Check for updates

A survey of feature extraction and fusion of deep learning for detection of abnormalities in video endoscopy of gastrointestinal-tract

Hussam Ali[1] · Muhammad Sharif[1] · Mussarat Yasmin[1] · Mubashir Husain Rehmani[2] · Farhan Riaz[3]

Abstract

A standard screening procedure involves video endoscopy of the Gastrointestinal tract. It is a less invasive method which is practiced for early diagnosis of gastric diseases. Manual inspection of a large number of gastric frames is an exhaustive, time-consuming task, and requires expertise. Conversely, several computer-aided diagnosis systems have been proposed by researchers to cope with the dilemma of manual inspection of the massive volume of frames. This article gives an overview of different available alternatives for automated inspection, detection, and classification of various GI abnormalities. Also, this work elaborates techniques associated with content-based image retrieval and automated systems for summarizing endoscopic procedures. In this survey, we perform a comprehensive review of feature extraction techniques and deep learning methods which were specifically developed for automatic analysis of endoscopic videos. In addition, we categorize features extraction techniques according to image processing domains and further we classify them based on their visual descriptions. We also review hybrid feature extraction techniques which are developed by the fusion of different kind of basic descriptors. Moreover, this survey covers various endoscopy data-sets available for the bench-marking of vision based algorithms. On the basis of literature, we explain emerging trends in computerized analysis of endoscopy. We also survey important issues, challenges, and future research directions to the development of computer-assisted systems for detection of maladies and interactive surgery in the GI tract.

Keywords Convolutional neural network (CNN) · Deep learning · Feature extraction · Gastrointestinal tract · Gastric cancer · Video endoscopy · Classification

✉ Hussam Ali
  hussamalics@gmail.com
  https://sites.google.com/site/hussamalics/

Extended author information available on the last page of the article

1 Introduction

The gastric burden is increasing with the fact of growing population and due to the formation of cancer in the gastrointestinal tract (GI) around the world. Every year nearly 0.7 million cases of specifically gastric cancer are reported (Siegel et al. 2015). It is estimated that cancer cases in both sexes are 24,590 deaths are estimated in 2015 and 10,720 in the United States alone. The worst conditions can be observed in developing countries (e.g. the Middle East and the Asian countries) (Swannell 2010; Organization et al. 2015). The normal clinical practice includes the intestinal biopsy (tissues sample of the mucosa is taken) of GI tract. Which are then analyzed by experts (under microscope), to see if there are any cancerous or abnormal cells exist. This is an invasive method for detection of gastric abnormalities and it requires high-level of expertise (Qi 2008). On the other hand, endoscopy is a less invasive method for screening GI tract (Kainuma et al. 2015). An endoscope is a flexible tube with a mounted camera, light source, and an accessory channel (Pennazio 2006). Moreover, an accessory channel can be used for cleansing of GI tract or inserting medical instruments. Therefore, the endoscope can also be used for the intestinal biopsy (Wallace and Keisslich 2010).

The GI tract can be categorized into several parts, starting from upper stomach parts GI tract have esophagus, stomach in middle, and duodenum as ending of stomach (upper GI tract), the jejunum, ileum (small-bowel), ending at the colon, and rectum (Carpi et al. 2011; Filip et al. 2011). Therefore, the endoscopy procedures refer to different names according to the target area of GI tract[1] e.g., for esophagus referred (esophagoscopy), area of stomach and duodenum (gastroscopy), rectum and sigmoid colon (proctoscopy), sigmoid colon (sigmoidoscopy), colon (colonoscopy) for whole GI tract (laparoscopy).

The endoscopic procedure helps the physician for the detection of gastric abnormalities in their early stages. Timely detection of chronic diseases can be cured with proper treatments. Thus, the screening process can be very useful for a substantial reduction in both, death-rate as well as the cost of treatment. Specifically, the deaths occur due to different gastrointestinal cancers, which can be cured if cancer was detected in its pre-malignant stage (Hamashima et al. 2015). Still, video endoscopy is a painful procedure, it also requires both time and expertise (Society 2016).

In contrast with the wired endoscopy, the wireless capsule endoscopy (WCE) is a painless tool than traditionally used white light video endoscopy (VE) for examining the internal cavity of the human body (Gastelum et al. 2015; Kim et al. 2005). Normally, a VE composed of a light source, a charged couple camera (CCD), and a video monitor (which is used to view the output of the endoscope).

1.1 Abnormalities in gastrointestinal tract

The GI tract is a crucial part of human body, it refers to stomach, small intestine and whole digestive system. The GI tract can be divided into an esophagus, stomach, small bowel and colon (Chu et al. 2015). Furthermore, the stomach has its own parts as the upper stomach, middle stomach and lower stomach (Miyahara et al. 2007). A brief introduction of the parts of a GI tract has given in Table 1.

There are many clinical conditions, including basic symptoms and mature diseases found in the GI tract. Some of these abnormalities of the digestive system are listed below. Moreover, some of them can be easily detected through a normal endoscopic procedure. However, there

---

[1] http://www.cancer.net/navigating-cancer-care/diagnosing-cancer/tests-and-procedures/types-endoscopy.

Table 1 Anatomy of the gastrointestinal tract

Disease	Description
Esophagus	When we eat food, the chewed food goes into the esophagus. The esophagus is a tube-like organ that used to carry food to the stomach (Khashab et al. 2016). The entry point of food into the stomach is called gastroesophageal (GE) junction, which is located under the lungs (Owens and Appelman 2014a)
Stomach	The stomach is a towel-like organ that holds food (Eberlein et al. 2002). After the food arrived from esophageal the special enzymes are produced to digest the food into that stomach (Rogy and Bünger 2015; Roukos et al. 2002)
Small-bowel	Secretion of stomach juices and the food are mixed. Then, the food goes into the entrance of small-intestine called duodenum (Callacondo-Riva et al. 2009). The small-intestine is like a narrow tube of approximately 6 meters in length. Due to complexity and length of small intestine, the screening is a difficult and time-consuming task (Axon 2008)
Colon	The last portion of GI tract, and the most prone area for having tumors and polyps. The colon cancer is the third most common disease found in the Western countries (Chu et al. 2015; Kiesslich and Neurath 2007). The colon is consisting of several parts, ascending colon, transverse colon, descending colon and rectum (Hurlstone et al. 2004)

are some diseases in GI tract (e.g., cancer, polyps and ulcers) that do not exhibit visible signs until it approaches an advance stage (Liedlgruber and Supervisor 2011).

(1) *Cancer* There are many types of cancer [e.g., adenocarcinoma (Kelley and Duggan 2003; Cobrin et al. 2006; Ueo et al. 2013; Gholami et al. 2015), lymphoma (Namikawa et al. 2005; Choi 2014), GIST (Korngold 2011), carcinoid tumor (Owens and Appelman 2014b), squamous cell carcinoma (Callacondo-Riva et al. 2009; Takita et al. 2005), and small cell carcinoma (Report 1990; Owens and Appelman 2014b)]. However, most of the cancers are started from an unusual growth of cells. The older cells do not die and unusual growth of cells forms cancer. Some common symptoms of gastric cancer include early satiety, abdominal pain, nausea, vomiting, bloating, weight-loss, and anorexia (Rogy and Bünger 2015). It is worth mentioning here, that cancer can start in any area of the GI tract.

(2) *Polyps* Polyps are similar to cancer, it is also an unusual mucosal growth and typically these are benign (Kato et al. 2010). On the other hand, there could be diminutive colorectal polyps, which are very dangerous if left untreated. Polyps can grow in any area of GI tract similar to cancer. However, mostly found in the colon and small-bowel (Hazewinkel et al. 2013).

(3) *Ulcer* The ulcer is also referred to a disease caused by the acid that is produced by the stomach itself. In peptic ulcer, the gastric cells are damaged with gastric juices. Normally ulcer appears in the duodenum, small intestine or in gastric lining (Karargyris and Bourbakis 2009b; Mountford et al. 1980; Jensen et al. 2016).

(4) *Helicobacter pylori* Helicobacter pylori (H. pylori) bacterium can be found in many areas of GI tract. The H. pylori causes inflammation in the mucosal wall. Therefore, an infection caused by this bacterium leads to various chronic abnormalities (e.g., cancer, ulcer and inflammation) (Ishihara et al. 2016; Leodolter et al. 2015; Kelley and Duggan 2003).

(5) *Inflammation* Inflammation refers to the condition of gastric abnormalities, involving dyspepsia, chronic gastritis and acid reflux, are normally associated with the inflammation of gastric lining. However, the main causes of inflammation of the GI tract are H. Pylori and hookworms (Peljto et al. 2016; Seidel and Burdick 1998; Wu et al. 2016).

(6) *Celiac disease* Celiac disease is one of the most difficult to diagnose because of its large number of symptoms. This is an autoimmune disorder in the small intestine, the intolerance of gluten found in wheat (Ciaccio et al. 2010; Boschetto et al. 2015; Gschwandtner et al. 2010).

(7) *Crohn's Disease* Crohn's disease is caused due to inflammation of the lining of the gut. This is an intestinal inflammatory disease, it may also cause a severe abdominal pain. In some cases ileum (part of small intestine) is effected from this disease (Pennazio 2006; Eliakim 2004).

(8) *Bleeding* Bleeding is another abnormality that is normally found in GI tract while screening. It may be caused by different other pathological conditions such as cancer, Crohn's disease, hepatitis c or ulcer (Lewis 2003; Jensen et al. 2016; Schlag et al. 2015).

(9) *Barrett's esophagus* Barrett's esophagus is a disease, specifically, associated with the esophagus. In Barrett's esophagus, the mucosal wall is damaged due to acid reflux disease, also known as Gastroesophageal reflux disease (GERD). However, many CADx supportive systems are designed for the diagnoses of Barrett's esophagus (Dattamajumdar et al. 2001; Shin et al. 2016).

1.2 Motivation: the need of computer-aided diagnosis (CADx) Systems

The endoscopy has several benefits, although, it comes along with certain trade-offs such as a huge number of frames are generated (video recordings) from the screening procedure of GI tract. If we consider the endoscopy of an individual, it can take up to 45 minutes to 8 hours to complete the screening procedure and approximately more than 10 thousand endoscopy frames are produced, depending on the target GI area. The time taken by the endoscopic process depends on the target GI area and skills of the gastroenterologist. A point to note here is that all endoscopic frames are not useful to the gastroenterologist because most of the frames are redundant, and only a small number of images may have some abnormal tissues (Sainju et al. 2014). Therefore, rest of images that not contain any abnormality can be discarded by observing each frame (Lehmann et al. 1999).

Besides, it is a difficult and lengthy process for doctors to observe each frame separately. Then, the abnormal frames can be easily overlooked by the medical experts. Therefore, the clinical practitioners demand such systems, that can automatically discover potential malignancies by analyzing the endoscopic frames.

Computer-aided diagnosing (CADx) systems are machine-vision based systems used for helping doctors in the analysis of endoscopic imaging data. In a typical CADx system, a decision is made on every frame based upon various characteristics (features), which were extracted from the frame. However, some systems are only a sub-part of a whole CADx system, the output of these systems is an image instead of decision (e.g., image enhancement and image compression) (Khan and Wahid 2014; Gu et al. 2015; Turcza and Duplaga 2011). Only for the purpose of the abnormality detection, a CADx system takes an image as input and returns a decision based on its characteristics, whether the image has a normal mucosal structure or some deformation, like (e.g., ulcer, blood, cancer, and polyps) (Albisser 2015). List of abbreviations and Acronyms is provided in Table 2.

The task of decision making in CADx involves many intermediate steps, by starting from the acquisition of images, and applying several pre-processing procedures (e.g., feature extraction, image segmentation, etc). There are several benefits of developing CADx systems for GI diseases. Ultimately, the patients, clinical practitioners, and medical students will gain assistance as follows.

Table 2 List of abbreviations and acronyms

Acronyms	Definition
3C	Color Curvelet Covariance
ACWE	Active Without Edges Model
ANN	Artificial Neural Network
BEEMD	Bi-dimensional Ensemble Empirical Mode Decomposition
BOF	Bag of Features
CADx	Computer-aided Diagnosis system
CBIR	Content-based Image Retrieval
CCD	Charge-Coupled Device
CD	Celiac Disease
CH	Chromoendoscopy
CLE	Confocal Laser Endoscopy
CMOS	Complementary Metal-Oxide Semiconductor
CNN	Convolutional Neural Network
DFTs	Discrete Fourier Transform
DT-DWT	Dual Tree- Discrete Wavelet Transform
DWFT	Discrete Wavelet Frame Transform
ELM	Extreme Learning Machine
EM	Endomicroscopy
GE	Gastroesophageal
GERD	Gastroesophageal Reflux Disease
GF-LBP	Gaussian Filtered- Local Binary Patterns
GIST	Gastrointestinal Stromal Tumors
GI tract	Gastrointestinal-tract
GLCM	Gray Level Co-occurrence Matrices
H.Pylori	Helicobacter Pylori
HDVE	High-definition Video Endoscopy
HSV	Hue Saturation Value
VLAD	Vector of Locally Aggregated Descriptors
IMF	Intrinsic Mode Functions
JLDCP	Jumping Local Difference Count Pattern
LBP	Local Binary Patterns
ME	Zoom/ Magnifying Endoscopy
SDMD	Symmetric Dense Micro-block Difference
NBI	Narrow Band Imaging
OC-LBP	Opponent Color-Local Binary Pattern
OCT	Optical Coherence Tomography
PHOG	Pyramidal Histogram oriented gradient
PNN	Probabilistic Neural Network
RGB	Red Green Blue
ROI	Region of Interest
RCLPB	Refined Completed LBP
SIFT	Scale Invariant Features Transform

Table 2 continued

Acronyms	Definition
SVM	Support Vector Machine
U-LBP	Uniform Local Binary Patterns
VE	Video Endoscopy
WCC	Wavelet Cross Co-occurrence
WCE	Wireless Capsule Endoscopy

- The endoscopic inspection time will be reduced for the gastroenterologist.
- Low cost of treatment, because of detection of cancer in its early stages.
- An increase in accuracy of a physician in predicting the stage of the gastric malignancies.
- CADx system can also be used for training the clinical staff and medical student without a need of an expert.

1.3 Comparison with the existing surveys

There are already existing surveys on different aspects of gastroenterology disease detection and there brief summarizes are presented in Table 3. Cho et al. (2011) have discussed various advancements in the optical technology of video endoscopy. Especially variation in NBI and CLE and provides heir trade-offs. The bestowed work by Song and Ang (2014) gives a brief description of various imaging modalities. In Beg and Ragunath (2015), also provides a good overview of the options that are available for the gastric endoscopy. Moreover, their work describes, how various endoscopy advancements can be useful in the detection different types of malignancies, and how these malignancies can be distinguished from the other pathological conditions ? A similar work had done by Leggett and Iyer (2015), where characteristics of various endoscopy technologies were compared and described in a great detail.

Liedlgruber and Uhl in Liedlgruber and Uhl (2011) provide statistics about the quantity of the work that has been done in the field of automated detection of gastrointestinal diseases using computer vision techniques. Their work provides only a brief overview of feature extraction techniques. They have the divided the endoscopic techniques on the basis of the level of mucosal intervention. In the same way, Keuchel et al. (2015) offers a review of quantitative measures of different pathological conditions and how they are helping in quantifying different clinical procedures. Moreover, a survey of numerous CADx developed for GI tract was presented. Their work provides a good overview of different aspects of GI tract (e.g., cleansing and pH level of GI tract). Also, they reviewed different diseases and abnormalities found in the GI tract. Furthermore, their work also provides a comparison of CADx on different parameters of descriptors, classifiers and data-sets used for testing CADx. Their work also describes the methods to measure the abnormalities, such as polyp, ulcer, inflammation, and cancer. Different types of CADx system are reviewed in Ogoveanu et al. (2015), with respect to their application in the investigation of GI tract with different gastrointestinal pathology. However, their work lacks a comparative view of different features extraction techniques. In above-mentioned surveys, there no single survey has tendency specifically toward feature extraction. In our work, we focus on the features extraction techniques that have been used in CADx systems which are specially developed for detection of gastric abnormalities.

Table 3 Comparison with the existing surveys on endoscopy and CADx. The surveys discussed (✓) or not included (×) the aspect of imaging technologies (I), segmentation (S), feature extraction methods (F), classification (C), accuracy measures (A), and data-sets (D)

Refs.	Year	Description	I	S	F	C	A	D
Cho et al. (2011)	2011	Gives an overview of gastrointestinal screening modalities with their position, FOV and accuracy parameters. Provides a good understanding of frames, their color and noise in endoscopic frames	✓	×	×	×	×	×
Song and Ang (2014)	2014	The summaries of options that are available for the early gastric cancer screening from traditional to image-enhanced endoscopy. All modalities are elaborated with great details and with their use in visualizing the micro-vascular structures	✓	×	×	×	×	×
Beg and Ragunath (2015)	2015	The available imaging modalities are compared on the basis of cost, difficulty in training and accuracy in lesion detection. With future directions, their work also gives the scientific justification of using (or not using) certain imaging technologies for the inspection of GI tract	✓	×	×	×	×	×
Leggett and Iyer (2015)	2016	The evolution of technologies for the visualization of GI tract was discussed in a great detail. Moreover, this study provides a great insight into the selection of imaging technology, which particularly best perform for detection of certain abnormalities	✓	×	×	×	×	×
Liedlgruber and Uhl (2011)	2011	Gives a brief overview of the endoscopic technologies and CADx systems. Specifically, they do not discuss features extraction techniques in a great detail. Also, their work do not described the segmentation and feature selection related to CADx	✓	×	✓	✓	✓	✓
Keuchel et al. (2015)	2015	The quantitative measures used in different GI-tract scenarios and occurrence of events in WCE are discussed in details. Moreover, an overview of CADx systems based on WCE and non-visual WCE is provided	×	×	×	×	✓	×
Ogoveanu et al. (2015)	2015	Offers an overview of profthe computer-based abnormality detection systems by employing capsule endoscopy. Also, it gives the advantages and disadvantage of using of WCE	×	✓	×	✓	×	×
Hegenbart et al. (2015)	2015	They have summarized different computer vision-based techniques exist specifically developed for detection of the CD. Moreover, they discuss the prominent imaging modalities used for screening of the GI tract	✓	×	×	✓	✓	×
Proposed study		Our work will discuss all aspects used in various CADx systems for GI tract screening. Moreover, the features extraction, feature learning, segmentation, data-sets, and classification techniques that have been used in CADx	✓	✓	✓	✓	✓	✓

1.4 Contributions of this article

The most significant task involved in the decision making of a CADx is the selection of an appropriate features extraction technique. As, these features later used for the segmentation, classification, and retrieval of the images. However, other sub-tasks such as image preprocessing (Figueiredo et al. 2018) and image segmentation also have an important role in the extraction of image descriptors. The main contributions of this paper listed as follows:

– In this paper, we review various endoscopy options available for screening of GI tract.
– We review and discuss various types of CADx with potential applications.
– We present a meticulous survey on feature extraction methods and classify them accordingly to their visual description and domain.
– We discuss strengths and limitations of different features extraction techniques specific various pathological conditions.
– We surveyed deep learning based representation learning techniques.
– This study also highlights the trends, open issues, and emerging challenges. Moreover, we review different publicly available endoscopy data-sets for testing vision-based CADx systems.

1.5 Organization of the paper

The rest of the article is organized as follows: Sect. 2 gives the introduction to CADx and describes its various applications. An overview of advancements in imaging modalities is given in Sect. 2. Moreover, this section also includes a brief overview of the anatomy of GI tract and different abnormalities found in the endoscopic images of GI tract. Then, it highlights the validation and accuracy measures used to asses the performance of CADx systems. Furthermore, the existing features extraction techniques in are reviewed in Sect. 3, 4, and 5 combined with the methods have been developed for the automated diagnosis abnormalities in GI tract via endoscopy videos. Feature extraction techniques developed in the spatial domain are discussed in Sect. 3 then Sect. 4 describes features extraction methods developed in the frequency domain of image processing. The automated features learning methods are reviewed in Sect. 5. Section 6 discusses the diverse range of endoscopic images data-sets that are publicly available, challenges and trends. Finally, this paper is concluded in Sect. 7.

2 Computer-aided diagnosis (CADx) system for GI tract : an overview

CADx systems are developed for the automatic detection of gastric abnormalities from the endoscopy of the GI tract. In recent years, a number of CADx systems have been developed. However, every CADx system has its own limitations and advantages.

2.1 Basic architecture of a CADx system and possible outputs

A number of sub-tasks are involved in a basic CADx system. The architecture of a CADx system with different applications is depicted in Fig. 1.

(1) *Preprocessing of Endoscopy frames* In context of gastric diseases, first endoscopy frames are pre-processed by different image processing methods. Images acquired from
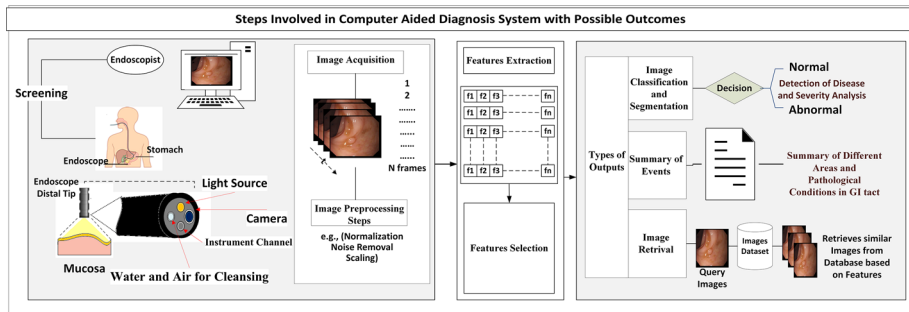
Fig. 1 The architecture of a computer aided diagnostic system

endoscopy normally suffer from different kinds of noises and variations such as, e.g., lens distortions, illumination invariance, scale invariance, rotation invariance, and specular highlights (Gueye et al. 2015; Tischendorf et al. 2010; Geng and Pahlavan 2015; Hafner et al. 2010b). Moreover, some other conditions like poor cleansing, bubbles, food presence, and instrument inclusion makes the automatic detection of lesions more challenging (Bejakovic et al. 2009). However, some of these issues can be treated with image pre-processing techniques. Therefore, the endoscopy frames were pre-processed by different technique according to the nature of the acquisition environment and noise. On the contrary, image pre-processing is also an important step in CADx of gastric diseases. The pre-processing step may involve frames normalization (Vieira et al. 2015), contrast enhancement (Song and Ang 2014), image compression (Khan and Wahid 2014), image scaling, image rotation, and color space transformation (Riaz et al. 2017). The image pre-processing is a crucial task, prior to features extraction, sometimes it includes the division of images into sub-images or removal of unnecessary frame's area (Alexandre et al. 2009).

After the image pre-processing, a compressed form of information has extracted, that are called features or descriptors. Then, the pixels of each image are represented by a feature vector (Vécsei et al. 2009). However, in some cases, after pre-processing, the region of interest (ROI) is selected and features are extracted from the segmented region (normally it is a lesion area). Further, these images are analyzed based on these extracted descriptors (Serpa-Andrade et al. 2016) . Even though, the lesions can also be segmented based on extracted features as illustrated in Riaz et al. (2013). Therefore, the lesion detection, retrieval, and classification of gastroenterology frames, tasks are performed based on the extracted features. In later sections, we have discussed the segmentation, features extraction, classification in detail.

(2) *Importance of Features Extraction* As mentioned earlier, endoscopic videos contain a large number of frames. However, these frames cannot be used directly for the task of classification and recognition due to computational limitations (curse of dimensionality) (Cong et al. 2016). There is a need to represent these images in a more compact form, while preserving their discrimination power. Sometimes features are used to store and retrieve images from database efficiently. Thus, features extraction can be used for data compression purposes (Bonnel et al. 2009). We have divided feature extraction techniques in three broad categories, features extraction method in spatial domain and frequency domain, these two categories are divided according to Liedlgruber and Uhl (2011). However, the third category includes both spatial and frequency domains methods, and requires multiple images to learn feature automatically.

Fig. 2 Types of CADx systems based on their outputs
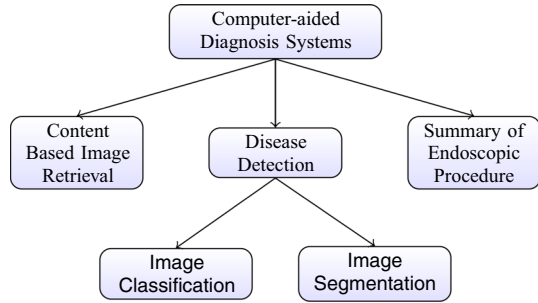
2.2 Potential outputs of a CADx system for GI tract

A number of systems have been developed for detection of abnormalities in the GI tract. We have categories them into three types, based on their respective outputs as described in Fig. 2.

