Study on Anomaly Image Detection for Capsule Endoscopy Diagnosis Support

March ２０１６

Erzhong Hu

Study on Anomaly Image Detection for Capsule Endoscopy Diagnosis Support

Graduate School of Systems and Information Engineering

University of Tsukuba

March ２０１６

Erzhong Hu

Contents

List of tables

List of figures

Chapter 1

Introduction

Nowadays, as gastrointestinal diseases and gastrointestinal cancers have become one of the most serious threats to worldwide human health, advanced examinational means are regarded very significant and indispensable in clinical diagnosis. Among a variety of examinational means, capsule endoscopy, a pill-shaped endoscopy technology, has attracted great attention and become broadly utilized during the past decade for the reason that it can provide visual and non-invasive examinations of the whole small intestine. Such kind of examination was almost impossible before.

However, capsule endoscopy has its own problem. Unlike the real-time examination of gastroscopy and colonoscopy, capsule endoscopy has a journey “alone” throughout the gastrointestinal tract, and captures video data including a large number of images. And thus, it then absorbs massive time for doctors to observe and make diagnosis for all the images. With regard to this issue, a variety of diagnosis support technologies covering different functions have been developed. Among them, anomaly detection is regarded as one of the most promising means

to remarkably improve the efficiency and quality of diagnosis.

Anomaly detection is desired to detect all anomalies including lesions and foreign matters from capsule endoscopy images. However, according to the investigation, in previous studies on anomaly detection for capsule endoscopy images, many approaches were put forward to detect only several specified kinds of lesions. It was always necessary to establish a classifier in respond to the specified target anomaly in advance. Although it is possible to employ a large number of training data involving different conditions, anyhow, such kind of approaches are available but limited only within the categories of specified lesions. Except for lesions, foreign matters, such as residue and bile, should also be filtered out because they are somewhat considered as the shelter of lesions in some conditions. So far, the detection of foreign matters has not been mentioned in most of the approaches. And thus, providing feasible anomaly detection based support for practical diagnosis of capsule endoscopy remains quite a difficult task.

By observation, it is found that normal images, containing only intestinal wall and lumen, usually comes out in relatively simple appearance, as compared with a variety of anomalies. Therefore, it is considered as a breakthrough to concentrate on the normal images instead of the various anomalies. If only the high similarity among normal images can be given a proper description, anomalies different from normal images are expected to be discovered accordingly. Meanwhile, unconcerned foreign matters such as residues could be detected in this way. Nevertheless, oversighted anomalies may occur if such kind of anomaly detection is adopted to various anomalies. This is for that some of the anomaly images, especially those with small or inconspicuous anomaly regions, probably possess quite high similarity to the normal images. In respond to this case, it is also necessary to take additional measures to decrease the oversights. To meet all the requirements above, a single-stage scheme is considered insufficient to settle all of the tasks.

With regard to these issues, therefore, this dissertation proposes a two-stage scheme involving anomaly degree evaluation and lesion oversight prevention, directing at the detection for all kinds of lesions and foreign matters, and maximally reduce the lesion oversights. First, in the stage of anomaly degree evaluation, various lesions and foreign matters are expected to be screened out by evaluating the anomaly degree of image data, which indicates the scale of difference distinct from the confirmed normal image data. To figure out the characteristics of various contents, in feature extraction, the proposed method utilizes Higher-order Local

Auto Correlation (HLAC), which is regarded as a qualified instrument and owning well suitability to address feature extraction for capsule endoscopy images. Furthermore, this dissertation also present techniques involving color space conversion and adjustment for feature extraction, aiming at facilitate the efficacy of anomaly degree evaluation for capsule endoscopy images. Second, in the following stage of lesion oversight prevention, a more specific detection method is developed to cope with images with small or inconspicuous lesion regions that are easily oversighted in the first stage. A geometric image descriptor, called LCE-HLAC features, is developed to make contribute on reinforcing the descriptive capacity for local details and realize more effective feature extraction for lesions of various appearance. To sum up, according to the unique design of each stage, the proposed scheme is considered available to deal with various anomalies without restriction of anomaly's categories, and having potential to maximally restrain oversighted lesions at the same time.

The remainder of this paper is organized as follows. The relevant materials of the background of research, starting from the brief introduction of gastrointestinal diseases and diagnosis, will be introduced in Chapter 2. Chapter 3 depicts the configuration and design of the proposed anomaly system. Chapter 4 and 5 describe the two stages of the proposed anomaly system in detail, and the experimental materials and discussions for each part. Following in Chapter 6, this dissertation also introduces the feasibility validation of the whole proposed anomaly detection system, by establishing general training datasets for the two stages of the proposed anomaly detection system. Chapter 7 follows at the last and gives the conclusions of the dissertation according to all the assumption and experimental validation results denoted in the foregoing chapters. After all, anomaly detection based diagnosis support is convinced to be very prospective and promotive for the practical application and future development of capsule endoscopy technology.

Chapter 2

Background

2.1 Gastrointestinal diseases and diagnosis

Gastrointestinal diseases, and gastrointestinal cancers in particular, are now among the most serious threats to human health. In 2012, gastrointestinal cancer, including both colorectal and stomach cancer, was reported as being the most frequently diagnosed form of cancer worldwide [1]. Domestic statistics by Japanese National Cancer Center [2] figured that, by 2015, incidence of colorectal cancer is expected to be the first place, while the stomach cancer will be the third place. In the same report, deaths from gastrointestinal cancer will increase beyond lung cancer and become the most threatening among all kinds of cancers.

Although cancer is deemed as severely dreadful disease because it is usually related with the length of life, most gastrointestinal cancer and other kinds of gastrointestinal diseases can be cured if appropriate diagnosis and treatment can be introduced in time. Another report was issued by Japanese Association of Clinical Cancer Centers [3] shown in Fig. 2.1.1. There is a comparison of five-year cancer

survival rates among the colorectal cancer, stomach cancer, and the lung cancer. Commonly, all of these three kinds of cancers have their staging defined from I to IV, that describe the severity of the cancer at the time of the diagnosis, taking into account the growth and size of the tumor and whether it has spread to adjacent organs [4]. Within Fig.2.1.1, the broken lines reveal that the survival rates of colorectal cancer and stomach cancer is obviously higher than lung cancer until stage III. More than 68% patients diagnosed before stage III have relatively good prognosis. However, for three kinds of cancers, the five-year survival rates decrease sharply below 15% if the progress comes to the stage IV. Thus, clearly, it is known that the intervention of early diagnosis is very important for gastrointestinal cancer and other kinds of gastrointestinal diseases.

In terms of examinational means, blood tests, radiography, ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) [5] and endoscopies are often utilized as important examination means [6]. Actually, in clinical application, a series of diagnostic procedures (as illustrated in Fig.2.1.2) are usually conducted in sequence before a definite diagnosis can be made. First, the laboratory tests, including blood test and stool test, are processed to check if there any abnormal sign (bacteria in the blood, hidden blood in the stool, etc.) in patient's blood and stool. The laboratory tests are effective to confirm the bacteria infection and the gastrointestinal bleeding. Second, the patient should accept imaging tests, such as radiography test upon the barium beefsteak meal (in which patient eats a meal containing liquid barium used to coat the inside of organs so that they will show up on an X-ray), ultrasound, and CT/MRI scans. Imaging tests can help to find the morphological abnormalities in gastrointestinal organs. For example, radiography test upon the barium beefsteak meal is powerful for many kinds of enteritis [7]. Endoscopic examinations follow after the tests above. Examination with gastroscopy or colonoscopy allows the doctor to view the inner wall of gastrointestinal tract. It can often help identify tumors, polyps, inflamed mucosa, ulcers and bleeding. Since such lesions can hardly observed by other technologies, endoscopic examinations are now playing significant and replaceable role in digestive diagnosis.

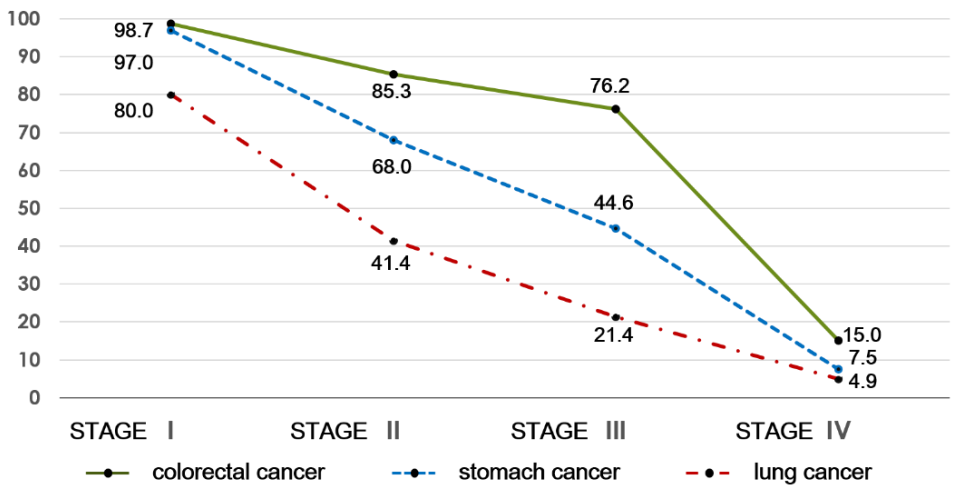


Figure 2.1.1. Five-year cancer survival rates (%), issued by Japanese Association of Clinical Cancer Centers in 2004.

Figure 2.1.2. Digestive diagnostic procedures.

2.2 Capsule endoscopy

Among endoscopic technologies, gastroscopy and colonoscopy are broadly and frequently used. During the examination by either gastroscopy or colonoscopy, it involve inserting a long, flexible lightening tube into the gastrointestinal tract. However, because small intestine is the longest and winding organ in the right middle section of gastrointestinal tract, it is hard for the endoscopic technologies above to reach deep into the small intestine and implement the examination overall. As the reference goes, the normal adult small intestine is about 400 cm in length, and consists of the duodenum (25-30 cm), and the jejunum (160-200 cm), and the rest part of ileum [8]. In contrast, gastroscopy and colonoscopy are generally not longer than 150cm. Therefore, for both of the endoscopic technologies, it is almost impossible to inspect the full length of small intestine [9]. As a solution to address this issue, capsule endoscopy, a novel endoscopy technology, has been developed.

Capsule endoscopy (CE) is a pill-shaped endoscopy technology (Fig.2.2.1) which was firstly invented in 2000 [10] and originally marketed in 2001. Since capsule endoscopy acquired its first approval by FDA (U.S. Food and Drug

Administration) in 2001, more than 2 million capsule endoscopies have been ingested worldwide [11]. In Japan, capsule endoscopy examination was firstly approved in 2007 [12], and so far, it has been introduced into more than 300 domestic medical facilities [13].

Within the capsule sell, there are miniaturized parts including optical lens, CMOS image sensors, LED illumination, battery, UHF-band transmitter and antenna. When works along the gastrointestinal tract, it is able to capture images at 2 to 6 frames per second [14], for about 8 hours, and simultaneously transmit the images outside to receiver set on the patient's body (Fig.2.2.2). And then, the images can be saved temporarily in the data recorder. Along with the movement of capsule endoscopy, the location information in gastrointestinal tract is also transmitted and recorded together with the image data, so that potential lesions could be located for further treatment.

Figure 2.2.1. Capsule endoscopy (PillCam® by Given Imaging Ltd.)

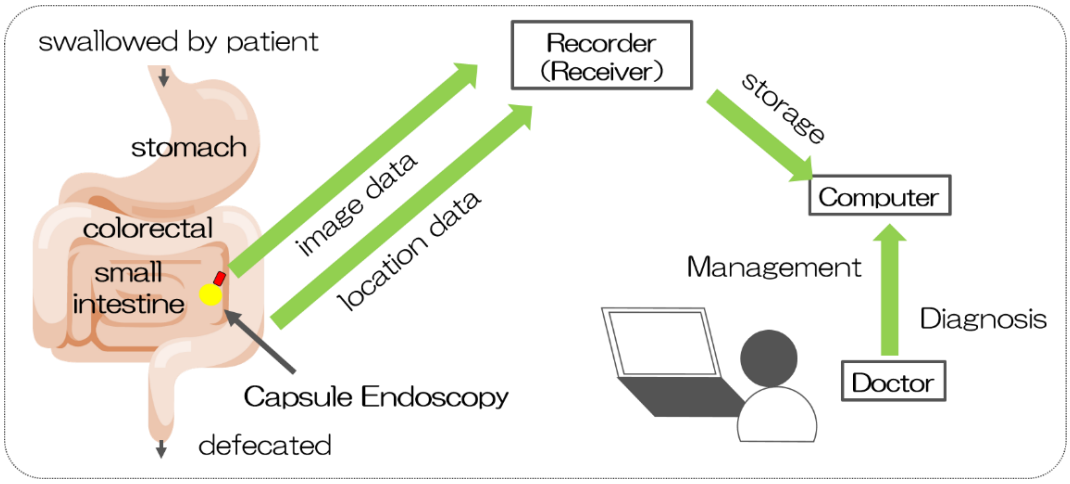


Figure 2.2.2. Diagnostic procedure of capsule endoscopy.

Due to the unique profile in size and flexibility, capsule endoscopy can go deep in through the whole gastrointestinal tract and have many. Specifically, capsule endoscopy also has a variety of advantages that could be summarized as below.

Good for patients:

(1) Due to the examination conducted by making patients simply swallow the capsule endoscopy, less physiological and mental burden are placed on patients compared to conventional endoscopy technologies.

(2) There is no need to stay in the medical facility, so that patients can go back to work or go home during the examination.

(3) As an optical method, there is no risk and worry of the radiation exposure during the examination.

Good for doctors:

(1) With capsule endoscopy, lesions difficult to be arrived by conventional endoscopy technologies can be fully captured and observed. This lead to advantages in terms of identifying the causation of bleeding, detecting polyps, inflammatory bowel disease, ulcers, and tumors of the small intestine. As reports said, capsule endoscopy has also been demonstrated particularly effective against some certain diseases, such as unexplained gastrointestinal bleeding and Crohn's disease [15, 16, 17].

(2) Examination of small intestine can be carried out completely by capsule endoscopy without doctor’s option. Thus, the experience and skill of doctor makes no difference to the examination results.

(3) Unlike the real-time examination by conventional endoscopies, all the video data captured by capsule endoscopy are saved, so that it is convenient for doctors to confirm the suspicions repeatedly if necessary.

As stated above, it is clear that capsule endoscopy is a patient-friendly tool that is not only capable to facilitate the examination of the whole small intestine, but also makes painless gastrointestinal examination possible. Actually, as to clinical application, capsule endoscopy is usually utilized when intestinal diseases in small intestine are suspected after the blood and stool tests, imaging tests and conventional endoscopy examinations.

2.3 Diagnosis support technologies for CE

A full-length CE video usually lasts for approximately 8 hours and contains in excess of 50,000 consecutive image frames [18], the examples of which are given by Fig. 2.3.1. On average, a doctor requires between 45 to 120 min to inspect and screen all the image data [19]. Actually, the diagnosis time is largely depending on both the doctor's skill and experience and also on image conditions. Some factors, such as the basic tone of mucosa, illumination conditions, and the amount of residue (of food or drug remaining in patient's gastrointestinal tract), all causes great impact on the diagnosis efficiency. Moreover, in the case if any part in the small intestinal are suspected to be anomalous, examination time tends to become longer, and then further adds workload on doctors [20].

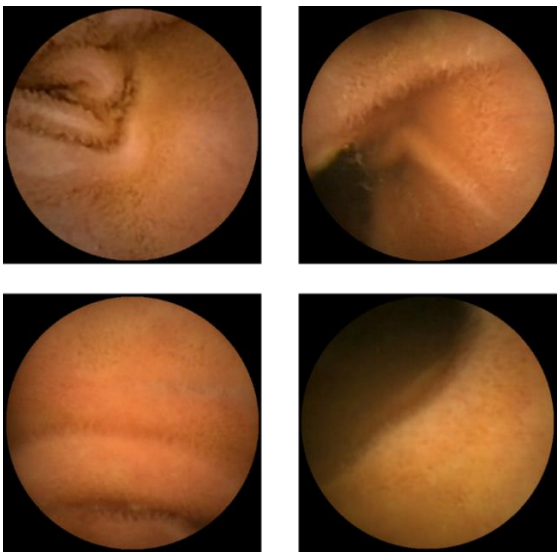


Figure 2.3.1. Examples of capsule endoscopy image (normal).

In order to release the doctor's workload and contribute to more exact diagnosis, many researches on computer-aided diagnosis (CAD), involving various kinds of automatic support technologies for capsule endoscopy diagnosis, have been activated in recent years. Among the automatic support technologies, bundled software developed by capsule endoscopy manufacturers have been firstly put into use to assist clinical diagnosis. For example, for the case of "RAPID®" [21], developed by Given Imaging Ltd. (the Israel company pioneered the capsule endoscopy technology), several functions are integrated in the software. For example, "Flexible spectral Imaging Color Enhancement" provides options of 3

types of spectral images, which make the view of elevated mucosa more distinct and visible. The “QuickView” mode allows doctor adjust the sampling rate of video, so that to reduce the reading time on request [22]. However, some reports have claimed that the accuracy levels of diagnosis are still restricted and unsatisfactory, and oversights may still occur due to lacking of efficient assistance [23]. Meanwhile, the “QuickView” mode are not regarded sufficiently benefiting the diagnosis [22]. Besides, some other pre-load softwares, such as “Endocapsule® 10 System” by Olympus (a Japanese company) and “MiroCam®” by IntroMedic (a Korean company), also provide similar functions mainly restricted within different view modes by measuring the scale of change between adjacent frames [24, 25].

Except for the bundled software developed by capsule endoscopy manufacturers, many other researches have also been inspired to provide various kinds of diagnostic support. Some of them focused on reducing diagnosis time by optimizing the control of video display. For example, Vu Hai et al developed an adaptive control of video display, on the basis of analysis of similarity of color feature and extraction of motion displacement between adjacent frames [26, 27]. Based on the developed techniques, the performances of this approach has also been validated by evaluating the required diagnostic time. As the result, it is indicated that the technique is of valuable assistance to medical doctors [8, 28]. Efforts put forward by Min Kook Choi et al. computed the meaningful motion changes of capsule endoscopy frames, and managed to remove the duplicated video frames [29]. Except of directly controlling video display or amount of frame data, several other studies have been carried on by suggesting significant information to help diagnosis. Bashar M. K. et al used local color histogram and Gauss Laguerre Transform (GLT) to identify non-bubbled frames and bubbled frames, respectively [30]. Bubbles are usually emerging and causing interference to diagnosis. By this means, informative frames without bubble in capsule endoscopy image were expected to be distinguished. In another work of Vu Hai et al, detection of small intestine contractions is achieved based upon the characteristics of geometric patterns [31]. Moreover, approach on detecting the landmark of adjacent organs by introducing a color model for gastrointestinal tract has been presented in [32].

These diagnosis support technologies, as denoted above, are all proposed to decrease doctor’s workload and improve the efficiency of diagnosis. Nevertheless, none of these technologies have direct concern to the possible anomalies in capsule endoscopy images. It means that, thus, these diagnosis support technologies are

regarded helpful, but too limited to take more important charge and share doctor's workload.

2.4 Anomaly detection for CE images

2.4.1 Anomaly detection

Except for the technologies depicted above, anomaly detection technologies are particularly concentrated among the diagnosis support technologies. Like the other technologies, anomaly detection aims at releasing doctor's workload and contribute to efficient diagnosis. Moreover, anomaly detection is considered quite meaningful and able to help doctors identify the lesions so that oversights could be reduced. Including but not limited in capsule endoscopy image domain, anomaly detection attracts attention of researchers in a variety of medical imaging domains, such as CT image [33, 34, 35], positron emission tomography (PET) image [36], mammograms image [37], ultrasound image [38, 39, 40], and histopathological image [41, 42].

Generally, there are two main ways of computer-aided diagnosis to provide diagnosis with assistant, screening and double check. Screening has an effect to automatically distinguish between the normal (negative) and anomaly (positive) before the doctor makes diagnosis [43]. In the case that suspicions of anomaly are sensed by the anomaly detection system, the suspected part of image data can be preferentially presented for doctor's observation. By this way, diagnosis time is able to be saved considerably. Another way of computer-aided diagnosis is double check, or named double reading. Double check aims to confirm if there are anomaly (false negative) existing in the data group which has been judged as normal by doctor [44]. Since diagnosis time is always limited in practical situation, oversights of anomaly can hardly be avoided completely. Thus, automatic double check has become desirable to decrease the oversights and improve the precision of diagnosis.

Since it has been stated in the section 2.3 that the full-length capsule endoscopy video usually contains many consecutive frames in excess of 50000, thus, it is reasonable that the diagnosis of capsule endoscopy image can also benefit from anomaly detection technique of suspected lesion frames. As illustrated in the paragraph above, in practice, the utility of the anomaly detection for capsule endoscopy image can make use of both of screening and double check (Fig.2.4.1.1).

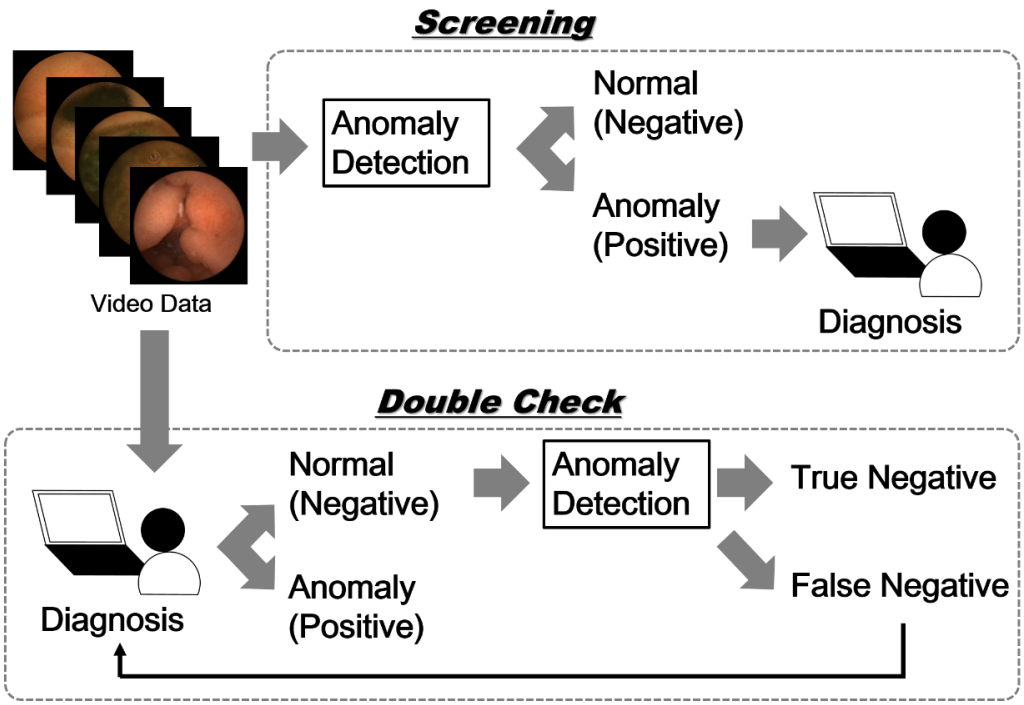


Figure 2.4.1.1. Prospective utility of anomaly detection for capsule endoscopy.

