DETECTION OF INTESTINAL BLEEDING IN WIRELESS CAPSULE ENDOSCOPY USING MACHINE LEARNING TECHNIQUES

A Thesis Submitted to the College of Graduate and Postdoctoral Studies

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

Md Hanif Ali Sohag

PERMISSION TO USE

In presenting the thesis in partial fulfillment 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/dissertation.

Requests for permission to copy or to make other use of the 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 S7N 5A9, Canada,

OR

Dean

College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan S7N 5C9, Canada.

ABSTRACT

Gastrointestinal (GI) bleeding is very common in humans, which may lead to fatal consequences. GI bleeding can usually be identified using a flexible wired endoscope. In 2001, a newer diagnostic tool, wireless capsule endoscopy (WCE) was introduced. It is a swallow-able capsule-shaped device with a camera that captures thousands of color images and wirelessly sends those back to a data recorder. After that, the physicians analyze those images in order to identify any GI abnormalities. But it takes a longer screening time which may increase the danger of the patients in emergency cases. It is therefore necessary to use a real-time detection tool to identify bleeding in the GI tract.

Each material has its own spectral 'signature' which shows distinct characteristics in specific wavelength of light [33]. Therefore, by evaluating the optical characteristics, the presence of blood can be detected. In the study, three main hardware designs were presented: one using a two-wavelength based optical sensor and others using two six-wavelength based spectral sensors with AS7262 and AS7263 chips respectively to determine the optical characteristics of the blood and non-blood samples.

The goal of the research is to develop a machine learning model to differentiate blood samples (BS) and non-blood samples (NBS) by exploring their optical properties. In this experiment, 10 levels of crystallized bovine hemoglobin solutions were used as BS and 5 food colors (red, yellow, orange, tan and pink) with different concentrations totaling 25 non-blood samples were used as NBS. These blood and non-blood samples were also combined with pig's intestine to mimic in-vivo experimental environment. The collected samples were completely separated into training and testing data.

Different spectral features are analyzed to obtain the optical information about the samples. Based on the performance on the selected most significant features of the spectral wavelengths, k-nearest neighbors algorithm (k-NN) is finally chosen for the automated bleeding detection. The proposed k-NN classifier model has been able to distinguish the BS and NBS with an accuracy of 91.54% using two wavelengths features and around 89% using three combined wavelengths features in the visible and near-infrared spectral regions. The research also indicates that it is possible to deploy tiny optical detectors to detect GI bleeding in a WCE system which could eliminate the need of time-consuming image post-processing steps.

ACKNOWLEDGMENTS

At first, I gratefully praise to almighty God for enabling me to complete my thesis work. I would like to express my most gratitude and appreciation to my supervisor, Dr. Khan A Wahid, for his continuous support and encouragement throughout the work. It was quite impossible for me to complete this work without his valuable expertise, advice, and encouragement.

Then, I would like to thank my thesis committee for their guidance and insightful comments.

Special thanks to Professor Julia Montgomery of Western College of Veterinary Medicine and Raelene Petracek of Prairie Swine Centre Inc. for providing blood samples to conduct the experiments. I am also thankful to all my lab colleagues who helped me a lot in conducting the experiments.

Finally, I would like to express thanks to my wife and parents for their mental support to complete my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ARL	Amount of reflected light
ASIC	Application-specific integrated circuit
ATL	Amount of transmitted light
AUC	Area under the curve
BIC	Bayesian information criterion
BS	Blood samples
CADSS	Computer-aided decision support scheme
CCD	Charge coupled device
CE	Capsule endoscopy
CMOS	Complementary metal-oxide semiconductor
CNN	Convolutional neural network
CP	Cut-off point
CWE	Conventional wired endoscopy
EDTA	Ethylene-diamine tetra-acetic acid
EM	Expectation-maximization
FWHM	Full-width half-max
GI	Gastrointestinal
Hb	Hemoglobin
LED	Light emitting diode
LGA	Land grid array
LPB	Local binary pattern
MAE	Mean absolute error
ML	Maximum likelihood
MLA	Machine learning algorithm
MSE	Mean squared error
NBS	Non-blood samples

NIR	Near-infrared
OGIB	Obscure gastrointestinal bleeding
PCA	Principal component analysis
RBC	Red blood cell
RF	Radio frequency
RGB	Red, green and blue
RMSE	Root mean squared error
RMSLE	Root mean squared logarithmic error
ROC	Receiver operating characteristic
SBI	Suspected blood indicator
SIFT	Scale-invariant feature transform
SOM	Self-organization map
SVM	Support vector machine
UV	Ultraviolet
WBC	White blood cell
WCE	Wireless capsule endoscopy
WCVM	Western college of veterinary medicine

CHAPTER 1

INTRODUCTION

1.1 Capsule Endoscopy

Wireless capsule endoscopy (WCE) is a procedure that captures images of the digestive tract using a tiny wireless camera. It is a non-invasive, painless endoscopic technique that can be used for diagnostic and other medical experiments to screen all components of gastrointestinal tract. In medical terms, capsule endoscopy is a method for observing the esophagus, stomach, small intestine, and colon within the GI tract [18]-[19]. WCE does not require any needles, fluids or instruments to be inserted into the body due to its non-invasive nature. In a vitamin-sized capsule, the patient swallows a capsule endoscopy (CE) camera. As the capsule travels through the digestive tract, the camera takes thousands of images that are transmitted to a data recorder worn around the waist using a belt. On the other hand, traditional endoscopy includes passing down the throat or through the rectum through a lengthy, flexible tube fitted with a video camera. There are different types of endoscopy procedures such as cystoscopy, colonoscopy, anoscopy, gastroscopy, etc. These procedures are classified on the basis of their region of operation and endoscopy application [20]. Prior to the development of the WCE, the small intestine could only be explored by invasive procedures (intraoperative enteroscopy) or poorly effective methods, such as small bowel series. Traditional endoscopy can create discomfort and have some risks. Over sedation, cramps, persistent pain, or even tissue perforation and minor internal bleeding are some of the major side effects of traditional endoscopies. In addition, it is also restricted to travel within the entire GI tract's abdomen and duodenum [27]. On the contrary, WCE a small electronic device and an advanced version of traditional endoscopy which has overcome these limitations [8]. The

extensive accessibility of WCE, which permits better visualization of the mucosa with few complications, has led to a revolution in small intestinal endoscopy and a substantial boost in CE signs. It helps doctors see within the small intestine - an area that is not easily reached through traditional endoscopy procedures. In patients suspected of bleeding and other kinds of abnormalities such as Crohn's disease in GI, WCE imaging is a must. It can be used for patients with the syndrome of polyposis and small bowl disorder. It also identifies celiac disease, inflammation, swelling, ulcers, bleeding, tumors, lesions, pre-cancer abnormalities, and blockages [6]. It can be endorsed to assess the muscle tube connecting the patient's mouth and belly (esophagus) to search for unusual, enlarged veins (varices).

A group of researchers invented WCE in Baltimore in 1989 and later introduced it as a commercial tool by Given Imaging Inc. PillCam SB2, which is a swallow-able wireless capsule, has been developed by Given Imaging, which has been the only product on the market since 2002. Since 2005, a competitor, Olympus Corporation, has been selling "EndoCapsule" in Europe and received FDA approval in September 2007. From an operational point of view, the difference between the two products is less. A miniaturized color camera and an optical dome with four white light LEDs are located at one end of the PillCam capsule. The camera captures over 50,000 images that are wirelessly transferred to a patient-worn storage device. The video feedback is then uploaded to a nearby storing device, where medical staff can review and examine endoscopic images. WCE technology has been refined over the years to provide superior resolution, increased battery life, and the ability to view various parts of the GI tract. Researchers have also been investigating automatic techniques for detecting abnormalities since the practical use of WCE in 2001 [9], [28]. These methods could reduce the observation time spent by the physicians on medical diagnosis and detection of anomalies [16].

1.2 Gastrointestinal Bleeding

Any form of bleeding that occurs in a tract of human GI is called gastrointestinal bleeding [1]. In several cases, it may be a life-threatening issue, if not properly diagnosed and addressed in time. Based on abnormality symptoms, clinicians conduct various medical exams to verify the nature and source of bleeding [4]. Endoscopy is one of the procedures for detecting GI bleeding. GI bleeding can mainly be classified as overt, occult or obscure bleeding.

1.2.1. Epidemiology of Obscure Gastrointestinal Bleeding (OGIB)

Obscure gastrointestinal bleeding is defined as persistent or recurrent bleeding associated with negative results from endoscopic evaluations of the upper and lower gastrointestinal tract. Depending on the existence or lack of clinically obvious bleeding, OGIB may be further categorized into overt or occult bleeding [43]. Overt OGIB is described as noticeable GI bleeding (e.g. melena or hematochezia) and may be classified as active or inactive bleeding since there is proof of continuing bleeding. OGIB is categorized as occult when there is no proof of visible GI bleeding such as unexplained anemia with iron deficiency suspected of causing GI blood loss [2], [37].

A prevalent clinical presentation of gastrointestinal bleeding is about 1 case per 1,000 people per year [38]. OGIB accounts for about 5% of all GI bleeding instances, with tiny intestines as the suspected cause. This has resulted in the new word "mid-gastrointestinal bleeding" being used to define bleeding between the papilla and the ileocecal valve [39].

Diagnosis and management of OGIB patients are especially difficult due to the small intestine's long and complicated loops [40]. The symptoms present may assist to guide the inquiry plan. Hematemesis may show upper GI swelling, while melena may suggest bleeding from the

nose to the large bowel anywhere. In contrast, hematochezia (the presence of blood and blood clots in the feces) suggests either a reduced GI bleeding or a rapidly higher GI bleeding.

1.2.2. Causes of Obscure Gastrointestinal Bleeding

Obscure gastrointestinal bleeding may result from any lesion throughout the GI tract, although most of it occurs in the small intestine and mainly involves vascular lesions [41]. The rate of OGIB is around 5% among all the patients with gastrointestinal bleeding [98]. Bleeding etiology in the small intestine is diverse and may depend on the patient's age. Tumors, meckel's diverticulum, dieulafoy's lesion, crohn's disease, celiac disease, etc. are the most common cause of OGIB in patients younger than 40 years. On the other hand, in patients 40 years older, angiectasia, nonsteroidal anti-inflammatory drug, enteropathy, vascular ectasias, celiac diseases are the most common causes of OGIB. There are still some uncommon causes of OGIB which includes hemobilia, hemosuccus pancreaticus, aortoenteric fistula, long-distance running, etc. [99].

1.3 Literature Review

In 1806, Phillip Bozzini, known as "Father of Endoscopy," demonstrated a light conductor (Lichtleiter) that allowed a direct view into the living body [21] - [22]. Antonin Jean Desormeaux first used the term 'endoscopy' in 1853 [19]. In 1954, "Fiberscope" was created as an optical unit that transmitted optical images along a flexible axis [23]. In 1955, a French bronchoscopy team broadcast on live television their endoscopic technique [24].

Dr. V. K. Zworykin launched the first wireless machine to send GI tract data in 1957. It was called "Radio Pill" and consisted of a plastic capsule with a length of 1.125 inches and a diameter of 0.4 inches. Signals were transferred according to the pressure variation in the GI Tract [29].

Many doctors and technologists have further created the endoscopic system and now the entire operation is miniaturized to include small camera/s and lighting equipment, communication module and processor [13], [25] - [26]. Microelectronic development, condensed batteries, and effective techniques have enabled the entire system to be transformed into a capsule [30], [27]. In 2001, a firm named Given Imaging launched a full capsule endoscopy with self-illumination, camera, battery, processor, and transmitter [9]. A data recorder device obtains all the related information from the capsule. A typical block diagram of WCE [31] system is shown in Fig. 1.1

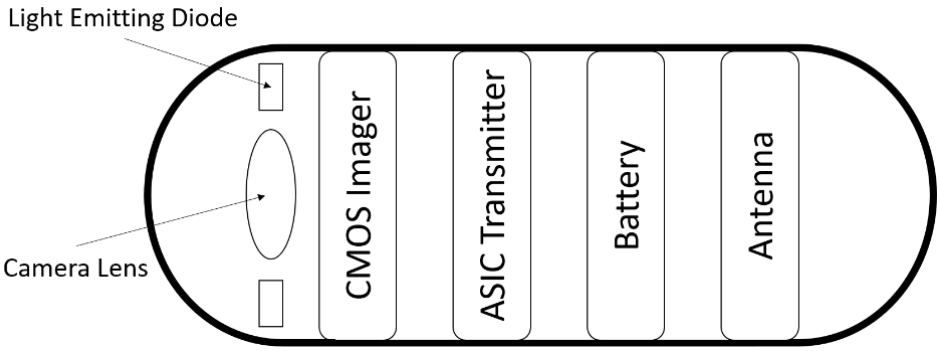


Figure 1.1: A typical block diagram of a wireless capsule endoscopy

The capsule sends the captured images to the data recorder via suitable wireless communication. Then the outputs will be analyzed to diagnose the defects. Fig. 1.2 shows the full operating steps of a WCE system.

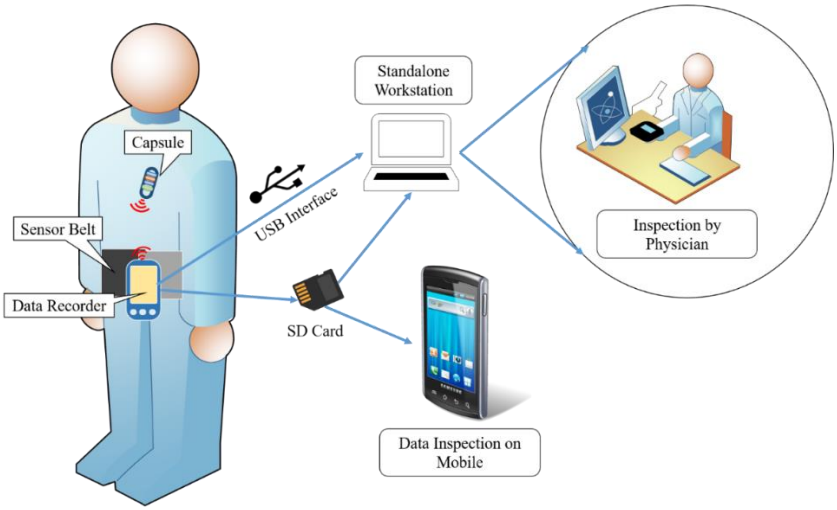


Figure 1.2: A complete system of wireless capsule endoscopy

Given Imaging's first capsule endoscope was PillCam SB, concentrating primarily on the small intestine. Given Imaging had also released some updated versions of its products like PillCam SB 2, SB 3, ESO 2, ESO 3, Colon 2. Some other companies also introduced their products onto the market like Intromedic, Olympus, JinShan Science, and CapsoVison. Table 1.1 shows a brief comparison of the existing capsule endoscopy products.

Table 1.1: Comparisons of capsule endoscopy available in the market

Device	**PillCam SB3**	**PillCam ESO2**	**PillCam Colon2**	**MiroCam**	**Endo Capsule**	**OMOM**	**CapsoCam**
Company name	Given Imaging		Intromedic	Olympus		JinShan	CapsoVision
Region of operation	Small bowel	Esophagus	Colon	Small bowel	Small bowel	Small bowel	Small bowel
Dimension (mm)	11.4 x 26.2	11 x 26	11 x 31	10.8 x 24.5	11 x 26	13 x 27.9	11 x 31
Mass (gm)	3.0	2.9	2.9	3.25	3.3	≤6	4
Battery life	≥8 hr	0.5 hr	10 hr	12 hr	12 hr	6-8 hr or longer	~15 hr
Field of view	156º	172º	172º	170º	160º	140º ±10	360º
Frame rate (fps)	2 or ~6	35	4 or 35	3	2	0.5 or 2	3 or 5
Type of image sensor	1 CMOS	2 CMOS	2 CMOS	1 CMOS	1 CCD	1 CMOS	4 CMOS
Resolution	256x256	256x256	256x256	320x320	1920x1080	640x480	1920x1080
Transmission	RF	RF	RF	HBC	RF	RF	USB

CCD, charge-coupled device; CMOS, complementary metal-oxide-semiconductor

RF, radiofrequency; HBC, human body communication

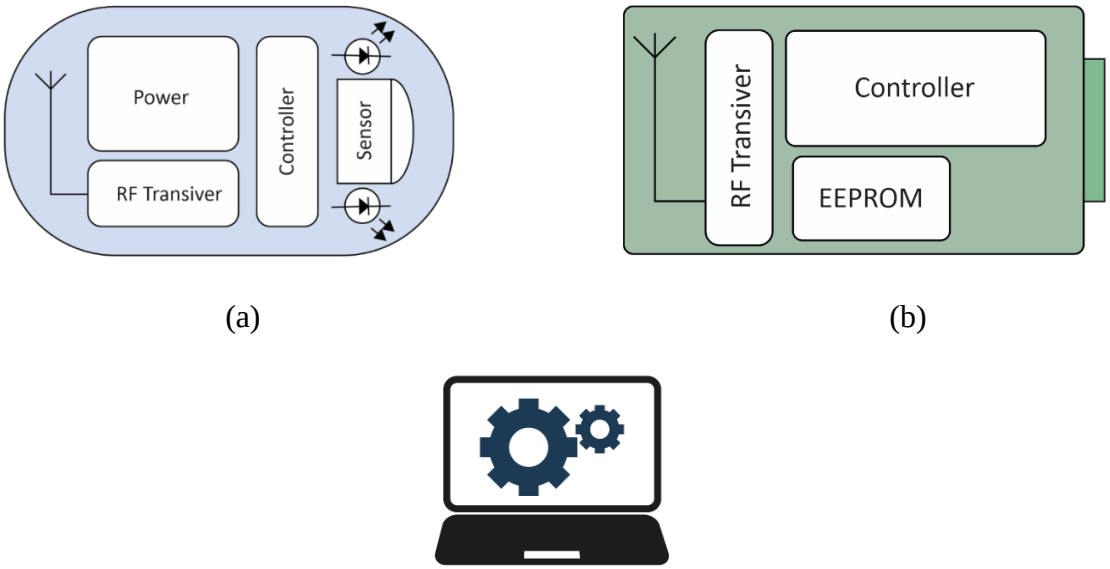
Given Imaging was the first who came up with a solution of automated bleeding detection using their PillCam SB2 capsule [72]. The capsule system had an embedded software called suspected blood indicator (SBI) which could mark the red areas in the frames as the potentially blood which eventually speeded up the bleeding detection diagnosis [14]-[15]. Furthermore, to explore the esophagus PillCam ESO was launched, and PillCam Colon was launched to examine the colon. Jinshan Science and Technology had also developed the OMOM impedance-pH capsule wireless monitoring system which integrated both pH sensor and impedance sensor to monitor patient's pH and impedance data continuously within the esophagus for diagnosing GI function disorders [97].

Most of these commercial capsule endoscopy systems only provide video or image capturing options. Also, there does not have any automated bleeding detection system in some of the systems to reduce the diagnosis time. On the other hand, the systems which have bleeding detection capabilities are usually based on the color properties of the blood which may provide wrong result as the blood color could vary from light red to dark red to brown. So, it could be very difficult for these commercial endoscopy devices to separate blood from non-blood samples since these devices are not using any optical sensors to observe the optical characteristics of blood. Therefore, in this research, an optical sensor-based blood detection system has been proposed using machine learning techniques to solve such problems.

1.4 Structure of a Typical Capsule Endoscopy System

The WCE system has three different parts - a) a capsule with a camera, b) a datalogger for receiving the capsule's data wirelessly, and c) a computer software, to download and analyze recording information. Accurate inspection of a WCE readings is time-consuming and requires concentrated, undivided attention, as anomalies can occur within the range of a very few frames.

On average, it takes approximately 1-2 hours to visualize all the images. However, some WCE device manufacturers have made an effort to develop computer aided software that could reduce the time required to analyze the WCE results. The three main components of a WCE system are shown in Fig 1.3



(a) (b)

(c)

Figure 1.3: Wireless capsule endoscopy components a) capsule, b) datalogger, c) computer software

The capsule generally has a modular structure, which consists of four layers- i) imaging sensor layer, ii) processing unit layer, iii) RF transceiver layer, and iv) power management layer. The overall block diagram of the capsule is shown in Fig. 1.3 (a). The modular structure of the capsule does not require separate modules and makes it more versatile for practical usage. The whole system could be used or modified for different purposes depending on its application. The capsule's size and power consumption should be in a limited range. Therefore, components were chosen considering performance, power requirement and packaging size during the designing process.

The data logger is capable of receiving the capsule's signal and then transmitting the data to a computer. The data logger usually has two modes of operation. First, when a data logger is plugged into a computer via USB port, it enters into the "Software" mode operation. In this mode, the data logger receives signals from the capsule and transmits it to the computer in real-time. On the other hand, when the data logger is powered up by an external power source without having any connection to the computer, the device enters into the "Standalone" mode. In standalone mode, the data logger still receives data from the capsule but stores them into the built-in memory like EEPROM or SD card. By connecting the data logger to the computer, the WCE software could read the stored image or sensor data.

The workstation software depends heavily on the WCE scheme. When administered to a human body, a typical endoscopy capsule operates for about 8 to 10 hours and produces around 50,000 frames. A doctor still has to sit before the computer to determine any faded or unreliable frame. While there could be some features that could assist the doctor to automate abnormality detection, the current systems are still inefficient and unreliable in providing such features. The information recorded on the data logger will be obtained at a workstation during post-processing, where the software is configured to evaluate the information recorded. The respective manufacturer of the commercial WCE system provides the analyzing software. Different WCE manufacturers have developed software with multiple features which are frequently maintained and updated. For example, PillCam develops and provides RAPID reader software which includes some advanced functionalities that help physicians with automatic bleeding detection.

1.5 Clinical Need and Target Population

Since bleeding is a common symptom for many GI diseases, it has major clinical importance for the diagnosis of the relevant diseases. Capsule endoscopy is usually performed only

after endoscopy or colonoscopy does not disclose the origin of bleeding to explore the blood loss. Conventional methods, like gastroscopy and colonoscopy, face problems and cause a lot of pain when examining the small intestine. On the contrary, because of WCE's non-invasive nature, it primarily is gaining popularity. It is safe, painless, free from infection and does not require sedation. It has a higher diagnostic efficiency than many other methods for assessing intestinal lumen and is capable of finding lesions. The main problem of WCE is the long period of medically-assessed bleeding detection steps in the entire GI tract. In addition, smaller bleeding regions and bleeding within a few frames sometimes may not be detected with naked eyes. Therefore, several researchers have given a lot of attentions to the automation of the bleeding detection system.

1.6 Thesis Objective

The goal of the thesis is to develop a machine learning classifier model that can be used to differentiate GI blood from other non-blood substances. To achieve the goal, the following research objectives are set:

a) To explore the spectrum of wavelengths for separating the blood and non-blood samples. These experiments could be carried out using two 6-channel spectral sensors- one in the visible spectral region (AS7262) and other in the infrared spectral region (AS7263).

b) To develop a machine learning classifier model for predicting the bleeding, from non-blood and blood samples by considering specified spectral wavelengths as the features for the model.

c) To compare the prediction results and overall accuracy of two different blood detection methods which predict the unknown samples based on two visible

spectral wavelengths, and three most significant spectral wavelengths from the combined visible and near-infrared light region respectively.