(1) *Content-based Image Retrieval (CBIR)* In a content-based image retrieval system, a query image is matched with the images in a database, comparison is based of descriptors of images. Then, a image or set of images with a similar characteristic are fetched from the database. Moreover, the images' database is contracted by extracted features corresponding to all images (André et al. 2010). However, it seems to be a simple image retrieval system, however, it can be used for helping the gastroenterologist by finding images with the same pathological conditions from a whole endoscopy sequence. A medical expert need to select one or two abnormal images from whole sequence of frames. After that, all frames from the video recording of an endoscopic procedure with a similar pathological conditions can be retrieved.

The similarity of images is a generic term, conversely, if we talk more specifically in the context of gastrointestinal diseases, can be images with similar pathological conditions (e.g., bleeding, ulcer, and inflammation) or endoscopic images from the same area of GI tract (Nosato et al. 2015). Therefore, we can refer a CBIR system as a CADx, instead of calling it as a component (Bonnel et al. 2009). The working of CADx typical CADx has shown in Fig. 1. The CADx with CBIR system takes a query image as input and search for its match, nonetheless, the output may contains a single image or set of images.

(2) *Disease Detection* Classification and segmentation are two basic machine learning problems. In context of disease detection, a system classifies or segments the lesion areas and it refers to a CADx system. However, some CADx systems only segment the diseased area in endoscopy images, but not classify them into several stages or disease (for example see Hwang and Celebi 2010). Moreover, the frames or lesions can be categorized by training models based on extracted features as describes below in detail.

(a) *Image Classification* Many CADx systems have been developed for the classification of the endoscopic images. These CADx systems only decide, whether an endoscopy frame belongs to abnormal or normal class. However, it can be a multi-level classification, where if a CADx system detects an image as abnormal further, it classifies the input image to grade or severity-level of disease (Yang et al. 2015).

Therefore, some classification systems can distinguish between multiple types of abnormalities (Nawarathna et al. 2014). Most of the classification systems are trained through extracting features from images and these images are labeled with their respective classes (also known as supervised learning). Different types of descriptors (described in Sects. 3, 4, and 5) are extracted from endoscopy images in form of numeric values and used to train the

classifiers. After completion of the training phase, the trained models are used to predict the images class without providing the respective class label. The CADx automatically assigns label to these images or video frames as shown in Fig. 1 according to their respective class. Although, the classification system could have real-time constraints for finding abnormal frames in a live video (Liedlgruber et al. 2011).

(b) *Image Segmentation* Image segmentation is an important and also a difficult process, specifically the in case of gastroenterology images. Because the dynamic imaging environment of the GI tract possess various challenges, as described in earlier sections. Image segmentation (also known as ROI selection or perceptual grouping of pixels) refers to a process of extracting a sub-image or set of pixels with similar characteristics from an image. In the context of CADx, these pixels represent the diseased area in a frame (Szczypiński et al. 2014). CADx are developed for the classification and segmentation of gastric images share some essential steps or components. They take images as input, perform pre-processing on these images, however some CADx use images without the pre-processing step. Although, this is a problem specific decision to pre-process the gastric images or not. After the image pre-processing, segmentation is performed for the selection of the abnormal area from endoscopic images (Rajivegandhi et al. 2015). However, many systems that only segment the abnormal areas in the images also provide annotations on these images (van der Sommen et al. 2014).

The segmentation of different irregularities with a CADx is a challenging job. There exist CADx for detection of other abnormalities such as e.g., polyps, bubbles, and blood, which can be found in GI tract discussed in later sections. The segmentation could be a real-time task or it can be done on a recorded video (Ševo et al. 2016). Figure 1 shows the architecture of a typical CADx.

The performance of these types of CADx is measured using parameters given in Table 4. These accuracy measures are used to measure the effectiveness of a system that has developed for classification or segmentation of gastric lesions in endoscopic frames.

(3) *Systems Output Summary of Endoscopic Procedures* There is also a third form of CADx that output not in form of a decision, neither an image. It gives the summary of an

Table 4 Accuracy measures used to evaluation the performance of classification and segmentation methods

#	Description	Formula
	Classification (Ali et al. 2017)	
1	True negative (TN), True positive (TP), False negative (FN), and False positive (FP).	
2	Accuracy (ACC)	$ACC = \frac{TP+TN}{P+N}$
3	Area under the curve (AUC)	$AUC = \int_{-\infty}^{\infty} TPR(T)FPR'(T)dT$
4	Sensitivity or true positive rate (TPR)	$TPR = Sensitivity = \frac{TP}{TP+FN}$
5	Specificity or true negative rate (TNR)	$TNR = Specificity = \frac{TN}{FP+TN}$
6	False positive rate (FPR)	$FPR = 1 - TNR$
	Segmentation (Riaz et al. 2013)	
1	Dice similarity coefficient (DSC), where annotated (A) image and segmented (S) image	$DSC = 2 \times \frac{A \cap S}{A+S'}$
2	F-measure, where precision (P) and recall	$F = 2 \times \frac{P \times R}{P+R}$

endoscopic procedure in form of a text document. The summary of GI tract is mostly used for WCE, due to a large amount of images and no control over the endoscopy movement (Bao et al. 2015). The system takes the images or set of frames in case of endoscopy video, and returns the summary with respect to another dimension which is time (Wang et al. 2016b). A summary contains temporal information, along with that, on which frame-interval contains a specific area of the GI-tract (Zou et al. 2015) or location of a diseased (abnormal) frame in the GI-tract. Information extracted by such systems could be crucial for the gastroenterologist in decision making for surgery or biopsy.

2.3 Summary and insights

In this section, an overview of CADx systems is given. CADx system are developed specifically for detection of diseases in GI tract. Also, a brief introduction of these components belongs to a typical CADx system is given. Various pre-processing tasks and the problems in exploration of GI environment are mentioned in a great detail. The CADx system are divided into three categories with respect to their respective outputs. Additionally, these invariants of CADx systems are explained. It is clear from our survey, that a small number of CADx systems exist for the image retrieval application. A large amount of work has been done on frames classification and segmentation. However, the segmentation of gastric lesion still needs much attention of researchers. The CADx systems that provide summaries of endoscopic procedures are also very rare in literature. Since development of such system is very complex in nature. Therefore, to develop these kinds of systems, one must have to consider various aspects and challenges in machine-vision.

2.4 Video endoscopy (VE)

The major categories of endoscopy include wired endoscopy with white light, the flexible wired standard endoscopy we referred in our paper as video endoscopy (VE). The second category in white light endoscopy is WCE which has been widely used for the inspection of the whole GI tract. Both of these technologies are described in upcoming subsections and the classification of these endoscopy techniques is presented in Fig. 3 and output and working with FOV shown in Fig. 4.

(1) *White Light Endoscopy* Endoscopy as we know, a procedure performed by a flexible tube like instrument having a mounted camera and light source on its distal tip (Thekkek et al. 2015). The gastroenterologist can have a good control over the movement of VE than WCE and can also perform cleansing by using the accessory channel while examining the GI tract (Beg and Ragunath 2015). Therefore, the standard VE still considered to be a most effective and less invasive way to discover small-size lesions and also used for biopsy of these lesions by employing the accessory channel (Iakovidis et al. 2006). (See Fig. 4a)

(2) *Wireless Capsule Endoscopy (WCE)* WCE is a more convenient way to intervene and inspect the GI tract. However, there are some issues involved with the usage of WCE for inspection of the gastric tract (see Fig. 4b). Primarily, an issue that is associate with WCE, is no control over the movement of camera (Liu et al. 2015). However, several methods have been proposed for controlling the movement of WCE (Carpi et al. 2011; Gao et al. 2009). Secondly, WCE is lacking an accessory channel and cannot be used for a real-time biopsy (Francisco et al. 2015). The third issue is the cost of the capsule, on top of the extra cost of screening by a medical expert. Because the capsule is disposable and it is used for only one time (Albisser 2015). Although with these issues, the WCE provides a painless solution to
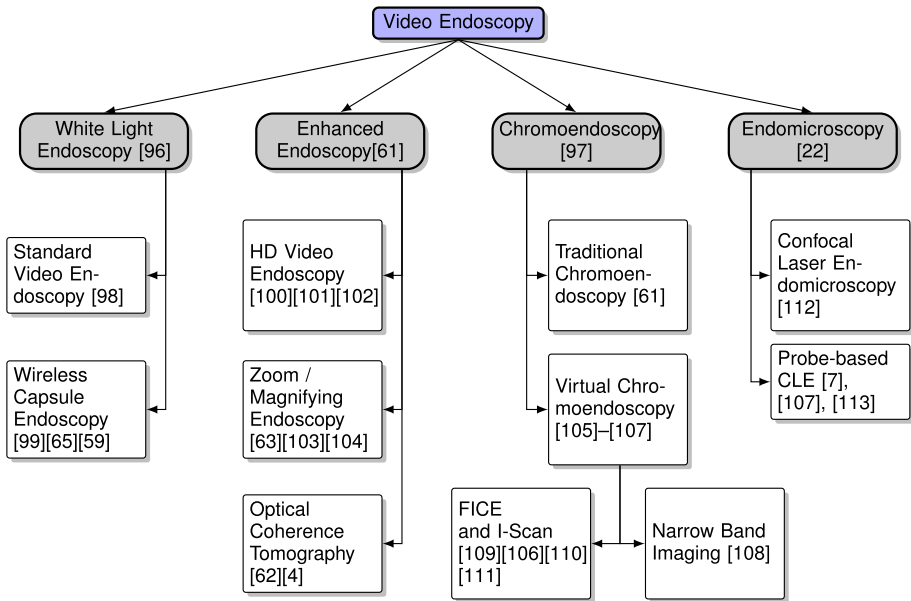
Fig. 3 Technological variations in the endoscopy for examining the GI tract in the pursuit of a better visualization of abnormal regions
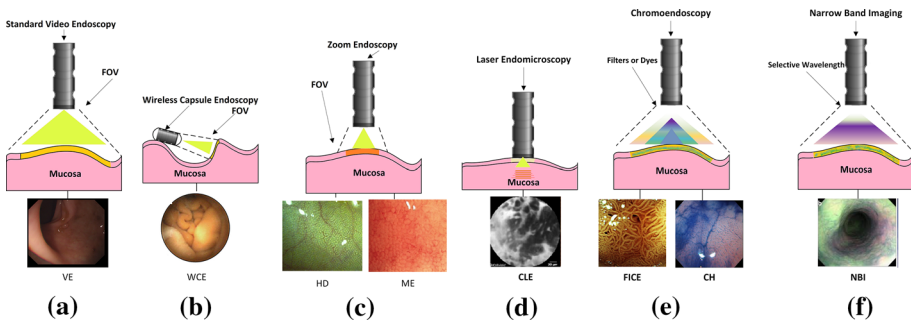

Fig. 4 Multiple endoscopy technologies and their respective FOVs, working, and outputs have depicted in this figure, **a** shows the working of a standard video endoscopy where the FOV is feasible for easy movement, **b** zoom endoscopy with reduced FOV not easy for the gastroenterologist to navigate through GI, **c** the WCE move along the GI tract without any navigation and it also has an uncontrolled FOV, **d** CLE with its mucosal intervention and a reduced FOV, **e** CH endoscopy and its two types of outputs, **f** NBI endoscopy and its output frame is visualized (some contents have adapted from Hegenbart et al. 2015)

the screening of the GI tract (Keuchel et al. 2015). There are some areas e.g. (small bowel) unreachable by using VE. Therefore, WCE is a less invasive option for such areas which are difficult to get to for VE (Ogoveanu et al. 2015).

2.5 Enhanced endoscopic technologies

VE is a standard definition video endoscopy which is equipped with the CCD camera with pixel resolution of more than 400,000 pixels per image. Moreover, the details preserved by

an endoscope frames are depending on the number of pixels (resolution) of this CCD camera. There are multiple advancements have been made for enhancing the visibility of the frames for the gastroenterologist and these are discussed in detail by Song and Ang (2014). However, a brief overview of some of these techniques has given as follows.

(1) *High-Definition Video Endoscopy (HDVE)* The advancements in the technology lead to the more density of transistor embedded in a single chip and resolution of the camera is also has increased. Advancements in the CCD technology and currently in the complementary metal-oxide semiconductor (CMOS) technology can allow embedding a large amount of pixels (more resolution) in a small chip. These chips are used in new high-definition (HD) endoscopes. The images produced by a HDVE have a resolution of 85 thousand to more than 1 million pixels (Bhat et al. 2014). Therefore, the area that is visible by standard VE can be magnified by 150 times or more in a HDVE. Moreover, the mucosal vascular structures are now more visible by using the HDVE instead of the standard VE (Penny et al. 2016; van der Sommen et al. 2014). The issue with the HDVE is its FOV (see Fig. 4c), the movement of endoscope become difficult because of a small FOV. Therefore, in recently developed endoscope, the medical expert can switch between VE and HDVE (Gotoda et al. 2016).

(2) *Zoom/Magnified Endoscopy (ME)* Magnifying endoscopy have similar benefits as the HDVE. Therefore, one can easily confused in the HDVE and ME due to their similar advantages, the output of ME is shown in Fig. 4c. However, the magnification can be achieved through some filter of lens in the standard VE (Leggett and Iyer 2015). The ME is also used for visualizing the mucosal structure in a large scale normally used with other imaging technologies (e.g., NBI and CH) (Muto et al. 2016; Lopez-Ceron et al. 2013).

(3) *Optical Coherence Tomography (OCT)* The issue with HDVE and ME is difficulty for the gastroenterologist in controlling the movement of the endoscope when the image has magnified. The solution to this problem is resolved by using optical coherence tomography (OCT) is an independent of endoscopy. it is an adaptive feature to standard VE (Qi 2008). It is used through the accessory channel of endoscopy. The OCT also provides great resolution, with and the VE helps in the movement in finding the abnormal area. OCT is a recent optical technique based on low-coherence interferometry. The OCT uses B-mode ultrasongraphy by focusing light beam on a target area and collect the scattered reflected light to construct a cross-sectional image (Beg and Ragunath 2015).

2.6 Chromoendoscopy (CH)

The Chromoendoscopy (CH) is traditionally used to investigate the mucosal structures using staining techniques. The mucosal surface is highlighted by sprinkling colourants over the mucosal surface and visualizing it under the light as shown in Fig. 4d. These dyes make the malignant area more prominent and reported useful in the gastric surveillance (Wong Kee Song et al. 2007).

(1) *Virtual Chromoendoscopy (VCH)* The digital or virtual chromoendoscopy (VCH) involves image processing algorithms and band-pass filters to give the effect of a dye-based (Traditional) chromoendoscopy (Kaltenbach et al. 2008). One advantage of VCH over the CH is the unnecessary cleaning. There is no need for spraying and suction for the cleansing of dyes for further endoscopic procedures (Buchner et al. 2010). Moreover, the gastroenterologist can switch between HDVE and VCH by using a single button, which makes it more user-friendly (Goetz et al. 2013).

(a) *Narrow Band Imaging (NBI)* Narrow Band Imaging is a type of chromoendoscopy which uses optical filters (digital image processing) and limited bandwidth characteristics of

light (Kodashima et al. 2014). NBI highlights the mucosal irregularities specially associated with polyps and dysplasia (see Fig. 4e). NBI uses as described above NBI is VCH it uses to filter light to improves the mucosal vascular structures, veins, and capillaries without dyes (Buchner et al. 2010).

(b) *FICE and I-Scan* FICE (Fuji-non Inc, Japan) system (Van Gossum 2015) and iSCAN systems (Pentax, Japan) (Leggett and Iyer 2015) use reflectance of light is processed instead of filtering the light that is coming from the source as done in NBI (see Fig. 4d). The reflected light is processed by a spectral estimation matrix circuit. The sensed high contrast color image, that is constructed through a combination of red, green and blue wavelengths (Coda 2014; Goetz et al. 2013; Nishimura et al. 2014).

(2) *Endomicroscopy (EM)* The most used imaging technology for visualization of mucosal vascular structures is CLE. Endomicroscopy is performed by focusing a beam of photons at various mucosal layers (Luck et al. 2004).

(a) *Confocal Laser Endomicroscopy (CLE)* The confocal laser endomicroscopy (CLE) can have many time magnified image than standard VE. Therefore, CLE provides details of inner layers of GI mucosa. The CLE system is endoscope based, we have denoted it as CLE and the other probe-based referred as pCLE (Gómez et al. 2010). The CLE uses standard VE with confocal imaging aperture. However, this system is now obsoleted (Francisco et al. 2015).

(b) *Probe-based CLE (pCLE)* The probe-based confocal laser endomicroscopy (pCLE) is a separate system that is inserted through the instrument channel of the VE. It contains its own processor, the images are acquisition is done by placement of probe on the mucosal layer (Wallace and Keisslich 2010; Buchner et al. 2010). The output of pCLE is depicted in Fig. 4f.

2.7 Summary and insights

The video endoscopy is a basic and minimal invasive tool for screening the GI tract. It is very useful in histology and surveillance of gastric disease. Currently, many advanced variations of video endoscopy exist as discussed in this section. Standard endoscopy is used more in practice. However, it has limited access to lower GI tract. Because the screening of small bowel is difficult due to the narrowness and twisty tract. Therefore, the WCE can be used, for screening far areas in the GI tract. There is no doubt, that these technology areas developed for assisting medical expert. However, adopting hybrid approaches (by combining the two different imaging technologies advantageously) can provide more flexibility in selecting FOV, desired magnification. Moreover, advancement in nanotechnology has opened the new pathways to gastrointestinal screening (Yan et al. 2012).

3 Features extraction techniques in spatial domain

In this paper, our main focus is on feature extraction techniques employed in CADx system for gastric diseases prevailing in the GI-tract. The features extraction is an important phase for detection of abnormalities besides the preprocessing and segmentation of endoscopic frames. Moreover, various training models are trained by extracting descriptors from endoscopy images (Ali et al. 2017). These trained models are further used for the segmentation and classification of gastric frames. There are several techniques have been proposed for features extraction. However, choosing an efficient feature extraction method depends on the nature
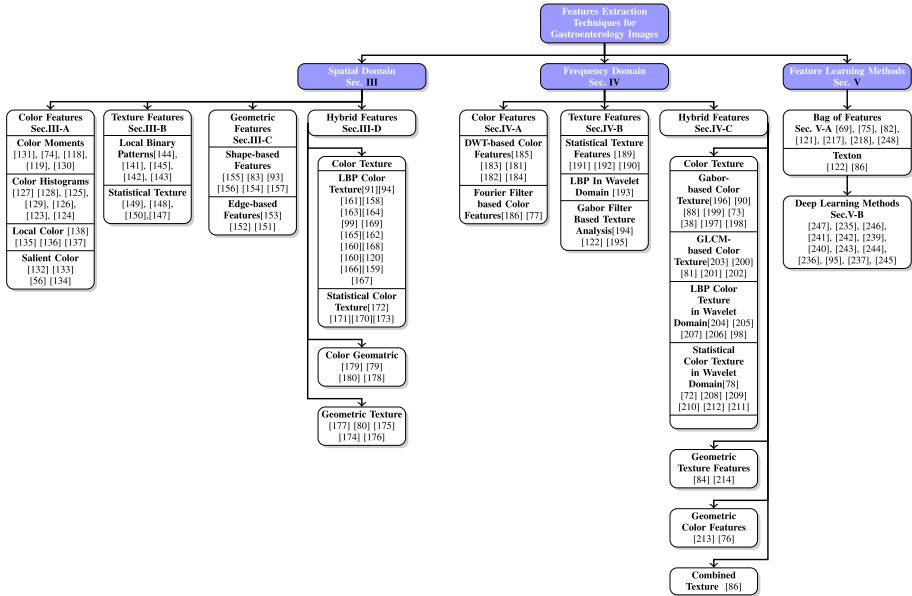
Features Extraction Techniques for Gastroenterology Images

Spatial Domain Sec. III | Frequency Domain Sec. IV | Feature Learning Methods Sec. V

Spatial Domain Sec. III

Color Features Sec.III-A
Color Moments [131], [74], [118], [119], [130]
Color Histograms [127] [128], [125], [129], [126], [123], [124]
Local Color [138] [135] [136] [137]
Salient Color [132] [133] [56] [134]

Texture Features Sec.III-B
Local Binary Patterns[144], [141], [145], [142], [143]
Statistical Texture [149], [148], [150],[147]

Geometric Features Sec.III-C
Shape-based Features [155] [83] [93] [156] [154] [157]
Edge-based Features[153] [152] [151]

Hybrid Features Sec.III-D
Color Texture Texture[91][94] [161][158] [163][164] [99] [169] [165][162] [160][168] [160][120] [166][159] [167]
Statistical Color Texture[172] [171][170][173]
Color Geomatric [179] [79] [180] [178]
Geometric Texture [177] [80] [175] [174] [176]

Frequency Domain Sec. IV

Color Features Sec.IV-A
DWT-based Color Features[185] [183] [181] [182] [184]
Fourier Filter based Color Features[186] [77]

Texture Features Sec.IV-B
Statistical Texture Features [189] [191] [192] [190]
LBP In Wavelet Domain [193]
Gabor Filter Based Texture Analysis[194] [122] [195]

Hybrid Features Sec.IV-C
Color Texture
Gabor-based Color Texture[196] [90] [88] [199] [73] [38] [197] [198]
GLCM-based Color Texture[203] [200] [81] [201] [202]
LBP Color Texture in Wavelet Domain[204] [205] [207] [206] [98]
Statistical Color Texture in Wavelet Domain[78] [72] [208] [209] [210] [212] [211]
Geometric Texture Features [84] [214]
Geometric Color Features [213] [76]
Combined Texture [86]

Feature Learning Methods Sec. V

Bag of Features Sec. V-A [69], [75], [82], [121], [217], [218], [248]
Texton [122] [86]

Deep Learning Methods Sec.V-B [247], [235], [246], [241], [242], [239], [240], [243], [244], [236], [95], [237], [245]

Fig. 5 Features extraction methods used in CADx for detection of gastric diseases, divided according to respective domains

of application, disease, and imaging modalities for which these were being extracted. If feature extracted for classification, it might be possible that these features are not suitable for segmentation or retrieval applications. In proceeding sections, the summaries existing of CADx systems are given and features extraction methods used in these CADx systems are discussed in great detail.

As we know, a digital image is represented by a 2D array of pixel values in the spatial domain of image processing. The spatial image processing refers to directly manipulation and analysis of these pixels. Consequently, for an early diagnosis of malignancies from gastric images, every pixel is precisely investigated (Gono et al. 2004). There are several feature extraction methods have been developed for the automated diagnosis of disease from endoscopic frames. We have classified them according to their perceptual information as described below. Furthermore, we have divided these categories according to their respective information into subcategories, which are used for classification of gastric lesions (see Fig. 5).

3.1 Color features in spatial domain

Color features are basic visual characteristics of images. Colors clues about the mixture of lights of different bandwidth from the visual spectrum. In the context of gastric frames, the colors are very important for visualizing the mucosal surface (in case of NBI and CH). Furthermore, the colors play an important role in the detection of clinical pathologies like an ulcer, bleeding, and inflammation, etc. (Cui et al. 2010; Li and Meng 2009a, b; Yuan et al. 2015a). Endoscopic frames are normally acquired in combination of three channels (red, green, and blue) RGB frames and every channel gives the intensity of a particular primary color. However, the endoscope generates gray images with a single intensity channel (for example see Fig. 4d). In a normal gastric environment, the images have a low contrast. The

reason for the low contrast of gastric frames is a less usage of color space (Riaz et al. 2012). Various color spaces are designed for representation of images and they have different applications accordingly. However, some notable color spaces are RGB, HSV, CIE-LAB, CIELUV, and CIEXYZ with three channels, but we do not say which one is effective for detection of specific malignancies. The features extraction methods involving color descriptors in spatial domain are summarized in Table 5.

(1) *Color Histograms* Basic information which can be extracted from a channel of an image is its histogram (Ghosh and Fattah 2017; Deeba et al. 2018). A histogram hints about the likelihood of a pixel intensity and gives a guess about the distribution of colors. RGB and YUV color space used by (Häfner et al. 2007) individual channels and combined RGB for the analysis of zoom endoscopy images. Similarly, the (Krishnan et al. 1998) used the HSI and RGB channels histogram. Additionally, used these color histogram features to train and ANN. RGB histogram features are employed for bleeding detection from WCE frames in Ghosh et al. (2015). Moreover, the HSV histogram color features were preferred over the use of RGB histogram features in some studies. WCE images which are normally represented in RGB color space can be transformed into other color spaces. Such as HSV color space, for better uniformity in colors. For instance, HSV color histograms were adopted for classification of bleeding frames in Poh et al. (2010). Local RGB color image histograms are used for computing threshold for the segmentation colonoscopy images in Tjoa et al. (2002) and CIE-LAB color difference method was used to minimize the error of segmentation.