Certainly, a reliable anomaly detection approach, which plays the most significant role in both of the procedures, is required as a preceding condition.

2.4.2 State-of-the-art studies

A large number of state-of-the-art methods for anomaly detection have been presented in recent years. Roughly, these methods can be divided into two broad categories, namely predefined-condition-based anomaly detection and geometric-feature-based anomaly detection.

As previously reported [20, 45, 46, 47], predefined-condition-based anomaly detection methods implement anomaly detection via pixel matching, for either single or multiple color components within the RGB, HSV, or CIE L*a*b* color spaces according to predefined conditions for specified lesions. One example [20] is ruled like this: in RGB color space, if R channel is between 170 and 200, and larger than G channel (G>150), meanwhile R-G is smaller than 40 with B larger than 100, the inspected pixel is regarded as part of mucosal ulcer. Such kind of anomaly detection adopting predefined-condition-based anomaly detection methods seems simple to be realized and understand. However, because pixel values are directly used to descript the attendance or absence of lesions, and

meanwhile, the lesion appearances vary considerably and are mutable, it is almost impossible to formulate a generalized definition even for all cases of one kind of given anomaly. Therefore, in the case that actual conditions mismatched the predefined definitions, lesions can be oversighted, so the accuracy levels of normal and anomaly discriminations cannot be guaranteed.

In contrast, geometric-feature-based anomaly detection methods appear to be a more promising direction. Such methods generally adopt specialized geometric feature extraction that utilizes color, texture, and shape features [48], and involve statistical classification techniques that can make identification for newly added normal or anomaly samples without matching with predefined conditions. For example, Li et al. [49, 50] adopted the chromaticity moment for the extraction of color features, and neural network classifier to realize the recognition. Giritharan et al. used dominant color descriptor for bleeding detection [51]. Barbosa et al. [52] proposed the extraction of texture features using a co-occurrence matrix combined with discrete wavelet transform, and also used a multilayer perceptron neural network classifier to conduct the statistical analysis. Similarly, Yeh et al. computed color coherence vector (CCV) for texture feature extraction [53]. Another popular texture descriptor is the local binary pattern (LBP) feature [54]. On the basis of the original LBP feature, the rotation-invariant uniform type of LBP (RIU-LBP) [55] was proposed and demonstrated effective in the capsule endoscopy image domain. Ulcer and tumor detection utilizing RIU-LBP feature, combined with SCM-based classification, has been reported in several issues [56, 57]. Compared to color wavelet covariance [58], LBP has been shown to be more effective for tumor detection. Besides, shape features have also been considered for anomalies that are known to exhibit distinctive figurations (such as polyps). Moreover, segmentation methods [59] and edge detection methods [60] have been applied.

Apart from these specific feature-based approaches, some studies have employed more complicated combinations of feature extraction techniques to obtain the characteristics of various lesions. For example, one study [61] employed MPEG-7 visual descriptors, including a variety of color, texture, and shape descriptors, for detecting bleeding, ulcers, and polyps. Besides, in the bleeding and ulcer detection method proposed by Szczypinski et al. [62], anomaly detection was executed using software [63] that calculated various statistics, such as image histograms, gradient magnitude, and grey-level co-occurrence matrices. A brief summary of the geometric features used in the approaches above is denoted as in Table 2.4.2.1.

Color	Chromaticity moment Dominant color descriptor
Texture	Co-occurrence matrix Color coherence vector Local binary pattern Color wavelet covariance
Shape	Segmentation Edge detection
Combination	MPEG-7 visual descriptors, Mazda software package

Table 2.4.2.1. Main geometric features used in CE anomaly detection.

Chapter 3

Proposed Anomaly Detection System

3.1 Issues and challenges

On one side, it has become clear that capsule endoscopy passes throughout the gastrointestinal tract and possesses advantages in terms of identifying the causation of bleeding, detecting polyps, inflammatory bowel disease, mucosal ulcers, and tumors of the small intestine. However, on the other side, since the movement of capsule endoscopy is completely guided by intestine peristalsis, unlike the operable conventional endoscopies, the whole journey of capture acts on capsule endoscopy's own attitude angle. And thus, the view of objects revealed in images are totally unconstrained and uncontrollable. With this in mind, it can be easily supposed that a variety of factors make it hard to draw a judgement by automatic anomaly detection.

First, because that the view of image differs greatly in variable lighting conditions, individual differences, seriousness of lesion and the capturing locations, in different cases, handling the lesion image or the lesion regions become quite hard

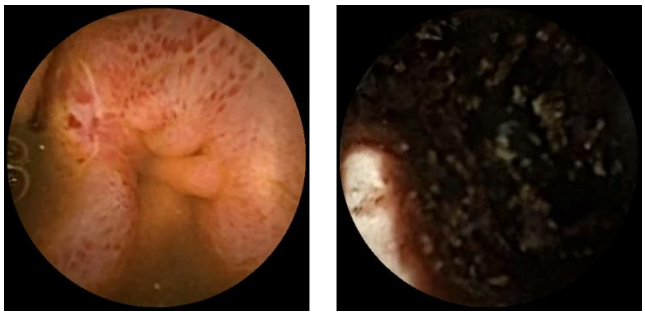


Figure 3.1.1. Various bleeding view in different situations.

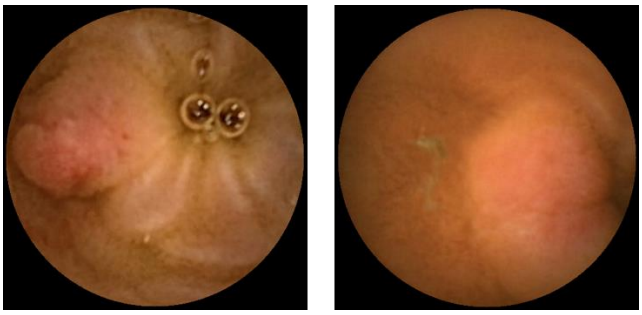


Figure 3.1.2. Gradual change between lesion region and surrounding normal region.

task. Two sets of images, Fig.3.1.1 and Fig.3.1.2, give examples for these kinds of problems. Fig.3.1.1 shows two examples of bleeding images. From this set of images, it is not difficult to find that though both images can be sorted as bleeding, the views are totally different. The left case, seems bright and clear, shows many small bleeding spots. Comparatively, in the right case, the intestinal tract has been filled with flowing dark blood. The status of mucosa of intestinal wall become almost impossible to be observed. In Fig.3.1.2, two examples of tumor image illustrate that it is hard to appropriately cut the lesion region out of the whole image, it is for that lesions do not always own clear contour, and the boundary between the lesion and the normal intestinal wall may involve very blurry and gradual change.

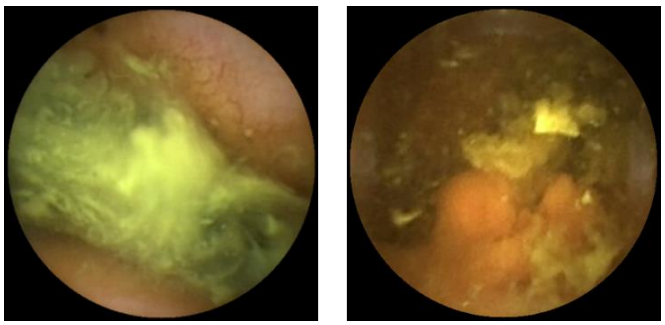
Respecting to the previous anomaly detection approaches, as referred in 2.4.2, the three types of features, color, texture, and shape, are generally regarded as tools to provide best descriptions for the detection of anomalies. However, the requirements of features vary according to the kind of lesion to be examined. For example, in order to detect tumor regions that usually have a reddish appearance with basically round contours, it is essential to employ suitable description schemes for both color and shape features. However, as already noted, most of the techniques for feature extraction have focused on only one of the three feature types. For

example, LBP and its improved form RIU-LBP are considered suitable for texture features, but it has no contribution to the description of the lesion's color and shape. To realize descriptions for more than one type of feature, it is often necessary to use a complicated combination of techniques for multiple feature extraction [61, 62].

Second, capsule endoscopy images usually contain various kinds of contents not limited within normal intestinal lumen region and lesion region. Besides, in particular, foreign matters such as residue and bile (Fig. 3.1.3), usually makes images hard to understand. In clinical diagnosis, actually, residue sometimes appears as the shelter of lesions in the same image simultaneously. For example, in Fig. 3.1.3 (b), there are polyps hidden behind some residue, but it becomes much difficult to observe the polyps since the interferential residue usually makes the whole image confusing. Therefore, according to doctor's opinion, it is necessary to inspect all of these suspicious images to insure no oversight happens. Summing up all the issues indicated above, the opinion can be drawn that there are two significant problems demanding prompt solution.

(1) For application to different kinds of lesions, previous researches focused on approaches aiming at specific lesions. Even though complicated combination of techniques for multiple feature extraction are adopted, it is considered incapable to cope with all kinds of lesions.

(2) In previous approaches, suspicious images including foreign matters, such as residue and bile, are not handled appropriately to be part of the anomaly detection. With regard to these issues, structuring a feasible anomaly detection system for practical diagnosis support remains a difficult task.



(a) (b)

Figure 3.1.3. Example of residues.

(a) Residues, (b) Residues with lesion hidden behind.

3.2 Objectives and motivation

Corresponding to these issues, this dissertation takes consideration of the request for a qualified anomaly detection system.

(1) All the lesions should be detected without establishing a complicated combination of feature extraction techniques for multiple lesion categories.

(2) Except for lesions, foreign matters such as residue, should also be considered appropriately as one of the targets of anomaly detection, so that oversighted lesions caused by foreign matters can be decreased in development.

Moreover, based upon the essential requirements above, an ideal anomaly detection method is also expected to possess excellent generalization ability, which is deemed indispensable to deal with various image data in variable lighting conditions or having different seriousness of lesion in different cases.

To meet all the requirements above, it is needed to discuss about the fundamental configuration of the anomaly detection scheme. Fig 3.2.1 gives a flowchart of the geometric-feature-based anomaly detection denoted in Chapter 2. As the principle of anomaly detection, it is always necessary to establish a classifier in respond to the target anomaly which has been specified in advance. For instance, bleeding detection needs a classifier trained by bleeding image data, while ulcer detection needs to use ulcer image data to train a particular ulcer classifier. Although it is possible to employ a large number of training data involving different conditions, such configuration, anyhow, is considered available but limited within only the category of specified lesion. To realize an anomaly detection system that can be adopted to various lesions and not limited to category of lesion, it is believed that the confine of establishment of the classifier for specified lesion should be regulated.

By observation, compared with a variety of anomalies, it is found that normal images, containing only intestinal wall and lumen, usually comes out in relatively simple appearance. Therefore, it is considered as a breakthrough to concentrate on the normal images instead of the various anomalies. If only the high similarity among normal images have a proper description, anomalies different from normal images are expected to be discovered accordingly. Meanwhile, unconcerned foreign matters such as residues could be detected in this way.

Nevertheless, a number of oversighted anomalies may occur if such kind of anomaly system is adopted to various anomalies. This is for that some of anomaly images, especially those with small or inconspicuous anomaly regions, probably

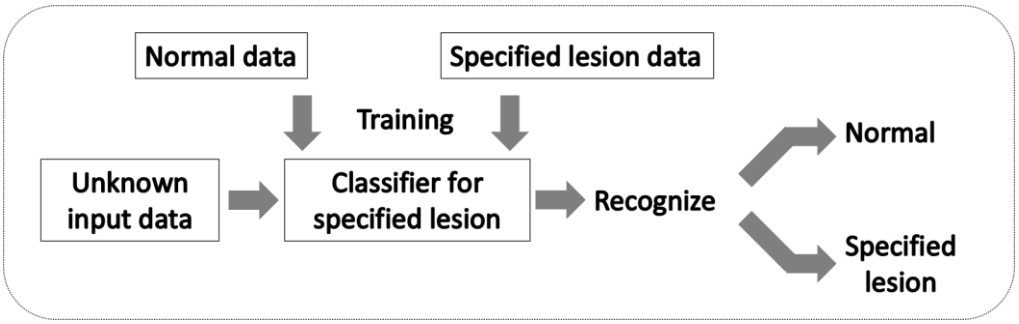


Figure 3.2.1. Flowchart of previous geometric-feature-based anomaly detection.

possess quite high similarity to the normal images. In respond to this case, it becomes necessary to take additional measures to decrease the oversights. A single-stage scheme is considered insufficient to settle all of the tasks.

3.3 Proposed anomaly detection system

Aiming at detecting all the lesions and foreign matters, and decreasing oversights of lesions, this dissertation proposes a two-stage scheme involving anomaly degree evaluation and lesion oversight prevention (Fig.3.3.1).

In the first stage of anomaly degree evaluation, it is expected to perform a role to distinguish image data with higher anomaly degree from those with lower degree. Specifically, lesions and residues that suggest to high anomaly degree are expected to be detected. To realize such anomaly degree evaluation, this research sets out from the idea by training the classifier with only the normal images. Afterwards, by measuring the scale of difference between unknown input image and the trained normal images, all suspicious contents, including various kinds of lesions and foreign matters are expected to be screen out. In contrast, normal images of intestinal lumen and wall, which have higher similarity to the trained normal images, are assigned with low anomaly degree.

The second stage following anomaly degree evaluation is named lesion oversight prevention, which is designed to prevent oversighted lesions. Although all kinds of lesions and foreign matters are desired to be detected in the first step, some lesion images, especially those with small region or inconspicuous appearance, are possibly remained in the group of low anomaly degree after anomaly degree evaluation, whereas the oversights of such lesions are still very risky. Considering that highly

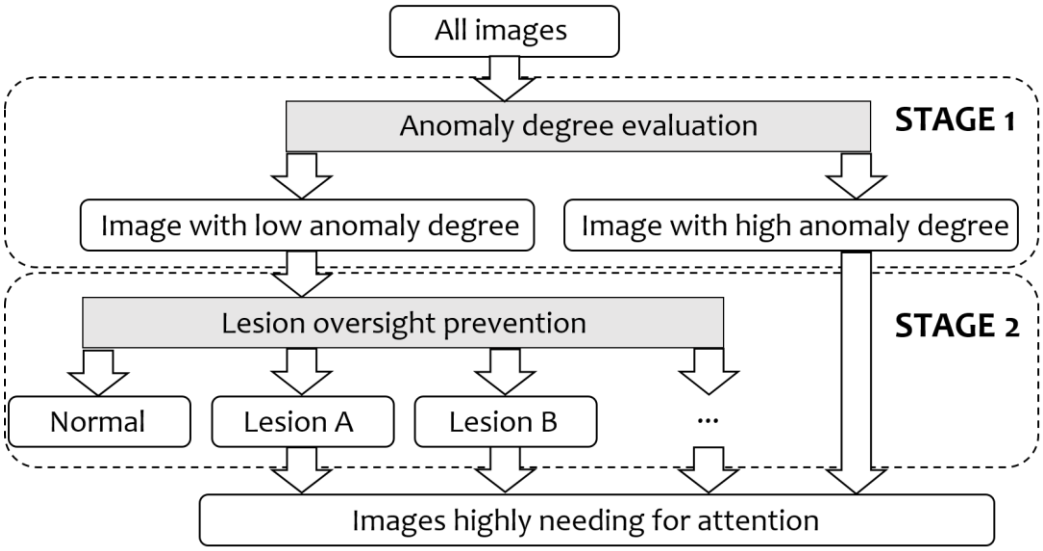


Figure 3.3.1. The proposed two-stage anomaly detection system

suspicious images should be filtered out for maximum efficacy. Both the two stages are included in the proposed anomaly detection system. Detailed information will be illustrated in the following Chapter 4, 5 and 6.

Chapter 4

Anomaly Degree Evaluation

4.1 Anomaly degree evaluation using Higher-order Local Auto-Correlation (HLAC)

4.1.1 Overview

The fundamental idea of anomaly degree evaluation, the first stage of the proposed system, employs training image datasets which contain only clear normal intestinal wall and lumen. By observing a large number of capsule endoscopy images, it is found that images of normal intestinal wall, even if collected from different patients, usually owns relatively simple and similar appearance (Fig.4.1.1.1). In contrast, as illustrated in 3.1, impeccable training for various anomalies is almost impossible, not only because there are numerical numbers of anomaly categories, but also mutable and ever-changing appearance even for same category but different state of anomalies. With regard of this fact, aiming at the

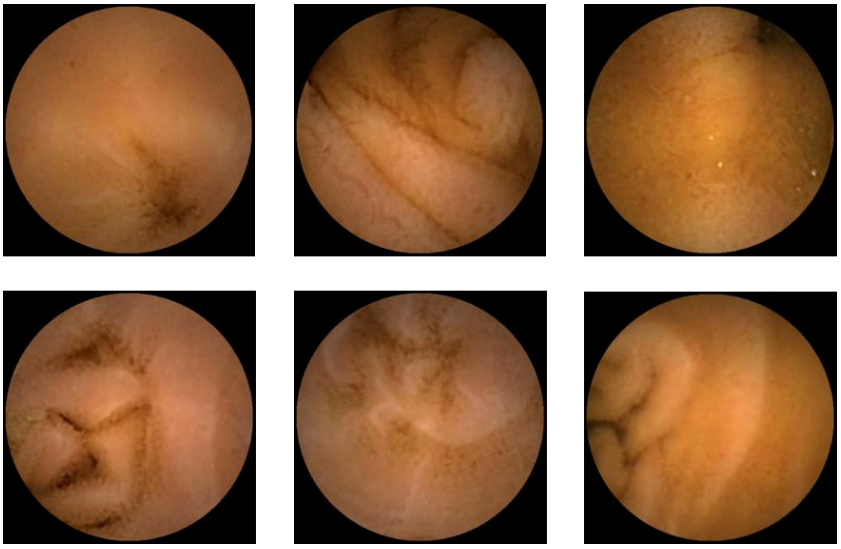


Figure 4.1.1.1. Clear normal intestinal image collected from different patients.

anomaly detection for all the lesions and foreign matters, the proposed method concentrates on describing the similarities among normal intestinal images rather than anomalies. And afterwards, anomaly detection could be interpreted that the difference between normal images and unknown test images, namely anomaly degree, is evaluated. By this means, unknown images containing anomaly contents are expected to be filtered out in accordance with the anomaly degree.

In terms of this idea, specifically speaking, the proposed anomaly degree evaluation includes two phases, training phase and testing phase (Fig.4.1.1.2). In training phase, images including only clear normal intestinal wall are employed as training samples. Since foreign matters involving residue and bile, have been considered as the detection target, the use of normal images that contain residue and bile should be avoided as much as possible. After pre-processing, Higher-order Local Auto-Correlation (HLAC) features [64] are extracted from these normal images are used to create a subspace representing the normal class. Details of this part will be given in the following section. Likewise, in testing phase, same image features are extracted from pre-processed testing images. Finally, anomaly degree is calculated in accordance with the scale of distance between the feature vector of testing image and the normal vector space.

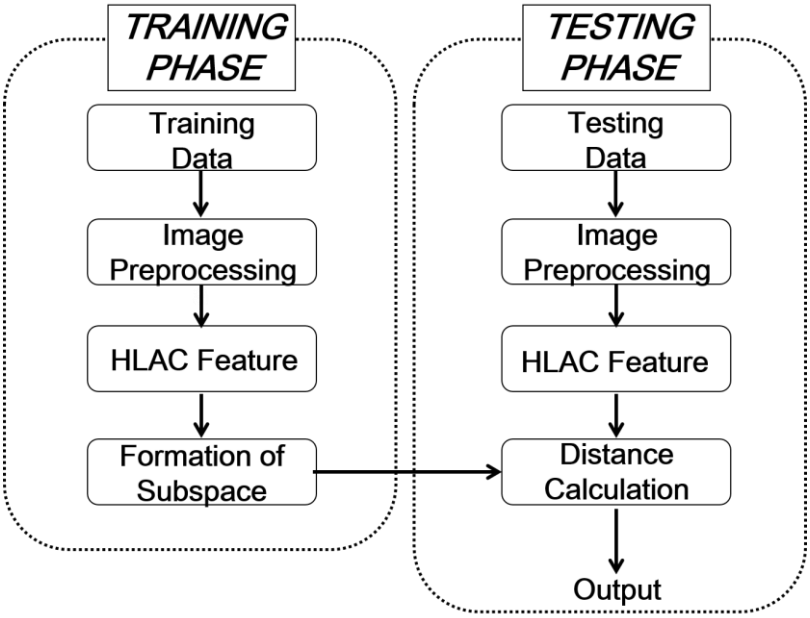


Figure 4.1.1.2. Flowchart of the proposed anomaly degree evaluation.

4.1.2 Higher-order Local Auto-Correlation (HLAC)

Higher-order Local Auto-Correlation (HLAC), develop by N. Otsu [64], is an image descriptor that has been demonstrated effective in many applications involving image classification, such as object counting [65], cancer detection from pathology images [66, 67], gesture recognition [68] and video surveillance [69].