1.7 Thesis Organization

The thesis is structured in eight chapters. A short description of each of the chapters and sections are given below:

Chapter 1: Introduction presents the overview of the WCE system, clinical needs of wireless endoscopy system, the motivation of the thesis, thesis organization.

Chapter 2: The chapter discusses the related literature works on detecting bleeding in the GI tract in the WCE systems. It also provides the overview of the systems which are being used in the commercial capsule endoscopy systems as well.

Chapter 3: The chapter represents the research motivation of the thesis. It also discusses the sample collections procedures which are used for validating the proposed methods. In order to obtain the best results from the experiment, all samples are prepared in an identical manner as far as possible. All the unforeseen deviations were also noted at the time of preparing the samples to discuss unforeseen outcomes.

Chapter 4: Proposed methods are discussed in the chapter. The basic working procedures of the developed sensors to detect bleeding are discussed. The hardware architectures of the developed sensors are clearly explained. The design criteria of the sensors, architectural design, and the results are presented in this chapter as well. The limitations of the sensors and expected results are also discussed in the chapter.

Chapter 5: Methods of selecting the most significant features for the classifier models are discussed in this chapter

Chapter 6: Development of the machine learning models and algorithms related to the pre-processing of the data before feeding into the machine learning model are discussed in the chapter. The complexity analysis, performance of the model, and model's accuracy are also discussed.

Chapter 7: The chapter discusses the comparative analysis of all the methods by incorporating the comparison between the existing methods.

Chapter 8: The chapter presents the conclusions regarding this research and also points out some recommendations for future directions related to the study.

CHAPTER 2

OVERVIEW OF BLEEDING DETECTION IN WCE

2.1 Introduction

Wireless capsule endoscopic recordings are now being inspected by using advance digital technologies, including analysis of images by using computers and other visual methods to detect GI bleeding. As this technique involves capturing different images in the gastrointestinal tract especially in the region of the small intestine, it has minimized the time of the doctors to interpret the traditional video recordings.

Capsule endoscopy is a non-invasive method commonly used to identify unknown internal GI bleeding which is not discovered by X-rays or ultrasounds [100]. The standardized method to investigate the existence of the disease involves capturing approximately 50,000 images of the GI tract by using camera sensors incorporated inside an endoscopic capsule [7],[10]-13]. The capsule is usually swallowed by the patient and excreted out through bowel movement. It usually takes about 1 to 2 hours to analyze the internal imaging of the GI tract by a physician [42]. The main task is to read and interpret those images to determine the existence and cause of bleeding. This method is unable to detect disease at the initial stage as it can miss smaller areas of GI bleeding or can have blurred images of the bleeding areas. The PillCam which is used as a suspected blood indicator had also shown an 80% inaccuracy in detecting and visualizing sensitive and smaller areas of bleeding within the gastrointestinal tract [71]. The main problem is that there is no confined area of bleeding and erosion of the lining of GI. For this reason, it was difficult for SBI PillCam to detect normal bleeding at an early stage of the disease. The color of blood even misinterprets the findings because the physicians might get confused between light or dark red to

brownish colored blood with interstitial fluid. The color of blood changes in different areas and it is also affected by the position of camera within the capsule [69].

2.2 Methods and Procedures

Recently many researches have been carried out about capsule endoscopic techniques. Various methods of identification of causes of the unknown GI tract bleeding were proposed by the researchers. Based on the research methodologies, the investigated research studies can be classified as:

- Signal processing methods
- Images and color processing techniques
- Use of artificial intelligence

All methodologies involving the processing of data in the frequency domain are included in the category of signal processing. Signal processing techniques include the Fourier transforming technique, wavelength transforming technique and band-bass signal transformation technique. Color and image processing techniques involve the transformation of color, segmentation etc. However, these methods are also called computer-vision technologies. On the other hand, the use of artificial intelligence includes neural networking and clustering of digital images.

Nowadays commonly used techniques for bleeding detection are based on pixel, image and patch. Statistical analysis and color resemblances are the main features of color and image processing techniques for the identification of GI bleeding [47]. This technique is less time consuming but unreliable in terms of the accuracy of results. On the contrary, the pixel-based technique has higher accuracy but it is cost-effective due to the use of neural networking and color properties [50] - [51]. However, the patch technique has better durability and speed. In the patch technique, images are split into patches or boxes by binary division method. Then from the patches,

those with the most cleared images are selected to identify the blood and non-blood areas [52] - [53].

The pattern and indications of pixel and image-based techniques have been described by many researchers in their studies. Statistical readings, colored models and histograms are image-based techniques. The ratio of range to color is based on pixel measurements.

There are numerous alternative methods discovered by the scientists to detect GI bleeding, their specificity and accuracy have also been assessed in many studies. The assessment of an individual method to obtain highly accurate results is discussed in this study. The advantages and limitations of different methods are also described. It is not a comparative study because each technique has a different method and every technique has different indications.

2.2.1. Blood Detection in WCE using Expectation-Maximization Clustering

Recently a new suspected blood indicator (SBI) system proposed by M2A software has shown better results but it cannot replace manual human methods in terms of precision at an early stage. It has been reported that the sensitivity and specificity of the Suspected blood indicator system is approximately 72 and 85 percent. To encounter this problem for the definitive diagnosis of GI bleeding, the expectation-maximization (EM) clustering method of statistics and Bayesian information criterion (BIC) [93] was suggested. The major drawback of EM cluster analysis is, it is insufficient information and inefficiency in terms of comparative studies.

However, the EM analysis is effective in the estimation of calculating the extent of bleeding to a maximum limit even if incomplete or obscure information is given. The maximum likelihood (ML) calculates the parameters with the given information. There are two steps in EM statistical analysis- E step and M step. In the expectation or E step, the missed information is estimated by using model parameters and observed information. In the M step, the probability data

is calculated by using the assumed missed information. The missed data calculated through the E step is used in the M step instead of using real information.

The main purpose of this method is to differentiate between colored pixels of blood and non-blood areas. As dark pixels have zero chrominance so it is quite difficult to distinguish between dark blood pixels and dark non-blood pixels. Some areas of bleeding are blue and some areas of dark blood pixels are red. That is why the same dark pixels are clustered together to have a comparative probability model for pixels of blood and pixels of non-blood areas.

The EM system is fed with the non-bleeding and bleeding pixels to identify accurately the bloody areas within the GI tract. There are also subcategories of bleeding pixels, such as dark red and bluish-red, that are to be distinguished. The Gaussian mixture density parameter categorizes each subcategory. Following are the steps to detect GI bleeding:

1. Removal of dark pixels of blood
2. Selection of pixels for blood detection
3. Tiny areas of blood are filtered from the regions to discriminate from the actual bleeding areas.

2.2.2. Color Range Ratio Technique Based on Pixels

A color descriptor technique based on the color spectrum of RGB pixel values was discussed in [60], which calculated the color range ratio for each pixel values. The main focus of the study was to separate the bleeding pixel values from the pixel values of non-bleeding areas. Blood color purity was used as an indicator in the pixel-based method. However, sometimes it failed to detect areas of bleeding. For this reason, the pixel values were explored to set a threshold level to detect GI bleeding. There were two categories in the color range ratio technique- one for accurate detection in bleeding areas and another for areas adjacent to bleeding. The research

suggested the RGB values for bleeding pixels to be in a certain range, such as - [(R=177 & R= 215) & (G=99 & G= 56) & (B= 52 & B= 20)].

2.2.3. Endoscopic Images Classification Based on Texture Patterns

Different methods of endoscopic images have been recognized by scientists to identify malignant areas such as chronic ulcers and E-coli [96]. For the extraction of texture features, a local binary pattern (LBP) is used and for the evaluation of the distribution of local binary pattern, the ratio of G- statistical log-likelihood is used, which tests the goodness of fit of observed frequencies to their expected frequencies [44]. The LBP system has been proven as a much accurate measure of textured images as it has developed many extensions for the sake of precision and sensitivity of complex cases. It is now known as a very helpful tool in many types of research. With the combination of image intensity and a standardized LBP system, more precise and fine images and results can be obtained. Initially, the LBP system does not incorporate color contrast intensity and image textures. A modified version of Kullback's criterion is G- statistics [96], which is used to indicate the probability index of samples coming from the same areas. The lower value of G-statistics, the more likelihood of its occurrence is present. self-organization map (SOM) is used for further classification. The greatest advantage of SOM is its uncontrolled nature, which is quite beneficial for the complex endoscopic images.

2.2.4. Methods Based on Deep Learning

Computer-aided decision support scheme (CADSS) systems in capsule endoscopy are being used in many types of research from the early years of the 2000 century [17]. CADSS is used to achieve higher precision values to diagnose a disease at an early stage. However, amendments in this technique have been made to improve image quality and deep extraction of data. The proportion rate of standard endoscopic technique and capsule endoscopy with CADSS is quite similar [17]. Still, there are many advantages of computer-aided technologies such as

higher efficacy, reduction of time to diagnose disorders at early stages, fine video quality, ability to detect minor hemorrhages [61]. Specific color and texture contrast-based techniques of CADSS systems are quite helpful in hemorrhagic and malignant lesions detection [49], [62] - [64]. Support vector machines (SVM) [62] - [64], binary classifiers [65], and neural networks [49] are also capable of classifying target classes in endoscopic images.

Although earlier techniques based on machine learning with invariant characteristics have shown concordant results for the identification of various lesions, those still have a very consistent ratio of design issues related to the lack of proper training and testing datasets. To overcome such problems, a few techniques related to deep learning-based capsule endoscopy have been suggested in table 2.1 [79] - [86].

Zou *et al.* had proposed a convolutional neural network (CNN) based technique to solve issues related to the digestive organ classification in capsule endoscopy [79]. In this concern, three possible classes of these issues include colon, stomach and small intestine, had been analyzed. In the comparison with Scale-invariant feature transform (SIFT) and SVM methods, the proposed techniques had shown a divine precision of almost 95.52% in 15,000 endoscopic images of 25 patients. In another technique by Segui *et al.* [80] for the classification of motility occurrences such as wrinkles, bubbles, transparent blob, wall and turbidity, the scored percentage of precision was 96.01 for 100K training and 29K testing images. On the contrary, the results of gained precisions were only around 82 to 83 percent in case of handcraft characteristics techniques such as SIFT [68], gist [87].

Table 2.1: State-of-the-Art deep learning-based methods for capsule endoscopy [92]

Article Cited	Class	No. of training/testing images	No. of patients or videos	Features	Accuracy	Specificity
[79]	Localization[a]	60K/15K	25 patients	Alexnet	95.5%	N/A
[80]	Scene classification[b]	100K/20K	50 videos	CNN	96.0%	N/A
[81]	Bleeding	8.2K/1.8K	N/A	Alexnet	99.9%	99.2%/ (N/A)
[82]	Haemorrhage	9,672/2,418	N/A	LeNet AlexNet GoogLeNet VGG-Net	100%	98.7%/(N/A)
[84]	Polyp	4,000 /(N/A)	35 patients	SSAE	98.0%	N/A
[86]	Various lesions[c]	465/233	1,063 volunteers	CNN	96.3%	90.7%/88.2%
[85]	Hookworm	400K/40K	11 patients	CNN	88.5%	84.6%/88.6%
[83]	Angiectasia	600/600	200 videos	CNN	N/A	100%/96%

CNN, convolutional neural networks; SSAE, stacked sparse autoencoder.

LeNet, AlexNet, GoogLeNet, VGG-net are CNN architectures

a) Localization, localization of stomach, small intestine, colon.

b) Scene classification, scene classification of bubble, wrinkle, turbid, wall, clear.

c) Various lesions, gastritis, cancer, bleeding, ulcer

In order to detect bleeding and haemorrhaging of positive images, AlexNet based deep learning-based technique had shown the best results with the precision percentage of 99.9% in 2,850 positive images [81]. Another technique had also shown a 100% precision in 390 positive images [82] to detect haemorrhaging or bleeding. GoogleNet that is the largest performance network and LeNet that is the smallest performance network has an undoubted performance discrepancy among the different deep learning networks up to 100% and 97.44% respectively. A CNN based network to detect GI angiectasia with semi segmentation algorithm was suggested in [83], which obtained a specificity of 96% and a sensitivity of 100%.

Although, the latest deep learning techniques have proved to be efficient in gaining result in some characteristics but in some instances, the hand-crafted characteristics have shown a better accuracy in results. In the case of medical modality classification, both the LBP [89] and Harris

corner [66] with SIFT features have provided more efficient classification results as compared to the CNN based technique with an inadequate training dataset [88].

One of the major drawbacks of deep learning methods is the need for a very huge dataset. In addition to the dataset problem, the deep learning-based technique has another issue of image overfitting. In case of small datasets, the problem of overfitting always doubles its problems. However, there are certain measures available in deep learning techniques to eliminate the problem of overfitting [90] - [91] which might make a deep learning-based approach more practical.

CHAPTER 3

RESEARCH MOTIVATION AND SAMPLE PREPARATION

3.1 Research Motivation

Bleeding is a common sign of many gastrointestinal diseases, and therefore the monitoring of bleeding is of great medical significance in the diagnosis of relevant diseases. A supporting diagnostic approach can be very helpful in this scenario. But it is a big challenge in the field of capsules endoscopy to locate and diagnose obscure GI bleeding. Most of the research in WCE images on automatic blood detection is primarily image-based or pixel-based which has a significantly higher false alarm rate since the real blood does not show a consistent pattern or distribution in all types of situations. Therefore, several other factors are directly involved in the detection of bleeding. The color properties of blood, for instance, are not always the same at various levels of oxygenation and may appear to be bright, dark or even black [67]. Blood can even be mixed with secreted fluids or digested food and may differ from normal blood [55]. Because of illumination light and camera angle, the images taken by the WCE camera may also appear completely different for blood [54]. To overcome this situation, the research proposes a sensor-based solution to detect gastrointestinal bleeding based on blood's optical properties with a higher true-positive (TP) rate while keeping the false-positive (FP) rate reasonably low.

3.2 Theory of Operation

Blood is a fluid which provides nutrients and oxygen in bodily cells and carries metabolic waste products from the cells. It is composed of red blood cells (RBC), plasma and white blood cells (WBC) of 45%, 54.3% and 0.7% [7]. RBC is also known as an erythrocyte containing hemoglobin, a protein containing oxygen within the RBC. Blood is red because of hemoglobin.

The blood color also depends on the level of blood oxygen saturation. The oxygenated blood is bright red whereas the deoxygenated blood is dark red [7] - [8]. The saturation of oxygen thus plays an important role in blood spectrophotometry. Therefore, the percentage of oxygen saturation depends on optical properties such as reflection, dispersion, transmission. The erythrocyte is discoid and has very strong light dispersion and absorption characteristics. The presence of RBC dominates the optical properties of the blood, and that is why these properties are also entirely dependent on blood dilution [9].

The spectrophotometer is a device that measures the intensity of an object's emitted or reflected light, transmitted light through any substance and few other optical properties [10]. There are few essential parts of this device, such as a light source, a filter that only allows a specific portion of the light spectrum, and a detector (photo-resistor, photodiode or phototransistor), which measures the intensity of reflected or transmitted light. Fig. 3.1 illustrates how spectrophotometers measure the intensity of the light [11]. The quantity of the reflected light from the sample's surface is measured as a percentage of the incident light. It is ideally 100 percent for a white surface and ideally 0 percent for each wavelength for a black surface. Substances other than black or white surface, however, have a different amount of transmitted light percentage at different wavelengths.

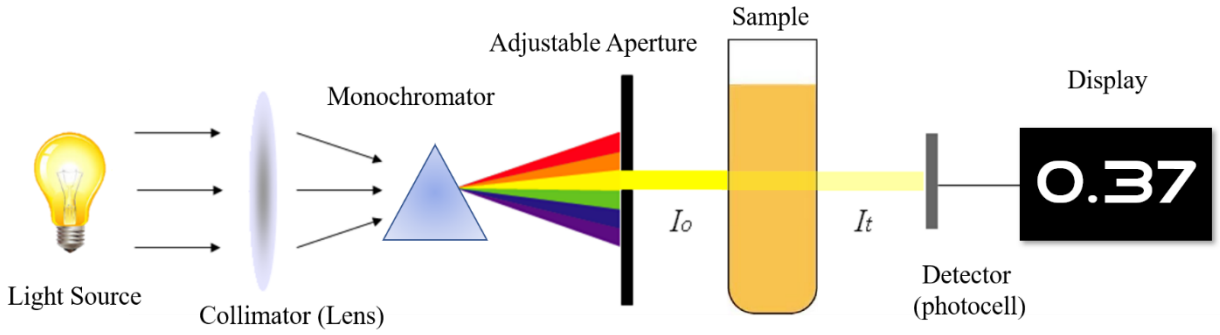


Figure 3.1: Basic working principle of a spectrophotometer [95]

In the interaction between blood and light, the internal structure of hemoglobin plays a significant role due to its reddish nature in the blood. Different portions of the light are transmitted, absorbed, scattered or reflected when the spectrophotometer emits white light on blood samples. The reflection can then be measured against the wavelengths varying from 400nm to 700nm with an increase of 10nm, such as 400nm, 410nm, 420nm, ...,680nm, 690nm, 700nm.

Due to the specific optical characteristics, all other substances will differ from the blood, and the amount of reflection will, therefore, be different. The reflected spectrum pattern is mainly used to distinguish between blood samples and non-blood samples. Instead of using the whole spectrum, however, it could be easier to optimize the decision-making parameters by using a few wavelengths of the spectrum. The initial objective was to find the optimum wavelength through this experiment with a spectrophotometer [35].

3.3 Related Works

In [35], two distinct data sets were used to explore the spectrum of light that could be used to distinguish blood samples from non-blood samples using a spectrophotometer. The research suggests that the proportion of the amount of reflected light (ARL) at two distinct wavelengths could be used to differentiate the samples. The ARL was also suggested as a parameter to separate the samples according to some of the past studies [3] – [5].

The ratio of ARL at two wavelengths (λa, λb), for any sample n, could be expressed as

$$R_n(\lambda_a, \lambda_b) = \frac{\text{ARL at } \lambda_a \text{ for sample n}}{\text{ARL at } \lambda_b \text{ for sample n}} \quad (3.1)$$

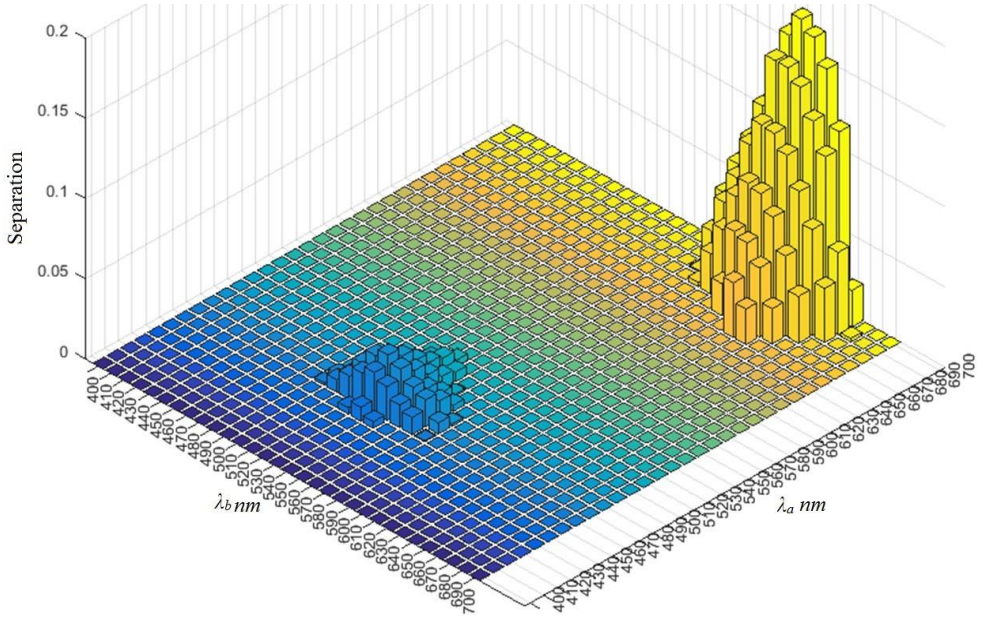


Figure 3.2: The ratio of ARL at two different wavelengths for separating blood from non-blood samples [35]

Fig. 3.2 shows the maximum separation of two distinct wavelength ranges (480nm, 530nm) and (700nm, 630nm). BS and NBS can easily be separated based on the vertical axis. The wavelengths (λ_a, λ_b) of the Eqn. 3.1 are illustrated using the two horizontal axes. A cut-off point (CP) of the ratio was determined to distinguish BS from the NBS. By subtracting the cut-off point, the normalized ratio is calculated from the ARL ratio. A sample may be considered as a blood sample or as a non-blood sample based on the following conditions-

$$\text{Blood samples: } R_{BS_i} >= \text{CP} \tag{3.2}$$

$$\text{Non-blood samples: } R_{NBS_i} < \text{CP} \tag{3.3}$$

Based on two different datasets, the research in [35] also suggests that, there is a better separation using the ARL value of (480nm, 530 nm) than the (700nm, 630nm). The cut-off point is 1.2093 for the (700nm, 630nm) range, which implies that if the proportion is higher than or equal to 1.2093, the sample is very probable to be a blood sample; otherwise, the sample would be a non-

blood sample. Table 3.1 demonstrates the cut-off points values and accuracy for the two different pairs of wavelengths.

Table 3.1: The value of the separating point and accuracy at different pairs of wavelengths

Pair of wavelengths	**ARL CP Value**	**Accuracy**
(700nm, 630nm)	1.2093	84%
(480nm, 530nm)	0.9189	98.48%

According to work in [35], two pairs of wavelengths have been suggested which can be used to detect bleeding in the GI tract. In this thesis, these wavelengths will be analyzed with the developed hardware design. In order to develop a proper working bleeding detection sensor, two other spectral sensors (AS2762 and AS7263) will also be tested to determine bleeding accurately and reliably in the GI tract.

3.4 Sample Preparation

For an endoscopic operation, the patient has to follow some periodic activities [2]. The patient will avoid eating at least 8 to 10 hours prior to endoscopy in order to obtain clear images of the GI tract. Sometimes the physician even asks for a laxative to flush out the small intestine which can improve photos. Therefore, in a clearer perspective, saliva, mucus, and few other digestive enzymes are the only possible accessible substances that could be discovered in an ordinary GI tract [70]. Blood or similar substances will be present in the event of an abnormality. The aim of this research is to use separate wavelengths of LEDs to distinguish blood from other substances. Three separate sensors were initially used in this study. Two different datasets containing blood samples and non-blood samples were used for each sensor. The blood samples are mainly solutions of bovine hemoglobin. As the patient has to fast for around 8-10 hours before the endoscopy, there is less chance of having other substances rather than blood or some food

colors which could still be in the GI tract for that long of fasting period of the patient. So, in this research only food colors were considered as the non-blood samples.