(2) *Color Movements* As we know, colors are very sensitive to illumination variations. In a dynamic environment like GI-tract images acquired under various lighting conditions, which poses new challenges for detection of gastric abnormalities while using colors as descriptors. Color moments are designed for dealing with color variations occur due to illumination changes. Using HSI color space and Tchebichef polynomials are used as basic functions for detection of bleeding and ulcer in WCE frames (Li and Meng 2009a). Similarly, in Li and Meng (2009b) bleeding regions are classified. Likewise, many feature extraction methods have been developed that uses different order statistics to represent the color features in RGB of WCE frames, initially converting them in to HSI color space (Lee et al. 2007). In the same way, precancerous lesions are detected from endoscopy frames by using 14 statistical features (Martinez-Herrera et al. 2016). Camera distortion in WCE is a big issue for many algorithms due to lack of control over the movement of the camera. The variable distance of camera to the mucosal wall causes undesired results. CIE-LAB color space separates light and uniform spaced channels. Hence, color moments are computed from channels of images as variances, entropy, and kurtosis for detection of ulcer and bleeding regions (Vieira et al. 2015). HSI color space is employed for color feature extraction due to its similarity to human visual perception system (Cui et al. 2010). Likewise, a number of color features are computed by computing statistical measures of different channels and their combinations for detection of bleeding from WCE frames.

(3) *Salient Color Features* For bleeding detection salient regions are detected by colors by transformation of RGB images into CMYK and CIE-LAB color space then first order moments are calculated to form features' set (Yuan and Meng 2015). Salient super pixels are identified for detection of bleeding regions by using color features in CIE-LAB color space (Iakovidis et al. 2015). RGB frames are transformed to HSV color space for color normalization, further, these images are converted back to RGB color space and color spectrum transformation has been performed for segmentation of bleeding regions (Jung et al. 2008)

(4) *Local Color* Local color information is extracted from images by dividing every image into small patches. Further features are extracted from these blocks. Pixel values of every patch are used as a color descriptor (Li and Meng 2007). Then, the local color features computed

from endoscopic images and additional every block is analyzed for potential bleeding by color values in Lau and Correia (2007). In some methods descriptors also may include pixel spatial location as well as color information as given in Alexandre et al. (2007). Similarly, in pixel values and spatial distances of different pixels are compared for segmentation of CH images in Riaz et al. (2009).

3.2 Texture features in spatial domain

Texture refers to a repetitive pattern in an image; In addition, it gives information about the characteristics of the surface of the image like e.g., coarseness and smoothness.

Many texture extraction techniques are proposed and applied for texture classification such as a Multi-scale Symmetric Dense Micro-block Difference (MSDMD) technique is introduced. It merges K-rotation with Gaussian distribution to experiment and utilize dense micro-block changes as local features to obtain pixel-level changes. Following, a High-order Vector of Locally Aggregated Descriptors (HVLAD) is formed, to encode the local descriptors to obtain a global descriptor. Finally, by fusing an average Spatial Pyramid Pooling, the multi-scale SDMD is carried out to produce an MSDMD-based texture descriptor (Dong et al. 2018b).

A multi-scale frequency and difference based representation (CDR) of image textures for classification is proposed. The local counting vector (LCV) is used to extract different types of textural formations employing the discrete local counting projection, while the differential excitation vector (DEV) is used to represent the variation of textures according to the differential excitation projection. Then, multiple texture features are formed by combining CDRs at various scales (Dong et al. 2018a).

Likewise, texture analysis is widely used in various fields. Specifically, texture analysis gained much importance in medical image analysis. Furthermore, texture features are very useful for detection of the cancerous region from endoscopic frames. The texture representation methods used to analyze endoscopic frames in spatial domain are discussed with their applications in upcoming paragraphs. An overview of these techniques is shown in Table 6.

(1) *Local Binary Patterns (LBP)* Local binary patterns (LBP) are very useful for representing images' texture. The simplest form LBP is calculated by comparing neighboring pixels with the central pixel and assigns it a binary code respectively. Additionally, these binary codes are transformed into decimals. The texture of image is represented locally, by computing occurrence of these codes by forming a histogram. The most important advantage of LBP is its rotation invariance (Constantinescu et al. 2015). Various advancements have been made to enhance the LBP's representational power. Therefore, multi-scale LBP is proposed which deals with illumination variations of WCE frames of small bowel (Li et al. 2011). The LBP with various neighborhood pixel size 8,16, and 24 are used with a combination of uniform LBP in Gross et al. (2009). Multiple variations of LBP e.g., LBP59, LBP256, LBP10, and difference based LBP are used for classification in Dahal et al. (2015). Uniform LBP is combined with vector quantization for feature extraction and then employed for classification of endoscopic frames (Lee et al. 2013).

A new LBP based features jumping and refined local pattern (JRLP) is presented for texture classification in Wang et al. (2018b). The local jumping information is extracted by first calculating jumping local difference count pattern (JLDCP) with second-order difference count pattern and diagonal difference count pattern to represent the jumping information further the detailed information of left by JLDCP is recorded to extract a refined completed

Table 5 A summary of color features extraction methods in spatial domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
Color histogram										
2018	Ghosh and Fattah (2017)	WCE	GI tract	Color histogram	Classification	Bleeding	k-NN	2300 images	Cross-validation 10	99.47%sen, 99.15%spec, and 97.85%acc
2018	Deeba et al. (2018)	WCE	GI tract	Color HSV and RGB	Classification	Bleeding	k-NN	8872 images	Cross-validation 5	99.6% acc
2016	Ghosh et al. (2015)	WCE	GI tract	Color histogram	Classification	Bleeding	k-NN	1000 images 200 images 200 of them show a sign of bleeding and other 800 as non-bleeding	LOO-CV	96.48%sen, 96.01% spec and 96.10%acc
2010	Poh et al. (2010)	WCE	GI tract	Adaptive color histogram	Classification	Bleeding	SVM + ANN	200 images (576 × 576)	Holdout	98.90%
2007	Häfner et al. (2007)	ZE	Colon	Luminance histogram Co-occurance channal histogram	Classification	Type1 colon lesion	k-NN	257 images 156 nonneoplastic and 101 neoplastic cases	Holdout	85%acc
2002	Tjoa et al. (2002)	VE	Colon	CIELAB and RGB Histogram	Segmentation	Polyps	None	None	None	None
1998	Krishnan et al. (1998)	VE	Colon	Color histogram	Classification	Bleeding	MLP NN + SOM	22 images for training 4 images for testing	Holdout	95%acc

Table 5 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
Color moments										
2016	Martinez-Herrera et al. (2016)	VE	GI tract	Statistical moments	Classification	Cancer	SVM, NN, k-nn	102 multispectral images 74 correspond to a healthy profile and 28 from precancerous diagnosis	LOO-CV	77%acc, 91%sen and 62%spec
2015	Vieira et al. (2015)	WCE	GI tract	Color a b histogram moments	Classification	Bleeding regions, Crohn's disease, suspected tumors, ulcers,polyps	SVM + MLP	2500 belong to normal and 700 to abnormal class	Cross-validation 10	97.4%sen 98.4%spec 98.0%acc image level 98.5%sen 98.8%spec 98.8%acc region level
2010	Cui et al. (2010)	WCE	GI tract	Six color features in HIS color space	Classification	Bleeding	SVM	6416 images 4 group 802 frames for training and 401 frames for testing	Holdout	97%acc
2009	Li and Meng (2009a)	WCE	Small bowel	Color moments	Classification	Bleeding, Ulcer	ANN	100 images of 10 patients	Cross-validation 4	84%spec and 92%sen
2007	Lee et al. (2007)	WCE	GI tract	Statistical color moments	Classification	Bleeding and other anomalies	None	5000 images	None	76%pre and 51%sen

Table 5 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
Salient color features										
2015	Yuan and Meng (2015)	WCE	GI tract	Salient color features	Classification	Bleeding and other anomalies	SVM	800 images and 400 images belongs to normal 400 images belongs to abnormal 200 images for training 200 images for testing	Holdout	95.89%acc, 98.77%sen and 93.45%spec
2015	Iakovidis et al. (2015)	WCE	GI tract	Color based CIE-LAB and selected first-order statistical features from various color components using SVM and saliency	Classification	Bleeding and other anomalies	SVM With RBF	252 cases (320 × 320)	Cross-validation 10	96%sen, 91%spec and 94%acc
2014	Khan and Wahid (2014)	WCE	GI tract	Color homogeneity, Red dominance	Compression	Multiple	None	None	None	None
2008	Jung et al. (2008)	WCE	GI tract	Color spectrum	Segmentation	Bleeding	None	1000 images	None	92.86%sen and 89.49%spec

Table 5 continued

Local color features

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2009	Riaz et al. (2009)	VE	GI tract	Pixel color	Segmentation	Cancer	Statistical classifier	144 images (518 × 481)	None	0.59DSC
2007	Li and Meng (2007)	WCE	GI tract	Local color	Classification	Multiple	SVM + MFNN	60 images 30 normal and 30 abnormal (256 × 256)	Holdout	65.2%sen and 82.5%spec
2007	Lau and Correia (2007)	WCE	GI tract	Local color	Segmentation	Bleeding	None	5 videos 401 frames	None	88.3% sensitivity
2007	Alexandre et al. (2007)	VE	Colon	RGB color features	Classification	Polyps	SVM	4620 images	Cross-validation 2	0.93AUC

Table 6 A summary of texture features extraction methods in spatial domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Data-set	Validation method	Results
Local binary patterns										
2015	Dahal et al. (2015)	VE	Colon	Difference based LBP	Classification	Ulcer	k-NN	207 images (720 × 480)	Cross-validation 10	90%acc
2015	Constantinescu et al. (2015)	WCE	Small bowel	LBP	Classification	Ulcer	SVM	344 images	Holdout	93.16%acc
2013	Lee et al. (2013)	VE (NBI-ME)	GI tract	LBP	Classification	Cancer	Unsupervised	57 images are used of which 27 images of abnormal lesions and 30 normal images (639 × 480)	None	46–100%pre and 39–87%sen
2011	Li et al. (2011)	WCE	Small bowel	Multi-scale local binary pattern (MS-LBP)	Classification	Cancer	SVM + KNN ensemble	1200 WCE-frames of 10 patients	Cross-validation 4	90.50%acc, 92.33%sen and 88.67%spec
2009	Gross et al. (2009)	NBI VE	Colon	LBP	Segmentation	Polyp	k-NN	102 polyp images	None	90%acc

Table 6 continued

Statistical texture features

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Data-set	Validation method	Results
2018	Ghosh et al. (2018)	WCE	GI tract	YIQ color space statistical features	Classification	Bleeding	SVM	100 images	Holdout	98%acc
2010	Magoulas et al. (2004)	VE	Colon	GLCM (energy-angular second moment, correlation, inverse difference moment, inverse difference moment)	Classification	Cancer	BP-ANN	80 images	Holdout	93%acc
2009	Ameling et al. (2009)	VE	Colon	GLCM 6 GLCM 16 LBP OC-LBP	Classification	Polyp	SVM	1736 images	Cross-validation 4	0.96AUC
2003	Maroulis et al. (2003)	VE	Colon	co-occurrence matrix texture features	Classification	Polyp	MLP-ANN	150 training 2809 testing (512 × 512)	Holdout	95%acc

LBP (RCLBP). The JRLP-based texture descriptors are created by combining both JLDCP and RCLBP.

(2) *Statistical Texture Features* Statistical measures are widely used to represent the texture of images. Statistics about intensity distribution delivers information about image's texture. In some developed methods, gray-level co-occurrence matrices (GLCM) are extracted by calculating the frequency of certain pixels in pairs. Further, several statistics (energy, contrast, correlation and homogeneity) are calculated from these matrices to represent the texture of images (Dahal et al. 2015). Similarly, Haralick's features are used in Bejakovic et al. (2009) for lesion detection from WCE frames. In Ghosh et al. (2018), a method presented to classify bleeding frames using statistical features computed from YIQ color space. In Ameling et al. (2009), texture features are extracted by computing GLCM for classification of colonoscopy images. These texture features are also used by Magoulas et al. (2004) where energy-angular second moment, entropy, inverse difference moment, and correlation are computed for analysis of WCE frames. In their previous work, they have used same features for detection of the colorectal lesions from VE videos (Maroulis et al. 2003).

3.3 Geometric features in spatial domain

(1) *Edge-Based Features* The gastric lesions normally do not own any particular shape or size. However, some lesions e.g., polyps can be represented through a geometric model. Therefore, the curvature of a lesion can be calculated by finding edges or contour of polyps (Krishnan et al. 1998a). As the same, in Kang and Doraiswami (2003) uses Sobel and Canny's edge detectors to model the shape of polyps. On the other hand, Celiac disease detection performed through the edge based features (Boschetto et al. 2016). For contraction detection wrinkle skeleton comprised of 14 features: 2 features of edge sharpness, a set of 4 local, a set of 8 directional features, and entropy related features (Spyridonos et al. 2006). The tensor gradients are calculated prior to determination of these features, the further statistical classifier is used for classification of endoscopic frames.

(2) *Shape-Based Features* More efficient methods for modeling shapes are used for analyzing pit-patterns from NBI images where fractal dimension, smooth spiral curve, Koch snowflake, Sierpinski triangle, and checkerboard (Häfner et al. 2015). An extension of higher order local auto-correlation (HLAC) features is used for retrieving multi-scale objects from optical colonoscopy images (Nosato et al. 2015). The HLAC method can calculate some geometrical features. Moreover, HLAC features represent the expressed characteristics for the whole endoscopic image, derived from the product-sum operations of the auto-correlation formula. In Hwang et al. (2007), elliptical shape-based features are used for detection of polyps in colonoscopy frames. Point-based SIFT features are employed for tracking the path of GI tract (Bao et al. 2015) which helps in the construction of 3D trajectory of WCE. Several geometric features with fuzzy logic used to classify the colon polyps (Krishnan and Goh 1999). Summaries of these methods are described in Table 7.

3.4 Hybrid features in spatial domain

In above-described feature extraction methods, descriptors extracted by processing frames in the spatial domain. These features extraction methods may work well in some specific conditions (e.g., for bleeding detection colors are good features). However, for detection of complicated abnormalities (e.g., cancer, polyps, and ulcers), these features have a limited discrimination power when they are separately used for CADx. They can be combined with

Table 7 A summary of geometric features extraction methods in spatial domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Data-set	Validation method	Results
Shape-based features										
2015	Häfner et al. (2015)	NBI VE	Colon	Fractal dimension	Classification	Polyp	k-NN	359 images of type A, 462 images of type B and 87 images of type C (800 × 900)	LOO-CV	88.2%acc
2015	Nosato et al. (2015)	VE	Colon	Higher order local auto-correlation (HLAC)	Retrieval	Multiple	None	100 images (800 × 600)	Cross-validation	None
2015	Bao et al. (2015)	WCE	GI tract	Point features SIFT	Summary	Multiple	None	None	None	None
2007	Hwang et al. (2007)	VE	GI tract	Elliptical shape	Segmentation	Polyp	None	8621 frames which consists of 815 polyp frames and 7806 normal frames	None	96%acc
1999	Krishnan and Goh (1999)	VE	Colon	Geometrical parameters	Segmentation	Multiple	Fuzzy rules	None	None	None

Table 7 continued

Edge-based features

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Data-set	Validation method	Results
2016	Boschetto et al. (2016)	CLE	Small bowel	Edge based features	Classification	Celiac disease	Linear NB + Quadratic NB	67 images (class distribution: 31 normal images, 17 with VA and 19 with CH) (1024 × 1024)	Cross-validation 10	94.03%acc
2006	Spyridonos et al. (2006)	WCE	Small bowel	Wrinkle skeleton comprised of 14 features: a set of 8 directional features, a set of 4 local entropy related features and 2 features of edge sharpness	Classification	Contraction	SVM With RBF	521 positive examples of images of contractions and 619 images not corresponding to intestine contractions	Holdout	90.84%sen and 94.43%spec
2003	Kang and Doraiswami (2003)	VE	GI tract	Sobel and Canny Edge detectors are used to model shape of polyps	Classification	Polyp	None	None	None	None
1998	Krishnan et al. (1998a)	VE	Colon	Contour-based	Segmentation	Polyp	None	None	None	None

other features to cope with various issues of gastric environment like rotation, scale, and illumination variation in the images. Two or more types of features are combined to achieve a better discrimination for lesions. Hybrid approaches used for automated diagnosis of gastric lesions are listed below. The overview of hybrid methods presented in Tables 8 and 9.

(1) *Color Texture Features in Spatial Domain* Red color and filter-based texture features are combined for finding inflammation area from endoscope frames in Ševo et al. (2016). Multiple abnormalities are detected through a combination of texture, color, and edge-based features in Zhao et al. (2015).

(a) *LBP-based Color Texture* Same as above, color and texture features are combined in Kodogiannis et al. (2007), where modified LBPs are extracted from RGB and HSV color channels of endoscopic frames and then combined. By doing so, color features are combined with texture information in Zhang et al. (2009). The method presented in Wang et al. (2016b) used I channel histogram from HSI, HV histogram from HSV, RGB histogram, Norm RGB histogram, RG histogram from the opponent histogram, and hue histogram for representation of endoscopy images and the LBP textures are combined. Similarly, color histogram, PHOG, and LBP texture from the super-pixels are extracted in Cong et al. (2015). In Szczypiński and Klepaczko (2009), for extraction of texture and color information from images, texture and color components Y, R, G, B, U, V, I, Q, color saturation, and hue are combined to obtain a comprehensive characterization of a colored texture. Color moments from the histogram of RGB and HSV color WCE image are extracted and combined with LBP features for representation of color texture (Emam et al. 2015). Gaussian filtered LBP (GF-LBP) features are extracted from endoscopic images. Moreover, colors are extracted from the pyramidal histogram of endoscopy images (Li et al. 2015). Endoscopic images are processed by dividing in patches then from these patches LBP features extracted. LBP features are extracted from HSV and RGB channels for a patch-based classification (Yao et al. 2010). Similar work was conducted by Liu et al. (2015), where LBP moments are combined with color moments. LBP texture spectra along with color histogram are combined to get texture color information in Kodogiannis and Lygouras (2008). Likewise, multiple texture descriptors, color features, and their combination are described in Letter (2007). Color histograms are combined with LBP and HSV color components histograms are added with different combinations of LBP features. Then the LBP with the central pixel of neighborhood of 8 and 16 pixels are extracted (Sousa et al. 2009). Uniform-LBP features are computed by accounting each channel of endoscopy frame for combining texture and color information (Li and Meng 2012). Endoscopy images are transformed from RGB to HIS color space and then chrominance moments are calculated from the histogram. For texture features, LBP features are extracted and then added with color features. Additionally, Tchebichef polynomials are used to model these color-texture features (Li and Meng 2009b).

(b) *Statistical Color Texture* Color and texture features are combined by computing dominant colors from GLCM of the images, 8 dominant colors are computed from every single image (Giritharan et al. 2008). A method bi-dimensional ensemble empirical mode decomposition (BEEMD) has proposed in Charisis et al. (2010) where intrinsic mode functions (IMFs) computed from each channel of the image to represent color texture features. Statistical moments energy, mean, standard deviation, skew, kurtosis, and entropy are computed from histograms of images by representing images in RGB and HIS color spaces. Moreover, a different combination of these moments was used for classification of abnormal colon images (Adler et al. 2012). In the same way, GLCM features combined with color features in Moccia et al. (2018).

Table 8 A summary of hybrid color texture features extraction methods in spatial domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
LBP color texture										
2016	Ševo et al. (2016)	VE	Colon	Edge-based texture calculated from filters and red color	Segmentation	Inflammation	Unsupervised	3 videos (768 × 576)	None	80%acc
2016	Wang et al. (2016b)	VE	Stomach	Histogram (HSV, RGB, LBP) 7 histogram	Summary	Cancer	Random forest	2610 frames, 790 over-saturated frames, 640 dark frames and 1158 obscure frames (768 × 576)	Holdout 90% 10%	95%acc
2015	Cong et al. (2015)	VE	Small bowel	Color and texture features	Classification	Cancer	SVM	3800 images of 1284 patient	Holdout	0.9542 AUC
2015	Zhao et al. (2015)	WCE	GI tract	Color and texture features, edges	Classification	Multiple	HMM	5029 images	Cross-validation 3	93.3%acc and 93.3 %sen
2015	Emam et al. (2015)	WCE	GI tract	LBP, image sequence reduction and summarizing. Color Moment RGB, Color Moment HSV, Color Histogram	Classification	Multiple	Unsupervised	10671 images (256 × 256)	None	88.3%acc
2015	Li et al. (2015)	WCE	Small bowel	Spatial pyramid based color histogram Gaussian based local binary pattern	Summary	Abnormality (summery)	Unsupervised	None	None	None

Table 8 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2015	Liu et al. (2015)	VE	GI tract	Color histogram, color moment, and LBP	Classification	Cancer	OB-SVM	99 images of 44 patients 75 normal stomach images, 194 images of early esophageal cancer and 39 normal esophageal images	Holdout	93%sen, 83%spec and 85%acc stomach 0.83, 0.88, 0.87 esophagus
2012	Li and Meng (2012)	WCE	Small bowel	Color and LBP texture from LAB, HSV	Classification	Cancer	SVM	1200 images, 600 normal and 600 abnormal frames (256 × 256)	Cross-validation 10	92.4%acc
2010	Yao et al. (2010)	VE	GI tract	Color + LBP	Classification	Bleeding and other anomalies	Gentle-boost	534 images of 278 patients which contains 20521 positive samples and 99467 negative samples (768 × 576)	Cross-validation 10	75.7%TPR and 86.4%TNR
2009	Szczypiński and Klepaczko (2009)	WCE	GI tract	Texture + Color	Classification	Ulcer	RBF classifier	109 and 258 (two classes) images for training 200 images for testing	Holdout	91%sen and 95%spec
2009	Zhang et al. (2009)	VE	GI tract	Color histogram and U-LBP	Classification	Ulcer	Gentle boost ensemble	2949 images of 403 patient (768 × 576)	Holdout	75.5%sen, 77.1%spec and 0.838AUC
2009	Sousa et al. (2009)	ME	GI tract	LBP color texture	Classification	Cancer	KNN + NB + DT + k-NN	176 CH images (518 × 481)	Cross-validation 10	91%acc

Table 8 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2009	Zhang et al. (2009)	VE	GI tract	RG Histogram, HS histogram, and LBP	Classification	Cancer	GentleBoost	2949 images of 413 patients (768 × 576)	Cross-validation 2	80%acc
2009	Li and Meng (2009b)	WCE	GI tract	Chrominance moment and LBP (from RGB,HSI)	Classification	Bleeding and other anomalies	MLP-NN	200 images (256 × 256)	Holdout	72.6%acc 83%spec and 62.3%sen
2008	Kodogiannis and Lygouras (2008)	WCE	GI tract	Statistical feature + Ri, Gi, Bi, Hi, Si, Vi + texture spectra in the chromatic and achromatic domain	Classification	Abnormality	Multiple-classifier approach using Fuzzy integral	140 images, 70 normal and 70 abnormal cases 35 for training 35 for testing	Holdout	94.28%acc
2007	Kodogiannis et al. (2007)	WCE	GI tract	Texture + Color	Classification	Multiple	Neuro-fuzzy classifier	140 images 70 normal 70 abnormal cases	Holdout	95%acc
2007	Letter (2007)	WCE	GI tract	Color and texture	Summary	Summery	SVC	60,000 images	Holdout	99%acc

Statistical color texture

2018	Moccia et al. (2018)	NBI	Colon	GLCM and Color	Classification	Frame types	SVM	700 images	Cross-validation 10	83%acc 0.89 AUC
2012	Adler et al. (2012)	VE	Colon	Color + Textured based features	Classification	Cancer	BP NN	Abnormal images and 12 normal images	Holdout	96.96%acc
2010	Charisis et al. (2010)	WCE	GI tract	Intrinsic mode functions BEEMD	Classification	Ulcer	SVM	80 images of 6 patients, 40 normal and 40 abnormal frames (576 × 576)	Cross-validation 10	95%acc
2008	Giritharan et al. (2008)	WCE	GI tract	HSV histogram + dominant color, and color co-occurrence	Classification	Bleeding	SVM ensemble	550 images from 5 videos and contain 400 positive 400 negative (256 × 256)	Holdout	93%spec 80%sen

Table 9 A summary of hybrid geometric color and geometric texture features extraction methods in spatial domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation Method	Results
Color geometric										
2017	Maghsoudi (2017)	WCE	GI tract	Super-pixel based clustering	Segmentation	Polyps	39 images	Holdout	94% acc	94% acc
2015	Huang et al. (2015)	WCE	Esophagus	SIFT + CCH	Classification	Gastroesophageal reflux disease	Hierarchical heterogeneous descriptor fusion support vector machine (HHDF-SVM)	147 images	Cross-validation 10	93.2%acc, 94.9%sen and 92.6%spec
2013	Riaz et al. (2013)	CH and NBI	Stomach	Edge-maps, creaseness, and color	Classification	Cancer	Normalized cut	142 CH and 224 NBI images	None	DSC of 0.69 and 0.77
2010	Figueiredo et al. (2010)	VE	Colon	Texture and color segmentation	Segmentation	Cancer	None	None	None	None
1997	Krishnan and Goh (1997)	VE	Colon	Color	Classification	Cancer	Fuzzy based edge detection	256 × 256 9 cases	None	None

Table 9 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation Method	Results
Geometric texture										
2015	Cong et al. (2016)	VE	GI tract	Histogram, PHOG and LBP features	Classification	Gastritis, cancer, bleeding and ulcer	Unsupervised	6000 images	None	74.31%acc
2015	Dunaeva et al. (2015)	VE (NBI-ME)	GI tract	Geometric and topological features	Classification	Cancer	Ensemble method (Adaboost)	90 frames (786 × 576)	LOO-CV	90%acc
2015	Yuan et al. (2015b)	WCE	GI tract	HOG, SIFT, and LBP (color and texture silency)	Classification	Ulcer	Locality-constrained linear coding (LLC)	170 ulcer and 170 normal images of 20 patients	Cross-validation 5	92.65%acc and 94.12%sen
2015	Zhang et al. (2016)	VE	Esophagus	Gray-level or local features + Global texture features	Classification	Cancer	SVM	66 patients with early esophageal cancer(mean age 53 years; age range 21-87 years; 32 male and 34 female total 91 patients (mean age 42 years; age range 32-75 years; 39 male and 52 female)	Cross-validation 10	93%acc, 89.4 %sen and 94 %spec

(2) *Geometric and Color Information Combined with Texture Features*

(a) *Geometric Texture* In Cong et al. (2016), three types of descriptors are combined as texture color and shape (LBP, RGB histogram, and PHOG feature respectively). Additionally, deep unsupervised features'selection is performed to select important features. In the same way, point-based features such as SIFT are merged with LBP and shape-based features HOG for classification of endoscopy frames in Yuan et al. (2015b). Likewise, topological features are calculated from statistical moments of the histogram of images. Then, the geometrical features are computed to find any potential the abnormal area from ME and NBI images in Dunaeva et al. (2015). In Zhang et al. (2016), clusters are formed based on the similarity of features and a high order kernel-based graph matching algorithm is proposed. A graph is represented by a combination of nodes and edges. Where in images pixels are denoted by nodes and edges are meant by the relation of similarity between these nodes. Similar work has performed in Maghsoudi (2017) where super-pixel algorithm is used for segmentation.

