Rukiye Nur KAÇMAZ

MINING COLONOSCOPY IMAGES FOR ABNORMALITY DETECTION

AGU 2020

MINING COLONOSCOPY IMAGES FOR ABNORMALITY DETECTION

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Ph.D.

By
Rukiye Nur KAÇMAZ
September 2020

GCRIS

MINING COLONOSCOPY IMAGES FOR ABNORMALITY DETECTION

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Ph.D.

By
Rukiye Nur KAÇMAZ
September 2020

GCRIS

SCIENTIFIC ETHICS COMPLIANCE

I hereby declare that all information in this document has been obtained in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name-Surname: Rukiye Nur KAÇMAZ
Signature :

GCRIS

REGULATORY COMPLIANCE

Ph.D. thesis titled Mining Colonoscopy Images for Abnormality Detection has been prepared in accordance with the Thesis Writing Guidelines of the Abdullah Gül University, Graduate School of Engineering and Science.

Prepared By	Advisor
Rukiye Nur KAÇMAZ	Prof. Bülent YILMAZ

Head of the Electrical and Computer Engineering Program

Assoc. Prof. Kutay İÇÖZ

GCRIS

ACCEPTANCE AND APPROVAL

Ph.D. thesis titled Mining Colonoscopy Images for Abnormality Detection prepared by Rukiye Nur KAÇMAZ has been accepted by the jury in the Electrical and Computer Engineering Graduate Program at Abdullah Gül University, Graduate School of Engineering and Science.

14 / 09 / 2020

JURY:

Advisor: Prof. Bülent YILMAZ

Member: Prof. Alper BAŞTÜRK

Member: Assist. Prof. Kasım TAŞDEMİR

Member: Assist. Prof. M. Zübeyir ÜNLÜ

Member: Assist. Prof. Zafer AYDIN

APPROVAL:

The acceptance of this Ph.D. thesis has been approved by the decision of the Abdullah Gül University, Graduate School of Engineering and Science, Executive Board dated ….. /….. / …….. and numbered ……..……. .

…….. /…….. / ……..

Graduate School Dean

Prof. Hakan USTA

GCRIS

ABSTRACT

MINING COLONOSCOPY IMAGES FOR ABNORMALITY DETECTION

Rukiye Nur KAÇMAZ
Ph.D. in Electrical and Computer Engineering Department
Supervisor: Prof. Bülent YILMAZ
September 2020

Detection of colon abnormalities is one of the most challenging tasks for gastroenterologists. However, the frames or videos obtained during the procedure are exposed to significant amount of unwanted artifacts such as motion artifact, specular reflection (SR), improper contrast levels, gastric juice and bubbles, or residuals. The images with such artifacts are called non-informative frames. In the first study, we investigated the effect of SR and use of image interpolation to remove SR in texture-based automatic polyp detection. We tested whether nearest neighbors, bilinear and bicubic interpolation methods caused any differences in terms of texture features and classification performance to discriminate polyps from the colon background. In the second study the main aim was to compare the performance of conventional machine learning and transfer learning methodologies in detecting non-informative frames. In machine learning part, we used gray level co-occurrence matrix, gray level run length matrix, neighborhood gray tone difference matrix, focus measure operators and three first order statistics, and random forest, support vector machines and decision tree approaches were used in the classification phase. In transfer learning part, we employed deep neural network architectures like AlexNet, SqueezeNet, GoogleNet, ShuffleNet, ResNet-18, ResNet-50, NasNetMobile, and MobileNet. The last study included the detection of colon abnormalities such as Crohn's, ulcerative colitis, cancer and polyp diseases on informative frames. The aim of this study was first to discriminate healthy frames from diseased ones, and to determine the disease types using both conventional machine learning and transfer learning approaches.
Keywords: Colonoscopy, artifacts, colon diseases, texture features, machine learning, transfer learning

ÖZET

ANORMALLİK TESPİTİ İÇİN VERİ MADENCİLİĞİ

Rukiye Nur KAÇMAZ
Elektrik ve Bilgisayar Mühendisliği Anabilim Dalında Doktora
Tez Yöneticisi: Prof. Dr. Bülent YILMAZ
Eylül 2020

Gastroentereloji uzmanları için kolon anormalliklerinin tespit edilmesi en zor görevlerden birisidir. Bununla birlikte işlem sırasında elde edilen görüntü veya videolar, hareket gürültüsü, yansıma gürültüsü (YG), uygun olmayan kontrast gürültüsü, mide öz suyu, baloncuklar veya kalıntılar içermektedir. Hastalık tespiti işlemi ise bilgi içeren temiz görüntüler ile yürütülmektedir. İlk çalışmada tekstür tabanlı otomatik polip tespitinde YG'nin etkisini ve YG'yi ortadan kaldırmak için kullanılan görüntü enterpolasyonunun kullanımı araştırıldı. Polipleri kolon arka planından ayırt etmek için, uygulanan en yakın komşular, bilineer ve bikübik enterpolasyon yöntemlerinin, tekstür özellikleri ve sınıflandırma performansı açısından herhangi bir farklılığa neden olup olmadığı test edildi. İkinci çalışmada temel amaç, bilgi taşımayan çerçeveleri tespit etmede geleneksel makine öğrenmesi ve transfer öğrenme yaklaşımlarının performanslarının karşılaştırılmasıydı. Makine öğrenmesi bölümünde, gri seviye eş oluşum matrisi, gri seviye koşu uzunluğu matrisi, komşuluk gri ton farkı matrisi, odak ölçüm operatörleri ve üç adet birinci derece istatistik kullanıldı. Sınıflandırma aşamasında rastgele orman, destek vektör makineleri ve karar ağacı yaklaşımları kullanılmıştır. Transfer öğrenme bölümünde derin sinir ağları olarak AlexNet, SqueezeNet, GoogleNet, ShuffleNet, ResNet-18, ResNet-50, NasNetMobile ve MobileNet tercih edildi. Son çalışma, bilgi taşıyan çerçevelerde Crohn, ülseratif kolit, kanser ve polip gibi kolon anormalliklerinin saptanmasını içermiştir. Bu çalışmanın amacı, öncelikle sağlıklı çerçeveleri hastalıklılardan ayırmak ve hem geleneksel makine öğrenmesi hem de transfer öğrenme yaklaşımlarını kullanarak hastalık türlerini belirlemekti.

