JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

# Class-Labeling Method for Designing a Deep Neural Network of Capsule Endoscopic Images Using a Lesion-Focused Knowledge Model

Ye-Seul Park* and Jung-Won Lee*

## Abstract

Capsule endoscopy is one of the increasingly demanded diagnostic methods among patients in recent years because of its ability to observe small intestine difficulties. It is often conducted for 12 to 14 hours, but significant frames constitute only 10% of whole frames. Thus, it has been designed to automatically acquire significant frames through deep learning. For example, studies to track the position of the capsule (stomach, small intestine, etc.) or to extract lesion-related information (polyps, etc.) have been conducted. However, although grouping or labeling the training images according to similar features can improve the performance of a learning model, various attributes (such as degree of wrinkles, presence of valves, etc.) are not considered in conventional approaches. Therefore, we propose a class-labeling method that can be used to design a learning model by constructing a knowledge model focused on main lesions defined in standard terminologies for capsule endoscopy (minimal standard terminology, capsule endoscopy structured terminology). This method enables the designing of a systematic learning model by labeling detailed classes through differentiation of similar characteristics.

## Keywords

Capsule Endoscopy, Class-labeling Method, Deep Learning, Knowledge Base (KB), Ontology

# 1. Introduction

Capsule endoscopy (CE) is one of the increasingly demanded diagnostic methods in recent years because, compared with general endoscopy, it has the advantage of being inserted into the human body without pain or discomfort [1]. In particular, because it enables a visualization of the entire digestive tract from the oral cavity to the anus, it can be used to observe the small intestine, which is difficult through a general endoscope. Thus, it may be used to identify gastrointestinal (GI) conditions in patients with small-bowel diseases such as Crohn's disease or celiac disease [2-4]. Generally, CE has three components: a disposable capsule serving to acquire or transmit an image, a terminal to receive or store the acquired signal, and an image processing unit used to analyze the stored image. Often, when CE is completed, the image stored in the receiver is transferred to the image processing unit, and the position of the capsule corresponding to the endoscopic frame or medical information on the lesion are analyzed. Currently, CE has been designed to automatically acquire such significant information through deep learning.

First, there are several studies to track the position of the capsule by learning the entire endoscopic image. As Fig. 1 shows, current software for capsule endoscope provides a service to track the position of the capsule. However, the technology for tracking capsules has poor completion because it is based on the differences between the electrodes transmitted from the capsule, complicating the tracking of the exact position due to overlapping of the received signals. Thus, studies are being conducted to track the location of capsules by applying deep learning. For example, [5] proposed a cascaded spatial-temporal deep framework, a method in which a frame with noise is extracted from the entire frame and subsequently classified into the mouth, esophagus, stomach, small intestine, and large intestine. However, because the digestive organs cannot be distinguished in the endoscopic image, these studies have not yet progressed actively. Second, studies are being conducted on automated algorithms to extract meaningful features such as polyps or tumors from capsule endoscopic images. In particular, studies on a deep learning framework that has achieved remarkable performance have been recently proposed. In [6], a deep neural network (DNN) model for recognizing polyps in capsule endoscopic images was designed to analyze or acquire various medical information. Furthermore, in [7], a study on detecting the gastrointestinal bleeding by using a convolutional neural network (CNN) was conducted. Finally, some studies to detect features specific to small-bowel-related diseases, the major target diseases of CE, have been conducted. [8] applied a deep learning technique to quantitatively evaluate the evidence frame for celiac disease, which may develop in the small intestine.
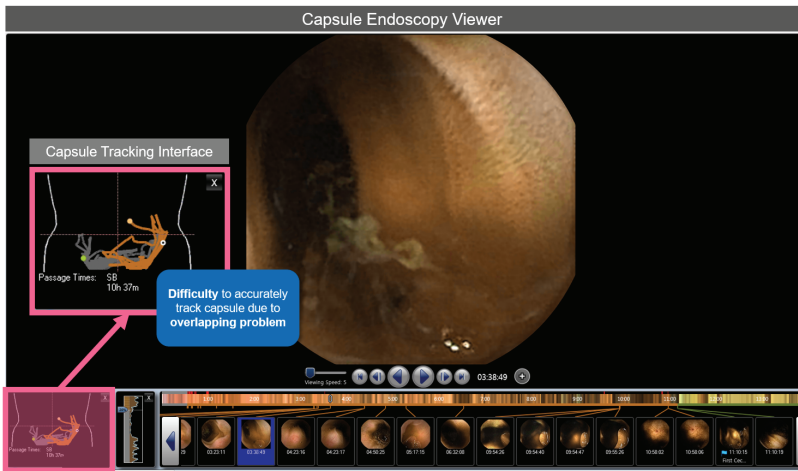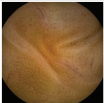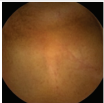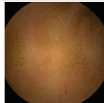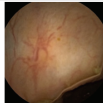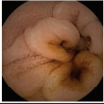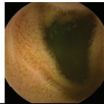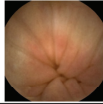


**Fig. 1.** Overlapping problem of tracking capsules identified in a current capsule endoscopy viewer (RAPID Reader Software v8 for PillCam).