(b) *Geometric Color* Color statistics such as mean, standard deviation are computed then combined with geometric parameters to detect polyps in colon images (Krishnan and Goh 1997). In the same way, heterogeneous color features are calculated by combining colors and point-based SIFT features in Huang et al. (2015). In Riaz et al. (2013), multiple features are combined for segmentation of lesions in endoscopic frames. Visual information, such as colors, edges, and textures are used to segment chromoendoscopy images. Similarly, an edge-based model is proposed named as Active Without Edges Model (ACWE) in Figueiredo et al. (2010). This method uses active-contours without edges model of Chan and Vese to segment the aberrant crypts foci to shape or structure. The prior medical knowledge confirms that the anomalous crypts'lines stain darker than normal crypts. Commonly, inside each focus, the shape of the crypts'clefts matches a similar pattern.

3.5 Summary and insights

In this section, we have discussed various feature extraction methods. These methods have been developed for the detection, classification, and summarization of gastric conditions from endoscopy videos. Moreover, these methods are based on techniques which are used to analyze images in the spatial domain of image processing. For better understanding, we have grouped these techniques in three broad categories as color, texture, and geometric features. Furthermore, these categories are expanded into subcategories according to feature extraction methods. In the context of gastrointestinal diseases diagnoses, colors are very important visual characteristics and colors play an important role in the detection of gastric ulcer, inflammation, and bleeding. On the other hand, texture features also provide a good description of malignancies like cancer, ulcer, polyp, and Celiac disease. As described earlier, in some methods, geometric features are used to establish a correlation between gastric lesions and a geometrical model. However, the selection of a suitable features extraction method for the diagnoses of gastric lesion highly dependent on its application, nature of imaging modalities, and type of gastric abnormalities. Moreover, researchers are trying to develop hybrid approaches to cope with multiple issues by combining similar or different types of features. Such as, colors are combined with texture to gain more discriminative power or colors are combined with geometric or shape-based features.

4 Features extraction techniques in frequency domain

As we know, in the spatial domain of image processing images are processed by a direct manipulation of pixels of an image. In most cases, images are used as they are without any transformation. Conversely, in frequency domain of image processing, every image is represented as combination of different frequency components (also known as Fourier analysis) or normally these components are the complex exponential. Therefore, images are not processed directly, rather they are first transformed into the frequency domain by using the Fourier transform or some other frequency transform and then features are extracted from the processed images in frequency domain. In this section, we will elaborate features extraction methods used to extract features by describing images in terms of its frequency components.

4.1 Color features in frequency domain

(1) *DWT-based Color Features* In the frequency domain, color features of endoscopic frames are extracted by transforming images by different methods. Color information is extracted through coefficient of wavelets by transformation of endoscopic image using pyramidal discrete wavelet transform (Pyramidal-DWT) in Häfner et al. (2009a). They have applied DT-DWT to the magnified endoscopic images for extraction of features containing scale, rotation, mean, and standard deviation from RGB color channels. Color Eigen sub-bands features are proposed in Kwitt and Uhl (2008). In contrast to the Pyramidal-DWT, six complex orientation sub-bands per decomposition scale are computed for feature vector construction (Kwitt et al. 2010). Where the color information is represented by extraction of features from color channels of the images in LAB color space. Then variance is calculated from de-correlated detail sub-bands of the stationary wavelet transform to represent features. Moreover, they have shown upright performance as compared with the performance DT-DWT based features. Similarly, the patch-based color features are extracted by transforming images in CIE-LAB color space and further computing DWT of three channels (Li et al. 2004). In addition, each block of endoscopic frames is divided into 16x16 blocks and the fractal dimension is computed from each block in Yamaguchi et al. (2015).

(2) *Fourier Filter Based Color Features* The images are transformed to Fourier domain later these endoscopy images are filtered using ring filters (Hafner et al. 2010a). Ring-shaped band-pass filters have been applied to get multi-scale analysis by selecting minimal and maximal ring width of 1 and 15, respectively. Similar kind of work is conducted in Vécsei et al. (2009), where ring-shaped filters of different size are used to filter images and further statistical information is computed from each channel of RGB frame. An overview of color feature extraction method has given in Table 10.

4.2 Texture features in frequency domain

As mentioned earlier, texture in spatial domain refers to characteristics of image surface (e.g., smoothness or coarseness). In the spatial domain, we have seen that most methods are based on statistical analysis of pixels or super-pixels. However, in the frequency domain, the images are processed, by first transforming an image into the frequency domain. Then, images are manipulated by performing some operations in the frequency domain to extract texture information. In most cases, statistical measures of outputs of operations are computed and represented as texture.

Table 10 A summary of color features extraction methods in frequency domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
DWT based color features										
2015	Yamaguchi et al. (2015)	VE	Esophagus	Fractal dimension + DWT	Classification	Cancer	Unsupervised	10 images (1024 × 1024)	None	None
2010	Kwitt et al. (2010)	VE	Colon	Dual-tree complex wavelet transform (DTCWT)	Retrieval	Multiple	Discriminant classifier	627 images	LOO-CV	96%acc, 94%sen and 97%spec
2009	Häfner et al. (2009a)	VE	Colon	Gaussian MRF + Color DWT pyramidal	Classification	Cancer	k-NN+ Bayes	286 neoplastic cases 198 non-neoplastic	LOO-CV	92.8%acc
2008	Kwitt and Uhl (2008)	VE	Colon	CIE-LAB, sub-band stationary wavelet transform	Classification	Polyps, Cancer	k-NN	484 images	LOO-CV	98.35%acc for 2 class and 90.50%acc for 6 class
2004	Li et al. (2004)	VE	Colon	Patch based + CIELAB + RGBLAB	Classification	Multiple	SVM + Ensemble	58 colonoscopic images, including 12 normal ones and 46 with abnormal regions	LOO-CV	83.4%acc
Fourier Filter based Color Features										
2010	Hafner et al. (2010a)	ZE	Colon	Fourier features	Classification	Cancer	SVM + discriminant analysis + Bayes	484 images of 53 patients (624 × 533)	LOO-CV	96.9%acc 2 class and 86.8%acc 6 class
2009	Vécsei et al. (2009)	ZE	Colon	Fourier features	Classification	Celiac	KNN, SVM, Bayes classifier	391 images	LOO-CV	94%acc

(1) *Statistical Features in Frequency Domain* Statistical methods of wavelet sub-bands have generally used for image classification tasks. Linear regression is used to model the descriptors of adjacent sub-bands. The regression residuals are then employed to represent the difference from a sample to a class of texture (Dong et al. 2015b). Similarly, a Heterogeneous and Incrementally Generated Histogram (HIGH) texture descriptors are modeled by wavelet coefficients by using four local features in wavelet sub-bands. Then, a non-negative multi-resolution vector (NMV) of the image is created by concatenating all sub-band textures. The low-dimensional basis of the linear subspace of NMVs is computed using Hessian regularized discriminative non-negative matrix factorization (Dong et al. 2015).

In Hassan and Haque (2015), the WCE frames are transformed through computing DFTs. Then, normalized co-occurrence matrices are computed by taking the log transform of the spectrum magnitude. Moreover, the texture features are represented by computing various statistics from the co-occurrence matrices of WCE images. Similarly, in Karkanis et al. (2001), four statistical measurements of GLCM were used by apply DWT on patches of images of different sizes, for lesion detection in endoscopy frames. A performance comparison of DT-DWT with Gabor wavelet transform (GWT) and DWT was presented in Häfner et al. (2009b). In the same way to overcome issues of shift invariance and direction selectivity in DWT a dual-tree complex wavelet transform (DT-DWT) is proposed in Häfner et al. (2008).

(2) *LBP in Frequency Domain* The images are transformed through the curve-let transformation for texture feature extraction. Furthermore, uniform-LBP are extracted form coefficient of the transformed domain to represent the texture of WCE image for ulcer classification in Li and Meng (2009c).

(3) *Gabor Filter Based Texture Analysis* In the same way, contour-let transform was performed on WCE frames and further, the log of Gabor filters was applied. Likewise, the texture features are represented by computing mean and standard deviation of filters' responses (Koshy and Gopi 2015). Gabor texture features with scale, rotation, and illumination variations are used by exploiting shift invariance properties (Riaz et al. 2011, 2012). Moreover, these texture features extraction methods are described in Table 11.

4.3 Hybrid features in frequency domain

Several types of features are combined together to create new hybrid features with additional discriminative power. These hybrid approaches are widely adopted in both domains. We have grouped these features extraction methods into a combination of basic features (e.g., color, texture, and geometric) and an overview of these hybrid approach is presented in Tables 12 and 13.

(1) *Colors Texture Features in Frequency Domain*

(a) *Gabor-based Color Texture* The local sample means and variances of the color component are combined with color channels, then Gabor- based texture features are joined with these color features (van der Sommen et al. 2014), then with same features random forest classifier used in Janse et al. (2016). Similarly, in Szczypiński et al. (2014) various color components and their ratios were used to analyze WCE images and homogeneous texture features are combined to get both color and texture information. In the same way, Log of Gabor filters (LoG) is used to extract texture features. For color texture features, LoG filters are applied to every channel of WCE frame in RGB color space (Karargyris and Bourbakis 2009b). In Coimbra et al. (2006), Gabor-based texture features are combined with scalable colors for partitioning WCE frames into various parts of GI tract. Then, again MPEG-7 features are used for classification of WCE frames (Coimbra and Cunha 2006). On the contrary, domi-

Table 11 A summary of texture features extraction methods in frequency domain for computer aided diagnosis of abnormalities in GI tract

Statistical texture features

Year	Author	Modality	Region	Feature	Task	Abnormality	Classifier	Dataset	Validation	Results
2015	Hassan and Haque (2015)	WCE	GI tract	DFT-based texture Descriptor	Classification	Bleeding	SVM	32 videos 600 bleeding and 600 non-bleeding frames 12 patients test set consisted of 860 bleeding and 860 non-bleeding frames (426 × 426)	Holdout	99.19%acc, 99.41%sen and 98.95%spec
2009	Häfner et al. (2009b)	ZE	GI tract	DT-CWT + Gabor wavelet texture	Classification	Cancer	k-NN	484 images	LOO-CV	93.39%acc
2008	Häfner et al. (2008)	ZE	Colon	Statistical + dual-tree complex wavelet transform	Classification	Cancer	1-nearest neighbor	484 images	LOO-CV	95.87%acc
2001	Karkanis et al. (2001)	VE	GI tract	DWT texture + statistical	Classification	Cancer	Multilayer feed-forward neural network	22472 images training and 32768 testing images (512 × 512)	Holdout	94%acc up to 99%acc,

LBP in wavelet domain

Year	Author	Modality	Region	Feature	Task	Abnormality	Classifier	Dataset	Validation	Results
2009	Li and Meng (2009c)	WCE	Small bowel	Texture features curvelet transform local binary pattern	Classification	Ulcer	SVM + ANN	100 images (256 × 256)	Cross-validation	93.28%acc, 92.37%spec and 91.46%sen

Table 11 continued

Gabor filter based texture analysis

Year	Author										
2015	Koshy and Gopi (2015)	WCE	GI tract	Contourlet transform and Log Gabor filter	Classification	Ulcer	SVM	137 images are used, of which 65 are ulcer and 72 are normal	Cross-validation	10	94.16%acc
2012	Riaz et al. (2012)	CH and NBI	GI tract	Gabor Texture auto-correlation for shift invariance of SGFS invariance properties of Gabor filters texton framework: texton AGF	Classification	Cancer	SVM	142 CH and 224 NBI images	Cross-validation	10	82%sen and 82%pec, 83%acc and 0.85 AUC
2011	Riaz et al. (2011)	CH and NBI	GI tract	Gabor textures	Classification	Cancer	Statistical classifier	176 CH images and 221 NBI	Cross-validation	10	83.9%acc

Table 12 A summary of hybrid color texture features extraction methods in frequency domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
Gabor-based color texture										
2016	Janse et al. (2016)	HDVE	Esophagus	Local Gabor and Color features	Classification	Cancer	RF	100 images of 39 patient data-set	LOO-CV	75%pre and 90%sen
2014	van der Sommen et al. (2014)	ZE	Esophagus	Local Color-and Gabor texture based features	Classification	Cancer	SVM	32 images (600 × 1200)	LOO-CV	95%sen 75%pre
2014	Szczypiński et al. (2014)	WCE	GI tract	Color and texture features	Classification	Ulcer, bleeding, polyps	SVM	None	None	98%sen/spec ratio
2010	Girgis et al. (2010)	WCE	Small bowel	MPEG-7 edge histogram descriptor (EHD) MPEG-7 dominant color descriptor (DCD)	Classification	Crohn's disease inflammation	SVM	47 videos 29 with disease 17 normal and 469 frames used for training 253 for validation and 277 for testing	Holdout	87%acc, 93%spec and 80%sen
2009	Bejakovic et al. (2009)	WCE	GI tract	Dominant Color descriptor homogeneous texture descriptor haralick texture features edge histogram descriptor	Classification	Crohn's disease and other	SVM	1685 frames with 188 lesions, 1231 normal images, and 266 extraneous images	Holdout	96.5%acc
2009	Karargyris and Bourbakis (2009b)	WCE	Small bowel	HSV + Gabor LOG + RGB	Classification	Bleeding and other anomalies	SVM + ANN	50 frames, 30 non ulcer images and 20 ulcer images	Holdout	75%sen and 73.3%spec
2006	Coimbra et al. (2006)	WCE	GI tract	Homogeneous texture + color	Classification	Abnormality (summery)	SVM	30,000 images	Cross-validation 3	90%acc

Table 12 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2006	Coimbra and Cunha (2006)	WCE	GI tract	MPEG-7 + color structure, local edge histogram, DCT and homogeneous texture	Classification	Ulcer, Bleeding, Polyps	Baysian	Ulcers (400 images from 21 events), polyps (899 images from 31 events) and 55,000 normal images blood (832 images from 17 events)	Cross-validation 2	87%acc
GLCM-based color texture										
2012	Sobri et al. (2012)	VE	GI tract	Texture and color + WT GLCM + texture features then combined with the features of Color Moment for image classification	Classification	Abnormality	SVM	220 ROIs	Holdout	87.3%acc
2009	Häfner et al. (2009c)	ZE	Colon	Wavelet texture + Color	Classification	Polyp	k-NN	484 images	LOO-CV	99.38%acc
2009	Bonnel et al. (2009)	WCE	Small bowel	Wavelet texture + color + SIDWT	Classification	Multiple	1D classifier	75 images 41 normal and 34 abnormal cases	LOO-CV	94.7%acc
2008	Lima et al. (2008)	WCE	Small bowel	Color wavelet covariance	Classification	Polyps, ulcer	NN RBF	2000 normal images, 23 bleeding images, 54 polyp images, 123 tumor images and 58 a priori undefined images, test set is about 85000	Holdout	93%sen, 95%spec and 94%acc

Table 12 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2008	Barbosa (2008)	WCE	Small bowel	Color channel wavelet transformation	Classification	Cancer	MLP NN	204 images for training 192 for testing 92 abnormal (256 × 256)	Holdout	98.7%sen and 96.6%spec
LBP color texture in wavelet domain										
2015	Mathew and Gopi (2015)	WCE	GI tract	LBP + CIE XYZ colour	Classification	Bleeding	k-NN	256 × 256 332 WCE images comprising 145 bleeding images and 187 normal images	Cross-validation 10	96.38%
2010	Li and Meng (2010a)	WCE	GI tract	LBP + DWT + color	Classification	Cancer	SVM	600 representative tumor CE images and 600 normal CE images from 12 patients' video	Cross-validation 2	88.3%acc
2010	Li and Meng (2010b)	WCE	GI tract	ULBP + Gabor texture + DWT	Classification	Bleeding and other anomalies	SVM	100 normal CE images, 100 tumor CE images, 100 bleeding CE images and 100 ulcer (256 × 256)	Cross-validation 2	100%acc tumor, 77%acc bleeding, 83%acc normal and 89%acc ulcer
2009	Li and Meng (2009d)	WCE	Small bowel	Color wavelet covariance	Classification	Cancer	SVM	300 images 150 abnormal and frames 150 normal (512 × 512)	Cross-validation 3	97.33%sen, 97.67%spec and 96.67%acc
2006	Iakovidis et al. (2006)	VE	GI tract	LBP + color + CWT	Classification	Cancer	SVM	6000 colon, 2600 gastric images for training and 9000 colon and 3900 gastric images for testing (320 × 240)	Holdout	94%acc

Table 12 continued

Statistical color texture in wavelet domain

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
2016	Serpa-Andrade et al. (2016)	VE	Esophagus	Fourier descriptor + HIS + statistical	Classification	Cancer	k-NN and RF	280 images 26 real cases of patients (10 healthy and 16 suffering from the disease)	Holdout	81%pre 86%sen and 72%spec
2010	Hafner et al. (2010b)	ZE	Colon	DT-CWT	Classification	Polyps	1-NN classifier	627 images	LOO-CV	99.4%acc
2010	Martins et al. (2010)	WCE	GI tract	DWT + DCT + color co-variance transform	Classification	Ulcer	GMM + HMM	600 images, 400 frames belong to normal and frames 200 abnormal class	Holdout	99%sen and 95.19%spec
2009	Barbosa et al. (2009)	WCE	Small bowel	3C	Classification	Cancer	MLP NN	600 frames and 400 normal frames 200 abnormal frames	Holdout	97.2%sen and 97.4%spec
2008	Vécsei et al. (2008)	VE	GI tract	Histogram + wavelet	Classification	Celiac disease	SVM + kNN + Bayes	391 Images	LOO-CV	98.5%acc and 100%sen
2008	Huang et al. (2008)	VE	GI tract	color RGB HIS YCB + wavelet	Classification	Helicobacter pylori + Ulcer + nonulcer dyspepsia	SVM	236 patients 130 of the study patients were defined as H. pylori-infected	Cross-validation 10	90%acc
2005	Iakovidis et al. (2005)	VE	GI tract	Texture spectrum, color Histogram LBP histogram color wavelet covariance	Classification	Polyps	Gaussian kernel SVM	1000 images (320 × 240)	Cross-validation 10	0.886AUC

Table 13 A summary of hybrid geometric color, and texture features extraction methods in frequency domain for computer aided diagnosis of abnormalities in GI tract

Year	Refs.	Endoscopy	GI area	Features	Application	Disease	Classifier	Dataset	Validation method	Results
Geometric texture features										
2018	Ali et al. (2018)	CH	Stomach	Gabor texture+ GLCM	Classification	Cancer	Multiple classifiers	176 images	Cross-validation 10	87.2% acc
2010	Hwang and Celebi (2010)	WCE	Colon	Gabor texture+ K-mean and Geometric features	Classification	Polyps	Unsupervised	128 images, 64 images with polyps and 64 normal	None	100%sen 81%spec
2009	Karargyris and Bourbakis (2009a)	WCE	Small bowel	SUSAN edge detector and log Gabor	Classification	Polyps	Fuzzy-SVM	Containing 10 with polyps and 40 normal total 50 WCE video frames	Holdout	100%sen and 67.5%spec
Geometric color features										
2015	Drozdzal et al. (2015)	WCE	Small bowel	Automatic methods for motility features	Segmentation	Cancer	Unsupervised	10 videos 256 × 256	None	None
2008	Alexandre et al. (2009)	VE	Colon	Color + Position + CWT + LBP	Classification	Polyps	SVM RBF-kernal	4620 images	Cross-validation 10	94.87AUC
Combined texture										
2014	Nawarathna et al. (2014)	WCE	Colon	Leung-Malik and LBP	Classification	Ulcer, bleeding, polyps	k-NN	800 images	Cross-validation 10	92%sen and 91.8%spec for WCE frames, 91%sen and 90.8% spec for VE

nant color descriptors and edge histogram features from the combination of MPEG-7 features were extracted. Additionally, these features were used for detection of Crohn's disease (Girgis et al. 2010). Gabor-based (homogeneous texture features) feature and other MPEG-7 features are used in the classification of WCE images with Crohn's disease (Bejakovic et al. 2009).

(b) *GLCM-based Color Texture* GLCM is extended in the frequency domain and wavelet cross co-occurrence (WCC) matrices have been proposed by Häfner et al. (2009c) where LUV color space is used to extract color and texture characteristics from every image by computing the statistics of GLCM which is computed from sub-bands of wavelets. In the same way, Bonnel et al. (2009) computed GLCM from wavelet domain and color information have mixed by exploiting RGB color channels. Similarly, GLCM based texture features are extracted from sub-bands of DWT by transforming every channel of image (Lima et al. 2008). In Barbosa (2008), GLCM features were extracted by transforming images by taking DWT of WCE frames, then computing all statistical measures from images in RGB and HSV color spaces. Correspondingly, wavelet-based GLCM is computed from images and then color moments are computed from every channel to form features' set in Sobri et al. (2012).

(c) *LBP Color Texture in Wavelet Domain* Images are converted from RGB color space to CIE-XYZ. Further, the LBP features are extracted from the contour-let transform of WCE images in Mathew and Gopi (2015). However, HSI and RGB color spaces are also tested for detection of bleeding. Although, features extracted from the transformation of images in CIE-XYZ color space have good performance. Likewise, in Iakovidis et al. (2006) color channels of an endoscopic frame first transformed by DFT and then GLCM is calculated in the wavelet domain. Furthermore, opponent color-local binary pattern (OC-LBP) features are extracted by each color channel and their intra color histograms. In Li and Meng (2010a), middle-level sub-band images result from DWT of images are used to extract texture information by LBP features. Furthermore, LBP features are extracted from three color channels of sub-band images. Similarly, in Li and Meng (2009d) DWT-based LBP are extracted from RGB and HSI and used for WCE images for classification. Moreover, in Li and Meng (2010b), color information is included by taking YCbCr color space into account for extraction of DWT-based LBP features for a color texture representation of endoscopic images.

