NBI	180	180

Table IV shows the kind of classifier consisting of each combination and correct/incorrect number of malignant and benign polyps with the voting processing. As evaluation of classification accuracy, each of *Sensitivity*, *Specificity*, *Accuracy*, *Positive Predictive Value* (*PPV*) and *Negative Predictive Value* (*NPV*) were calculated by the following formula.

TABLE IV. CLASSIFICATION RESULT

	Malignant		Benign	
	True	False	True	False
White Light	149	31	132	48
Dye	94	86	167	13
NBI	59	121	158	22
White Light + Dye	130	50	156	24
White Light + NBI	52	128	140	40
Dye + NBI	122	58	152	28
White Light + Dye + NBI	118	62	153	27
Poll Result	152	28	164	16

True Positive (*TP*) represents numbers that classified malignant as malignant. *False Negative* (*FN*) represents numbers that classified malignant as benign. *False Positive* (*FP*) represents numbers that classified benign as malignancy. *True Positive* (*TP*) represents numbers that classified benign as benign.

Sensitivity represents the validity that classified malignant as malignant. *Specificity* represents the validity that classified benign as benign. *Accuracy* represents the whole validity. *PPV*

TABLE V. ACCURACY EVALUATION

	Sensitivity	*Specificity*	*Accuracy*	*PPV*	*NPV*
White Light	75.6	80.9	78.0	82.7	73.3
Dye	87.8	66.0	72.5	52.2	92.7
NBI	72.8	56.6	60.2	32.7	87.7
White Light + Dye	84.4	75.7	79.4	72.2	86.6
White Light + NBI	56.5	52.2	53.3	28.8	77.7
Dye + NBI	81.3	72.3	76.1	67.7	84.4
White Light + Dye + NBI	81.3	71.1	75.2	65.5	85.0
Poll Result	90.4	85.4	87.7	84.4	91.1

represents positive predictive value that classified malignant as malignant. *NPV* represents positive predictive value that classified benign as benign.

$$Sensitivity = \frac{TP}{TP+FP} \quad (4)$$

$$Specificity = \frac{TN}{FN+TN} \quad (5)$$

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \quad (6)$$

$$PPV = \frac{TP}{TP+FN} \quad (7)$$

$$NPV = \frac{TN}{FP+FN} \quad (8)$$

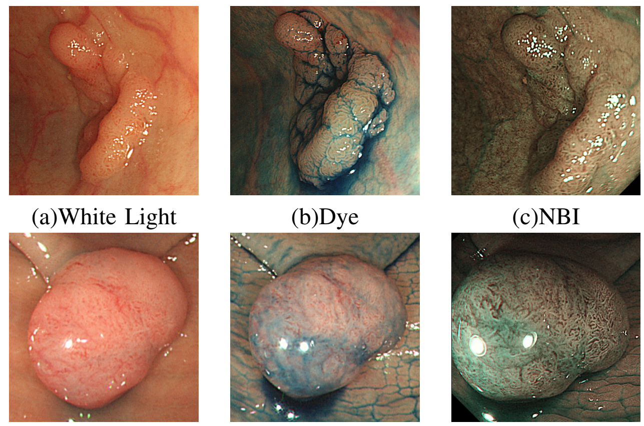
From Table IV, it is shown that the proposed method least misclassified the malignant polyps. In addition, Table V shows that both Sensitivity as validity of malignant polyp classification and PPV as predictive value of malignant polyp were obtained with high accuracy. When a malignant polyp was classified as a benign polyp, there would be a delay in polyp extraction that could become life-threatening. From these results, it is shown that the proposed method is useful for polyp diagnosis. Furthermore, the accuracy as the validity from all classifications shows high value in the proposed method. Error classification examples of benign polyp (a) (b) (c) and malignant polyp (d) (e) (f) are shown in Figure 5. A benign polyp has usually a round shape and a malignant polyp has a uneven shape with some feature on blood vessel. However, the polyps in Figure 5 have the opposite features and there is some possibility that this example is an incorrect classification result.

IV. CONCLUSION

In this paper, multiple CNN-SVM classifiers were constucted using three kinds of endoscope images taken by regular endoscope. The paper proposed a highly accurate classification method by integrating the results based on the voting processing. The effectiveness of the proposed method was confirmed via experiments using actual endoscopic images to classify malignant and benign polyps with CNN features and multiple SVM classifiers. As future work, some improvement is needed to reduce the misclassified polyps by increasing the number of dataset and constructing a specialized CNN model for endoscope images with fine tuning to get higher accuracy.

ACKNOWLEDGMENT

Iwahori's research is supported by Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (C) (23500228) and Chubu University Grant.



(a)White Light (b)Dye (c)NBI

(d)White Light (e)Dye (f)NBI

Figure 5. Example of Error Classification

REFERENCES

[1] Y. Mori et al., "Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos)." Gastrointestinal endoscopy, vol. 81, no. 3, 2015, pp. 621–629.

[2] M. Misawa et al., "Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy." Gastroenterology, 2016.

[3] H. Hiroaki, I. Yuji, and K. Kunio, "Automatic polyp detection from endoscope image using likelihood map based on edge information (in japanese)," in IEICE Technical Report, vol.115, no.401, 2015, pp. 193–198.

[4] U. Hiroyasu, O. Tsubasa, Yuji, Iwahori, and K. Kunio, "Automatic polyp region detection using watershed algorithm (in japanese)," in WiNF 2016, 2016, pp. B–22X.

[5] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "Cnn features off-the-shelf: an astounding baseline for recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014, pp. 806–813.

[6] J. Donahue et al., "Decaf: A deep convolutional activation feature for generic visual recognition." in ICML, 2014, pp. 647–655.

[7] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, 1998, pp. 2278–2324.

[8] B. Schölkopf and A. J. Smola, Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press, 2002.

[9] D. G. Lowe, "Object recognition from local scale-invariant features," in Computer vision, 1999. The proceedings of the seventh IEEE international conference on, vol. 2. IEEE, 1999, pp. 1150–1157.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

 ISBN: 978-1-61208-534-0