The explanation of HLAC starts by the definition. *N*-th order HLAC is described by the following auto-correlations:

$$\boldsymbol{R}_N(\boldsymbol{a}_1, \cdots \boldsymbol{a}_N) = \sum_{r} I(\boldsymbol{r}) I(\boldsymbol{r} + \boldsymbol{a}_1) \cdots I(\boldsymbol{r} + \boldsymbol{a}_N) \tag{1}$$

where I is an intensity image and $\boldsymbol{r}$ is the location vector, so that $I(\boldsymbol{r})$ is the reference pixel. $\boldsymbol{a}_i (i = 1,2, \cdots \boldsymbol{N})$ corresponds to the displacement vector surrounding $\boldsymbol{r}$. Changing the order N leads to changes in the numbers of terms in $I(\boldsymbol{r}) \cdots I(\boldsymbol{r} + \boldsymbol{a}_N)$. To balance efficacy and the computational burden, N is generally set as $N \in \{0, 1, 2\}$ [28]. Fig.4.1.2.1 shows the gray-scale mask patterns that represent combinations of the reference pixel and surrounding correlative pixels when computing the auto-correlations. Specifically, zeroth-order HLAC obtains the intensity of the reference pixel, and first-order and second-order HLAC calculate auto-correlations between the reference pixel and surrounding correlative pixels,

indicating the apparent edge curves and evaluating curvature, respectively. Overlapping is also considered, such that squared and cubic auto-correlations of the central pixel itself are also calculated. Moreover, mask patterns arranged on the same rows can also take into consideration uniformity within various orientations (horizontal, vertical, and diagonal) and generate a variety of mask patterns based on different N and $\boldsymbol{a}_i$ values. In the case of gray-scale HLAC, the dimensionality is raised up to 35 in accordance with the auto-correlation configurations. Given the definition, HLAC benefits from a number of advantages below, which are all regarded as suitable for anomaly detection of capsule endoscopy image.

(1) Low computational cost. Due that only addition and multiplication operations are involved in the calculation, HLAC feature extraction leads to very low computational cost. Therefore, it has the potential to deal with high-speed feature extraction to meet the requirement of the large number of capsule endoscopy images.

(2) Shift invariance. No location information is recorded so that feature extraction is independent of object's location. This property is also appropriate to capsule endoscopy image because the location of anomaly objects emerge is uncertain, and due to shift invariance, it makes no impact on feature extraction and anomaly detection results.

(3) Robustness to noise. From the definition formula (1), the estimation of auto-correlation can be derived as below:

$$E(\boldsymbol{R}) = E[(\boldsymbol{S}_1 + \boldsymbol{n}_1)(\boldsymbol{S}_2 + \boldsymbol{n}_2)] = E(\boldsymbol{S}_1\boldsymbol{S}_2) + E(\boldsymbol{n}\boldsymbol{S}) + E(\boldsymbol{n}_1\boldsymbol{n}_2) \quad (2)$$

where $\boldsymbol{S}$ is the valid image signal, and $\boldsymbol{n}$ indicates the noise signal. Generally, when noise follows the gauss distribution, there is the estimation

$$E(\boldsymbol{n}) = 0 \quad (3)$$

Thus,

$$E(\boldsymbol{R}) = E(\boldsymbol{S}_1\boldsymbol{S}_2) + E(\boldsymbol{n}_1\boldsymbol{n}_2) \approx E(\boldsymbol{S}_1\boldsymbol{S}_2) \quad (4)$$

As to capsule endoscopy image, since the image quality captured by capsule endoscopy is largely limited by the illumination condition and integrated image sensor, noise is widespread in all the images. Under this circumstances, the robustness to noise, evidently, can protect the feature extraction from the effect of noise signal.

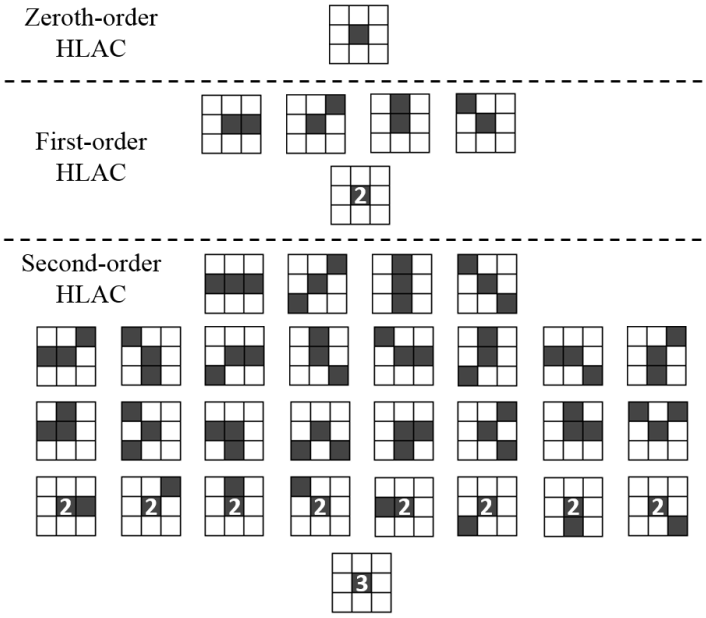


Figure 4.1.2.1. Gray-scale HLAC mask patterns.

As to the anomaly detection for capture endoscopy images, HLAC features can be extracted within a 3 by 3 local area as indicated above. However, actually, since the pixel values change gradually and may have no large obvious difference within such a 3 by 3 local area, HLAC is sometimes regarded to have a too detailed view to realize the effective measurement of the targeted objects. With such kind of concern, it is considered necessary to make adjustments to let the feature extraction become more adaptable to capture endoscopy images.

One of the commonly-used means is to alternate the size of feature extraction itself. The displacement interval can be raised so that the mask patterns are extended to a larger scale as shown in Fig.4.1.2.2. The displacement interval denotes the displacement, or the distance from one pixel to a horizontal or vertical neighboring pixel. A larger displacement interval, intuitively, can inspect pixels further depart from the others and evaluate greater change but less correlation among these pixels. Accompanied with the adjustment of the displacement interval, combination of different displacement intervals can be further adopted. In the preliminary investigation, displacement intervals 1, 2, 4, 6, 8 are tried, and the combination of displacement intervals 1, 4, and 6 has been demonstrated suitable according to the comparative results. In this way, HLAC mask patterns in the size of 3×3, 7×7, and 11×11 are adopted simultaneously, so that a three-time long HLAC feature vector are obtained by connecting the HLAC features in three sizes directly.

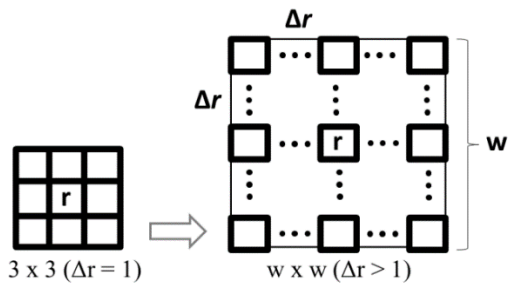


Figure 4.1.2.2. HLAC mask patterns with different displacement intervals.

4.1.3 Invariant transformation for HLAC features

In the view of capsule endoscopy images, directionality are considered making no sense to anomaly detection, because of the inconstant and uncontrollable movement and rotation of capsule endoscopy passing throughout the gastrointestinal tract. With regard to this fact, the proposed anomaly degree evaluation makes use of rotation and inversion invariant transformation of HLAC features [70], to eliminate the impact of rotation and inversion.

Rotation and inversion invariant transformation of HLAC handles mask patterns in same shape and different directions equivalently. In the light of this rule, 35-dimention gray-scale HLAC mask patterns are divided into 8 groups. The results of

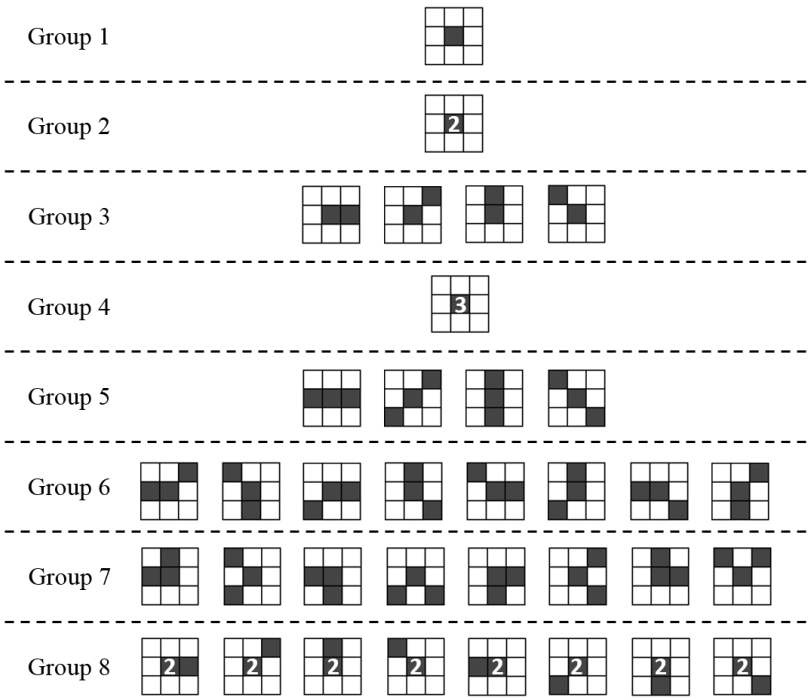


Figure 4.1.3.1. Rotation and inversion invariance transformation for HLAC.

auto-correlation calculation in same group are summated, so that the 35-dimension vectors are restructured to 8-dimension vectors (Fig.4.1.3.1). By this means, feature extraction is also benefited from rotation and inversion invariance.

4.1.4 Image processing

4.1.4.1. Brightness filter

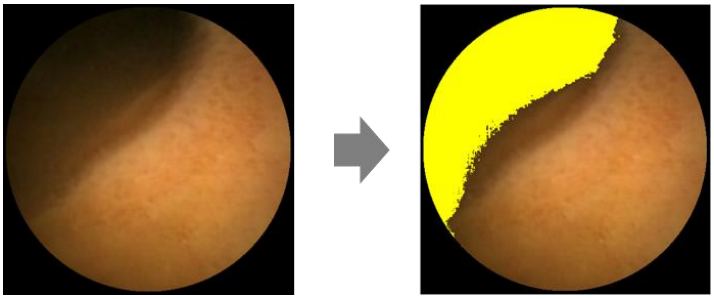
Acting as the first step of image processing, a brightness filter is presented to exclude the dark regions and excessive bright regions. During the period capsule endoscopy works in small intestine, illumination is powered by LED light. On one side, objects far from the LED illumination, such as the lumen regions in the images usually have lower brightness than intestinal wall regions. On the opposite side, bubbles and intestinal juice near to the LED illumination sometimes appears excessive bright when reflection happens. For anomaly detection, these dark regions and excessive bright are uninformative. And thus, the brightness filter is proposed to discriminate and removing these dark regions and excessive bright regions. By analysis upon a large number of images, the brightness filter is introduced below as depicted.

$$Y = 0.299R + 0.587G + 0.114B \tag{5}$$

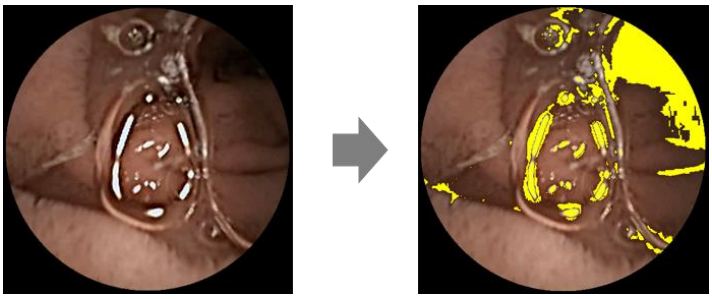
$$R = G = B = \begin{cases} 0 & (230 \leq Y \; or \; Y \leq 55) \\ No\ change & (55 < Y < 230) \end{cases} \tag{6}$$

The proposed anomaly detection complies with the principle that if a region of pixels are available should be decided by they are visible for human eyes or not. Doubt may come out that if the anomalies may occur within the dark region but the system fails to sense them when the brightness filter is adopted, and hence, anomaly detection losses the potential to provide further assistance to diagnosis. Actually, there is no need of worry because capsule endoscopy sequentially proceeds forward throughout the intestinal tract, so that objects in dark regions at certain moment can be captured when illumination arrives later.

Fig. 4.1.4.1 gives two sets of examples of the processing by the brightness filter. The first set (Fig. 4.1.4.1(a)) shows it conducted exclusion of dark region of intestinal lumen, while the other set (Fig. 4.1.4.1(b)) shows how excessive bright bubble regions got discriminated after brightness filter introduced.



(a)



(b)

Figure 4.1.4.1. Examples processed by the brightness filter.
(a) Dark region. (b) Excessive bright region.

4.1.4.2. Color space

The proposed method employs HSV color space. Although the RGB color space is the most widely used color space for image analysis, red, green, and blue components are highly correlated and are not perceptually uniform for human eyes [71, 72]. The HSV color space involves hue, saturation and value, which can be obtained from the RBG color space:

$$\mathrm{Hue}' = \begin{cases} 60° \times \frac{\mathrm{G-B}}{\mathrm{MAX-MIN}} + 0°, & \text{if MAX} = \mathrm{R} \\ 60° \times \frac{\mathrm{B-R}}{\mathrm{MAX-MIN}} + 120°, & \text{if MAX} = \mathrm{G} \\ 60° \times \frac{\mathrm{R-G}}{\mathrm{MAX-MIN}} + 240°, & \text{if MAX} = \mathrm{B} \\ 0, & \text{if R} = \mathrm{G} = \mathrm{B} \end{cases} \tag{7}$$

$$\mathrm{Hue} = \begin{cases} \mathrm{Hue}', & \text{if Hue}' \geq 0 \\ 360° + \mathrm{Hue}', & \text{if Hue}' < 0 \end{cases} \tag{8}$$

$$\mathrm{Saturation} = \frac{\mathrm{MAX-MIN}}{\mathrm{MAX}} \tag{9}$$

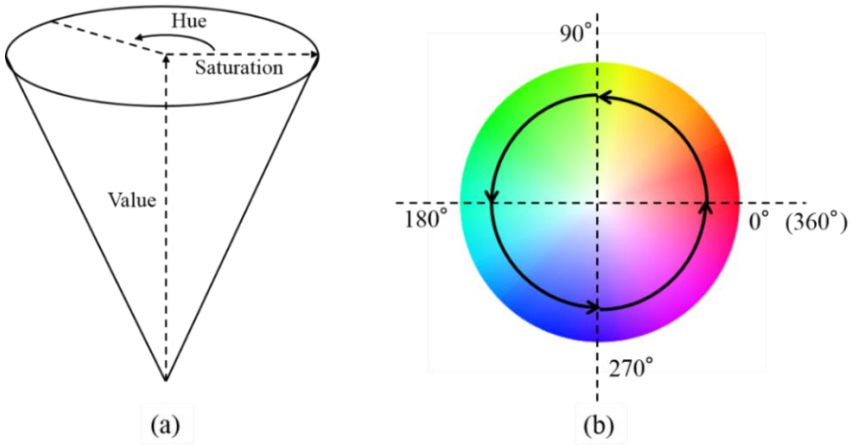


Figure 4.1.4.2. (a) HSV color space model. (b) Hue channel.

$$\text{Value} = \text{MAX} \tag{10}$$

where MAX is the largest value among R, G and B, and MIN is the smallest. Hue channel is the description indicating differential between various colors corresponding to chroma. Saturation channel refers to intensity of color and it is also deemed as the proportion of a solid color mixed with white color. Value channel represent the brightness of color. As an indication of color variety, the hue channel is independent of brightness, as the plane of the hue channel is always orthogonal to the value (brightness) axis (Fig.4.1.4.2.). This property is particularly suitable for CE images. For all forms of highlights, shades, and shadows in the intestinal lumen, the hue channel is essentially unaffected.

4.1.5 Subspace method

To verify the performance of the proposed image pre-processing and feature extraction. In the part of anomaly degree evaluation, subspace method is introduced to calculate the anomaly degree.

First of all, for the dimensionality reduction and statistical analysis of learning images, principal components analysis (PCA) is utilized in the proposed method. PCA is one of methods for multivariate statistical analysis. Among a population of samples, characteristics of individual are determined by many variables. By PCA, a smaller number of representative vectors are obtained to describe the whole population.

In the calculation, suppose there is a set of sample vectors $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2 \dots \mathbf{X}_R\} =$

$\{\mathbf{V}_1, \mathbf{V}_2 \ldots \mathbf{V}_M\}^T$, containing *M*-dimension vectors of *R* learning images. As the first step, all vectors of **X** are normalized to $\mathbf{X}_{norm}$:

$$\mathbf{S} = \{\mathbf{S}_1, \mathbf{S}_2 \ldots \mathbf{S}_M\}^T = \{\sigma_{\mathbf{V}_1}, \sigma_{\mathbf{V}_2} \ldots \sigma_{\mathbf{V}_M}\}^T \tag{11}$$

$$\mathbf{U} = \{\mathbf{U}_1, \mathbf{U}_2 \ldots \mathbf{U}_M\}^T = \{\mu_{\mathbf{V}_1}, \mu_{\mathbf{V}_2} \ldots \mu_{\mathbf{V}_M}\}^T \tag{12}$$

$$\mathbf{X}_{norm} = \{\frac{\mathbf{X}_1-\mathbf{U}}{\mathbf{S}}, \frac{\mathbf{X}_2-\mathbf{U}}{\mathbf{S}} \ldots \frac{\mathbf{X}_M-\mathbf{U}}{\mathbf{S}}\} \tag{13}$$

where **S** is a vector consist of all standard deviations of $\mathbf{V}_1, \mathbf{V}_2 \ldots \mathbf{V}_M$, while **U** includes all means of $\mathbf{V}_1, \mathbf{V}_2 \ldots \mathbf{V}_M$.

The autocorrelation matrix $\mathbf{C}$ of $\mathbf{X}_{norm}$ is obtained as following:

$$\mathbf{C} = \frac{\mathbf{X}_{norm}\mathbf{X}_{norm}^{\mathbf{T}}}{R-1} \tag{14}$$

The eigenvectors $\mathbf{U}=[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M]$ are calculated by:

$$\mathbf{U}^T\mathbf{C}\mathbf{U} = \mathbf{\Lambda} \tag{15}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_S)$ is the eigenvalue matrix.

The variable order *K* of the normal subspace is calculated according to the contribution rate η_K described in (12). When η_K exceeds a certain threshold *TH* (usually set as 0.99, 0.999…) with the smallest $K=K_1$, the order is fixed as K_1.

$$TH \leq \eta_K = \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{M} \lambda_i} \tag{16}$$

And next, with the order K_1, normal subspace $\mathbf{U}_K\mathbf{U}_K^T$ is formed based on the eigenvectors $\mathbf{U}_K = \{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ ($\mathbf{u}_1, \cdots, \mathbf{u}_k$ are called first principal component to k^{th} principal component). The distance $d_\perp$ between normalized feature vector $\mathbf{x}_{norm}=(\frac{\mathbf{x}-\mathbf{U}}{\mathbf{S}})$ of test image **x** and the normal subspace is defined as:

$$d_\perp = \mathbf{x}_{norm}^T\left(\mathbf{I}_M - \mathbf{U}_K\mathbf{U}_K^T\right)\mathbf{x}_{norm} \tag{17}$$

where $\mathbf{I}_M - \mathbf{U}_K\mathbf{U}_K^T$ ($\mathbf{I}_M$ is the M-dimension identity matrix) indicates the ortho-complement subspace of normal subspace.

As shown in Figure 4.1.5.1, in accordance to each testing image, the deviation distance between the feature vector of testing samples and the normal space are

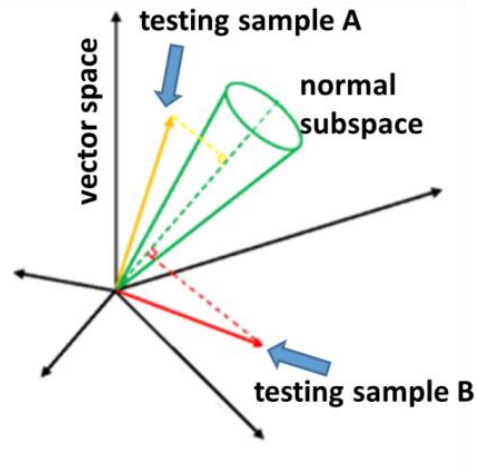


Figure 4.1.5.1. Subspace method.

calculated. In term of higher anomaly degree indicated by larger deviation distance, images containing anomaly are expected to be detected.

4.1.6 Experiments

4.1.6.1. Experimental data and procedures

In this section, capsule endoscopy images (480×480, bitmap, 24bit, RGB) are derived from one patient's video data and all of the images have been tagged in line with doctor's indexing beforehand. In the experiments, images of four types of contents are divided into two sets: training set and testing set.

Since all kinds of lesions and foreign matters (residues) that suggest to high anomaly degree are expected to be detected. To realize the anomaly degree evaluation, training images contain only clear normal intestinal wall and lumen. In term of this view, as depicted in Table 4.1.6.1, 354 images for patient 1, and 166 images for patient 2, containing only clear normal intestinal wall and lumen are selected as the learning images. To evaluate the performances of the proposed method, several categories of unused testing images are grouped according to the contents. Specifically, for patient 1, three categories contains normal intestinal wall and lumen (421 images), residue (120 images), and tumor (458 images). For patient 2, normal intestinal wall and lumen (100 images), residue (250 images), and tumor (310 images), are included in the testing datasets. According to these datasets, normal intestinal wall and lumen images are expected to be recognized as normal with low anomaly degree, while residue, tumor and bleeding images are expected

	Normal wall and lumen		Residue	Tumor	Bleeding
Patient 1	775		120	458	NA
	Training 354	Testing 421			
Patient 2	266		250	NA	310
	Training 166	Testing 100			

Table 4.1.6.1. Capsule endoscopy image dataset for 4.1.

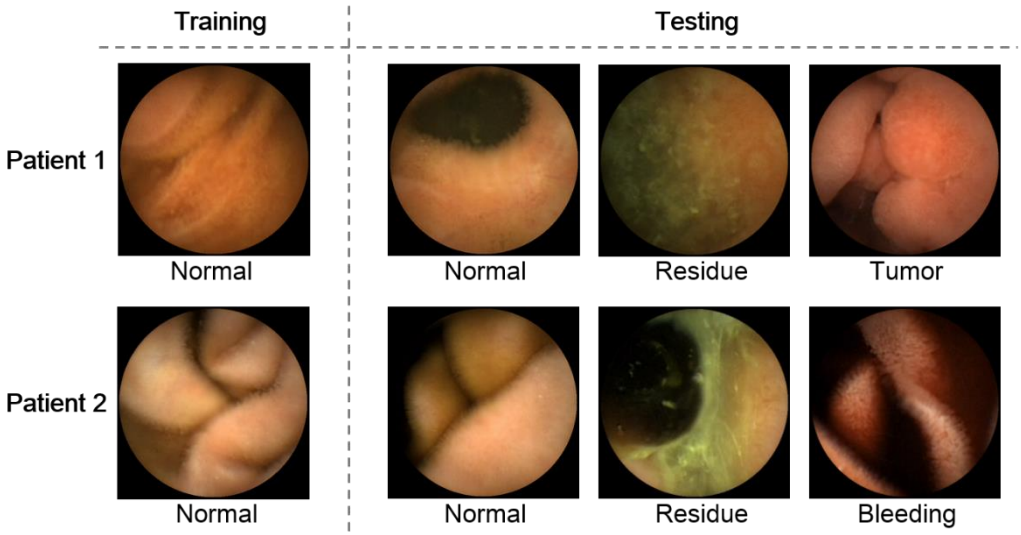


Figure 4.1.6.1. Examples of training and testing image for 4.1.

to be recognized as anomaly with high anomaly degree. In this part of experiments, anomaly degree evaluation is implemented individually for both patients. Fig.4.1.6.1 gives examples of all learning and testing sets.