Anahtar kelimeler: Kolonoskopi, gürültü, kolon hastalıkları, tekstür özellikleri, makina öğrenmesi, transfer öğrenme

Acknowledgements

Firstly, I would like to thank my research advisor Prof. Bülent YILMAZ. His office was always open whenever I had a problem or needed advice about my research. Especially, I am so grateful for his endless support and wise advice on my career plans.

I am particularly grateful to my husband Metin Sarper KAÇMAZ, who makes invaluable contributions to my life and has the most significant share in every success I have achieved so far.

I am so grateful to my dear friends Gizem ARGIN and Refika Sultan DOĞAN who always support me and share our wonderful times. I have had one of the most fun and quality times of my life with them in the last two years.

Finally, I must express my deep gratitude to my dear family for providing me with endless support and continuous encouragement. Nothing would have been possible without them.

GCRIS

Table of Contents

1 INTRODUCTION 10

1.2 OBJECTIVES AND SCOPE 11

2 BACKGROUND 13

2.1 COLON ANATOMY 13
2.2 COLONIC DISEASES 14
2.3 COLONOSCOPY TYPES 16
2.3.1 Virtual Colonoscopy 16
2.3.2 Wireless Capsule Endoscopy 17
2.3.3 Optical Colonoscopy 18
2.4 COLON IMAGE PROCESSING 19
2.4.1 Pre-processing 19
2.4.2 Feature extraction and Selection 20
2.4.3 Classification 29
2.4.4 Transfer Learning 29

3 Study 1:Specular Reflection On Polyp Detection In Colonoscopic Images 41

3.1 EFFECT OF SPECULAR REFLECTION ON POLYP DETECTION 41
3.2 LITERATURE REVIEW 42
3.3 METHOD 43
3.4 RESULT 49

4 Study 2: Non-Informative Frame Elimination In Colonoscopic Images 54

4.1 NON-INFORMATIVE FRAME ELIMINATION 55
4.2 LITERATURE REVIEW 55
4.3 METHOD 58
4.4 RESULT 60

5 Study 3: Abnormality Detection In Colonoscopic Images 65

5.1 ABNORMALITY DETECTION 65
5.2 LITERATURE REVIEW 66
5.3 METHOD 67
5.4 RESULT 69
5.4.1 Binary (Healthy Frames vs. Frames with Disease) Classification 70
5.4.2 Multi-label Classification 71

6 Conclusions And Future Prospects 73

6.1 CONCLUSIONS 73
6.2 CONTRIBUTION TO GLOBAL SUSTAINABILITY 79
6.3 FUTURE PROSPECTS 80
BIBLIOGRAPHY 83

List of Figures

Figure 2.1.1 Colon Structure…….…….. 13
Figure 2.2.1 Crohn's Disease and Ulcerative Colitis……..…..15
Figure 2.2.2 Polyp and Cancer……15
Figure 2.3.1.1 Virtual Colonoscopy and Output Image (Reconstructed)……16
Figure 2.3.2.1 Wireless Capsule Endoscopy and Output Image……17
Figure 2.3.3.1 Conventional Colonoscopy and a Sample Colonoscopy Image……18
Figure 2.4.3.1 Naïve Bayes Classifier …….……..31
Figure 2.4.3.2 Logistic Regression Classifier …….……..32
Figure 2.4.3.3 Decision Tree Classifier ……..….…… 32
Figure 2.4.3.4 Random Forest Classifier ……..……...…….……...…33
Figure 2.4.3.5 k-NN Classifier ……...…….……34
Figure 2.4.3.6 SVM Classifier …….……...……34
Figure 2.4.4.1 AlexNet Architecture …….…….…....……..…35
Figure 2.4.4.2 SqueezeNet Architecture …….…….…...……...36
Figure 2.4.4.3 GoogleNet Architecture …….……...……..….37
Figure 2.4.4.4 ShuffleNet Architecture …….……...……..…38
Figure 2.4.4.5 ResNet Architecture ……...……..…38
Figure 2.4.4.6 MobileNet Architecture ……...……...…....……..…39
Figure 2.4.4.7 NasNetMobile Architecture ……...……..…40
Figure 3.3.1 Three Different Polyps And Their Ground Truth Counterparts From CVC-Clinic DB Database……..……..……..……..…….…….…...44
Figure 3.3.2 Original Image……..……...45
Figure 3.3.3 Cropped Image……..……...45
Figure 3.3.4 Ground Truth Tiles……..……..……...45
Figure 3.3.5 Tiles With Reflection……...…....45
Figure 3.3.6 Four Different Size "Real" Reflections Added On The Tiles…….....46
Figure 4.3.1 Shows Four Images As Examples Of Informative And Non-Informative Frames From Our Image Database……..59
Figure 4.3.2. Adaptive Histogram Equalization Input And Output Images……..59
Figure 5.3.1 Sample Images From Our Dataset……...……....69