Figure 3.3: Few of the blood solutions with different dilutions

All solutions are stored in test tubes. The sample tube dimensions were 40ml. Samples of bovine hemoglobin were produced from crystallized hemoglobin. Fig. 3.3 shows some of the blood samples. Many of the reagents used in research are solutions that need to be bought or prepared. The exact concentration value is very important in this research. Therefore, the solution concentration and its preparation method must be as accurate as possible. This chapter discusses useful data on the fundamental ideas and guidelines for preparing blood samples using Sigma H2625 hemoglobin.

3.4.1. Preparation of Bovine Hemoglobin Blood Samples

Hemoglobin (Hb) is a complex protein that contains mainly iron molecule. The ordinary erythrocyte concentration of Hb is 34%. It is the most significant respiratory protein in vertebrates because of its ability to transport oxygen from the lungs to body tissues. It also helps to expel carbon dioxide from the body. The molecule of hemoglobin iron also plays an important role in

assisting red blood cells to gain their disc-like shape, helping them move through blood vessels readily.

Sigma H2625 is the crystallized bovine hemoglobin used in the preparation of samples for hemoglobin. Hemoglobin is generally expressed in grams per blood deciliter (g/dL). A lower blood hemoglobin level is directly associated with a low oxygen level, while a higher hemoglobin level is associated with a higher oxygen level [46]. Biological and biochemical researches rely entirely on the ability to detect a range of substances concentrations. In this research, understanding the concentration of hemoglobin in the blood is very important. So, it is required to reliably and accurately calculate how much hemoglobin crystal solid in a certain volume of distilled water would need to be dissolved to prepare the proper amount in a volume of 1 ml of blood.

The amount of a substance is often measured by how much it weighs – i.e. in g or mg or kg etc. The concentration is the amount divided by the volume it is dissolved in.

$$concentration = \frac{amount}{volume} \tag{3.4}$$

For example, if 0.9g of sodium chloride is dissolved in 100mL of water, it is referred to as 0.9% w/v since it is 0.9g weight divided by 100mL volume and "w/v" stands for weight per volume." The concentration of this solution could be described as 0.009g/mL or 9mg/mL. Alternatively, if 0.9g is dissolved in 100mL, then 9g would be dissolved in 1000ml (=1L) and the concentration could also be written as "9g/L"

In this research, the reagents are mixed as percent solutions since dry reagents of hemoglobin are used. When preparing solutions from dry reagents, the same mass of any reagent is used to make a given percent concentration although the molar concentrations would be different. In general,

$$Weight\ percent\ (w/v) = \frac{mass\ of\ solute\ (g)}{volume\ of\ solution\ (ml)} \times 100 \tag{3.5}$$

So, for this research, 100ml of 10% solution of dry hemoglobin reagent would contain 10g dry hemoglobin crystal in a final volume of 100ml.

Ten blood samples were prepared in this research work. In order to do that, firstly the weight of an empty test tube was assessed with a scientific measuring scale. After that, the weight of the test tube along with the bovine hemoglobin crystal was again measured. The difference was the mass of hemoglobin between these two measurements. Finally, 100mL of water was poured into the test tube. Then the test tube was correctly shut and shaken. Table 3.2 shows the list of blood samples prepared for the experiment.

Table 3.2: List of blood samples

Sample Number	Sample Description
1	Hemoglobin -1 %
2	Hemoglobin -2 %
3	Hemoglobin -5 %
4	Hemoglobin -8 %
5	Hemoglobin -10 %
6	Hemoglobin -13 %
7	Hemoglobin -15 %
8	Hemoglobin -17 %
9	Hemoglobin -20 %
10	Hemoglobin -40 %

3.4.2. Preparation of Non-blood Samples

Different kinds of food colors were used as non-blood samples in this research. The samples contain red, tan, yellow, pink and orange food colors with different levels of concentrations. For each of the food colors, 5 different levels of concentration were considered.

Figure 3.4: Artificial food colors as non-blood samples for the experiment

The non-blood samples of twenty-five food colors were shown in Fig. 3.4. In Table 3.3, the full list of NBS samples were also provided.

Table 3.3: List of non-blood samples

Sample Number	**Food Color Sample Description**
1-5	Red (5 different concentrations)
6-10	Tan (5 different concentrations)
11-15	Yellow (5 different concentrations)
16-20	Pink (5 different concentrations)
21-25	Orange (5 different concentrations)
Total	25 Non-blood samples

3.4.3. Preparation of Blood and Non-blood Samples with Pig's Intestine

Since testing on human requires additional permission and authorization, it is wise to use other animal which has similar gastrointestinal tract and similar digestive pattern like humans. In this case, pig's intestine can be used for experimental purposes as it is an omnivorous representative with metabolically similar gastrointestinal functions close to humans [59]. Though the pigs have narrower intestinal lumens, sparse transverse folds and wider but less prominent villi compared to humans, it can be used for this research due to its omnivorous characteristics. For this

research, tissue from pig's small intestine was used in vitro to provide a practical test environment of the lumen wall which could also be considered as a viable resource in the development of the sensor.

In order to prepare the samples using pig's intestine, the full intestine was cut into several pieces.

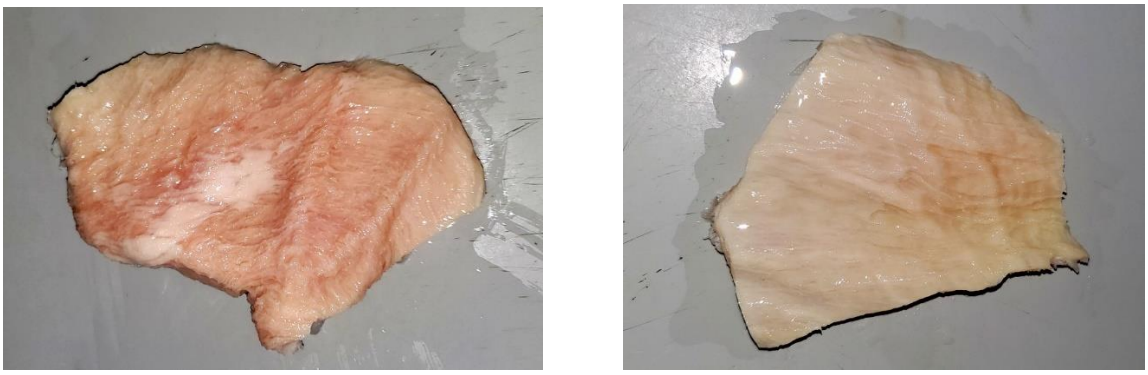
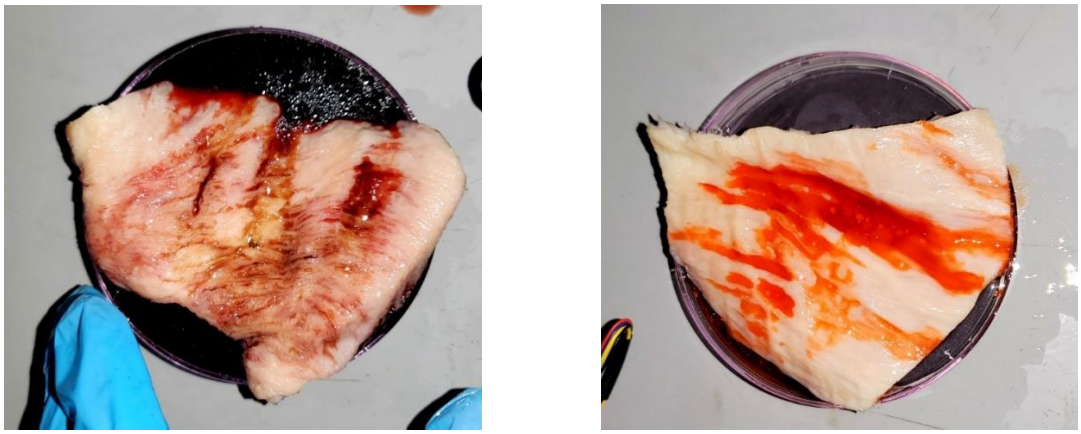


Figure 3.5: Some of the pieces of the whole pig's intestine for the experiment

The blood and non-blood samples were mixed carefully with the pig's intestine to mimic the real test environment. Fig 3.6 (a) shows the piece of intestine with blood sample and Fig 3.6 (b) shows the piece of intestine with non-blood food color sample.



(a) (b)

Figure 3.6: Sample preparation a) pig's intestine mixed with blood sample, b) pig's intestine mixed with non-blood (tan color) sample

After including the pig's intestine, the blood and non-blood samples were 22 and 32 respectively. Table 3.4 and Table 3.5 describes the blood and non-blood samples

Table 3.4: List of blood samples with pig's intestine

Sample Number	**Sample Description**
1	Hemoglobin ((Hb)-1 %
2	Hemoglobin -2 %
3	Hemoglobin - 5 %
4	Hemoglobin - 8 %
5	Hemoglobin -10 %
6	Hemoglobin - 13 %
7	Hemoglobin - 15 %
8	Hemoglobin - 17 %
9	Hemoglobin - 20 %
10	Hemoglobin - 40 %
11	Intestine with Hb - 1 %
12	Intestine with Hb - 2 %
13	Intestine with Hb - 5 %
14	Intestine with Hb - 8 %
15	Intestine with Hb - 10 %
16	Intestine with Hb - 13 %
17	Intestine with Hb - 15%
18	Intestine with Hb - 17 %
19	Intestine with Hb - 20 %
20	Intestine with Hb - 40 %
21	Intestine with Hb - random %
22	Intestine with Hb - random %
Total	22 Blood Samples

Table 3.5: List of non-blood samples with pig's intestine

Sample Number	Sample Description
1-5	Red (1-5 levels concentration) color
6-10	Tan (1-5 levels concentration) color
11-15	Yellow (1-5 levels concentration) color
16-20	Pink (1-5 levels concentration) color
21-25	Orange (1-5 levels concentration) color
26	Intestine Only
27	Intestine with 4^{th} level red color
28	Intestine with 5^{th} level red color
27	Intestine with 4^{th} level tan color
28	Intestine with 3^{rd} level tan color
27	Intestine with 3^{rd} level pink color
28	Intestine with 4^{th} level yellow color
Total	32 Non-blood samples

*higher level means higher concentration

3.5 Justification of Sample Selection

Blood specimens from bovine hemoglobin solutions were prepared for the dataset. There was also a broad variation in the range of blood concentration from very lower to higher blood concentrations. The range of blood sample concentrations were varied from 1% to 40% solutions. The broad range of solutions were taken into consideration to determine where the developed sensor provides better results or fails. Food colors were considered as non-blood samples primarily because those could somehow mimic the gastrointestinal wall color and blood color. In addition, some of the past researches used similar NBS for their study [74], [77] - [78].

Although the sensor is designed to identify the bleeding of the human intestine, the blood samples used in this research were produced of bovine hemoglobin. The reasons behind not using human blood samples are primarily owing to the absence of equipment in the workstation and the

restricted authorization that only enables the animal blood samples to be handed over. Since there are many similarities of hemoglobin and blood substances between human and other mammals, the results of the assessment using animal blood samples are therefore anticipated to be comparable to those of human blood [32].

CHAPTER 4

PROPOSED METHODS

The following chapter provides an overview of the proposed bleeding detection methods and then provides data preparation, data collection, working steps information of the proposed methodologies. Data preparation consists of steps to prepare blood and non-blood samples. Sigma H2625 crystallized bovine hemoglobin was used for the blood samples. Different artificial food colors liquid solutions are used to prepare the non-blood samples. The collection of data involves measuring the reflected light values for various wavelengths of light transmitted from the optical sensors. The working procedure describes the system for three different methods developed using microcontroller and Raspberry Pi. The processing unit for the first method was developed using an 8-bit RISC microcontroller, whereas in the last two methods Raspberry Pi was used as the processing unit.

4.1 Method 1: Optical LED Sensor based on Two Specific Wavelengths

The working principle of the sensor for blood presence detection is based on blood optical properties. According to the research in [35], the two sets of wavelengths were selected. The first set includes LEDs of 700nm and 625nm, and the second set includes LEDs of 475nm and 530nm. The main idea for gastrointestinal bleeding recognition in real-time includes the development of an optical sensor that can be swallowed to detect effective bleeding in the esophagus and stomach or attaching it close to the gastrointestinal tract wall, close to a prospective source of bleeding for continuous monitoring. The system is battery-powered and includes two specific wavelength LEDs, an AVR microcontroller, photodiode and related peripheral circuits for blood presence

detection. Fig. 4.1 shows the functional block diagram of the optical sensor-based bleeding detection system.

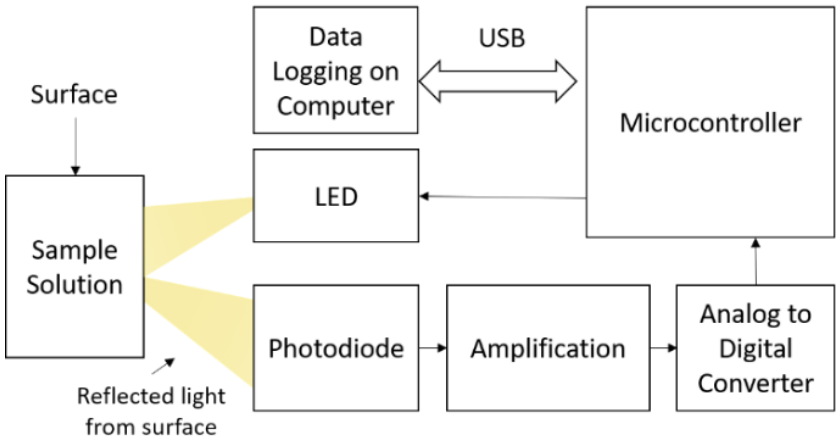


Figure 4.1: Block diagram of the blood detection sensor system

4.1.1. Photodetector

A photodetector is mainly a light or electromagnetic radiation sensor [45]. The sensor's p-n junction converts light photons into energy. Some photodetector examples are photodiodes and phototransistors. It is a light-weight sensor that transforms light energy into electric power or current voltage proportionate to the incident optical power.

In the research, photodiodes were used as photosensors. A photodiode is a form of p-n junction semi-conductive device. In order to produce an electric current, the photodiode accepts light energy as input. It is sometimes named as a light detector or photosensor as well. It works in reverse bias, i.e. the p–side of the photodiode is linked to the negative battery terminal (or energy supply) and n–side of the positive battery terminal. Silicon, Germanium, Indium Gallium Arsenide Phosphide and Indium Gallium Arsenide are typical photodiode materials. Internally, there are optical filters constructed into a photodiode and a surface area. The response time of a photodiode

usually gets slower with the increase of its surface area. As shown in Fig 4.2 below, it has two terminals. The smaller terminal acts as a cathode whereas the longer terminal as anode.

Figure 4.2: A photodiode

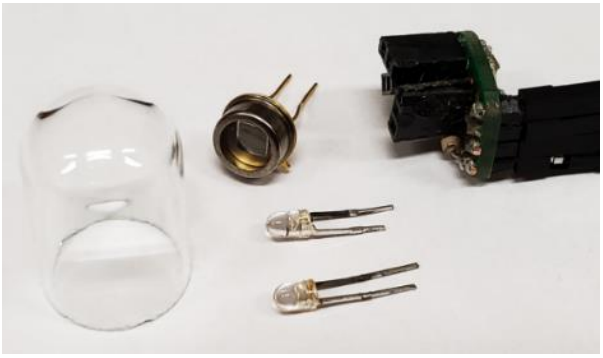
In some spectral region (range of optical wavelengths), the photodetector has to be adaptive according to application. In some situations, the responsivity needs to be continuous or at least specified within a certain range of wavelengths. In certain wavelength ranges, it is also necessary to have zero response. A photodiode that is sensitive only to visible wavelengths of light and provides zero response to the other wavelengths of light was used in the experiment.

4.1.2. Working Principle and Developed Sensor System of Method 1

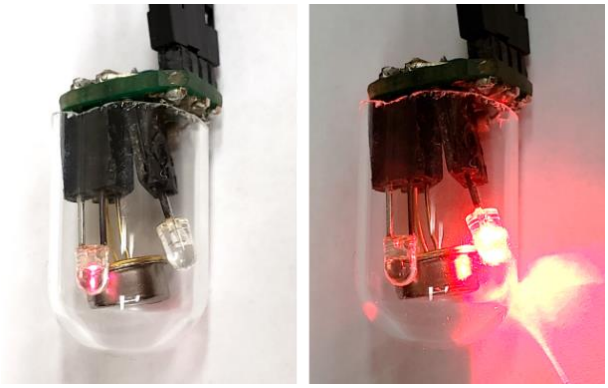
The main principle of the method is based on the amount of reflected light theory. When light is reflected by the presence of blood cells in the optical path in a container, it provides the reflected values for different concentrations of blood. The sensitivity and accuracy depend on the sensitivity and stability of the intensity measurement system and on the fluctuations of the incident light respectively. The purpose of this study is to propose a simple method for improvement of both the sensitivity and stability of the blood detection system, without using complicated equipment, and to describe the construction of a practical bleeding sensor system.

The prototype was developed using an 8-bit RISC microcontroller operating at a clock speed of 16MHz to detect the ARL of the samples and a high-speed epitaxy PIN photodiode with

an active region of 13mm^2. The photodiode has a quick response time which can detect the pulsed light readily. Fig. 4.3 shows the sensor's fundamental structure with all of its components.



(a)



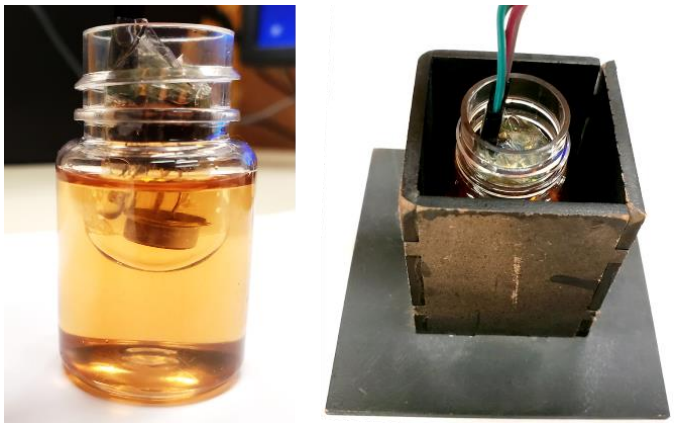
(b)

Figure 4.3: Developed prototype a) all parts of the blood detection sensor b) blood detection sensor in operation

4.1.3. The Procedure of Data Collections for Method 1

To justify the proposed algorithm with the sensor prototype, 10 blood samples of 1%, 2%, 5%, 8%, 10%, 13%, 15%, 17%, 20% and 40% dilutions of crystallized bovine hemoglobin blood were prepared. Five food colors with different dilutions totally 25 samples were also included as non-blood samples for the testing.

Two sets of LEDs were individually tested to separate the blood samples from the non-blood samples. The first set contains LEDs with wavelengths of 700nm and 625nm, and the second set contains wavelengths of 475nm and 530nm. The LEDs were placed very close to the sample surface so that only the specified light on the wavelength was transmitted to the sample surface. Then the microcontroller toggled those LEDs with a defined interval and reflected light from the sample's surface was measured using the high-speed epitaxy PIN photodiode. It is very important to measure the reflection for the specific wavelength of lights for an accurate result. If other light sources are present around the samples, the final results may not be as accurate as expected. That is why the entire experiment was carried out in a dark and closed environment using a black box shown in Fig. 4.4.



(a) (b)

Figure 4.4: a) Sensor in the non-blood solution, b) Taking readings of the samples in a black box environment

The microcontroller was connected to the computer using the serial communication port using a USB to TTL converter device. The RX and TX pins of the USB to TTL converter are alternately connected to the microcontroller's TX and RX pin for capturing the data to the computer transmitted by the microcontroller. For each sample, around 20 reflected sample data of the two specific wavelengths of each set of LEDs were stored. So, for the first set of LEDs, 200

sample data for 10 blood samples and 500 sample data for 25 non-blood samples for each set of LEDs were collected.

4.2 Method 2: Optical Spectral Sensor in Visible Regions

In order to understand the core of this method, spectrometry needs to be clarified. The light that the eyes see is, in fact, a part of what the sun sends and it travels within a broad frequency range called a spectrum. Human eyes can only sense a tiny part of the solar spectrum which is known as a visible spectrum or visible light. By chance, the visible spectrum is precisely the sub-set of high-intensity frequencies. In fact, as an outcome of the evolution of human eyes for thousands of years, the eyes can only detect the most significant portion of the spectrum of light that hits the planet.

When light strikes on an object and it is not completely absorbed, it falls apart and it is an optical phenomenon called reflection. If this reflected light is in the range of human's sight and the eye absorbs some of the visible spectrum wavelengths, only then the associated color of that spectrum can be seen. So, when someone sees a green leaf, it means that the leaf absorbed every wavelength except the green component of electromagnetic radiation with a wavelength of about 520nm.

4.2.1. Working Principle and Developed Sensor System of Method 2

The spectral detector works in the visible light area on the basis of the six wavelengths of light. The SparkFun AS7262 visible spectral sensor breakout introduces spectroscopy like the same way as a spectrophotometer [56], making it simpler to assess and characterize how distinct materials absorb and reflect distinct light wavelengths than ever before. It is a cost-effective multi-spectral sensor-on-chip solution designed for spectral ID applications. This highly integrated

device provides 6-channel multi-spectral sensing of approximately 430nm to 670nm in visible wavelengths with a full-width half-max (FWHM) of 40nm that can be roughly translated into red, orange, yellow, green, blue, and violet color. The AS7262 uses deposited interference filter technology to integrate Gaussian filters into conventional CMOS silicon and is packaged in a land grid array (LGA) package [57]. The impulse response of the filter is a Gaussian function and removes the overshoot to a step function input while minimizing the rise and fall time. This package offers an integrated aperture to regulate the light that enters the range of sensors. Access to control and spectra information is introduced either through the I2C register set or through a serial UART with a high-level AT spectral set. The board also has several methods to illuminate items that will attempt to evaluate for more precise reading of the spectroscopy. There is an onboard LED specifically selected to obtain more precise spectral measurements and also to evaluate surface reflectivity. The 6-channels can either be read as raw 16-bit values or as calibrated floating-point values via the I2C bus. There is also a temperature sensor on-board that can be used to read the chip's temperature. The 6 spectral sensor-module (AS7262) from SparkFun [94] was used to conduct the initial research which is shown in Fig 4.5.