4.1.6.2. Evaluation criteria

In the experiment, the anomaly degree values of all the training image data are calculated. And afterwards, a threshold T determined by the mean μ and standard deviation σ of the anomaly degrees values of training data themselves, is employed to decide if the input testing image data is normal or anomaly.

In order to measure the efficacy of the proposed method, the directly predicted results that encompass the fundamental evaluation criteria sensitivity and specificity are obtained. The definitions of sensitivity and specificity are given as following formulas:

$$\text{Sensitivity} = \frac{\text{Number of correct positive predictions}}{\text{Number of positives}} \times 100(\%) \qquad (18)$$

$$\text{Specificity} = \frac{\text{Number of correct negative predictions}}{\text{Number of negatives}} \times 100(\%) \tag{19}$$

Specificity is an indicator of normal images (negatives), while sensitivity is an indicator of anomaly images (positives). Both high sensitivity and high specificity correspond to excellent performance for binary classification.

4.1.6.3. Experimental results and discussions

The detection experimental results are shown in Table 4.1.6.3.1 and Table 4.1.6.3.2. First of all, as to the anomaly degree evaluation for patient 1, sensitivity for anomalies is 76.82%, while the specificity is 97.93%. Specifically, 120 residue images are all identified as anomaly. 324 out of 458 tumor images, and 410 out of 421 normal images are correctly detected. For patient 2, the proposed method only missed 5 images of residue and correctly recognized all the normal and bleeding image data. The sensitivity and specificity for patient 2 are 99.11% and 100%, respectively. Totally, results for both patients are posted into one table (Table 4.1.6.3.3). By the proposed anomaly degree evaluation method of anomaly degree evaluation, 999 out of 1138 anomaly images are detected so that the sensitivity of 87.79% is derived eventually. Meanwhile, specificity is well controlled at 97.89%. Accordingly to the results, it can be inferred that the proposed anomaly degree evaluation method based on anomaly degree evaluation is fundamentally valid and effective for capsule endoscopy image.

	Sensitivity	Specificity
Normal	NA	97.39%(410/421)
Residue	100.00%(120/120)	NA
Tumor	70.74%(324/458)	NA
Total	76.82%(444/578)	97.39%(410/421)

Table 4.1.6.3.1. Anomaly degree evaluation results for patient 1.

	Sensitivity	Specificity
Normal	NA	100.00%(100/100)
Residue	98.00%(245/250)	NA
Bleeding	100.00%(310/310)	NA
Total	99.11%(555/560)	100.00%(100/100)

Table 4.1.6.3.2. Anomaly degree evaluation results for patient 2.

	Sensitivity	Specificity
Normal	NA	97.89%(510/521)
Residue	98.65%(365/370)	NA
Tumor	70.74%(324/458)	NA
Bleeding	100.00%(310/310)	NA
Total	87.79%(999/1138)	97.89%(510/521)

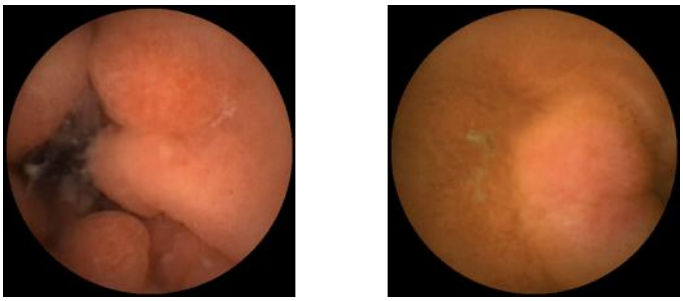
Table 4.1.6.3.3. Combined anomaly degree evaluation results for both patients.

4.2 Improved anomaly degree evaluation

4.2.1 Overview

The proposed anomaly degree evaluation has been demonstrated effective for two patients' data including normal, residue, tumor and bleeding. Especially for images of residue and bleeding, high accuracies are achieved and almost none of these anomaly data are oversighted. However, with regard to tumor images, only 70.74% was successfully detected, remaining around 30% images failed to be recognized as anomaly. The partition of oversights has surpassed the allowance range of a feasible system.

By observation of the tumor image datasets, it is found that the images belong to the one section, within which large tumor regions, as shown in Fig. 4.2.1.1(a), appear and occupy most of the image view, all possess high anomaly degree. At the same time, it has become clear that these regions present rich redness and differs from normal intestinal wall obviously. In contrast, when observe the images failed to be detected, it can be learned that these images (Fig. 4.2.1.1(b)) belong to



(a) (b)

Figure 4.2.1.1. Examples of successfully detected and failed tumor images.

different section from the above. Within these images, smaller tumor regions take the place of large ones, and meanwhile, pixels in the tumor regions have less-reddish values.

The dispersion of observations draw the conclusion that the anomaly detection method based on anomaly degree evaluation proposed in 4.1 seems powerful to tumors with obvious and large regions, but weak in detection of smaller regions with low redness. To address this issue, several breakthrough points are figured out as illustrated below.

(1) The representation of the color information suitable to capsule endoscope image is highly required, so that the tumor regions can have distinct description by image features.

(2) Feature extraction is requested to respond to anomaly regions with different sizes.

As the answer of these required specifications, in this section, a novel image processing method using non-linear conversion of HSV color space, and HLAC feature extraction based on image segmentation, are both introduced into the proposed anomaly degree evaluation additionally.

4.2.2 Non-linear conversion of HSV color space

Based on evaluations of a large number of CE images, Fig. 4.2.2.1(a) shows the distribution of their hue components. The overwhelming majority of non-zero components span the range of [10, 80], leaving the proportions of components within the range of (0, 10) and (80, 360) at less than 7.1×10^{-4} and 8.3×10^{-5}, respectively. Clearly, the narrow range of the hue component distribution leads to very limited deviations within the numerical representations along the hue axis,

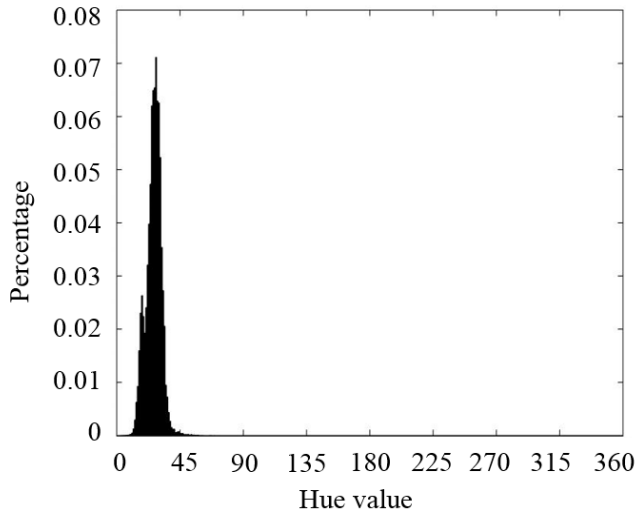
especially for reddish lesion-related components. These constrained deviations are then delivered to geometric features when feature extraction is performed on the hue channel image, and thus, generate quite limited discriminability. Meanwhile, Fig. 4.2.2.1(b) shows that relatively reddish components in lesion regions incline toward to 0, having a narrow distribution near to that of the normal components, while hue components in most normal regions have a much wider distribution. Obviously, the numerical discrimination for the lesion components are severely restricted in the ordinary HSV color space.

In order to balance the unevenly utilized range along the hue axis and to further improve discrimination for bleeding and tumor detection, the following non-linear conversion is introduced:

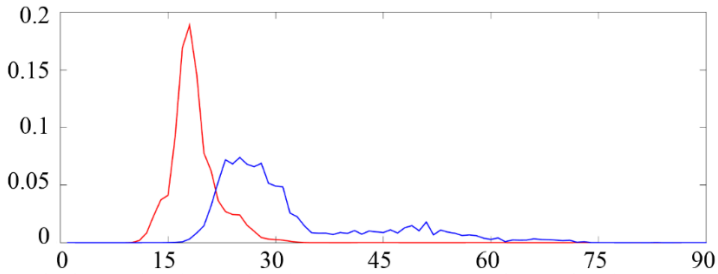
$$\mathrm{Hue}_{\mathrm{after}} = \begin{cases} 0 & (\text{ if } \mathrm{Hue}_{\mathrm{before}} = 0 \text{ or } 80 < \mathrm{Hue}_{\mathrm{before}} \leq 360) \\ 360 & (\text{ if } 0 < \mathrm{Hue}_{\mathrm{before}} < 10\ ) \\ 1.0877^{(80-\mathrm{Hue}_{\mathrm{before}})} & (\text{if } 10 \leq \mathrm{Hue}_{\mathrm{before}} \leq 80) \end{cases} \tag{20}$$

where $\mathrm{Hue}_{\mathrm{before}}$ is directly calculated from the RGB color space. To make use of the full width along the hue axis, the proposed non-linear conversion involves reversal and expansion. First, the components for $\mathrm{Hue}_{\mathrm{before}}$ within the range of [10, 80] are inversed to $80 - \mathrm{Hue}_{\mathrm{before}}$. Then, the full [0, 360] range is used by expanding the inversed components from the [10, 80] section with the exponential function $1.0877^{(80-\mathrm{Hue}_{\mathrm{before}})}$, so that the circular hue channel model can be fully utilized. As shown in Fig. 4.2.2.2, the proposed non-linear conversion has been proposed to balance the unevenly utilized range along the hue axis, and enhance the numerical discrimination for normal and lesion components.

Fig. 4.2.2.3 give the effect when the proposed non-linear conversion is conducted. Firstly, in Fig. 4.2.2.3(a), it reveals that the components within the narrow range have been mapped to the entire hue axis, so that the discrimination among the components become more explicit. Secondly, in 4.2.2.3(b), it is found that the range of normal and lesion components, as a result, have been expanded to the same level of width. Clearly, the numerical discrimination for lesion and normal components have been reinforced after the non-linear conversion.



(a)



(b)

Figure 4.2.2.1. (a) Ordinary hue components. (b) Ordinary hue components in normal and lesion regions.

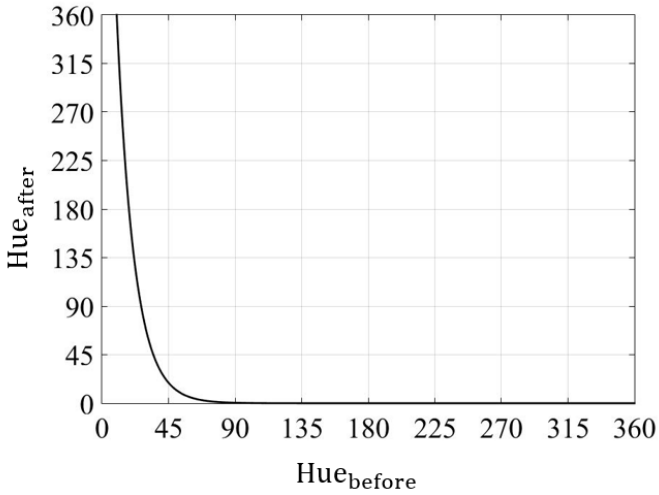
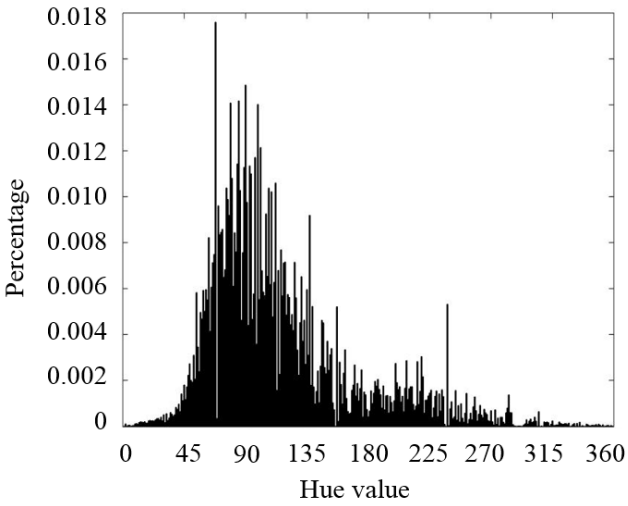
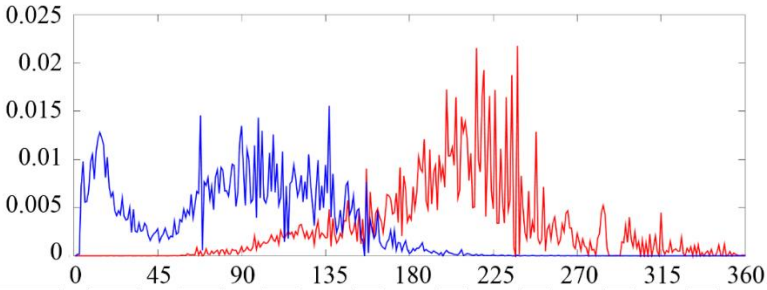


Figure 4.2.2.2.Non-linear conversion function.



(a)



(b)

Figure 4.2.2.3. (a) Converted hue components. (b) Converted hue components in normal and lesion regions

4.2.3 Image segmentation

In feature extraction, HLAC features are obtained on the coverage of whole image so that small local lesion regions make limited contribution to be reflected in feature vectors. With this issue concerned, this dissertation additionally adopts image segmentation for feature extraction in order to detect anomaly in small local regions so that such lesion regions can to be detected by this means. Specifically, one image is divided into small equal unit areas, and adjunct $n \times n$ unit areas consist of a subarea, which is regarded as a new objective region for anomaly degree evaluation. Since these subareas share overlapping areas, even though the lesion regions get split into small unit areas, it can still be reflected in more than one subarea. When conduct anomaly degree evaluation from one image, anomaly

degree of all the subareas are calculated. In the case that more than one subarea recorded as the representative. And by this means, if either of the subareas is affirmed as anomaly, the whole image is identified as anomaly.

4.2.4 Experiments

4.2.4.1. Individual performance evaluation

In the first experiment of this part, performance of the improved anomaly degree evaluation, including the non-linear conversion of HSV color space and HLAC feature extraction based on image segmentation, is implemented on individual patients' data. Proceeding after 4.1, datasets of 2 patients (Table 4.2.4.1.1) are all the same as in section 4.1. By using the same datasets, it is possible to make intuitive comparison of the improved method and the previous one.

Sensitivity and specificity are also employed in the results. As illustrated in Table 4.2.4.1.2, by using the proposed improved anomaly degree evaluation, the detection performance for the tumor images with low anomaly degrees are greatly enhanced. The sensitivity for tumor images has risen by 21%, from 70.7% to 91.7%. Among the tumor images successfully detected this time, many of the images are found owning smaller tumor regions with low redness as depicted in 4.2.1. It indicates that when the non-linear conversion of HSV color space and HLAC feature extraction based on image segmentation adopted, color representation for tumor regions become more distinct from the normal, and meanwhile small regions can be sensed within the segmented subareas. The sensitivity for residue images decreased from 98.6% to 95.7%. The accuracy for bleeding detection maintained 100% with all of the 310 bleeding images correctly detected. After introduced the improved method, specificity fell slightly from 97.9% to 96.4%. From the presented results, it has been made clear that in the individual evaluation performed separately upon two patients' data, the proposed improved anomaly degree evaluation show more capacity to cope with the tumor images, especially those owing smaller regions with low redness.

	Normal wall and lumen		Residue	Tumor	Bleeding
Patient 1	775		120	458	NA
	Training 354	Testing 421			
Patient 2	266		250	NA	310
	Training 166	Testing 100			

Table 4.2.4.1.1. Capsule endoscopy image dataset for experiments in 4.2.4.1.

	Sensitivity		Specificity	
	Before	Improved	Before	Improved
Normal	NA	NA	97.9% (510/521)	96.4%(-1.5%) (502/521)
Residue	98.6% (365/370)	95.7%(-2.9%) (354/370)	NA	NA
Tumor	70.7% (324/458)	91.7%(+21%) (420/458)	NA	NA
Bleeding	100% (310/310)	100%(±0%) (310/310)	NA	NA
Total	87.8% (999/1138)	95.3%(+7.5%) (1084/1138)	97.9% (510/521)	96.4%(-1.5%) (502/521)

Table 4.2.4.1.2. Results of evaluation on individual data.

4.2.4.2. Multi-patient performance evaluation

In 4.1.1, images of normal intestinal wall and lumen collected from different patients have shown higher similarity than various anomalies. Nevertheless, individual difference is still regarded as a great issue for automatic anomaly detection. By observing the normal and anomaly images of a large number of patients, it is found that gastrointestinal mucosa tone usually changes does not only depend on what kind of lesion exists, but the amount of bile and intestinal juice as well. Without doubt, appearance of lesion regions varies much more severely than the normal. For example, Fig. 4.2.4.2.1 presents several tumor images collected from different patients' video data. Apparently, color, size and shape of tumor regions in each images are quite variable, and meanwhile, the regions of intestinal wall show different tone as well. In the right upper image, intestinal wall looks more yellowed. It is just for the reason that bile and intestinal juice is more found than in other patients' data. Therefore, actually, to realize a feasible anomaly detection system, it is far from sufficient even that the proposed anomaly detection approach

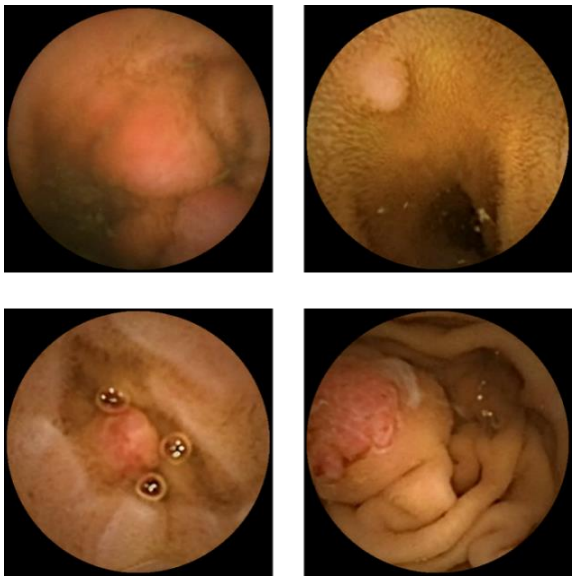


Figure 4.2.4.2.1. Example of various tumors image collected from different patients.

has been demonstrated well-performed when utilized upon individual patient's data. To meet the practical requirements, for a qualified approach, it is also necessary to have good generalization to deal with various image of multi-patient data.

Therefore, in the second part, to further confirm the capacity of improved anomaly degree evaluation method, experiments are implemented on a larger datasets, which are composed of 20 patients' image data (Table 4.2.4.2.1). Specifically speaking, datasets of 6 out of 20 patients only own normal images with none of anomaly images included. And this normal dataset, first of all, is reserved for training the normal subspace, exclusively in the training phase. Besides, 4 patients' data contain both normal and bleeding images, while 3 patients' data contain both normal and tumor images, respectively. At last, images of residue are collected from 4-patient dataset, 3-patient dataset, and other patient's data differing from the formers. With consideration of the numbers of images, a training dataset is obtained on 1962 normal images, and testing datasets containing 2193 bleeding images, 74 tumor images, 1906 residue images, and 2905 normal images.

	Normal	Normal &bleeding	Normal &tumor	Residue
Number of patient	6	4	3	10
Number of normal images	1962	910	1995	NA
Number of anomaly images	NA	2193	74	1906

Table 4.2.4.2.1. Datasets utilized in 4.2.4.2.

Moreover, in order to measure the contribution of each technique within improved anomaly degree evaluation, experiments are divided into 4 groups by adjusting the utilization of the non-linear conversion of HSV color space and the image segmentation. Detailed combinations are illustrated in Table 4.2.4.2.2. Group 1 utilizes neither non-linear conversion of HSV color space nor the image segmentation. Group 2 utilizes non-linear conversion of HSV color space but no image segmentation. Group 3 is opposite to Group2. The image segmentation is adopted but non-linear conversion of HSV color space is exclude. Group 4 takes advantage of both the non-linear conversion of HSV color space and the image segmentation simultaneously.

Results are presented in Table 4.2.4.2.3. To make the presentation more understandable, accuracies for various kinds of data obtained by using different groups of techniques are directly used instead of separated sensitivity and specificity. In general, aiming at all categories of anomalies, Group 4 using both the non-linear conversion of HSV color space and the image segmentation achieved the highest accuracy at 98.17%. 4119 of 4173 anomaly images including bleeding, tumor and residue are successfully detected. As to each of the anomaly datasets, accuracies for bleeding, tumor and residue are valued as 99.27% (2177/2193), 87.84% (65/74), and 98.48% (1877/1906), respectively. Within normal dataset, 81.62% normal images are correctly recognized. Across the four groups, it can be found that the most important results, performances on bleeding and tumor datasets, both show advances when improved anomaly degree evaluation techniques are adopted. For bleeding datasets, although there are no severe changes among different groups, the accuracy is enhanced step by step in the case that one or both

	Non-linear Conversion of HSV Color Space	Image Segmentation
Group 1	no	no
Group 2	yes	no
Group 3	no	yes
Group 4	yes	yes

Table 4.2.4.2.2. Four groups of improved anomaly degree evaluation techniques.

