REGION-COLOR BASED AUTOMATED BLEEDING DETECTION IN CAPSULE ENDOSCOPY VIDEOS

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Electrical and Computer Engineering

University of Saskatchewan

Saskatoon, SK, Canada

By

SONU SAINJU

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering

57 Campus Drive

University of Saskatchewan

Saskatoon, Saskatchewan

Canada, S7N 5A9

ABSTRACT

Capsule Endoscopy (CE) is a unique technique for facilitating non-invasive and practical visualization of the entire small intestine. It has attracted a critical mass of studies for improvements. Among numerous studies being performed in capsule endoscopy, tremendous efforts are being made in the development of software algorithms to identify clinically important frames in CE videos. This thesis presents a computer-assisted method which performs automated detection of CE video-frames that contain bleeding.

Specifically, a methodology is proposed to classify the frames of CE videos into bleeding and non-bleeding frames. It is a Support Vector Machine (SVM) based supervised method which classifies the frames on the basis of color features derived from image-regions. Image-regions are characterized on the basis of statistical features. With 15 available candidate features, an exhaustive feature-selection is followed to obtain the best feature subset. The best feature-subset is the combination of features that has the highest bleeding discrimination ability as determined by the three performance-metrics: accuracy, sensitivity and specificity. Also, a ground truth label annotation method is proposed in order to partially automate delineation of bleeding regions for training of the classifier.

The method produced promising results with sensitivity and specificity values up to 94%. All the experiments were performed separately for RGB and HSV color spaces. Experimental results show the combination of the mean planes in red and green planes to be the best feature-subset in RGB (Red-Green-Blue) color space and the combination of the mean values of all three planes of the color space to be the best feature-subset in HSV (Hue-Saturation-Value).

ACKNOWLEDGMENTS

This research was supported by Grand Challenges Canada (GCC) Star in Global Health and Natural Science and Engineering Research Council of Canada (NSERC). I would like to acknowledge Canada Foundation for Innovation (CFI) and Canadian Microelectronics Corporation (CMC) for providing the lab infrastructure and development tools.

I owe a debt of gratitude to my supervisors Dr. Khan Wahid and Dr. Francis M. Bui, without whom this research project would not have been possible. I am most grateful for their guidance and valuable advice. I also thank them for providing a supportive environment for the research.

My special appreciation goes to my husband for his kindness and encouragement. I also acknowledge the members of my research group for their help and support.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS

Following are the publications related to this work by the author:

- Sainju Sonu, Francis M. Bui, and Khan A. Wahid. "Automated Bleeding Detection in Capsule Endoscopy Videos Using Statistical Features and Region Growing." *Journal of medical systems* 38.4 (2014): 1-11.

- Sainju Sonu, Francis M. Bui, and Khan Wahid. "Bleeding detection in wireless capsule endoscopy based on color features from histogram probability." *Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference* on. IEEE, 2013.

LIST OF ABBREVIATIONS

3D	Three Dimensional
ANN	Artificial Neural Network
CAD	Computer Aided Diagnosis
CE	Capsule Endoscopy
CIE	International Commission on Illumination
CWC	Color Wavelet Covariance
DCT	Discrete Cosine Transform
EM	Expectation-Maximization
FDA	Food and Drug Administration
FN	False Negative
FP	False Positive
GI	Gastrointestinal
HSI	Hue-Saturation-Intesity
HSV	Hue-Saturation-Value
LBP	Local Binary Pattern
MLP	Multilayer Perceptron
NN	Neural Network
OGIB	Obscure Gastrointestinal Bleeding
PDF	Probability Density Function
RBF	Radial Basis Function
RGB	Red-Green-Blue
ROI	Region of Interest
RR	Red Ratio

SBI	Suspected Blood Indicator
SFFS	Sequential Floating Forward Selection
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
VCE	Video Capsule Endoscopy
VSCH	Vector Supported Convex Hull Method
WCE	Wireless Capsule Endoscopy

Chapter 1
Introduction

This chapter introduces the scope of computer-aided methods in the automated detection of bleeding in Capsule Endoscopy (CE) images. It begins with a brief introduction to the CE technology and its scope in the diagnosis of bleeding related abnormalities. The motivation behind the use of region-based classification technique is its ability to utilize information from the local regions of the video frames. The limited exploration of this technique in the existing works related to bleeding detection has also provided a motivation in investigating the approach in this field. The chapter presents the motivation of the thesis by providing the past and current research scenarios in the use of various types of classification techniques. The objectives and organization of the thesis are finally presented at the end.

1.1 Capsule Endoscopy

Capsule Endoscopy (CE), also called Video Capsule Endoscopy (VCE) or Wireless Capsule Endoscopy (WCE) is a unique endoscopic technique which was first developed by an Israel based company called Given Imaging. It uses a different approach to the examination of the GI tract than the thin lighted-tube approach which is inserted down the throat. Though the standard endoscopic techniques are able to perform both diagnostic and limited therapeutic functions, these techniques are painful and uncomfortable to the patients. CE technology was thus developed as an innovative way to perform non-invasive examination of the GI tract. This technique employs a pill-shaped miniaturized device which captures and transmits the images from the interior of the GI tract of the patients. The original CE developed by Given Imaging was the first CE device to be approved

by the US Food and Drug Administration (FDA) in 2001 AD [1]. Over subsequent years, CE systems from other manufactures also entered the commercial market. Currently, three CE systems are approved by the FDA for the diagnosis of the small bowel: *PillCam SB2*, Given Imaging Ltd, Yoqneam, Israel; *Endocapsule*, Olympus America Inc., Center Valley, Pennsylvania; and *MicroCam*, IntroMedic Company Ltd, Seoul, Korea [1]. *PillCam Colon* of the Given Imaging has also been recently cleared by the FDA for the visualization of the colon [2]. Given Imaging has developed CE systems for visualizing esophagus.

1.1.1 CE in the Diagnosis of Bleeding Related Abnormalities

GI tract bleeding [3] accounts for approximately 300,000 hospitalizations per year in the United States [4]. In majority of the cases, the source of bleeding can be found with the traditional diagnostic methods. However, the source of bleeding cannot be readily identified in approximately 5% of patients [5]. These bleeding incidences with unknown origin are called Obscure Gastrointestinal Bleeding (OGIB). The small bowel has been known to be one of the prime locations of lesions leading to OGIB. The visualization of the entire small intestine is not possible with upper endoscopy and colonoscopy. Technologies like CE and push enteroscopes have recently been advanced to make the small intestine accessible for examination [6]. Also, among these two methods, it was studied that CE is preferred by patients because of less pain and more ease and comfort associated with the procedure [7]. Other works have also shown that CE has larger range of access to the small intestine, and thus can identify more bleeding sites than push enteroscopes [7]. CE has thus established itself as a gold standard tool in the examination of the small bowel.

1.1.2 Scope of Computer Aided Bleeding Detection Systems in CE

Standard endoscopy and capsule endoscopy both have the ability to view the GI tract directly. However, the two systems have different operation scenarios. In standard endoscopy procedures, the diagnosis process is controlled by the physician. Given that the whole process is viewed by the physician in real time anyway; manual marking of suspicious regions is not time consuming. But in case of CE, the pill-shaped device moves along the GI tract in a manner similar to food, which is through natural peristalsis. Thus, the diagnosis process is not controllable as the camera cannot be moved according to the desire of the endoscopist. The recorded video is then viewed later for examination. The real time viewing of the process is infeasible as the complete process lasts for an average of 8 hours. Also, the process produces around 55,000 frames per patient per examination and it requires around 2 hours of time for a physician to inspect the video [8].

Thus, there is a high scope of computer assisted diagnosis tools in CE for making the diagnosis more accurate, reliable and fast. The inspection time of the video can be greatly reduced if the suspicious frames could be recognized by a computer system. The suspicious frames thus selected could then be presented to the physician for the final decision. Given Imaging (Yoqneam, Israel) provides a tool called Suspected Blood Indicator (SBI) that detects red pixel in the images [9]. However, studies have shown that performance of SBI is not sufficient to screen all types of bleeding in the GI tract [10], [11]. This has motivated a lot of studies in the development of computer assisted diagnosis tools to automatically detect bleeding areas in capsule endoscopy images.

1.2 Research Motivation

A number of studies related to Computer Aided Diagnosis (CAD) in capsule endoscopy are pattern recognition problems that perform classification of images into different classes according to the image features. Image features characterize certain properties of images. Features that describe image properties such as color, texture, brightness, contrast etc. have been widely studied and used in pattern recognition involving images. Classification itself can be of various types based on the identification area, which is the area of image that is used in the calculation of feature. It can range from whole image and image fragments to even the smallest image element i.e. the image pixel. The choice of an identification area for the mathematical formulation of the features depends on its appropriateness to the aim of the image analysis. According to the identification area, classification is performed in four ways: pixel-based, patch-based, image-based and region-based.

Pixel based methods analyze and classify every pixel in the images. Such a scheme was followed in [12], [13]. Both of these methods require multi-dimensional feature vector for processing every pixel in an image. Classification based on this kind of feature extraction scheme requires high computation cost for a single image, which gets even worse with higher resolution images. Since thousands of images need to be examined for a complete diagnosis, these methods do not seem to be effective for real application. Also, features derived from single pixel do not account for relations that might exist between the pixel and its local and global neighborhood.

Image based methods generate features utilizing the area of the whole image. The features thus characterize the whole images which are then subject to classification. In [14], [15], [16], image-based features like local color histogram, Color Wavelet Covariance (CWC), spatial pyramids are utilized that can carry both global and local information. However, this kind of classification

scheme is highly likely to misclassify images containing small bleeding area. Also, delineation of the bleeding areas is not possible with this kind of scheme.

An intermediate area of identification is thus generally preferred in capsule endoscopy images where features are generated from image patches. Image patches are derived by dividing an image into blocks of fixed shape and size. One such example is to divide an image into rectangular blocks of 8x8 pixels. These image patches characterize local image features and are computationally less rigorous than pixel based schemes. Authors in [17] performed a classification based on image patches by dividing images (256 × 256 pixel resolution) into rectangular blocks of 30×30 pixels and similarly, in [18] classification based on blocks of 10×10 pixel on images of 400×400 resolution is performed. The authors in [19] point out that circular ROIs provide a good inclusion of abnormality. They used overlapping circular Regions of Interest (ROIs) in the calculation of features and perform classification based on the circular ROIs.

However, [20] suggests that lesions are generally of arbitrary shapes and sizes which makes fixed sized rectangular or circular patches unsuitable to serve as effective ROIs. Fixed shaped and sized ROIs are prone to background noise especially in case of smaller lesions. The authors in. [20] further show that all the features used in the study performed better classification when the features were derived from the image regions. Image regions are image partitions of arbitrary shape and size which are derived by dividing an image into several pixel groups such that the individual groups contain pixels which are similar to each other with respect to some criteria. In [20], mean-shift segmentation was used to divide images into homogeneous regions and then features such as 3D histogram, LBP (Local Binary Pattern), MS-CWC, were calculated to detect lesion in gastroscopic images. Thus, a promising area of study has been opened to explore region based classification approach in detecting bleeding areas too. The segmentation scheme used in [20]

involves optimization of multiple parameters for optimal segmentation performance. It thus provides us a motivation to propose a region selection algorithm which requires minimal parameters and optimization. So, in this thesis, a region-based classification system was developed to accurately classify CE frames into bleeding and non-bleeding classes by using histogram related features.

1.3 Thesis Objectives

The overall objective of this thesis is to develop an efficient and accurate method to detect the frames of CE videos that contain possible bleeding occurrences by:

i) Extracting region features with bleeding discrimination capabilities.

ii) Devising a region-based classification system that can classify the frames of the CE videos into bleeding and non-bleeding classes.

iii) Selecting the best features from the available ones in order to achieve high classification performance.

1.4 Thesis Organization

This thesis consists of six chapters. Chapter 2 presents the overview of the existing methods that have been performed to automate bleeding detection in CE videos and images. Several feature computation techniques and classifiers are presented in the chapter along with the corresponding works that adopted the technique in their methodology. In chapter 3, the overview of the proposed methodology is illustrated. The chapter then presents the feature extraction techniques that were used in the proposed method. It also illustrates the feature-selection and classification approaches that were adopted in the thesis. Chapter 4 presents all the results and experiments that were

performed to tune the parameters involved in order attain accurate detection of the bleeding frames in CE videos. Finally, in chapter 5, a summary of the accomplishments is presented along with various recommendations for potential future.

.

Chapter 2

Bleeding Detection in CE Images

This chapter offers a survey on some representative methodologies in the area of bleeding detection in CE, specifically identifying those techniques which are relevant to fulfill the research objectives of this field. After a thorough evaluation of the published works, it was found that almost all the existing methodologies are classification methods that make use of Artificial Neural Networks (ANN), Support Vector Machines (SVM) and other machine learning methods in order to adapt to the problem at hand.

Figure 2- 1 shows the basic stages of a generalized classification system [21]. A classification problem starts with feature extraction and is followed by the feature extraction and classifier design stages. The system is then evaluated. It is, however, subject to redesign at any stage in order to improve the overall performance. This chapter presents various techniques adopted by the existing algorithms of bleeding detection in relation to the basic stages of a general classification system as presented in Figure 2- 1.

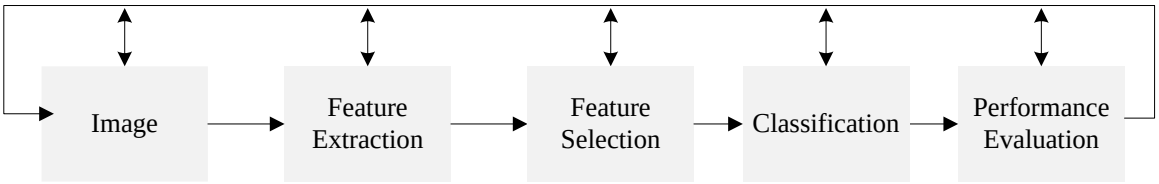

Figure 2- 1 Basic stages of a classification system.

2.1 Feature Extraction

Measurable quantities must be derived in order to perform computer analyses in images or their fragments. Image features are numerical indicators that characterize color, texture, contrast, brightness and other aspects of image appearance. Feature extraction is thus the mathematical formulation of image properties. Three major aspects must be considered for feature extraction: color space selection; identification area; and type of features. The types of identification area and their corresponding advantages and limitations have been covered in section 1.2. It was argued that region-based features improve bleeding detection in CE video frames as the features are able to provide information from the local levels of image which are important in characterizing the image. Also, the image regions, which are used in the derivation of region-based features, are better suited for delineating the boundaries of bleeding areas. The classification time of a single frame can also be reduced significantly with region-based classification schemes than with the pixel-based ones.

Selection and transformation to an appropriate color space is a crucial part of feature extraction. The relevance of a color feature to a recognition problem is greatly influenced by the choice of color space. Color spaces aid in the specification of colors in some standard way. Typically, a color space constitutes of a 3D coordinate system where each point represents a constituent color of the model. RGB (Red-Green-Blue) and HSI (Hue-Saturation-Intensity) are the two most frequently used color spaces in bleeding detection applications in CE. The existing research works regarding bleeding discrimination have been performed using RGB only [23], [24], using HSI only [17], [25], [26] and using both RGB and HSI [12], [15]. Other color spaces such as CIE (International Commission on Illumination) Lab, CIE LUV, YCbCr, YUV have also been utilized by some studies [19], [20].

2.1.1 Features for Bleeding Discrimination in CE images

As in Figure 2- 2, bleeding regions appear different from their surrounding tissues mainly in term of color. Color is one of the important cues used by clinicians to discriminate between bleeding and normal regions in CE images. Thus, a lot of studies have shown the usefulness of color features in this stream [12], [13], [15], [[25]-[27]]. Among the works performed with color related features, a few of them used the values of color pixel either directly or along with some modifications in order to characterize the color property. For example, [12] used feature vector *x = (R, G, B, H, S, I)* for every pixel in the image; where, *R*, *G*, and *B* are the red, green and Blue color values of the pixel in RGB color space and *H*, *S*, and *I* are the Hue, Saturation and Intensity values of the pixels in HSI color space. Similarly, [13] used Red Ratio (RR) features which were calculated for every picture element. The authors defined three RR features as follows: $F_1 = \frac{R(i,j)}{G(i,j)}$, $F_2 = \frac{R(i,j)}{B(i,j)}$, $F_3 = \frac{R(i,j)}{R(i,j)+G(i,j)+B(i,j)}$; where *R(i,j), G(i,j), B(i,j)* are the red, green and blue pixel values of the pixel at the position (i, j) of the image.