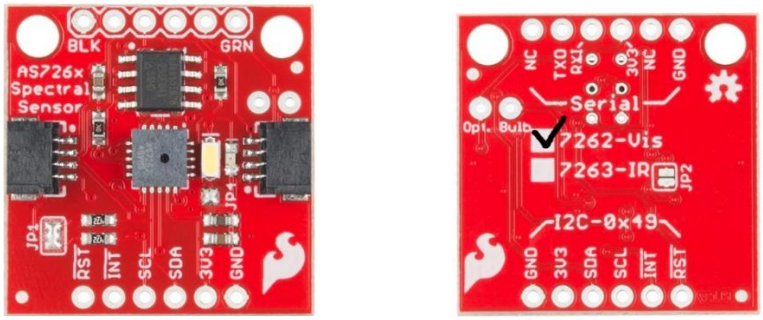


Figure 4.5: Backside and the front side of AS7262 sensor [94]

For the initial testing of the second method, the experiment was conducted using the SparkFun’s AS7262 sensor module. The sensor was connected to a multiplexer which was connected to the Raspberry Pi. Python programming language was used to read the 6 spectral data

from the sensor. The 6 spectral data from each of the blood and non-blood samples were then stored in a CSV file for further analysis. Fig. 4.6 shows the complete arrangement for conducting the experiment with the spectral sensor.

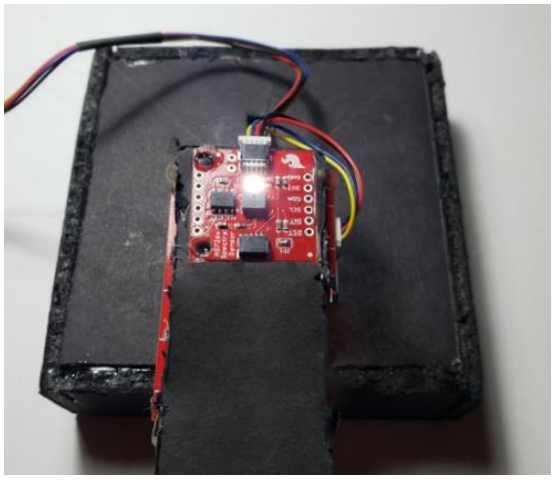


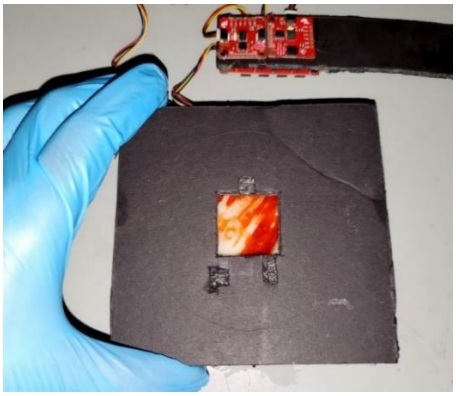
Figure 4.6: Developed bleeding detection system using AS7262 module

4.2.2. The Procedure of Data Collections for Method 2

In order to test the second method, blood and non-blood samples were kept in a plastic petri dish. Then the whole petri dish was covered using a dark box in order to make sure no other spectral wavelength was present at the time of the experiment. Only a certain portion of the black box is kept open in order to insert the AS7262 sensor module and get the reflected light values of the 6 spectral wavelengths.

(a) (b)

(c)



(d)

Figure 4.7: Experimental setup a) blood sample b) non-blood sample c) pig's intestine mixed with non-blood sample d) complete dark box arrangement for collecting accurate data

In this setup, multiple AS7262 modules can also be connected and can be controlled separately as well. Fig. 4.8 shows the data collection procedures using the AS7262 module. Using this experimental arrangement, the data can be easily stored in the computer for further processing. A complete dark box environment was ensured at the time of collecting the data to collect accurate reflected spectral data using the system. The AS7262 module's white LED was also kept on in order to get accurate data in the dark environment.

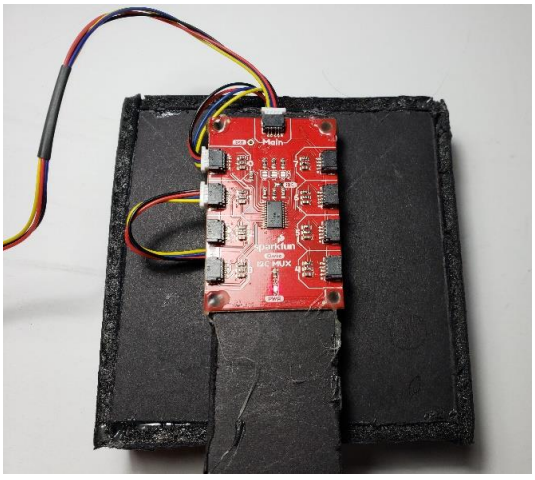


Figure 4.8: Data collection using AS7262 module while sensor facing towards surface of the samples in a dark box environment

4.3 Method 3: Optical Spectral Sensor in Near-infrared Regions

The near-infrared (NIR) optical spectroscopy provides an easy analytical method for characterizing materials in various applications. For the initial testing of the third method, the experiment was conducted using the SparkFun's AS7263 [73] sensor module. The AS7263 is a digital 6-channel multi-spectral sensor for spectral identification in the NIR light wavelengths. AS7263 consists of 6 independent optical filters whose spectral response is defined in the NIR wavelengths from approximately 600nm to 870nm with FWHM of 20nm. An integrated LED driver with programmable current is also provided for electronic shutter applications.

Through nano-optic embedded interference filter technology, the AS7263 incorporates Gaussian filter in a typical CMOS silicon and has been bundled with an LGA package, offering an optimized illumination to handle the light entering the sensor screen. Controlling and spectral access to data is carried out either via the I2C register set or via the UART serial using a spectral high-level AT command set [58].

The working method is similar to the method shown in section 4.2 as it uses the same hardware design components with an exception only to the spectral chip which is AS7263 rather than the AS7262. The overall developed sensor system using both AS7262 and AS7263 is shown in Fig 4.9

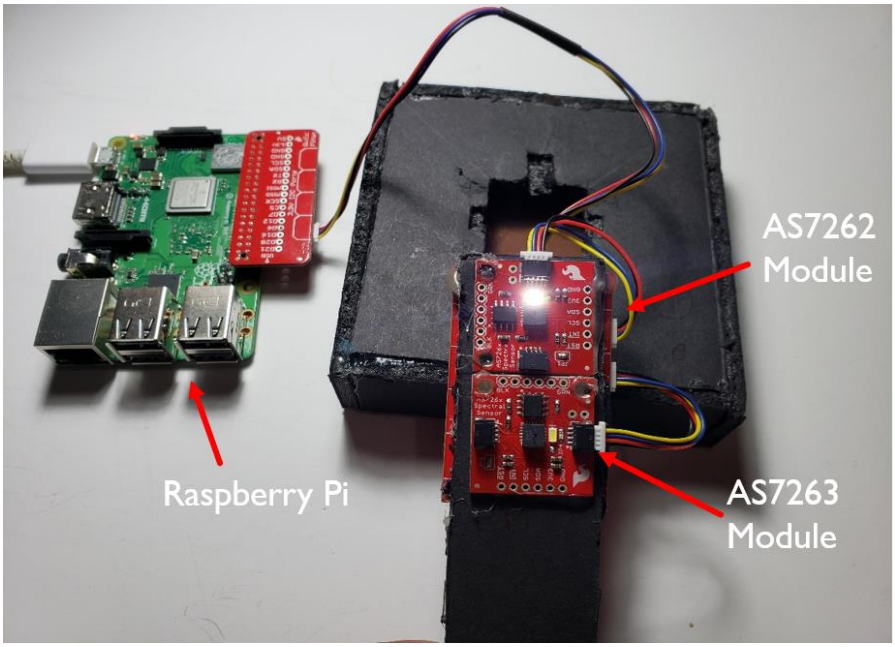


Figure 4.9: Developed bleeding detection system using AS7262 and AS7263 modules and Raspberry Pi

CHAPTER 5

AUTOMATED FEATURE SELECTION

5.1 Overview of the Features of the Methods

The aim of this research is to reduce the number of wavelengths to a minimum for the optimal hardware design of the bleeding sensor. In this respect, the feature selection methods mentioned in section 5.2 can be quite supportive. It helps considering only specified LEDs associated with the most significant wavelengths in the final prototype design and retains the sensor accuracy in a reasonable range. Therefore, the bulky AS7262 or AS7263 modules can be substantially reduced to a few LEDS and can be incorporated easily into the WCE system.

In machine learning, features are the characteristics of a problem on which the model predicts results. Features typically are numerical. But, in syntactic pattern recognition structural features such as strings and graphs are used. For example, the selling price of a house can be predicted by the size of the house feature. Features are a column of input data which are also called attributes or dimensions. There may be several features tagging a particular problem set. In order to improve the exactness of the model, it is necessary to choose the features that are more applicable to the issue. It also decreases model's uncertainty as the least important function details can be removed.

Feature selection is one of the main pre-processing phases of developing a machine learning model. Different algorithms can be used to achieve the most important features. The features used for training the models in this research are mainly the reflectance value of light from the samples at different wavelengths. The features which are used to build the models for the three different methods are shown in Table 5.1

Table 5.1: Overview of the features of different methods

Method Type	No of Features	Features	Wavelength Region	Blood Samples	Non-blood Samples	Data per Samples
Method 1 (with food colors)	2	Ratio and the average value of the reflected light for 700nm and 625nm wavelength	Visible	10	25	20
Method 2 (with food colors)	6	Reflected light value of 450nm, 500nm, 550nm, 570nm, 600nm, 650nm wavelength	Visible	10	25	20
Method 3 (with food colors)	6	Reflected light value of 610nm, 680nm, 730nm, 760nm, 810nm, 860nm wavelength	Near infrared	10	25	20
Method 2 (with food colors and pig's intestine)	6	Reflected light value of 450nm, 500nm, 550nm, 570nm, 600nm, 650nm wavelength	Visible	22	32	20
Method 3 (with food colors and pig's intestine)	6	Reflected light value of 610nm, 680nm, 730nm, 760nm, 810nm, 860nm wavelength	Near infrared	22	32	20

From Table 5.1, it can be seen that, there are two types of datasets-

a) **First Dataset**: 25 non-blood food color samples and 10 hemoglobin blood samples. There are total 35 blood and non-blood samples. Total observations are 700 since 20 observations per sample.

b) **Second Dataset**: 32 non-blood food color samples and 22 hemoglobin blood samples with pig's intestine. The total blood and non-blood samples are 54. So, the total observations are 1080 since 20 observations per sample.

5.2 Feature Selection Methods

The selection of features is one of the key concepts in machine learning that significantly impacts on the model's efficiency. The data features used to train machine learning models affect the results immensely. Irrelevant or somewhat less relevant features may affect the efficiency of the model adversely. The first and most important step of every model design must therefore be the feature selection and the data cleaning.

Selection of features requires automatic or manual choosing of features that are most relevant to the predicted component or desired results. When the data set contains irrelevant features, models can learn from irrelevant features and be less precise.

A strong understanding of feature selection is required in order to achieve greater precision for machine learning models. A good grasp of these strategies leads to better performing models. The clear knowledge of the underlying structure and data properties contributes to a better understanding of the algorithms. There are two main reasons why function selection is performed before data modeling:

1. Reduction of features, reduces overfitting and improves model generalization. There will be fewer opportunities to take noise-based decisions with fewer redundant data.
2. To better understand the characteristics and their relation to response variables. It also decreases training times, since less data points minimize the difficulty of algorithms and train algorithms quicker.

Such two objectives are often at odds and therefore require other approaches: based on the data at hand, sometimes a correct form of selection for the first reason may not be necessarily appropriate for the second and vice versa. But what also appears to be happening is that people are

indiscriminately utilizing their preferred approach or what is more easily accessible from their preference device, in specific methods more appropriate for the first reason to achieve the goal of the second's.

In the study, some of the most common approaches for selecting features are applied and observed side-by-side on the sample dataset. According to Table 5.1, there are only 2 features in method 1. So, feature selection method was not used for method 1. But in the case of method 2 and method 3, there are 12 features in total with 6 features each. So, in order to find the appropriate features, three feature selection methods are used- a) Univariate feature selection, b) Random forest feature selection, c) Principal component analysis (PCA) feature selection.

5.2.1. Univariate Feature Selection

Univariate feature selection independently tests each feature to assess the intensity of the feature's relationship with the response factor. These methods are easy to use and understand and are particularly suitable for better comprehension of data. Statistical measures can be used to pick the most important features of the input function by comparing each feature to the target variable.

When the relation between one feature and the response variable is evaluated, the other features are overlooked, which is the explanation why it is referred to as "univariate". Every feature has its own test score. Eventually all test scores are measured and features that have the highest scores have been chosen. The Python Scikit-learn presents feature selections as objects to execute the process of transformation. According to [48], there are 3 basic selection objects in Scikit-learn: a) SelectKBest, b) SelectPercentile and c) GenericUnivariateSelect.

In 'SelectKBest' method, only the highest scoring features are selected and the rest of the features are ignored. In case of 'SelectPercentile', it eliminates almost all the popular predictive univariate tests on each feature except the maximum user-defined scoring rate such as false

positive, false rate of discovery or family-wise errors. However, the ‘GenericUnivariateSelect’ offers a configurable filtering technique, enabling a hyper-parameter search estimator to select the best univariate selection strategy.

Such objects have a scoring function as feedback that returns univariate p-values (or just SelectKBest and SelectPercentile scores). The scoring functions for regression are f_regression, mutual_info_regression and scoring functions for classification are chi2, f_classif, mutual_info_classif.

In this research, blood and non-blood are needed to be classified separately. So, according to the discussion, ‘chi2’ scoring function is used with ‘SelectKBest’ univariate selection method.

5.2.2. Random Forest Feature Selection

One of the most popular machine learning algorithms is the random forest. It is so effective because it usually has good predictive quality, low overfit and is easy to interpret. The fact that, it is simple to derive the value of every parameter on the tree decision because of this interpretability. In other terms, the contribution of each parameter to the decision can be easily calculated by this method.

Feature selection using the random forest comes under the category of embedded methods. Embedded methods combine the qualities of filter and wrapper methods. They are implemented by algorithms that have their own built-in feature selection methods. Some of the benefits of embedded methods are:

- Highly accurate.
- Easy to generalize.
- Easy to interpret.

Random forests consist of 4 to 12 hundred decision trees, each of them built over a random extraction of the observations from the dataset and a random extraction of the features. Not every tree sees all the features or all the observations, and this guarantees that the trees are de-correlated and therefore less prone to over-fitting. Each tree is also a sequence of yes-no questions based on a single or combination of features. At each node, the tree divides the dataset into 2 branches, each of them hosting observations that are more similar among themselves and different from the ones in the other branch. Therefore, the importance of each feature is derived from how "pure" each of the branch is.

For classification, the measure of impurity is either the 'Gini' impurity or the information gain/entropy and for regression the measure of impurity is variance. Therefore, when training a tree, it is possible to compute how much each feature decreases the impurity. The more a feature decreases the impurity, the more important the feature is. In random forests, the impurity decrease from each feature can be averaged across trees to determine the final importance of the variable. To give a better intuition, features that are selected at the top of the trees are in general more important than features that are selected at the end nodes of the trees, as the top splits lead to bigger information gains.

5.2.3. Principal Component Analysis (PCA) Feature Selection

PCA is a linear transformation technique, mostly used for feature extraction and dimensional reduction. It enables the simplification of high-dimensional data while preserving patterns and trends by converting the details into less dimensional feature summaries. Other popular applications of PCA include exploratory data analyzes and signal de-noise in equity trading, and genome-data analysis and bioinformatic gene expression levels. It allows us to identify patterns on the basis of feature correlation in the dataset. It is unsupervised and a clustering-like

learning method [34]. It always shows patterns with no previous knowledge as to whether the samples originate in various treatment groups or whether they are phenotypical. In short, PCA attempts to detect maximum variance in high dimensional data and projects them in a new subspace with the same or less sizes than the original one. It is often used to decrease the dimensionality of a large data set in order to make machine learning more practical where the original data is inherently large in size such as in image recognition. This converts a sequence of observations of potentially correlated variables into values of uncorrelated variables called principal components using an orthogonal transformation. It emphasizes on variation by bringing out strong patterns in a dataset which makes data easy to explore and visualize as well.

5.3 Selection of Appropriate Features from the Samples

There are often completely irrelevant, trivial and unimportant features in a high-dimensional dataset. In contrast to the essential characteristics, the role of this category of features has always been less significant for predictive modelling. They could also have null contributions. Such characteristics cause a number of issues such as the excessive allocation of resources to additional features, which is a disturbance for which the design is extremely poor and which in effect inhibits effective modelling.

In this research, there are six features of reflected light values consisting of 450nm, 500nm, 550nm, 570nm, 600nm, 650nm wavelength for the 2^{nd} method and six features of reflected light values consisting of 610nm, 680nm, 730nm, 760nm, 810nm and 860nm wavelength for the 3^{rd} method. But all of the 12 features may not be equally contributing to the greater performance of the model. So, it is better to find out the most significant features from the dataset which will be effective in building the model and will enhance the performance of a machine learning model.

5.3.1. Features Selection using Univariate Feature Selection

According to Table 5.1, it can be seen that the datasets correspond to classification tasks on which the model needs to predict blood samples based on total 12 features for method 2 and method 3. In order to select the appropriate features for method 2 and method 3, the second dataset was used which had blood samples and non-blood samples with pig's intestine.

There was a total of 1080 observations in the second dataset. The first task was to load the dataset. The 6 different features labeled into the outcomes of 1 and 0, where 1 stands for the blood, and 0 denotes for non-blood. The dataset was checked for missing values before starting the selection. First, for the "univariate feature selection", a Chi-Squared statistical test for non-negative features will be used to select two of the best features from the dataset. The Scikit-learn library provides the SelectKBest class that can be used with a suite of scoring function to select a specific number of features, in this case, the scoring function is Chi-Squared. Fig. 5.1 and 5.2 show the scores for each attribute and the 2 attributes chosen- 500nm and 550nm for the 2nd method and 610nm for the 3rd method. These scores will help further in determining the best features for training the model.

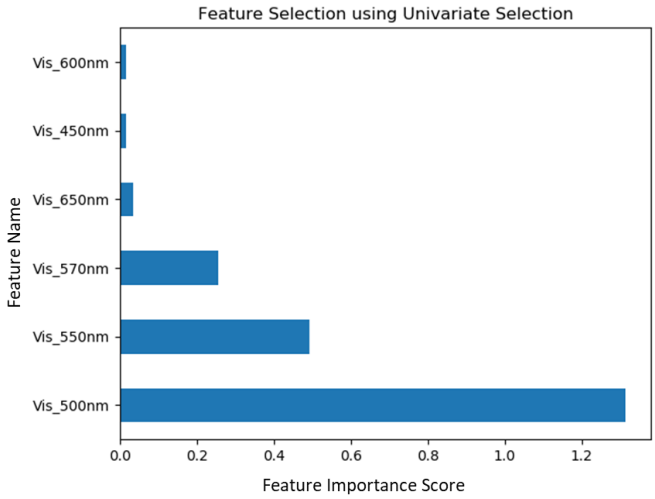


Figure 5.1: Feature selection using univariate selection for method 2

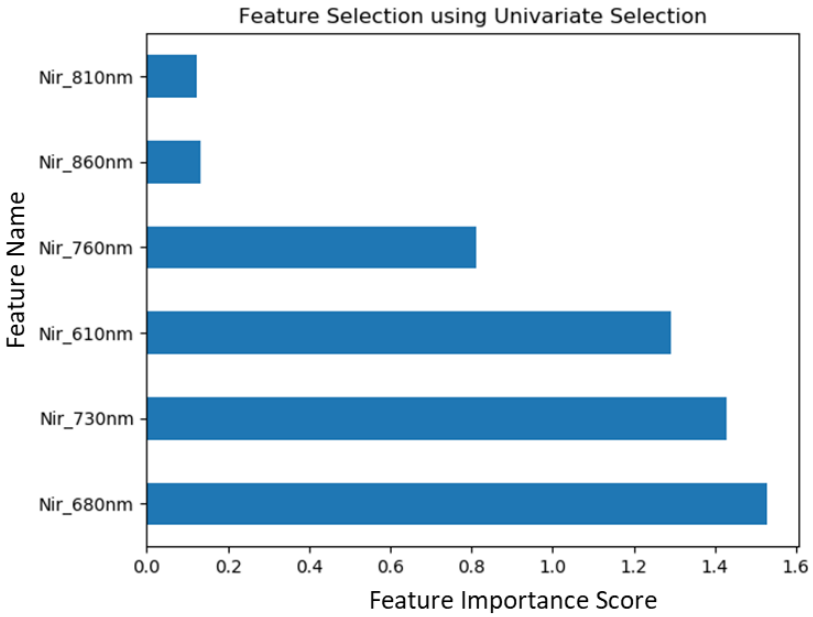


Figure 5.2: Feature selection using univariate selection for method 3

5.3.2. Features Selection using Random Forest Feature Selection

During training time of random forest, it considers different decision trees. The actual prediction is taken into account after analyzing all the probabilities from those trees. The features can be found by observing the decrease in accuracy. If any of the features are removed and the accuracy increased, then the rest of the features are important than the removed one. In practice, it is not removed from the model rather it randomly permutes the feature and measures the decrease in accuracy on an out of bag sample. There is actually a second way of computing feature importance often called “Gini importance”. Each function value is computed as a sum of the number of splits that contain the element, proportionally to the amount of the samples it divides, according to the Gini importance. In the case of a random forest regressor, the importance is assessed by measuring the reduction in node impurity by the probability to reach that node. The probability of a certain node can be computed according to the number of samples reaching the node divided by the total number of samples. The feature with higher probability is more

significant than the others. The method was implemented by Python Scikit-learn library to find the best feature [36].

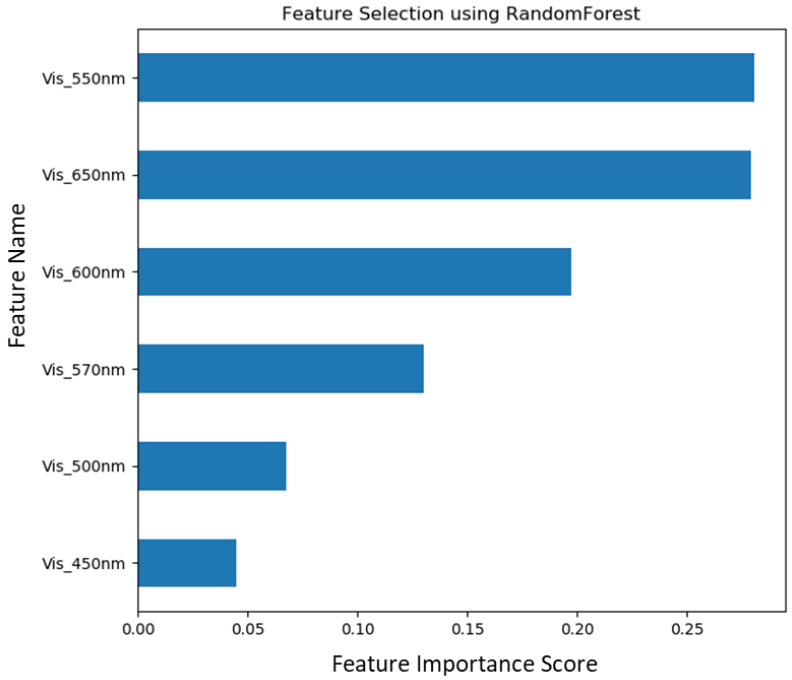


Figure 5.3: Feature selection using random forest for method 2

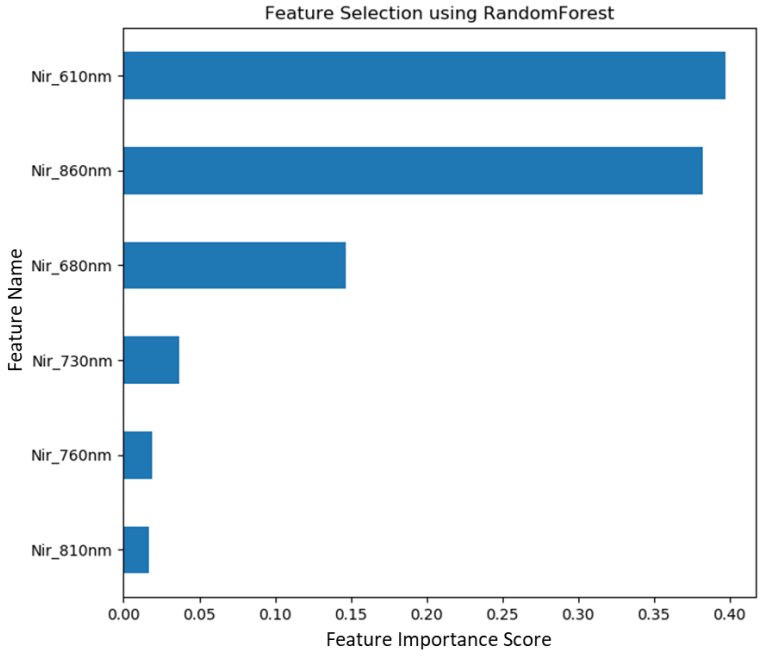


Figure 5.4: Feature selection using random forest for method 3

From Fig. 5.3 and 5.4, it is clear that the highest score is 550nm for 2nd method and 610nm for the 3rd method.