However, the training images that are used in the majority of learning models do not consider various attributes (degree of wrinkles, presence of valves, etc.) that can be obtained from the capsule. As Fig. 2 shows, the capsule endoscopic image can be distinguished by the location of the organ, and by the attributes that can be identified in the image. For example, properties such as wrinkles, bubbles, and excreta (type 4, feces) are some features that can distinguish images in addition to endoscopic position. Grouping and labeling images with similar features improves the performance of machine learning or deep learning models. In other words, these characteristics should be considered when designing a learning model. Nevertheless, in conventional approaches, these attributes are included or limited to one

"normal class". Similarly, "abnormal classes" contain a few lesioned frames; hence, class labeling is performed without distinguishing the characteristics of certain lesions (e.g., normal or bleeding).

Therefore, in this paper, we propose a class-labeling method that can be used to design a learning model by constructing a lesion-focused knowledge model. The proposed method describes some discovery information ("findings") in capsule endoscopy, based on two CE-related standards—the minimal standard terminology (MST) for gastrointestinal endoscopy [9] and the capsule endoscopy structured terminology (CEST) [10]. These provide specific criteria for the classes that can be defined in the output layer of a learning model such as a DNN. In other words, it enables a systematic design of a learning model and improves its performance by distinguishing the attributes or characteristics of the capsule endoscopic images.



**Fig. 2.** Various properties of capsule endoscopic frames.

## 2. Related Work

In recent years, image processing techniques using DNN models (among which are CNN models optimized for image) have been activated due to their superior performance in complicated calculations. This is also true in studies on CE, which is a field that relies on medical images. The majority of studies related to capsule endoscopic imaging are dominated by techniques that provide specific information obtained from images or associated with a particular disease. In this paper, related works are divided into the following two perspectives, as shown in Fig. 3.

First, there are general-purposed studies on the entire GI tract, which can be photographed using capsule endoscopy without distinction of specific diseases or organs. [5] proposed a method to track the location of the capsule by using a cascading deep learning framework that separates noisy frames from the full dataset and classifies frames into the mouth, stomach, small intestine, and colon. However, capsule endoscopic frames do not indicate distinctive similarities between organs; thus, creating an approach by cascaded model is still difficult. Thus, the majority of studies are being actively conducted to identify similarities between lesions. [6] used the deep learning technique to detect polyps, which can be observed in mucous membranes in the gastrointestinal tract. [11] surveyed recent studies on polyp

detection methods by classifying them as shape-based, texture-based, and hybrid techniques. According to [11], pre-processing was performed to extract features (widely known as "hand-crafted" features) and image classification was performed using the features extracted by a past technique. Moreover, studies have been actively conducted on CNNs, which perform image classification by using automatically learned features. Additionally, studies have been conducted to observe bleeding in the GI lining through capsule endoscopic frames, and similar studies have been conducted to detect bleeding in the GI mucosa [12-15]. Furthermore, [16] was conducted to detect the presence of parasites in the GI tract by identifying a localized feature called "tubular".
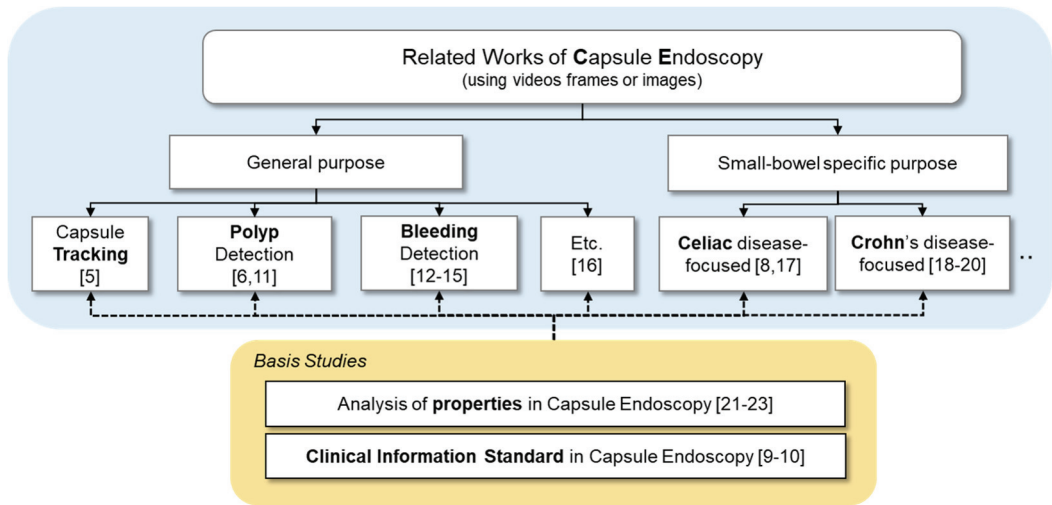