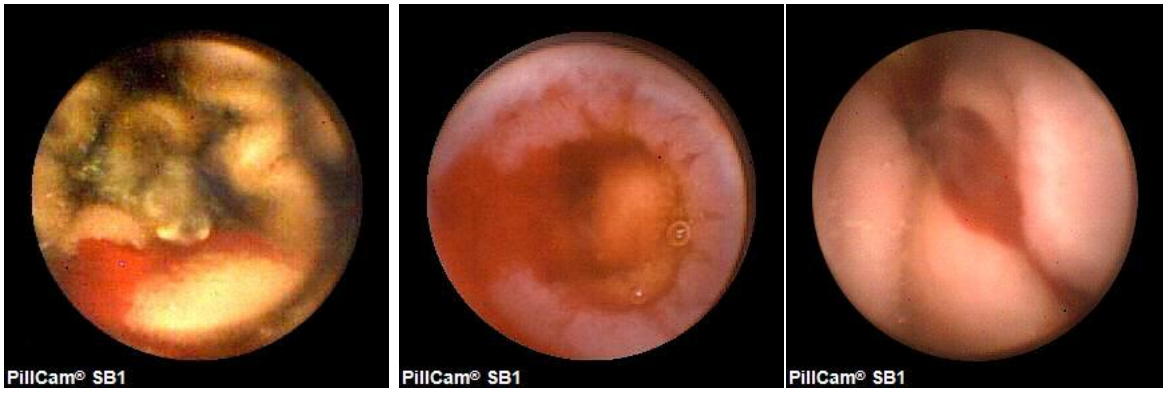

Figure 2- 2 CE Images with visible bleeding areas. The images are taken from ref. [28].

Usage of pixel values at every location is not an effective way of representing image color. Histogram presents an effective way to represent color distribution of images. It is a direct way to show how individual color levels are occupied in an image. However, global histograms are rarely used as they do not provide local information regarding color distribution. Images are, thus, divided into non-overlapping blocks of fixed size, so that local histograms can individually be built from the blocks. Also, inclusion of all the color levels requires a large structure, due to which sampling of each color channels is performed in order to achieve smaller structures. Such a color feature, proposed by the authors in [29], has been used in [15] to characterize images for bleeding detection. The authors in [15] divided an image of 576×576 pixels into 9 separate blocks and computed local histogram from each. It divided H, S, and V channels of HSV space uniformly into 12, 5 and 8 bins respectively. The feature vector computed in such a way consisted 225 components. Similarly, [25] used 3D histogram with 32×32×24 sampling in the HSV color space to create the 3D feature structure. The structure was compressed using 3D Discrete Cosine Transform (DCT) into 286 components long feature vector.

Besides the color features, a number of works has shown the usability of texture features that are calculated on the color channels for bleeding discrimination [14], [17]. The textural features characterize the pattern of pixels in an image and their relation to the surrounding. Authors in [14] claim that Color Wavelet Covariance (CWC) works suitably in the classification of CE images including bleeding images. CWC coefficients are the covariances of the second order textural measures between different channels of the used color space. In order to calculate CWC coefficients, each color channel is applied with three level Discrete Wavelet Transformation (DWT) images and upon using only the second level wavelet detailed channels, 9 sub images are obtained for each image. Then, for each sub image, co-occurrence matrices at four different

directions are determined which results in 36 matrices. The co-occurrence matrices contain the spatial dependence of the gray level intensities or the color intensities. Haralick proposed fourteen different texture measures from the co-occurrence matrices [30]. However in this study, only angular second moment, correlation, inverse difference moment and entropy were extracted from each of the 36 matrices, which resulted in 144 wavelet features, Finally, covariance of the same features in different color channels were calculated resulting in 72 feature components per image. CWC features were used for bleeding detection in [15] and [31] too. However, [15] claims that the color features derived from HSV histograms outperform the CWC features regardless of the choice of the classifying system. Furthermore, the calculation of the co-occurrence matrices takes longer time inducing longer time requirements for the calculations of CWC [20].

2.2 Classification

In machine learning, there is no universal classifier. The choice of the classifier depends on the available data and the application. Existing bleeding detection methods in CE images mostly employ supervised pattern recognition scheme. As the name suggests, supervised learning methods exploit a priori known information to design a classifier which, after being trained, can predict the labels of unseen data. Two types of classifiers are common in supervised pattern recognition: i) Based on Bayes decision theory; ii) Based on optimization of cost function.

2.2.1 Classifiers based on Bayes Decision Theory

Given M classes: $\omega_1, \omega_2, \ldots, \omega_M$, and an instance of unknown label represented by a feature vector x, Bayesian decision theory states that, x is assigned to the class ω_i if $P(\omega_i|x)>P(\omega_j|x)$, $\forall j \neq$

i. From Bayes rule: $P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{p(x)}$, where $p(x|\omega_i)$ is the class conditional probability density function (PDF) and p(x) is the pdf of *x*. Upon using the Bayes rule and given that p(x) is positive and same for all classes, Bayes decision rule can also be written as $p(x|\omega i)P(\omega_i) > p(x|\omega_j)P(\omega_j))$, $\forall j \neq i$. However, the underlying pdfs are not known and they have to be estimated from the available data. Bayesian based classifier is used in [23] to select blood pixels. It models the unknown probability models for blood and non-blood pixels via mixture model and Expectation Maximization (EM) algorithm.

2.2.2 Classifiers based on Optimization of Cost Function

The authors in [32] also claim that the classifiers based on the optimization of cost function perform better in presence of limited data. Also, the usage of the latter type of classifier is common in bleeding detection in CE images. These classifiers try to determine a decision surface separating the classes by directly using the data form the training set rather than the pdfs of the underlying data [32]. Among such schemes, Neural Networks (NNs) and Support Vector Machines (SVMs) are widely used in the computer systems built for CE images.

2.2.2.1 Neural Networks

A Neural Network is the network of basic perceptron elements which are also called neurons, analogous to the terminology used in neuroscience to refer to the basic building blocks of the brain. Multilayer Perceptron (MLP) or a feedforward neural network is a widely used neural network in real world problems which require multiple layers of neurons in order to achieve minimum error. In an MLP, neurons are connected as seen in Figure 2- 3, which is a two layer network. Here, the inputs, outputs and the hidden variables are represented by nodes and the links between the nodes

represent weight parameters. The links between the nodes and x_o and z_0 are the bias parameters. Overall network function for the network shown in Figure 2- 3 is then given as in equation (2.1) [33]. Where, $\boldsymbol{w}$ is the vector formed by grouping all weight and bias parameters. h(.) is a differentiable nonlinear activation function which transforms the input unit activations to hidden units as $z_j = h(a_j)$. Similarly, each output unit activation is transformed generally by a logistic sigmoid function σ (.) as $y_k = \sigma(a_k)$. Thus, the neural network model is a non-linear function which maps a set of input variables $\{x_i\}$ to the output variables $\{y_k\}$ controlled by a weight vector, $\boldsymbol{w}$. The weight vector is adjusted by reducing the difference between the inputs and their desired outputs.

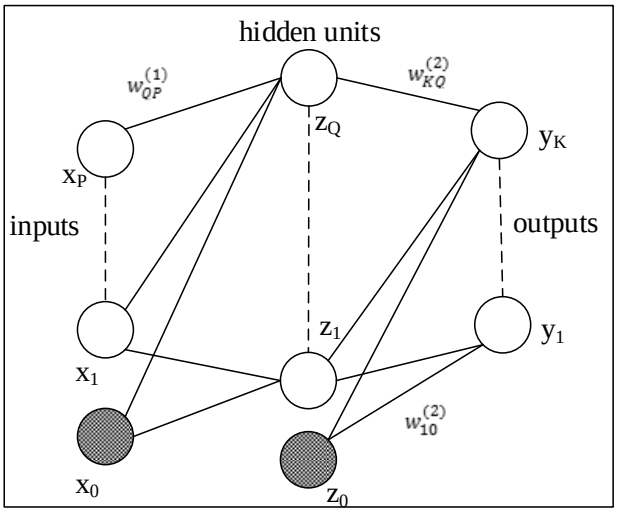

Figure 2- 3 A two layer feedforward neural network.

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma(\underbrace{\sum_{j=1}^{Q} w_{kj}^{(2)} h\left(\underbrace{\sum_{i=1}^{P} w_{ji}^{(1)} x_i + w_{j0}^{(1)}}_{a_j}\right) + w_{k0}^{(2)}}_{a_k}) \quad (2.1)$$

Because of powerful models and simple algorithms, multilayer neural networks find their application in many real world problems including bleeding detection in CE images. Use of the neural network in bleeding detection is common. In [13], neural networks were used to perform pixel-wise classification scheme. Similarly, classification was performed in [12] using a probabilistic neural network. Patch based classifications were performed in CE images in [17] and it also performed experiments to the performance of the classification system by varying the number of neurons in the hidden layer.

2.2.2.2 Support Vector Machines (SVM)

Support Vector Machines (SVMs) are supervised learning models which constructs an optimal hyperplane to classify data into different classes. A linearly separable two class case is presented in Figure 2- 4. A hyperplane can be described by the equation: $g(x) = \boldsymbol{w}^T\boldsymbol{x} + w_0 = 0$, that classifies the feature vectors x_i, *i=1, 2, …, N*, into two classes ω_1 and ω_2. As seen in Figure 2- 4, lines drawn parallel to the separating line are the supporting hyperplanes and the distance between them is called the margin. Width of the margin is constrained by support vectors which are the data points that are closest to the separating hyperplane. An issue of concern here, is that there exists an infinite number of possible hyperplanes between the data of the two classes. Every hyperplane is defined by its direction, determined by $\boldsymbol{w}$, and its position, w_0. Since, the optimal hyperplane is the one that separates the high probability density areas of two classes with maximum possible margin between them, the goal is to determine the direction that provides the maximum margin. Please refer [34] for the mathematical details.

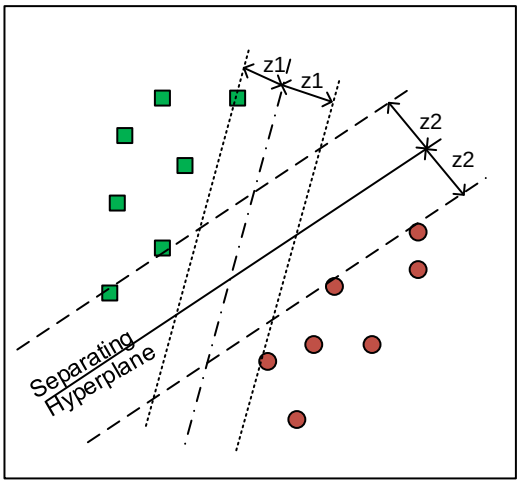

Figure 2- 4 Training a SVM requires finding an optimal hyperplane. Two such hyperplanes are shown above which separate members of two classes. Class 1 is shown as green squares and class 2 is shown as red circles.

SVM is one of the learning models which has been effectively used in many applications including pattern recognition in capsule endoscopy videos. It has been used in [35] to detect small bowel polyps and ulcers and to detect tumors in [36]. Another study [16], uses SVM with Color Invariant Descriptor and Pyramid of Color Invariant Histograms to detect bleeding images. Similarly in [25], SVM is used to detect blood frames by using HSI (Hue, Saturation and Intensity) and LBP (Local Binary Pattern) features. In [15], classifications of the CE frames were performed using both SVM and NN. The study reports slightly better classification with SVM than with NN. However, the performance of SVM and NN depend on many parameters and their direct comparison is difficult. Both NN and SVM have been successfully used in image analysis in CE proving the usability of both techniques.

2.3 Performance Measures

Performance evaluation in a classification problem includes determining the classification performance of the system. Sensitivity, specificity and accuracy are three ideal criteria for measuring classification performance. Sensitivity is the probability of correctly labeling the members of positive cases. Specificity is the probability of correctly labeling negative cases. Accuracy is the probability of correctly labeling both positive and negative cases. They are given as in equation (2.2) [22].

$$sensitivity, (Sen.) = \frac{TP}{(TP + FN)}$$

$$specificity, (Spec.) = \frac{TN}{(TN + FP)}$$

$$accuracy, (Acc.) = \frac{(TP + TN)}{(TP + FN + TN + FP)} \tag{2.2}$$

TP, FP, TN, FN are the counts of True Positive, False Positive, True Negative and False Negative events respectively. Almost all the bleeding detection methods have utilized the aforementioned metrics for system evaluation. However, all the methods calculate these measures using their own database. The unavailability of standard dataset or image database for studies related to CE images and videos makes it difficult to perform direct comparison between different methods. Thus, this thesis doesn't provide quantitative comparison between the methods.

Chapter 3

Proposed Methodology

In this chapter, the overview of the proposed bleeding detection method is presented at first and then, the feature computation, classification, feature selection and performance evaluation schemes utilized in the proposed methodology are presented. Feature extraction scheme includes the steps that are required to characterize images with measurable quantities. These quantities, called features, are calculated using the image data. The chapter first presents the features and the color spaces that were selected to characterize the image regions. It then presents the region selection schemes adopted to partition images into constituent regions. The adopted region selection method allows fast image segmentation with minimal human intervention. The ground truth annotation method is then proposed. This method automates the annotation method partially. Finally, classification, feature selection schemes are presented. Performance measures adopted for evaluation are presented at the end of the chapter.

3.1 Overview of the Proposed Bleeding Detection Method

A region based classification scheme is proposed in this thesis for bleeding detection in CE images. This classification system makes use of region features to classify image regions into bleeding and non-bleeding classes. Like any other supervised classification scheme, the proposed methodology contains two basic stages: the classifier learning stage and the application stage. Figure 3- 1 illustrates an overview of the method. The classifier learning stage determines the optimized classifier through the use of data with known class labels and the application stage makes use of the optimized classifier to classify different sets of data with unknown class labels.

At the classifier learning stage, the classifier is presented with the training data by the use of which the classifier learns the discriminating plane between the classes. The classifier is then tested to with a set of test images to evaluate the classification performance of the classifier. The classification performance of a classifier is significantly affected by the features used in the training and testing. Since, multiple features are calculated from the image regions, feature selection is performed to optimize the classifier performance. Multiple learning and testing operations are performed to optimize the classifier. And, metrics providing numerical measures of the classification performance of the classifier are used to evaluate the system at every stage of feature selection process.

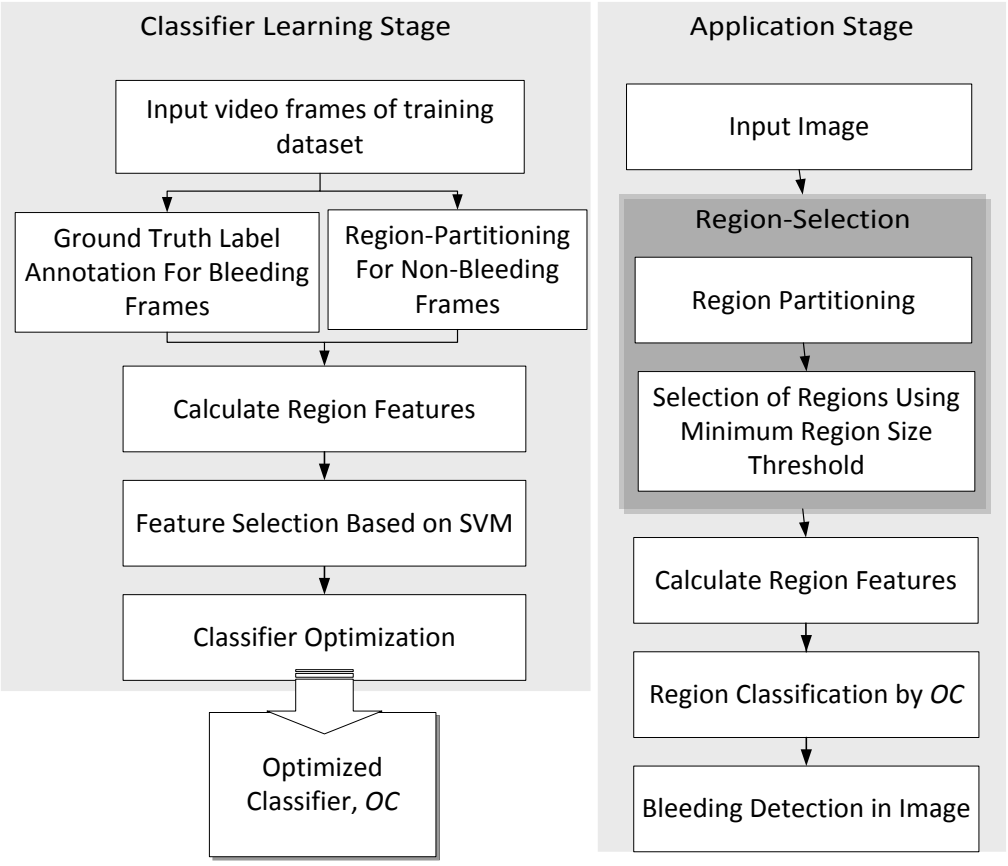

Figure 3- 1 Overview of the proposed bleeding detection method.

The training and testing data used in both the stages are the features obtained from image regions. Automatic region selection method, as described in section 3.5.1, is applied to automatically segment the image into constituent regions in the application stage. Similarly, non-bleeding images are also segmented using the same method for obtaining ground truth labels for non-bleeding class. However, the ground truth labels for bleeding regions are labeled from the proposed semi-automated ground truth label annotation method.

The partitioning of image into constituent regions, selecting suitable color spaces and image features and finally calculating those features are the steps required for characterizing images or image regions. These steps jointly form the image feature extraction scheme of the proposed method, which will further be discussed in this chapter.