(d) *Statistical Color Texture in Wavelet Domain* Second order statistics are computed from the Color Curvelet Covariance (3C) of images, in 3C images are transformed by DCT and converting images from RGB to the HSV color space. Furthermore, the coefficients are modeled by Gaussian Mixture Model (GMM) (Martins et al. 2010). As the same in Barbosa et al. (2009), where 3C is used to form the features' set for detection of tumors in small-bowel. In Serpa-Andrade et al. (2016), a combination of statistical features, texture features with color are combined, texture features using DCT in HSI color space are extracted. The final feature set is formed by combining of the Hu moment and Fourier descriptors. ME frames are transformed from DT-CWT then, texture features are computed from six level sub-band images. Furthermore, statistics or Wiebull parameter are computed for representation of the pit-pattern characteristics of gastric lesions in Hafner et al. (2010b). In the same way, texture information is combined with features of wavelet domain by computing statistics from sub-bands of transformed images for detection of the frames with Celiac disease (Vécsei et al. 2008). Statistical measures like, mean, variance, standard deviation, skewness, and kurtosis statistics were computed from sub-bands of images. Moreover, second order statistics include entropy, energy, inverse difference moment, contrast, and co-variance are computed from sub-bands. A color channel his-

togram has been used for merging color information into textures. In Iakovidis et al. (2005), 3C are used for texture color extraction by using discrete wavelet frame transform (DWFT) instead of DWT. As well colors features are included by computing features from three color channels of the images. For completing this task, endoscopic images are converted to multiple color spaces (RGB color image was converted to HSV and YCbCr color spaces), then these images in different color spaces are transformed with DWT (Huang et al. 2008).

(2) *Geometric Information Combined with Texture Features* A watershed-based method for segmentation of polyps is proposed in Hwang and Celebi (2010). However, a marker selection technique proposed by combining Gabor texture and k-mean clustering for polyp shape analysis. Motility is a term used to represent the contraction of the muscles that unite and drive contents in the gastrointestinal (GI) tract. Therefore, in Drozdzal et al. (2015) contraction of GI tract is detected for classification of a specific area in WCE frames, for edges detection, Gabor filters are used. The contraction detector and lumen size estimation are used for detection of stable motility (Drozdzal et al. 2015). SUSAN's edge detector and LoG is used for detection of edges of polyps in colonoscopy frames and produce crisp segments (Karargyris and Bourbakis 2009a). Furthermore, geometric parameters like center curvature are used for clustering and segmenting the polyps. Grid-based color and position information is merged and compared with other features extraction techniques like LBP and CWC, for the detection of polyps (Alexandre et al. 2009).

(3) *Texture Features Combined with Other textures* In some studies, texture features are combined with other type texture features to increase the discrimination power.

A multi-scale rotation-invariant representation (MRIR) texture based on multiscale wavelet transform splits the magnitude pattern (MP) mapping of texture and the sign pattern (SP) mapping of texture employed as a step function. The step function fits the wavelet sub-bands of the MP for calculating the sampled directional mean vectors (SDMVs) of the sub-bands and concatenated with frequency vectors (FVs) of SP mappings to form MRIR vector to get textural representation (Dong et al. 2017). Likewise, in Nawarathna et al. (2014), texture features extracted from Gaussian filters of multiple shapes and sizes are combined with LBP features named as Leung-Malik LBP (LM-LBP). These features are then used for detection of multiple pathological conditions from endoscopy images.

4.4 Summary and insights

In this section, we have given an extensive review of features extraction methods which process images in the frequency domain. In frequency domain, images are first transformed to Fourier domain where images are represented in terms of the frequency component. Furthermore, by processing frequency components of images and different features can be extracted from images. These methods well-performed for detection of abnormalities from endoscopy. Yet, frequency methods are computational extensive due to the overhead of transformation. Even though by using fast and efficient methods of transformation can reduce the computational complexity of these methods. For instance, we have seen the color texture approaches in the frequency domain have gained much focus of researchers. In contrast, geometrical methods not efficiently applied in the frequency domain. Since, in the frequency domain, geometric characteristics cannot be preserved. Whereas from gastric abnormalities, polyps and contraction of gastric tract have a specific geometric shape and therefore can be easily targeted in the spatial domain.

5 Feature learning methods

In previous sections, we have categories features extraction methods based on their respective domain. In this section, we will discuss some features learning approaches, where the feature are extracted and represented without any explicit method. Although these features are calculated based on some input parameters. Features learning or representation learning methods automatically find representation from input frames, which is needed for classification or recognition tasks. It does not requires to explicitly defining feature set or attributes. Several features learning methods have been developed for the fact that images are preprocessed for classification or learning tasks and in some recognition problems, the features can be defined explicitly. Thus, its beneficial to discovering representation of images based on input data without hard-coding the features extraction algorithms. These methods can be divided into two categories: first the bag of features (BOF) approach where a dictionary of visual words is learned by some already existing specific feature extraction method. Secondly, deep learning based features extraction methods where the power of the neural network is employed for extraction of important characteristics of images. A summary of these methods has been given in Table 14.

5.1 Bag of features

BOF model is widely used in various classification tasks. In BOF, features of images are treated as word in documents. A dictionary or code-book is learned for computing occurrence of each visual word exists in the images. Furthermore, this histogram is used as a feature vector for every image (Yuan et al. 2017b). A resembling work is done in Gueye et al. (2015), where SIFT features are extracted from every frame and then a dictionary is formed by collecting distinct visual words from every image. The vocabulary is formed by a clustering algorithm like k-means. Furthermore, a histogram of these words from every image is computed to represent features by means of vector quantization. Moreover, these features are used for classification of colon images for abnormalities. Similarly, densely sampled SIFT features are computed from local features with the BOF model for the classification of lesion patterns in endoscopic frames in Miyaki et al. (2015). K-mean is used for forming visual vocabulary and used to quantize feature set for each image. Furthermore, cancerous regions and normal surrounding tissue patches are used to form a code-book. Likewise in André et al. (2010), BOF approach is used with a dense detector and a bi-scale SIFT description for retrieval of pCLE imagery from an image database.

In Yuan et al. (2015a), visual word based color histogram features from RGB, HSV, YCbCr, CMYK, and LAB color spaces are tested for bleeding detection from endoscopic frames. Moreover, two level of saliency is used for extraction of the bleeding area. The illumination of elements have a good separation in CIE-LUV color space from colors. Therefore, the normalization in colors, illumination component L is filtered using a homomorphic filtering. Furthermore, a vocabulary is formed by clustering visual words, and then the adaptive color histogram is formed by means of these color words (Riaz et al. 2017).

(1) *BOF of Texture (Texton)* Texton refers to a visual vocabulary of words formed with texture features as described in Riaz et al. (2012), where Gabor texture features are used to create a dictionary of texture features. Similarly, in Lung Malik based Gabor filter's bank is used by Nawarathna et al. (2014) to extract texton from WCE for detection containing multiple abnormalities. Moreover, they have mixed LBP features with texton using the code-book model.

Table 14 Features learning methods for computer-aided diagnoses of abnormalities in GI tract

Bag of features based methods

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Dataset	Validation method	Results
2017	Yuan et al. (2017b)	WCE	GI tract	Colors with BOF model	Classification	Multiple	SVM	1680 images	Cross-validation 5	90.78%acc
2017	Riaz et al. (2017)	CH	GI tract	Color CIE-LUV adapted histogram with BOF model	Classification	Cancer	Multiple	130 images	Cross-validation 10	88%sen and 0.93AUC
2016	Wang et al. (2016a)	VE	Esophagus	Colors with BOF model	Classification	Multiple	Online metric learning	12000 images of 424 peoples	Cross-validation 10	0.93AUC
2015	Miyaki et al. (2015)	NBI	GI tract	SIFT with BOF model	Classification	Cancer	SVM	587 patches of gastric cancer and 503 patches images of gastric skin	Holdout	86.5%acc
2015	Yuan et al. (2015a)	WCE	GI tract	Color with BOF model YCbCr historam	Classification	Bleeding	k-NN + SVM	400 frames with bleeding and 2000 normal frames (256 × 256) total 2400 WCE frames	Cross-validation 10	95.75%acc, 92%sen, 96.5%spec and 0.97AUC

Table 14 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Dataset	Validation method	Results
2015	Gueye et al. (2015)	WCE	Colon	SIFT with BOF model	Classification	Multiple	SVM	450 normal, 450 abnormal and second data-set 600 normal-600 abnormal of WCE images	Holdout	98.25%acc
2010	André et al. (2010)	CLE	GI tract	Dense SIFT with BOF Model	Retrieval	Retrieval	k-NN	1036 images	LOO-CV	89%acc
Deep learning and ANN based methods										
2018	Mahmood et al. (2018)	VE and synthetic	GI tract	FCN	Classification	Multiple	ANN	20,000 images	Holdout	78.7%acc
2018	Iakovidis et al. (2018)	VE and WCE	GI tract	WCNN	Classification	Multiple	ANN	10,000 images	Holdout	96%acc
2018	Hirasawa et al. (2018)	VE	GI tract	CNN	Classification	Cancer	ANN	130,00 images	Holdout	92%sen and 98%acc
2018	Turan et al. (2018)	WCE	GI tract	Recurrent CNN	Summary	Trajectory	ANN	40,000 images	Holdout	None
2018	He et al. (2018)	WCE	GI tract	CNN	Classification	Hookworms	ANN	440,000 images	Holdout	88.5% acc
2017	Zhang et al. (2017)	WCE	Colon	CNN	Classification	Polyps	SVM	1970 VE 332 NBI images	Cross-validation 3	99.4% acc

Table 14 continued

Year	Refs.	Endoscopy	GI area	Features	Application	Diseases	Classifier	Dataset	Validation method	Results
2017	Garcia et al. (2017)	CLE	GI tract	CNN	Classification	Polyps	ANN	3257 images	Holdout	94.0% acc
2017	Yuan et al. (2017a)	Colonoscopy	Colon	CNN	Classification	Polyps	ANN	3463 frames	Holdout	91.09% acc
2017	Nan et al. (2017)	CLE	GI tract	Patch-based CNN	Segmentation	Cancer	Ensemble	1400 images	Holdout	91.09% acc
2015	Yu et al. (2015)	WCE	GI tract	CNN	Classification	Multiple	NELM	60,000 images for training and 15,000 images for testing (480×480)	Holdout	97.25%acc
2015	Zou et al. (2015)	WCE	GI tract	CNN	Summary	Multiple	DCNN	25 cases 60,000 training images 15,000 testing images (480×480)	Holdout	95%acc
2011	Pan et al. (2011)	WCE	GI tract	Color features using PNN	Classification	Bleeding	Probabilistic neural network	50000 pairs (768×530)	Holdout	93.1%sen and 85.6%spec
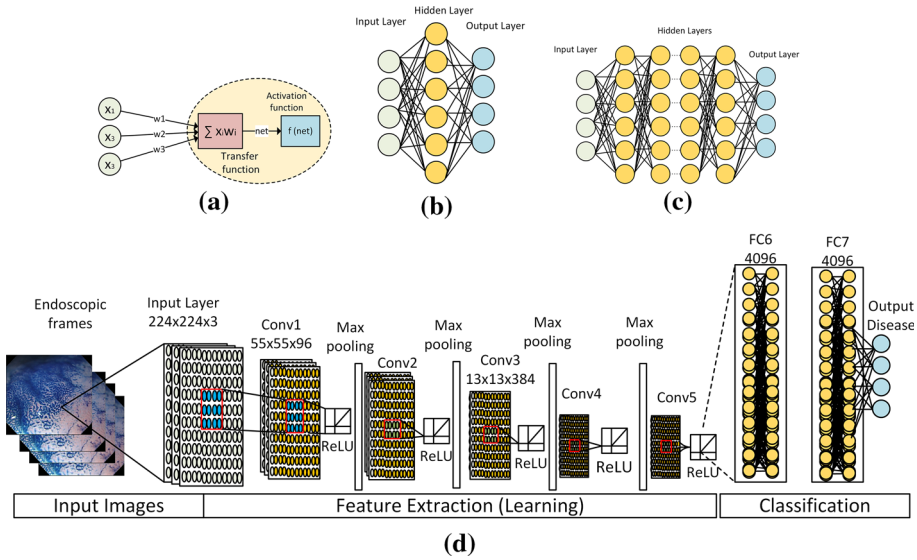
Fig. 6 Concepts of artificial neural network (ANN) and deep learning: **a** a basic single layer neural network with activation function. Normally the activation function is a sigmoid real-value function or ReLU for better convergence, **b** a standard ANN with one hidden layer, **c** an abstract model of deep ANN with many hidden layers, **d** an example of deep convolutional neural network (some contents adapted from Badrinarayanan et al. 2017; Schmidhuber 2015; Hu et al. 2018; Krizhevsky et al. 2012)

In BOF methods, we have to explicitly define the parameters for learning the representation (for example size of bins of the histogram and the types of features which are extracted from images). However, there are methods which does not require explicit type definition of descriptors. These methods learn the representation based on the input data. Next, we will discuss the applications of a revived field of Artificial Neural Networks in the extraction of features and segmentation.

5.2 Deep learning methods

A simple Artificial Neural Network (ANN) is a network of connected processing units called neurons. A simple neuron is a real-valued threshold based function which gets activated through input value received from weights input links. The intensity of each link is multiplied with its weight and weighted sum of input links is passed to a neuron which has a threshold value as shown in Fig. 6a. Whenever the input value of a neuron is more than its threshold value, its activated (Schmidhuber 2015). An ANN with a single layer of neurons only can learns a simple function like the classification of data into only two categories (a linearly separable data). For more complex tasks, hidden layers play their role in providing flexibility in learning more non-linear decision boundary as shown in Fig. 6b.

In the early days of ANN, the addition of more hidden layers is restricted by the available computation power, number of weights learned by an ANN also increased with number of hidden units therefore, a huge amount of parameters are learned during a training session of ANN. Thus, it is computation extensive task and such systems was not easily available at that time. The other limitation of training a Deep ANN was the availability of large data-sets for training because small data-set cause under-fitting of the Deep ANN. Deep Learning is

a recent trend in the field of ANN which has revolutionized almost every area of life. Deep Learning is performed by increasing the hidden layer in an ANN as shown in Fig. 6b.

Deep learning can be used in many fields. Therefore, choosing the number of hidden layers, type of hidden layers, their connections type, and output layer units are application dependent. Deep learning can be generative or unsupervised when non-labeled data or target class is not available (e.g., Deep Belief Networks (DBN), Restricted Boltzmann Machine (RBM), Deep Boltzmann Machine (DBM), regularized Auto-encoders, etc.). The supervised (discriminative) Deep learning models are useful when we have class labels with data (e.g., Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) etc). The hybrid models of Deep learning also exist, for instance, for speech recognition the output probabilities of a neural network into a Hidden Markov Model (HMM) (Deng and Yu 2014).

In medical image analysis four successful deep models are as follows: CNN, Fully convolutional network (FCN), AE, and DBN (Hu et al. 2018). A CNN is based on convolutional layers ReLU activation function layer, pooling layer (max pooling can be average pooling), and fully connected layers as shown in Fig. 6d. Convolutional layers learns representations which amplify aspects of the input that is important for discrimination suppress irrelevant variations. For example, endoscopy frame are composed of pixels and each pixel from each channel of frame fed to the input layer of CNN. Then, the first layer of CNN normally learns features related to edges in a particular location and orientation thus provide translation and rotation invariant description of images. The pooling layers are used for down-sampling features ultimately reduces the dimension of the feature representation. Similarly, the Auto-encoder is a unsupervised model for learning feature with a low dimension (Badrinarayanan et al. 2017).

In FCN model up-sampling is used instead of down-sampling and the de-convolutional layer is used. FCN normally used for pixel classification (segmentation). Generally, CNN has shown its excellent performance in image recognition problems. However, the input of CNN structure is bounded by comparatively small images due to the fully connected layers (a huge number of weights). It reduces its ability to be directly applied to large input images. Alternatively, FCN does not owns any fully-connected (FC) layer and it can be applied to images of virtually any sizes compared to CNN (LeCun et al. 2015).

Sources of parsimony in the deep neural networks in object recognition is due to variations of extraneous factors in input images, such as scale, area and angle variations (Kondor 2008). These sources of deformations can be represented by symmetry groups (A symmetry group is a set of transformations that preserve the identity of an object and obey the group axioms) (Gens and Domingos 2014). Therefore, these sets of composable variations preserve the information of target class. The deep convolutional neural networks can have shift invariance by computing descriptors by using weights model in each part of the frames. However, a convolution layer uses fewer parameters than a fully connected layer and preserves many useful transformations but the CNN fails or minimally cope other groups of symmetries (Cohen and Welling 2016). Other group of symmetries, such as smoothness, adaptability, generality, equivariance/invariance, depend on restrictions imposed during learning (Anselmi et al. 2017).

In the classification task, transformation symmetries express equivalence classes that record part of the intraclass variations. It also keeps the output (class labeling) distributions by implying a quotient space up to transformation, points remain equal and representation in invariant to transformation. However, the hypothesis space is restricted to this quotient description space is essential for learning from high-dimensional data by decreasing the examples' complexity of training (the size of the labeled training set) (Gu et al. 2015). By using pooling and convolution, CNN has explicit parametrization for translation equivariance

and robustness (filtering with local kernels and pooling). The pooling gradually decreases the dimensionality to reduce the number of parameters and calculation in Deep CNN. This lessens the training time and managing over-fitting. After pooling, dimensionality should not be too high or too low. When dimensionality after dimension reduction is too large, dimension reduction is meaningless. Many vital features will be discarded in the process when dimensionality after dimension reduction is too small (Cao et al. 2019; Mallat 2016).

There are multiple theories regarding the cause of translation invariance in CNN. One idea is that translation invariance is due to the increasing receptive field size of neurons in successive convolution layers. Another possibility is that invariance is due to the pooling operation. Some suggest that it is due to data augmentation while training CNN. There is widespread consensus in the literature that CNNs are capable of learning translation-invariant representations (Kauderer-Abrams 2017; Mallat 2016).

By using weight sharing, CNN can learn more complex transformation beyond translations by explicitly learning the symmetry or convolution group when new data is provided for training. It learns different properties of representation for instance sparsity, weight-sharing topologies and locality rather of handcrafting them (Anselmi et al. 2017).

In practice, designing a desirable model requires trial and error. The design of the deep neural network (types of layers, number of layers, number of units in a single layer, connection setting, activation functions and various training parameters) are not the only decisions we have to make; also the optimization algorithm and its parameters interplay tightly with these choices. The specific dataset and the chosen loss function also define the loss surface along which we want to optimize. There are a lot of hyperparameters involve in design and infinitely many ways to create a deep neural network. Therefore, it is not feasible to automate neural network model selection. We design it manually and every one has its own way of designing a deep neural network. A better way is to mimic the design of a model that has been developed for a similar application and tweak according to our requirements.

As CNN architectures are covariant to translations with convolutions, CNN can linearize the operation of very complex nonlinear transformations in high dimensions. To calculate invariants to shifts and linearize diffeomorphisms, different scales can be separated and non-linearity is applied by cascading filters, computing a wavelet transform and point-wise contractive nonlinearity. Linearization is a strategy used in machine learning to reduce the dimension with a linear projector. CNN gradually contract the representation space. Such operations are defined by linear operators which belong to groups of local symmetries. We can avoid the curse the variability of input data, the capacity to approximate the output class. Moreover, the various group of symmetries discussed in Anselmi et al. (2017); Dieleman et al. (2016).

Similar models use power of the Deep NN to learn a representation model for endoscopy frames (Pogorelov et al. 2017). The abstract level of understanding or representation is created automatically in the hidden layer, where each layer contains different level of abstraction. The images are directly fed to neural network moreover, a large number of annotated images are needed for training the ANN. However, the images computer generated images can be used for training the of convolutional neural network (CNN) as suggested by Ahn et al. (2018); Mahmood et al. (2018). Normally, the training procedure is computationally exhaustive and requires lots of resources. On the contrary, a CNN was employed for learning features from WCE in Yu et al. (2015). It is a hybrid method (named as HCNN-NELM) where a CNN and extreme learning machine (ELM) are combined for features learning and classification tasks. The CNN layers are used to extract visual information at different abstract levels. Additionally, this information is used for classifier's learning tasks using ELM. In the same way, deep CNN is used for classification of digestive organs of WCE frames (Zou et al. 2015). In addition, the SVM classifier was trained on extracted features (learned by the CNN). In Pan et al.

(2011), bleeding frames are detected using a probabilistic neural network (PNN). Here, the color features are extracted from directional pixel values of individual channels of RGB and HSI color spaces. The best part of the deep learning models is they can be used easily for other similar recognition applications using transfer learning (Sevakula et al. 2018). As hookworms (He et al. 2018), polyps (Zhang et al. 2017), and cancer (Hirasawa et al. 2018) is detected in WCE images by using CNN. However, (Turan et al. 2018) used recurrent CNN for estimating the trajectory of wireless capsule in the GI tract. Laser based endomicroscopy images are analyzed with CNNs (Garcia et al. 2017; Nan et al. 2017) for detection of abnormal areas. AlexNet is used in Yuan et al. (2017a) for classification of polyps images and above 90% accuracy is achieved. In some methods, the gastric cancer is detected by modalities other than endoscopy as in Gibson et al. (2018). CT-scan images are used and multiple Deep NN models along with proposed V-Dense network model. Similarly, a FCN based CNN-CRF is trained on synthetic data-set along with real data-set due to unavailability of data due to privacy issue in Mahmood et al. (2018) In Iakovidis et al. (2018), a Weakly Supervised Convolutional Neural Network (WCNN) is proposed with Deep Saliency Detection (DSD) algorithm. The localization is performed and 96% highest accuracy is achieved on VE and 88% on WCE frames.

5.3 Summary and insights

In this section, we have discussed several automatic feature learning methods. As we discussed earlier, in features learning methods features are learned based on images data. Furthermore, these important feature then extracted from data for classification or image retrieval task. In BOF model, the extracted features can be any kind of features texture, colors, or point-based local features as SIFT. Moreover, these features can be extracted by using both domains as in Texton features. On the other hand, deep learning methods use the power of multi-layer ANN for learning abstract information from images. CNN have multiple layers each layer is for a different level of abstraction. Although, these methods are better in performance. However, the training of CNN is computationally intensive and require special hardware. Moreover, A large amount of annotated images are also required for training CNNs.

6 Trends, challenges, and future research directions

Advancements in surgical-vision techniques have revolutionized the surgical procedures and ultimately provides computer-assisted interventions. A successful CADx system requires efficient features extraction and image representation methods. Therefore, there is a need to design such features extraction methods those deal with dynamics of the gastric environment and provides a better description of the gastric lesion.

In this section, we review the option available for enhancing the performance of existing systems.

6.1 Hybridization and fusion of features

It is apparent from the extensive literature review, that much of features extraction methods have a tendency towards development hybrid features as shown in Fig. 7. In both domains, single visual characteristics of gastric images are not much developed because of uncertain nature of gastric lesions. Moreover, there are many methods, we have found in literature which are composed of basic color and texture information extraction. The least development has
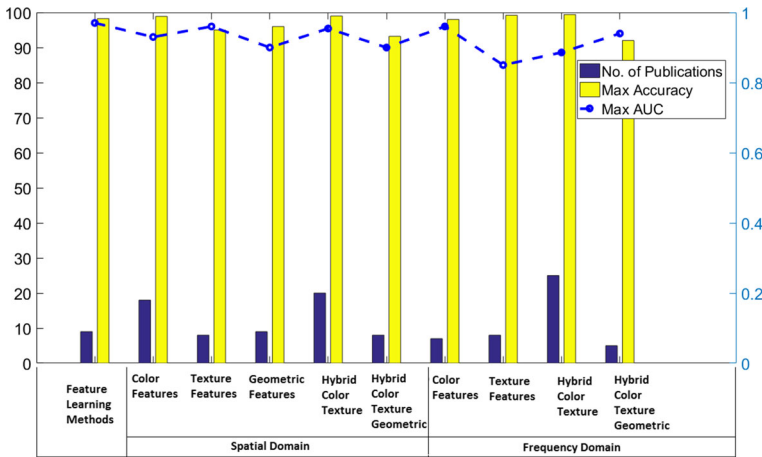
Fig. 7 A study of post-1997 publications in the computer aided diagnosis of endoscopic images. Normalized trends in publications containing phrases "gastrointestinal", "abnormality detection", "endoscopy", "feature extraction", "image-retrieval", and "classification", containing IEEE, Elsevier, and Springer Publications. It shows the tendency of researchers toward development of features extraction methods specific domains and it is clear from this graph that hybridization of features gaining much attention of researchers. Also, maximum accuracy and AUC achieved by these features extraction methods in CADx depicted in the graph

been made in the extraction of geometric information along with texture and color feature. However, the extraction of features depends on nature of abnormalities. Even though, features can be combined to deal with multiple diseased conditions in endoscopic gastrointestinal frames and able to cope with various imaging conditions.