**Fig. 3.** Several studies on deep learning of capsule endoscopy.

However, CE is more medically valuable in the screening small-intestine-related diseases that are difficult to identify using general endoscopic examination. Thus, the second set of studies focused on frames of the small intestine from among the whole frames. Some of them focused on celiac disease, which can be found in the small intestine. [8] applied a DNN model to quantitatively evaluated celiac disease. [17] did not apply the deep learning technique, but studied the classification of CE image patterns to perform a comparative analysis between patients with celiac disease and normal patients. Other studies focused on Crohn's disease, which is a characteristic disease of the small intestine. Therefore, several studies conducted on evaluating and analyzing capsule endoscopic frames to obtain information associated with Crohn's disease have been conducted [18-20]. However, as the number of patients who have been examined using CE internationally is very small, collecting image data that can be utilized for machine learning is limited. Thus, some studies have analyzed the various characteristics of images that can be used for training. [21] analyzed some cases in various papers and extracted comprehensive characteristics that can be utilized as major features. [22] conducted studies to define various attributes of image datasets to detect gastrointestinal diseases through CE. In addition, [23] applied a support vector machine (SVM) or Siamese neural network method to remove redundant frames from a large number of frames and to provide summary information on capsule endoscopic frames.

However, such studies proposed a list of features that can be obtained in images, and analysis of the similarities or differences between features is insufficient. Additionally, as they do not consider standards

such as MST and CEST, which define clinical information that can be obtained from capsule endoscopes; hence, they lack a semantic relationship to clinical information. Therefore, we propose a class-labeling method to enable the systematic design of DNNs based on a lesion-focused knowledge model in which associativity or hierarchical structure of information are defined. Although not in the same field, [24] proposed a method of learning radiology images using label ontology and [25] suggested a detection model of medical texts based on an ontology describing the semantic similarity of medical words in 2019. Similarly, [26] illustrated a "Deep-Onto" network for surgical workflow and context recognition in 2019. As such, recent studies have attempted to improve learning performance by using ontology as a knowledge model. Similarly, we defined a lesion-focused knowledge model using ontology, and it is discussed in Section 3.

# 3. Lesion-Focused Knowledge Model for CE

In this section, we investigate the main lesion information based on the major anatomy of interest and findings that can be observed in capsule endoscopic images. Additionally, we propose the lesion-focused knowledge model using ontology.

## 3.1 Anatomy of Gastrointestinal Tract

Capsule endoscopy is a test to observe the condition of the GI tract. Therefore, the main concern of CE is the observation of digestive organs. Therefore, we analyze an anatomical knowledge of the GI tract in this subsection to build a knowledge model that can be used in deep learning. The details of these components are shown in Fig. 4. First, the main anatomical landmarks are divided into 4 parts (esophagus, stomach, small intestine, and large intestine) whose details are as follows:

- The *esophagus* is a tube that conveys food from the mouth to the stomach, and the endoscope capsule travels through it. The capsule often navigates through the esophageal section within a few tens of seconds.
- The *stomach* is a digestive organ in the form of a pocket in which the food coming down through the esophagus is digested. It is acidic through the secretion of gastric acid, and it can be observed in a wide area because it is the form of a pocket.
- The *small intestine* is composed of the duodenum, ileum, and jejunum, and is an organ in which various nutrients are chemically decomposed and absorbed. It is characterized by the presence of villi to absorb large amounts of nutrients.
- The *large intestine* consists largely of the cecum, colon, etc., and water is absorbed in it. Additionally, residues (such as feces) after the nutrients have been absorbed, can also be observed.

The following is the structure of the GI junction that enables the anatomical landmarks to be identified. The entire GI tract can be distinguished by three types of valves, whose detailed explanation is as follows.

- The *Z-line* is the boundary between the esophagus and the mucous membrane that constitutes the stomach, and a part that is sharply narrowed by a striated muscle in the entrance to the esophagus is observed.

- The *pyloric valve* is the membrane at the border between the stomach and small intestine, and it participates in regulating the transfer of food. Similar to the Z-line, there is are sharp and narrowing muscles in the vicinity of the pyloric valve.
- The *ileocecal valve* is a valve with a protruding structure, located on the left and right walls of the boundary between the cecum and the colon. Its main function is to limit the backflow of the contents of the colon to the ileum.
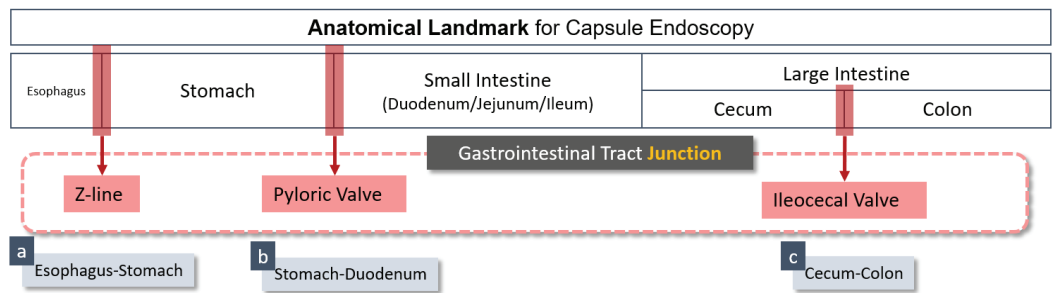