3.2 Feature Extraction: Statistical Features

Bleeding manifests itself by appearing redder in color than its surrounding or tends to have lower intensity than normal tissues. Color is one of the important cues used by clinicians to discriminate between bleeding and normal regions in capsule endoscopic images. However, variation in illumination, variation in the shades and extent of bleeding, presence of other fluids and debris pose complexity in detecting bleeding in capsule endoscopy images. Thus, bleeding identification based on limiting threshold of color values might be erroneous. Hence, features that hold information on the color distribution of images are useful to identify bleeding and non-bleeding regions.

Histogram is a graphical representation of frequency of occurrences of intensity levels of images or image regions. Histograms are widely used in image analysis due to its simplicity and robustness to variations in scale and translation. In color images, color distribution of each plane

can be modeled by its individual histogram. Histograms of the planes of a color space model the color distribution of images represented in that space. For image analysis, numerical indicators like mean, standard deviation, entropy, skew and energy can be calculated from the histograms of the color planes. These measures characterize the histograms of the image color planes, and thus convey information about the color distribution of images. These features have been successfully used in image classification and analysis purposes in studies [37], [38]. The histograms and these statistical features do not provide local color information of the image if they are directly obtained from the whole image. But, if the features are measured from smaller parts of the image, the local color information can be extracted. These statistical measures were thus calculated from the color channels of the different image regions which were obtained from the appropriate region selection method.

For simplicity and uniformity, all the images used in this work were converted to an 8-bit image. For 8-bit images, each color plane consists of 256 intensity levels and hence there would be 256 different entries in the histograms of each plane. Thus, if *L* is the number of bins in each histogram, width of each bin, w, is given by256/*L*. The features were calculated from the first order histogram probability, *P(g)*, of the different color channels of the image regions which is given as in equation (3.1) [39].

$$P(g(i)) = \frac{N(g(i))}{NumPixel} \tag{3.1}$$

Where, $g(i) = \frac{w-1}{2} + w * i$; *[i=0, 2 … L-1]*.

Here, *NumPixel* is the total number of pixels in the image region being considered and *N(g(i))* is the number of pixels with the level value *g(i)*. *g(i)* is the value of the bin center, given as in equation (3.1). Following are the statistical measures which were used as the region features:

3.2.1 Mean (*M*)

Mean of an image or any region of an image gives its average brightness. It is calculated as given in equation (3.2) [39].

$$\mathrm{M} = \sum_{i=0}^{L-1} g(i)\, P(g(i)) \qquad (3.2)$$

3.2.2 Standard deviation (*S*)

Variance of an image region gives the spread of the pixel intensities around the mean intensity value. A higher value of variance denotes a higher contrast of pixel intensities in the region; whereas, a lower value indicates a lower contrast in the image region. Square root of variance gives the standard deviation. It is defined as given in equation (3.3) [39].

$$\mathrm{S} = \sqrt{\sum_{i=0}^{L-1} (g(i) - \mathrm{M})^2\, P(g(i))} \qquad (3.3)$$

3.2.3 Entropy (*N*)

Entropy is a measure which gives uncertainty or information content. Complex images, in which pixel values change unexpectedly, have larger entropy values. It is defined as follows in equation (3.4) [39].

$$\mathrm{N} = -\sum_{i=0}^{L-1} P\big(g(i)\big)\ \log_2 P(g(i)) \tag{3.4}$$

3.2.4 Skew (*K*)

Skew of a probability distribution gives a measure of its asymmetry. A positive value of skew means the distribution has a longer tail at the left side and similarly a negative sign implies a longer tail at the right side of the distribution. It is defined as follows in equation (3.5) [39].

$$\mathrm{K} = \frac{1}{S^3}\sum_{i=0}^{L-1}(g(i) - \mathrm{M})^3 P(g(i)) \tag{3.5}$$

3.2.5 Energy (*E*)

Energy tells us about the way the pixel levels are distributed in a region. An energy value of 1 suggests that the region has a constant pixel level. Accordingly, it is understood that, as the value of energy gets lower, the pixel values are distributed over more number of color pixel levels. It is defined in equation (3.6) [39].

$$\mathrm{E} = \sum_{i=0}^{L-1}[P(g(i))]^2 \tag{3.6}$$

3.3 Feature Extraction: Color Space Selection

Two color spaces: RGB (Red-Green-Blue) and HSV (Hue-Saturation-Value) were used for representing the images and for calculating the statistical features. All the regions were attributed

with five features listed as in section 3.2 in all three channels of RGB and HSV color space. All the used features and their corresponding symbols are summarized in Table 3- 1. RGB color space was utilized in this work as it doesn't require any conversion. A lot of research works have successfully utilized RGB color space in various applications in image analysis including bleeding detection. HSV color space, on the other hand, is another widely used color space in image analysis. Its relevancy in recognition problems is attributed to the resemblance of its approximation to the intuitive human color perception. The following sections define the color spaces and the conversion between the two.

3.3.1 RGB

RGB is a color model which represents color as a combination of the three primary colors Red, Green and Blue. This is an additive color model where the primary colors are added to produce the secondary colors of light. The color subspace of interest is a cube as shown in Figure 3- 2 [40]. Red, green and blue are at the three corners of the cube with black at the origin of the cube and white at the farthest corner of the cube from the origin.

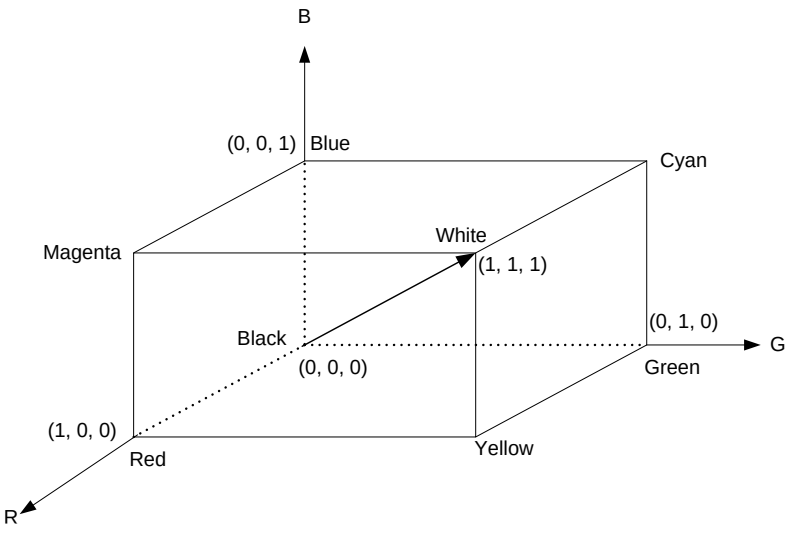

Figure 3- 2 RGB color space cube [40].

3.3.2 HSV

HSV is another widely used color model which is intuitively based on human color perception. It is attributed with separate chromatic and non-chromatic planes. Hue and Saturation are its chromatic planes. Hue refers to the true color. It is the dominant wavelength in the spectral distribution of a color patch. Saturation represents purity of color. It is the amount of white light mixed with hue. Value is the non-chromatic plane which represents illumination level. Figure 3-3 shows a HSV hexacone which is in the subset of HSV space pertaining to the valid RGB values [40]. Hue H ranges from 0° to 360° beginning and ending with the red color with all the intermediate colors in between. Saturation S varies from 0 on the central axis (V-axis) to 1 on the edges of the cone. Value V is the vertical axis which ranges from 0 (apex of the cone) to 1 (base of the cone).

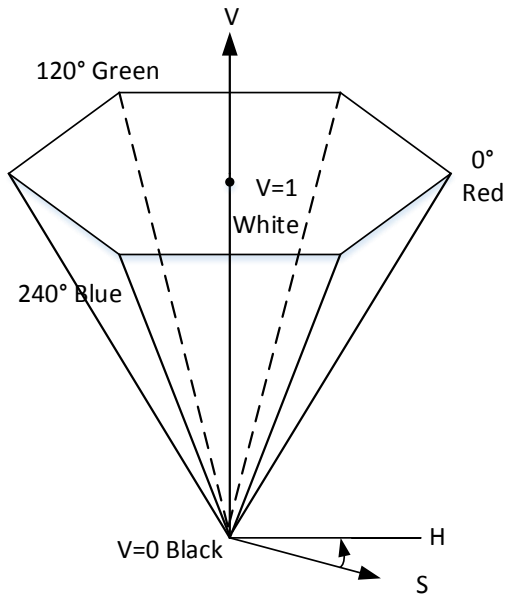

Figure 3- 3 HSV color space hexacone [40].

The RGB to HSV transformation is as follows in equation (3.7) [35]:

$$h = \begin{cases} 0, & if\ \ max = min \\ \left(60\ \times \dfrac{g-b}{max-min} + \ 0^{\circ}\right) mod\ 360^{\circ}, & if\ max = r \\ 60\ \times \dfrac{b-r}{max-min} + \ 120^{\circ}, & if\ max = g \\ 60 \times \dfrac{r-g}{max-min} + \ 240^{\circ}, & if\ max = b \end{cases}$$

$$s = \begin{cases} 0, & if\ max = 0 \\ \dfrac{max-min}{max} = 1 - \dfrac{min}{max}, & otherwise \end{cases}$$

$$v = max$$

(3.7)

Where,

r, g, b are the normalized values [0, 1] in the RGB color space.

max and *min* are the largest and smallest values among r, g, or b in a color.

Table 3- 1 Summary of feature symbols

	Symbols		
Feature Name	**Red Hue**	**Green Saturation**	**Blue Value**
Mean (M)	RM HM	GM SM	BM VM
Standard deviation (S)	RS HS	GS SS	BS VS
Entropy (N)	RN HN	GN SN	BN VN
Skew (K)	RK HK	GK SK	BK VK
Energy (E)	RE HE	GE SE	BE VE

3.4 Feature Extraction: Regions

Images were partitioned into constituent regions of arbitrary shapes and sizes in order to extract features from them. There are different definitions for the term *region* for varying context. The proposed method utilizes color information of the pixels to partition an image into its constituent regions. For this method, region is defined as a pixel set which is specified by using a class membership function defined in a color space. Such region formation is performed in the chosen color space.

In the proposed method, minimum variance color quantization was performed to obtain different pixel groups. Generally, quantization reduces the number of colors in an image by first selecting a set of colors that represent the color gamut of the image and then computing appropriate mapping from the original color space to the representative colors in the reduced space. For example, for an RGB image, the color cube is partitioned into smaller boxes and then all colors that fall into each box are mapped to the color value at the center of the box. Minimum variance quantization takes the distribution of input variable into account while partitioning the color cube. Also, it groups pixels based upon the variance between their color values. Please refer to [41] for the algorithm of the method. The pixels are grouped such that they would have small variance from the representative color of the group. The quantization scheme is thus able to produce regions containing similar colored pixels, however, the number of represented colors after quantization plays a vital role in the type of regions formed.

Even though the grouping of the pixels take place in the color space, it is preferable to extract regions in the image that correspond to recognizable object surfaces. It was empirically found out that, with smaller values like 10 to 15 colors, extracted regions corresponded to object surfaces in the image. But, such small values mean large groups of pixels with large variances. Large variances

in turn can cause dissimilar pixels to be grouped together. This is harmful especially when small bleeding areas are present as they are likely to be included within the surrounding non-bleeding are. Again, if the representative colors after quantization are more than 30, meaningful regions would not be extracted. Also, large number of regions means high computational costs. Thus, 20 to 30 regions were determined to be useful to region partitioning in bleeding detection. Finally 24 number of colors was selected as it provided a reasonable tradeoff between accuracy and complexity.

3.5 Feature Extraction: Region Selection Schemes

Region selection was performed to partition images into regions which are homogeneous in color. Features are crucial in characterizing image regions and in discriminating between the bleeding and non-bleeding classes. Region selection must thus produce image regions which align with the nature of the image properties that needs to be attributed for the problem at hand. Since the aim is to characterize the color distribution of the regions, the region selection scheme based on the color similarity of the pixels in the image was followed. Also, according to the requirement of the system, the images were partitioned in two ways: automated and semi-automated.

3.5.1 Automatic Region Selection

Fully automated region selection is required in the application stage of the methodology where the video frames are fed to the system for bleeding detection. Any form of human intervention in the application stage is highly undesirable and even impractical. An automatic region selection method as shown in Figure 3- 4 was thus followed. This method has two main steps: region partitioning and selection of regions using minimum threshold for region size.

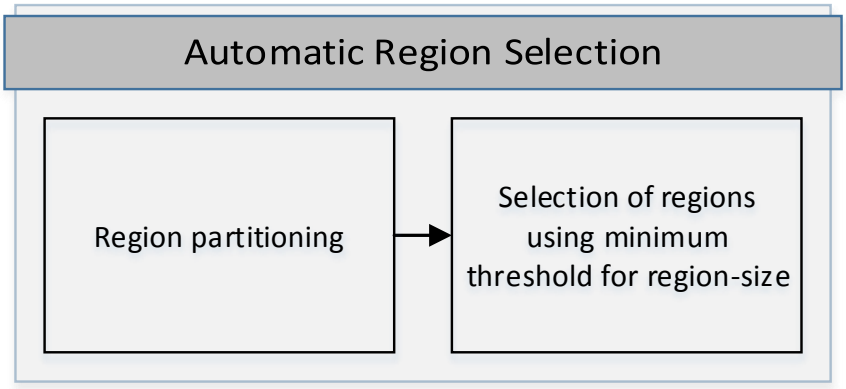

Figure 3- 4 Automatic region selection.

3.5.1.1 Region partitioning

Color segmentation and color quantization are different processes, however; if color quantization is performed in order to significantly reduce the number of colors in the new image, then it can be thought of as a rough segmentation [42]. Following this concept of segmentation, constituent homogeneous regions were selected from an RGB image using the steps:

i) Smooth the image with an averaging filter of size 3x3.

ii) Obtain a new image with a reduced color palette of size *s* by using minimum variance quantization technique on the color palette of the smoothed image of size *O*. The size of the reduced palette *s*=24 was empirically found to be a good estimate of the number of regions in images as it provided a reasonable tradeoff between accuracy and complexity. Figure 3- 5 shows sample images and the resulting images are shown in pseudo-color for better visualization of the different constituent regions of image.

Figure 3- 5 Region partitioning. Left: Original color images; Right: Images with 24 regions obtained after quantizing the original image at the left. Different regions of the images are represented with different colors.

iv) Group all the pixels with the same color value as one region. Thus, there would be *s* number of regions in the resulting image. However, use original pixel values while calculating features.

3.5.1.2 Selection of regions using threshold for minimum region size

Formation of constituent image regions is an important step in the proposed methodology. To further improve the speed of bleeding detection process, regions which are not useful in bleeding discrimination can be omitted in further calculations. The region sizes were analyzed to determine whether or not any difference exists between bleeding and non-bleeding regions in terms of their sizes. If true, the regions with sizes that are outside of the range of bleeding regions can totally be ignored, and thus the required number of calculations can be reduced too.

3.5.2 Proposed Semi Automated Method for Ground Truth Annotation of Bleeding Regions

Training dataset presents the classifier with the features of the different classes involved. For features based on regions, preparation of training data requires the delineation of regions of interest at first. The training data for non-bleeding class was simply created by using the automatic region selection method. The automatic region selection scheme is easy and fast in non-bleeding case as there is no need to select correct regions from the available pool of image regions because none of the image regions consist of bleeding pixels. But in case of bleeding images, the region selection scheme would produce a pool of image regions in which only a selected regions would be the actual bleeding regions. In order to obtain only the bleeding regions, manual intervention is required to select the correct regions. Hence, a new method with minimal human intervention was proposed to annotate bleeding regions in images. This method is applied on successive frames of a video. The scheme makes use of the similarity between the consecutive frames of the capsule endoscopy video.

Figure 3- 6 demonstrates the flowchart of this scheme. The basic idea of the scheme is to obtain single bleeding region from the consecutive bleeding frames of a video by performing seeded region growing separately in all the frames. However, the initial starting point or seed needs to be provided only the first frame. For simplicity, the term ‘current frame’ is used for the frame whose seed has been determined and the seed’s validity has been confirmed. The process gets started from the first seed at the first bleeding frame. The location of first seed, *S1*, was manually selected by the user. It should be within the bleeding area and preferably at the center of a region or far from the edges. After growing a region at any current frame, the centroid of the region, and the average color of the region, *M1*, was determined. As shown in Figure 3- 6, the coordinates of the

region-centroid at the current frame (green points encircled by dashed green circle) was used as the location of seed for the next frame (blue points encircled by solid blue circle). Before starting the region growing process from this position at the next frame, the validity of the latest seed must be evaluated. This was performed by checking whether the color distance between the possible seed pixel at the next frame and the average color, *M1*, of the region grown at the current frame was lower than a fixed threshold. If not, then another pixel in the immediate 5x5 neighborhood of this position was inspected to find a pixel whose color is within the threshold distance of the average color, *M1*, of the current frame. This step affirms the feasibility of the seed, *S2*, belonging to a bleeding area in the next frame and also accounts for slight changes occurring between the successive frames. Also, before moving on to the next frame, distance between the average color of the regions of the current, and previous frame were checked. This step increases the reliability of the scheme by stopping the process whenever average color of grown region starts to significantly differ from previously grown regions.