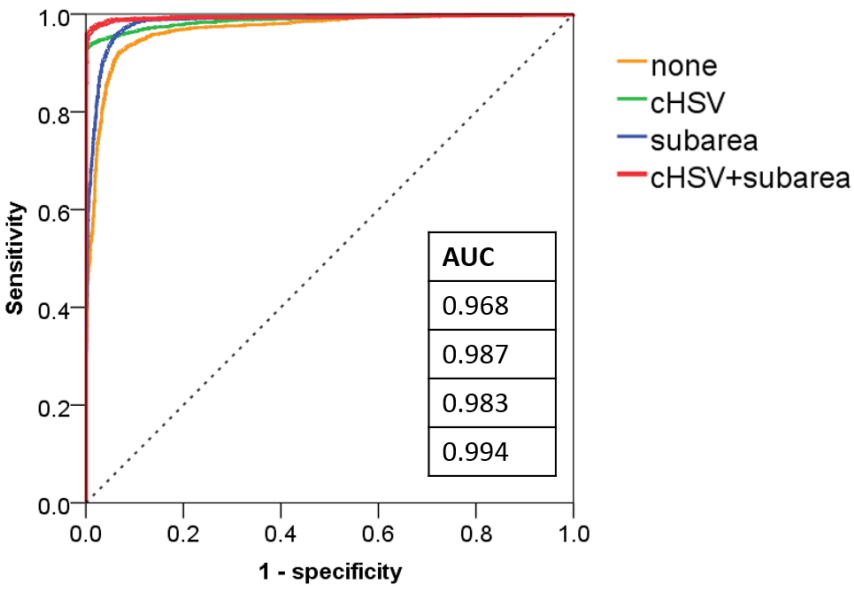
of the techniques are involved. Tendency of the change happens same for the tumor datasets. Nevertheless, the enhancement seems much more obvious since the accuracy has been raised from 58.11% (43/74) up to 87.84% (65/74). This refers that the effectiveness for bleeding and tumor detection has been indeed enhanced in contrast of the previous method. As to results for normal dataset, the best performance is obtained by Group 3. However, the accuracy fell down from 88.43% to 81.62% after the non-linear conversion of HSV color space was added. The reason is considered that non-linear conversion made color components along the Hue axis more distinct, and because of this, even in normal images, subareas with color components slightly close to those in anomaly regions may have larger anomaly degree. Therefore, some of the normal images are identified as anomaly by mistake.

	Normal	Bleeding	Tumor	Residue	All anomalies
Group 1	76.49% (2222/2905)	97.26% (2133/2193)	58.11% (43/74)	98.74% (1882/1906)	97.24% (4058/4173)
Group 2	83.86% (2436/2905)	97.40% (2163/2193)	59.46% (44/74)	97.85% (1865/1906)	97.58% (4072/4173)
Group 3	88.43% (2569/2905)	98.59% (2162/2193)	60.81% (45/74)	98.74% (1882/1906)	97.99% (4089/4173)
Group 4	81.62% (2371/2905)	99.27% (2177/2193)	87.84% (65/74)	98.48% (1877/1906)	98.17% (4119/4173)

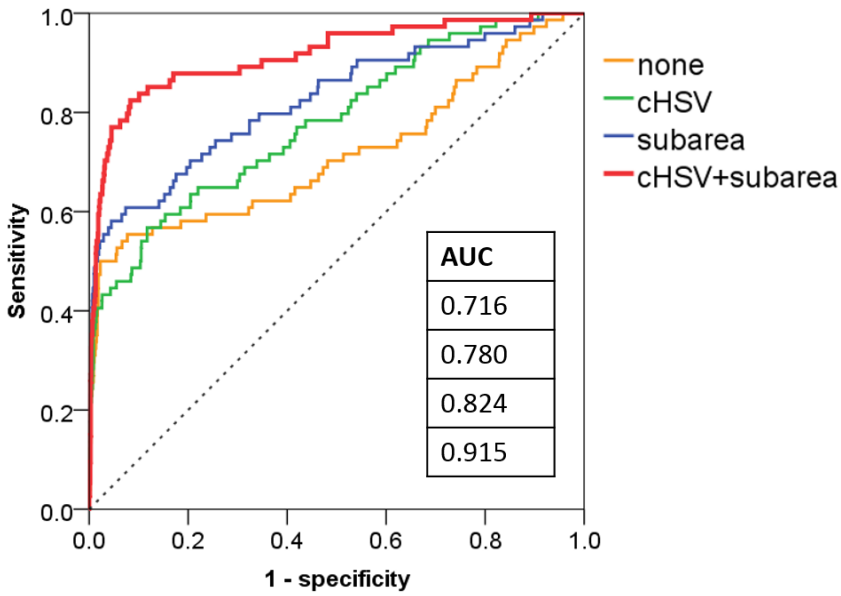
Table 4.2.4.2.3. Results of multi-patient performance evaluation.

Furthermore, ROC curves are obtained for all normal and anomaly test samples. ROC curves were traced out by changing the threshold for anomaly degree. Without a fixed threshold, the relationship between sensitivity and specificity can be expressed more objectively. Accordingly, in the results relating to the ROC curves, curves with large AUC corresponding to the dynamic combination of high sensitivity and specificity are expected.

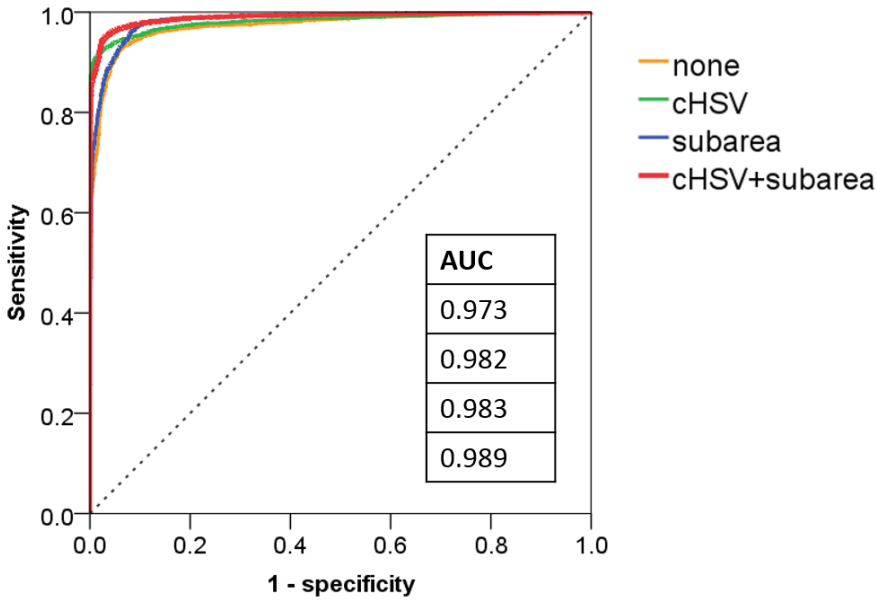
The ROC curves shown in Fig. 4.2.4.2.2 work out an agreement with Table 4.2.4.2.3. It needs to be explained that in ROC curves, legend "cHSV" is short for conversion of HSV color space, and "subarea" indicates the image segmentation. In Fig. 4.2.4.2.2(a), the improved anomaly degree evaluation, presented by the red line, has the largest AUC value of 0.994. This means the improved anomaly degree evaluation obtained most remarkable and balanced performance for both the normal and bleeding images. Situation of tumor datasets (Fig. 4.2.4.2.2(b)) is same. The AUC values have markedly risen from 0.716 to 0.915. With all of normal, bleeding, tumor and residue datasets counted, as shown in Fig. 4.2.4.2.2(c), the advantages of improved anomaly degree evaluation continues to be proved as AUC value of 0.989 surpasses all the other groups. In summary, by the experiments implemented on multi-patient datasets, not only the capacity of each newly developed techniques are demonstrated to be contributory to anomaly detection, but the generalization of the improved anomaly degree evaluation method has been confirmed when it is used upon anomaly datasets collected from multiple different patients.



(a)



(b)



(c)

Figure 4.2.4.2.2. ROC curves for multi-patient performance evaluation. (a) Bleeding. (b) Tumor. (c) All anomalies (bleeding, tumor and residue).

4.3 Summary

In this chapter, the first stage of the anomaly detection system, namely anomaly degree evaluation, is developed on the basis of the principle that, it is promising to detect various anomalies by measuring the scale of difference between unknown image and the training images of normal intestinal wall and lumen. In the section 4.1, the anomaly degree evaluation is proposed by taking advantage of Higher-order Local Auto Correlation for feature extraction, and subspace method as the classifier. Besides, the proposed method implements image pre-processing procedures to fulfill suitable recognition for capsule endoscopy images. In the section 4.2, aiming at reducing oversights of anomaly images, moreover, an improved scheme involving a non-linear color space conversion of HSV color space and image segmentation is introduced. In the experiments, the effectiveness of improved anomaly degree evaluation is validated by both of individual and multi-patient evaluation.

Chapter 5

Lesion Oversight Prevention

5.1 Overview

As the stage of proposed anomaly detection system, lesion oversight prevention is presented to prevent oversighted lesions possibly remained after anomaly degree evaluation, in particular, for the lesion images with small or inconspicuous lesion regions are easily oversighted.

Previous studies [73, 74] focused on the various anomalies of CE images using feature extraction based on higher-order local auto-correlation (HLAC). Analyses have revealed that HLAC can encompass the three features of color, texture, and shape, and that it has a number of advantages, including low computational cost, shift invariance, and robustness to image noise. However, in the case of anomaly detection from CE images with small or inconspicuous lesion regions, problems appeared that HLAC was less effective and local small-scale lesions can be easily overlooked and thus lead to false-negative predictions.

Regarded as the stage to restrain oversight, the lesion oversight prevention is desirable to have a generalizable scheme for more mutable lesion cases. In order to

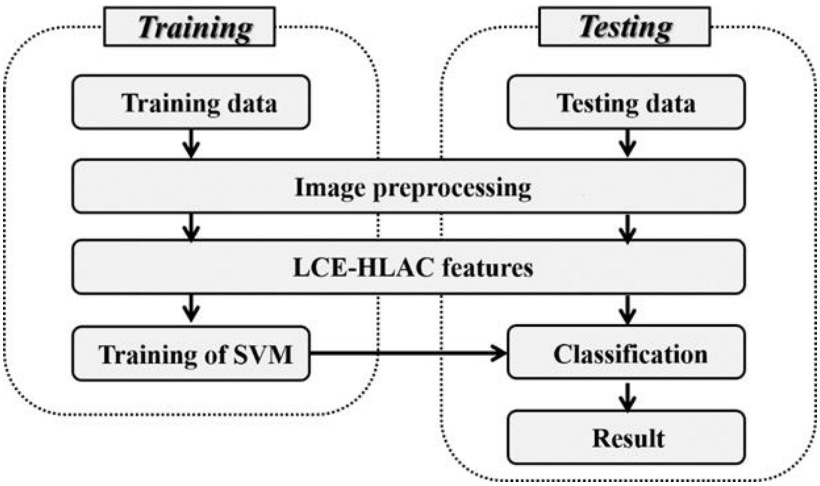


Figure 5.1.1. Processes of proposed lesion oversight prevention.

solve this problem, in this stage, the dissertation will introduce an improved geometric descriptor called LCE-HLAC (Local Contrast Enhanced Higher-order Local Auto-Correlation), which is considered as a prospective solution to remedy the low descriptive capacity for local details by HLAC and realize more effective feature extraction for lesions of various appearance.

The proposed lesion oversight prevention consists of two phases same as anomaly degree evaluation, namely training and testing (Fig. 5.1.1), and three main processes, namely image pre-processing, geometric feature extraction, and classification. During both stages, to cope with imbalances within the selected color space, image pre-processing involving a non-linear conversion of the color space is executed. Then, the proposed approach for geometric feature extraction is applied to obtain image features. The features extracted from the training dataset, which include both normal and anomalous images, are used to train the support vector machine (SVM) classifier, which then classifies the test dataset, which also consists of normal and anomalous images.

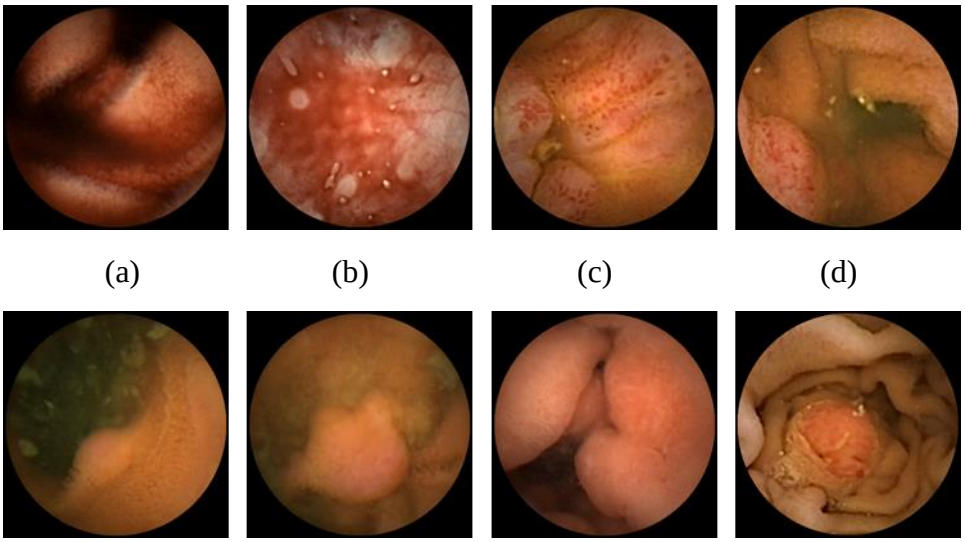
5.2 Issues and objectives

Due to the restriction of available data, same as in Chapter 4, frequently occurring lesions including bleeding and tumors, which are usually associated with serious lesions in clinical diagnosis are regarded as objects with high necessity to be detected, and thus, employed in this section.

Bleeding images for different patients can sometimes have different appearances

due to the conditions of bleeding (blood's tone and quantity) and individual variations. For example, in Figs. 5.2.1(a) and (b) more shed blood can be observed than in Figs. 5.2.1(c) and (d). Such distinctions are strongly affected by the seriousness of bleeding, and also by the distance between the bleeding and the location where the image is captured. Images taken closer to the bleeding spot often contain more shed blood.

Tumors are defined as abnormal masses of tissue attached to the lumen surface, which can be either benign or malignant. Bowel polyps, such as those shown in Figs. 5.2.1(e), (f), and (g), are benign tumors. Figures 5.2.1(e) and (f) present the most common single polyps, which are not as harmful as malignant tumors. Single polyps are usually reddish in appearance, and are either partially or completely round. Polyps collected from a segment affected with the Cronkhite-Canada syndrome (Fig. 5.2.1(g)) are also benign tumors. However, unlike single polyps, multiple polyps are widespread. To prevent benign tumors from becoming malignant, clinically, early detection for benign tumors is regarded very important. The tumor shown in Fig. 5.2.1(h) is another case of a tumor selected from our database. It has been diagnosed as duodenum papilla cancer, which is a kind of malignant tumor that must be treated soon using clinical procedures. From investigating various kinds of tumor, tumor images can vary in appearance even more than bleeding images, in terms of tumor area, size, shape, and color across cases.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 5.2.1. Examples of (a)-(d) bleeding and (e)-(h) tumor images.

In contrast to its efficacy for detecting anomalies within large areas, as in Fig. 5.2.1(g), the HLAC feature is somewhat deficient in coping with some images that include smaller local anomalies, as in Figs. 5.2.1(e), (f), and (h). Since HLAC calculates a summation of uniformly weighted auto-correlations across the entire image and lacks detailed representations for small-scale anomalous regions, features extracted from these local anomalous regions tend to be submerged and are thus not fully reflected within the feature. Given this concern, it is necessary to develop an approach for feature extraction that is more efficient for the local details associated with small-scale anomalies, and yet still retains the major advantages of HLAC. Local anomalies usually have morbid appearances within small lesion regions, with distinct changes from the surrounding normal intestinal wall. As a solution, the evaluations of local contrast, which are particularly appropriate for the detection of small-scale anomalies, is introduced into the HLAC feature. This extension is called LCE-HLAC. By employing LCE-HLAC, it is possible to formulate for peculiar pixels or areas that differ from their surroundings.

5.3 Local Contrast Enhanced Higher-order Local Correlation (LCE-HLAC)

5.3.1 Foundation and interception

To evaluate the local contrast, the auto-correlations are divided according to intensity differences between $I(\boldsymbol{r})$ and neighboring pixels $I(\boldsymbol{r}+\boldsymbol{a})$, which are measured by a set of functions. When N is restricted as $N \in \{0, 1, 2\}$. In each case, the auto-correlations for LCE-HLAC are defined, respectively, as follows:

$$\boldsymbol{R}_{\mathrm{N}=0}\left(f_{k_0}\right) = \sum_{\boldsymbol{r}\in I} I(\boldsymbol{r})\ f_{k_0}\left(I(\boldsymbol{r}), \overline{I(\boldsymbol{r}+\boldsymbol{a})}\right), \quad k_0 \in \{1,2,3\} \tag{21}$$

$$\boldsymbol{R}_{N=1}\left(\boldsymbol{a}_1, f_{k_1}\right) = \sum_{\boldsymbol{r}\in I} I(\boldsymbol{r}) I(\boldsymbol{r}+\boldsymbol{a}_1)\ f_{k_1}\left(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_1)\right), \quad k_1 \in \{1,2,3\} \tag{22}$$

$$\boldsymbol{R}_{N=2}\left(\boldsymbol{a}_1, \boldsymbol{a}_2, f_{k_1}, f_{k_2}\right)$$

$$= \sum_{\boldsymbol{r}\in I} I(\boldsymbol{r}) I(\boldsymbol{r}+\boldsymbol{a}_1)\, I(\boldsymbol{r}+\boldsymbol{a}_2)\, f_{k_1}\left(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_1)\right) f_{k_2}\left(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_2)\right),$$

$$k_1, k_2 \in \{1,2,3\} \tag{23}$$

where f_k corresponds to mutually exclusive functions.

$$f_1\big(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_i)\big) = \begin{cases} 1 & , \mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i) > T \\ 0 & , \mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i) \leq T \end{cases} \tag{24}$$

$$f_2(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_i)) = \begin{cases} 1 & , \mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i) < -T \\ 0 & , \mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i) \geq -T \end{cases} \tag{25}$$

$$f_3(I(\boldsymbol{r}), I(\boldsymbol{r}+\boldsymbol{a}_i)) = \begin{cases} 1 & , |\mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i)| \leq T \\ 0 & , |\mathrm{I}(\boldsymbol{r}) - \mathrm{I}(\boldsymbol{r}+\boldsymbol{a}_i)| > T \end{cases} \tag{26}$$

For $\boldsymbol{a}_i \in \boldsymbol{R}^N$, f_1, f_2, and f_3 execute the respective comparisons between all pairings of neighboring pixels and the reference pixel. Since f_1, f_2, and f_3 represent mutually exclusive conditions, at any given point, only one of the three functions can be activated and conduct the corresponding calculation of auto-correlations. T is the judgment threshold. Smaller (larger) T values make feature extraction more (less) sensitive to local contrasts. Since T is a dominant factor on sensitivity to local contrast information, preliminary discussions about determining its thresholds are provided in the following section. In the case of $N = 0$, given that only the reference pixel is involved in auto-correlation, this dissertation introduces $\overline{I(\boldsymbol{r}+\boldsymbol{a})}$, which indicates the average intensity of d surrounding pixels, to replace $I(\boldsymbol{r}+\boldsymbol{a}_i)$.

$$\overline{I(\boldsymbol{r}+\boldsymbol{a})} = \frac{1}{d}\sum_{j=1}^{d} I\big(\boldsymbol{r}+\boldsymbol{a}_j\big) \tag{27}$$

LCE-HLAC can be thought of an extension of HLAC. Fig. 5.3.1.1(a) shows examples of mask patterns when N is 0, 1, and 2, respectively. In each situation, the mask patterns for auto-correlation are expanded according to the evaluations for local contrasts by the judgment functions f_k. In the zeroth-order case ($N = 0$), apart from intensity, the description of sharply fluctuating pixels, possibly associated with part of a small anomalous area or discrete anomalous pixels, is clearly distinguished from those referring to insignificant changes. First- and second-order LCE-HLAC benefit from further enhancements for evaluating edge curves. For example, Fig. 5.3.1.1(b) demonstrates that some diagonal second-order LCE-HLAC pattern masks are capable of obtaining edge information when scanning beyond the boundaries of a tumor region. Due to the obvious difference between normal and tumor sides, comparisons executed by the judgment functions f_1, f_2,

and f_3 sort the auto-correlations of object pixels into different pattern masks. In this way, edge characteristics are easily reflected within the feature vector. Moreover, because LCE-HLAC comprehensively covers all pixel intensities, co-occurrence relations within various patterns, and contrastive relations for the reference and surrounding pixels, it effectively encompasses all three feature types of color, texture, and shape. Similar to HLAC feature extraction, to fit the scale of targeting objects in capsule endoscopy, the size of LCE-HLAC mask pattern is decided as 19×19 in the same way which was denoted in Chapter 4.

In terms of computational complexity, considering all the dimensions shown in Table 10, the feature vector of LCE-HLAC involves a higher degree of dimensionality when extended from the original HLAC. Given that various contrast conditions are being considered, the dimensions of the LCE-HLAC feature vectors increase from 35 to 225. Actually, little additional computational cost is incurred along with the increase of dimensionality, because the mutually exclusive nature of the judgment functions means that auto-correlation calculations are executed sparsely under exclusive conditions, and no overlapping calculations are involved. However, the comparative calculation between central pixel and surrounding or neighboring pixels has become the reason of more practical computational cost. When working under MATLAB platform (MATLAB 2015a, Windows 7, Intel 4790K, 32GB), feature extraction by LCE-HLAC needs around 10~12 times the computing time of the original HLAC. But in other computational environment such as Visual C++ (Microsoft Visual Studio 2013, Windows 7, Intel 4790K, 32GB), it seems that the computational cost becomes no longer a problem because LCE-HLAC feature can be extracted at about 24 fps from the 480×480 images. If more optimization can be involved, high-speed processing is considered definitely possible to cope with even the real-time capsule endoscopy image data.

Besides, naturally, LCE-HLAC also inherits the other HLAC advantages of shift invariance, additivity, and robustness to noise. The materials below give a demonstration for noise robustness of LCE-HLAC.