**Fig. 4.** Components of the gastrointestinal tract anatomy.

## 3.2 Findings in Capsule Endoscopic Images

Standards terminologies such as MST or CEST state some medical "findings" that can be observed from a capsule endoscopic frame [9,10]. The findings referred to here are generic features that can be identified in the image and have medical characteristics. Thus, these may be features of a disease that can be observed in the gastrointestinal tract or features that are not related to diseases but can be distinguished from normal features. As Fig. 5 shows, the main findings of CE can be classified according to the lesions on the cross section of the GI tract or the degree of elevation (shape) of the lesions.
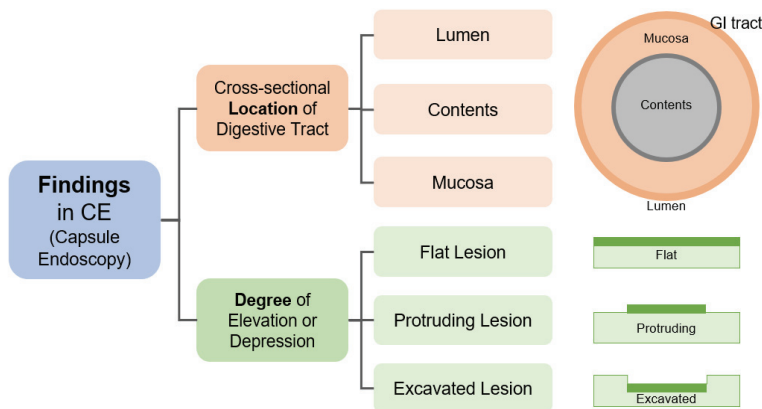


**Fig. 5.** Taxonomy of the main findings in capsule endoscopic images.

- Cross-sectional location of digestive tract: The first category is the classification of lesions from the perspective of the "cross-sectional" view in the GI tract. It has three types: mucous membrane (mucosa), internal (contents), and external (lumen). The "mucosa" category includes lesions or

findings that can occur in the mucosa. The "contents" category includes lesions or findings in the content that can be confirmed inside the GI tract. In the "lumen" category, findings that can occur as the shape of the GI tract is distorted are included.

- Degree of elevation or depression: The second category refers to the classification of lesions by their protruding degree. These are largely protruding, excavated, and flat lesions.

Here, we define the lesion-focused knowledge model by using ontology. It has two major components: gastrointestinal anatomy and findings. As Fig. 6 shows, these two elements are connected by various associations ("*canBeShownIn*", "*hasIndividual*", and "*hasSubclass*"). In particular, Fig. 6 is an extension of the "Dilated" class which is a subclass of "GIfindings" ("CrossSectionalLocation"–"Lumen").
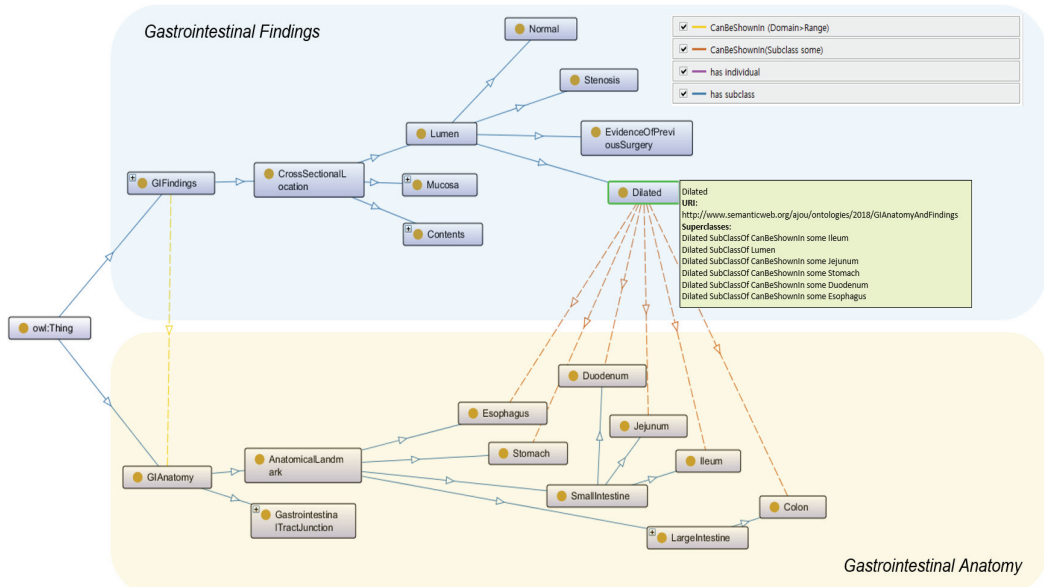


**Fig. 6.** Lesion-focused knowledge model for capsule endoscopy (by ontology).

# 4. Class-Labeling Method for Deep Learning

In this section, we propose a class-labeling method for deep learning based on the lesion-focused knowledge model defined in Section 3. First, a class label for a capsule endoscopic image is defined, and a color-based analysis is performed on a lesioned frame requiring a more detailed analysis. Based on this, a class-labeling method for a systematic design of a DNN is completed.