Figure 6.3.1 Real-Time Processing Pipeline Of Automatic Disease Detection On Colonoscopy Videos……..…..81

GCRIS

List of Tables

Table 2.4.2.1 The subfamilies of Focus Measure Operators (FMOs)…….29

Table 3.4.1 Summary of t-test results for the comparison of features obtained from reflection free tiles and interpolated tiles that included different size reflections…….51

Table 3.4.2 Summary of f-measure values for the comparison of interpolation methods and classification approaches to automatically discriminate tiles with polyps from the healthy background tiles with and without specular reflection of different sizes……. 52

Table 3.4.3. Accuracy of the best f-measure results for bilinear, nearest neighbor, and bicubic interpolation methods……..53

Table 4.2.1 The depth, number of parameters and input size of image for different type of transfer learning models……..……57

Table 4.4.1. Machine learning results of 5 different texture features using decision tree algorithm……..62

Table 4.4.2. Machine learning results of FMOs subcategories using decision tree algorithm……..62

Table 4.4.3 Transfer learning results of informative and non-informative discrimination…….……63

Table 4.4.4. The best results of both machine learning and transfer learning algorithm ……63

Table 5.3.1 Frame number of healthy and diseases databases……68

Table 5.4.1.1Performance metrics of machine learning results for binary classification……70

Table 5.4.1.2 Performance metrics of transfer learning results for binary classification……....71

Table 5.4.2.1 Performance metrics of conventional machine learning results of multi label classification……..….72

Table 5.4.2.2 Performance metrics of transfer learning results of disease type detection classification……..…..72

Table 6.1.1 Results of previous works related with non-informative frames……....75

Table 6.1.2 Results of previous works related with abnormality……..78

GCRIS

To my lovely family

GCRIS

Chapter 1

1 Introduction

1.1 General

Human digestive system is examined in two parts as upper and lower. The large intestine (colon) is located in the lower part of the digestive system. This part of our body may have some abnormalities such as Crohn's and Ulcerative Colitis which are also called as inflammatory bowel diseases (IBD). In addition to this kind of diseases, polyp is the most common disease in the large intestine. If polyps are not diagnosed in early stage, they may lead to cancer. Abovementioned large intestine diseases significantly reduce not only life quality but also survival rate [1, 2]. According to American Cancer Society Statistics predictions, 104,610 people will suffer from colon cancer, and also it is estimated that approximately half of these people (53,200) will lose their life in 2020 because of colon cancer diseases. An important reason for the late diagnosis of colon cancer is that people do not want to undergo colonoscopy until they feel a serious discomfort. In order to detect these problems, three different methods are preferred; wireless capsule endoscopy (WCE), virtual colonoscopy (VC) and optical or conventional colonoscopy (CC) [3]-[4]. In WCE a small pill like capsule is swallowed by the patient. The capsule captures thousands of images/frames while moving in the digestive tract including colon and wirelessly sends them to an external recorder. After this process is completed the capsule is egested from body. The disadvantage of this method is that the experts have to cope with approximately 60,000 images. WCE is time-consuming, however it may be preferred in the future because of its non-invasiveness once better fast image processing approaches are developed aiming WCE frames [2]. VC is an imaging procedure that uses computed tomography and computers to produce two-(axial slices) and three dimensional

(reconstructed) images of the colon. VC is not preferred frequently because it uses x-ray, and polyps or suspected tissues can only be detected but not removed during the same operation [3]. In case of detecting any kind of abnormalities such as polyp or any other differentiation, experts want to remove a piece of tissue (biopsy) and send to pathology. Neither VC nor WCE have an equipment to remove abnormal tissue from the large intestine. Only CC has suitable equipment to take a biopsy during the operation [4]. CC is a procedure that enables an expert to evaluate inside of the colon using a long, flexible tube with a camera and a light source at the tip. It is the most popular technique because experts have a chance to examine the colon surface both in real-time and offline. Due to the folded structure of the colon the examination of its surface can be difficult in real-time, however during CC, videos or images/frames can be recorded, and thus experts can evaluate the patient's status even after the operation (offline assessment). Thus, enhancement of conventional endoscopy properties may be evaluated as a noteworthy improvement. When all of these pros and cons are taken into account, detecting abnormalities using conventional colonoscopy is the main aim of this thesis. In this thesis, we will investigate methodologies to detect abnormalities on the images or frames from the videos. Before the frames are processed, we will develop a system to eliminate noisy/non-informative frames. This computer-aided system will help the experts to minimize diagnostic errors.

1.2 Objectives and Scope

Although abnormality detection on colonoscopic images has been a hot topic for a long time, no complete solution has been found yet. In the previous works, polyps, inflammation, abnormal vascularization, and bleeding can be listed as some of the abnormalities that emerge on colon tissue. Most of the previous studies have preferred studying on one type of abnormality detection. In this study we aimed to detect multi-disease abnormalities such as polyp, IBD, and cancer. In addition to this, our purpose was not only detecting abnormalities but also detecting inconvenient/non-informative frames automatically. The number of frames extracted from videos is high, and all of these frames are not convenient to perform processing on. Thus, we need to remove inconvenient/non-informative frames. These images are called inconvenient/non-

informative because specular reflection, residual, movement and camera focusing problem may affect image quality and visibility. We studied on CC images, but automatic non-informative frame elimination is also needed for WCE images. During WCE procedure approximately 60,000 images are obtained and it is not easy to detect disease from mixed dataset (both informative and non-informative included). Disease should be detected from clear/informative frames.