5.3.3. Features Selection using PCA Feature Selection

PCA uses linear algebra to transform the dataset into a compressed form. Generally, this is called a data reduction technique. A property of PCA is that it can automatically choose the number of dimensions or principal components in the transformed result. In this case, PCA will be used to guess 2 principal components. The python Scikit-learn library has PCA class to implement the feature selection. PCA uses variance-ratio to get relevant features. Which tells how much information (variance) can be attributed to each of the principal components. By using the attribute "explained_variance_ratio_", From Fig. 5.5, it can be observed that the first principal component 550nm contains around 0.51% of the variance and the second principal component 500nm contains 0.86% of the variance for method 2.

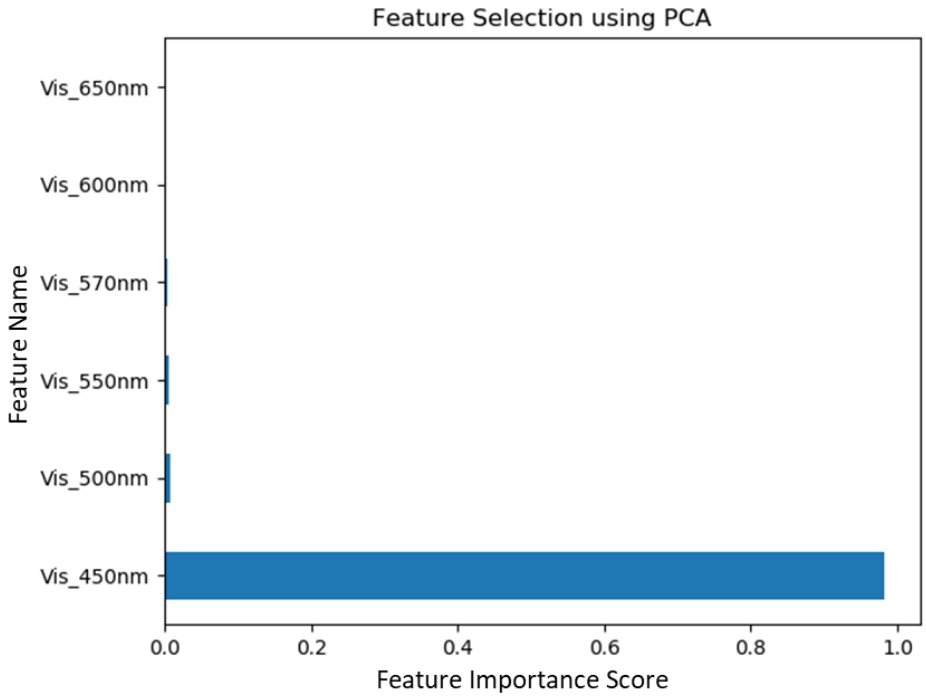


Figure 5.5: Feature selection using PCA for method 2

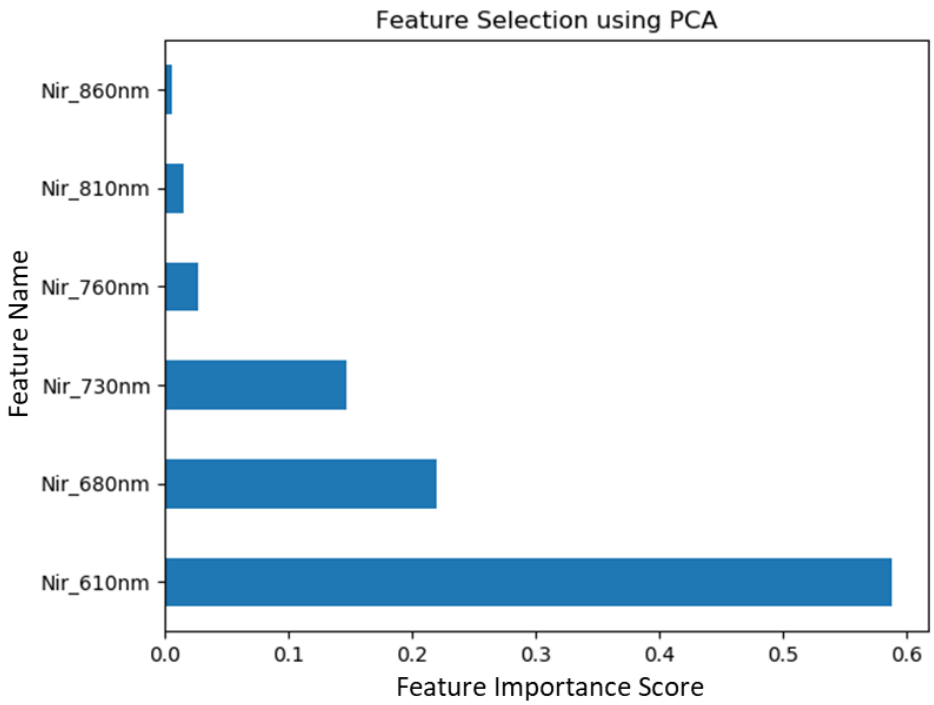


Figure 5.6: Feature selection using PCA for method 3

For method 3, from Fig. 5.6, it can also be observed that the first principal component 610nm contains around 58.81% of the variance and the second principal component 680nm contains 21.95% of the variance.

Table 5.2: Summary of the selected features

Method Type	**No of Features**	**Features**	**Univariate Feature Selection**	**Random Forest Feature Selection**	**PCA Feature Selection**	**Most Significant Features**
Method 2	6	450nm, 500nm, 550nm, 570nm, 600nm, 650 nm	**500nm, 550nm,** 570nm	**550nm**, 650nm, 600nm, **500nm**	450nm, **500nm, 550nm**	**500nm, 550nm**
Method 3	6	610nm, 680nm, 730nm, 760nm, 810nm, 860nm	680nm, 730nm, **610nm**	**610nm,** 860nm, 680nm	**610nm,** 680nm, 730nm	**610nm**

The significant features were used to train the final model for this research. From Table 5.2, it can be seen that the most significant features are 500nm and 550nm for the 2nd method and 610nm for the 3rd method. Table 5.3 shows the final two methods with the most significant features for building machine learning models.

Table 5.3: Summary of the methods with selected features

Method Type	**No of Features**	**Most Significant Features**
Method 1	2	ARL value and average value for (700nm, 625nm)
Method 2 and 3 Combined	3	500nm, 550nm, 610nm

There are mainly 3 different hardware setups in this research. In method 1 hardware setup, two LEDs, an 8-bit microcontroller and photodiode were used. For the method 2 and method 3 hardware setups, Raspberry Pi was used as the processing unit and AS7262 spectral sensor in visible wavelength range for method 2 and AS7263 spectral sensor in near infrared wavelength range for method 3 were used as the reflected light detectors for the blood and non-blood samples.

Table 5.4: Summary of the experiments

Experiment Type	Method Type	Dataset	No of Features	Most Significant Features	Total Samples
Experiment 1	Method 1	Food colors and hemoglobin samples	2	ARL value and average value for (700nm, 625nm)	25 non-blood samples, 10 blood samples
Experiment 2	Method 2 and 3 combined	Food colors and hemoglobin samples	3	500nm, 550nm, 610nm	25 non-blood samples, 10 blood samples
Experiment 3	Method 2 and 3 combined	Food colors and hemoglobin samples with pig's intestine	3	500nm, 550nm, 610nm	32 non-blood samples, 22 blood samples

According to Table 5.3, the three hardware setups were then divided into two methods based on the number of features considered for the final experiments. From Table 5.1, it can be observed that, there are two different datasets- one with 25 non-blood food color samples and 10 hemoglobin samples whereas the second one with 32 non-blood color sample with pig's intestine and 22 hemoglobin blood sample along with pig's intestine. Therefore, the final experiments for the method 1 and combined method 2 and 3 were analyzed on both of these two datasets. The summary of the final three experiments were shown in Table 5.4.

CHAPTER 6

MACHINE LEARNING MODELS AND THEIR PERFORMANCE EVALUATION

6.1 Proposed Framework

Dataset is taken as input to the classification methods followed by performance evaluation as depicted in Fig. 6.1. Four types of classification methods are used to analysis the results. The experiments are done by using Python and Scikit-learn library.

The training process of a model is the process of feeding data into a neural network and letting it learn the patterns of the data by itself. The training process takes in the data and pulls out the best features of the dataset. During the training process for a supervised classification task, the network is passed both the input features and the response labels of the training data. However, during testing, the network is only fed with input features.

In the testing process, the learnt patterns of that network are tested. The features are passed to the network, and the network then predict the proper response labels. The data for the network is divided into training and testing sets. The output classifier or model can not be evaluated with the same training dataset since the it has already learnt the characteristics of that dataset during its training process. So, it would be extremely biased if it is being tested against the same dataset for further analysis.

Therefore, it is necessary to distinguish the training and the testing data from the dataset in order to train and evaluate the model. But during the development phase, it may not be always possible to have so many details. The obvious solution in such situation is to split the dataset into two sets, one for training and the other for testing; and it has to be done before starting the model training. It is also required to make sure splitting the data should be in a proper manner**.** In order

to achieve that task, for this research, the whole datasets were divided into training and testing datasets. The testing datasets were kept fully separated. The training datasets had no influence of the testing datasets. The models were trained in such a way that it could provide accurate validation results for the fully separated testing datasets.

It is important to compare the performance of multiple machine learning algorithms consistently. The key to a fair comparison of machine learning algorithms is ensuring that each algorithm is evaluated in the same way on the same data. In order to ensure that, each algorithm will be evaluated on a consistent test harness.

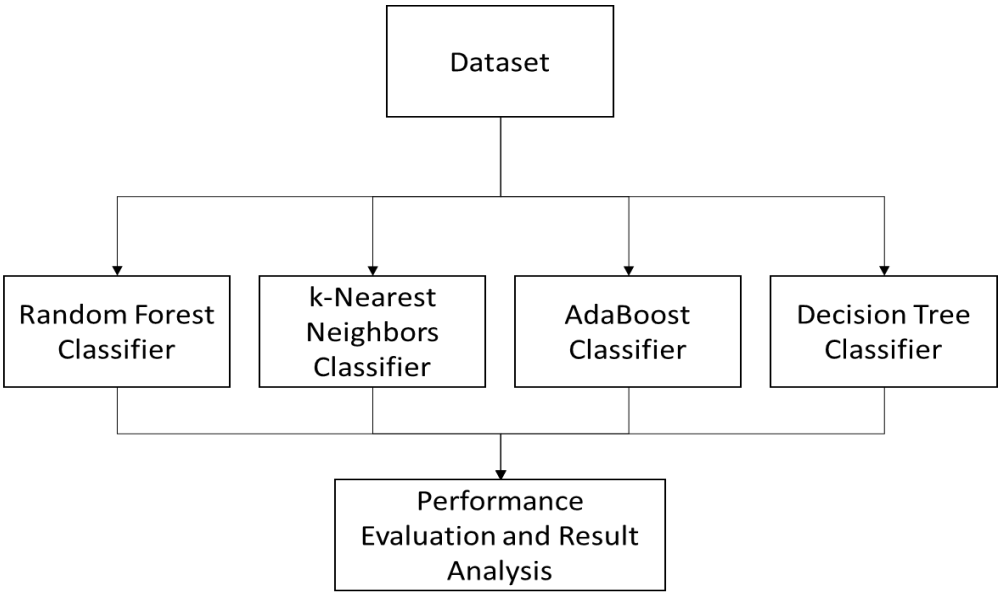


Figure 6.1: Proposed experimental framework

Four different algorithms were used for comparison in this research:

- k-Nearest Neighbors Classifier
- Random Forest Classifier
- AdaBoost Classifier
- Decision Tree Classifier

These four machine learning algorithms are very handy, popular and important, particularly for text analytics and general classification. Some of the algorithms are discussed below-

6.1.1. k-Nearest Neighbors Classifier

k-Nearest Neighbors (k-NN) is one of the best machine-learning algorithms that is very basic, flexible, easy to understand and scalable. It has succeeded in a large number of classification and regression problems, despite its simplicity. It is in the area of supervised learning and is widely applicable to the analysis of patterns, data mining, intrusion detection, economics, safety, political science, handwriting, video recognition, etc. It is appropriate for statistical problems of both classification and regression. It is primarily used in the industry for statistical classification problems, though.

It is an algorithm for non-parametrical and lazy learning. Non-parametric means for the underlying assumptions on data representation, there is no inference. In other words, the model structure is calculated from the dataset. In its training phase it uses the entire dataset. Whenever a prediction is made for an unseen data case, it scans for k-most like instances through the whole training data and eventually returns the decision with the most similar instance. It implies that when k is close to 1, the class label is given to the nearest neighbor. Nevertheless, if K is 5, the algorithms choose the next five most closely related data points and categories them according to groups represented by the corresponding five items, by a majority vote.

k-NN is often used in search applications where related objects are being searched, such as finding similar items. Algorithms say that object is one of them, if it is close to its neighbors. Like the same way, apple is most likely a fruit, for instance, when apple is more similar to peach, pear, and cherry (fruit) than monkey, cats or rats (animals).

6.1.1.1. Steps of k-Nearest Neighbors Classifier

The k-NN algorithm utilizes 'feature similarities' to determine the values in new datasets, ensuring the new data point will be given a value based on how well that correlates with the points in the training set. For developing or building a classifier, the process of data collection, preparation, and then testing is common in the entire machinery learning area. A dataset is necessary to implement every algorithm. Therefore, the training and the test data must be prepared during the first phase of k-NN. The algorithm's other steps can be explained as below-

a) Select an integer number K of the nearest neighbors
b) Calculate the Euclidean distance of K number of neighbors between test data and each row of the training data.
c) Take the K nearest neighbors according to the calculated Euclidean distance and sort them in ascending order.
d) Count the number of the data points in each category among these K neighbors
e) The category for which the neighbor's number is highest is applied to the new data points.

6.1.2. Random Forest Classifier

Random forest is a type of supervised, ensemble-based learning algorithm. Ensembled algorithms combine more than one same or different type of algorithms for the classification of object. There are many individual decision-making trees created on a randomly divided dataset. The collection of decision trees is also known as the forest. Every individual forest tree in a random forest provides a class prediction and the class with the most votes becomes the final prediction of the model which is shown in Fig. 6.2.

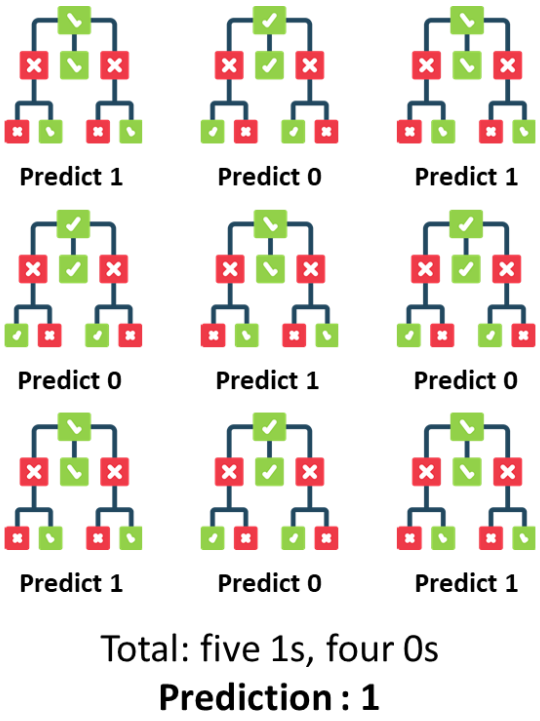


Figure 6.2: Random forest classifier's prediction

For classification and regression, which provides an estimation of the relations between a dependent variable and one or more independent variables, random forests can be used. It includes various applications, such as recommendation engines, classification of images and selection of features. It is ideal for classifying trustworthy lenders, detecting fraudulent activity and forecasting diseases.

For almost every predictive issue (even non-linear ones), random forests classifier is a good choice. It is a rather new learning machine strategy that was developed in the 90's by Bell Labs. Compared to other nonlinear classification algorithms, it is easier and more powerful.

6.1.2.1. Steps of Random Forest Classifier

Because of the number of decision-makers involved in the process, random forests classifier is recognized as an exceptionally precise and stable tool. The overfitting

problem does not impact it since all forecasts were made on average, which cancels out the biases. The classifier works as follows-

a) Pick random dataset samples.

b) Build a sample decision tree and extract the outcome of each decision tree.

c) Take out a vote on each outcome that is predicted.

d) Pick the prediction outcome as the final prediction with the most votes

6.2 Performance Matrices

Performance evaluation in a classification problem includes determining the classification performance of the system. A binary classifier predicts all data instances of a test dataset as either positive or negative. This classification (or prediction) produces four outcomes –

1. True positive (TP): correct positive prediction
2. False positive (FP): incorrect positive prediction
3. True negative (TN): correct negative prediction
4. False negative (FN): incorrect negative prediction

Sensitivity/recall, specificity, precision, and accuracy are the most ideal criteria for measuring classification performance. In order to analyze the performance of a classifier model, these matrices are a must to understand.

Sensitivity (SN) or recall (REC) is the possibility that representatives of the successful categories are appropriately identified. It is determined by dividing the right number of positive expectations by the total number of positive predictions. The highest sensitiveness is 1.0, while the worst response is 0.0 which also indicates how appropriate the model in determining positive predictions.

$$recall(REC) = sensitivity\ (SN) = \frac{TP}{TP+FN} \tag{6.1}$$

Probability of accurately identifying negative situations is specificity (SP). SP is calculated by dividing the amount of correct negative predictions by total negative numbers. It is also referred to as a true negative rate. The best feature is 1.0 and the worst 0.0 respectively. This explains how well the model can eliminate false alarms.

$$specificity\ (SP) = \frac{TN}{TN+FP} \tag{6.2}$$

Precision (PREC) decides the successful labeling quality. Precision quantifies the number of predictions in positive category that really belong to the positive class. The correct number of positive predictions divided by the overall positive predictions is considered as the precision. The highest accuracy is 1.0, while the worst is 0.0 and also indicates how many positive classifications are relevant.

$$precision\ (PREC) = \frac{TP}{TP+FP} \tag{6.3}$$

Accuracy (ACC) is the likelihood that accurately mark all positive and negative events. It is measured as the amount of all correct observations divided by the total dataset size. The highest accuracy is 1.0, while the worst is 0.0.

$$accuracy\ (ACC) = \frac{TP+TN}{TP+TN+FN+FP} \tag{6.4}$$

Almost all bleeding detection methods have utilized the aforementioned metrics for system evaluation. However, all the methods calculate these measures using their specified classifier models.

6.3 Development and Analysis of Models for Experiment 1

In order to choose the best feature for the classifier model of the experiment 1, the collected sample data (10 blood samples and 25 non-blood samples) were observed carefully. From the observation, it appeared that the average and the ratio of the reflected light data from the two wavelengths might be used as features to separate the samples easily from the non-blood. In the experiment, approximately 20 reflected sample data were collected for each of the 10 blood samples and 25 non-blood samples using both LED sets. After that, the ratio of the ARL of the two specific wavelength LEDs was calculated for both sets of LEDs. The cut-off point of the ratio for the 35 samples was determined by considering the mean value of the ratio of the 20 reflected sample data. 700nm and 625nm LED set were used mainly for the experiment as from Fig 3.2, it can be observed from the first LED set (475nm, 530nm) and second LED set (700, 625nm), the second LED set provides better separation between the ARL of ratio than the first LED set. So, it is wiser to choose the first LED set (700nm, 625nm) for the developed prototype to separate the blood and non-blood samples.

On the basis of the above experimental framework shown in Fig. 6.1, four classifier models were evaluated using python to carry out the binary classification task in which the blood samples were considered to be "1" while the non-blood samples were considered to be "0" in the classification. All the models of the experimental framework used the above mentioned two features to train the blood detection model and calculate the probability scores. The models were compared based on different parameters of the models such as training accuracy, testing accuracy, precision scores, recall scores which are shown in Fig 6.3 to Fig. 6.6 respectively. For Exp. 1, the k-NN classifier performs better than the other algorithms which is shown in Fig 6.4. The training accuracy for k-NN classifier was 92.50% which is lower than the all the other classifiers but the

test accuracy is higher than all the classifier which can be realized from the Fig. 6.4. Precision is another factor that indicates how many of the positively classified samples are relevant to the model. From Fig 6.5 it can be seen clearly that the k-NN classifier provides the best results out of all the algorithms. The models' sensitivity or recall score is shown in Fig 6.6 which indicates how good the model in detecting positive samples. k-NN provides a recall value of 83.33% which also seems higher or equivalent compared to the other algorithms. All the performance matrices of the algorithms can also be visualized in Table 6.1. From the table, it is clear that the k-NN classifier provides the best results for test accuracy, precision, AUC matrices. So, it can be surely concluded that the k-NN is the best algorithm for classifying blood and non-blood samples for Exp. 1.