## 4.1 Definition of Class Label in Capsule Endoscopic Images

Here, we define and classify significant super-classes of capsule endoscopic images. There are three super-classes in a capsule endoscopic image set:

- Normal Class: Because capsule endoscopy is conducted for 12 to 14 hours, the majority of the frames are clean and without lesions. In such a scenario, the detailed class can be distinguished by

the position and properties of the endoscopic image (wrinkles, bubbles, etc.). Particularly, a detailed classification according to the position in the GI tract (esophagus, stomach, duodenum, etc.) can be defined to track the position of the capsule.

- Abnormal Class: It is composed of frames where lesions or special findings that can be observed in a capsule endoscope exist. This can be defined in detail according to the classification of the "findings" defined in Fig. 5.
- Non-discriminable Class: It corresponds to a scenario in which the image captured by the capsule endoscope is not properly photographed due to low power, transmission or reception problems, and a large amount of foam.

## 4.2 Special Case Study: Abnormal Class (Lesioned Frames)

In this subsection, we analyze the similarity of several frames with lesions based on previously analyzed GI anatomy and findings. In this study, 32 lesion cases were collected according to six categories defined in the abnormal class (three categories classified by cross-sectional location, and the other three by elevation degree). Table 1 shows the detailed types and number of collected cases. Through these cases, we applied the K-means clustering method, which is an unsupervised learning method, to the similarity analysis. The feature used in clustering was the average value of the red green blue (RGB) channel of each image, and we reduced the dimension of images through this method. As a result, two clusters in the first category (Cluster A, Cluster B) and two clusters in the second category (Cluster C, Cluster D) were derived as shown in Fig. 7. The detailed results are as follows.

**Table 1.** Specification of sample images (lesioned frames of capsule endoscopy)

| | Class | Lesion type | Count |
|---|---|---|---|
| C* | Lumen | {Stenosis, Dilation, Scars, Sign of Previous Surgery} | 6 |
| | Contents | {Red Blood, Clot, Hematin, Feces, Foreign Body} | 6 |
| | Mucosa | {Missing Villi, Hemorrhagic} | 2 |
| E* | Flat | {Angiectasias, Spot} | 7 |
| | Protruding | {Venous Lake, Nodule, Polyp} | 10 |
| | Excavated | {Ulcer} | 1 |

C*=cross-sectional location of digestive tract, E*=degree of elevation or depression.
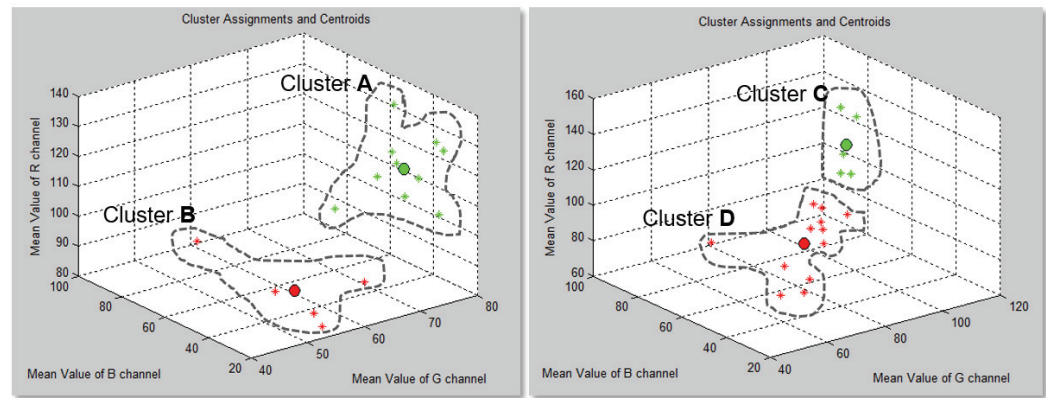


**Fig. 7.** Results of similarity analysis based on lesion-focused capsule endoscopic images (binary clustering).

- Cluster A: It is the first cluster indicating similarity in the case images of the lumen, contents, and mucosa classified on the basis of the cross section of the GI tract. Dilation and stenosis were mainly included, and it demonstrated similarity to the G channel.
- Cluster B: Cluster B indicated similarity to the R channel, with the main cases being Missing Villi, Red Blood, and Hemorrhagic.
- Cluster C: Cluster C was one of the lesioned image clusters according to the degree of lesion elevation and was mostly composed of frames such as flat-type lesions. At this time, we concluded that some images of elevated lesions such as "polyps" were incorrectly clustered.
- Cluster D: Finally, in Cluster D, most of the protruding lesions were included.