In this study, we studied on two different types of open source datasets. First dataset was comprised of the colon images from "CVC-ClinicDB" database prepared in the Hospital Clinic of Barcelona, Spain [5]. This dataset was used for the first part of this thesis which is the effect of interpolation on specular reflection for polyp detection. Second dataset included the frames extracted from the videos related with different types of colon diseases downloaded from https://www.gastrointestinalatlas.com/index.html link.

This thesis has three main parts. In the first part of this thesis, we investigated the effect of image interpolation on polyp detection on images with specular reflection. Two research questions were posed: How are the texture features used in automatic detection of polyps affected by the interpolation on specular reflections? (2) If they are affected by the interpolation approach does it really affect the classification performance? [6]. In the second part of this thesis, we studied methods to eliminate non-informative colonoscopy frames. In that part, we used conventional machine learning and transfer learning approaches [6]. In the last part, we explored abnormality detection problem from informative frames using conventional machine and transfer learning methods. Each part of this thesis will be elaborated separately in the following chapters.

Chapter 2

2 Background

2.1 Colon Anatomy

Colon which is also called large intestine is the last part of gastrointestinal part. It starts with cecum and end with anal part (Figure 2.1.1). Small intestine sends digested food to the large intestine. Colon is approximately 150 cm and examined in four parts which are ascending, transverse, descending and sigmoid. Large intestine encloses small intestine. Ascending part of colon is the beginning of colon. This part of colon starts just above cecum. The colon ascending is followed by the transverse colon. Transverse colon ends up with left colic flexure part and colon continues with descending part. Last part of colon is the sigmoid colon and it is about 40 cm long. Sigmoid colon is followed by rectum and anus respectively [7].

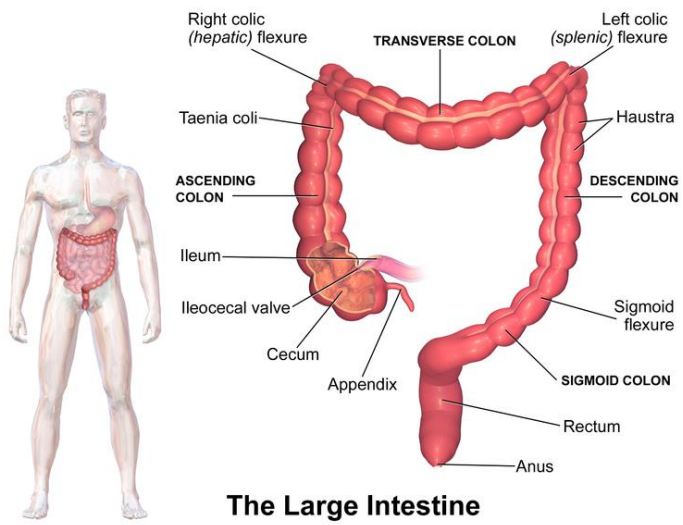


Figure 2.1.1 Colon structure [8].

2.2 Colonic Diseases

Colonic diseases arise from the colon's inability to perform its normal task over time. Diseases can be either benign or malign. The most common colon diseases are inflammatory bowel diseases (IBD), polyp, and cancer. IBD has two types as Crohn's and ulcerative colitis (UC). Experts sometimes have difficulty in discriminating Crohn's and UC. Generally, patients have complaints like diarrhea and abdominal pain. These diseases can be distinguished by their location. Crohn's disease may take place at any part of the digestive tract but UC generally occurs in the colon and rectum (Figure 2.2.1). These diseases can be cured with medical or surgical approaches. The other two diseases we have studied is polyp and colon cancer (Figure 2.2.2). The polyps are adenomatous or hyperplastic. While adenomatous polyp types are called malign, hyperplastic polyp types called benign. The risk of a polyp turning into cancer is high for malign type polyps when compared to the benign type polyp. However, when experts detect a polyp in a colon during the colonoscopy procedure, they tend to remove it. There are two reasons for this. First, it is not possible to know whether the polyp is benign or malign without removing the polyp using todays' technologies and secondly since polyps are likely to turn into cancer overtime, not removing the polyp may pose a risk to the patient's life. Cancerous part in the colon can be removed with surgery. In addition, a colonoscopy as a screening test may decrease the colon cancer risk [9].