Region Growing. Figure 3- 7 demonstrates the steps of a seeded region growing technique which starts with an initial point. For simplicity, the latest point that was added to the growing region was called the seed in this algorithm. After the addition of every new pixel, the average color of the region was updated. The *n*-connected neighbors of the seed were added to the neighbor list of the region. Each neighbor of the region was then compared against the mean of the region in order to determine whether it meets the growing condition or not. The region thus iteratively grows in all four directions until all of its neighbors fail the growing criteria.

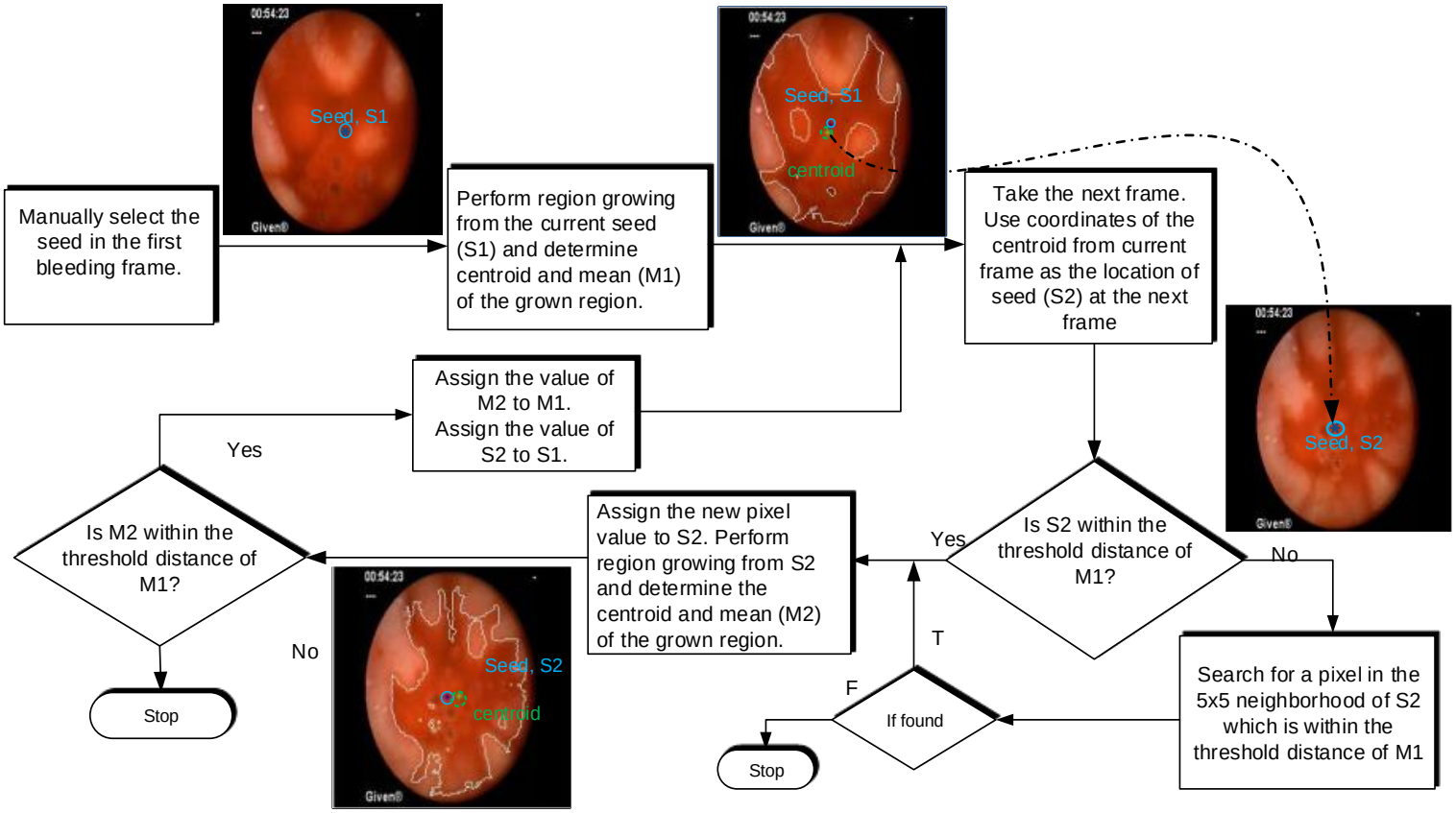

Figure 3- 6 Semi-automatic ground truth annotation: the first two CE frames are shown alongside the steps of the flowchart. Blue solid circle encloses the seed of the current frame and green dashed circle encloses the centroid the current grown-region.

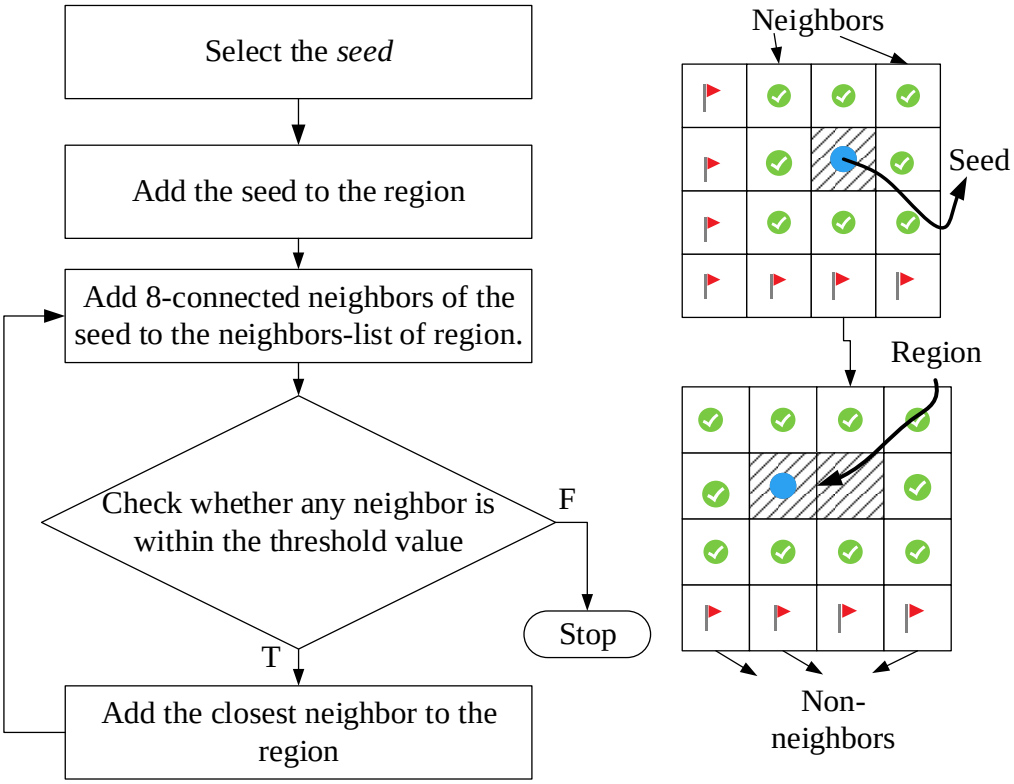

Figure 3- 7 Left: Steps of region growing algorithm, Right: Illustration of two consecutive iterations of region growing process where following symbols are used: blue dot: seed, green tick: neighbor of region, red flags: non-neighbor of region, hashed pixel: member of region.

A distance measure in the coordinate system of a color model gives the closeness of pixels in terms of color. Since differences between colors cannot be accurately perceived from distance measures based on RGB values [43], HSV color space was used for measuring color difference. The distance measure as described in [44] was used. For any two points (H1, S1, I1) and (H2, S2, I2), their color difference is given in as equation (3.8),

$$d = \sqrt{(d_C)^2 + (d_I)^2}$$

Where,

$$d_I = |I_1 - I_2|$$

$$d_C = \sqrt{{S_1}^2 + {S_2}^2 - 2S_1 S_2 \cos\theta}$$

$$\theta = \begin{cases} |H_1 - H_2| & if\ |H_1 - H_2| < 180^\circ \\ 360^\circ - |H_1 - H_2| & if\ |H_1 - H_2| > 180^\circ \end{cases}$$

(3.8)

3.6 Support Vector Machine for Classification

Based on the successful usages of SVM in similar applications, it was adopted in this study too. The classification of bleeding and non-bleeding patterns is a demanding application and the linearly separable case shown at Figure 2- 4 is not sufficient. Such complex problems require the design of non-linear classifiers. However, given *N* points which are non-linearly separable in an *l*-dimensional space, mapping them into higher dimensional space can increase the probability of finding a space where the points are linearly separable. Adopting such an approach, SVM uses appropriate nonlinear functions in order to preprocess data to a higher dimensional space where the data from the two classes can be separated by a hyperplane [45]. Furthermore, architecture of SVM supports the use of kernels that allows the mappings even into an infinite space.

Also, there could be cases where the perfect separation of the points of the two classes might negatively affect the generalizing abilities of the SVM classifier. In such cases, some flexibility in the classification is allowed by introducing a cost parameter, *C* in the SVM models that controls the tradeoff between cost of misclassification and forcing rigid margins. The optimal value of *C* is determined along with other parameters of the SVM model in the training stage. A rigorous search called the grid search was adopted in the proposed method to determine the optimal parameters. In this search, each parameter is varied on a range of values taken in geometric steps and the model

is evaluated at each stage. Thus, if there are two parameters in the model, with each varying in a range consisting of 10 intervals, then there would be 10×2 steps in the search.

3.7 Feature Selection

A major challenge in pattern recognition is the large number of features which is also called the curse of dimensionality. Addition of features obviously increases the computational complexity and it can increase the discriminative power of the classifier too. However, it might also lead to no or little gain in classification because of the correlation between the features and sometimes it might even cause a loss due of addition of noise. Also, more the number of features, more will be the number of classifier parameters. Thus, in presence of limited number of training patterns, it becomes imperative to keep the number of features to a sufficient minimum in order to achieve good generalization ability [21]. In regards to these issues in pattern recognition, feature selection was performed to optimize classification performance. Given a set of features, feature selection chooses subsets of features which would reduce the feature space dimension without compromising discriminative power of the classifier.

There exists two approaches for feature subset selection: scalar feature selection and feature vector selection. In scalar feature selection, the features are ranked individually on the basis of their discrimination capabilities. Then, a subset of features with the best performance is selected to form the feature vector. This selection scheme is very simple but it fails to hold important information regarding correlation between the features. Thus, feature vector selection approach is widely adopted for complex problems and for highly correlated features. Again, there are two ways of performing feature vector selection: filter approach and wrapper approach. The two approaches

differ in terms of the optimality rule which is a criterion used for determining the optimality of feature subsets.

- Filter Approach: The optimality rule in this approach is independent of classifier [46].
- Wrapper Approach: The optimality rule is dependent to classifier. For all the combinations possible, performance of the classifier is evaluated and the one giving the best performance is selected [46].

In this study, wrapper approach was adopted for feature selection. An exhaustive feature selection was followed to find the best feature subset which involves evaluation of classifier performance after each classifier training. Figure 3-8 presents the scheme of this feature selection approach. 15 features, as listed in Table 3- 1, were used in this algorithm by combining them in all possible ways. These features were considered in pairs, triplets and up to groups of fifteen, in all possible combination. The whole experiment was thus divided into 15 sub parts according to the size of the feature subset, *n*. In each case of size *n*, 15 features can be combined in *15Cn* number of ways. For each combination, an SVM was trained and tested with the corresponding subset of features selected from the available features. The SVM was further evaluated on image level with the usage of a set of 100 test images. Since, the classifier classifies only the regions, image level classification was performed in a slightly different manner. In order to classify an image, all of its constituent regions were classified at first. If any one of them was classified as a bleeding region, then the image was categorized as a bleeding image. The classification performance of the method was evaluated on the basis of the results of image level classification on the 100 test images. Every

classifier was trained and tested with the same set of images. The combination of features resulting in the best image level performance of the classifier was finally selected as the best feature subset.

3.8 Performance Measures

The results obtained from the classification algorithm were analyzed for evaluating the performance of the proposed bleeding detection algorithm. There exists four possible cases about the detection result. A bleeding image can be truly detected as a bleeding image and a non-bleeding image can truly be identified as a non-bleeding image. If bleeding is considered to be a positive event and non-bleeding to be a negative event, the above mentioned two cases are called True Positive (TP) and True Negative (TN) detections respectively. Likewise, if a bleeding image is falsely identified as a non-bleeding image, then this is the case of False Negative (FN) detection and if a non-bleeding image is falsely identified as a bleeding image, it is called False Positive (FP) detection. From these values, various measures like sensitivity, specificity and accuracy can be calculated. Sensitivity, specificity and accuracy are three ideal criteria for assessing the classification performance of the proposed algorithm and these measures are widely used in literature regarding evaluation of pattern recognition systems. They are described as in equation (2.2). By considering only the sensitivity value, important information on the correct classification of the negative events, will be missed. Similarly, specificity value alone will also not be able to provide complete assessment of the classification performance. Further, if the number of positive and negative cases differs significantly, accuracy alone can be misleading because it can be biased to either sensitivity or specificity depending on the number of positive and negative cases. For example, if the number of positive cases is significantly less than the number of negative cases, accuracy will be biased to specificity. Hence no single measure among the three is enough for

complete assessment of classification performance, because of which, the proposed method makes use of all of them.

Various circumstances in practical scenarios play a role in determining whether a high sensitivity or a high specificity should be preferred over the other. For example, in bleeding detection, physicians would prefer having a false positive hit rather than actually missing the positive event. However, a lot of false positive hits can be annoying and might reduce. Also, there can be cases where bleeding is present in a batch of adjacent frames where some beginning or end frames could only contain faint traces of bleeding and could be missed. These false negatives could lead to a low value of sensitivity but it could be tolerable in this case.

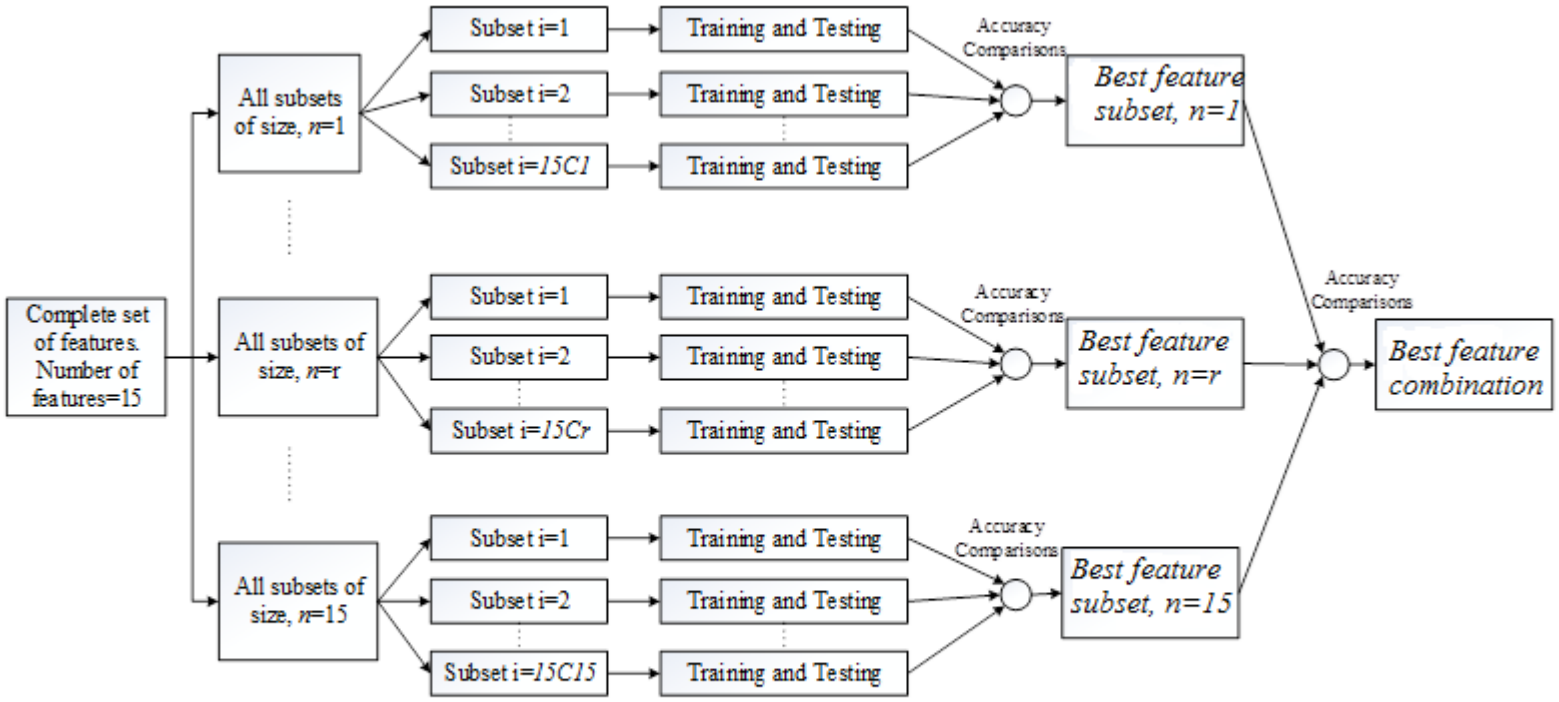

Figure 3-8 Exhaustive feature selection scheme followed in the proposed method for bleeding detection.