6.2 Usage of endoscopic technologies

The literature review confirms that the WCE is an emerging technology and now has widely used by the practitioners for the screening of whole GI tract and specifically small-intestine as shown in Fig. 8. It is fact that WCE has uncontrolled movement and screen procedure is normally unattended, a large number of frames are generated and a few frames are useful for a gastroenterologist. Therefore, more methods are developed for detection of abnormality from WCE images. Flexible wired VE is more used for screening the GI parts which are easily accessible e.g., esophagus and colon. However, the enhancements are normally used with VE because of presence of instrument channels and more control over movement of the camera.

6.3 Dynamics of images acquisition

In the gastrointestinal environment, ideal conditions for image acquisition are very rare. On the other hand, camera distortion and specular reflections are very common in an endoscopic frame sequence. Moreover, uncontrolled movement of endoscopic camera leads to scale, rotation, and illumination invariance. Presence of air bubbles, poor cleansing, the presence of food, and gastric juices are some challenges. Gastrointestinal area variations have poses novel challenges to the automatic detection of gastric diseases. The accuracies of existing CADx systems were presented in Table 15. It shows a changed performances for diagnoses of images in different GI areas.
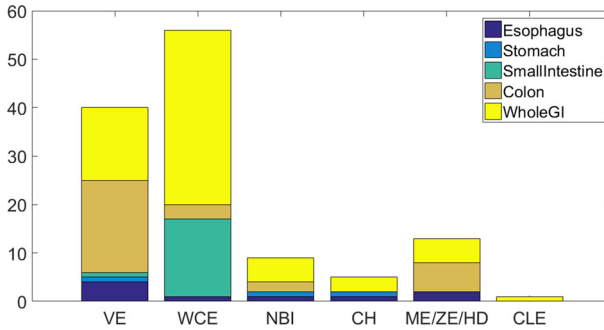
Fig. 8 Number of articles shows the usage of imaging modalities for automated detection of abnormalities in specific GI areas. Post-1997 publications in the area of computer aided diagnosis of endoscopic images

6.4 Insufficient color space

Gastrointestinal color images do not possess a sufficient color space to provide a better discrimination for abnormal regions. However, for detection of ulcer and bleeding in frames, colors have a significant importance. Specifically, in imaging technologies like CH and NBI, the involvement of colors in the detection of lesions will also increase. Consequently, owning a lack of color space utilization, colors cannot be completely ignored. Transformation of images in other color spaces may have increased the discrimination power of color features. In literature, we have seen, that many color spaces are employed for different discrimination tasks.

6.5 Color space transformations

Color spaces are basic constructs to represent images. The selection of the best color space to represent an image is a difficult and application dependent task. However, extensive analysis can be conducted to figure out color space which can help in diagnoses of gastric diseases. Many researchers have conducted experimentation for classification and segmentation of gastric images by transforming images into different color spaces prior to feature extraction task (Riaz et al. 2017). The usage of HSV and CIE-LAB are in many cases worked well due to their color uniformity and better separation over the whole visible light spectrum (Sousa et al. 2009). Still we are not sure which color space better represents the heterogeneous endoscopy images. However, analysis of feature extraction techniques can be conducted by transforming images into different color spaces and analyzing their discrimination power for a specific endoscopy imaging technology or disease.

6.6 Gastric lesions with a specific geometric structure

Gastric lesions do not have any specific shape or geometric structure. However, some lesions like polyps have an elliptical shape. Moreover, variations in gastrointestinal environment effect the shape of polyps and often it appears to be random. Therefore, it is become more challenging to geometrically model any gastric lesion in terms of specific descriptors.

Table 15 Accuracy of computer aided diagnoses methods according with respect to area of GI tract

Accuracy(%)	Esophagus	Stomach	Small-bowel	Colon	Whole GI
100–95	–	Wang et al. (2016b)	Li and Meng (2009d)	Gueye et al. (2015), Krishman et al. (1998), Maroulis et al. (2003), Adler et al. (2012), Kwitt et al. (2010), Kwitt and Uhl (2008), Häfner et al. (2010a), Häfner et al. (2008), Häfner et al. (2009c), Häfner et al. (2010b)	Hirasawa et al. (2018), Turan et al. (2018), Zhang et al. (2017), Garcia et al. (2017), Nan et al. (2017), Yuan et al. (2017b), Yuan et al. (2015a), Yu et al. (2015), Zou et al. (2015), Ghosh et al. (2015), Poh et al. (2010), Vieira et al. (2015), Cui et al. (2010), Yuan and Meng (2015), Hwang et al. (2007), Kodogiannis et al. (2007), Letter (2007), Charisis et al. (2010), Hassan and Haque (2015), Karkanis et al. (2001), Bejakovic et al. (2009), Mathew and Gopi (2015), Vécsei et al. (2008)
95–90	Huang et al. (2015), Zhang et al. (2016)	–	Constantinescu et al. (2015), Boschetto et al. (2016), Li and Meng (2012), Li and Meng (2009c), Bonnel et al. (2009), Lima et al. (2008)	Dahal et al. (2015), Gross et al. (2009), Magoulas et al. (2004), Häfner et al. (2009a), Vécsei et al. (2009)	Iakovidis et al. (2015), Zhao et al. (2015), Sousa et al. (2009), Kodogiannis and Lygouras (2008), Giritharan et al. (2008), Dunaeva et al. (2015), Yuan et al. (2015b), Häfner et al. (2009b), Koshy and Gopi (2015), Coimbra et al. (2006), Iakovidis et al. (2006), Huang et al. (2008)
89–85	–	–	Girgis et al. (2010)	Häfner et al. (2007)	Miyaki et al. (2015), André et al. (2010), Emam et al. (2015), Liu et al. (2015) Coimbra and Cunha (2006), Sobri et al. (2012), Li and Meng (2010a), Li and Meng (2010b)
85–80	–	–	–	Häfner et al. (2015), Ševo et al. (2016), Li et al. (2004)	Zhang et al. (2009), Riaz et al. (2012), Riaz et al. (2011)
≤ 79	–	–	–	–	Martinez-Herrera et al. (2016), Li and Meng (2009b), Cong et al. (2016)

6.7 Designing more generic and image-adaptive features

It is obvious that above-mentioned issues can be addressed by developing such feature extraction methods which can deal with these imaging variations (Lucas et al. 2018). Computer-aided diagnoses will more accurate if the extracted descriptors are more robust to scale, rotation, and illumination invariance (Ahn et al. 2018). Moreover, the scale and rotation variations can be dealt with kernel-based feature extraction methods (e.g., LBP and SIFT). However, illumination variations are easily coped by utilization of different color space representations of images. The texture in the lesion also have some repetitive shapes. The combination of texture and shape-based features can be a better option.

6.8 Generic features

We have seen many feature extraction methods. However, these methods are application dependent detect only a specific disease. Some features are applicable for classification, these may not appropriate for image retrieval or segmentation tasks. Moreover, most of these feature extraction methods are developed for general recognition applications. Also, there is a need to develop methods that well-perform specifically on gastric images. We can encounter multiple gastric abnormalities in a single endoscopic sequence and most of the features are used to represent a single type of lesions (e.g., cancer, ulcer, or bleeding). However, it is a necessity to design more robust, generic, imaging modality, GI area, and application independent features for representation of gastric frames.

6.9 The power of deep neural network

After so much research in field of image processing and machine learning, we still do not know what attributes best represent the abnormalities in endoscopy frames. There are lots of methods for manual feature extraction have been proposed but no one claims to be more generalized. These manual features do not best cope with the versatility of images acquisition and dynamic conditions of gastric tract (Shichijo et al. 2017). Therefore, deep learning came into the big picture. The deep convolutional neural network shuns the need of manual electing features for representing image (Sharma et al. 2017) (Razzak et al. 2018). With multi dimensional applications of deep learning, deep learning methods can be used for segmentation of images as well (Guo et al. 2017).

Representation learning is one of the central issues in machine learning. However, without fully understanding the work of deep neural network, we still able to use it for learning efficient image representations. In this paper, our focus is on features learning. Moreover, the deep neural network can be used in different fields everyday life, for instance, translate text, recognize music, poetry, painting, predict behaviours of humans and calculate the quantum energy of molecules. The understanding of these groups of symmetries is an issue that goes far beyond the applications of learning. If we accomplish to specify them one day, we will better understand the geometry of the data in large dimensions. But this geometry is underlying many scientific problems.

(1) *Segmentation using Deep Learning* Deep neural network can be used for anatomical partitioning of endoscopy images (Hirasawa et al. 2018; Lai 2015). Several methods have been suggested by the researchers segment areas in medical images (e.g., for brain tumor de Brébisson and Montana 2015). Moreover, some already trained models (e.g., SegNet Badrinarayanan et al. 2017 and Wang et al. 2018a) can be fine-tuned for segmentation of

cancer areas in endoscopic frames of GI tract (Nan et al. 2017). Various types of mucosal structures can also be segmented using deep convolutional network (e.g., example Separator-Net and Object-Net presented in Kainz et al. 2017). Random fields used in modeling human perception and can also be helpful when mixed with CNN as idea presented in Arnab et al. (2018).

(2) *Texture Feature Extraction Using Deep Learning* Basic texture feature extraction methods like Gabor- based, LBP, and GLCM can be combined with deep learning methods to represent the gastric anomalies in a better way. Gabor-based texture descriptors can be learned from images using different orientation and scales (Luan et al. 2017). We can perform same for LBP, GLCM, and color-texture hybrid features. Moreover, other features like color and geometric for gastric images can be learned using deep learning methods (Komura and Ishikawa 2017). Gabor filters can be used to optimize the energy and computation of CNNs, as Sarwar et al. (2017) shows that the convolutional layers have representation like randomly tuned Gabor filters.

6.10 Publicly available expert annotated images data-sets and challenges

Many computer-aided diagnostic challenges organized by different researchers around the world. These challenges and there corresponding publicly available image data-sets have summarized as below:

(1) *Challenge on Analysis of Images to Detect Abnormalities in Endoscopy (AIDA-E)* There is an increased burden of gastrointestinal diseases around the world. It is challenging to develop methods for screening the GI tract for potential abnormalities and early diagnoses of tissue malignancies. Several researchers are still interested in the automatic detection of gastric abnormalities and working on different scales and optical technologies. Therefore, the aim of this challenge is to provide standard data-sets and benchmarks, so that the performance new developed methods can be compared on common the parameters. This challenge has many sub-challenges under its umbrella and having different needs and endoscopic imagery for testing vision-based algorithms, summarized as follows:

(a) *Gastric Chromoendoscopy Images in Cancer Surveillance* This challenge is about detection of gastric abnormal frames for classification of chromoendoscopy images the classification is based on the taxonomy provided by Ribeiro (2005). The CH images belong to three groups; Group 1 contains CH images with normal mucosa. Group II have images with mucosal abnormal condition metaplasia. Group III has images of patients which were suffering from dysplasia. The aim of this challenge is to classify these images based on color, shape and irregular texture (see Chromogastro[2] in Table 16).

(b) *Esophagus Micro-Endoscopy Images in Barrett's Surveillance* This challenge is about early detection of cancer from BE which is a premalignant state of mucosal cells. In BE, the normal cells are replacing with metaplastic cells containing goblet cells. The main challenge is to classify each CLE image according to histologically. The images will be classified gastric metaplasia (GMP), intestinal metaplasia or proper Barrett's esophagus (BAR), or neoplasia (NPL) (see CLE_barrett[3] in Table 16).

(c) *Confocal Endoscopy in Celiac Imaging* A CLE images data-set containing various pathologies of CD has provided for automatic detection of mucosal damage. Small-bowel mucosa damaged from mild or with increased intraepithelial lymphocytes and Crypt Hyperplasia (CH) to more severe damage referred to villous Atrophy (VA). The aim of this challenge

---

[2] https://aidasub-chromogastro.grand-challenge.org.

[3] https://aidasub-cleceliachy.grand-challenge.org/description/.

Table 16 Publicly available clinical endoscopy images data-set for testing and comparison of performance CADx

Dataset	Frames/videos	Endoscopy	Disease	GI. area	Description
Chromogastrov[a]	176 frames	CH	Cancer	Stomach	518 × 481 (*.png)
CLE_barrett[b]	262 frames	CLE	Cancer	Esophagus	1024 × 1024 (*.jpg)
CLE_celiachy[c]	181 frames	CLE	Celiac	Small bowel	1024 × 1024 (*.jpg)
CVC-ColonDB[d]	612 frames	WCE	Polyp	Colon	574 × 500 (*.tif)
ETIS-Larib[e]	196 frames	VE	Polyp	Colon	1225 × 996 (*.tif)
ASU-Mayo Clinic[f]	20 videos	VE	Polyp	Colon	–
Gastric_Data[g]	800 frames	VE	Abnormal	Stomach	489 × 409
HD_barrett[h]	100 frames	HDVE	Cancer	Esophagus	1600 × 1200
CVC-VideoClinicDB[i]	18 videos	WCE	Polyp	Colon	384 × 288 (*.bmp)
CVC-DB[j]	912 frames	WCE	Polyp	Colon	574 × 500 (*.tif)
CVC-ClinicHDSegment[k]	168 frames	HDVE	Polyp	Colon	1920 × 1880 (*.png)

[a] https://aidasub-chromogastro.grand-challenge.org
[b] https://aidasub-clebarrett.grand-challenge.org
[c] https://aidasub-cleceliachy.grand-challenge.org/description/
[d] https://polyp.grand-challenge.org/site/Polyp/CVCClinicDB/
[e] https://polyp.grand-challenge.org/site/Polyp/EtisLarib/
[f] https://polyp.grand-challenge.org/site/Polyp/AsuMayo/
[h] https://endovissub-barrett.grand-challenge.org/data_description/
[i] https://endovissub2017-giana.grand-challenge.org/polypdetection
[j] https://endovissub2017-giana.grand-challenge.org/polypsegmentation/
[k] https://endovissub2017-giana.grand-challenge.org/polypsegmentation/

to develop a CADx system to classify images in either a normal mucosa or villous atrophy (VA), crypt hypertrophy (CH) of both (VACH). Each classified image showed a normal mucosa, villous atrophy or crypt hypertrophy, as increase severity of CD damage to the intestinal mucosa (see CLE_celiachy[4] in Table 16).

(2) *Endoscopy Vision Challenge* In this challenge, data-sets has been provided for endoscopic surgical vision related tasks. These data-sets are developed for performance comparisons and bench-marking of different vision-based algorithms. Computer vision-based methods including 3D surface reconstruction, lesion surveillance, tracking, and surgical instruments segmentation from endoscopic frames or videos. Some sub-challenges of this grand challenge are as follows:

(a) *Automatic Polyp Detection in Colonoscopy Videos* This challenge is about polyp segmentation and tracking in colonoscopy videos, it also provides ground truth values along with colon images by indicating polyp pixels. Moreover, the challenging task can be divided into two sub-task, first is segmentation of polyps with more accuracy and the second sub-task is the detection of frames with polyps and take account of the occurrences of polyps in the whole video sequence (see CVC-ColonDB[5] ETIS-Larib[6] ASU-Mayo Clinic[7] in Table 16).

(b) *Detection of Abnormalities in Gastroscopic Images* In this challenge, 800 gastroscopic images are provided for detection of abnormality in these images. In training data, 260 are abnormal frames and 205 are normal frames. In testing set 129 are normal and 104 abnormal frames (see Gastric_Data[8] in Table 16).

(c) *Early Barrett's Cancer Detection* HD endoscopic frames are provided to test algorithms which are developed for detection of Barret's cancer. HDVE images of 39 patients are gathered where 17 have cancer and 22 are healthy are grouped into 2 sets, 50 images have cancer and 50 have no cancer present in them. The challenge is to detect images with cancer (see HD_barrett[9] in Table 16).

(d) *Gastrointestinal Image ANAlysis (GIANA)* This challenge is also about segmentation and tracking polyps from endoscopic frames. It comes with two data-sets one is for detection and another data-set is for segmentation of polyps. 300 images for training and 612 images for testing the algorithms developed for polyp frame detection. Moreover, 18 video sequence for polyp detection and segmentation data-set contains 168 frames (see CVC-VideoClinicDB[10] CVC-DB[11] CVC-ClinicHDSegment[12] in Table 16).

(d) *Instrument Segmentation and Tracking* This challenge has two parts one is segmentation of surgical instruments and the other part is tracking of these segmented surgical instruments in the whole video sequence. Images data-sets both for tracking and segmentation of surgical instruments have been provided. Moreover, this images data-set contains two types rigid and robotic instruments for segmentation (see data description[13])

---

[4] https://aidasub-clebarrett.grand-challenge.org.

[5] https://polyp.grand-challenge.org/site/Polyp/CVCClinicDB/.

[6] https://polyp.grand-challenge.org/site/Polyp/EtisLarib/.

[7] https://polyp.grand-challenge.org/site/Polyp/AsuMayo/.

[8] https://endovissub-abnormal.grand-challenge.org/data/.

[9] https://endovissub-barrett.grand-challenge.org/data_description/.

[10] https://endovissub2017-giana.grand-challenge.org/polypdetection.

[11] https://endovissub2017-giana.grand-challenge.org/polypsegmentation/.

[12] https://endovissub2017-giana.grand-challenge.org/polypsegmentation/.

[13] https://endovissub-instrument.grand-challenge.org/data/.

7 Conclusion

Computer-aided diagnosis (CADx) is the future of clinical practices. By using a less invasive endoscopy to observe the gastrointestinal tract is the most reasonable approach for screening. As a result, efficient computer vision techniques are needed for the detection of abnormalities from the endoscopic sequence. Moreover, advancements in surgical-vision techniques will revolutionize the surgical procedures and ultimately provides computer-assisted interventions. A successful CADx system requires efficient features extraction and image representation methods. Therefore, there is need to design such features extraction methods those deal with dynamics of the gastric environment and provides a better description of the gastric lesion. In this paper, we have presented a survey of existing CADx systems have been developed for detection of gastric abnormalities. Moreover, we have reviewed them based on their feature extraction techniques. These features extraction techniques are grouped based on their respective domain and descriptors. We have also mentioned various endoscopy modalities, and abnormalities. This survey also has discussed various open issues, trends, and challenges. Moreover, image data-sets associated with these medical-imaging challenges in the field of computer-assisted endoscopy are described.

References

Adler S, Hassan C, Metzger Y, Sompolinsky Y, Spada C (2012) Accuracy of automatic detection of small-bowel mucosa by second-generation colon capsule endoscopy. Gastrointest Endosc 76(6):1170–1174

Ahn J, Loc HN, Balan RK, Lee Y, Ko J (2018) Finding small-bowel lesions: challenges in endoscopy-image-based learning systems. Computer 51(5):68–76

Albisser Z (2015) Computer-aided screening of capsule endoscopy videos. Ph.D. Dissertation, University of Oslo

Alexandre L, Casteleiro J, Nobre N (2007) Polyp detection in endoscopic video using SVMs. In: 11th european conference on principles and practice of knowledge discovery in databases KPKDD, vol 4702, pp 358–365

Alexandre LA, Nobre N, Casteleiro J (2008) Color and position versus texture features for endoscopic polyp detection. In: 1st International conference on biomedical engineering and informatics: new development and the future, BMEI, vol 2, pp 38–42

Ali H, Sharif M, Yasmin M, Rehmani MH (2017) Computer-based classification of chromoendoscopy images using homogeneous texture descriptors. Comput Biol Med 88:84–92

Ali H, Yasmin M, Sharif M, Rehmani MH (2018) Computer assisted gastric abnormalities detection using hybrid texture descriptors for chromoendoscopy images. Comput Methods Programs Biomed 157:39–47. https://doi.org/10.1016/j.cmpb.2018.01.013

Ameling S, Wirth S, Paulus D, Lacey G, Vilarino F (2009) Texture-based polyp detection in colonoscopy. In: Bildverarbeitung für die Medizin, pp 346–350

André B, Vercauteren T, Perchant A, Buchner AM, Wallace MB, Ayache N (2010) Introducing space and time in local feature-based endomicroscopic image retrieval, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 5853 LNC, pp 18–30

Anselmi F, Evangelopoulos G, Rosasco L, Poggio T (2017) Symmetry regularization. CBMM Memo 063, June 2017

Arnab A, Zheng S, Jayasumana S, Romera-paredes B, Kirillov A, Savchynskyy B, Rother C, Kahl F, Torr P (2018) Conditional random fields meet deep neural networks for semantic segmentation XX(Xx):1–15

Axon A (2008) Is diagnostic and therapeutic endoscopy currently appropriate? Suggestions for improvement. Best Pract Res Clin Gastroenterol 22(5):959–970

Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 39(12):2481–2495

Bao G, Pahlavan K, Mi L (2015) Hybrid localization of microrobotic endoscopic capsule inside small intestine by data fusion of vision and RF sensors. IEEE Sens J 15(5):2669–2678

Barbosa D, Ramos J, Lima CS (2008) Detection of small bowel tumors in capsule endoscopy frames using texture analysis based on the discrete wavelet transform. In: Annual international conference of the IEEE engineering in medicine and biology society, EMBS, pp 3012–3015

Barbosa DJ, Ramos J, Correia JH, Lima CS (2009) Automatic detection of small bowel tumors in capsule endoscopy based on color curvelet covariance statistical texture descriptors. In: Annual international conference of the IEEE on engineering in medicine and biology society, pp 6683–6686

Beg S, Ragunath K (2015) Image-enhanced endoscopy technology in the gastrointestinal tract: What is available? Best Pract Res Clin Gastroenterol 29(4):627–638

Bejakovic S, Kumar R, Dassopoulos T, Mullin G, Hager G (2009) Analysis of Crohn's disease lesions in capsule endoscopy images. In: IEEE international conference on robotics and automation, pp 2793–2798

Bhat YM, Abu Dayyeh BK, Chauhan SS, Gottlieb KT, Hwang JH, Komanduri S, Konda V, Lo SK, Manfredi MA, Maple JT et al (2014) High-definition and high-magnification endoscopes. Gastrointest Endosc 80(6):919–927

Bonnel J, Khademi A, Krishnan S, Ioana C (2009) Small bowel image classification using cross-co-occurrence matrices on wavelet domain. Biomed Sig Process Control 4(1):7–15

Boschetto D, Mirzaei H, Leong RWL, Tarroni G, Grisan E (2015) Semiautomatic detection of villi in confocal endoscopy for the evaluation of celiac disease. In: Annual international conference of the IEEE engineering in medicine and biology society, EMBS, pp 8143–8146

Boschetto D, Di Claudio G, Mirzaei H, Leong R, Grisan E (2016) Automatic classification of small bowel mucosa alterations in celiac disease for confocal laser endomicroscopy. In: SPIE medical imaging, vol. 9788. International Society for Optics and Photonics, pp.978 809–978 809

Buchner AM, Shahid MW, Heckman MG, Krishna M, Ghabril M, Hasan M, Crook JE, Gomez V, Raimondo M, Woodward T, Wolfsen HC, Wallace MB (2010) Comparison of Probe-Based Confocal Laser Endomicroscopy With Virtual Chromoendoscopy for Classification of Colon Polyps. Gastroenterology 138(3):834–842

Callacondo-Riva D, Ganoza-Salas A, Anicama-Lima W, Quispe-Mauricio A, Longacre TA (2009) Primary squamous cell carcinoma of the stomach with paraneoplastic leukocytosis: a case report and review of literature. Hum Pathol 40(10):1494–1498

Cao Z, Mu S, Xu Y, Dong M (2019) Image retrieval method based on CNN and dimension reduction. arXiv preprint arXiv:1901.03924

Carpi F, Kastelein N, Talcott M, Pappone C (2011) Magnetically controllable gastrointestinal steering of video capsules. IEEE Trans Biomed Eng 58(2):231–234

Charisis V, Tsiligiri A, Hadjileontiadis LJ, Liatsos CN, Mavrogiannis CC, Sergiadis GD (2010) Ulcer detection in wireless capsule endoscopy images using bidimensional nonlinear analysis. In: Annual international conference of the IEEE on engineering in medicine and biology society (EMBC), pp 236–239

Cho WY, Jang JY, Lee DH (2011) Recent advances in image-enhanced endoscopy. Clin Endosc 44(2):65–75

Choi BI (2014) Radiology illustrated: gastrointestinal tract, vol 2. Springer, Berlin

Chu QD, Zibari GB, Gibbs JF (2015) Surgical oncology: a practical and comprehensive approach. Surgical Oncology a Practical and Comprehensive Approach, pp 3–678

Ciaccio EJ, Tennyson CA, Lewis SK, Bhagat G, Green PH (2010) T1199 distinguishing patients with celiac disease by quantitative analysis of videocapsule endoscopy images. Gastroenterology 138(5):S510