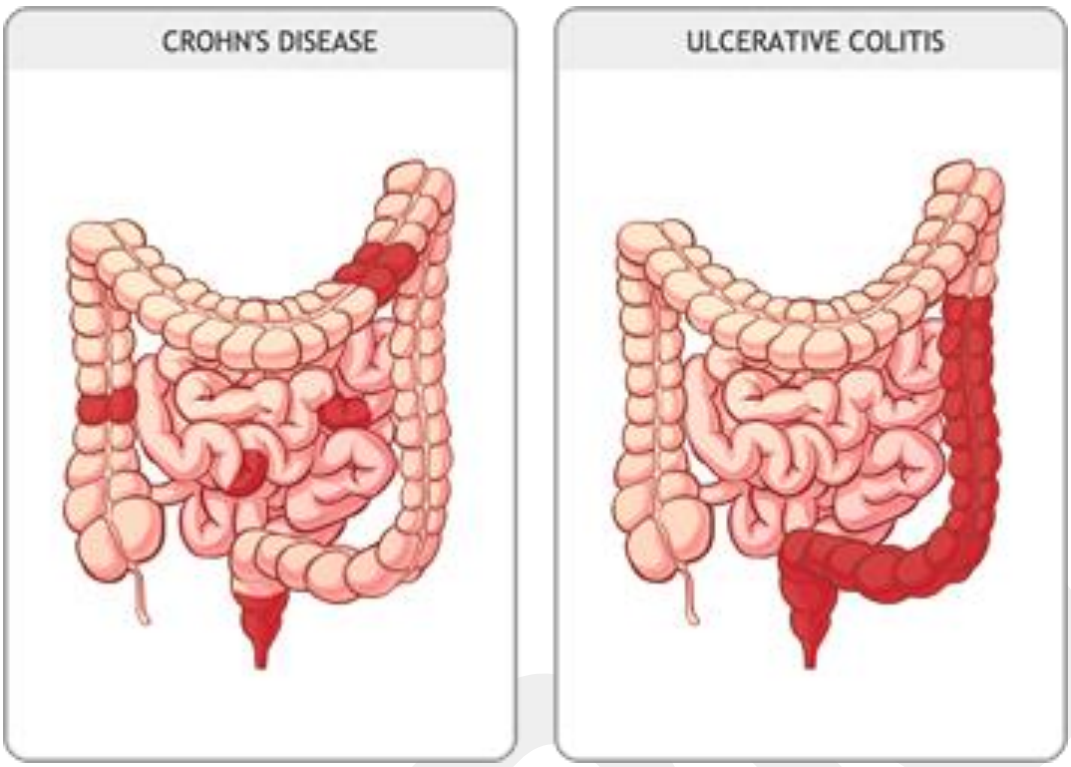


Figure 2.2.1 Crohn's disease and ulcerative colitis [10].

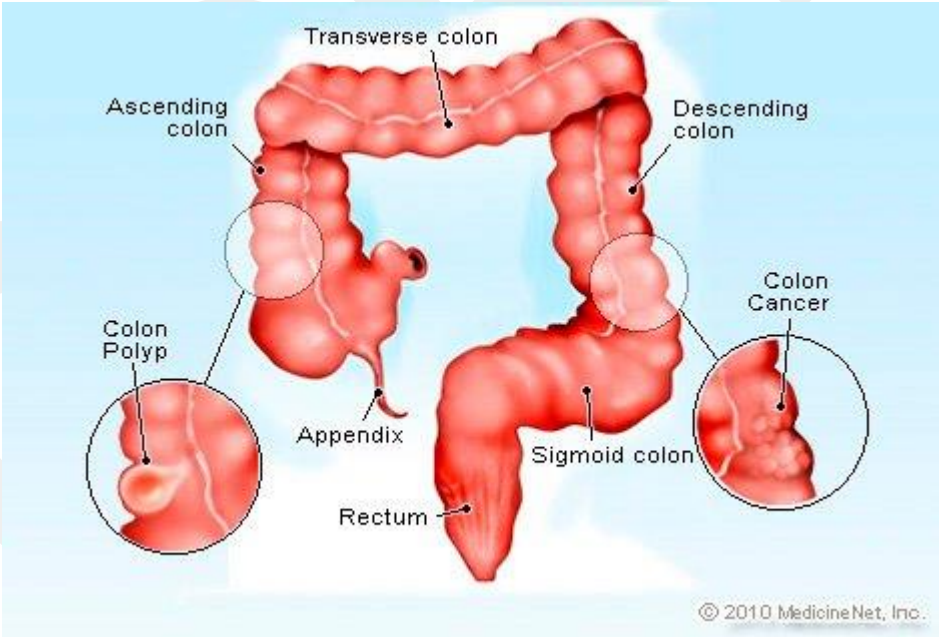


Figure 2.2.2 Polyp and cancer [11].

2.3 Colonoscopy Types

2.3.1 Virtual Colonoscopy

Virtual colonoscopy (VC) is also called computed tomography colonoscopy because it uses X-rays to create the image of colon (Figure 2.3.1.1). This method has some advantages and disadvantages. For example, the procedure takes 10-15 minutes, requires no sedation, and is less invasive than conventional colonoscopy. It is an alternative method for people who cannot have conventional colonoscopy. Besides all these benefits, it has some drawbacks. During this procedure a small amount of air through a short tube is send to the colon in order to perform monitoring properly. This air may cause a rupture in the colon. In addition, this method is only used for diagnosis. If the expert detects a polyp or any other abnormal tissue, s/he will direct the patients to the conventional colonoscopy to remove it. When this method is compared to the other two colonoscopy methods, X-rays are used only in this technique. Although patients are exposed to low dose X-rays, the cancer risk can be increased. Because of these reasons, virtual colonoscopy is not widely used [12].

Figure 2.3.1.1 Virtual colonoscopy and output image (reconstructed) [13], [14].

2.3.2 Wireless Capsule Endoscopy

Wireless capsule endoscopy is the least invasive method among colonoscopy types. In this method a small camera is located in a small capsule and the patient swallow it (Figure 2.3.3.1). This capsule travels through the digestive system and as long as there is no unexpected situation, capsule is removed from the body by defecation. If the capsule cannot be removed with defecation in two weeks, experts remove it with operation. In addition to this, during this procedure people also carry a recording device that is wirelessly connected to the capsule. The number of recorded images is approximately 60,000. Even though this approach is painless and less invasive, it is expensive and records many images. In order to detect any kind of abnormality, experts should choose their area of interest among from whole digestive system. In addition to this, this method does not allow expert to take a biopsy or remove the polyp, and only contains imaging and diagnosis. Although the shape of a polyp or an abnormal structure gives an indication about the disease to experts, they need to send the removed tissues (biopsy specimens) to the pathology department in order to obtain a definitive diagnosis [15].

Figure 2.3.2.1 Wireless capsule endoscopy and output image [16], [17].