Chapter 4

Results and Experiments

This chapter presents all the experiments performed and the results obtained. All the experiments were carried out as computer simulations which were performed in MATLAB [47]. And all the videos and images used in this study were taken from [28] and from the DVD resource available in [48]. The LIBSVM package [49] was used to implement the Support Vector Machine (SVM) models in our methodology.

Two sets of experiments were performed. A set of experiments were performed for selecting the best feature subspaces for bleeding detection in RGB and HSV color space. The best feature subsets were further tested with three different videos. Also, another set of experiments were performed to determine appropriate minimum region size threshold by examining the size histograms of bleeding and non-bleeding regions. The effect of this parameter in the classification performance of the methodology was also observed.

4.1 Feature Extraction

Classification of video frames was performed on the basis of the statistical features as listed in Table 3- 1. Classification ability of each possible combination of features was evaluated. In such an evaluation, the feature vectors used in training and testing the SVM, consists of only the features that constitute the feature combination being evaluated. Figure 4-1 presents scatter plots of some random feature combinations of the train and test data. These plots show that only selective feature combinations produce differentiable cases of bleeding and non-bleeding classes. These outcomes

also show that appropriate feature combinations among the available set of features must be selected in order to classify the image regions correctly.

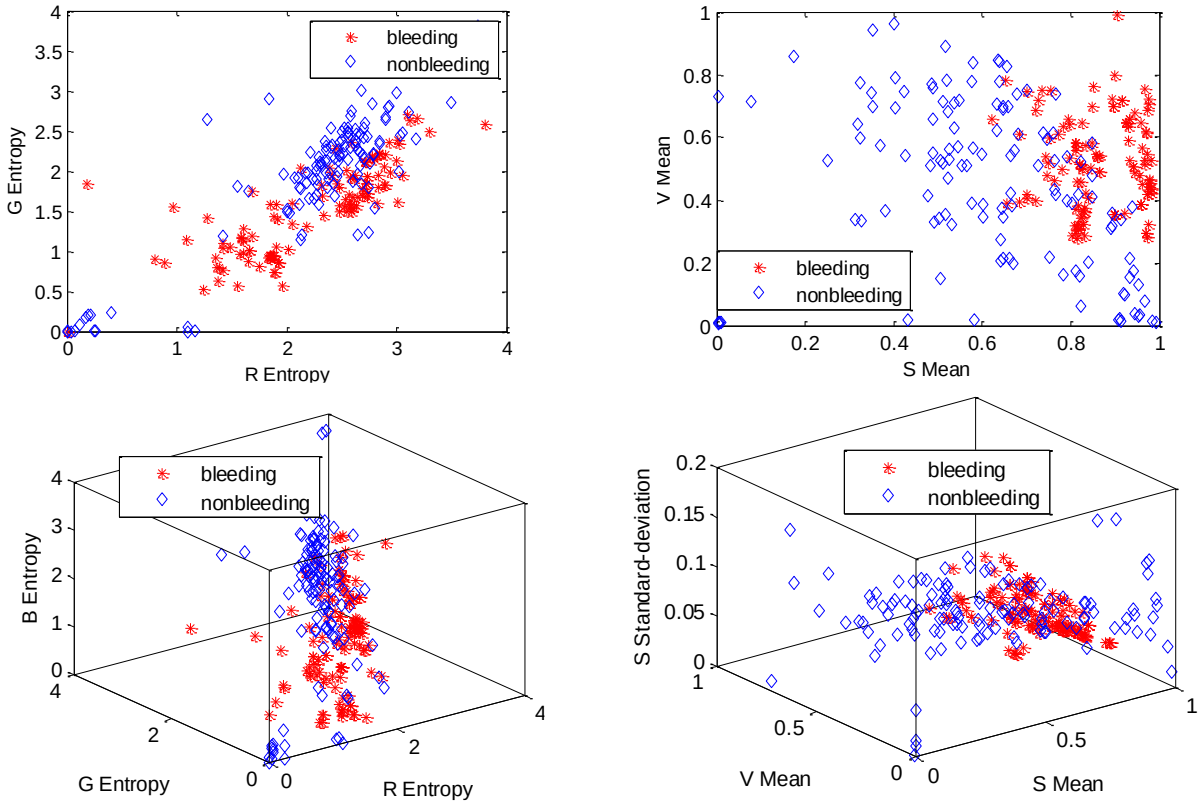

Figure 4-1 Scatter plots of random feature combinations in RGB and HSV color spaces.

4.2 Training and Testing of the SVM

Radial Basis Kernel function (RBF) was used for SVM training. Optimal value for cost, *C*, was searched in the range $\log_2 C = [-5, -4, \dots, 10]$ and gamma, G in the range: $\log_2 G = [-15, -14, \dots, 1]$. Classification performance of a model, trained with each possible combination of *C* and *G*, was evaluated with 5-fold cross validation. The parameter values resulting in the highest cross validation accuracy were selected as the optimal parameters and were then used to train a new model using the whole training set.

All the SVM models were trained and tested using the set of training data consisting of 248 instances. The 248 instances consist of equal number of bleeding and non-bleeding feature instances/vectors. Each of the bleeding instance is derived from a single bleeding region of a frame with bleeding. Thus, 124 bleeding frames were used to create the 124 bleeding instances. The bleeding instances were created from 10 videos taken from [48] by using the ground truth annotation method described in section 3.5.2 . Similarly, the non-bleeding instances were derived from the 124 regions of 25 non-bleeding frames. The non-bleeding frames were also taken from ref. [48].

After training and testing, image level test was performed for each SVM with a set of 100 test images in order to calculate the metrics for performance evaluation. The test image set consists of 50 bleeding images and 50 non-bleeding images. The feature combination producing the best classification performance was the best combination of features. The best feature combinations in RGB and HSV color spaces were further tested with three independent videos. Our set of test images contains fewer images but these are independent images with varying amount and type of bleeding events. Thus, the test set is adequate enough for roughly estimating the performance of the proposed methodology. This dataset can be extended in future if more data is available.

4.3 Feature Selection

For selecting the best feature subsets, bleeding discrimination ability of each possible feature combination was separately evaluated as described in section 3.7. At first, best combinations of features for each subset size n [n=1 to 15] were obtained. Since it is difficult to compare the features based on all three measures, accuracy was used to sort out the best feature combination for each size n at first. Accuracy was chosen because the numbers of positive and negative cases

in our test dataset are equal in this experiment which makes it unbiased toward sensitivity and specificity. Other measures were examined afterwards for ensuring the best performance in discriminating both the classes.

Table 4-1 and Table 4-2 demonstrate the best combination of features for all subset size *n* [*n*= 1, 2, .., 15] in RGB and HSV color spaces respectively. Feature combinations which produce high sensitivity value do not necessarily ensure high classification performance. For example, in Table 4-1, the sensitivity value for the best combination of subset size *s*=1 is 100% but its specificity is 0%. What these values mean is that, among the 50 bleeding images, all of them were classified into their true class but all the non-bleeding images were misclassified. Thus in this case, all the images in the test set were classified as the positive class, no matter what their actual labels were. Hence those combinations, which produced all three performance measures with highest values and with minimal differences among themselves, were selected as the best feature subsets. In cases where there were multiple of such subsets, any one of them is shown in Table 4-1and Table 4-2 for simplicity. After selecting the best feature subset for all sizes, the same scheme as earlier is utilized for determining the best subset among themselves. In case of multiple feature subsets achieving the highest level of performance, the one with the smallest size is selected. For example, 94% is the highest level of classification performance obtained by the RGB features in this experiment; but, among the selected feature subsets, the one with RM and GM is of the smallest size and is thus the finally selected feature combination for the RGB color space. This scheme is also presented by Figure 4-2 and Figure 4-3 which demonstrate the plots of the best performance values for all subset sizes, *n*, in RGB and HSV spaces. Both the color spaces were able to attain high accuracy of 94%.

Table 4-1 Feature selection results for RGB color space

Size	Acc.	Sens.	Spec.	Feature Combination
1	0.5	1.0	0.0	GM
2	0.94	0.94	0.94	RM, GM
3	0.94	0.94	0.94	RM, GM, GN
4	0.94	0.94	0.94	RM, GM, GN, BE
5	0.94	0.94	0.94	RM, GM, BN, RK, GK
6	0.94	0.94	0.94	RM, GM, BN, RK, GK, RE
7	0.94	0.94	0.94	RM, GM, BN, RK, GK, RE, BE
8	0.94	0.94	0.94	RM, GM, BN, RK, GK, RE, BE, GE
9	0.93	0.94	0.92	RM, GM, BM, BS, RN, GN, BN, RK, GK
10	0.93	0.94	0.92	All except RS, GS, BS, RN, RK
11	0.93	0.96	0.9	All except RS, GS, GE, BE
12	0.93	0.96	0.9	All except RS, GS, BE
13	0.93	0.96	0.9	All except RS, GS
14	0.87	1	0.74	All except GS
15	0.79	1	0.58	All

Table 4-2 Feature selection results for HSV color space

Size	Acc.	Sens.	Spec.	Feature Combination
1	0.72	0.92	0.52	HM
2	0.89	0.92	0.86	HM, SM
3	0.94	0.98	0.9	HM, SM, VM
4	0.91	0.94	0.88	HM, SM, SS, SE
5	0.9	0.96	0.84	HM, SM, VM, SS, HE
6	0.89	0.94	0.84	HM, SM, VM, HS, SS, SE
7	0.9	0.96	0.84	HM, SM, VM, HS, SS, VS, SE
8	0.86	0.96	0.76	HM, SM, VM, HS, SS, VS, SE, VE
9	0.84	0.98	0.7	HM, SM, VM, HS, SS, VS, HN, HK, VE

10	0.82	0.94	0.7	All except SN, VN, SK, VK, VE
11	0.75	0.94	0.56	All except SN, VN, SK, VK
12	0.68	0.94	0.42	All except VN, SK, VK
13	0.54	0.98	0.1	All except VK, HE
14	0.54	1.00	0.08	All except VK
15	0.48	0.96	0	All

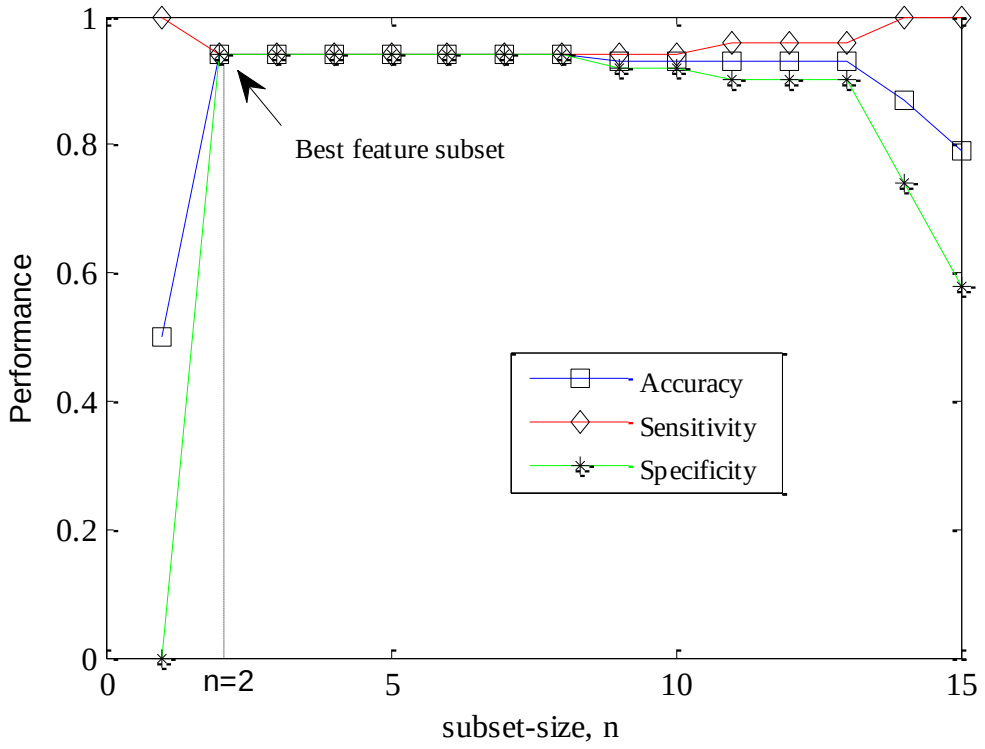

Figure 4-2 Performance comparison of the best feature subsets of all sizes in RGB color space.

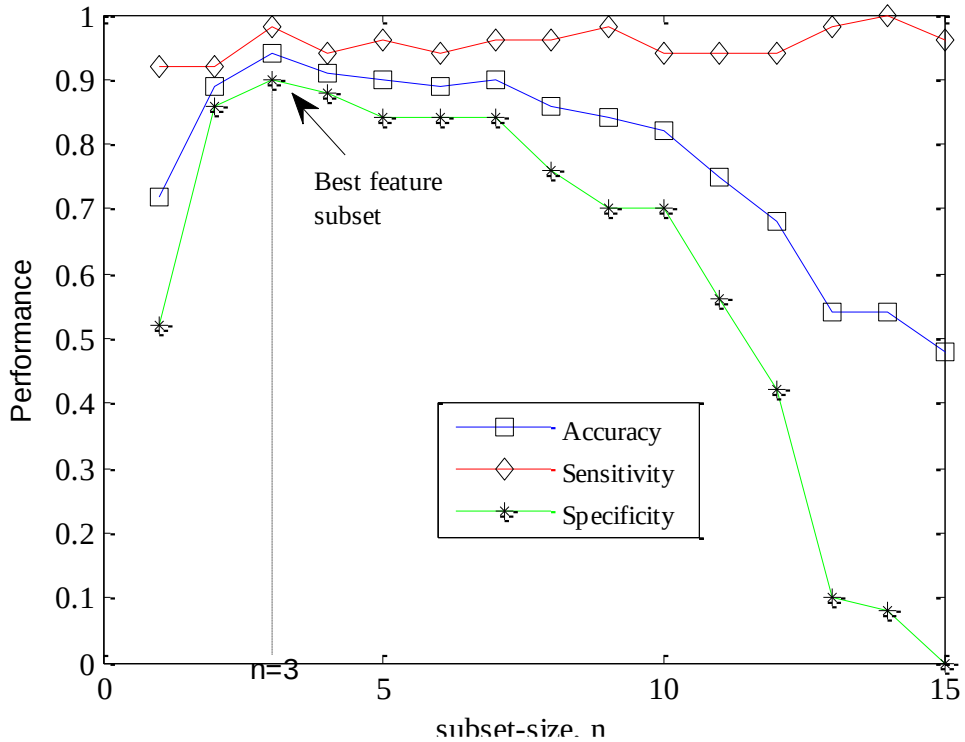

Figure 4-3 Performance comparison of the best feature subsets of all sizes in HSV color space.

4.3.1 Results of Feature Selection in RGB Color Space

Table 4-1 shows that in RGB color space, none of the features are able to discriminate bleeding singlehandedly. It was noted that upon using all the available features, very poor classification performance of bleeding detection was obtained. But, the combination of *RM* and *GM* could achieve highest performance level attained in this set of experiments with accuracy, sensitivity and specificity of 94%. Thus, all the three performance measures of this combination are high which ensures an all-round performance in the discrimination of the bleeding and non-bleeding regions. Moreover, all the selected feature subsets, which are of sizes higher than n=2, contain the combination of *RM* and *GM*. Thus, it can be inferred that the selected subsets for sizes higher than n=2 result from further addition of selective features to the combination of *RM* and *GM*. However,

it was noticed that high level of performance was achievable only up to subset size n=8. Performance started to gradually decline afterwards.

Thus, the mean values of the red and green planes are the two most relevant features for bleeding detection in the RGB color space. It is also interesting to inspect the scatter plot of the features in the reduced feature space of the best feature subset. Figure 4-4 presents the feature instances of the bleeding and non-bleeding classes with *RM* in the x-axis and *GM* in the y-axis. The plot shows the presence of a relation between the mean of the red and green color channels of the image regions which provide discriminative capabilities towards the bleeding discrimination problem.