Cobrin GM, Pittman RH, Lewis BS (2006) Increased diagnostic yield of small bowel tumors with capsule endoscopy. Cancer 107(1):22–27

Coda S (2014) An investigation of the diagnostic potential of autofluorescence lifetime spectroscopy and imaging for label-free contrast of disease. Ph.D. dissertation, Thesis submitted for the award of Doctor of Philosophy (Ph.D.) Imperial College of Science, Technology and Medicine

Cohen T, Welling M (2016) Group equivariant convolutional networks. in: International conference on machine learning, pp 2990–2999

Coimbra MT, Cunha JS (2006) MPEG-7 visual descriptors–contributions for automated feature extraction in capsule endoscopy. IEEE Trans Circuits Syst Video Technol 16(5):628–637

Coimbra M, Campos P, Cunha JS (2006) Topographic segmentation and transit time estimation for endoscopic capsule exams. In: IEEE international conference on acoustics, speech and signal processing, ICASSP, vol 2, pp II–II

Cong Y, Wang S, Liu J, Cao J, Yang Y, Luo J (2015) Deep sparse feature selection for computer aided endoscopy diagnosis. Pattern Recognit 48:907–917

Cong Y, Wang S, Fan B, Yang Y, Yu H (2016) UDSFS: Unsupervised deep sparse feature selection. Neurocomputing 196:150–158

Constantinescu AF, Ionescu M, Rogoveanu I, Ciurea ME, Streba CT, Iovanescu VF, Artene SA, Vere CC (2015) Analysis of wireless capsule endoscopy images using local binary patterns. Appl Med Inf 36(2):31

Cui L, Hu C, Zou Y, Meng MQH (2010) Bleeding detection in wireless capsule endoscopy images by support vector classifier. In: IEEE International conference on information and automation, ICIA, pp 1746–1751

Dahal A, Oh J, Tavanapong W, Wong J, De Groen PC (2015) Detection of ulcerative colitis severity in colonoscopy video frames. In: International workshop on content-based multimedia indexing

Dattamajumdar AK, Blount PL, Myers JA, Proctor AH, Goldman BH, Reid BJ, Martin RW (2001) A low-cost fiber-optic instrument to colorimetrically detect patients with Barrett's esophagus for early detection of esophageal adenocarcinoma. IEEE Trans Biomed Eng 48(6):695–705

de Brébisson A, Montana G (2015) Deep neural networks for anatomical brain segmentation. arXiv preprint arXiv:1502.02445

Deeba F, Islam M, Bui FM, Wahid KA (2018) Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection. Biomed Sig Process Control 40:415–424. https://doi.org/10.1016/j.bspc.2017.10.011

Deng L, Yu D et al (2014) Deep learning: methods and applications, Foundations and trends®. Signal Process 7(3–4):197–387

Dieleman S, De Fauw J, Kavukcuoglu K (2016) Exploiting cyclic symmetry in convolutional neural networks. arXiv preprint arXiv:1602.02660

Dong Y, Tao D, Li X (2015a) Nonnegative multiresolution representation-based texture image classification. ACM Tran Intell Syst Technol (TIST) 7(1):4

Dong Y, Tao D, Li X, Ma J, Pu J (2015b) Texture classification and retrieval using shearlets and linear regression. IEEE Trans Cybern 45(3):358–369

Dong Y, Feng J, Liang L, Zheng L, Wu Q (2017) Multiscale sampling based texture image classification. IEEE Signal Process Lett 24(5):614–618

Dong Y, Feng J, Yang C, Wang X, Zheng L, Pu J (2018a) Multi-scale counting and difference representation for texture classification. Vis Comput 34(10):1315–1324

Dong Y, Wu H, Li X, Zhou C, Wu Q (2018b) Multiscale symmetric dense micro-block difference for texture classification. IEEE Trans Circuits Syst Video Technol. https://doi.org/10.1109/TCSVT.2018.2883825

Drozdzal M, Seguí S, Radeva P, Malagelada C, Azpiroz F, Vitrià J (2015) Motility bar: a new tool for motility analysis of endoluminal videos. Comput Biol Med 65:320–330

Dunaeva O, Edelsbrunner H, Lukyanov A, Machin M, Malkova D (2015) The classification of endoscopy images with persistent homology. In: 16th International symposium on symbolic and numeric algorithms for scientific computing, SYNASC. Elsevier, pp 565–570

Eberlein T, Reid B, Hawk E (2002) Report of the stomach/esophageal cancers progress review group. National Cancer Institute, Bethesda, MD

Eliakim R (2004) Wireless capsule video endoscopy: three years of experience. World J Gastroenterol 10(9):1238–1239

Emam AZ, Ali YA, Ben Ismail MM (2015) Adaptive features extraction for capsule endoscopy (CE) video summarization. In: International conference on computer vision and image analysis applications, ICCVIA

Figueiredo IN, Figueiredo PN, Stadler G, Ghattas O, Araujo A (2010) Variational image segmentation for endoscopic human colonic aberrant crypt foci. IEEE Trans Biomed Eng 29(4):998–1011

Figueiredo IN, Leal C, Pinto L, Figueiredo PN, Tsai R (2018) Hybrid multiscale affine and elastic image registration approach towards wireless capsule endoscope localization. Biomed Sig Process Control 39:486–502. https://doi.org/10.1016/j.bspc.2017.08.019

Filip D, Yadid-Pecht O, Andrews CN, Mintchev MP (2011) Self-stabilizing colonic capsule endoscopy: pilot study of acute canine models. IEEE Trans Med Imaging 30(12):2115–2125

Francisco SIM, Sousa RG, Coimbra MT (2015) Recognition of Cancer using random forests as a bag-of-words approach for gastroenterology. Ph.D. Dissertation, Masters thesis, Universidade do Porto, 2015. (Cited on pages 77 and 78)

Gao M, Hu C, Chen Z, Liu S, Zhang H (2009) Finite-difference modeling of micromachine for use in gastrointestinal endoscopy. IEEE Trans Biomed Eng 56(10):2413–2419

Garcia E, Hermoza R, Castanon CB, Cano L, Castillo M, Castanñeda C (2017) Automatic lymphocyte detection on gastric cancer IHC images using deep learning. In: IEEE 30th international symposium on computer-based medical systems (CBMS). IEEE, pp 200–204

Gastelum A, Mata L, Brito-de-la Fuente E, Delmas P, Vicente W, Salinas-Vázquez M, Ascanio G, Marquez J (2015) Building a three-dimensional model of the upper gastrointestinal tract for computer simulations of swallowing. Med Biol Eng Comput 54:525–534

Geng Y, Pahlavan K (2015) On the accuracy of RF and image processing based hybrid localization for wireless capsule endoscopy. In: IEEE wireless communications and networking conference (WCNC), pp 464–469

Gens R, Domingos PM (2014) Deep symmetry networks. In: Advances in neural information processing systems, pp. 2537–2545

Gholami S, Janson L, Worhunsky DJ, Tran TB, Squires MH, Jin LX, Spolverato G, Votanopoulos KI, Schmidt C, Weber SM, Bloomston M, Cho CS, Levine EA, Fields RC, Pawlik TM, Maithel SK, Efron B, Norton JA, Poultsides GA (2015) Number of lymph nodes removed and survival after gastric cancer resection: an analysis from the US Gastric Cancer Collaborative. J Am Coll Surg 221(2):291–299

Ghosh T, Fattah S, Shahnaz C, Kundu A, Rizve M (2015) Block based histogram feature extraction method for bleeding detection in wireless capsule endoscopy. In: IEEE eegion 10 conference TENCON. IEEE, pp 1–4

Ghosh T, Fattah SA, Wahid KA (2017) CHOBS: color histogram of block statistics for automatic bleeding detection in wireless capsule endoscopy video. IEEE J Transl Eng Health Med 6:1–12

Ghosh T, Fattah SA, Wahid KA, Zhu WP, Ahmad MO (2018) Cluster based statistical feature extraction method for automatic bleeding detection in wireless capsule endoscopy video. Comput Biol Med 94(September 2016):41–54. https://doi.org/10.1016/j.compbiomed.2017.12.014

Gibson E, Giganti F, Hu Y, Bonmati E, Bandula S, Gurusamy K, Davidson B, Pereira SP, Clarkson MJ, Barratt DC (2018) Automatic multi-organ segmentation on abdominal CT with dense v-networks. IEEE Trans Med Imaging 37(8):1822–1834

Girgis HZ, Mitchell BR, Dassopoulos T, Mullin G, Hager G (2010) An intelligent system to detect Crohn's disease inflammation in wireless capsule endoscopy videos. In: 7th IEEE international symposium on biomedical imaging: from nano to macro, ISBI, pp 1373–1376

Giritharan B, Yuan X, Liu J, Buckles B, Oh J, Tang SJ (2008) Bleeding detection from capsule endoscopy videos. In: 30th Annual international conference of the IEEE in engineering in medicine and biology society, EMBS, pp 4780–4783

Goetz M, Malek NP, Kiesslich R (2013) Microscopic imaging in endoscopy: endomicroscopy and endocy-toscopy. Nat Rev Gastroenterol Hepatol 11(1):11–18

Gómez V, Buchner AM, Dekker E, Van Den Broek FJC, Meining A, Shahid MW, Ghabril MS, Fockens P, Heckman MG, Wallace MB (2010) Interobserver agreement and accuracy among international experts with probe-based confocal laser endomicroscopy in predicting colorectal neoplasia. Endoscopy 42(4):286–291

Gono K, Obi T, Yamaguchi M, Ohyama N, Machida H, Sano Y, Yoshida S, Hamamoto Y, Endo T (2004) Appearance of enhanced tissue features in narrow-band endoscopic imaging. J Biomed Opt 9(3):568

Gotoda T, Uedo N, Yoshinaga S, Tanuma T, Morita Y, Doyama H, Aso A, Hirasawa T, Yano T, Uchita K et al (2016) Basic principles and practice of gastric cancer screening using high-definition white-light gastroscopy: eyes can only see what the brain knows. Dig Endosc 28(S1):2–15

Gross S, Stehle T, Behrens A, Auer R, Aach T, Winograd R, Trautwein C, Tischendorf J (2009) A comparison of blood vessel features and local binary patterns for colorectal polyp classification. In: SPIE medical imaging. International Society for Optics and Photonics, pp 72 602Q–72 602Q

Gschwandtner M, Liedlgruber M, Uhl A, Vecsei A (2010) Experimental study on the impact of endoscope distortion correction on computer-assisted celiac disease diagnosis. In: 10th IEEE international conference on information technology and applications in biomedicine, ITAB, pp 1–6

Gu Y, Xie X, Li G, Sun T, Wang D, Yin Z, Zhang P, Wang Z (2015) Design of endoscopic capsule with multiple cameras. IEEE Trans Biomed Circuits Syst 9(4):590–602

Gueye L, Yildirim-Yayilgan S, Cheikh FA, Balasingham I (2015) Automatic detection of colonoscopic anomalies using capsule endoscopy. In: IEEE International conference on image processing (ICIP), pp 1061–1064

Guo Y, Liu Y, Georgiou T, Lew MS (2017) A review of semantic segmentation using deep neural networks. Int J Multimed Inf Retr 7(2):87–93

Häfner M, Kendlbacher C, Mann W, Taferl W, Wrba F, Gangl A, Vécsei A, Uhl A (2007) PIT pattern classification of zoom-endoscopic colon images using histogram techniques. In: 7th nordic signal processing symposium, NORSIG 2006, pp 58–61

Häfner M, Kwitt R, Wrba F, Gangl A, Uhl A (2008) One-against-one classification for zoom-endoscopy images. In: 4th IET international conference on advances in medical, signal and information processing, (MEDSIP), pp 4–7

Häfner M, Gangl A, Liedlgruber M, Uhl A, Vécsei A, Wrba F (2009a) Combining Gaussian Markov random fields with the discrete wavelet transform for endoscopic image classification. In: 16th International conference on digital signal processing, DSP, pp 1–6

Häfner M, Kwitt R, Uhl A, Gangl A, Wrba F, Vécsei A (2009b) Feature extraction from multi-directional multi-resolution image transformations for the classification of zoom-endoscopy images. Pattern Anal Appl 12(4):407–413

Häfner M, Kwitt R, Uhl A, Wrba F, Gangl A, Vécsei A (2009c) Computer-assisted pit-pattern classification in different wavelet domains for supporting dignity assessment of colonic polyps. Pattern Recognit 42(6):1180–1191

Hafner M, Brunauer L, Payer H, Resch R, Gangl A, Uhl A, Wrba F, Vécsei A (2010a) Computer-aided classification of zoom-endoscopical images using fourier filters. IEEE Trans Inf Technol Biomed 14(4):958–970

Hafner M, Uhl A, Vecsei A, Wimmer G, Wrba F (2010b) Complex wavelet transform variants and discrete cosine transform for scale invariance in magnification-endoscopy image classification. In: 10th IEEE international conference on information technology and applications in biomedicine, pp 1–5

Häfner M, Tamaki T, Tanaka S, Uhl A, Wimmer G, Yoshida S (2015) Local fractal dimension based approaches for colonic polyp classification. Med Image Anal 26(1):92–107

Hamashima C, Shabana M, Okada K, Okamoto M, Osaki Y (2015) Mortality reduction from gastric cancer by endoscopic and radiographic screening. Cancer Sci 106(12):1744–1749

Hassan AR, Haque MA (2015) Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos. Comput Methods Programs Biomed 122(3):341–353

Hazewinkel Y, López-Cerón M, East JE, Rastogi A, Pellisé M, Nakajima T, Van Eeden S, Tytgat KMAJ, Fockens P, Dekker E (2013) Endoscopic features of sessile serrated adenomas: validation by international experts using high-resolution white-light endoscopy and narrow-band imaging. Gastrointest Endosc 77(6):916–924

He J-Y, Wu X, Jiang Y-G, Peng Q, Jain R (2018) Hookworm detection in wireless capsule endoscopy images with deep learning. IEEE Trans Image Process 27(5):1

Hegenbart S, Uhl A, Vécsei A (2015) Survey on computer aided decision support for diagnosis of celiac disease. Comput Biol Med 65:348–358

Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J et al (2018) Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric Cancer 21(4):653–660

Huang C-R, Chung P-C, Sheu B-S, Kuo H-J (2008) Helicobacter pylori-related gastric histology classification using support-vector-machine-based feature selection. IEEE Trans Inf Technol Biomed 12(4):523–531

Huang C-R, Chen Y-T, Chen W-Y, Cheng H-C, Sheu B-S (2015) Gastroesophageal reflux disease diagnosis using hierarchical heterogeneous descriptor fusion support vector machine. IEEE Trans Biomed Eng PP(99):1

Hurlstone DP, Cross SS, Adam I, Shorthouse AJ, Brown S, Sanders DS, Lobo AJ (2004) Efficacy of high magnification chromoscopic colonoscopy for the diagnosis of neoplasia in flat and depressed lesions of the colorectum: a prospective analysis. Gut 53(2):284–90

Hu Z, Tang J, Wang Z, Zhang K, Zhang L, Sun Q (2018) Deep learning for image-based cancer detection and diagnosis—a survey. Pattern Recognit 83:134–149

Hwang S, Celebi ME (2010) Polyp detection in wireless capsule endoscopy video based on image segmentations and geometric feature. In: IEEE international conference on acoustics speech and signal processing, (ICASSP), vol 7, pp 678–681

Hwang S, Oh J, Tavanapong W, Wong J, De Groen PC (2007) Polyp detection in colonoscopy video using elliptical shape feature. In: International conference on image processing, ICIP, vol 2, pp 465–468

Iakovidis DK, Maroulis DE, Karkanis SA, Brokos A (2005) A comparative study of texture features for the discrimination of gastric polyps in endoscopic video. In: IEEE Symposium on computer-based medical systems, pp 575–580

Iakovidis DK, Maroulis DE, Karkanis SA (2006) An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. Comput Biol Med 36(10):1084–1103

Iakovidis DK, Chatzis D, Chrysanthopoulos P, Koulaouzidis A (2015) Blood detection in wireless capsule endoscope images based on salient superpixels. In: Annual international conference of the IEEE engineering in medicine and biology society (EMBS), pp 731–734

Iakovidis DK, Georgakopoulos SV, Vasilakakis M, Koulaouzidis A, Plagianakos VP (2018) Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification. IEEE Trans Med Imaging 37(10):2196–2210

Ishihara K, Ogawa T, Haseyama M (2016) Helicobacter pylori infection detection from gastric X-ray images using KLFDA-based decision fusion. In: IEEE 4th global conference on consumer electronics, GCCE, pp 204–205

Janse MHA, van der Sommen F, Zinger S, Schoon EJ, de With PHN (2016) Early esophageal cancer detection using RF classifiers. In: Proceedings of SPIE 9785, vol 9785, pp 97 851D-1–8

Jensen DM, Ohning GV, Kovacs TOG, Ghassemi KA, Jutabha R, Dulai GS, Machicado GA (2016) Doppler endoscopic probe as a guide to risk stratification and definitive hemostasis of peptic ulcer bleeding. Gastrointest Endosc 83:129–136

Jung YS, Kim YH, Lee DH, Kim JH (2008) Active blood detection in a high resolution capsule endoscopy using color spectrum transformation. In: International conference on biomedical engineering and informatics: new development and the future, BMEI 2008, vol 1, pp 859–862

Kainuma M, Furusyo N, Urita Y, Nagata M, Ihara T, Oji T (2015) The association between objective tongue color and endoscopic findings: results from the Kyushu and Okinawa population study (KOPS). BMC Complement Altern Med 15:372–379

Kainz P, Pfeiffer M, Urschler M (2017) Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. PeerJ 5:e3874

Kaltenbach T, Sano Y, Friedland S, Soetikno R (2008) American gastroenterological association (AGA) institute technology assessment on image-enhanced endoscopy. Gastroenterology 134(1):327–340

Kang J, Doraiswami R (2003) Real-time image processing system for endoscopic applications. In: Canadian conference on electrical and computer engineering CCECE, toward a caring and humane technology, vol 3, pp 1469–1472

Karargyris A, Bourbakis N (2009a) Identification of polyps in wireless capsule endoscopy videos using Log Gabor filters. In: IEEE/NIH life science systems and applications workshop, LiSSA, pp 143–147

Karargyris A, Bourbakis N (2009b) Identification of ulcers in wireless capsule endoscopy videos. In: IEEE international symposium on biomedical imaging: from nano to macro, ISBI, no. 3, pp 554–557

Karkanis SA, Iakovidis DK, Karras D, Maroulis D (2001) Detection of lesions in endoscopic video using textural descriptors on wavelet domain supported by artificial neural network architectures. IEEE Int Conf Image Process 2:833–836

Kato M, Kaise M, Yonezawa J, Toyoizumi H, Yoshimura N, Yoshida Y, Kawamura M, Tajiri H (2010) Magnifying endoscopy with narrow-band imaging achieves superior accuracy in the differential diagnosis of superficial gastric lesions identified with white-light endoscopy: a prospective study. Gastrointest Endosc 72(3):523–529

Kauderer-Abrams E (2017) Quantifying translation-invariance in convolutional neural networks. arXiv preprint arXiv:1801.01450

Kelley JR, Duggan JM (2003) Gastric cancer epidemiology and risk factors. J Clin Epidemiol 56(1):1–9

Keuchel M, Kurniawan N, Baltes P, Bandorski D, Koulaouzidis A (2015) Quantitative measurements in capsule endoscopy. Comput Biol Med 65:333–347

Khan TH, Wahid KA (2014) White and narrow band image compressor based on a new color space for capsule endoscopy. Signal Process Image Commun 29(3):345–360

Khashab MA, El Zein M, Kumbhari V, Besharati S, Ngamruengphong S, Messallam A, Abdelgalil A, Saxena P, Tieu AH, Raja S, Stein E, Dhalla S, Garcia P, Singh VK, Pasricha PJ, Kalloo AN, Clarke JO (2016) Comprehensive analysis of efficacy and safety of peroral endoscopic myotomy performed by a gastroenterologist in the endoscopy unit: a single-center experience. Gastrointest Endosc 83(1):117–125

Kiesslich R, Neurath MF (2007) Endomicroscopy is born-do we still need the pathologist? Gastrointest Endosc 66(1):150–153

Kim B, Lee S, Park JH, Park J-O (2005) Design and fabrication of a locomotive mechanism for capsule-type endoscopes using shape memory alloys (SMAs). IEEE/ASME Trans Mechatron 10(1):77–86

Kodashima S, Fujishiro M, Ono S, Niimi K, Mochizuki S, Asada-Hirayama I, Konno-Shimizu M, Matsuda R, Minatsuki C, Nakayama C, Takahashi Y, Sakaguchi Y, Yamamichi N, Tanaka C, Koike K (2014) Evaluation of a new image-enhanced endoscopic technology using band-limited light for detection of esophageal squamous cell carcinoma. Dig Endosc 26(2):164–171

Kodogiannis VS, Lygouras JN (2008) Neuro-fuzzy classification system for wireless-capsule endoscopic images. Eng Technol pp 620–628

Kodogiannis VS, Boulougoura M, Lygouras JN, Petrounias I (2007) A neuro-fuzzy-based system for detecting abnormal patterns in wireless-capsule endoscopic images. Neurocomputing 70(4–6):704–717

Komura D, Ishikawa S (2017) Machine learning methods for histopathological image analysis. arXiv preprint arXiv:1709.00786

Kondor IR (2008) Group theoretical methods in machine learning. Columbia University, New York

Korngold E (2011) GISTs—gastrointestinal stromal tumors, Elisabetta AM de Lutio di Castelguidone, Ed. Springer, vol. 42, no. 3

Koshy NE, Gopi VP (2015) A new method for ulcer detection in endoscopic images. In: 2nd international conference on electronics and communication systems (ICECS). IEEE, pp 1725–1729

Krishnan S, Goh P (1997) Quantitative parametrization of colonoscopic images by applying fuzzy technique. In: 19th Annual international conference of the IEEE on engineering in medicine and biology society, vol 3, pp 1121–1123

Krishnan S, Goh P (1999) Region labeling of colonoscopic images using fuzzy logic. In: Proceedings of the first joint BMES/EMBS conference, vol 2, p 1149

Krishnan S, Yang X, Chan K, Kumar S, Goh P (1998a) Intestinal abnormality detection from endoscopic images. In: 20th Annual international conference of the IEEE engineering in medicine and biology society biomedical engineering towards the year 2000 and beyond vol 2, no. 2, pp 895–898

Krishnan S, Yap C, Asari K, Goh P (1998b) Neural network based approaches for the classification of colono-scopic images. In: 20th annual international conference of the IEEE on engineering in medicine and biology society, vol 3, pp 1678–1680

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pp 1097–1105

Kwitt R, Uhl A (2008) Color eigen-subband features for endoscopy image classification. In: IEEE international conference on acoustics, speech and signal processing, pp 589–592

Kwitt R, Uhl A, Häfner M, Gangl A, Wrba F, Vécsei A (2010) Predicting the histology of colorectal lesions in a probabilistic framework. In: IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW), pp 103–110

Lai M (2015) Deep learning for medical image segmentation. arXiv preprint arXiv:1505.02000

Lau PY, Correia PL (2007) Detection of bleeding patterns in WCE video using multiple features. In: Annual international conference of the IEEE engineering in medicine and biology, pp 5601–5604

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436

Lee J, Oh J, Shah SK, Yuan X, Tang SJ (2007) Automatic classification of digestive organs in wireless capsule endoscopy videos. In: ACM symposium on applied computing, SAC, pp 1041–1045

Lee T-C, Lin Y-H, Uedo N, Wang H-P, Chang H-T, Hung C-W (2013) Computer-aided diagnosis in endoscopy: a novel application toward automatic detection of abnormal lesions on magnifying narrow-band imaging endoscopy in the stomach. In: 35th annual international conference of the IEEE in engineering in medicine and biology society, EMBC, pp 4430–4433

Leggett CL, Iyer PG (2015) Mucosal imaging advanced technologies in the gastrointestinal tract. Tech Gastrointest Endosc 17(4):161–170

Lehmann TM, Gönner C, Spitzer K (1999) Survey: interpolation methods in medical image processing. IEEE Trans Biomed Eng 18(11):1049–1075

Leodolter A, Alonso S, González B, Ebert MP, Vieth M, Röcken C, Wex T, Peitz U, Malfertheiner P, Perucho M (2015) Somatic DNA hypomethylation in H. pylori-associated high-risk gastritis and gastric cancer: enhanced somatic hypomethylation associates with advanced stage cancer. Clin Transl Gastroenterol 6(4):e85

Letter N (2007) News Letter. Am J Physiol 18(24):3–4

Lewis BS (2003) The utility of capsule endoscopy in obscure gastrointestinal bleeding. Tech Gastrointest Endosc 5(3):115–120