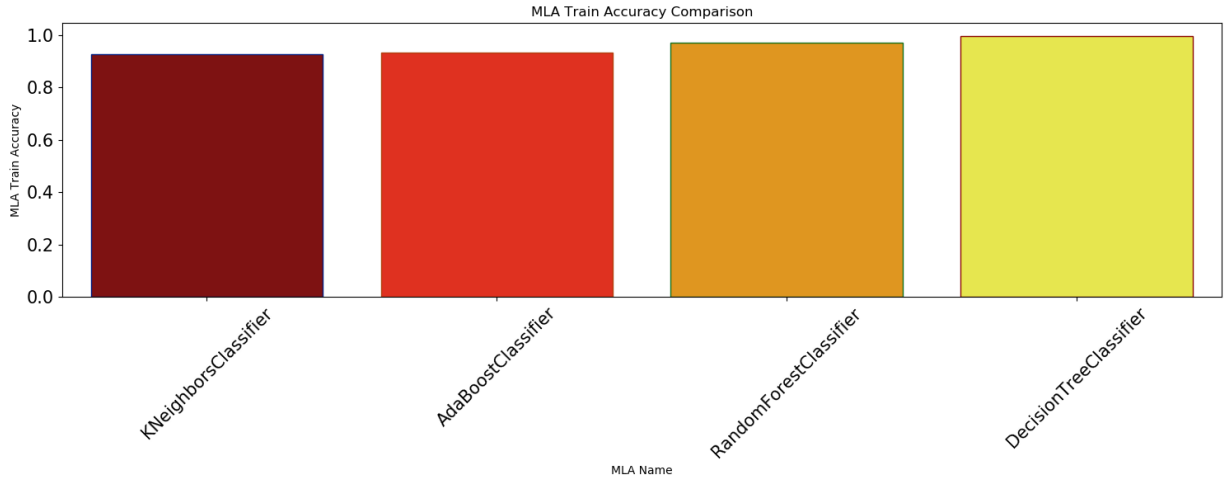


Figure 6.3: Machine learning algorithms' train accuracy comparison (Experiment 1)

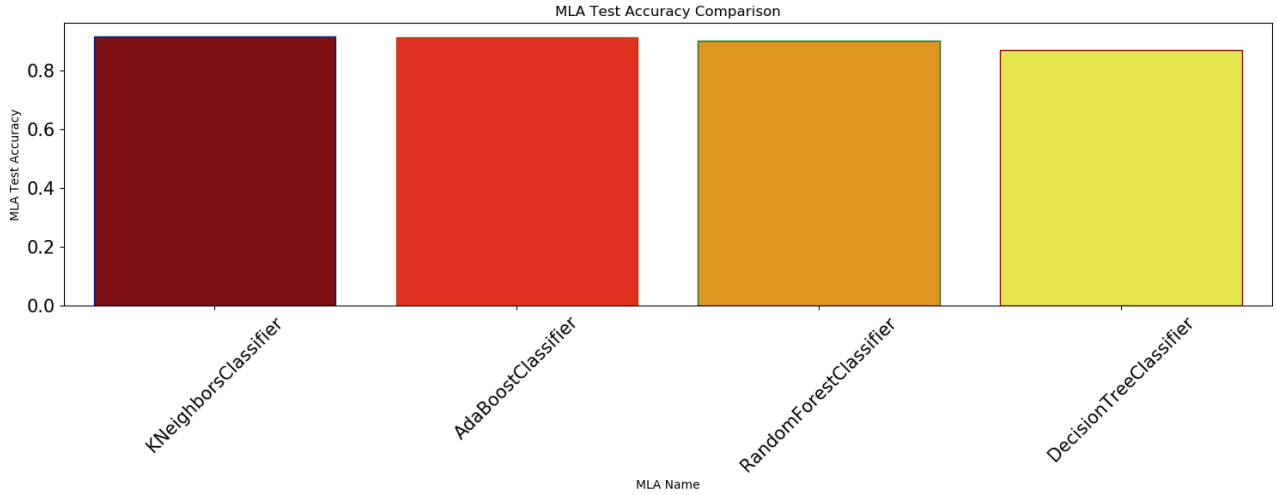


Figure 6.4: Machine learning algorithms' test accuracy comparison (Experiment 1)

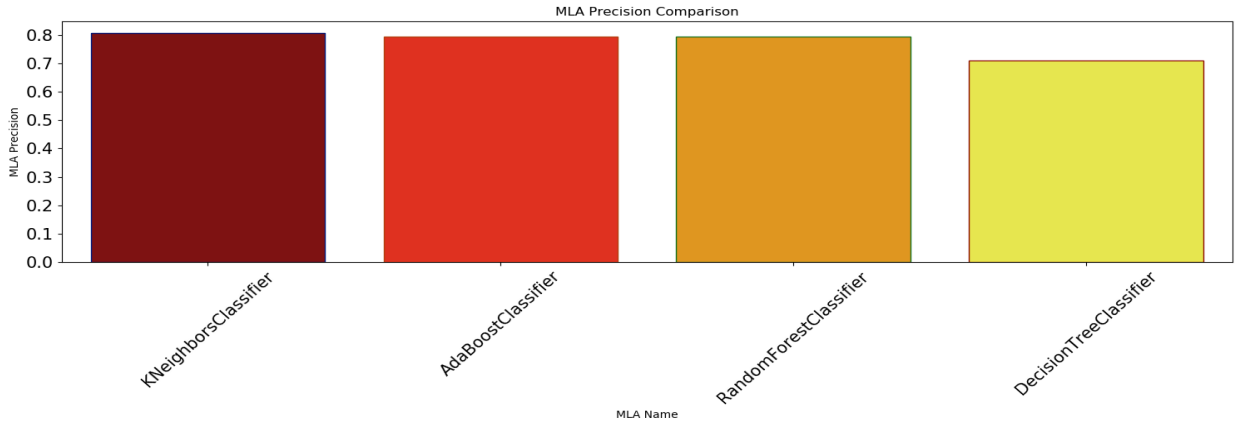


Figure 6.5: Machine learning algorithms' precision comparison (Experiment 1)

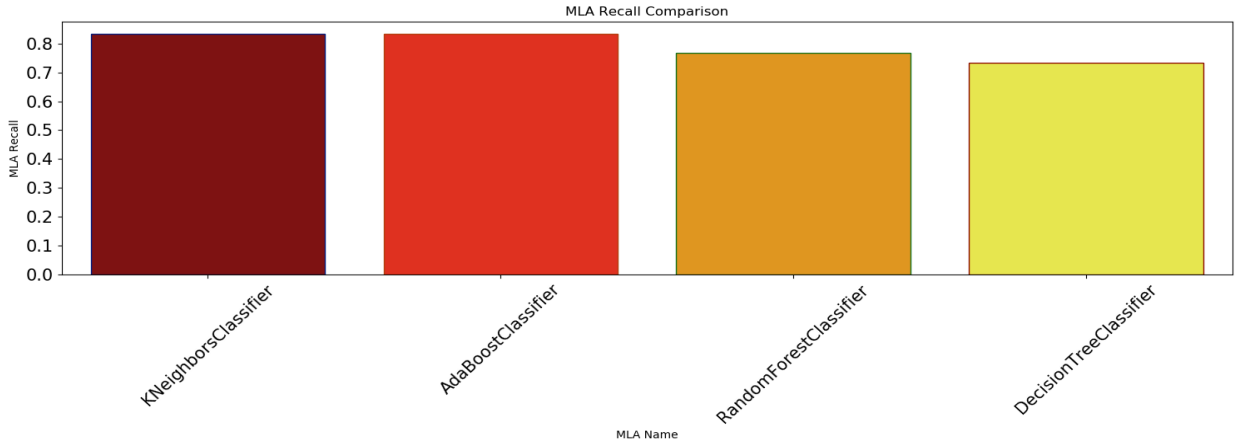


Figure 6.6: Machine learning algorithms' recall comparison (Experiment 1)

The samples were considered to be blood or non-blood samples based on the high probability calculated by the trained model. From the initial Spectrophotometer experiment, it is found that the first set of LEDs with 700nm and 625nm performs better than the second set of LEDs with 475nm and 530nm. So, in this experiment, 700nm to 625nm reflected value dataset was used to figure out the best result from the machine learning model. According to the analysis of four different models, the highest testing accuracy was found to be 91.54% for the k-NN classifier. Table 6.1 shows the full comparison between different machine learning algorithms.

Table 6.1: Analysis of machine learning algorithms' performance using different performance matrices for Exp. 1

MLA Name	MLA Train Accuracy	MLA Test Accuracy	MLA Precision	MLA Recall	MLA AUC
KNeighborsClassifier	0.925	0.9154	0.8064	0.8333	0.8866
AdaBoostClassifier	0.9341	0.9115	0.7936	0.8333	0.8841
RandomForestClassifier	0.9727	0.8962	0.7894	0.750	0.8450
DecisionTreeClassifier	0.9955	0.8692	0.7096	0.7333	0.8216

The model predicts the blood samples from the non-blood samples with good accuracy but it fails to separate at some of the samples. The model mostly failed to separate the blood samples having a low dilation of less than 1-2% and some non-blood samples which had a strong color similar to blood with higher light absorption characteristics.

The confusion matrix is very useful to understand the performance of any classifier. It gives information about predicted and actual values. Here TP is the number of positive examples that are correctly classified, TN is the number of negative examples that are correctly classified, FN is

the number of positive examples that are incorrectly classified as negative and FP is the number of negative examples that are incorrectly classified as positive. There are different types of performance matrices are used in this research (1) Overall Accuracy, (2) Specificity, (3) Recall, (4) Precision, (5) ROC curve (AUC). The specificity is the proportion of the TN and (TN+FP) and with the higher specificity fewer positive cases are labeled as negatives, so this ratio can be regarded as the percentage of negative cases correctly classified as belonging to the negative class. The proportion of cases that are TP for all the cases that are positive in the diagnostic tests (TP+FN) is called sensitivity. It ranges from 0 to 1 and an attribute that is perfectly correlated to the class provides a value of 1.

The confusion matrix is shown in Fig. 6.7 to describe the performance of the k-NN model on the test dataset. From the confusion matrix, it can be clearly understood that out of the 200 non-blood samples 188 samples are correctly predicted, and out of the 60 blood samples 50 samples are correctly predicted which gives an overall accuracy of 91.54% for the model.

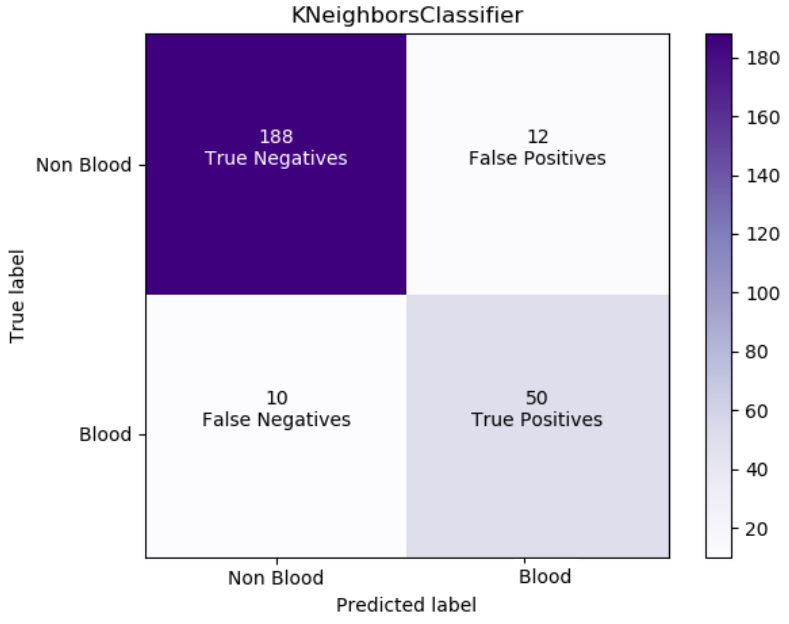


Figure 6.7: Confusion matrix of the k-NN classifier (Experiment 1)

Summary of Experiment 1:

Training data: 7 blood samples, 15 non-blood samples (20 data/sample)

7*20=140, 15*20=300. Total= 440 samples

Fully separated testing data: 3 blood samples, 10 non-blood samples (20 data/sample)

Blood=3*20=60,

Non-blood=10*20=200,

Total= 260 samples

k-NN Classifier's Performance Details:

Blood/Positive= 60 (True Pos=50, False Neg=10)

Non-blood/Negatives=200 (True Neg=188, False Pos=12)

Test accuracy=91.54%

6.4 Development and Analysis of Models for Experiment 2

The models for the second and third methods with the AS7262 and AS7263 modules are combined for the proper development of the classifier and only 3 most significant features out of the 12 features were selected. For each of the spectral wavelengths, 20 sample data were collected for the training. So, around 700 sample data (500 data values for 25 non-blood food color samples and 200 data values for 10 blood samples) were used for the training and testing. According to Table 5.3, three features - 500nm, 550nm, and 610nm wavelength will be used for the development and analysis of the models based on the experimental framework shown in Fig. 6.1. Four classifier models were evaluated using Python to carry out the binary classification task in which the blood samples were considered to be "1" while the non-blood samples were considered to be "0" in the classification. The models' training accuracy, testing accuracy, precision scores, recall scores are

shown in Fig 6.9 to Fig. 6.11 respectively. For the experiment 2, reflected light values were used as the features for developing the classifier model. Out of the four classifier models, k-NN classifier performs better than the other algorithms in the validation data. Fig 6.9 and Fig 6.10 indicate that k-NN provides a training accuracy or 97.95% and test accuracy of 89.23% which are higher than most of the algorithms. Fig 6.11 indicates k-NN also provides higher precision value which is around 0.7857. This score indicates the model's positively classified predictions are very relevant to the actual value. From Fig 6.12, it is clear that the k-NN classifier algorithm is very good at detecting positive values out of all the classifier algorithms. The overall summary of all the classifier models can be observed in Table 6.2. So, it is clear that the k-NN algorithm is the best approach to differentiate the blood and non-blood samples for the experiment 2.

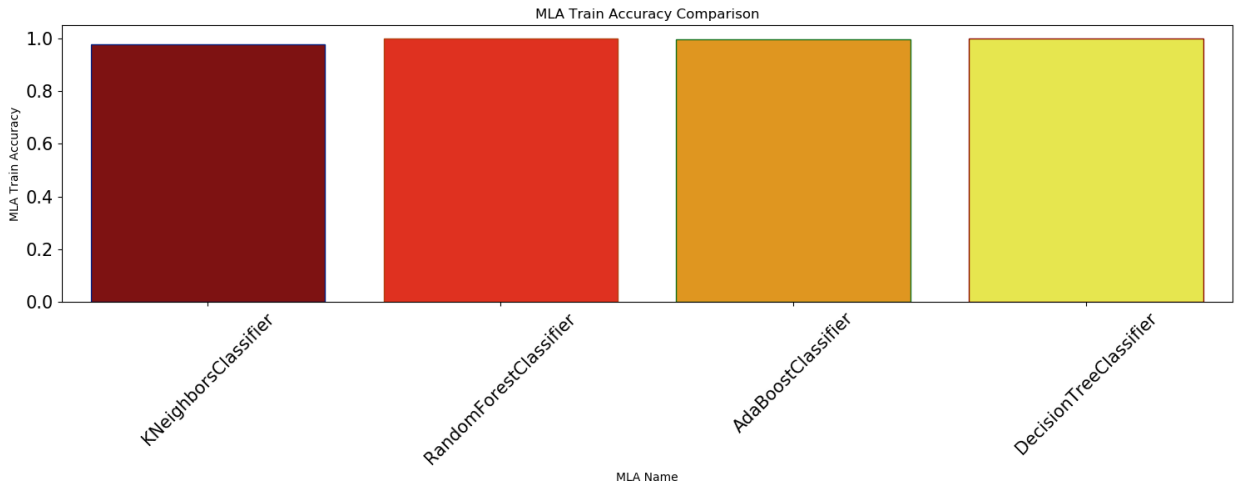


Figure 6.8: Machine learning algorithms' train accuracy comparison (Experiment 2)

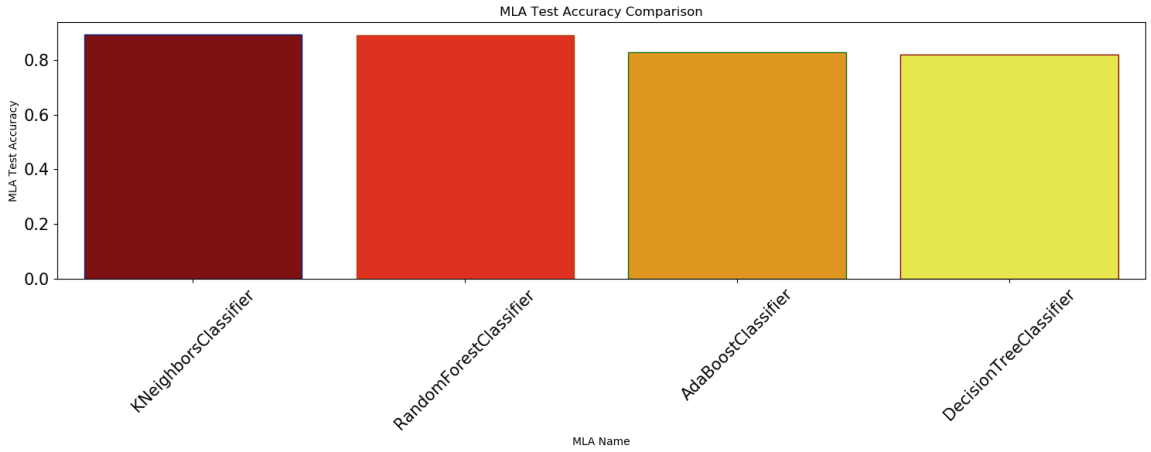


Figure 6.9: Machine learning algorithms' test accuracy comparison (Experiment 2)

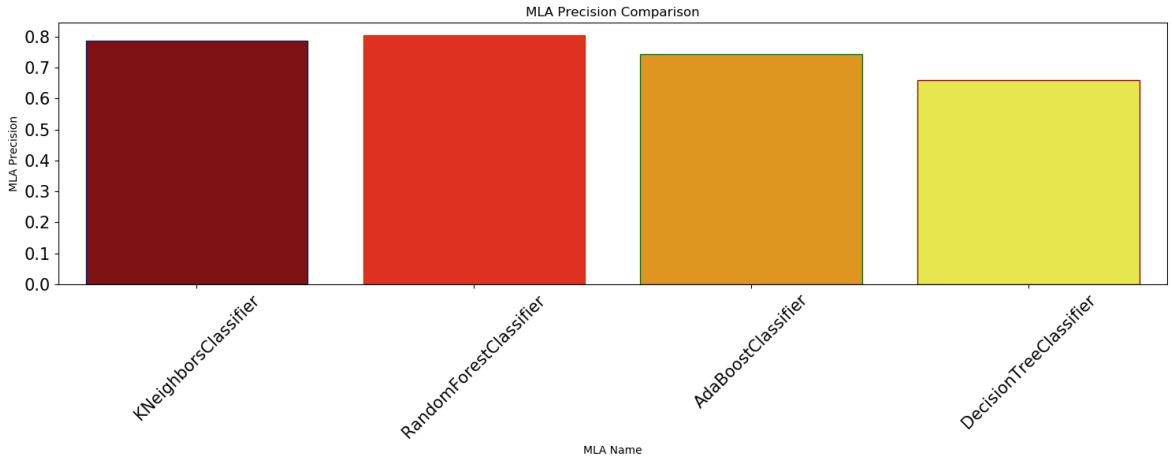


Figure 6.10: Machine learning algorithms' precision comparison (Experiment 2)

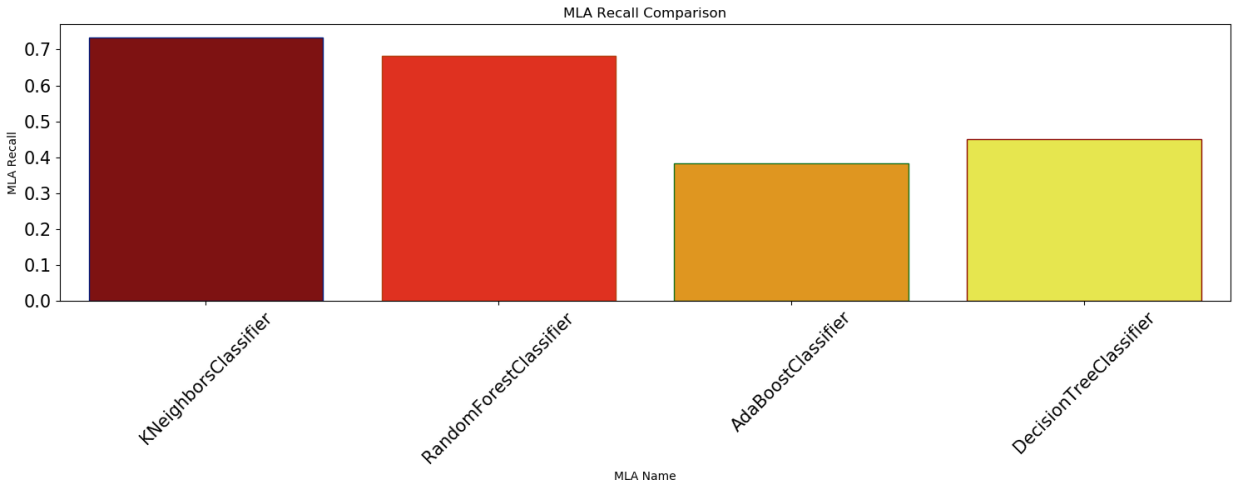


Figure 6.11: Machine learning algorithms' recall comparison (Experiment 2)

From Table 6.2, it can easily be seen that the k-NN classifier performed really well in separating the blood and non-blood samples with a testing accuracy of 89.23%. It can also be seen that some blood samples tend to give false results but the rate is not so much. Therefore, it can be considered as a good classifier model.

Table 6.2: Analysis of machine learning algorithms performance using different performance matrices for Exp. 2

MLA Name	**MLA Train Accuracy**	**MLA Test Accuracy**	**MLA Precision**	**MLA Recall**	**MLA AUC**
KNeighborsClassifier	0.9795	0.8923	0.7857	0.7333	0.8366
RandomForestClassifier	1.0	0.8885	0.8039	0.6833	0.8166
AdaBoostClassifier	0.9977	0.8269	0.7419	0.3833	0.6716
DecisionTreeClassifier	1.0	0.8192	0.6585	0.45	0.69

The confusion matrix of the k-NN model is shown in Fig 6.13. From that figure, it can be understood clearly that, out of the 200 samples of non-blood samples 188 samples were correctly predicted and 12 samples were wrongly predicted. On the other hand, for the 60 blood samples, all the samples were predicted correctly. So, the k-NN classifier model provides an accuracy of around 89.23%.