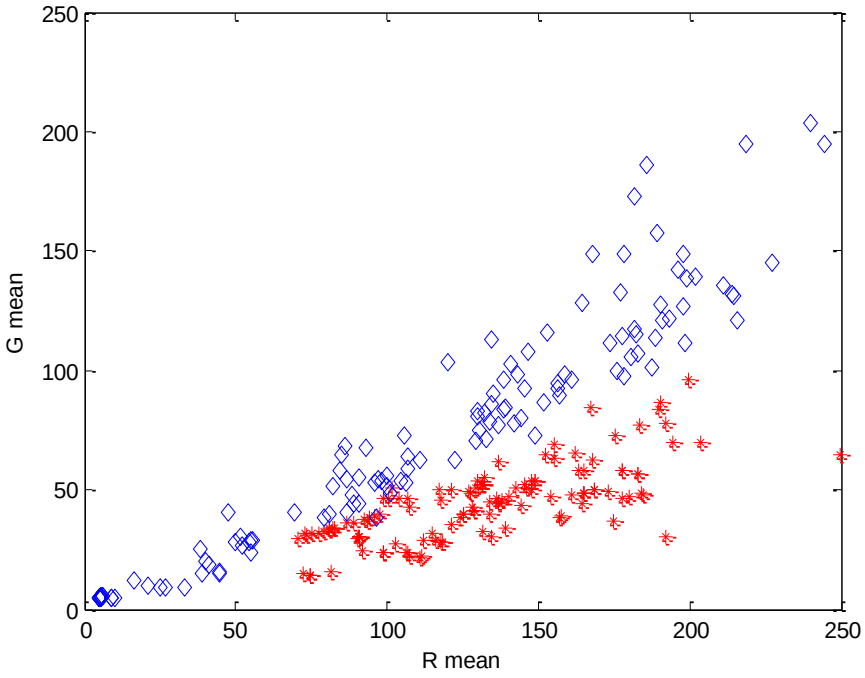

Figure 4-4 Scatter plots of the feature instances in the best feature subset of RGB color space. Bleeding instances are shown as red star and non-bleeding instances are shown as blue diamonds.

4.3.2 Results of Feature-Selection in HSV Color Space

Table 4-2 presents the results of feature selection in HSV color space. It shows that *HM* is the most important feature in the HSV color space and it is singly able to achieve a fair amount of bleeding discrimination. The combination of *HM*, *SM*, and *VM* produced values of accuracy, sensitivity and specificity as high as 94%. Thus, all the three performance measures of this combination are high ensuring an all-round performance in the discrimination of the bleeding and non-bleeding regions. Moreover, almost all other best feature subsets of size higher than n=3, contain the combination of *HM*, *SM*, and *VM*. But, further addition of features into this feature subset resulted in the decline of the performance. Thus, the combination of mean values of the hue, saturation and value planes is the most relevant feature subset for bleeding detection in the HSV color space. Figure 4-5 presents the feature instances of the bleeding and non-bleeding classes with *HM* in the x-axis, S*M* in the y-axis and *VM* in the z-axis. The plot shows a clear separability of the bleeding and non-bleeding instances in the shown feature space.

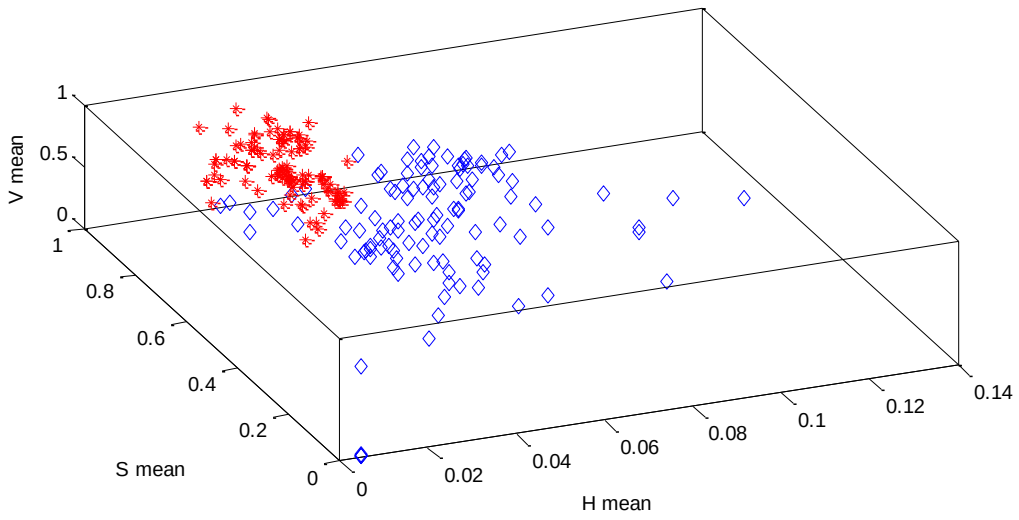

Figure 4-5 Scatter plots of the feature instances in the best feature subset of HSV color space. Bleeding instances are shown as red star and non-bleeding instances are shown as blue diamonds.

4.3.3 Comparison of the Feature Subsets in RGB and HSV Color Space

It is noteworthy to compare the performances of the feature subsets in RGB and HSV spaces. Figure 4-6 and Figure 4-7 show the number of combinations in each subset size, *n*, which produce accuracy values above 50% in RGB and HSV color space respectively. These plots portray the differences in the nature of the feature subsets in RGB and HSV color spaces. It was noted that the number of combinations, with accuracy higher than 50%, is much larger in RGB than HSV for every subset size. The occurrences of numerous combinations with comparable performances are because of many redundant features. For example the feature subset containing *RM, GM, GE* has accuracy of 94 %, which is exactly the same with the feature combination containing *RM, GM, BE*. Here *GE* and *BE* can be seen as two redundant features. There are numerous such redundant features, and thus there are many feature combinations that produce the same or comparable performances. For equivalent level of performance, there are significantly higher numbers of such feature combinations with RGB color space than with HSV. This shows that the RGB color space has significantly larger number of highly correlated features.

It was also noted that the addition of selective higher order features to the best feature subset provides no more useful information in the decision making. This is true for both RGB and HSV spaces; but in HSV, the higher order statistical features tend to affect the decision making negatively. This explains why the sorted accuracy value for each subset size in RGB space decreases gradually from their maximum point as shown in Figure 4-6; whereas, this decrease is very steep in HSV, as shown in Figure 4-7. Some bleeding detection results are presented in Figure 4-8. These are obtained using the best feature subsets of RGB and HSV.

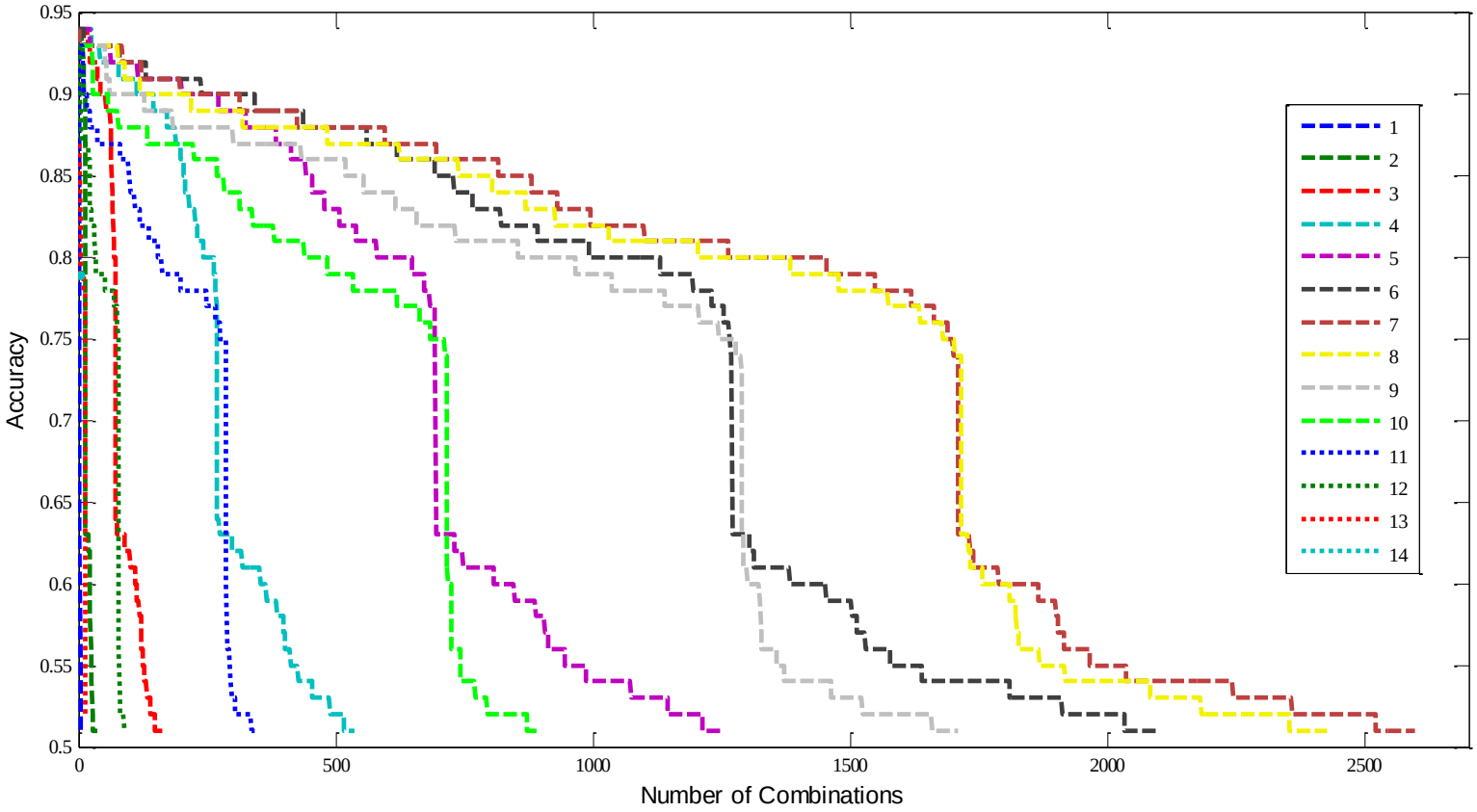

Figure 4-6 Plot of the accuracy values sorted in descending order up to 50% for all feature subset size, *n*, in RGB color space.

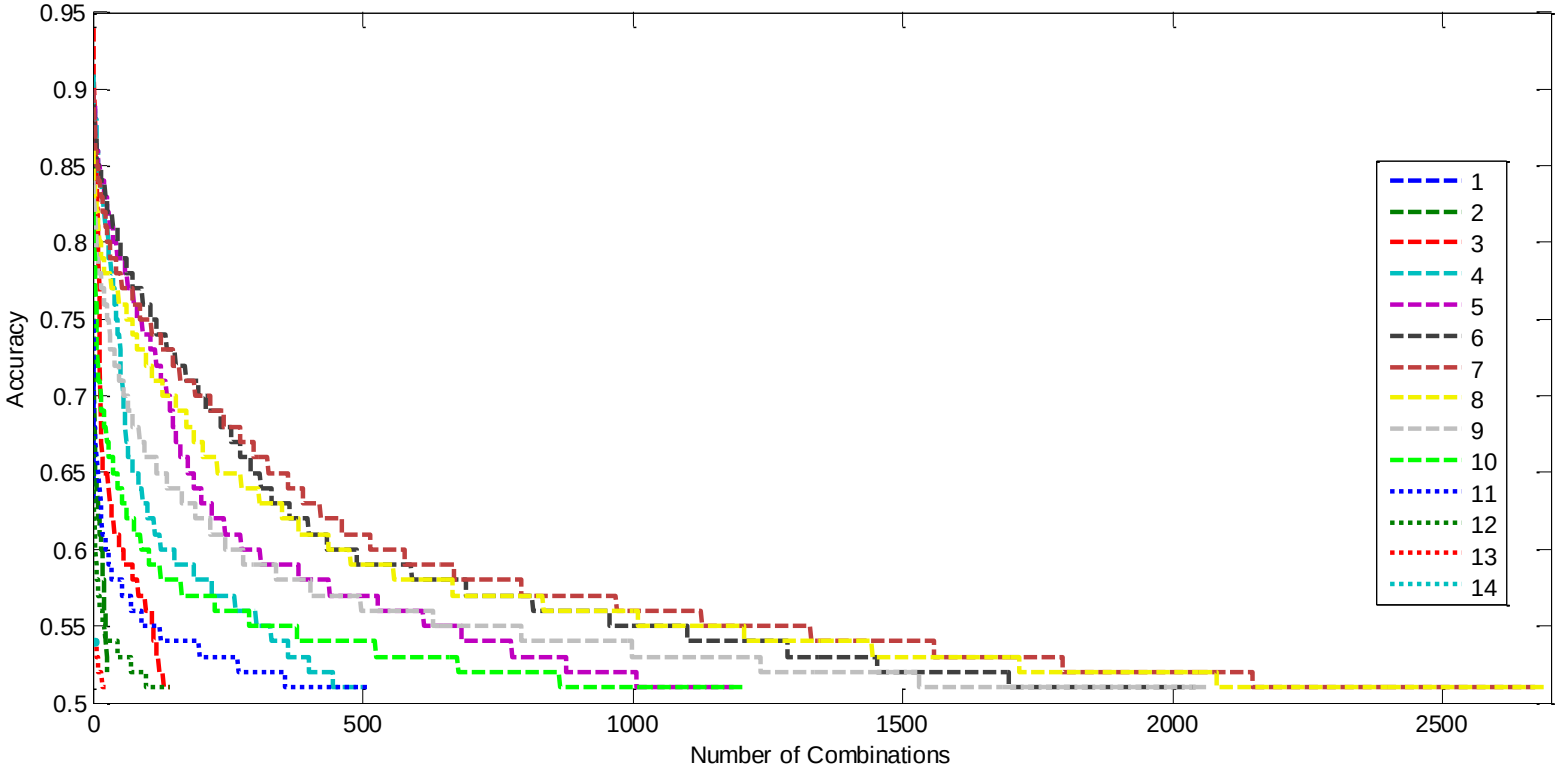

Figure 4-7 Plot of the accuracy values sorted in descending order up to 50% for all feature subset size, *n*, in HSV color space.

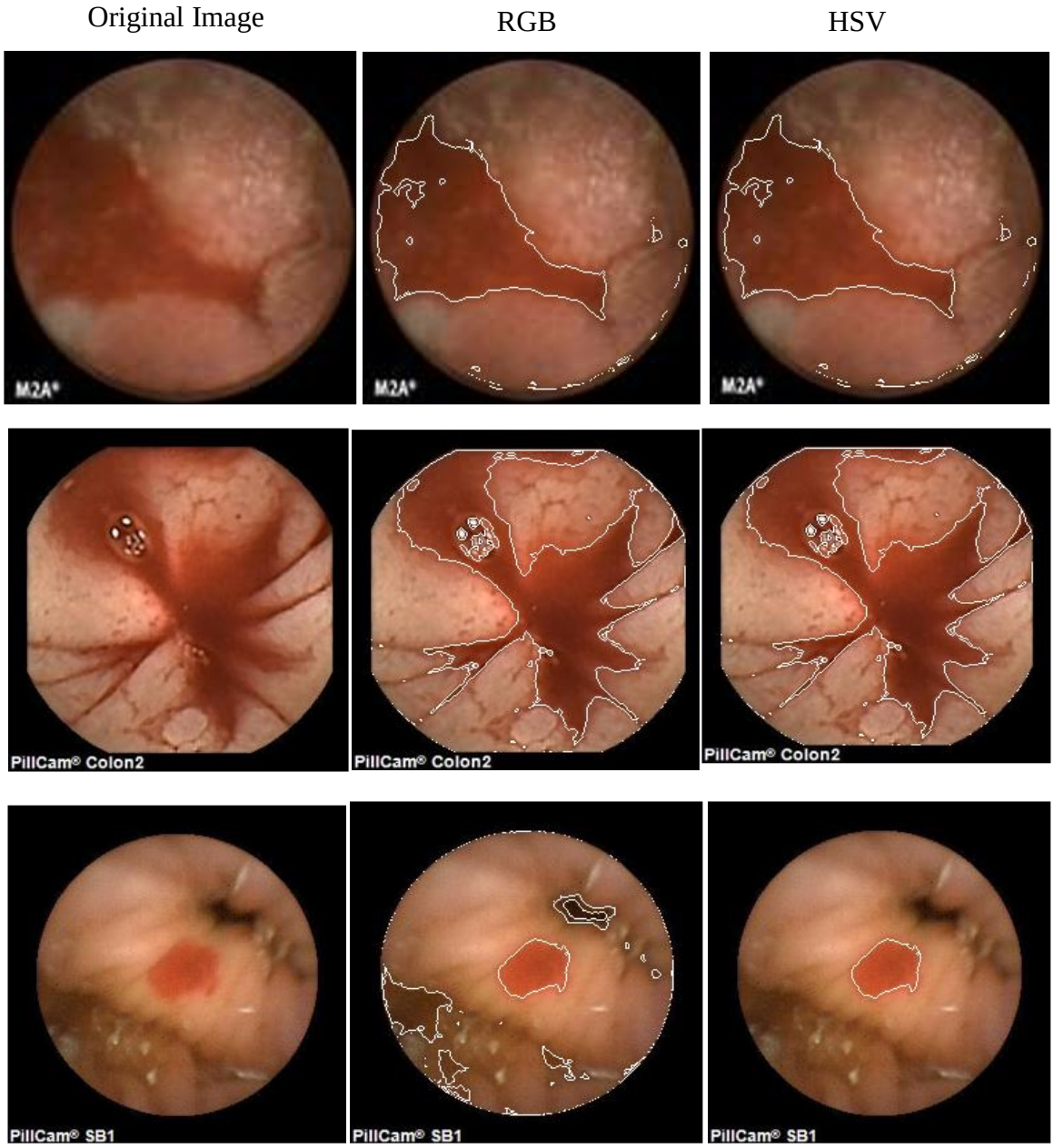

Figure 4-8 Bleeding detection performed by the corresponding best feature subsets in RGB and HSV color spaces. The detected bleeding regions are delineated with white lines.