Taking the first order as an example, the estimation of LCE-HLAC can be described as:

$$E(\boldsymbol{R}) = E[(\boldsymbol{S}_1 + \boldsymbol{n}_1)(\boldsymbol{S}_2 + \boldsymbol{n}_2)\, f_{k_1}((\boldsymbol{S}_1 + \boldsymbol{n}_1), (\boldsymbol{S}_2 + \boldsymbol{n}_2))] \qquad (28)$$

As demonstrated in equation (2), Chapter 4, the formula can be transformed to:

$$E(\boldsymbol{R}) \approx \boldsymbol{E}(\boldsymbol{S}_1\boldsymbol{S}_2)\boldsymbol{E}[\, f_{k_1}((\boldsymbol{S}_1+\boldsymbol{n}_1),(\boldsymbol{S}_2+\boldsymbol{n}_2))] \tag{29}$$

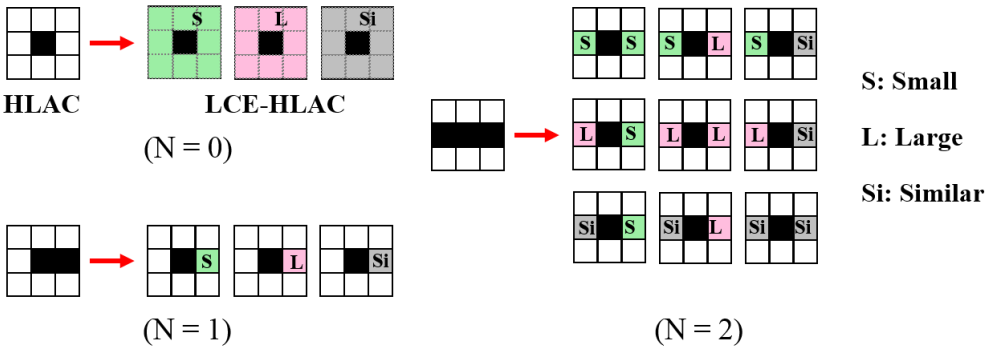
The latter part $\boldsymbol{E}[\, f_{k_1}((\boldsymbol{S}_1+\boldsymbol{n}_1),(\boldsymbol{S}_2+\boldsymbol{n}_2))]$ is for the estimation of the mutually exclusive functions. Since only difference value is evaluated by the mutually exclusive functions, it can be referred as below:

$$\boldsymbol{E}[\, f_{k_1}((\boldsymbol{S}_1+\boldsymbol{n}_1),(\boldsymbol{S}_2+\boldsymbol{n}_2))] = \boldsymbol{E}[(\boldsymbol{S}_1+\boldsymbol{n}_1)-(\boldsymbol{S}_2+\boldsymbol{n}_2)] =$$

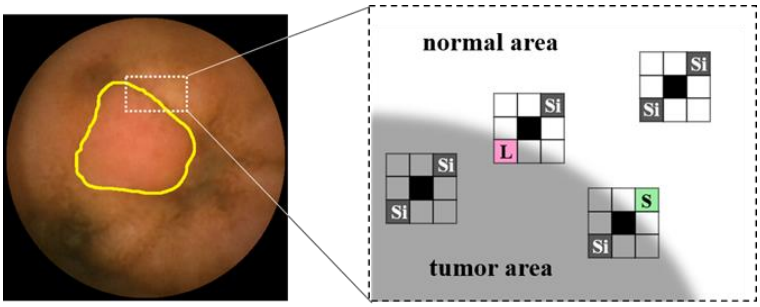
$$\boldsymbol{E}(\boldsymbol{S}_1)-\boldsymbol{E}(\boldsymbol{S}_2)+\boldsymbol{E}(\boldsymbol{n}_1)-\boldsymbol{E}(\boldsymbol{n}_2) = \boldsymbol{E}(\boldsymbol{S}_1)-\boldsymbol{E}(\boldsymbol{S}_2) = \boldsymbol{E}[\, f_{k_1}(\boldsymbol{S}_1,\boldsymbol{S}_2)] \tag{30}$$

Therefore, the estimation of LCE-HLAC $E(\boldsymbol{R})$ can be derived as

$$\boldsymbol{E}(\boldsymbol{R}) \approx \boldsymbol{E}(\boldsymbol{S}_1\boldsymbol{S}_2)\boldsymbol{E}[\, f_{k_1}(\boldsymbol{S}_1,\boldsymbol{S}_2)] = \boldsymbol{E}[\boldsymbol{S}_1\boldsymbol{S}_2\, f_{k_1}(\boldsymbol{S}_1,\boldsymbol{S}_2)] \tag{31}$$



(a)



(b)

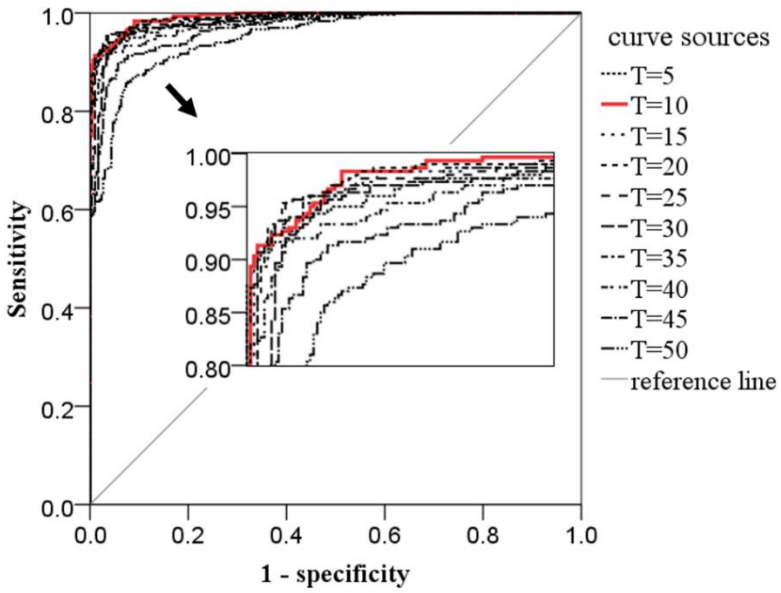
Figure 5.3.1.1. Example of how LCE-HLAC mask patterns extract edge features.

	HLAC	LCE-HLAC
Zeroth-order	1	3
First-order	5	15
Second-order	29	207
Total	35	225

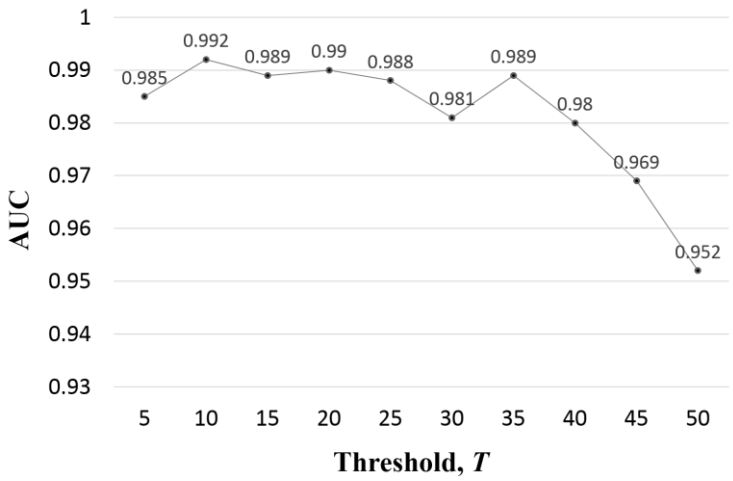
Table 5.3.1.1. Dimensionality of HLAC and LCE-HLAC.

5.3.2 Determining threshold T

In order to acquire an applicable *T*, this dissertation utilize a small dataset of CE images (480×480, bitmap, 24bit, RGB). As the detection of tumor images is generally regarded to be a more difficult task than bleeding detection, this work created a dataset including 600 normal and 300 tumor randomly selected images, and employed principal component analysis (PCA) to construct a subspace of the feature vectors and evaluate the performances of alternative *T* values. Specifically, feature vectors extracted from normal images are firstly used to construct a normal feature subspace. Then, the distances between the LCE-HLAC feature vectors for the remaining 300 normal images and 300 tumor images and the established subspace are calculated and equated as distinction degrees. These distinction degrees and then analyzed so that to produce their receiver operating characteristic (ROC) curves (Fig. 5.3.2.1(a)). Finally, the area under the curve (AUC) is then calculated to determine the best-performing threshold *T*, which corresponds to the largest AUC value. Based on our investigations of various images, *T* is assigned with a set of values ranging from 5 to 50. As shown in Fig. 5.3.2.1(b), $T = 10$ leads to the largest AUC, *T* values of below 40 lead to no significant differences, and *T* values of over 40 tend to be detrimental to performance. In our experiments, LCE-HLAC with $T = 10$ is applied to validate its efficacy and for comparison with other approaches.



(a)



(b)

Figure 5.3.2.1. Performance obtained with various threshold T values in preliminary PCA validations. (a) ROC curves and (b) comparison of AUC values.

5.4 SVM-based classification

SVM [75] is one of the most popular state-of-the-art classifiers. The principle behind SVM is to construct a hyperplane on which image feature vectors can be separated by the maximized margin between classes [76]. Consider a set of data samples of the form $\{(y_1, \boldsymbol{z}_1), (y_2, \boldsymbol{z}_2), \ldots, (y_n, \boldsymbol{z}_n)\}$. $y_i \in \{-1, +1\}$ indicates the labels of the classes that the feature vectors $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n$ $(\boldsymbol{z}_i \in \boldsymbol{R}^N)$ belong to. In this work, for anomaly detection, positive samples are valued +1, while negative

samples that correspond to the normal class are valued -1. Then, the hyperplane can be written as:

$$\mathbf{z} \in F: (\mathrm{w} \cdot \mathbf{z}) - b = 0 \tag{32}$$

where w is the weight vector and b is a non-negative parameter. If data samples can be separated linearly, then two parallel hyperplanes, $\boldsymbol{H}_1$: $(\mathrm{w} \cdot \mathbf{z}) - b = -1$ and $\boldsymbol{H}_2$: $(\mathrm{w} \cdot \mathbf{z}) - b = +1$, should exist. Since the distance between the two hyperplanes (margin) is $\frac{2}{\|\mathrm{w}\|}$, to obtain the maximized margin, calculations for satisfying the parameters w and b are required. Accordingly, the problem is equivalent to minimizing $\|\mathrm{w}\|^2/2 = \mathrm{w}^t\mathrm{w}/2$, with constraints on points $y_i(\mathrm{w} \cdot \boldsymbol{z}_i) - b \geq 1$, i =1, 2, …, n. By introducing the Lagrange multiplier $\alpha_i \geq 0$, i =1, 2,…, n, the problem is further transformed to one of finding a solution that minimizes the following formula:

$$L(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \, \alpha_j y_i y_j \boldsymbol{z}_i{}^t \boldsymbol{z}_j \tag{33}$$

As a result, the proper parameter w^* can be expressed as:

$$\mathrm{w}^* = \sum_{i=1}^{N} \alpha_i{}^* y_i \, \boldsymbol{z}_i \tag{34}$$

where $\alpha_i{}^*$ is the solution for Eq. (15). Finally, the resultant y on a new sample vector $\boldsymbol{z}_{n+1}$ is calculated as follows:

$$y = \mathrm{sgn}(\sum_{i=1}^{N} \alpha_i{}^* y_i \, \boldsymbol{z}_i{}^t \boldsymbol{z}_{n+1} - b\,) \tag{35}$$

When the input data are inseparable by a linear SVM as above, a kernel function $k(\boldsymbol{z}_i,\boldsymbol{z}_j) = (\Phi(\boldsymbol{z}_i) \cdot \Phi(\boldsymbol{z}_j))$ is applied [77]. $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ are n-dimensional inputs. Φ is a map from the n-dimension to an m-dimensional space (m > n). Commonly, kernel functions include linear functions, polynomial functions, and radial basis functions.

5.5 Experiments

5.5.1 Experimental data

Capsule endoscopy data were collected from videos for 28 clinical patients. Observations and diagnoses were carried out by three experienced doctors. For each patient, the expert doctors compiled detailed reports including diagnostic comments and recording time ranges for normal/lesion segments. Based on these reports, this research first sampled video segments that had been documented as containing certain kinds of data, and then, extracted from these segments complete sets of images (480×480, bitmap, 24bit, RGB), before excluding any seriously distorted images, such as those full of bubbles, strong bile, residues, or taken under extreme lighting conditions. The remaining images were subsequently categorized and labeled. In this way, a complete dataset of 69 video segments that yielded 5,642 normal images from 33 segments, 5,476 bleeding images from 22 segments, and 1,164 tumor images from 14 segments (Table 5.5.1.1) is obtained. Before being used, the data for each category were double checked by the doctors to ensure that all the image data were consistent with their opinions. Note that all of these segments were separated so that the diversity of the dataset could be guaranteed.

Within an entire capsule endoscopy video, suspected lesions usually occur in partial sections, and so, generally speaking, it is much easier to obtain normal data than lesion data, which is reflected in the imbalance of normal, bleeding, and tumor

Category	Normal	Bleeding	Tumor
Number of patients	14	14	8
Number of segments	33	22	14
Number of images	5642	5476	1164
Number of images used for bleeding detection	5476	5476	-
Number of images used for tumor detection	1164	-	1164

Table 5.5.1.1. Categories of experimental data in 5.5.

images within our database. However, to eliminate any influences of this imbalance that may cause failures for small partitions of the training data, while simultaneously fully utilizing all the lesion data, this work created non-overlapping training and testing datasets that consisted of the same number of normal images as lesion images (Table 5.5.1.1). For each category, the normal/lesion images are first separated from each patient into two equal parts. Then one part for each patient is assigned to create the training dataset and assigned the other part to create the testing dataset. In accordance with the principle of SVM, all normal images and lesion images are labeled as -1 and +1, respectively.

5.5.2 Procedures and evaluation criterion

In order to demonstrate the performance of the proposed color space conversion and the feature extraction method, analyses were conducted for experiments on bleeding and tumor detection. In evaluating the performance of the proposed method, this work also employed the basic HLAC feature and the RIU-LBP feature for comparison. For evaluating the proposed color space conversion, feature extraction was implemented on both the proposed non-linear conversion of the HSV color space (cHSV, the same below) and the ordinary HSV color space. Moreover, CIE L*a*b*, another color space that separates color from intensity, was utilized as another control group. Given that the V (value) channel of the HSV color space and the L channel of the CIE L*a*b* color space both refer to brightness, which is strongly related to lighting conditions, this paper excluded these channels and only utilized the H-S channels and a*-b* channels, respectively. This dissertation concatenated the feature vectors obtained from each channel to create a long feature vector in the experiments.

In order to measure the efficacy of the proposed method and to compare with other control methods, the directly predicted results that encompass the fundamental evaluation criteria sensitivity and specificity are first obtained. And furthermore, the ROC curves for all normal and anomaly test samples are drawn accordingly. SVM was employed to estimate positive probability for all test samples. ROC curves were traced out by changing the threshold for positive probabilities. Accordingly, in the results relating to the ROC curves, curves with large AUC corresponding to the dynamic combination of high sensitivity and specificity are expected.

5.5.3 Results and discussions

Summaries of results for bleeding and tumor detection are presented in Tables 5.5.3.1 and 5.5.3.2 respectively. In Table 5.5.3.1, there are no conspicuous differences among the different groups of approaches. The best performance in terms of total accuracy (98.80%) was obtained by the combination of the original HLAC feature and the proposed cHSV color space model. However, the highest sensitivity (99.49%), which depicts the best bleeding detection result, was obtained by the proposed method using the LCE-HLAC feature and color space conversion. For all groups, the average accuracies are larger than 97.48%, with no sensitivity and specificity values of less than 96.90% and 96.09%, respectively. Such high accuracies clearly indicate that all method groups had considerable efficacies for bleeding samples. The ROC curves, as shown in Fig. 5.5.3.1(a), verify this result.

With respect to the detection of tumors, as presented in Table 5.5.3.2, the results yielded by the different methods diverge more widely than those in Table 5.5.3.1. The best performance was obtained by the proposed method, with a sensitivity of 96.74%, a specificity of 88.14%, and a total accuracy of 92.44%. In terms of sensitivity, tumor detection based on the combination of the LCE-HLAC feature and the proposed cHSV color space was 3.1 percentage points better than the best results obtained by the HLAC group. The highest specificity (94.33%) was obtained when the LCE-HLAC feature and the original HSV color space were adopted. This indicates that although the efficacies in terms of total accuracy and sensitivity achieved using the LCE-HLAC feature and cHSV color space surpass those obtained upon the other color spaces, false-positive predictions were best controlled by using the original HSV color space.

For the first group of LCE-HLAC, in addition to being the best accuracy combination, the method using the ordinary HSV space reached 91.41% for total accuracy and 88.49% for sensitivity. However, decreases emerged when the LCE-HLAC feature extraction was combined with the CIE L*a*b* color space. With the CIE L*a*b* color space, only 79.21% of tumor images were detected correctly. The situation for the group utilizing the original HLAC feature is similar. The HLAC feature yields matched performances adopting both the ordinary HSV color space and the cHSV color space, but the sensitivity suffers a severe drop to 74.40% when the CIE L*a*b* color space is introduced. For the RIU-LBP feature group, however, unlike the groups of LCE-HLAC and HLAC, both the CIEL*a*b* and

original HSV color spaces performed similarly for RIU-LBP feature extraction. It can also be inferred that the proposed cHSV color space is adaptive for RIU-LBP feature, since it obtained a total accuracy of 88.75% and a sensitivity of 90.89%, which are slightly higher than those for the other color spaces. Generally speaking, the color space did no substantially affect the RIU-LBP group's performances.

Fig. 5.5.3.1(b) presents ROC curves for all groups. Consistent with the predicted results of accuracy, the proposed method, with both the LCE-HLAC feature and color space conversion, had the largest AUC value (0.977, see Table 5.5.3.2). From this result, it can be indicated that the consequences of both false-positive predictions and false-negative predictions are well restrained by the proposed method.

Moreover, the results for tumor detection vary greatly according to the technique adopted. From further analysis, it is identified that the images with multiple tumor regions, such as in Fig. 5.5.3.2(a), tend to have concentrations of strong reddish components. In contrast to Figs. 5.5.3.2(b), (c), and (d), which show smaller local tumors, the histogram of cHue components in Fig. 5.5.3.2(a) shows an obvious shifting to the right side. However, the majority of cHue components in Figs. 5.5.3.2(b), (c), and (d) are similar to those in the normal intestinal image, such as in Fig. 5.5.3.2(e). Moreover, pixel values of local tumors are assumed to be depicted only within the small region beyond 160. The divergent detection results indicate that LCE-HLAC is the most suitable for the detection of local tumors. Compared to HLAC, the edge characteristic between tumor region and the adjacent normal lumen region are more effectively extracted by LCE-HLAC.

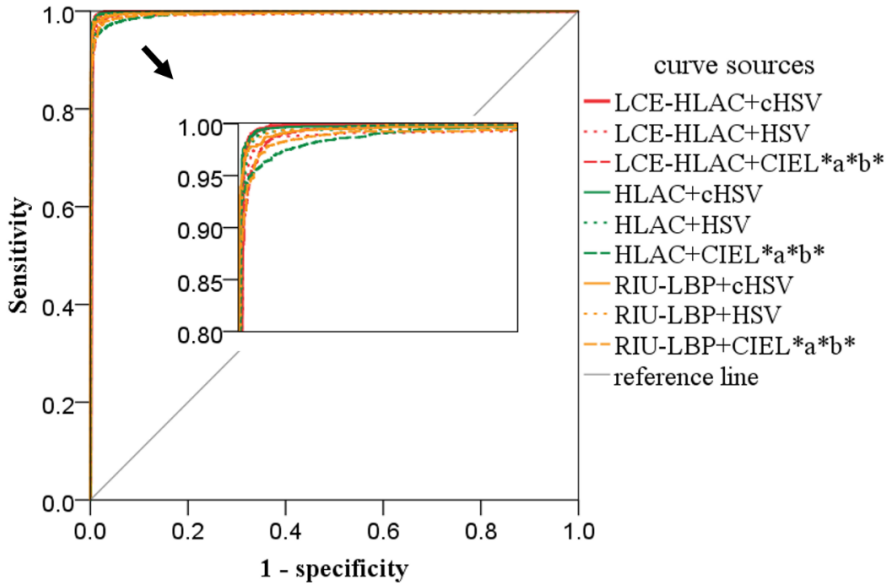
These results clearly demonstrate that the combination of LCE-HLAC feature extraction and non-linear conversion for the HSV color space is efficient for both bleeding and tumor detection. Its performance is superior to that for the control conditions consisting of various technique combinations for tumor detection.

		Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
LCE-HLAC	cHSV	98.54	99.49	97.59	0.997
	HSV	97.94	98.39	97.48	0.991
	CIE L*a*b*	97.48	98.87	96.09	0.996
HLAC	cHSV	98.80	99.20	98.39	0.997
	HSV	98.47	99.09	97.85	0.994
	CIE L*a*b*	97.66	96.90	98.43	0.994
RIU-LBP	cHSV	98.08	97.81	98.36	0.997
	HSV	98.32	98.98	97.66	0.997
	CIE L*a*b*	98.21	99.12	97.30	0.994

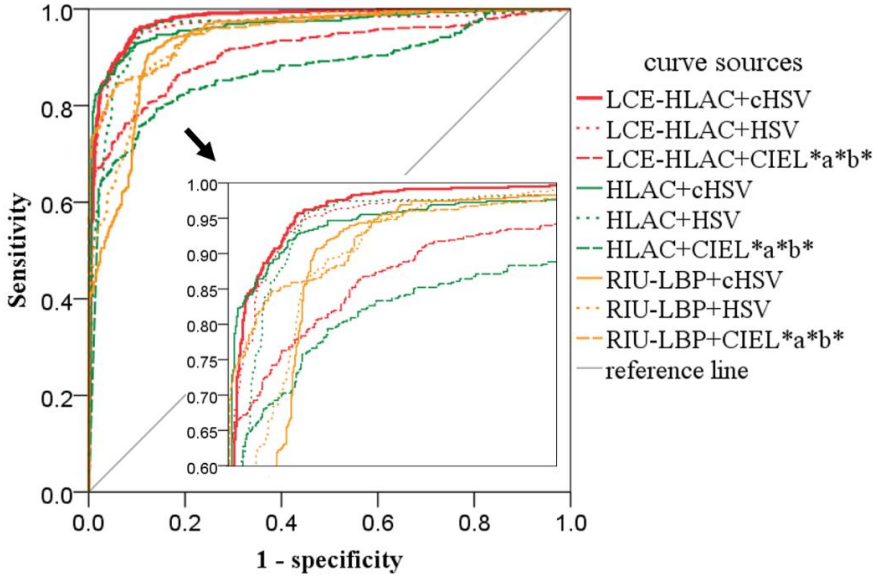
Table 5.5.3.1. Classification results for bleeding detection in terms of accuracies and AUC values (areas under ROC curves).

		Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
LCE-HLAC	cHSV	92.44	96.74	88.14	0.977
	HSV	91.41	88.49	94.33	0.965
	CIE L*a*b*	84.02	79.21	88.83	0.913
HLAC	cHSV	91.07	93.13	89.00	0.965
	HSV	91.67	93.64	89.69	0.961
	CIE L*a*b*	82.05	74.40	89.69	0.874
RIU-LBP	cHSV	88.75	90.89	86.60	0.938
	HSV	87.46	88.66	86.25	0.944
	CIE L*a*b*	86.68	87.97	85.40	0.956

Table 5.5.3.2. Classification results for tumor detection in terms of accuracies and AUC values (areas under ROC curves).



(a)



(b)

Figure 5.5.3.1. ROC curves for (a) bleeding, (b) tumor detection.

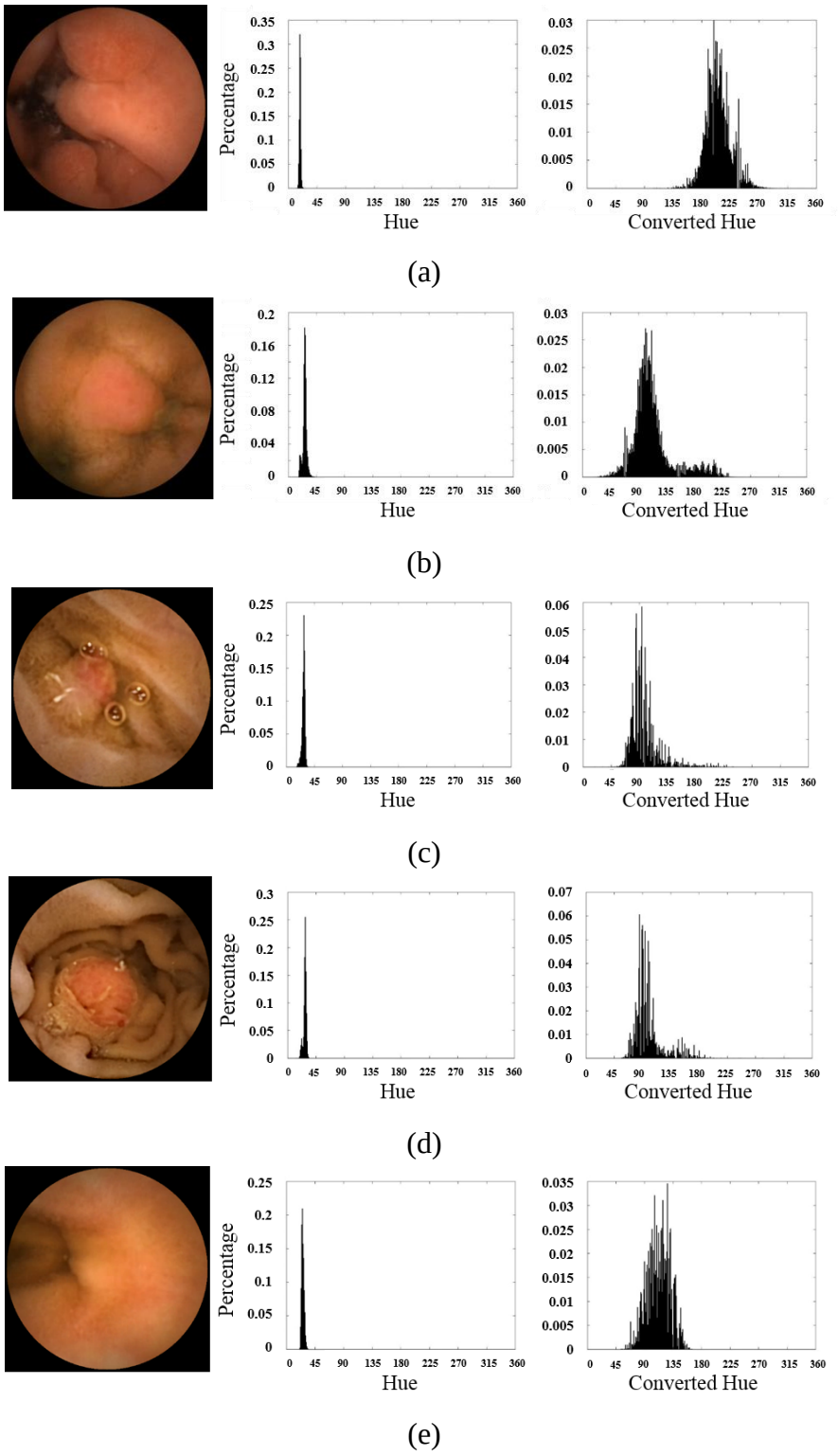


Figure 5.5.3.2. Histogram analysis. (a) Tumor image with multiple tumor regions, (b)-(d) tumor images with local tumor regions, and (e) normal intestinal image.

5.6 Summary

In this chapter, the lesion oversight prevention aiming to restrain oversight, especially for the lesion images with small or inconspicuous lesion regions, is proposed by introducing an improved geometric descriptor called LCE-HLAC (Local Contrast Enhanced Higher-order Local Auto-Correlation). LCE-HLAC is developed to inspect the co-occurrence of neighboring (surrounding) pixels with different contrasting relationship. By this means, LCE-HLAC is expected to enhance the description for local details by HLAC, and realize more effective feature extraction for lesions of various size and appearance. Combined with the previously proposed techniques such as the non-linear conversion of HSV color space, LCE-HLAC has been demonstrated efficient and superior to control conditions for both bleeding and tumor detection.

Chapter 6

Feasibility Validation of the Proposed Anomaly Detection System

6.1 Overview

In Chapter 4 and 5, the two-stage scheme system involving anomaly degree evaluation and lesion oversight prevention has been introduced in turn. With regard to the feasibility of the proposed system, it is necessary to validate how it performs in a practical situation. Therefore, this section will principally present several experiments, which are simulative but completely in the same form of the practical usage.

As explained in Chapter 3, the first stage of anomaly degree evaluation is to measure the difference between the input image and the normal intestinal image. The images owning high anomaly degree and low anomaly degree are then classified into two categories. Afterwards, the images with low anomaly degree will be rechecked by the second stage, lesion oversight prevention, to eliminate oversights of lesion images with small or inconspicuous lesion regions.

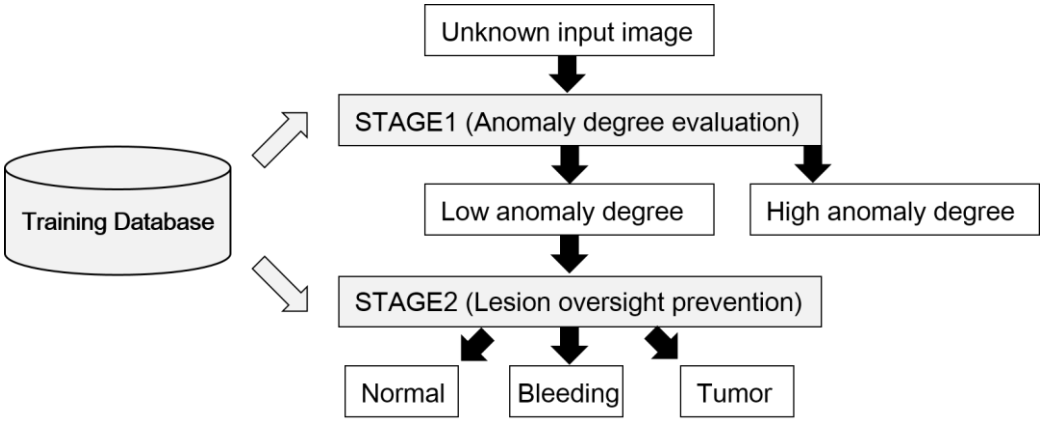


Figure 6.1.1. Feasibility validation by establishing a general training database.

As the expected usage of the proposed two-stage scheme, images with high anomaly degree after the first stage, and all images judged as certain kind of anomaly in the second stage, will be suggested to doctor for further observation. In the light of the configuration of the proposed system, no matter under either the screening mode, or the double check mode, the unused image data collected from other patients will be imported into the proposed system and processed throughout both stages.

With regard to the design of the anomaly detection system, in advance, it is necessary to construct a general database as explicated in Fig. 6.1.1, to train the normal subspace in the first stage, and the SVM-based lesion recognizer in the second stage. By using the training database, the system can be constructed and adjusted to conduct self-governed anomaly detection.

6.2 Experiments

6.2.1 Experimental data

The feasibility validation experiments employ 28 patients' data in total as expressed in Table 6.2.1.1. According to the source of data, all images (480×480, bitmap, 24bit, RGB) are five categories: normal, bleeding, tumor, normal&bleeding, and normal&tumor. The categories mean that only one kind of specified content is included within one patient's data, or two kinds of contents are both belong to one patient's data. Since detection for residue images has been sufficiently discussed in Chapter 4, relevant validation is not involved in this section. Among the employed

	Normal	Bleeding	Tumor	Normal &bleeding	Normal &tumor
Number of patients	6	10	5	4	3
Number of normal images	1962	NA	NA	910	1995
Number of anomaly images	NA	2358	1090	2193	74
Index of datasets	A1	B1	C1	Normal: A2 Bleeding: B2	Normal: A3 Tumor: C2
Usage	Training Database	Training Database	Training Database	Testing	Testing

Table 6.2.1.1. Datasets utilized feasibility validation in 6.2.

datasets, meanwhile, images in each datasets are further indexed to facilitate the use for different purpose. Indexes “A”, “B” and “C” indicate normal images, bleeding images and tumor images respectively. In this way, 1962 normal images from 6 patients are indexed as A1. 2358 bleeding images from 10 patients are indexed as B1. 1090 tumor images collected from 5 patients are named C1. Besides, within 4 patients’ data containing both of normal and bleeding, 910 normal images are employed and indexed as A2, while 2193 bleeding images are indexed as B2. Similarly, 1995 normal image belong to 3 patients’ data are named A3, remaining 74 tumor images with the index C2.

With regard to the training database of the proposed anomaly detection system, A1, B1 and C1 are used to establish the experimental database. Specifically for the two stages, A1 is not only employed to construct the normal subspace in the first stage, but also used to train the SVM classifier in the second stage. Meanwhile, since only normal training data is demanded in the first stage, the datasets of B1 and C1 are just used in the second stage. As to the evaluation, datasets containing A2, B2, A3 and C2 are treated as testing input data. 2905 normal images, 2193 bleeding images and 74 tumor images, collected from 7 patients in all, are included in these data.

6.2.2 Settings and evaluation criterion

Throughout the proposed anomaly detection system, all anomaly images are expected to be detected, and at the same time, false positive cases (normal images

identified as anomaly) is supposed to be reduced as much as possible. Within the anomaly detection, parameters can be adjusted within the training database and the classifiers. On this occasion, in the first step, threshold T that divides high anomaly degree and low anomaly degree is ruled by anomaly degree values of all the normal images themselves in training dataset A1. By setting different threshold T, the output data and statistical results of the first stage vary accordingly, and thus, the input data of the second stage will also be changed. With consideration of this issue, in the first step, threshold T is determined by altering a fixed number P:

$$P = \frac{\text{Number of training images(Anomaly degree} \leq T)}{\text{Number of training images(All)}} \tag{36}$$

where P indicates the percentage of training images possessing anomaly degree values no larger than the threshold T. For example, when P is set as 0.9, the threshold will be precisely equivalent or larger than 90% of the training images' anomaly degrees. In practice, threshold for the first stage is set by adopting P=0.9 and 0.95, and in this way, accuracies for lesion images (true positive, sensitivity) and normal images (true negative, specificity) can be obtained under respective conditions.

In the second stage, bleeding classifier trained by datasets A1 and B1, and tumor classifier trained by A1 and C1, are applied to all the low-anomaly-degree output images. Parameters in the both classifiers should be optimized upon the performance of the training database. After this, the images contain undetected bleeding and tumor images, and all normal images identified as "normal" in the first stage. As the use of bleeding classifier and tumor classifier are parallel as depicted in Figure 3.3.1 (Chapter 3), each classifier can export its own true negative rate for normal images, and true positive rate for either bleeding images or tumor images. It is noteworthy that bleeding images imported into the tumor classifier are possible to be detected as tumor, and conversely, tumors image may be detected as bleeding image as well. On the principle that all suspicions are suggested to the doctor, such phenomenon is no harmful but beneficial to the effectiveness. At the same time, true negative predictions determined by each part are summarized as the final true negative output. That means only the normal images recognized as "normal" in all parts of the system are finally exported as "normal". Throughout the whole system, at last, total accuracies for every dataset are calculated as the results of feasibility validation of the proposed anomaly detection system.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Number of testing images}} \tag{37}$$

6.2.3 Results and discussions

Experimental results for P=0.9 and P=0.95 are respectively denoted in Table 6.2.3.1 and 6.2.3.2. In both tables, accuracies for separated datasets and stages are all recorded. Firstly in the case of P=0.9, it is observed that the propose anomaly detection system demonstrates performances for bleeding images, tumor images and normal images as 99.86%, 94.59% and 72.63%, in accuracy, respectively. Viewing from stage 1, 99.27% of bleeding images, 87.84% of tumor images, and 81.62% of normal images have been corrected recognized. After lesion oversight prevention is introduced, accuracies for bleeding and tumor datasets are both improved as 81.25% oversighted bleeding images and 55.56% oversighted tumors are detected additionally. In particular, it is noticed that within the correctly detected bleeding images, apart from the 10 images discovered by bleeding classifier, tumor classifier has also contributed to finding more bleeding images, though these images are identified as the member of tumor class. The total accuracy for normal dataset decreases from 81.62% to 72.63%, this is for the reason that each part generates false-positive predictions, and so that, the true-negative percentages are finally multiplied to be a smaller value.

By reviewing 6.2.3.2, it is found that in the case of P=0.95, accuracies for bleeding, tumor and normal datasets are achieved as 99.82%, 93.24% and 80.00%, respectively. Similarly, the stage 2 has been demonstrated able to reduce lesion oversights, and improve the total accuracies by 1.01 percentage points and 13.51 percentage points for bleeding and tumor datasets respectively. However, in the stage 2, bleeding classifier and tumor classifier have both contributed to the crossing detection (one kind of lesion is detected as the other), with 3 bleeding images and 2 tumor images successfully filtered out. With regard to the normal datasets, accuracy decreases from 91.98% to 80.00% due to the false-positive predictions emerge in the stage 2.

From the Table 6.2.3.1 and 6.2.3.2, it can be inferred that the setting of P has enormous influence on the performances, especially for the normal dataset. In the stage 1, a larger P makes smaller allowance for false-positive predictions. In the stage 2, more false-positive predictions have been brought into the system when the number of input data increases along with P. Therefore, the total accuracy for

normal images is indirectly, definitely raised up by a larger *P*. When *P* is altered from 0.9 to 0.95, the accuracy of stage 1 is increased by 10.36 percentage points, and the total accuracy after the stage 2 is increased by 7.37 percentage points.

Setting a larger *P* can lead to higher accuracy for normal datasets, but actually, the accuracies for bleeding and tumor detection have shown no apparent difference. In the stage 1, although the accuracy for tumor dataset falls down from 87.84% to 79.73%, it is denoted that the stage 2 is able to make up the difference by detecting more tumor images when *P* is raised up from 0.9 to 0.95. In practical use, the setting of *P* should be taken consideration with the property of training database (e.g. similarity of training images) and the doctor's demand (e.g. the partition of normal images expected to be filtered out in advance).

Some examples of bleeding images and tumor images, which are failed to be detected by the system, are shown in Fig. 6.2.3.1. The reason why the proposed anomaly detection system fails to detect the bleeding images seems due that there is almost no fresh blood visible within the image as the first image shows, but only dark blood spots that may indicate the bleeding has ever happened at earlier time before the examination. Such dark blood spots have fade in color and become too dark to be figured out by the proposed system. Tumor images are oversighted for the reason may be ascribed to that the tumor region is revealed at the center of the intestinal lumen, and meanwhile, the view of tumor image is at the time of intestinal contraction. Tumor regions is not only partly hidden by intestinal wall, but show low brightness and inconspicuous color.

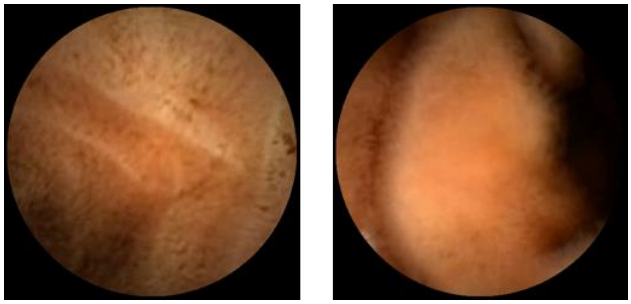
In future practical use, although a small part of lesion cases could not be detected by evaluating the single image, many other methods remain available to make the doctor notice the lesion image. The solution could be established, for example, by introducing evaluation between a series of consecutive image frames instead of one single image. If one out of the series are detected as lesion, neighboring images should also be highly concentrated accordingly.

	Bleeding	Tumor	Normal (bleeding detection)	Normal (tumor detection)
Stage 1	99.27% (2177/2193)	87.84% (65/74)	81.62%(2371/2905)	
Stage 2	81.25% (10+3/16)	55.56% (5+0/9)	94.39% (2238/2371)	92.11% (2184/2371)
Total	99.86% (2190/2193)	94.59% (70/74)	77.04% (2238/2905)	75.18% (2184/2905)
			72.63%(2110/2905)	

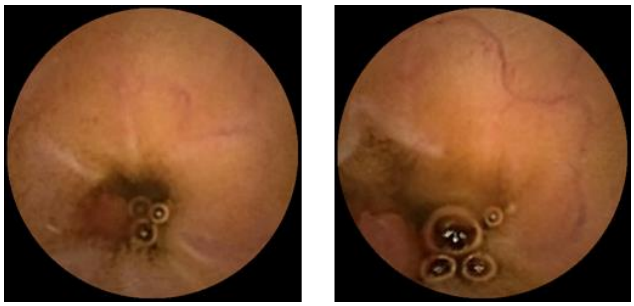
Table 6.2.3.1. Accuracies for all separated datasets and stages (*P*=0.9).

	Bleeding	Tumor	Normal (bleeding detection)	Normal (tumor detection)
Stage 1	98.81% (2167/2193)	79.73% (59/74)	91.98%(2672/2905)	
Stage 2	84.62% (19+3/26)	66.67% (8+2/15)	91.09% (2434/2672)	93.45% (2497/2672)
Total	99.82% (2189/2193)	93.24% (69/74)	83.79% (2434/2905)	85.96% (2497/2905)
			80.00%(2324/2905)	

Table 6.2.3.2. Accuracies for all separated datasets and stages (*P*=0.95).



(a)



(b)

Figure 6.2.3.1. Examples of oversighted bleeding and tumor images. (a) Bleeding (b) Tumor.

6.3 Summary

The two-stage scheme of proposed anomaly detection system for capsule endoscopy images has been presented in Chapter 4 and 5, and validation experiments for each part has been implemented to give evidence to support the hypothesis. With the goal of providing support for practical diagnosis, in this chapter, experiments are introduced to simulate the actual diagnosis situation. To make the system qualified to cope with unknown input image data, a training database, containing both normal dataset and different kinds of anomaly datasets, is constructed in advance. In the experiments, the proposed system has shown powerful capable to detect bleeding and tumor images from the unused testing data of a number of patients. Meanwhile, false-positive has also been well controlled, so that the lightening of diagnosis workload can be remarkably guaranteed.

Chapter 7

Conclusion

This dissertation has presented a study on anomaly image detection for capsule endoscopy images. To overcome the drawbacks of previous approaches and realize a more feasible anomaly detection for diagnosis support, the proposed system was developed with a two-stage scheme including anomaly degree evaluation and lesion oversight prevention. The first stage of anomaly degree evaluation was proposed to perform a role to distinguish image data with higher anomaly degree from those with lower degree. Lesions and residues that suggest to high anomaly degree are expected to be detected. The second stage of lesion oversight prevention was designed to prevent oversighted lesions that are easily oversighted in the first stage.

The anomaly degree evaluation made use of Higher-order Local Auto Correlation (HLAC) feature and focused on image processing methods. Subspace method was adopted to measure the anomaly degree and distinct normal images of clear intestinal wall and lumen from anomaly images of lesions and foreign matters (residue). In addition, the non-linear conversion for HSV color space and image segmentation have been further proposed to enhance the description of lesion and normal regions. In experiments, individual evaluation conducted on two patients' data reached 95.3% in sensitivity, and 7.5 percentage points improvement in

comparison with the initial proposed method. Afterwards, experiments of multi-patient evaluation, which involved different patients' data in training phase and testing phase, also confirmed the advance when compared to the approach without the non-linear conversion and feature extraction upon the image segmentation, according to the results denoted by both the numerical accuracies and ROC curves.

The second stage of lesion oversight prevention introduced a geometric image feature called LCE-HLAC (Local Contrast Enhanced Higher-order Local Auto-Correlation) was proposed for feature extraction. To deal with feature extraction for various lesions, especially the lesion images with small or inconspicuous lesion regions, LCE-HLAC reinforced extraction for detailed color, texture and shape features by enclose local contrast into the ordinary HLAC, so that to present the boundary and texture characteristics effectively. To demonstrate the efficacy of LCE-HLAC, a dataset of capsule endoscopy images collected from 28 patients' cases were employed. In contrast to methods based on HLAC and RIU-LBP features, the experimental results demonstrated that LCE-HLAC feature extraction owns superiority for the objective detection tasks. As denoted in statistical results, the proposed method had the best performance (98.54% accuracy) for bleeding detection and (92.44% accuracy) tumor detection.

Feasibility validation of the proposed anomaly detection system followed lastly. In the experiments throughout the whole system, training database including both normal and anomaly datasets were prepared for both stages in advance. Results shown that the system was powerful to deal with unused input testing data, in the situation similar to the practical diagnosis procedure. By the proposed system, in the case of P=0.95 (the true negative rate for training image is 0.95), 99.82% of bleeding images and 93.24% of tumor images were successfully detected, while 80% of normal images were corrected identified. Both the remarkable sensitivity for anomaly images and specificity for normal images were regarded as the insurance of practical feasibility of the proposed anomaly detection system.

The proposed anomaly detection system has been demonstrated satisfying the essential requirements for detecting various anomalies without defining or restraining the objective anomalies within certain categories. Moreover, the system has enabled detection for images containing other foreign matters such as residue. These measures all compensated for the lack of solution to foreign matters in the state-of-the-art researches, and make the system more equipped and feasible for practical application.