2.3.3 Optical or Conventional Colonoscopy

Optical or conventional colonoscopy (CC) is a widely used technique to detect colon diseases. In this method unlike virtual colonoscopy a long tube with a camera and light source at the distal part is used to detect diseases and record colon images (Figure 2.3.3.1). In addition, optical colonoscopy particularly stands out with its catheter, as this catheter allows specialist to remove polyps or abnormalities and send to the pathology department. Moreover, during this process radiation is not used unlike VC [18]. However, optical colonoscopy has some disadvantages such as the long procedure duration which is approximately 40-45 minutes, being invasive and painful, and requiring sedation. However, definitive diagnosis is possible only with this approach. Using optical colonoscopy, experts can take an image or record a video of the abnormal region. Colon has a folded structure thus examination of colon is not easy for experts. Recorded videos or images can be useful to re-evaluate patients' situation after the operation.

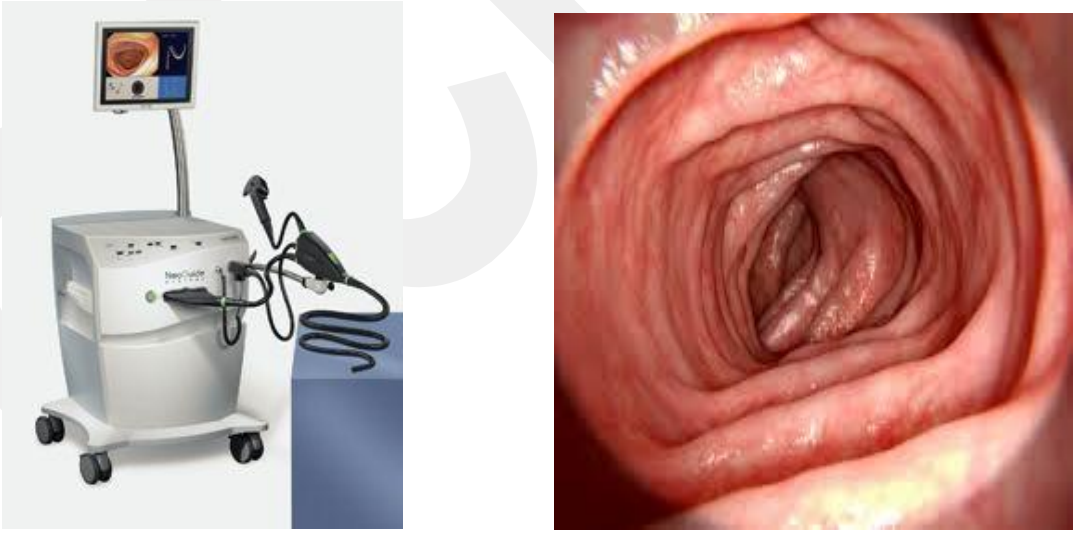


Figure 2.3.3.1 Optical or conventional colonoscopy and one sample colonoscopy image [19],[20].

2.4 Colon Image Processing

2.4.1 Pre-processing

Raw colonoscopy images obtained for diagnosis or other purposes generally need pre-processing because they include different kinds of artifacts. In order to use these images in machine learning (ML) or neural network (NN) algorithms, images should be as clear as possible in order to increase the accuracy of the automatic diagnosis. Pre-processing approaches cannot be grouped under only one title but the most frequently used ones can be listed as follows: Firstly, color conversion can be used to decrease complexity. Studying with RGB images requires a high capacity computer and processing RGB images takes a longer time than grayscale images thus color conversion can be useful. Converting images into grayscale generally does not affect the object or disease detection negatively. Color images are represented with 3 channels but grayscale images with two channels. Redundant pixels are removed in grayscale images and thus complexity decreases. However, in some problems color images are necessary. When one looks at the image and if the image is not affected significantly even if it is grayscale, one can work with grayscale images. Secondly, standardization is an important point to prepare data to further processing steps. Collected images may not always be at the same size because they may be obtained from different devices or sources. Resizing all available images to a standard size of the user's choice frequently solves this problem. Thirdly, data augmentation is another technique for pre-processing. This method is applied to increase the number of data to be used in a study by imposing data variation by scaling, rotation or affine transformation of the available images. Data augmentation may increase the accuracy rate because the model will be trained with different types of images. Lastly, different from above techniques, many methods can be applied in ML and NN to prepare the dataset. For example, edge detection, noise reduction, background removing, smoothing, opening or closing. The pre-processing methodology one will choose depends entirely on the status of the data to be used in further steps [21].

2.4.2 Feature extraction

Pre-processing step is followed by feature extraction part. Feature extraction part is important in order to classify images. Features can be defined as interpretation of images mathematically. Features are extracted to determine different classes [22]. Using raw data is really time consuming and managing it is really difficult especially for colourful images thus features are extracted from images. This method makes image processing easier. Also, managing raw data is not only waste of time but also required high computing sources. However, feature extraction methods allow images to be reduced in size and processed. In addition to these beneficial steps, feature extraction step enables removing redundant information. All of these steps effect learning speed of computer [23]. Many feature extraction methods have been applied in image processing approaches. Feature extraction can be grouped under three headings which are color, shape and texture features. Color space gives information about image using intensity values. Color features can be histogram intersection method, zernike chromaticity distribution moments and color histogram. Shape features can be binary image algorithm, horizontal and vertical segmentation. Texture feature extraction methods are gray level co-occurrence matrix, edge detection, laws texture [24]. Although abnormality detection on colonoscopic images has been a hot topic for a long time, no complete solution has been found yet. In the previous works, polyps, inflammatory bowel diseases, cancer can be listed as some of the abnormalities that emerge on colon tissue. In order to detect polyps, some researchers preferred shape-based detection while some of them selected texture-based approaches [25]-[30]. Furthermore, severity of ulcerative colitis detection is a study that can be found in the literature, they preferred to use local binary pattern method as texture features [31]. In recent years, apart from these studies, researchers have focused on detection of abnormalities that included more than one disease at the same time. For example; Krishnan et al. published a study related to the intestinal abnormality detection focusing on only polyps and tumors. Edge detection and curvature computation methods were used to detect the tumor and the polyp [32]. Gueye et al. worked on wireless capsule images in order to detect abnormalities eliminating non-informative frames manually. They used not only texture-based methods but also shape and color based approaches [33].