Fig. 8 shows example frames of lesion images corresponding to each cluster, and the above analysis result can be confirmed.
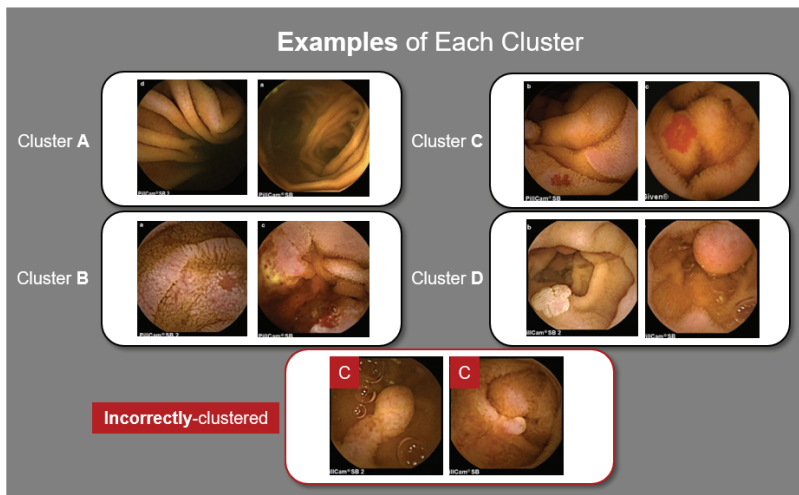


**Fig. 8.** Examples of each cluster (instance analysis of clustering).

# 5. Discussion: Application of the Class-Labeling Method

In this paper, we define and classify significant information that can be observed in capsule endoscopy images into normal, abnormal, and non-discriminable classes. In addition, a collection of frames for special cases, which have medical significance, and detailed analysis are performed for the abnormal class. Based on this, a class label as shown in Fig. 9 can be defined, and the defined class can be utilized in the output layer of the capsule endoscopic image learning model as follows.

- Labeling normal class according to various attributes of frames: The normal frame of the capsule endoscopy can be class-labeled according to the position of the capsule and the property information (wrinkle, plane, etc.) observed in the frames. We designed a CNN to classify GI landmarks (stomach, duodenum, jejunum, large intestine) as shown in Fig. 2, consisting of two convolutional layers, two fully connected layers, and an output layer. When we labeled four

classes, the accuracy was approximately 33.5%. However, when we labeled twelve classes (four GI landmarks by three types (plain, wrinkle, and bubble) in same condition, the accuracy was approximately 61.1%, an improvement of approximately 27.6%.

- Labeling abnormal class according to specificity of lesion and color similarity: Abnormal frames of capsule endoscopy were defined according to the location or type of lesion. Thus, when learning a capsule endoscopic frame, class-labeling in lesioned frame can be conducted in detail by using these hierarchies. In addition, this enables designing DNNs with consideration to various methods such as separating or subdividing class labels of learning data by applying the analyzed color-similarity results. For example, we can define "Flat-Excavated Lesion class" comprising "Spot", "Angiectasia", and "Ulcer", due to spatial-similarity in the G channel. However, this characteristic has a limitation in validation due to difficulty in collecting lesion data. Therefore, we have been gathering CE data that contain disease information.

- Eliminating indeterminate or redundant frames: Class labels for non-determinable frames can be used to distinguish the outlier frames. For example, a frame that has many bubbles or feces can be used as obstacle data in learning. To validate this characteristic, we designed a CNN consisting of ten convolutional layers, five pooling layers, three fully-connected layers, and one output layer. Its purpose was to recognize GI junctions (z-line, pyloric valve, and ileocecal valve) and learn from eleven patients' dataset (253,003 frames). Each patient had a noise image of at least 1% and up to 36%, and the elimination operation of indeterminate or redundant frames improved learning performance from 86% to 92%.
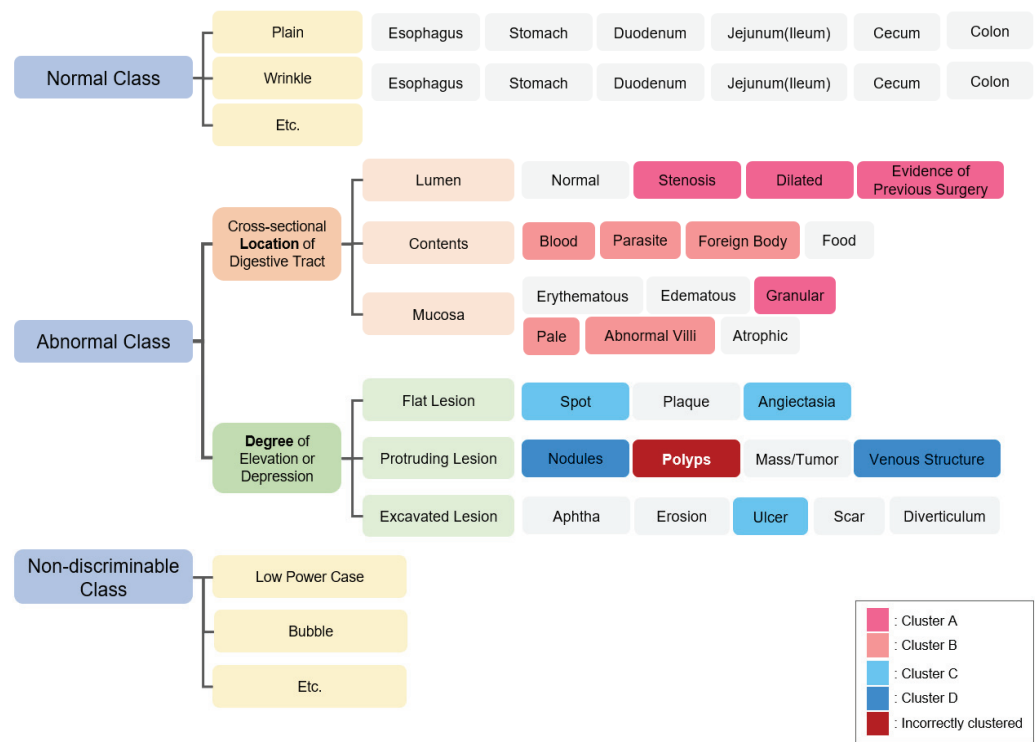