On the basis of the proposed anomaly detection system, many other extensional diagnosis support technologies for capsule endoscopy are believed possible to be realized and become meaningful means to support the diagnosis. For example, the anomaly degree evaluation can be directly used to mark the highly-suspected and low-suspected video segments, so that it is able to suggest the priority according to which doctor's examination should be conducted. Since capsule endoscopy examination is usually introduced following the examination results of blood test, radiology imaging and conventional endoscopies, settings for training database and the both classifiers in both stages can be appreciate adjusted in the light of doctor's expectation and demand based on the foregone results. The proposed anomaly detection system can be also utilized as a novel measure for automatic video display control. Video segments with high anomaly degree can have lower frame rate, while those with low anomaly degree can proceed in higher frame rate. Such kind of anomaly-detection-based control may provide more reliable and effective assistance than that only built on evaluating the change among the successive images.

Anomaly-detection-based control is possible to make contribution in wider domain beyond the diagnosis supporting technologies. It may also make contribution to the capsule endoscopy hardware. For instance, if real-time identification for the contents captured by capsule endoscopy could be realized, it can be utilized as the foundation to control the capturing framerate and balance of power supply. Moreover, if self-propelled capsule (which is also being researched worldwide) is realized, anomaly-detection-based control could even help to adjust the capturing angle and distance from the target.

Last but not the least, as many studies on other applied disciplines, the proposed anomaly detection system for capsule endoscopy, anyway, still needs plenty of validation to meet full requirements of clinical application. Therefore, it is strongly expected that more researchers can be inspired, so that the anomaly detection based diagnosis support can be further developed and become an indispensable parts of capsule endoscopy technology someday.

References

[1] International Agency for Research on Cancer, World Cancer Factsheet 2012, 2013.

[2] National Cancer Center (Japan), Prediction of cancer incidence number and death number 2015. (Japanese only)

[3] Japanese Association of Clinical Cancer Centers, Cancer Survival Rates at Japanese Association of Clinical Cancer Centers 2004. (Japanese only)

[4] Cyllene R. Morris, Cody N. Ramirez, Sara N. Cook et al., Cancer Stage at Diagnosis, California Cancer Registry, Jun. 2013.

[5] University of Rochester Medical Center, "Digestive diagnostic procedures," http://www.urmc.rochester.edu/encyclopedia/content.aspx?ContentTypeID=85&ContentID=P00364. Accessed 2 Jan. 2016.

[6] D. Xynopoulos, A. A. Mihas, E. Paraskevas, D. Dimitroulopoulos, D. M. Heuman and A. A. Mihas, "Small bowel tumors," Annals of Gastroenterology, 15: 18-35, 2002.

[7] Masaki Taruishi, Yusuke Saito, Kenichiro Kozawa et al., "Small intestine inspection method: small intestine X-ray inspection," Stomach and Intestine (Igaku-Shoin), 43: (4), 2008. (Japanese only)

[8] Khursheed N. Jeejeebhoy, "Short bowel syndrome: a nutritional and medical approach," Canadian Medical Association Journal, 166 (10): 1297-1302, May 14, 2002.

[9] J-F. Rey, G. Gay, A. Kruse and R. Lambert, “ESGE guideline for video capsule endoscopy,” Endoscopy, 36: 656-658, 2004.

[10] G.Iddan, Gavriel Meron, Arkady Glukhovsky and Paul Swain, “Wireless capsule endoscopy,” Nature, 405:417, May. 2000.

[11] Piotr R. Slawinski, Keith L. Obstein and Pietro Valdastri, “Emerging issues and future developments in capsule endoscopy”, Techniques in Gastrointestinal Endoscopy, 17(1): 40-46, Jan. 2015.

[12] Akiko Shiotani, Keisuke Honda, Makiko Kawakami et al., “Five-year single center experience of small bowel capsule endoscopy,” Kawasaki Medical Journal, 38(3): 97-105, 2012.(Japanese only)

[13] http://www.nomudake.com/index.html, Accessed 3 Jan. 2016.

[14] Martin Keuchel, Friedrich Hagenmuller and Hisao Tajiri, Video Capsule Endoscopy: A Reference Guide and Atlas, Springer-Verlag Berlin Heidelberg, 2014.

[15] Jonathan A. Leighton, Stuart L. Triester, Virender K. Sharma, “Capsule endoscopy: a meta-analysis for use with obscure gastrointestinal bleeding and crohn's disease,” Gastrointestinal Endoscopy Clinics of North America, 16 (2): 229-250, Apr. 2006.

[16] A. K. Hara, J. A. Leighton, R. I. Heigh, V. K. Sharma, A. C. Silva, G. De Petris, J. G. Hentz and D. E. Fleischer, “Crohn disease of the small bowel: preliminary comparison among CT enterography, capsule endoscopy, small-bowel follow-through, and ileoscopy,” Radiology, 238: 128-134, 2006.

[17] G. A. Doherty, A. C. Moss and A. S. Cheifetz, “Capsule endoscopy for small-bowel evaluation in Crohn's disease,” Gastrointestinal Endoscopy, 74: 167-175, 2011.

[18] M. Delvaux and G. Gay, “Capsule endoscopy: techniques and indications,” Best Practice & Research Clinical Gastroenterology, 22: 813-837, 2008.

[19] Y. Yagi, H. Vu, T. Echigo, R. Sagawa, K. Yagi, M. Shiba, K. Higuchi and T. Arakawa, “A diagnosis support system for capsule endoscopy,” Inflammopharmacology, 15: 78-83, 2007.

[20] T. Gan, J. C. Wu, N. N. Rao, T. Chen and B. Liu, “A feasibility trial of computer-aided diagnosis for enteric lesions in capsule endoscopy,” World Journal of Gastroenterology, 14: 6929-6935, 2008.

[21] Cristiano Spada, Maria Elena Riccioni and Guido Costamagna, “RAPID® access real-time device and rapid access software: new tools in the

armamentarium of capsule endoscopy," Expert Review of Medical Devices, 4:4, 2007.

[22] Koulaouzidis A., Smirnidis A., Douglas S. et al, "QuickView in small-bowel capsule endoscopy is useful in certain clinical settings, but QuickView with Blue Mode is of no additional benefit," European Journal of Gastroenterology & Hepatology. 24(9):1099-1104, Sep. 2012.

[23] C. Signorelli, F. Villa, E. Rondonotti, C. Abbiati, G. Beccari and R. de Franchis, "Sensitivity and specificity of the suspected blood identification system in video capsule endoscopy," Endoscopy, 37: 1170-1173, 2005.

[24] http://www.olympus.co.jp/jp/news/2013a/nr130204capsulej.jsp, Accessed 3 Jan. 2016.

[25] http://www.intromedic.com/eng/sub_company4_2.html, Accessed 3 Jan. 2016.

[26] Hai Vu, Tomio Echigo, Ryusuke Sagawa et al., "Controlling the display of capsule endoscopic video for diagnostic assistance," IEICE Transactions on Information and Systems, E92-D (3), Mar. 2009.

[27] Hai Vu, Tomio Echigo, Ryusuke Sagawa et al., "Adaptive control of video display for diagnostic assistance by analysis of capsule endoscopic images," Proceeding of the 18th International Conference on Pattern Recognition, 3:980-983, Aug. 2006.

[28] Hai Vu, Tomio Echigo, Ryusuke Sagawa et al., "Evaluating the control of the adaptive display rate for video endoscopy diagnosis," Proceeding of the 2008 IEEE International Conference on Robotics and Biomimetics, 14-17, Dec. 2008.

[29] Min Kook Choi, Hyun Gyu Lee and Ryan You et al, "Motion analysis for duplicate frame removal in wireless capsule endoscopy video," International Journal of Computer, Electrical, Automation, Control and Information Engineering, 4:2, 2010.

[30] M. K. Bashar, K. Mori, Y. Suenaga et al., "Detecting informative frames from wireless capsule endoscopic video using color and texture features," Medical Image Computing and Computer-Assisted Intervention, 11(2): 603-610, 2008.

[31] Hai Vu, Tomio Echigo, Ryusuke Sagawa et al., "Detection of contractions in adaptive transit time of the small bowel from wireless capsule endoscopy videos," Computers in Biology and Medicine, 39:16-26, 2009.

[32] Hai Vu, Yasushi Yagi and Tomio Echigo, "Color analysis for segmenting digestive organs in VCE," Proceeding of 20th International Conference on

Pattern Recognition, 2468-2471, 2010.

[33] Peter M. A. van Ooijen, Joost Dorgelo, Felix Zijlstra et al, “Detection, visualization and evaluation of anomalous coronary anatomy on 16-slice multidetector-row CT,” European Radiology, 14(12): 2163-2171, Dec. 2004.

[34] Yongbum Lee, T. Hara, H. Fujita et al., “Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique,” IEEE Transactions on Medical Imaging, 20(7): 595-604, Jul. 2001.

[35] Binsheng Zhao, “Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm,” Journal of Applied Clinical Medical Physics, 4(3), 2003.

[36] Charles M. Laymon and James Edwin Bousher, “Anomaly detection and artifact recovery in pet attenuation-correction images using the likelihood function,” IEEE Journal of Selected Topics in Signal Processing, 7(1), Feb. 2013.

[37] Edward J. Kendall, Michael G. Barnett and Krista Chytyk-Praznik, “Automatic detection of anomalies in screening mammograms,” BMC Medical Imaging, Dec. 13, 2013.

[38] Karen Drukker, Maryellen L. Giger, Karla Horsch et al., “Computerized lesion detection on breast ultrasound,” Medical Physics, 29(7): 1438-1446, Aug. 2002.

[39] P. S. Hiremath and Jyothi R. Tegnoor, “Automatic detection of follicles in ultrasound images of ovaries using edge based method,” IJCA Special Issue on RTIPPR (2):120–125, 2010.

[40] Yudai Yamazaki, Hirokazu Nosato, Masaya Iwata et al., “Anomaly detection based on density estimation of normal data in cone-restricted subspace”, IPSJ Transactions on Mathematical Modeling and its Applications, 8(1): 28-37, 2015.(Japanese only)

[41] Jia Qu, Hirokazu Nosato, Hidenori Sakanashi et al., “Computational cancer detection of pathological images based on an optimization method for color-index local auto-correlation feature extraction,” Proceeding of IEEE 2014 11th International Symposium on Biomedical Imaging, 822-825, Apr. 2014.

[42] Demir, Cigdem and Bulent Yener, “Automated cancer diagnosis based on histopathological images: a systematic survey,” Rensselaer Polytechnic Institute Tech. TR-05-09, 2005.

[43] Arthur Petrosian, Heang-Ping Chan and Mark A. Helvie, “Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture

analysis," Physics in Medicine and Biology, 39:2273-2288, 1994.

[44] Destounis S. V., DiNitto P., Logan-Young W. et al., "Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? Initial experience," Radiology. 232:578–584, 2004.

[45] B. Penna, T. Tillo, M. Grangetto, E. Magli and G. Olmo, "A technique for blood detection in wireless capsule endoscopy images," Proceeding of 2009 European Signal Processing Conference, 1865-1868, 2009.

[46] P. Y. Lau, "Detection of bleeding patterns in WCE video using multiple features," Proceeding of International Conference of IEEE Engineering in Medicine and Biology Society, 5601-5604, 2007.

[47] A. A. Al-Rahayfeh and A. A. Abuzneid, "Detection of bleeding in wireless capsule endoscopy images using range ratio color," International Journal of Multimedia and its Applications, 2.2: 1-10, 2010.

[48] Yingju Chen and Jeongkyu Lee, "A review of machine-vision-based analysis of wireless capsule endoscopy video," Diagnostic and Therapeutic Endoscopy, Vol. 2012, ID 418037, 2012.

[49] B. Li and M. Q. Meng, "Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments," Computers in Biology and Medicine, 39: 141-147, 2009.

[50] B. Li and M. Q. Meng, "Computer aided detection of bleeding regions for capsule endoscopy images," IEEE Transactions on Biomedical Engineering, 56: 1032-1039, 2009.

[51] Balathasan Giritharan, Xiaohui Yuan, Jianguo Liu et al., "Bleeding detection from capsule endoscopy videos," Proceeding of International Conferences of IEEE Engineering in Medicine and Biology Society, 4780-4783, 2008.

[52] D. J. Barbosa, J. Ramos and C. S. Lima, "Detection of small bowel tumors in capsule endoscopy frames using texture analysis based on the discrete wavelet transform," Proceeding of International Conferences of IEEE Engineering in Medicine and Biology Society, 3012-3015, 2008.

[53] Jinn-Yi Yeh, Tai-Hsi Wu, and Wei-Jun Tsai, "Bleeding and ulcer detection using wireless capsule endoscopy images," Journal of Software Engineering and Applications, 7: 422-432, 2014.

[54] T. Ojala, M. Pietika¨inen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, 29(1): 51-59, 1996.

[55] T. Ojala, Matti Pietikäinen, and Topi Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7): 971-987, Jul. 2002.

[56] B. Li and M. Q. Meng, “Texture analysis for ulcer detection in capsule endoscopy images,” Image and Vision Computing, 27: 1336-1342, 2009.

[57] B. Li and M. Q. Meng, “Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection,” IEEE Transactions on Information Technology in Biomedicine, 16: 323-329, 2012.

[58] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras and M. Tzivras, “Computer-aided tumor detection in endoscopic video using color wavelet features,” IEEE Transactions on Information Technology in Biomedicine, 7: 141-152, 2003.

[59] S. Hwang and M. E. Celebi, “Polyp detection in wireless capsule endoscopy videos based on image segmentation and geometric feature,” Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing, 678- 681, 2010.

[60] J. Lee, J. H. Oh, S. K. Shah, X. Yuan and S. J. Tang, “Automatic classification of digestive organs in wireless capsule endoscopy videos,” Proceeding of ACM symposium on Applied computing, 1041-1045, 2007.

[61] M. T. Coimbra and J. P. S. Cunha, “MPEG-7 visual descriptors-contributions for automated feature extraction in capsule endoscopy,” IEEE Transactions on Circuits and Systems for Video Technology, 16: 628-637, 2006.

[62] P. Szczypiński, A. Klepaczko, M. Pazurek and P. Daniel, “Texture and color based image segmentation and pathology detection in capsule endoscopy videos,” Computer Methods and Programs in Biomedicine, 113: 396-411, 2014.

[63] P. Szczypiński, M. Strzelecki, A. Materka and A. Klepaczko, “Mazda—a software package for image texture analysis,” Computer Methods and Programs in Biomedicine, 94: 66-76, 2009.

[64] N. Otsu and T. Kurita, “A new scheme for practical flexible and intelligent vision systems,” Proceeding of IAPR Workshop on Computer Vision, 431-435, 1988.

[65] T. Kobayashi, T. Hosaka, S. Mimura, T. Hayashi and N. Otsu, “HLAC approach to automatic object counting,” Proceeding of Bio-inspired, Learning and Intelligent Systems for Security, 40-45, 2008.

[66] J. Qu, H. Nosato, H. Sakanashi, K. Terai and N. Hiruta, “Cancer detection from pathological images using higher-order local autocorrelation feature,” Proceeding of IEEE International Conference on Signal Processing, 1198-1201, 2012.

[67] Hirokazu Nosato, Hidenori Sakanashi, Masahiro Murakawa et al., “Histopathological diagnostic support technology using higher-order local autocorrelation features,” Proceeding of IEEE Symposium on Bio-inspired Learning and Intelligent Systems for Security, 61-65, 2009.

[68] Takio Kurita and Satoru Hayamizu, “Gesture recognition using HLAC features of PARCOR images”, IEICE Transactions on Information and Systems E86-D(4): 719-726, 2003.

[69] S. Mimura, K. Itoh, T. Kobayashi, T. Takigawa, A. Tajima, A. Sawamura and N. Otsu, “The cow gait recognition using CHLAC,” Proceeding of Bio-inspired, Learning and Intelligent Systems for Security, 56-57, 2008.

[70] H. Nosato, T. Kurihara, H. Sakanashi, M. Murakawa, T. Kobayashi, T. Furuya, T. Higuchi, N. Otsu, K. Terai and N. Hiruta, “An extended method of higher-order local autocorrelation feature extraction for classification of histopathological images,” IPSJ Transactions on Computer Vision and Applications, 3:211-221, 2011.

[71] C. H. Li and P. C. Yuen, “Regularized color clustering in medical image database,” IEEE Transactions on Medical Imaging, 19: 1150-1155, 2000.

[72] M. Mete, Color analysis, In Delineation of malignant areas in histological images of head-neck cancer (28-31), University of Arkansas at Little Rock, 2008.

[73] E. Hu, H. Nosato, H. Sakanashi and M. Masahiro, “Anomaly detection for capsule endoscopy images using higher-order local auto correlation features,” Proceeding of IEEE International Conference on Systems, Man, and Cybernetics, 2289-2293, 2012.

[74] E. Hu, H. Nosato, H. Sakanashi and M. Masahiro, “A modified anomaly detection method for capsule endoscopy images using non-linear color conversion and higher-order local auto-correlation (HLAC),” Proceeding of International Conference of IEEE Engineering in Medicine and Biology Society, 5477-5480, 2013.

[75] V. N. Vapnik, Statislical Learning Theory, NewYork: Wiley-Interscience, 1998.

[76] T. Onoda, “Summary of support vector machine,” Operations Research, 225-

230, 2001.

[77] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, The kernel trick, In Numerical Recipes: The Art of Scientific Computing (3rd edition) (889-892), New York: Cambridge University Press, 2007.

Ethical approvals

This study has been carried out in compliance with the procedure for handling ergonomics experiments in the National Institute of Advanced Industrial Science and Technology (AIST), and in response to the approval of the Committee on Ergonomic Experiments in AIST. The experiments used the endoscopic image data, which were obtained with the approval of the Ethics Committee of the Faculty of Medicine, Toho University, and with the written informed consent from all patients. All of the endoscopic image data were provided to AIST after unlinkable anonymization in order to fully protect the patient's personal information.

List of publications

1. Erzhong Hu, Hirokazu Nosato, Hidenori Sakanashi, Masahiro Murakawa , “Preliminary Study toward Anomaly Detection in Capsule Endoscopy Images based on Higher-order Local Auto Correlation” ,情報処理学会第 74 回全国大会論文集, 名古屋工業大学, 2012 年 3 月.

2. Erzhong Hu, Hirokazu Nosato, Hidenori Sakanashi and Masahiro Murakawa, “Anomaly Detection for Capsule Endoscopy Images Using Higher-order Local Auto Correlation Features”, Proc. of 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2012), pp.2289-2293, Oct. 2012.

3. Erzhong Hu, Hirokazu Nosato, Hidenori Sakanashi and Masahiro Murakawa, “A modified anomaly detection method for capsule endoscopy images using non-linear color conversion and higher-order local auto-correlation (HLAC) ”, Proc. of IEEE Int. Conf. Engineering in Medicine and Biology Society (EMBC 2013), pp.5477-5480, Jul. 2013.

4. Erzhong Hu, Hidenori Sakanashi, Hirokazu Nosato, Eiichi Takahashi, Yasuo Suzuki, Ken Takeuchi, Hiroshi Aoki and Masahiro Murakawa, “Bleeding and tumor detection for capsule endoscopy images using an improved geometric feature,” Journal of Medical and Biological Engineering. (Accepted)

Acknowledgements

First of all, I would like to express my deepest appreciation to Prof. Masahiro Murakawa (Team Leader of Artificial Intelligence Applications Research Team, AIST & Professor of the University of Tsukuba) and Prof. Hidenori Sakanashi (Senior Researcher of Artificial Intelligence Applications Research Team, AIST & Associate Professor of the University of Tsukuba). This research is carried out under the direct of Prof. Masahiro Murakawa and Prof. Hidenori Sakanashi. After I came to Japan, they have made so much effort to push forward and encourage my study, and have done everything to help me to become more skilled and adaptable. Meanwhile, they have set an example for me and their extraordinary professional dedication has infected me so profoundly.

I would also show my gratitude to Prof. Koichi Mizutani (Professor of the University of Tsukuba). As my supervisor in University of Tsukuba, he is always keeping concerned with my coursework, paperwork and daily affairs all the time. During the period of finishing this dissertation, Prof. Mizutani has also given me a lot of advice that is very significant and helpful to the work.

Special thanks are expressed to Prof. Hirokazu Nosato (Senior Researcher of Artificial Intelligence Applications Research Team, AIST & Associate Professor of Toho University), who has helped me throughout my research by introducing techniques, making suggestions, and, acquiring and administrating the capsule

endoscopy data. Without Prof. Hirokazu Nosato's kindhearted help, this research would not have progressed so far.

My gratitude also goes to Yasuo Suzuki M.D., Ken Takeuchi M.D. and Hiroshi Aoki M.D. at Toho Univ. Sakura Medical Center. The respectful medical doctors have taken a lot of time out from their busy schedules, to collect and provide the capsule endoscopy data, and share the concerns to the research. All of their work has provided convenience for the progress of the research.

I also would like to thank the dissertation committee and professors, including Prof. Masahiro Murakawa, Prof. Hidenori Sakanashi, Prof. Koichi Mizutani, Prof. Kiyoyuki Chinzei (Deputy Director of Health Research Institute, AIST & Professor of the University of Tokyo), and Prof. Yasuyo Kita (Senior Researcher of Intelligent Media Research Group, AIST & Professor of the University of Tsukuba). Thanks to these professors for reviewing my dissertation and providing precious comments.

I would also like to appreciate all the other researchers in the Artificial Intelligence Applications Research Team including Dr. Eiichi Takahashi, Dr. Masaya Iwata, Dr. Yuji Kasai, and Dr. Jun Ogata, who have spent their valuable time to share opinions to my research. Besides, in the same research team, Ms. Ayako Suzuki and Ms. Eriko Nishino have helped me a lot with the office procedures, and thus, I also would like to express my deep gratitude to them.

My appreciation is also expressed to all past and current members of the team, including but not limited to Dr. Jiaxing Ye, Yudai Yamazaki, Dr. Hirotaka Nosato, Dr. Fumio Sakabe. Their kindhearted help and precious advice have greatly promoted my research.

I would like to thank the University of Tsukuba and Ministry of Education, Culture, Sports, Science and Technology (MEXT) scholarship, for accepting me as an international student and providing essential financial support during my study in Japan.

Lastly, I would like to acknowledge my family for all their love and encouragement. For my parents who raised me with a love of science and supported me in my pursuits. And most of all for my loving, encouraging, and patient wife Jia Qu, whose faithful support during my research is so appreciated.