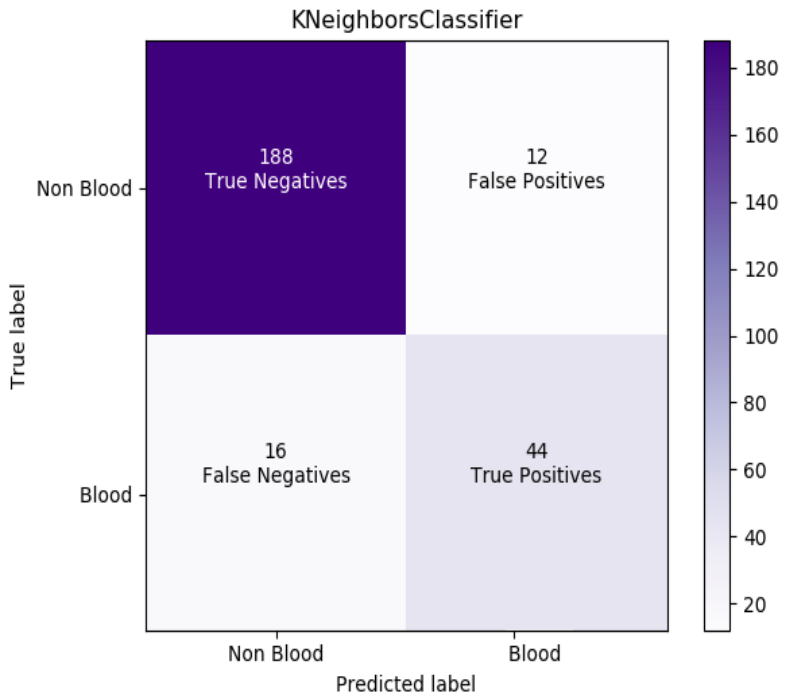


Figure 6.12: Confusion matrix of k-NN classifier (Experiment 2)

Summary of Experiment 2:

Training data: 7 blood samples, 15 non-blood samples (20 data/sample)

7*20=140, 15*20=300. Total= 440 samples

Fully separated testing data: 3 blood samples, 10 non-blood samples (20 data/sample)

Blood=3*20=60,

Non-blood=10*20=200,

Total= 260 samples

k-NN Classifier's Performance Details:

Blood/Positive= 60 (True Pos=44, False Neg=16)

Non-blood/Negatives=200 (True Neg=188, False Pos=12)

Test accuracy=89.23%

6.5 Development and Analysis of Models for Experiment 3

In case of experiment 3, machine learning models are developed based on the only 3 most significant features (500nm,550nm and 610nm). The difference between the experiment 2 and experiment 3 is basically with the dataset. In experiment 3, more blood and non-blood samples were added to represent the real experimental scenario by including pig's intestine with the existing dataset. There were 32 non-blood samples and 22 blood samples in experiment 3 compared to the 25 non-blood samples and 10 blood samples in experiment 2. For each of the samples around 20 sample data were collected for the training. Three top most selected features- 500nm, 550nm, and 610nm wavelengths were used for the development and analysis of the four classifier models. The k-NN classifier and decision tree classifier models perform better than the other two models. Fig 6.15 and Fig 6.16 indicate that k-NN classifier provides a training accuracy of 98.33% and test accuracy of 89.05% which are higher than the other algorithms. Fig 6.18 indicates that the k-NN classifier also provides a higher recall value of 0.8875 than the other classifiers. The k-NN has also shown a higher AUC value of 0.8899 which indicates that it is good at separating the different classes. The overall summary of all the classifier models can be observed in Table 6.3. Therefore, it can be seen that the k-NN classifier is very efficient at differentiating the blood and non-blood samples using the 3 specified spectral wavelengths.

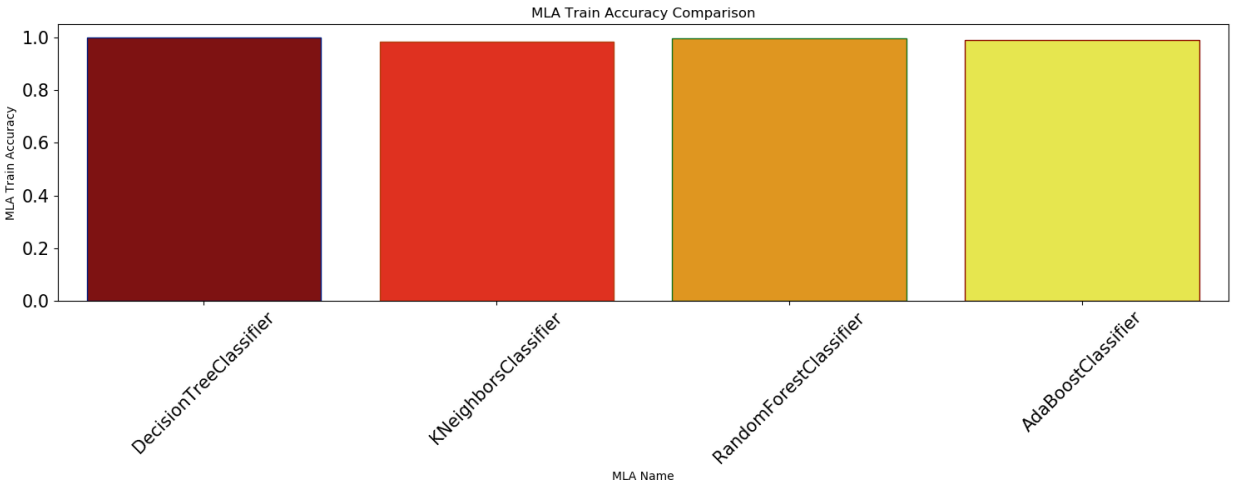


Figure 6.13: Machine learning algorithms' train accuracy comparison (Experiment 3)

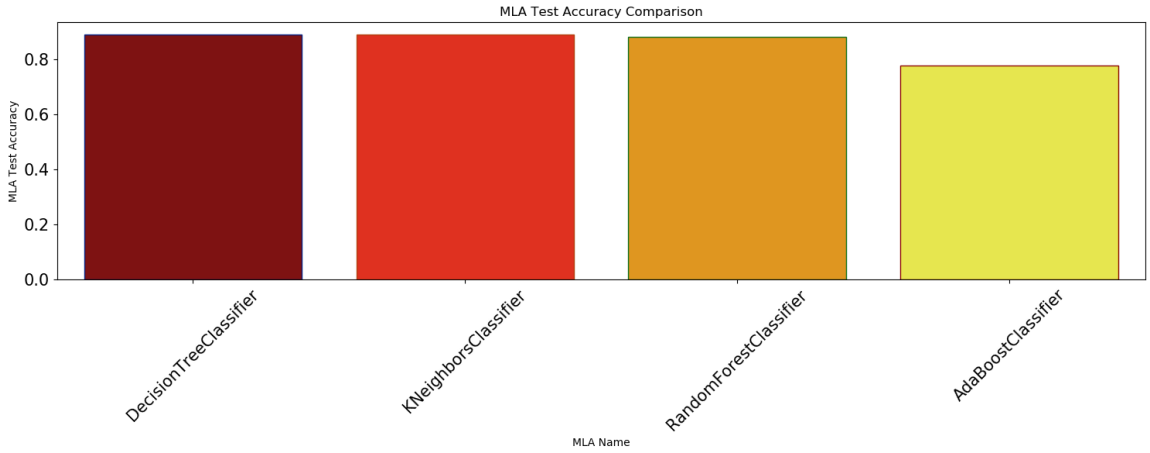


Figure 6.14: Machine learning algorithms' test accuracy comparison (Experiment 3)

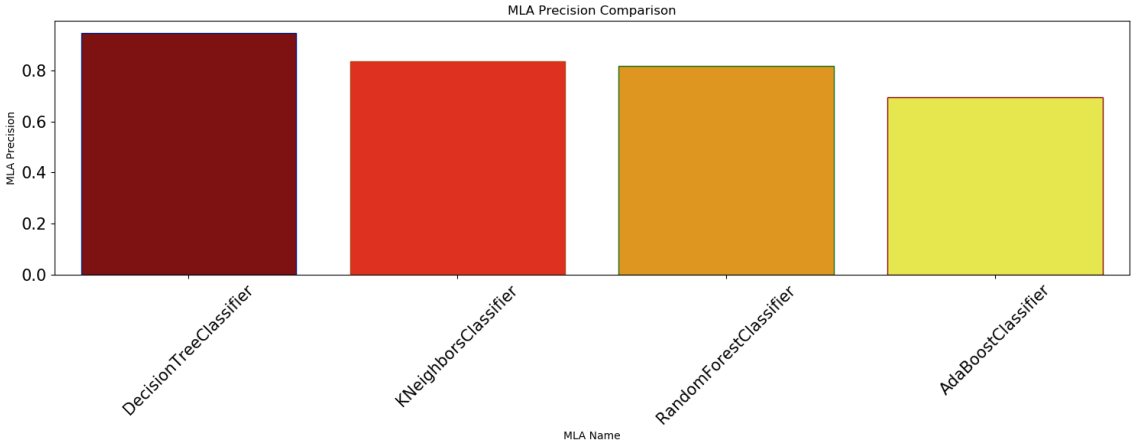


Figure 6.15: Machine learning algorithms' precision comparison (Experiment 3)

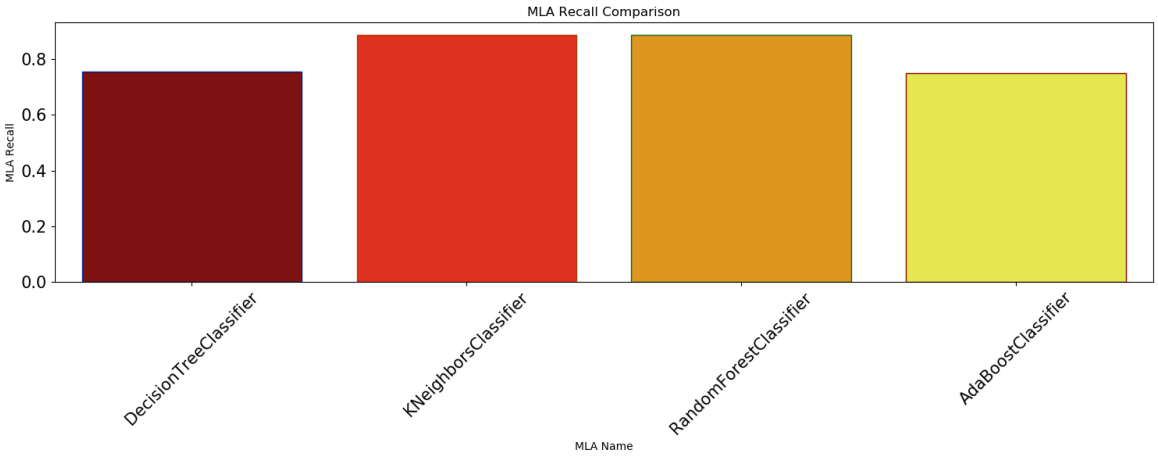


Figure 6.16: Machine learning algorithms' recall comparison (Experiment 3)

The k-NN performed really well in separating the blood and non-blood samples with a testing accuracy of 89.05%. Though the for some blood samples with lower concentration it tends to provide false results, still it can be considered as a good classifier model.

Table 6.3: Analysis of machine learning algorithms performance using different performance matrices for Exp. 3

MLA Name	MLA Train Accuracy	MLA Test Accuracy	MLA Precision	MLA Recall	MLA AUC
KNeighborsClassifier	0.9833	0.8905	0.8352	0.8875	0.8899
DecisionTreeClassifier	1.0	0.8905	0.9453	0.7562	0.8646
RandomForestClassifier	0.9985	0.8810	0.8160	0.8875	0.8822
AdaBoostClassifier	0.9909	0.7786	0.6936	0.75	0.7730

The confusion matrix of the k-NN model is shown in Fig 6.19. Out of the 260 samples of non-blood samples 232 samples were correct and 28 samples were wrong. In case of blood samples, 142 blood samples predicted correctly predicted out of the 160 samples which indicates an accuracy of around 89.05%.

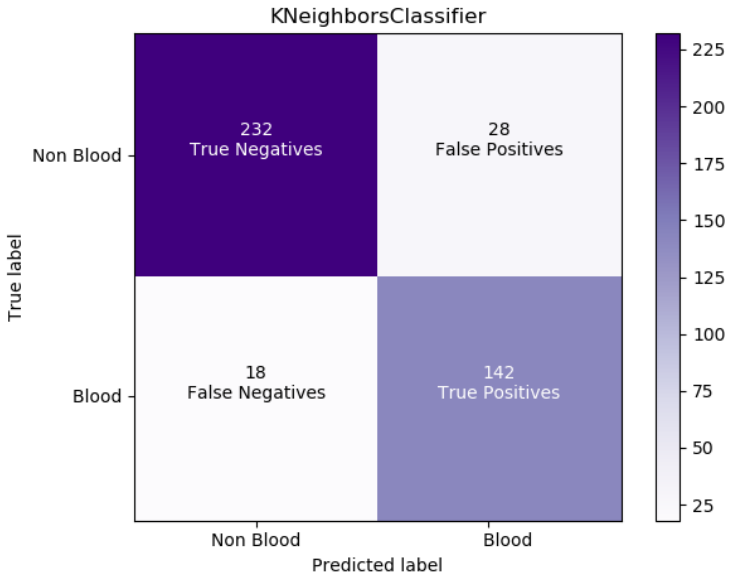


Figure 6.17: Confusion matrix of k-NN classifier (Experiment 3)

Summary of Experiment 3:

Training Data: 7 blood samples and 7 blood samples with pig's intestine, 15 non-blood Samples (20 data/sample) and 4 non-blood samples with pig's intestine (20data/sample)

14*20=280, 19*20=380. Total= 660 samples

Fully Separated Testing Data: 3 blood samples and 5 blood samples with pig's intestine, 10 non-blood Samples and 3 non-blood samples with pig's intestine

Blood=8*20=160,

Non-blood=13*20=260,

Total= 420 samples

k-NN Classifier's Performance Details:

Blood/Positive= 160 (True Pos=142, False Neg=18)

Non-blood/Negatives=260 (True Neg=232, False Pos=28)

Test Accuracy=89.05%

CHAPTER 7

COMPARATIVE ANALYSIS

The chapter described the analyses and the results of the models of the three different experiments. The experiments were done to explore the ways to separate the blood samples from the non-blood samples. The summary of the findings will be summarized in this chapter and then a comparison with some previous works will also be presented here.

7.1 Evaluation Metrics

Evaluating the model's accuracy, is an essential part of the process in developing any machine learning model since it describes how well the model is performing in its predictions. Evaluation metrics sometimes also change according to the problem type. In this research, three separate methods with different hardware components were used to detect blood using optical sensors. For the first method, two LEDs of 700nm and 625nm wavelength were used and for the other methods, 6-spectral wavelength visible sensor modules and 6-spectral wavelength near-infrared sensor modules were used. Two different datasets of blood and non-blood samples was used to determine the best method for detecting blood in the gastrointestinal tract. Four different classifier models were evaluated to determine the performance for all the three experiments.

Linear or non-linear models can be a typical example of binary classification problem which contains only true and false. After training all the models, those models are evaluated by checking the error rates evaluation metrics in prediction. The errors represent how much the models are making mistakes in prediction. The basic concept of accuracy evaluation is comparing the original response with the predicted one.

The MSE, MAE, RMSE, RMSLE, and R-Squared metrics are mainly used to evaluate the prediction error rates and model performance in the model's analysis.

- Mean absolute error (MAE) represents the difference between the original and predicted values extracted by averaged the absolute difference over the data set.
- Mean squared error (MSE) represents the difference between the original and predicted values extracted by squared the average difference over the data set.
- Root mean squared error (RMSE) is the error rate by the square root of MSE.
- R-squared (coefficient of determination) represents the coefficient of how well the values fit compared to the original values. The value from 0 to 1 interpreted as percentages. The higher the value is, the better the model is.
- There is a variation of RMSE that is sometimes more convenient to use - root mean squared logarithmic error (RMSLE). The only difference from the RMSE is that instead of using output, y values directly, it uses the logarithm of them.

The above metrics can be expressed as,

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \tag{7.1}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \tag{7.2}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \tag{7.3}$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \tag{7.4}$$

$$RMSLE(y,\hat{y}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\log_e(y_i+1) - \log_e(\hat{y}_i+1)\right)^2} \quad (7.5)$$

Where, $\hat{y} - predicted\ value\ of\ y\ and\ \hat{y} - mean\ value\ of\ y$

In this research work, three different experiments with different hardware designs were conducted to detect the bleeding in the gastrointestinal tract. 4 different classifier models were used to train the collected data using the three different hardware setups. The performances of all the experiments were measured using different performance matrices i.e. (RMSE, RMSLE, R-Squared, MSE, MAE, and ROC curve). Table 7.1 shows the results of classifiers' performance using various types of performance matrices.

Table 7.1: Results of final classifier's performance using different performance matrices

Experiment Type	**Classifier Name**	**Features**	**R-Squared**	**MAE**	**MSE**	**RMSE**	**RMSLE**	**ROC Curve (AUC%)**	**Test Accuracy (%)**
Experiment 1 (without pig's intestine)	k-NN classifier	700nm 625nm	0.523	0.0846	0.0846	0.2908	0.201	88.66	91.54
Experiment 2 (without pig's intestine)	k-NN classifier	500nm 550nm 610nm	0.393	0.1076	0.1076	0.3281	0.2274	83.66	89.23
Experiment 3 (with pig's intestine)	k-NN classifier	500nm 550nm 610nm	0.535	0.1095	0.1095	0.3309	0.2293	88.99	89.05

All of the experiments performed well in predicting the blood form the samples. In both experiment 2 and experiment 3, the k-NN classifier performed almost similar. The experiment 2 without pig's intestine sample data provides an overall accuracy of 89.23% whereas the experiment 3 with pig's intestine provides 89.05% which are very close to each other and can be used for practical environment for distinguishing blood from non-blood samples. It can also be observed from the ROC curve that, the experiment 3 with pig's intestine provides better performance with an AUC value of 88.99%. It could be because the experiment 2 used 10 hemoglobin blood samples and 25 food colors as the non-blood samples whereas in the experiment 3 with pig's intestine, 22 blood samples and 32 non-blood samples were used to detect the bleeding which provided more reliable information related to the bleeding.

7.2 Experimental Analysis with Different Concentrations of Blood

From the above three experiments, it can be seen that k-NN classifier performs better than the other classifiers. Furthermore, among all the three experiments, the 3rd experiment seems more reliable and efficient as it had considered the pig's intestine for training and testing the blood and non-blood samples. So, it is wise to consider the 3rd experiment as the most valid one. In order to explore the limitations of the 3rd experiment model, it should be tested with different concentrations of blood samples as well. In order to achieve that the blood samples were separated into three different groups based on their concentrations- low, medium, and high concentrations, which is shown in Table 7.2.

Table 7.2: List of blood samples with different level of concentrations

Group Number	Concentration Level	Blood Sample Description	Blood Sample Observations
1st Group	Low	Hemoglobin-1 % Hemoglobin -2 % Hemoglobin -5 %	60
2nd Group	Medium	Hemoglobin -8 % Hemoglobin -10 % Hemoglobin -13 % Hemoglobin -15 %	80
3rd Group	High	Hemoglobin -17 % Hemoglobin -20 % Hemoglobin -40 %	60

These three groups of blood concentrations were then tested with the 3rd experiment's developed k-NN classifier model to evaluate the performance of the model. The result is shown in Table 7.3.

Table 7.3: Summary of the results with different level of blood concentrations

Group Number	Blood Sample Observations	k-NN Classifier Predictions	Accuracy (%)
1st Group	60	True positive – 43 False negative – 17	71.67
2nd Group	80	True positive – 77 False negative – 3	96.25
3rd Group	60	True positive – 53 False negative – 7	88.33

It can be observed from the table that, when the blood concentration is lower than or equal to 5% hemoglobin, the model does not perform quite well. In case of medium concentrations of blood ranging from 8 to 15 %, the model performed really well with an accuracy of 96.25%. On the other hand, the model performed moderate in case of high concentrations of blood samples. The model does not perform well in high concentrations in contrast to the medium concentrations of blood. Since the developed model was trained on high levels of red, tan colored non-blood samples, it could not properly distinguish the higher concentrations of blood samples properly and

get confused with the non-blood samples. Therefore, the model struggled to get the right predictions for the high concentrations of blood. However, the overall accuracy of the k-NN classifier in predicting medium and high concentrations of blood samples is quite good.

7.3 Summary Comparison

In recent years, several studies have been conducted on the identification of bleeding with several optical and imagery instruments. Table 7.4 offers a contrast between the proposed work and different existing research works.

Table 7.4: Summary of proposed work and existing works

Article cited	Sensor Type	Measures	Parameter	BS	NBS	Comments/ Fails to separate
[76]	Violet (415nm) and red (720nm) LED	Transmission anchored in the GI wall	Ratio	1	5	N/A
[74]	TCS230D	Internal reflection	Two factors	13 Hemoglobin	0	N/A
[77]	LED (415nm,720nm)	Transmission	Ratio	05 0.1%, 1%, 5%, 10%, 50% human blood	19	Except 1% to 10% of blood
[75]	TCS 3200	Internal reflection	HSL	08 Normal human blood	02 Intestinal juice	Fails in low concentration Max/ (128, 256, 512)
[77]	Spectrophotometer	Transmission	Ratio	05 0.1%, 1%, 5%, 10%, 50% human blood	19	Except 1% to 10% of blood
[78]	LED (415nm,720nm)	Transmission	Ratio	06 5%, 7.7%, 13.3%, 15.4%, 33%, 40% human blood	Water	

Proposed Method (1st Experiment)	LED (700nm,625nm)	Reflection	Ratio, average	10 1%, 2%, 5%, 8%, 10%, 13%, 15%, 17%, 20%, 40% hemoglobin	25 Food colors	Accuracy of 91.54%
Proposed Method (2nd Experiment)	AS7262 (500nm,550nm) and AS7263 (610nm)	Reflection	Spectral wavelengths (2 in visible region, 1 in near infrared region)	10 1%, 2%, 5%, 8%, 10%, 13%, 15%, 17%, 20%, 40% hemoglobin	25 Food colors	Accuracy of 89.23%
Proposed Method (3rd Experiment)	AS7262 (500nm,550nm) and AS7263 (610nm)	Reflection	Spectral wavelengths (2 in visible region, 1 in near-infrared region)	22 1%, 2%, 5%, 8%, 10%, 13%, 15%, 17%, 20%, 40% hemoglobin and 12 hemoglobin with pig's intestine	32 Food colors with pig's intestine	Accuracy of 89.05%. Lower accuracy at 5% or lower blood concentrations

There does not have enough things to compare with the previous research articles as the dataset used in various studies was completely different. If a quantitative comparison is attempted, it can result in confusion between the studies. Therefore, specific analysis of works such as accuracy, detail, responsiveness, and little more is challenging.

The analysis may rely on only a few things like the number of samples used, validity of the proposed model, operating area and operating mode. Firstly, there is much more test samples than any other previous works. Furthermore, this analysis used two separate datasets for the training and validation of the classifier models, while no study showed the validation by using a separate dataset. Furthermore, the sensors used in this study can detect any blood level except very low,

similar to the previous works. A combination of multi-spectral sensors can be considered to diagnose bleeding more efficiently in the visible, near-infrared and ultraviolet areas.

The proposed experimental methods have been tested with different concentrations of blood samples and various non-blood samples with and without pig's intestine. The methods have been able to distinguish the blood and non-blood samples with an accuracy of 91.54% using the 2 features without the pig's intestine, 89.23% and 89.05% with the three most significant features of combined visible and near-infrared spectral wavelengths including and excluding pig's intestine respectively. It was also found that the k-NN classifier did not perform well for 5% or lower concentrations of blood samples. However, it performed well for higher concentrations of blood samples.