In this study, in order to perform feature extraction, we used four popular second order texture analysis approaches called gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), neighborhood gray tone difference matrix (NGTDM), focus measure operators (FMOs) and three first order statistics, such as kurtosis, standard deviation, and skewness. The gray level co-occurrence matrix (GLCM), which describes the relationship between neighbouring pixels, indicates the frequency of image brightness recurrence at a certain distance and direction. If the co-occurrence matrix is denoted by P, the value of P (i, j) specifies how many times i coincides with the value of j in some specified positional relations. i and j are the pixel values. But in general, this distance between pixels is regarded as a "one-pixel distance" and a co-occurrence matrix is formed accordingly. Some of the attributes extracted from the GLCM are energy, autocorrelation, contrast correlation, difference, and homogeneity [34].

$$Energy = \sum_{i,j=0}^{N-1} (P_{ij})^2 \quad \textbf{(1)}$$

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \quad \textbf{(2)}$$

$$Entropy = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \quad \textbf{(3)}$$

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \quad \textbf{(4)}$$

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad \textbf{(5)}$$

$$Kurtosis = \frac{\sum_{i=1}^{n}(X_i - X_{avg})^4}{s^4} \quad \textbf{(6)}$$

$$S.deviation = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad \textbf{(7)}$$

$$Skewness = \frac{\sum_{i}^{N}(X_i - \bar{X})^3}{(N-1)*\sigma^3} \quad \textbf{(8)}$$

In the gray level run length matrix (GLRLM) approach, a set of consecutive pixels with the same gray level value in the specified direction forms a gray level sequence. Run length is the number of pixels in each gray level sequence. The presence of a large number of neighboring pixels at the same gray level represents a coarse-grained texture, while a small number of neighboring pixels have the same gray level represents a finer texture with a faster change. If GLRLM is denoted by P, the value of P (i, j) indicates how many times the gray level i has occurred in length j. This description is for angularly different directions, but the most common use is horizontal direction [35]. The most common of the attributes created using the GLRLM matrix are short-run emphasis (SRE gives greater importance to short sequence lengths of any gray level), long-run emphasis (LRE gives greater importance to long string lengths of any gray level), gray level non-uniformity (GLN, the smallest value when the array lengths are balanced in a balanced manner), run-length uniformity (RLU, takes the smallest value when the sequence lengths are balanced), run percentage (RP), high gray level run emphasis (HGRE, emphasis the high gray level) and low gray level run emphasis (LGRE, emphasis low gray level).

$$SRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)}{j^2}}{N_r(\theta)} \quad \textbf{(9)}$$

$$LRE = \frac{\sum_{i=1}^{N_g}P(i,j|\theta)\,j^2}{N_r(\theta)} \quad \textbf{(10)}$$

$$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{Nr}P(i,j)^2\right)}{N_z} \quad \textbf{(11)}$$

$$RLU = \frac{\sum_{j=1}^{Nr}\left(\sum_{i=1}^{Ng}P(i,j|\theta)\right)^2}{N_r(\theta)^2} \quad \textbf{(12)}$$

$$RP = \frac{N_r(\theta)}{N_p} \quad \textbf{(13)}$$

$$LGRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{Nr}\frac{P(i,j)}{i^2}}{N_z} \quad \textbf{(14)}$$

$$HGRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{Nr}P(i,j)i^2}{N_z} \quad \textbf{(15)}$$

GCRIS

Neighborhood gray tone difference matrix (NGTDM) is a matrix of columns that accommodates elements up to the number of tones and allows one to extract texture-related properties. It was first proposed by Amadasun and King in 1989 and has been used in different problems. The attributes that are extracted from this vector are coarseness, contrast, busyness, complexity, and strength [36].

$$coarseness = \left[\sum_{i=1}^{G\max} \{ p_i . s_i \} \right]^{-1} \quad \textbf{(16)}$$

$$busyness = \frac{\sum_{i=1}^{G\max} \{ p_i . s_i \}}{\sum_{i=1}^{G_{\max}} \{ \sum_{j=1}^{G\max} \{ \left| i.p_i - j.p_j \right| \} \}} ; p_i \neq 0; p_j \neq 0 \quad \textbf{(17)}$$

$$complexity = \left[\sum_{i=1}^{G\max} \left\{ \sum_{j=1}^{G\max} \left\{ \frac{\left| i - j \right| . (p_i . s_i + p_j . s_j)}{n . (p_i + p_j)} \right\} \right\} \right] p_i \neq 0; p_j \neq 0 \quad \textbf{(18)}$$

$$contrast = \left[\frac{1}{N_g . (N_g - 1)} . \sum_{i=1}^{G_{\max}} \left\{ \sum_{j=1}^{G_{\max}} \{ p_i . p_j . (i - j)^2 \} \right\} \right] . \left[\frac{1}{n} . \sum_{i=1}^{G_{\max}} \{ s_i \} \right] \quad \textbf{(19)}$$

$$strength = \frac{\sum_{i=1}^{G_{\max}} \left\{ \sum_{j=1}^{G_{\max}} \left\{ (p_i + p_j) . (i - j)^2 \right\} \right\}}{\sum_{i=1}^{G_{\max}} \{ s_i \}} ; p_i \neq 0; p_j \neq 0 \quad \textbf{(20)}$$