**Fig. 9.** Definition of class labels in capsule endoscopy and application of color-similarity analysis.

# 6. Conclusion

In this paper, we proposed a class-labeling method that can be used in the output layer of a DNN model by constructing a knowledge model from major lesion information defined in related standards of capsule endoscopy. First, we defined the lesion-focused knowledge model that considers the anatomy of the gastrointestinal tract and findings in capsule endoscopy. Subsequently, we collected 32 main lesion frames defined in MST and CEST and analyzed color-based similarities. Finally, we extracted the major attributes of capsule endoscopy and completed a class-labeling method. This method enables systematic learning through the application of various properties. In the future, we intend to conduct a study on a DNN model that can distinguish capsule position or specific lesions (tumor, polyp, etc.).

# Acknowledgement

# References

[1] D. Moneghini, G. Missale, L. Minelli, and R. Cestari, "P.18.2 upper and lower gastrointestinal lesions overlooked at conventional endoscopy and further diagnosed with small bowel capsule endoscopy: the crucial role of endoscopic experience in patients with obscure gastrointestinal bleeding," *Digestive and Liver Disease*, vol. 48(Suppl. 2), p. e216, 2016.

[2] A. Bar-Gil Shitrit, B. Koslowsky, D. M. Livovsky, D. Shitrit, K. Paz, T. Adar, S. N. Adler, and E. Goldin, "A prospective study of fecal calprotectin and lactoferrin as predictors of small bowel Crohn's disease in patients undergoing capsule endoscopy," *Scandinavian Journal of Gastroenterology*, vol. 52, no. 3, pp. 328-333, 2017.

[3] M. Fischer, S. Siva, J. M. Wo, and H. M. Fadda, "Assessment of small intestinal transit times in ulcerative colitis and Crohn's disease patients with different disease activity using video capsule endoscopy," *AAPS Pharmscitech*, vol. 18, no. 2, pp. 404-409, 2017.

[4] M. Lujan-Sanchis, E. Perez-Cuadrado-Robles, J. Garcia-Lledo, J, F. J. Fernandez, L. Elli, V. A. Jimenez-Garcia, et al., "Role of capsule endoscopy in suspected celiac disease: a European multi-centre study," *World Journal of Gastroenterology*, vol. 23, no. 4, pp. 703-711, 2017.

[5] H. Chen, X. Wu, G. Tao, and Q. Peng, "Automatic content understanding with cascaded spatial–temporal deep framework for capsule endoscopy videos," *Neurocomputing*, vol. 229, pp. 77-87, 2017

[6] Y. Yuan and M. Q. H. Meng, "Deep learning for polyp recognition in wireless capsule endoscopy images," *Medical Physics*, vol. 44, no. 4, pp. 1379-1389, 2017.

[7] P. Li, Z. Li, F. Gao, L. Wan, and J. Yu, "Convolutional neural networks for intestinal hemorrhage detection in wireless capsule endoscopy images," in *Proceedings of 2017 IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, China, 2017, pp. 1518-1523.

[8] T. Zhou, G. Han, B. N. Li, Z. Lin, E. J. Ciaccio, P. H. Green, and J. Qin, "Quantitative analysis of patients with celiac disease by video capsule endoscopy: a deep learning method," *Computers in Biology and Medicine*, vol. 85, pp. 1-6, 2017.

[9]  L. Aabakken, B. Rembacken, O. LeMoine, K. Kuznetsov, J. F. Rey, T. Rosch, G. Eisen, P. Cotton, and M. Fujino, "Minimal standard terminology for gastrointestinal endoscopy–MST 3.0," *Endoscopy*, vol. 41, no. 8, pp. 727-728, 2009.