CHAPTER 8

CONCLUSION AND FUTURE WORK

Three separate studies were proposed and the findings were discussed in this chapter. The chapter concludes the entire study and provides suggestions for potential improvements in the future.

8.1 Conclusion

The purpose of the research is to propose a method related to development of a bleeding detection sensor for the WCE system. Nevertheless, owing to similarity in blood components [32] the same sensing and procedures could be used to classify human GI bleeding. Therefore, blood samples used in the analysis are not human blood but bovine hemoglobin.

From Table 7.1, the experimental setup with only two different wavelength LEDs (700nm and 625nm) provides a better result than the combined hardware setup using the AS7262 and AS7263 spectral sensors. According to experiment 2 and experiment 3, it can be seen that the outcomes using food colors with and without pig's intestine provide an accuracy of 89.23% and 89.05% respectively which are quite similar. This kind of bleeding detection sensor system may be used with the existing WCE system by replacing the existing camera system which takes a long time to detect bleeding.

The three spectral wavelengths- 500nm,550nm, and 610nm of the AS7262 and AS7263 chip provide a better bleeding observation in both scenarios of with and without pig's intestine which may not be possible by an imaging sensor that only analyzes the color properties of the blood rather than the spectral properties of blood.

Finally, it has been shown that the experiments with AS7262 and AS7263 sensors provide an overall accuracy of around 89% which could solve the problem of bleeding detection in all scenarios. Though the k-NN model is not good at predicting blood samples of 5% or lower concentrations, it has shown great efficiency and improved performance compared to the other existing methods.

8.2 Recommendation for Future Work

In this section, some recommendations have been provided to improve the design to detect the human GI bleeding which was beyond the scope of the thesis.

1. The experiments carried out in this analysis, take place in a controlled environment. In spite of some of the experiments were done with real pig's intestine, the results might be slightly different in real scenarios. Thus in-vivo sample testing with sensors is needed to confirm that the proposed sensors and the corresponding operating mode can be used in a living body environment.

2. The sensor, processor, transmission modules and power modules are becoming smaller and more efficient day by day. Therefore, several sensors might be combined into the endoscopic system for higher accuracy with more analytics.

3. The scale of the proposed sensors is slightly greater than that of capsule devices available on the market. It needs to develop technologies further to limit the size to a minimum.

4. One important point in miniaturizing the system is to design an IC which will combine all modules into a single integrated die. Thus, it would take less space and consume less power.

5. In addition to the current features, other features like pH level, glucose level, etc., could be incorporated to detect GI tract abnormalities. If there could have a triggering feature in the capsule which would change the camera's frame rate or wake up the capsule from power saver mode, then it could significantly save power consumption.

6. In addition, there could be some research on getting the accurate position of the WCE capsule which will provide the exact bleeding location inside the body when bleeding is detected by the sensors.

REFERENCES

[1] "Gastrointestinal (GI) Bleeding," [Online]. Available: https://www.niddk.nih.gov/health-information/digestive-diseases/gastrointestinal-bleeding. [Accessed: 01-Aug-2019].

[2] B. S. M. Kim *et al.*, "Diagnosis of gastrointestinal bleeding: A practical guide for clinicians," in *World J. Gastrointest. Pathophysiol.*, vol. 5, no. 4, pp. 467–478, Nov. 2014.

[3] M. C. Bateson and I. A. D. Bouchier, "Gastrointestinal Bleeding," in *Clinical Investigations in Gastroenterology*, Cham: Springer International Publishing, 2017, pp. 103–108.

[4] A. D. Pryor, T. N. Pappas, and M. S. Branch, *Gastrointestinal bleeding: a practical approach to diagnosis and management*, Springer-Verlag New York, 2010.

[5] H. Z. Girgis, B. R. Mitchell, T. Dassopoulos, G. Mullin, and G. Hager, "An intelligent system to detect Crohn's disease inflammation in wireless capsule endoscopy videos," in *2010 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2010 - Proceedings*, 2010, pp. 1373–1376.

[6] R. Kozarek, *Endoscopy in Small Bowel Disorders.* Springer International Publishing, 2016.

[7] B. Li, M. Q. H. Meng, and J. Y. W. Lau, "Computer-aided small bowel tumor detection for capsule endoscopy," in *Artif. Intell. Med.*, vol. 52, no. 1, pp. 11–16, May 2011.

[8] G. Pan and L. Wang, "Swallowable Wireless Capsule Endoscopy: Progress and Technical Challenges," in *Gastroenterol. Res. Pract.*, vol. 2012, 2012, Art no. 841691.

[9] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," in *Nature*, vol. 405, no. 6785, pp. 417, May 2000.

[10] Y. Yuan and M. Q. H. Meng, "Polyp classification based on Bag of Features and saliency in wireless capsule endoscopy," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2014, pp. 3930–3935.

[11] R. Kumar *et al.*, “Learning disease severity for capsule endoscopy images,” in *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, 2009, pp. 1314–1317.

[12] Y. Yuan, B. Li, and M. Q. H. Meng, “Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video,” in *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 2, pp. 624–630, Mar. 2016.

[13] M. Keroack, “Video capsule endoscopy,” in *Curr. Opin. Gastroenterol.*, vol. 20, no. 5, pp. 474-81, Sep. 2004.

[14] Y. Fu, W. Zhang, M. Mandal, and M. Q. H. Meng, “Computer-aided bleeding detection in WCE video,” in *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 2, pp. 636–642, Mar. 2014.

[15] T. C. Lee, Y. H. Lin, N. Uedo, H. P. Wang, H. T. Chang, and C. W. Hung, “Computer-aided diagnosis in endoscopy: a novel application toward automatic detection of abnormal lesions on magnifying narrow-band imaging endoscopy in the stomach,” in *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 4430–4433

[16] A. R. Hassan and M. A. Haque, “Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos,” in *Comput. Methods Programs Biomed.*, vol. 122, no. 3, pp. 341–353, Dec. 2015.

[17] M. Liedlgruber and A. Uhl, "Computer-Aided Decision Support Systems for Endoscopy in the Gastrointestinal Tract: A Review," in *IEEE Reviews in Biomedical Engineering*, vol. 4, pp. 73-88, Nov. 2011.

[18] P. C. De Groen, “History of the Endoscope [Scanning Our Past],” in *Proc. IEEE*, vol. 105, no. 10, 2017, pp. 1987–1995.

[19] A. Di Ieva, M. Tam, M. Tschabitscher, and M. D. Cusimano, “A journey into the technical evolution of neuroendoscopy,” in *World Neurosurg.*, vol. 82, no. 6, pp. 777-789, Dec. 2014.

[20] B. Krans, “Endoscopy: Purpose, Procedure & Types,” [Online]. Available: https://www.healthline.com/health/endoscopy. [Accessed: 01-Aug-2019].

[21] R. M. E. Engel, “Philipp Bozzini—The Father of Endoscopy,” in *J. Endourol.*, vol. 17, no. 10, pp. 859–862, Dec. 2003.

[22] “Bozzini and the Lichtleiter - Looking into the Body - Diagnosis - History of Urology - EAU European Museum of Urology,” [Online]. Available: http://history.uroweb.org/history-ofurology/diagnosis/looking-into-the-body/. [Accessed: 01-Aug-2019].

[23] H. H. Hopkins and N. S. Kapany, “A flexible fibrescope, using static scanning,” in *Nature*, vol. 173, no. 4392. pp. 39–41, 1954.

[24] C. Nezhat, “Nezhat's History of Endoscopy,” [Online]. Available: http://laparoscopy.blogs.com/endoscopyhistory/chapter_08/. [Accessed: 01-Aug-2019].

[25] Z. Li, Z. Liao, and M. McAlindon, *Handbook of capsule endoscopy*. Springer Netherlands, 2014.

[26] D. J. Hass, *Capsule endoscopy: a guide to becoming an efficient and effective reader*. Springer, 2017.

[27] L. J. Sliker and G. Ciuti, “Flexible and capsule endoscopy for screening, diagnosis and treatment,” in *Expert Rev. Med. Devices*, vol. 11, no. 6, pp. 649–666, Nov. 2014.

[28] P. Valdastri, M. Simi, and R. J. Webster, “Advanced Technologies for Gastrointestinal Endoscopy,” in *Annu. Rev. Biomed. Eng.*, vol. 14, no. 1, pp. 397–429, Jul. 2012.

[29] V. K. Zworykin, “A ‘Radio Pill’,” in *Nature*, vol. 179, no. 4566, p. 898, 1957.

[30] S. C. Mukhopadhyay, “Wearable sensors for human activity monitoring: A review,” in *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, Mar. 2015.

[31] M. H. Floch, “Capsule Endoscopy,” in *J. Clin. Gastroenterol.*, vol. 42, no. 7, 2008.

[32] B. H. A. Maas, A. Buursma, R. A. J. Ernst, A. H. J. Maas, and W. G. Zijlstra, “Lyophilized bovine hemoglobin as a possible reference material for the determination of hemoglobin derivatives in human blood,” in *Clin. Chem.*, vol. 44, no. 11, pp. 2331–2339, Nov. 1998.

[33] “Spectra in the Lab.” [Online]. Available: https://www.ifa.hawaii.edu/~barnes/ASTR110L_F05/spectralab.html. [Accessed: 28-Nov-2019].

[34] N. Altman and M. Krzywinski, “Points of Significance: Clustering,” in *Nature Methods*, vol. 14, no. 6, pp. 545–546, May 2017.

[35] A. H. Khan, “Design of An Optical Sensor to Detect Human Intestinal Bleeding for Capsule Endoscopy,” M.S Thesis, Dept. Electrical and Computer Eng., University of Saskatchewan, Saskatoon, Canada, 2018.

[36] “Feature importances with forests of trees — scikit-learn 0.22.1 documentation,” [Online]. Available: https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html. [Accessed: 19-Dec-2019].

[37] ASGE Standards of Practice Committee *et al.*, “The role of endoscopy in the management of obscure GI bleeding,” in *Gastrointest. Endosc.*, vol. 72, no. 3, pp. 471–479, Sep. 2010.

[38] L. S. Friedman and P. Martin, “The problem of gastrointestinal bleeding,” in *Gastroenterol. Clin. North Am.*, vol. 22, no. 4, pp. 717–721, Dec. 1993.

[39] C. Ell and A. May, “Mid-gastrointestinal bleeding: Capsule endoscopy and push-and-pull enteroscopy give rise to a new medical term,” in *Endoscopy*, vol. 38, no. 1, pp. 73–75, Jan. 2006.

[40] V. Singh and J. A. Alexander, “The evaluation and management of obscure and occult gastrointestinal bleeding,” in *Abdominal Imaging*, vol. 34, no. 3. pp. 311–319, Jun. 2009.

[41] S. F. Pasha, A. K. Hara, and J. A. Leighton, “Diagnostic evaluation and management of obscure gastrointestinal bleeding: a changing paradigm,” in *Gastroenterol. Hepatol. (N. Y).*, vol. 5, no. 12, pp. 839–850, Dec. 2009.

[42] A. C. Stein, A. Appannagari, I. Habib, C. E. Semrad, and D. T. Rubin, “A Rapid and Accurate Method to Detect Active Small Bowel Gastrointestinal Bleeding on Video Capsule Endoscopy,” in *Dig. Dis. Sci.*, vol. 59, no. 10, pp. 2503–2507, Oct. 2014.

[43] G. S. Raju, L. Gerson, A. Das, and B. Lewis, “American Gastroenterological Association (AGA) Institute Technical Review on Obscure Gastrointestinal Bleeding,” in *Gastroenterology*, vol. 133, no. 5, pp. 1697–1717, Nov. 2007.

[44] “G likelihood-ratio test.” [Online]. Available: https://influentialpoints.com/Training/g-likelihood_ratio_test.htm. [Accessed: 13-Apr-2020].

[45] H. J. Haugan, S. Elhamri, F. Szmulowicz, B. Ullrich, G. J. Brown, and W. C. Mitchel, “Study of residual background carriers in midinfrared InAsGaSb superlattices for unfcooled detector operation,” in *Appl. Phys. Lett.*, vol. 92, no. 7, Feb. 2008, Art no. 071102.

[46] J. A. Collins *et al.*, “Relating oxygen partial pressure, saturation and content: the haemoglobin-oxygen dissociation curve,” in *Breathe*, vol. 11, no. 3, pp. 194–201, Sep. 2015.

[47] T. Ghosh, S. K. Bashar, S. A. Fattah, C. Shahnaz and K. A. Wahid, "An automatic bleeding detection scheme in wireless capsule endoscopy based on statistical features in hue space," in *17th International Conference on Computer and Information Technology (ICCIT)*, Dhaka, 2014, pp. 354-357.

[48] “1.13. Feature selection — scikit-learn 0.22.1 documentation,” [Online]. Available: https://scikit-learn.org/stable/modules/feature_selection.html. [Accessed: 16-Dec-2019].

[49] G. Pan, G. Yan, X. Qiu, and J. Cui, “Bleeding detection in wireless capsule Endoscopy based on probabilistic neural network,” in *J. Med. Syst.*, vol. 35, no. 6, pp. 1477–1484, Dec. 2011.

[50] N. Bourbakis, S. Makrogiannis, and D. Kavraki, “A neural network-based detection of bleeding in sequences of WCE images,” in *Proceedings - BIBE 2005: 5th IEEE Symposium on Bioinformatics and Bioengineering*, 2005, pp. 324–327.

[51] C. K. Poh *et al.*, “Multi-level local feature classification for bleeding detection in Wireless Capsule Endoscopy images,” in *2010 IEEE Conference on Cybernetics and Intelligent Systems*, 2010, pp. 76–81.

[52] P. Y. Lau and P. L. Correia, “Detection of bleeding patterns in WCE video using multiple features,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2007, pp. 5601–5604.

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

[54] Y. Zheng *et al.*, "Single-image vignetting correction using radial gradient symmetry," in *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, Anchorage, Alaska, 2008, pp. 1-8.

[55] J. Forbes, Ed., “Progress of Human Anatomy and Physiology,” in *British and Foreign Medical Review: Or, Quarterly Journal of Practical Medicine and Surgery*, 1845, pp. 265–270.

[56] “AS726X NIR/VIS Spectral Sensor Hookup Guide - learn.sparkfun.com,” [Online]. Available: https://learn.sparkfun.com/tutorials/as726x-nirvi/all. [Accessed: 12-Dec-2019].

[57] “AS7262 6-Channel Spectral ID Device - Spectral Sensing | PMT,” [Online]. Available: https://www.pmt-fl.com/spectral-sensing/as7262. [Accessed: 12-Dec-2019].

[58] “AS7263 6-Channel Spectrometer - AMS | Mouser Canada,” [Online]. Available: https://www.mouser.ca/new/ams/ams-as7263-spectral-id-device/. [Accessed: 12-Dec-2019].

[59] T. T. Kararli, “Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals,” in *Biopharm. Drug Dispos.*, vol. 16, no. 5, pp. 351–380, Jul. 1995.

[60] A. A. Al-Rahayfeh and A. A. Abuzneid, “Detection of Bleeding in Wireless Capsule Endoscopy Images Using Range Ratio Color,” in *Int. J. Multimed. Its Appl.*, vol. 2, no. 2, pp. 1–10, May 2010.

[61] D. K. Iakovidis and A. Koulaouzidis, “Software for enhanced video capsule endoscopy: challenges for essential progress,” in *Nat. Rev. Gastroenterol. Hepatol.*, vol. 12, no. 3, pp. 172–186, Mar. 2015.

[62] D. K. Iakovidis and A. Koulaouzidis, “Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software,” in *Gastrointest. Endosc.*, vol. 80, no. 5, pp. 877–883, Nov. 2014.

[63] 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,” in *Conf. Proc. Annual. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annual. Conf.*, 2011, pp. 6643–6646.

[64] A. Karargyris and N. Bourbakis, “Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos,” in *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2777–2786, Oct. 2011.

[65] A. V. Mamonov, I. N. Figueiredo, P. N. Figueiredo, and Y. H. Richard Tsai, “Automated polyp detection in colon capsule endoscopy,” in *IEEE Trans. Med. Imaging*, vol. 33, no. 7, pp. 1488–1502, Mar. 2014.

[66] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Proceedings of the Alvey Vision Conference 1988*; Romsey, UK. Romsey, 1998, pp. 147–151

[67] “Is blood really blue? - Science World,” [Online]. Available: https://www.scienceworld.ca/blood-really-blue/. [Accessed: 10-Dec-2019].

[68] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proceedings of the IEEE International Conference on Computer Vision*, Kerkyra, Greece, vol. 2, 1999, pp. 1150–1157.

[69] Y.G. Lee and G. Yoon, “Real-Time Image Analysis of Capsule Endoscopy for Bleeding Discrimination in Embedded System Platform,” in *Int. J. of Biomed. and Biolog. Engineering,* vol. 5, no. 11, pp. 583–587, 2011.

[70] K. Mikolajczyk and C. Schmid, “Scale & Affine Invariant Interest Point Detectors,” in *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, Oct. 2004.

[71] C. Signorelli *et al.*, “Sensitivity and specificity of the suspected blood identification system in video capsule enteroscopy,” in *Endoscopy*, vol. 37, no. 12, pp. 1170–1173, Dec. 2005.

[72] A. Novozámský *et al.*, “Automatic blood detection in capsule endoscopy video,” in *J. Biomed. Opt.*, vol. 21, no. 12, pp. 126007.1–126007.8, Dec. 2016.

[73] “SparkFun Spectral Sensor Breakout - AS7263 Visible (Qwiic),” [Online]. Available: https://www.sparkfun.com/products/14351. [Accessed: 01-Aug-2019].

[74] H. Liu *et al.*, “An intelligent electronic capsule system for automated detection of gastrointestinal bleeding,” in *J. Zhejiang Univ. Sci. B*, vol. 11, no. 12, pp. 937–943, Dec. 2010.

[75] P. Qiao, H. Liu, X. Yan, Z. Jia, and X. Pi, “A Smart Capsule System for Automated Detection of Intestinal Bleeding Using HSL Color Recognition,” in *PLoS One*, vol. 11, no. 11, Nov. 2016, Art no. e0166488

[76] S. Schostek and M. O. Schurr, “The HemoCop Telemetric Sensor System- Technology and Result of in-vivo Assessment,” in *Stud Health Technol Inform.*, vol. 177, pp. 97–100, 2012.

[77] S. Schostek *et al.*, “Telemetric real-time sensor for the detection of acute upper gastrointestinal bleeding,” in *Biosens. Bioelectron.*, vol. 78, pp. 524–529, Apr. 2016.

[78] S. Schostek *et al.*, "Volunteer Case Series of a New Telemetric Sensor for Blood Detection in the Upper Gastrointestinal Tract: The HemoPill," in *Dig. Dis. Sci.*, vol. 61, no. 10, pp. 2956–2962, 2016.

[79] Y. Zou, L. Li, Y. Wang, J. Yu, Y. Li, and W.L. Deng, "Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Singapore. Piscataway (NJ), 2015, pp. 1274–1278.

[80] S. Seguí *et al.*, "Generic feature learning for wireless capsule endoscopy analysis," in *Comput Biol Med.*, vol. 79, pp. 163–172, Dec. 2016.

[81] X. Jia and M. Q. H. Meng, "A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016, pp. 639–642.

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

[83] R. Leenhardt *et al.*, "A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy," in *Gastrointest. Endosc.*, vol. 89, no. 1, pp. 189–194, 2019.

[84] Y. Yuan and M.Q. Meng, "Deep learning for polyp recognition in wireless capsule endoscopy images," in *Med Phys.*, vol. 44, pp. 1379–1389, Apr. 2017.

[85] J. Y. He, X. Wu, Y. G. Jiang, Q. Peng, and R. Jain, "Hookworm Detection in Wireless Capsule Endoscopy Images with Deep Learning," in *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2379–2392, May 2018.

[86] D.K. Iakovidis, S.V. Georgakopoulos, M. Vasilakakis, A. Koulaouzidis, and V. P. Plagianakos, “Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification,” in *IEEE Trans Med Imaging*., vol. 37, pp. 2196–2210, Oct. 2018.

[87] A. Oliva and A. Torralba, “Modeling the shape of the scene: a holistic representation of the spatial envelope,” in *Int J Comput Vis*., vol. 42, pp. 145–175, May 2001.

[88] S. Khan and S. P. Yong, “A comparison of deep learning and hand-crafted features in medical image modality classification,” in *2016 3rd International Conference on Computer and Information Sciences (ICCOINS);* 2016 Aug 15-17; Kuala Lumpur, Malaysia. Piscataway (NJ). 2016, pp. 633–638.

[89] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, July 2002.

[90] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” in *J Mach Learn Res*., vol.15, pp. 1929–1958, Jan. 2014.

[91] M. Cogswell, F. Ahmed, R. B. Girshick, C. L. Zitnick, and D. Batra, “Reducing Overfitting in Deep Networks by Decorrelating Representations,” in *CoRR*, vol. abs/1511.06068, Jun. 2016.

[92] Y. Hwang, J. Park, Y. J. Lim, and H. J. Chun, “Application of Artificial Intelligence in Capsule Endoscopy: Where Are We Now?,” in *Clin. Endosc*., vol. 51, pp. 547–551, Nov. 2018.

[93] S. Hwang, J. Oh, J. Cox, S. J. Tang, and H. F. Tibbals, “Blood detection in wireless capsule endoscopy using expectation maximization clustering,” in *Medical Imaging 2006: Image Processing*, vol. 6144, Mar. 2006, Art no. 61441P

[94] “SparkFun Spectral Sensor Breakout - AS7262 Visible (Qwiic),” [Online]. Available: https://www.sparkfun.com/products/14347. [Accessed: 01-Aug-2019].

[95] S. Borman, “The Expectation Maximization Algorithm: A short tutorial,” [Online]. Available: http://www.seanborman.com/publications/EM_algorithm.pdf. [Accessed: 01-Aug-2019].

[96] P. Wang, S. M. Krishnan, C. Kugean, and M. P. Tjoa, “Classification of endoscopic images based on texture and neural network,” in *Annual Reports of the Research Reactor Institute, Kyoto University*, 2001, pp. 3691–3695.

[97] Z. Li *et al.*, “The current main types of capsule endoscopy,” in *Handbook of Capsule Endoscopy*, Springer Netherlands, 2014, pp. 5–45.

[98] A. Szold, L. B. Katz, and B. S. Lewis, “Surgical approach to occult gastrointestinal bleeding,” in *Am. J. Surg.*, vol. 163, no. 1, pp. 90–93, Jan. 1992.

[99] S. Palimaka, G. Blackhouse, and R. Goeree, “Capsule Endoscopy in the Assessment of Obscure Gastrointestinal Bleeding: An Economic Analysis,” in *Ont. Health Technol. Assess. Ser.*, vol. 15, no. 2, pp. 1–32, Feb. 2015.

[100] M. Muñoz-Navas, “Capsule endoscopy,” in *World Journal of Gastroenterology*, vol. 15, no. 13, pp. 1584–1586, Apr. 2009.