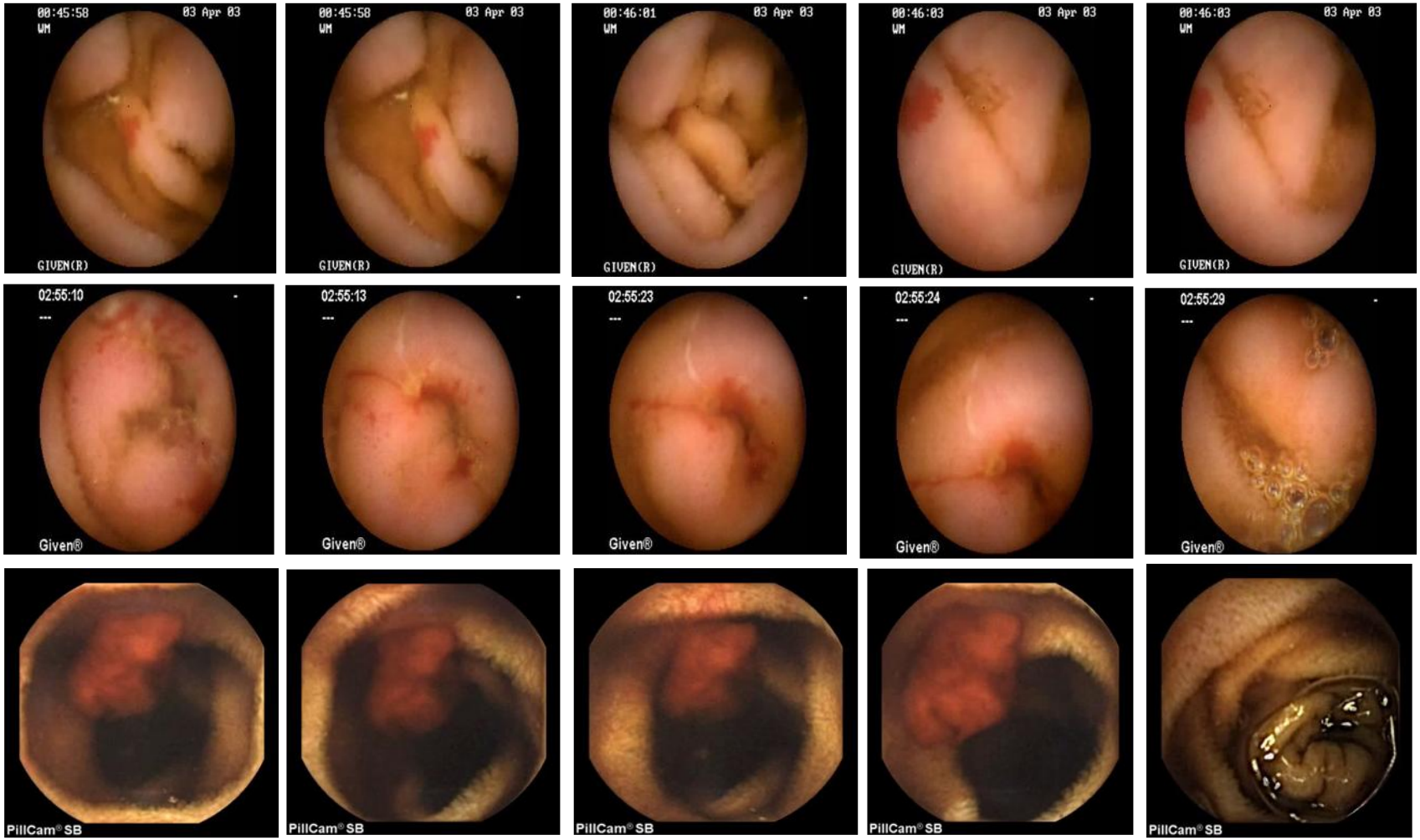

Figure 4-9 Samples of frames of Video 1, Video 2 and Video 3 in row 1, row 2 and row 3 respectively.

4.4 Test of the Feature Selection Results

The best feature subsets were determined on the basis of a test set containing limited number of images. Thus, the relevancy of the best feature subsets were further tested in different video frames in the application stage of the methodology. As described in section 3.1, an application stage of the proposed methodology makes use of the optimal classifier which is trained and tested using the best feature subsets. Three independent CE videos were used to evaluate the performance of bleeding discrimination in each video individually. In order to evaluate the classification performance, all frames of each video were extracted and then classified in the application stage by using the optimal classifier.

Table 4-3 Performance of the best feature subsets in three independent videos in RGB and HSV color spaces.

	RGB			HSV		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Video 1	0.871	0.786	0.878	0.854	0.928	0.848
Video 2	0.918	0.994	0.876	0.992	1.000	0.988
Video 3	0.988	0.800	1.000	0.988	0.800	1.000

Some frames of the three videos are shown in Figure 4-9 to help in the comparative analysis of the performance of the features in the three videos. Video 1 has 589 numbers of frames with 42 bleeding and 547 non-bleeding frames. Video 2 has 500 numbers of frames with 177 bleeding images and 323 non-bleeding images. Video 3 has 428 numbers of frames with 25 bleeding and 403 non-bleeding frames. The classification performance of the best combination of features on bleeding detection in three videos are presented in Table 4-3. The feature subsets provide good classification performance in Video 2 and Video 3 with accuracy as high as 91.8% and 98.8% respectively. Lighting along with the amount of bleeding seem to affect the performance of the proposed detection algorithm. For example, Video 2 consists of frames with multiple bleeding

regions with large and clear bleeding areas due to which the classification performance was remarkably good in this video. Also, improper lighting conditions seem to affect the sensitivity in case of Video 3 in extreme conditions, especially when bleeding regions fall within the dark areas. Overall, good discrimination of bleeding and non-bleeding regions in this experiment has further corroborated the relevancy of the best feature subsets in the proposed bleeding discrimination methodology.

4.5 Tuning of Region Size Threshold

As described in section 3.5.1, sizes of the image regions can be used to further improve the system performance in regards to the number of required calculations. The aim here is to find the range of region sizes that occur predominantly in the bleeding category only. In order to determine the valid range of region sizes for the bleeding class, at first the frequency of occurrences of various region sizes in the two classes were examined. Then, suitable region size threshold was determined by analyzing the effects of region size thresholds to the classification performances.

At first, histograms of region sizes were created for bleeding and non-bleeding regions of the test images which are shown in Figure 4-10 and Figure 4-11 respectively. The histograms show that bleeding and non-bleeding occur in a large overlapping range of region sizes. But, it can be inferred that the bleeding regions rarely occur with sizes less than 200 pixels whereas non-bleeding regions are highly probable to occur with sizes lesser than 100 pixels.

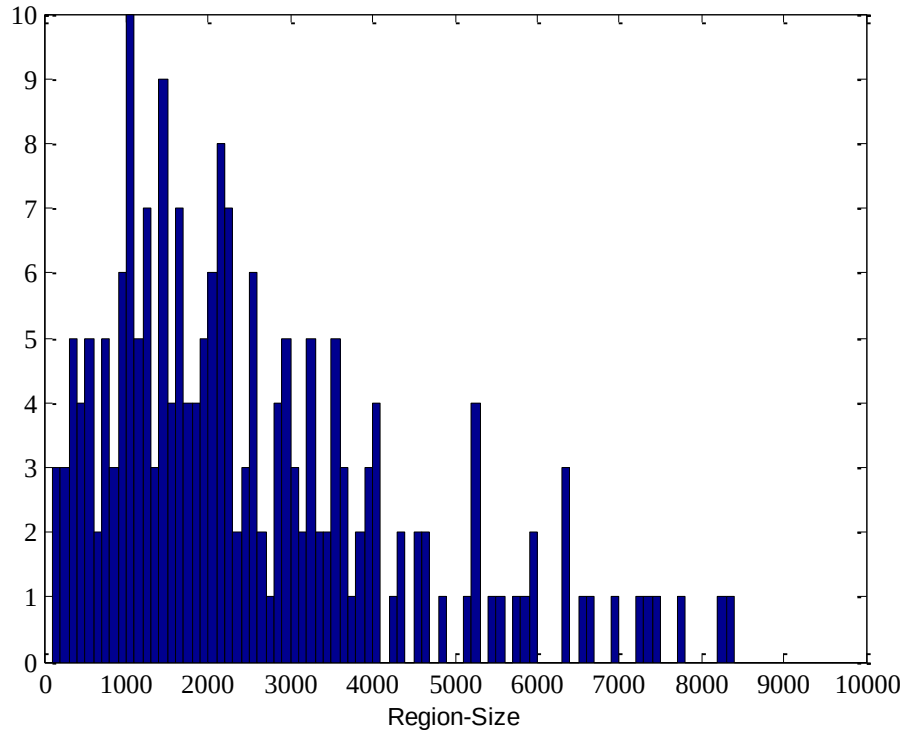

Figure 4-10 Histogram of region sizes for bleeding regions of test images.

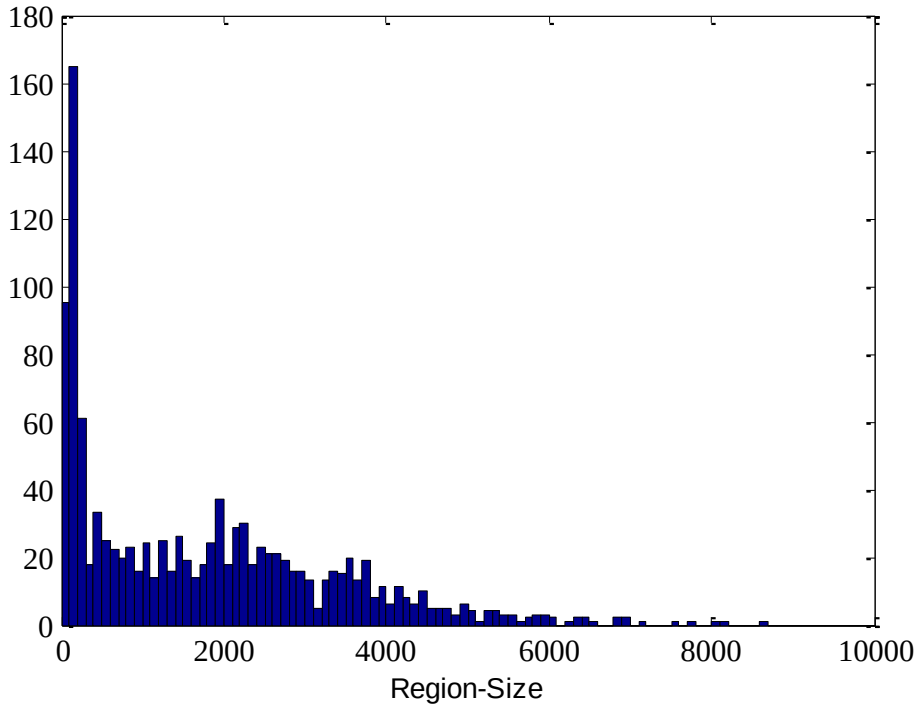

Figure 4-11 Histogram of region sizes for non-bleeding regions of test images.

Table 4-4 Comparison of performance and time in RGB color space by varying the region size threshold

Region size threshold	Video 1			Video 2			Video 3		
	Sens.	Spec.	Time (sec)	Sens.	Spec.	Time (sec)	Sens.	Spec.	Time (sec)
0	0.786	0.878	712	0.994	0.876	605	0.800	1.000	521
100	0.786	0.878	710	0.994	0.876	601	0.800	1.000	518
200	0.500	0.878	706	1.000	0.873	595	0.800	1.000	512
300	0.405	0.878	703	0.938	0.873	593	0.800	1.000	510
400	0.262	0.878	693	0.938	0.873	588	0.800	1.000	508
500	0.119	0.879	691	0.932	0.876	586	0.800	1.000	507
1000	0.119	0.925	671	0.864	0.876	574	0.800	1.000	493
2000	0.000	0.982	595	0.689	0.938	506	0.600	1.000	455
4000	0.000	1.000	419	0.175	1.000	320	0.000	1.000	345

Table 4-5 Comparison of performance and time in HSV color space by varying the region-size threshold

Region size threshold	Video 1			Video 2			Video 3		
	Sens.	Spec.	Time (sec)	Sens.	Spec.	Time (sec)	Sens.	Spec.	Time (sec)
0	0.929	0.848	354	1.000	0.988	302	0.800	0.988	223
100	0.929	0.848	350	1.000	0.988	298	0.800	0.988	222
200	0.643	0.848	349	1.000	0.988	297	0.800	0.988	220
300	0.548	0.848	343	1.000	0.988	294	0.800	0.988	220
400	0.405	0.848	340	1.000	0.988	290	0.800	0.988	219
500	0.405	0.848	340	1.000	0.988	290	0.800	0.988	218
1000	0.262	0.890	332	1.000	0.988	286	0.800	0.988	214
2000	0.000	0.969	305	1.000	0.988	262	0.800	0.988	199
4000	0.000	1.000	230	0.057	1.000	185	0.000	1.000	158

Another experiment was performed to study the effect of ignoring small regions to the performance of the methodology by using the same three different videos as in section 4.4. As earlier, frames of individual videos were classified by using the most optimized classifier in RGB and HSV spaces. Here region partitioning was followed by region selection using various region size thresholds. Separate simulations were performed for the region size thresholds of 0, 100, 200, 300, 400, 500, 1000, 2000 and 4000. Table 4-4 and Table 4-5 demonstrate that ignoring the small regions could affect the sensitivity of this methodology, especially for videos with smaller bleeding regions. Video 1, has multiple frames consisting of a single, small bleeding region as shown in

Figure 4-9. Thus, performance for Video 1 was negatively affected by ignoring smaller regions and the effect increased with the increase in the region size threshold. But, Video 2 and Video 3 have larger bleeding regions because of which the performances for these videos were not affected until very large regions were omitted. Also, it was noted that none of the performances were affected till the region size threshold of 100 pixels.

Table 4-4 and Table 4-5 also show the total time required by the bleeding detection process for all three videos. In order to observe the reduction in time caused by the reduction in the number of regions, time for completing the bleeding detection process in each video can be compared against each other for different minimum region size thresholds. Thus, it was observed that by ignoring smaller regions, reduction in time can be achieved without compromising performance. Thus, from the size histograms presented in Figure 4-10, Figure 4-11and from Table 4-4 and Table 4-5, it can be concluded that the regions of size lesser than 100 pixels can safely be omitted in the bleeding detection process, in order to reduce the number of calculations.

4.6 Performance Comparison to Other Existing Methods

In order to demonstrate the classification performance of the proposed method, it was compared with existing state of the art methods in bleeding detection. In Table 4-6, state of the art methods are enlisted in order of the time of their publication. All types of classification methods according to identification area are covered in Table 4-6. Similar values of performance metrics are seen to be reported by the enlisted methods. Further, there is no clear indication of the superiority of any method from their reported values of performance metrics. Thus, various aspects of the methods must be considered in addition to the performance metrics while comparing the methods. Readers

can also refer to [50] for additional reviews of computer aided decision support systems for GI tract including detection of bleeding pathology.

In Table 4-6, references [12], [23] and [51] are pixel based classification methods. Even with the use of simple features, this type of method is computationally exhaustive. Likewise, the authors in [31] use covariance of second order statistical features (CWC) based on discrete wavelet transform and then perform classification between the abnormal and normal classes with the use of texton boost classifier. This method achieves a reasonably good classification performance but the results of the method are determined from a database of only 100 WCE images. Even though, it is an image based method, the used features are computationally intensive. The authors in this study have themselves claimed the high computational cost requirement of the method. In contrast, the proposed method requires first order statistical features which demand significantly lower computational requirements than the CWC features

In [17], [19]and [26], identification areas which are intermediate to the pixel and image based areas were used. This type of method derives features from fixed shape and size image patches, which are highly likely to introduce noise, especially for smaller bleeding regions. These methods have reported similar values of the performance metrics except [19] which reports a perfect classification performance values. However, these are patch level classification results which are not as conclusive as image level classification results. Authors in [19] also provide an image level classification result of 3.5% false positive rate. But this result is determined on the basis of only one patient's video. Video of a single patient covers limited types of bleeding occurrences and results based on a single video do not provide conclusive result for image level classification. However, the image level classification results of the proposed method are based on 100 independent images from standard public database sources. The results of the proposed method

are further supported by additional experiments performed with three independent video sequences consisting of 1512 frames. Different types of videos were used in the experiment to make it fair and Table 4-3 shows that the method is able to provide 92.9% sensitivity even with small bleeding areas in video frames.

Similar to our method, [52] is a region-based classification method. Though the method is able to achieve good classification performance, it involves large number of parameters in various processes like edge detection, morphological operations and superpixel segmentation. Although the authors provide optimal values for these parameters, some human intervention might be required to optimize these parameters in varying scenarios because of their significant effect to the performance of the method. However, the proposed method requires the optimization of only one parameter which has significant effect to the system. Thus, the proposed method requires minimal human intervention, making it suitable for automatic applications.