Li B, Meng MQH (2007) Analysis of the gastrointestinal status from wireless capsule endoscopy images using local color feature. In: Proceedings of the international conference on information acquisition, ICIA, pp 553–557

Li B, Meng MQH (2009a) Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments. Comput Biol Med 39(2):141–147

Li B, Meng MQH (2009b) Computer-aided detection of bleeding regions for capsule endoscopy images. IEEE Trans Biomed Eng 56(4):1032–1039

Li B, Meng MQH (2009c) Texture analysis for ulcer detection in capsule endoscopy images. Image Vis Comput 27(9):1336–1342

Li B, Meng MQ (2009d) Small bowel tumor detection for wireless capsule endoscopy images using textural features and support vector machine. In: IEEE/RSJ international conference on intelligent robots and systems, IROS, pp 498–503

Li B, Meng MQH (2010a) Tumor CE image classification using SVM-based feature selection. In: IEEE/RSJ international conference on intelligent robots and systems, IROS, pp 1322–1327

Li B, Meng MQH (2010b) Capsule endoscopy images classification by color texture and support vector machine. In: IEEE international conference on automation and logistics, ICAL, pp 126–131

Li B, Meng MQH (2012) Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection. IEEE Trans Inf Technol Biomed 16(3):323–329

Li P, Chan KL, Krishnan SM, Gao Y (2004) Detecting abnormal regions in colonoscopic images by patch-based classifier ensemble. Int Conf Pattern Recognit 3:774–777

Li B, Meng MQ-H, Lau JY (2011) Computer-aided small bowel tumor detection for capsule endoscopy. Artif Intell Med 52(1):11–16

Li B, Jin H, Yang C, Xu G (2015) A novel color textural feature towards capsule endoscopy video summary. In: IEEE international conference on information and automation, ICIA, pp 766–769

Liedlgruber M, Uhl A (2011) Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review. IEEE Rev Biomed Eng 4:73–88

Liedlgruber M, Supervisor T (2011) Computer-aided classification of endoscopic images from the gastrointestinal tract. Ph.D. Dissertation, Faculty of Natural Sciences, University of Salzburg

Liedlgruber M, Uhl A et al (2011) A summary of research targeted at computer-aided decision support in endoscopy of the gastrointestinal tract. Department of Computer Sciences, University of Salzburg, Austria. http://www.cosy.sbg.ac.at/research/tr.html, Technical Report, vol 1

Lima C, Barbosa D, Ramos J, Tavares A, Monteiro L, Carvalho L (2008) Classification of endoscopic capsule images by using color wavelet features, higher order statistics and radial basis functions. In: 30th Annual international conference of the IEEE on engineering in medicine and biology society, EMBS, pp 1242–1245

Liu DY, Gan T, Rao NN, Xu GG, Zeng B, Li HL (2015) Automatic detection of early gastrointestinal cancer lesions based on optimal feature extraction from gastroscopic images. J Med Imaging Health Inf 5:296–302

Lopez-Ceron M, Van Den Broek FJC, Mathus-Vliegen EM, Boparai KS, Van Eeden S, Fockens P, Dekker E (2013) The role of high-resolution endoscopy and narrow-band imaging in the evaluation of upper GI neoplasia in familial adenomatous polyposis. Gastrointest Endosc 77(4):542–550

Luck B, Maitland K, Collier T, Sung K-B (2004) Confocal microscopy [detecting and diagnosing cancers]. IEEE Potentials 23:14–17. https://doi.org/10.1109/MP.2004.1266933

Luan S, Zhang B, Chen C, Cao X, Ye Q, Han J, Liu J (2017) Gabor convolutional networks. arXiv preprint arXiv:1705.01450

Lucas A, Iliadis M, Molina R, Katsaggelos AK (2018) Using deep neural networks for inverse problems in imaging. IEEE Signal Process Mag 35(1):20–36

Maghsoudi OH (2017) Superpixels based segmentation and SVM based classification method to distinguish five diseases from normal regions in wireless capsule endoscopy. arXiv preprint arXiv:1711.06616

Magoulas GD, Plagianakos VP, Vrahatis MN (2004) Neural network-based colonoscopic diagnosis using on-line learning and differential evolution. Appl Soft Comput 4(4):369–379

Mahmood F, Chen R, Durr NJ (2018) Unsupervised reverse domain adaptation for synthetic medical images via adversarial training. IEEE Trans Med Imaging 37(12):2572–2581

Mallat S (2016) Understanding deep convolutional networks. Philos Trans R Soc A Math Phys Eng Sci 374(2065):20150203

Maroulis DE, Iakovidis DK, Karkanis SA, Karras DA, Maroulis DE, Iakovidis DK, Karkanis SA, Karras DA (2003) Cold: a versatile detection system for colorectal lesions in endoscopy video-frames. Comput Methods Programs Biomed 70(2):151–166

Martinez-Herrera SE, Benezeth Y, Boffety M, Emile JF, Marzani F, Lamarque D, Goudail F (2016) Identification of precancerous lesions by multispectral gastroendoscopy. Signal Image Video Process 10(3):455–462

Martins MM, Barbosa DJ, Ramos J, Lima CS (2010) Small bowel tumors detection in capsule endoscopy by Gaussian modeling of color curvelet covariance coefficients. In: Annual international conference of the IEEE engineering in medicine and biology society, EMBC, pp 5557–5560

Mathew M, Gopi VP (2015) Transform based bleeding detection technique for endoscopic images. In: 2nd international conference on electronics and communication systems, ICECS, pp 1730–1734

Miyahara R, Niwa Y, Matsuura T, Maeda O, Ando T, Ohmiya N, Itoh A, Hirooka Y, Goto H (2007) Prevalence and prognosis of gastric cancer detected by screening in a large Japanese population: data from a single institute over 30 years. J Gastroenterol Hepatol (Aust) 22(9):1435–1442

Miyaki R, Yoshida S, Tanaka S, Kominami Y, Sanomura Y, Matsuo T, Oka S, Raytchev B, Tamaki T, Koide T, Kaneda K, Yoshihara M, Chayama K (2015) A computer system to be used with laser-based endoscopy for quantitative diagnosis of early gastric cancer. J Clin Gastroenterol 49(2):108–115

Moccia S, Vanone GO, Momi ED, Laborai A, Guastini L, Peretti G, Mattos LS (2018) Learning-based classification of informative laryngoscopic frames. Comput Methods Programs Biomed 158:21–30. https://doi.org/10.1016/j.cmpb.2018.01.030

Mountford RA, Brown P, Salmon PR, Alvarenga C, Neumann CS, Read AE (1980) Gastric cancer detection in gastric ulcer disease. Gut 21(1):9–17

Muto M, Yao K, Kaise M, Kato M, Uedo N, Yagi K, Tajiri H (2016) Magnifying endoscopy simple diagnostic algorithm for early gastric cancer (MESDA-G). Dig Endosc 28(4):379–393

Namikawa T, Kobayashi M, Okabayashi T, Ozaki S, Nakamura S, Yamashita K, Ueta H, Miyazaki J, Tamura S, Ohtsuki Y (2005) Primary gastric small cell carcinoma: report of a case and review of the literature. Med Mol Morphol 38(4):256–261

Nan Y, Coppola G, Liang Q, Zou K, Sun W, Zhang D, Wang Y, Yu G (2017) Partial labeled gastric tumor segmentation via patch-based reiterative learning. arXiv, no. December, 2017. https://arxiv.org/ftp/arxiv/papers/1712/1712.07488.pdf

Nawarathna R, Oh J, Muthukudage J, Tavanapong W, Wong J, de Groen PC, Tang SJ (2014) Abnormal image detection in endoscopy videos using a filter bank and local binary patterns. Neurocomputing 144:70–91

Nishimura J, Nishikawa J, Nakamura M, Goto A, Hamabe K, Hashimoto S, Okamoto T, Suenaga M, Fujita Y, Hamamoto Y, Sakaida I (2014) Efficacy of i-scan imaging for the detection and diagnosis of early gastric carcinomas. Gastroenterol Res Pract

Nosato H, Sakanashi H, Takahashi E, Murakawa M (2015) Method of retrieving multi-scale objects from optical colonoscopy images based on image-recognition techniques. In: IEEE biomedical circuits and systems conference: engineering for healthy minds and able bodies, (BioCAS), pp 1–4

Ogoveanu IR, Iurea MEC, Treba CTS, Ovanescu VFI, Ere CCV, Constantinescu AF, Ionescu M, Rogoveanu I, Ciurea ME, Streba CT, Iovanescu VF, Vere CC (2015) Wireless capsule endoscopy in correlation with software application in gastrointestinal diseases. Current Health Sci J 41(2):89–94

Organization WH et al (2015) Cancer factsheet 2015. http://www.who.int/mediacentre/factsheets/fs297/en/. Accessed 1 Mar 2015

Owens SR, Appelman HD (2014a) Atlas of Esophagus and stomach pathology. Springer, Berlin

Owens SR, Appelman HD (2014b) Carcinoid tumor of the stomach. Springer, Berlin

Pan G, Yan G, Qiu X, Cui J (2011) Bleeding detection in wireless capsule endoscopy based on probabilistic neural network. J Med Syst 35:1477–1484

Peljto M, Milici AJ, Lange H, Krueger J, Aeffner F, Young GD, Martin NT (2016) Digital image analysis of inflammatory cells and mediators of inflammation. U.S. Patent 9,298,968

Pennazio M (2006) Capsule endoscopy: Where are we after 6 years of clinical use? Dig Liver Dis 38(12):867–878

Penny HA, Mooney PD, Burden M, Patel N, Johnston AJ, Wong SH, Teare J, Sanders DS (2016) High definition endoscopy with or without I-scan increases the detection of celiac disease during routine endoscopy. Dig Liver Dis 48(6):644–649

Pogorelov K, Riegler M, Halvorsen P, Griwodz C, Lange T, Randel K, Eskeland S, Dang-Nguyen D-T, Ostroukhova O, Lux M, Spampinato C (2017) A comparison of deep learning with global features for gastrointestinal disease detection. In: CEUR workshop proceedings, vol 1984, pp 8–10

Poh CK, Htwe TM, Li L, Shen W, Liu J, Lim JH, Chan KL, Tan PC (2010) Multi-level local feature classification for bleeding detection in wireless capsule endoscopy images. In: IEEE conference on cybernetics and intelligent systems, CIS, pp 76–81

Qi X (2008) Computer-aided diagnosis of early cancers in the gastrointestinal tract using optical coherence tomography. Ph.D. Dissertation, Case Western Reserve University

Rajivegandhi C, Shree ND, Khan S, Abinaya B (2015) Detection Of peptic ulcers based on thresholding and watershed segmentation. In: 3rd international conference on signal processing, communication and networking (ICSCN). IEEE, pp 1–5

Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. In: Classification in bioapps. Springer, Cham, pp 323–350

Report C (1990) Small cell carcinoma of the stomach case report and review of the literature. Dig Dis Sci 35(4):513–518

Riaz F, Ribeiro MD, Coimbra MT (2009) Quantitative comparison of segmentation methods for in-body images. In: 31st Annual international conference of the IEEE engineering in medicine and biology society: engineering the future of biomedicine, EMBC, pp 5785–5788

Riaz F, Areia M, Silva FB, Dinis-Ribeiro M, Nunes PP, Coimbra M (2011) Gabor textons for classification of gastroenterology images. In: ieee international symposium on biomedical imaging: from nano to macro. IEEE, pp 117–120

Riaz F, Silva FB, Ribeiro MD, Coimbra MT (2012) Invariant gabor texture descriptors for classification of gastroenterology images. IEEE Trans Biomed Eng 59(10):2893–2904

Riaz F, Silva FB, Ribeiro MD, Coimbra MT (2013) Impact of visual features on the segmentation of gastroenterology images using normalized cuts. IEEE Trans Biomed Eng 60(5):1191–1201

Riaz F, Hassan A, Nisar R, Dinis-Ribeiro M, Coimbra MT (2017) Content-adaptive region-based color texture descriptors for medical images. IEEE J Biomed Health Inf 21(1):162–171

Ribeiro M (2005) Clinical endoscopic and laboratorial assessment of patients with associated lesions to gastric adenocarcinoma,. Ph.D. Dissertation, Faculdade de Medicina da Universidade do Porto, Ph.D. thesis

Rogy MA, Bünger MA (2015) The historical perspective of gastric cancer. In: Gastric cancer. Springer, pp 3–21

Roukos DH, Agnantis NJ, Fatouros M, Kappas AM (2002) Gastric cancer: introduction, pathology, epidemiology. Gastric Breast Cancer 1(1):1–3

Sainju S, Bui FM, Wahid KA (2014) Automated bleeding detection in capsule endoscopy videos using statistical features and region growing. J Med Syst 38(4):25

Sarwar SS, Panda P, Roy K (2017) Gabor filter assisted energy efficient fast learning convolutional neural networks. In: 2017 IEEE/ACM international symposium on low power electronics and design (ISLPED), IEEE, pp 1–6

Schlag C, Menzel C, Nennstiel S, Neu B, Phillip V, Schuster T, Schmid RM, Von Delius S (2015) Emergency video capsule endoscopy in patients with acute severe GI bleeding and negative upper endoscopy results. Gastrointest Endosc 81(4):889–895

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117

Seidel R, Burdick JS (1998) Gastric leiomyosarcoma presenting as a gastric wall abscess. Am J Gastroenterol 93(11):2241–2244

Serpa-Andrade L, Robles-Bykbaev V, Gonzalez-Delgado L, Moreno JL (2016) An approach based on Fourier descriptors and decision trees to perform presumptive diagnosis of esophagitis for educational purposes. In: 2015 IEEE International autumn meeting on power, electronics and computing (ROPEC). IEEE

Sevakula RK, Singh V, Verma NK, Kumar C, Cui Y (2018) Transfer learning for molecular cancer classification using deep neural networks. IEEE/ACM Trans Comput Biol Bioinf. https://doi.org/10.1109/TCBB.2018.2822803

Ševo I, Avramović A, Balasingham I, Elle OJ, Bergsland J, Aabakken L (2016) Edge density based automatic detection of inflammation in colonoscopy videos. Comput Biol Med 72:138–150

Sharma H, Zerbe N, Klempert I, Hellwich O, Hufnagl P (2017) Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Comput Med Imaging Gr 61:2–13

Shichijo S, Nomura S, Aoyama K, Nishikawa Y, Miura M, Shinagawa T, Takiyama H, Tanimoto T, Ishihara S, Matsuo K et al (2017) Application of convolutional neural networks in the diagnosis of helicobacter pylori infection based on endoscopic images. EBioMedicine 25:106–111

Shin D, Lee MH, Polydorides AD, Pierce MC, Vila PM, Parikh ND, Rosen DG, Anandasabapathy S, Richards-Kortum RR (2016) Quantitative analysis of high-resolution microendoscopic images for diagnosis of neoplasia in patients with Barrett's esophagus. Gastrointest Endosc 83(1):107–114

Siegel RL, Miller KD, Jemal A (2015) Cancer statistics, 2015. CA A Cancer J Clin 65(1):5–29

Sobri Z, Amylia H, Sakim M (2012) Texture color fusion based features extraction for endoscopic gastritis images classification. Int J Comput Electr Eng 4(5):674–678

Society AC (2016) Stomach Cancer. www.cancer.org. [Online]. http://www.cancer.org/acs/groups/cid/documents/webcontent/003141-pdf.pdf

Song M, Ang TL (2014) Early detection of early gastric cancer using image-enhanced endoscopy: current trends. Gastrointest Interv 3(1):1–7

Sousa D-RMAM, André, Coimbra M (2009) Identifying cancer regions in vital-stained magnification endoscopy images using adapted color histograms. In: 16th IEEE international conference on image processing (ICIP), pp 681–684

Spyridonos P, Vilariño F, Vitrià J, Azpiroz F, Radeva P (2006) Anisotropic feature extraction from endoluminal images for detection of intestinal contractions. In: International conference on medical image computing and computer-assisted intervention, MICCAI, vol 9, pp 161–168

Swannell R (2010) World Cancer Report 2014, The Globe, no. 1, pp 6–7

Szczypiński P, Klepaczko A (2009) Selecting texture discriminative descriptors of capsule endpscopy images. In: 6th international symposium on image and signal processing and analysis, ISPA, pp 701–706

Szczypiński P, Klepaczko A, Pazurek M, Daniel P (2014) Texture and color based image segmentation and pathology detection in capsule endoscopy videos. Comput Methods Programs Biomed 113(1):396–411

Takita J, Kato H, Miyazaki T, Nakajima M, Fukai Y, Masuda N, Manda R, Fukuchi M, Kuwano H (2005) Primary squamous cell carcinoma of the stomach: a case report with immunohistochemical and molecular biologic studies. Hepatogastroenterology 52(63):969–974

Thekkek N, Lee MH, Polydorides AD, Rosen DG, Anandasabapathy S, Richards-Kortum R (2015) Quantitative evaluation of in vivo vital-dye fluorescence endoscopic imaging for the detection of Barrett's-associated neoplasia. J Biomed Opt 20(5):056002

Tischendorf JJW, Gross S, Winograd R, Hecker H, Auer R, Behrens A, Trautwein C, Aach T, Stehle T (2010) Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. Endoscopy 42(3):203–207

Tjoa MP, Krishnan SM, Doraiswami R (2002) Automated diagnosis for segmentation of colonoscopic images using chromatic features. In: IEEE Canadian conference on electrical and computer engineering, CCECE, vol 2, pp 1177–1180

Turan M, Almalioglu Y, Araujo H, Konukoglu E, Sitti M (2018) Deep EndoVO: a recurrent convolutional neural network (RCNN) based visual odometry approach for endoscopic capsule robots. Neurocomputing 275:1861–1870. https://doi.org/10.1016/j.neucom.2017.10.014

Turcza P, Duplaga M (2011) Low power FPGA-based image processing core for wireless capsule endoscopy. Sens Actuators A Phys 172(2):552–560

Ueo T, Yonemasu H, Yada N, Yano S, Ishida T, Urabe M, Takahashi K, Nagamatsu H, Narita R, Yao K, Daa T, Yokoyama S (2013) White opaque substance represents an intracytoplasmic accumulation of lipid

droplets: immunohistochemical and immunoelectron microscopic investigation of 26 cases. Dig Endosc 25(2):147–155

van der Sommen F, Zinger S, Schoon EJ, de With PHN (2014) Supportive automatic annotation of early esophageal cancer using local gabor and color features. Neurocomputing 144:92–106

Van Gossum A (2015) Image-enhanced capsule endoscopy for characterization of small bowel lesions. Best Pract Res Clin Gastroenterol 29(4):525–531

Vécsei A, Fuhrmann T, Uhl A (2008) Towards automated diagnosis of celiac disease by computer-assisted classification of duodenal imagery. In: 4th IET international conference on advances in medical, signal and information processing, MEDSIP. IET, pp 1–4

Vécsei A, Fuhrmann T, Liedlgruber M, Brunauer L, Payer H, Uhl A (2009) Automated classification of duodenal imagery in celiac disease using evolved Fourier feature vectors. Comput Methods Programs Biomed 95(2):S68–S78

Vieira PM, Ramos J, Lima CS (2015) Automatic detection of small bowel tumors in endoscopic capsule images by ROI selection based on discarded lightness information. In: 37th Annual international conference of the IEEE on engineering in medicine and biology society (EMBC), pp 3025–3028

Wallace MB, Keisslich R (2010) Advances in endoscopic imaging of colorectal neoplasia. Gastroenterology 138(6):2140–2150

Wang S, Cong Y, Fan H, Liu L, Li X, Yang Y, Tang Y, Zhao H, Yu H (2016a) Computer-aided endoscopic diagnosis without human specific labeling. IEEE Trans Biomed Eng 9294:1

Wang S, Cong Y, Cao J, Yang Y, Tang Y, Zhao H, Yu H (2016b) Scalable gastroscopic video summarization via similar-inhibition dictionary selection. Artif Intell Med 66:1–13

Wang G, Li W, Zuluaga MA, Pratt R, Patel PA, Aertsen M, Doel T, David AL, Deprest J, Ourselin S, Vercauteren T (2018a) Interactive medical image segmentation using deep learning with image-specific fine-tuning. IEEE Trans Med Imaging 37(7):1562–1573

Wang T, Dong Y, Yang C, Wang L, Liang L, Zheng L, Pu J (2018b) Jumping and refined local pattern for texture classification. IEEE Access 6:64416–64426

Wong Kee Song LM, Adler DG, Chand B, Conway JD, Croffie JMB, DiSario JA, Mishkin DS, Shah RJ, Somogyi L, Tierney WM, Petersen BT (2007) Chromoendoscopy. Gastrointest Endosc 66(4):639–649

Wu X, Chen H, Gan T, Chen J, Ngo C-W, Peng Q (2016) Automatic hookworm detection in wireless capsule endoscopy images. IEEE Trans Med Imaging 35(7):1741–1752

Yamaguchi J, Yoneyama A, Minamoto T (2015)Automatic detection of early esophageal cancer from endoscope image using fractal dimension and discrete wavelet transform. In: 12th International conference on information technology-new generations, (ITNG). IEEE, pp 317–322

Yan R, Park J-H, Choi Y, Heo C-J, Yang S-M, Lee LP, Yang P (2012) Nanowire-based single-cell endoscopy. Nat Nanotechnol 7(3):191–196

Yang X, Ye X, Slabaugh G (2015) Multilabel region classification and semantic linking for colon segmentation in CT colonography. IEEE Trans Biomed Eng 62(3):948–959

Yao R, Zhang S, Yang W, Cheng S, Chen Y (2010) Abnormality detection on gastroscopic images using patches assembled by local weights. In: 2010 International conference on medical image analysis and clinical application, MIACA 2010, pp 38–41

Yu J-S, Chen J, Xiang Z, Zou Y-X (2015) A hybrid convolutional neural networks with extreme learning machine for wce image classification. In: IEEE international conference on robotics and biomimetics (ROBIO), pp 1822–1827

Yuan Y, Meng M-H (2015) Automatic bleeding frame detection in the wireless capsule endoscopy images. In: IEEE international conference on robotics and automation, ICRA, pp 1310–1315

Yuan Y, Li B, Meng Q (2015a) Bleeding frame and region detection in the wireless capsule endoscopy video. IEEE J Biomed Health Inf 2194:1–1

Yuan Y, Wang J, Li B, Meng MQH (2015b) Saliency based ulcer detection for wireless capsule endoscopy diagnosis. IEEE Trans Med Imaging 34(10):2046–2057

Yuan Z, IzadyYazdanabadi M, Mokkapati D, Panvalkar R, Shin JY, Tajbakhsh N, Gurudu S, Liang J (2017a) Automatic polyp detection in colonoscopy videos. In: Medical imaging 2017: image processing, vol. 10133. International Society for Optics and Photonics, p 101332K

Yuan Y, Yao X, Han J, Guo L, Meng MQH (2017b) Discriminative joint-feature topic model with dual constraints for WCE classification. IEEE Trans Cybern 48(7):2074–2085

Zhang S, Yang W, Wu YL, Yao R, Cheng SD (2009) Abnormal region detection in gastroscopic images by combining classifiers on neighboring patches. In: International conference on machine learning and cybernetics vol 4, pp 2374–2379

Zhang Z, Bai L, Ren P, Hancock ER (2016) High-order graph matching kernel for early carcinoma EUS image classification. Multimed Tools Appl 75(7):3993–4012

Zhang R, Zheng Y, Mak TWC, Yu R, Wong SH, Lau JY, Poon CC (2017) Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. IEEE J Biomed Health Inf 21(1):41–47

Zhao Q, Mullin GE, Meng MQ-H, Dassopoulos T, Kumar R (2015) A general framework for wireless capsule endoscopy study synopsis. Comput Med Imaging Gr 41:108–16

Zou Y, Li L, Wang Y, Yu J, Li Y, Deng WJ (2015) Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network. In: IEEE international conference on digital signal processing (DSP), pp 1274–1278

Affiliations

Hussam Ali[1] [ORCID] **· Muhammad Sharif[1] · Mussarat Yasmin[1] · Mubashir Husain Rehmani[2] · Farhan Riaz[3]**

Muhammad Sharif
muhammadsharifmalik@yahoo.com

Mussarat Yasmin
mussaratabdullah@gmail.com

Mubashir Husain Rehmani
mshrehmani@gmail.com

Farhan Riaz
farhan.riaz@ceme.nust.edu.pk

[1]   Department of Computer Science, COMSATS University Islamabad, Wah Campus, Islamabad, Pakistan

[2]   Department of Computer Science, Cork Institute of Technology, Cork, Ireland

[3]   Department of Computer & Software Engineering, CE&ME, National University of Sciences and Technology (NUST), Islamabad, Pakistan