[10] L. Y. Korman, M. Delvaux, G. Gay, F. Hagenmuller, M. Keuchel, S. Friedman, et al., "Capsule endoscopy structured terminology (CEST): proposal of a standardized and structured terminology for reporting capsule endoscopy procedures," *Endoscopy*, vol. 37, no. 10, pp. 951-959, 2005.

[11] B. Taha, N. Werghi, and J. Dias, "Automatic polyp detection in endoscopy videos: a survey," in *Proceedings of 2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, Innsbruck, Austria, 2017, pp. 233-240.

[12] X. Jia and M. Q. H. Meng, "A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images," in *Proceedings of 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, 2016, pp. 639-642.

[13] X. Jia and M. Q. H. Meng, "A study on automated segmentation of blood regions in wireless capsule endoscopy images using fully convolutional networks," in *Proceedings of 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI)*, Melbourne, Australia, 2017, pp. 179-182.

[14] X. Jia and M. Q. H. Meng, "Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and CNN features," in *Proceedings of 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, Korea, 2017, pp. 3154-3157.

[15] X. Li, H. Zhang, X. Zhang, H. Liu, and G. Xie, "Exploring transfer learning for gastrointestinal bleeding detection on small-size imbalanced endoscopy images," in *Proceedings of 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, Korea, 2017, pp. 1994-1997.

[16] X. Wu, H. Chen, T. Gan, J. Chen, C. W. Ngo, and Q. Peng, "Automatic hookworm detection in wireless capsule endoscopy images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 7, pp. 1741-1752, 2016.

[17] E. J. Ciaccio, C. A. Tennyson, G. Bhagat, S. K. Lewis, and P. H. Green, "Classification of videocapsule endoscopy image patterns: comparative analysis between patients with celiac disease and normal individuals," *Biomedical Engineering Online*, vol. 9, article no. 44, 2010.

[18] S. Bejakovic, R. Kumar, T. Dassopoulos, G. Mullin, and G. Hager, "Analysis of Crohn's disease lesions in capsule endoscopy images," in *Proceedings of 2009 IEEE International Conference on Robotics and Automation*, Kobe, Japan, 2009, pp. 2793-2798.

[19] R. Kumar, Q. Zhao, S. Seshamani, G. Mullin, G. Hager, and T. Dassopoulos, "Assessment of Crohn's disease lesions in wireless capsule endoscopy images," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 355-362, 2011.

[20] E. Gal, A. Geller, G. Fraser, Z. Levi, and Y. Niv, "Assessment and validation of the new capsule endoscopy Crohn's disease activity index (CECDAI)," *Digestive Diseases and Sciences*, vol. 53, no. 7, pp. 1933-1937, 2008.

[21] S. Segui, M. Drozdzal, G. Pascual, P. Radeva, C. Malagelada, F. Azpiroz, and J. Vitria, "Generic feature learning for wireless capsule endoscopy analysis," *Computers in Biology and Medicine*, vol. 79, pp. 163-172, 2016.

[22] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, et al., "Kvasir: a multi-class image dataset for computer aided gastrointestinal disease detection," in *Proceedings of the 8th ACM on Multimedia Systems Conference*, Taipei, Taiwan, 2017, pp. 164-169.

[23] J. Chen, Y. Zou, and Y. Wang, "Wireless capsule endoscopy video summarization: a learning approach based on Siamese neural network and support vector machine," in *Proceedings of 2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016, pp. 1303-1308.

[24] K. Yan, Y. Peng, V. Sandfort, M. Bagheri, Z. Lu, and R. M. Summers, "Holistic and comprehensive annotation of clinically significant findings on diverse CT images: learning from radiology reports and label ontology," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, 2019, pp. 8523-8532.

[25] T. Zheng, Y. Gao, F. Wang, C. Fan, X. Fu, M. Li, Y. Zhang, S. Zhang, and H. Ma, "Detection of medical text semantic similarity based on convolutional neural network," *BMC Medical Informatics and Decision Making*, vol. 19, article no.156, 2019.

[26] H. Nakawala, R. Bianchi, L. E. Pescatori, O. De Cobelli, G. Ferrigno, and E. De Momi, ""Deep-Onto" network for surgical workflow and context recognition," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 4, pp. 685-696, 2019.

**Ye-Seul Park**  https://orcid.org/0000-0003-2584-7489

She is in course of Ph.D. in Department of Electrical and Computer Engineering at Ajou University, Suwon, Korea. She received B.S. and M.S. degrees in Electrical and Computer Engineering from Ajou University in 2015 and 2017, respectively. Her research interests include data modeling for medical images, deep learning (machine learning) of medical images, clinical big data-based medical software validation, ontology and embedded software.

**Jung-Won Lee**  https://orcid.org/0000-0001-8922-063X

She is a professor of the Department of Electrical and Computer Engineering at Ajou University, Korea, since 2006. For 2003–2006, she was the BK professor in Ewha Womans University. She received Ph.D. degree in Computer Science Engineering from Ewha Womans University, Korea in 2003. She was a researcher of LG Electronics and did an internship in the IBM Almaden Research Center, USA. Her research interests include mobile context awareness, ontology, medical data modeling and embedded and software.