In Table 4-6, an analytical comparison of the enlisted methods is also provided in terms of their computational costs requirements to perform image level classification. Two sub steps are focused in the comparison: feature extraction and segmentation/ region partitioning. Following aspects were taken into account for labelling the feature extraction into high, medium and low categories: feature complexity, feature vector size per identification area and number of identification area per video frame. For example, both [52] and the proposed method utilize simple features but, [52] divides an image into a total of 400 regions in 576x576 images which is much higher than the proposed method which produces less than 24 regions in 256x256 images. Thus, feature extraction in the proposed method has significantly lower computational costs than in [52]. Likewise, computational costs for region partitioning schemes were also labelled into the same three categories: high, medium and low. The region partitioning method used in the proposed method is

based on color quantization which is computationally less rigorous than other segmentation methods.

Thus, the method proposed in this work has low computational requirements in both feature extraction and region partitioning. It has exhaustive feature selection but that is required only once and doesn't affect in the application phase. Moreover, the method has an additional level of automation than all other methods as it employs a semi-automatic ground truth label annotation method.

Table 4-6 Comparison of existing state-of-the-art methods in bleeding detection.

Ref.	Identification-area	# images and videos	Source of data	Features	#components / identif. area	Feature Selection	Ground truth label	Classification	Computational Cost **		Sens.	Spec.
									Feature	Region-partn.		
[23]	pixel	15222 frames from 3 videos	Single Hospital	RGB pixel values	3/ pixel	None	Manual	EM clustering	H	None	92.0	98.0
[51]	pixel	2,000 frames from 2 videos	N/A.	Pixel values	N/A	None	N/A	Thresholding	H	None	92.86	89.49
[17]	patch	400 frames from 20 videos	Single Hospital	various	N/A	None	Manual	ANN	M	L	*91.6	*93.6
[26]	patch	N/A	Single Hospital	Adaptive color histogram	110/ patch	None	N/A	ANN	H	L	91.8	93.5
[12]	pixel	14630 frames from 150 videos	Single Hospital	RGB and HSV values	6 /pixel	None	Manual	ANN	H	None	93.1	85.8
[31]	image	100 images	Jin shan co.	CWC	72 / image	None	N/A	Texton Boost	H	None	89.1	82.3
[19]	patch	538 frames plus 1 video	N/A	various	2494 / patch	Anova, SFFS, VSCH	Manual	SVM, VSCH	H	L	*86.0 -100.0	*98.6-100.0
[52]	region	5000 frames from 20 videos	Single Hospital	R/B, R/G, R/(R+G+B)	3 /region	None	Manual	SVM	M	M	99.0	94.0
ours	region	348 frames plus 3 videos	Standard public database	histogram features	15/ region	Exhaustive	semi-automated	SVM	L	L	94.0	94.0

* marked sensitivity and specificity values are patch level classification results. All other values are image level classification performances.

** Computational costs are labelled as high, medium and low categories, which are represented by the symbols: H, M, and L respectively.

Chapter 5
Conclusion and Future Work

5.1 Summary of Accomplishments

Bleeding detection in CE videos is a problem of interest, and thus has attracted a critical mass of research. A number of works have been performed in this field by devising supervised classification systems which apply intelligent techniques to classify the frames into different classes. This thesis argues that the region-based classification approach is suitable to this application as it provides information from the local regions of the video frames. The main objective of the thesis is to use relevant feature extraction, feature selection and classification techniques in order to efficiently perform bleeding detection in CE.

The proposed bleeding detection method was presented in the thesis for detecting the frames of the video which have possible bleeding occurrences in order to assist physicians in the examination of gastrointestinal tract. The proposed method performs a supervised classification of the video frames on the basis of statistical features that were derived from the regions of frames. The proposed method uses region selection technique based on color quantization in order to divide the image into regions of homogeneous color. This region selection method is fast and requires minimal parameter optimization. Statistical measures were then derived from the first order histogram probability of the color channels of the regions. RGB and HSV color channels were utilized in the proposed work.

The experiments and results were then presented. The use of all available 15 features resulted in poor classification in both RGB and HSV color spaces. Thus, an exhaustive search was performed for determining the best combination of features. The combination which achieved

maximum sensitivity and specificity of 94% were selected as the best feature subsets. Additional experiments to test the relevancy of these selected feature subsets also produced consistently good performance with three independent videos. Thus through feature selection, high classification performance was achieved. Furthermore, the best feature subsets in RGB and HSV were only of size two and three respectively. With the use of these subsets, the method requires less number of features per image region thus, reducing the classification time per video frame.

Also, the regions of the bleeding and non-bleeding classes were experimented in this thesis. The results determined that the regions occurring in small sizes, with 100 or less number of pixels, can totally be ignored as they are not likely to be of any importance. Thus by ignoring small regions, the number of regions per frame was reduced, which in turn reduced the number of classifications required to class per frame.

A semi-automatic method was also proposed to annotate the ground truth labels of bleeding class. Manual annotation of the ground truth labels for bleeding requires manual delineation of the whole bleeding area. Since large training set is always desirable, manual annotation could be tedious. The proposed annotation method addresses this problem by requiring the user to input only a single point in the bleeding area. After receiving the location of the starting point and a threshold value, the method then grows a region containing similar neighboring pixels within the given threshold.

Overall, the proposed method combines multiple synergistic approaches which makes it suitable for the detecting CE frames containing bleeding. The proposed method achieves high classification performance with the region-color features derived from first order probability histograms. Thus, the region partition method and the extracted color features are relevant to bleeding detection. Further, Table 4-6 illustrates various features of the proposed method like low

computational cost, high classification ability, added level of automation, and usage of standard data sources that has proved its novelty and relevance in classifying bleeding and non-bleeding regions in CE images. Minimal human intervention and low computational requirements of the method also make it feasible for practical implementation. Thus, the method shows great promise in assisting physicians in bleeding detection in CE videos.

5.2 Recommendations for Future Works

This section provides a few recommendations which can be explored in future for improving in the classification performance of the method:

1. Preprocessing techniques were not explored in this work. Ref. [25] has used detection of specular highlights in air bubbles which improved the classification performance of their method. Such techniques which perform image enhancement, removal of very dark areas and highlights can be integrated into the proposed method to improve the performance of the system.
2. The methodology classifies the video frames into two classes: bleeding and non-bleeding. A classification system with an extra class can be designed which would take ambiguous and difficult cases into account.
3. Other image segmentation techniques can be tested and compared with the automatic region selection technique used in the methodology.
4. The bleeding detection process can be sped by addressing the correlation that is present among the adjacent frames of the video.

.

LIST OF REFERENCES

[1] A. Wang, S. Banerjee, B. A. Barth, Y. M. Bhat, S. Chauhan, K. T. Gottlieb, V. Konda, J. T. Maple, F. Murad, P. R. Pfau, D. K. Pleskow, U. D. Siddiqui, J. L. Tokar, and S. A. Rodriguez, "Wireless capsule endoscopy," *Gastrointest. Endosc.*, vol. 78, no. 6, pp. 805–15, Dec. 2013.

[2] "Given Imaging Receives FDA Clearance for PillCam® COLON in Patients Following Incomplete Colonoscopy." [Online]. Available: http://www.givenimaging.com/en-us/Innovative-Solutions/Capsule-Endoscopy/pillcam-colon/Pages/COLON-Press-release.aspx. [Accessed: 15-April-2014].

[3] P. Valdastri, M. Simi, and R. J. Webster III, "Advanced technologies for gastrointestinal endoscopy," *Annu. Rev. Biomed. Eng.*, vol. 14, pp. 397–429, 2012.

[4] H. Hussain, S. Lapin, and M. S. Cappell, "Clinical scoring systems for determining the prognosis of gastrointestinal bleeding," *Gastroenterol. Clin. North Am.*, vol. 29, no. 2, pp. 445–464, 2000.

[5] L. B. Katz, "The role of surgery in occult gastrointestinal bleeding," in *Seminars in gastrointestinal disease*, 1999, vol. 10, no. 2, pp. 78–81.

[6] S. L. Triester, J. a Leighton, G. I. Leontiadis, D. E. Fleischer, A. K. Hara, R. I. Heigh, A. D. Shiff, and V. K. Sharma, "A meta-analysis of the yield of capsule endoscopy compared to other diagnostic modalities in patients with obscure gastrointestinal bleeding," *Am. J. Gastroenterol.*, vol. 100, no. 11, pp. 2407–18, Nov. 2005.

[7] B. S. Lewis and P. Swain, "Capsule endoscopy in the evaluation of patients with suspected small intestinal bleeding: Results of a pilot study," *Gastrointest. Endosc.*, vol. 56, no. 3, pp. 349–353, Sep. 2002.

[8] W. A. Qureshi, "Current and future applications of the capsule camera," *Nat. Rev. drug Discov.*, vol. 3, no. 5, pp. 447–450, 2004.

[9] J. M. Buscaglia, S. A. Giday, S. V Kantsevoy, J. O. Clarke, P. Magno, E. Yong, and G. E. Mullin, "Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study," *Clin. Gastroenterol. Hepatol.*, vol. 6, no. 3, pp. 298–301, Mar. 2008.

[10] S. Liangpunsakul, L. Mays, and D. K. Rex, "Performance of Given suspected blood indicator," *Am. J. Gastroenterol.*, vol. 98, no. 12, pp. 2676–8, Dec. 2003.

[11] C. Signorelli, F. Villa, E. Rondonotti, C. Abbiati, G. Beccari, and R. de Franchis, "Sensitivity and specificity of the suspected blood identification system in video capsule enteroscopy," *Endoscopy*, vol. 37, no. 12, pp. 1170–1173, 2005.

[12] G. Pan, G. Yan, X. Qiu, and J. Cui, “Bleeding detection in Wireless Capsule Endoscopy based on Probabilistic Neural Network,” *J. Med. Syst.*, vol. 35, no. 6, pp. 1477–84, Dec. 2011.

[13] Y. Fu, M. Mandal, and G. Guo, “Bleeding region detection in WCE images based on color features and neural network,” in *Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on*, 2011, pp. 1–4.

[14] C. S. Lima, D. Barbosa, J. Ramos, A. Tavares, L. Monteiro, and L. Carvalho, “Classification of endoscopic capsule images by using color wavelet features, higher order statistics and radial basis functions,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2008, pp. 1242–5, Jan. 2008.

[15] P. C. Khun, “Feature selection and classification for Wireless Capsule Endoscopic frames,” *2009 Int. Conf. Biomed. Pharm. Eng.*, pp. 1–6, Dec. 2009.

[16] G. Lv, G. Yan, and Z. Wang, “Bleeding detection in wireless capsule endoscopy images based on color invariants and spatial pyramids using support vector machines,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2011, pp. 6643–6, Jan. 2011.

[17] B. Li and M. Q.-H. Meng, “Computer-aided detection of bleeding regions for capsule endoscopy images,” *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1032–9, Apr. 2009.

[18] Y. Lee and G. Yoon, “Real-Time Image Analysis of Capsule Endoscopy for Bleeding Discrimination in Embedded System Platform,” *World Acad. Sci. Eng. Technol.*, pp. 2526–2530, 2011.

[19] P. Szczypiński, A. Klepaczko, M. Pazurek, and P. Daniel, “Texture and color based image segmentation and pathology detection in capsule endoscopy videos,” *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 396–411, 2014.

[20] K. Sun, Y. Wu, X. Lin, S. Cheng, Y.-M. Zhu, and S. Zhang, “Mean Shift-based Lesion Detection of Gastroscopic Images,” in *Proceedings of the Second Sino-foreign-interchange Conference on Intelligent Science and Intelligent Data Engineering*, 2012, pp. 167–174.

[21] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Fourth. Academic Press, 2009, p. 7.

[22] Emily Watt MLIS et al., “Toward medical decision making,” in *Medical Imaging Informatics*, R. K. T. Alex A.T. Bui, Ed. Springer US, 2010, pp. 403–438.

[23] S. Hwang, J. Oh, J. Cox, S. J. Tang, and H. F. Tibbals, “Blood detection in wireless capsule endoscopy using expectation maximization clustering,” in *Proc. of SPIE*, 2006, p. 61441P–61441P–11.

[24] J. Ma, T. Tillo, B. Zhang, Z. Wang, and E. G. Lim, “Novel training and comparison method for blood detection in wireless capsule endoscopy images,” in *Medical Information and Communication Technology (ISMICT), 2013 7th International Symposium on*, 2013, pp. 56–60.

[25] M. W. Mackiewicz, M. Fisher, and C. Jamieson, “Bleeding detection in wireless capsule endoscopy using adaptive colour histogram model and support vector classification,” *Proc. SPIE*, vol. 6914, p. 69140R–69140R–12, 2008.

[26] C. K. Poh, L. Li, and J. Liu, “Multi-level local feature classification for bleeding detection in Wireless Capsule Endoscopy images,” *2010 IEEE Conf. Cybern. Intell. Syst.*, pp. 76–81, Jun. 2010.

[27] Y. Lee and G. Yoon, “Improvement of Blood Detection Accuracy using Image Processing Techniques suitable for Capsule Endoscopy,” pp. 1146–1149, 2012.

[28] “capsuleendoscopy.org.” [Online]. Available: http://www.capsuleendoscopy.org/Pages/default.aspx. [Accessed: 01-Aug-2013].

[29] M. K. Bashar, K. Mori, Y. Suenaga, T. Kitasaka, and Y. Mekada, “Detecting informative frames from wireless capsule endoscopic video using color and texture features,” *Med. Image Comput. Comput. Assist. Interv.*, vol. 11, no. Pt 2, pp. 603–10, Jan. 2008.

[30] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans. Syst. Man. Cybern.*, vol. 3, no. 6, pp. 610–621, Nov. 1973.

[31] X. Liu, J. Gu, Y. Xie, J. Xiong, and W. Qin, “A new approach to detecting ulcer and bleeding in Wireless capsule endoscopy images,” in *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, 2012, pp. 737–740.

[32] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach: A Matlab Approach*. Academic Press, 2010, p. 29.

[33] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006, pp. 227–229.

[34] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 2009, pp. 119–142.

[35] A. Karargyris and N. Bourbakis, “Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2777–2786, 2011.

[36] S. a Karkanis, D. K. Iakovidis, D. E. Maroulis, D. a Karras, and M. Tzivras, “Computer-aided tumor detection in endoscopic video using color wavelet features,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 3, pp. 141–52, Sep. 2003.

[37] S. Sergyan, “Color histogram features based image classification in content-based image retrieval systems,” in *Applied Machine Intelligence and Informatics, 2008. SAMI 2008. 6th International Symposium on*, 2008, pp. 221–224.

[38] M. S. Banu and K. Nallaperumal, “Analysis of color feature extraction techniques for pathology image retrieval system,” in *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on*, 2010, pp. 1–7.

[39] S. E. Umbaugh, Y.-S. Wei, and M. Zuke, “Feature extraction in image analysis. A program for facilitating data reduction in medical image classification,” *Eng. Med. Biol. Mag. IEEE*, vol. 16, no. 4, pp. 62–73, 1997.

[40] “Color Models.” [Online]. Available: http://software.intel.com/sites/products/documentation/hpc/ipp/ippi/ippi_ch6/ch6_color_models.html. [Accessed: 04-Nov-2014].

[41] P. Heckbert, *Color image quantization for frame buffer display*, vol. 16, no. 3. ACM, 1982.

[42] L. Lucchese and S. K. Mitra, *An Algorithm for Fast Segmentation of Color Images*. Springer Verlag, 1999, pp. 110–119.

[43] N. Ikonomakis, K. N. Plataniotis, and A. N. Venetsanopoulos, “A Region-based Color Image Segmentation Scheme,” *Proc. SPIE Vis. Commun. Image Process.*, vol. 3653, pp. 1202–1209, 1999.

[44] D. C. Tseng and C. H. Chang, “Color segmentation using perceptual attributes,” in *Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol. IV. Conference D: Architectures for Vision and Pattern Recognition,*, 1992, no. 1, pp. 228–231.

[45] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000, p. 49.

[46] Sergios Theodoridis and K. Koutroumbas, *Pattern Recognition*, Fourth. Academic Press, 2009, pp. 284–285.

[47] “The MathWorks, Inc.” [Online]. Available: http://www.mathworks.com/.

[48] D. O. F. and D. R. Cave, *Capsule Endoscopy*. Saunders Elsivier, 2008.

[49] C.-C. Chang and C.-J. Lin, “{LIBSVM}: A library for support vector machines,” *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[50] M. Liedlgruber and A. Uhl, “Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review.,” *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 73–88, Jan. 2011.

[51] Y. S. Jung, Y. H. Kim, D. H. Lee, and J. H. Kim, “Active Blood Detection in a High Resolution Capsule Endoscopy using Color Spectrum Transformation,” *2008 Int. Conf. Biomed. Eng. Informatics*, pp. 859–862, May 2008.

[52] Y. Fu, W. Zhang, M. Mandal, and M. Q.-H. Meng, “Computer-aided bleeding detection in WCE video,” *IEEE J. Biomed. Heal. informatics*, vol. 18, no. 2, pp. 636–42, Mar. 2014.