The focus measure operators (FMOs) we used in this study included different types of feature extraction methods that were grouped into six subcategories/families (Table 2.4.2.1). These categories were the gradient operator, Laplacian, the discrete wavelet transform (DWT), image statistics, the discrete cosine transform (DCT) based methods and a category that included miscellaneous approaches. The first subcategory included the DCT based methods based on the energy ratio and the reduced energy ratio. The DCT was employed to detect the images with the motion artifact, because the DCT coefficients depicted information about the spatial frequency distribution of the image. We know that the motion artifact causes the loss of high frequency components on the images [37]. The second subcategory included the DWT based methods such as the sum of wavelet coefficients, the variance of wavelet coefficients, and the ratio of wavelet coefficients. The wavelet coefficients were also preferred to extract information about the spatial frequency components, and especially the edges became more detectable using this approach [38]. The third subcategory consisted of four Laplacian-based methods; the modified Laplacian, the energy of Laplacian, the variance of Laplacian, and the diagonal of Laplacian. This technique was used to emphasize the intensity changes of an image by the help of the second derivative [39]. The fourth feature extraction method was based on the image statistics that consisted of the gray level variance, gray level local variance, normalized gray level, histogram entropy, and the histogram range [40]. The fifth subcategory included the gradient-based methods, which were the Gaussian derivative, the energy of gradient, threshold gradient, squared gradient, Tenengrad and the Tenengrad variance. Gradient-based operators assume that clear or focused images have more edges than the blurred images that is why we used this approach to calculate the first derivative of pixels on the images. In addition to this, Tenengrad was used to calculate the magnitude of the image gradient [41]. The last category was called miscellaneous, and included the absolute central moment, Brenner’s method, contrast, curvature, steerable filters, spatial frequency, Vollath's correlation and the Helmli's mean method [42].

$$EnergyRatio = \frac{\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)^2}{F_{0,0}^2}; (u,v) \neq 0 \quad \textbf{(21)}$$

$$\mathrm{Re}\,ducedEnergyrRatio = \frac{F_{0,1}^2 + F_{1,0}^2 + F_{2,0}^2 + F_{1,1}^2 + F_{0,2}^2}{F_{0,0}^2} \quad \textbf{(22)}$$

$$SumofWaveletCoefficients = \sum_{i,j\in\Omega_D} \left|W_{LH1}(i,j)\right| + \left|W_{HL1}(i,j)\right| + \left|W_{HH1}(i,j)\right| \quad \textbf{(23)}$$

$$VarianceofWaveletCoefficients = \sum_{i,j\in\Omega_D} (W_{LH1}(i,j) - \mu_{LH1})^2 + \sum_{i,j\in\Omega_D} (W_{HL1}(i,j) - \mu_{HL1})^2 + \sum_{i,j\in\Omega_D} (W_{HH1}(i,j) - \mu_{LL1})^2 \quad \textbf{(24)}$$

$$RatioofWavelet = \frac{M_H{}^2}{M_L{}^2}$$

$$M_H^2 = \sum_k \sum_{i,j\in\Omega_D} W_{LHk}(i,j)^2 + W_{HLk}(i,j)^2 + W_{HHk}(i,j)^2 \quad \textbf{(25)}$$

$$ML^2 = \sum_k \sum_{i,j\in\Omega_D} W_{Llk}(i,j)^2$$

$$ModifiedLaplacian = \sum_{i,j\in\Omega(x,y)} \Delta_m I(i,j) \quad \textbf{(26)}$$

$$EnergyofLaplacian = \sum_{i,j\in\Omega(x,y)} \Delta_m I(i,j)^2 \quad \textbf{(27)}$$

$$GrayLevelVariance = \sum_{(i,j)\in\Omega(x,y)} (I(i,j) - \mu)^2 \quad \textbf{(28)}$$

$$GrayLevelLocalVariance = \sum_{(i,j)\in\Omega(x,y)} (L_v(i,j) - \bar{L}_v)^2 \quad \textbf{(29)}$$

$$HistogramEntropy = \sum_{k=1}^{L} P_k \log(P_k) \quad \textbf{(30)}$$

$$HistogramRange = \max(k \,|\, H > 0) - \min(k \,|\, H > 0) \quad \textbf{(31)}$$

$$GaussianDerivative = \sum_{(i,j)} (I * \Gamma_x)^2 + (I * \Gamma_y)^2 \quad \textbf{(32)}$$

$$GradientEnergy = \sum_{(i,j)\in\Omega(x,y)} (I_x(i,j)^2 + I_y(i,j)^2) \quad \textbf{(33)}$$

$$ThresholdGradient = \sum_{(i,j)\in\Omega(x,y)} |I_x(i,j)|, |I_x(i,j)| \geq T \quad \textbf{(34)}$$

$$Tenengrad = \sum_{(i,j)\in\Omega(x,y)} (G_x(i,j)^2 + G_y(i,j)^2) \quad \textbf{(35)}$$

$$TenengradVariance = \sum_{(i,j)\in\Omega(x,y)} (G(i,j) - \bar{G})^2 \quad \textbf{(36)}$$

$$AbsoluteCentralMoment = \sum_{k=1}^{L} \lfloor k - \mu \rfloor P_k \quad \textbf{(37)}$$

$$Brenner'sFocusMeasure = \sum_{i,j} |I(i,j) - I(i,2j)|^2 \quad \textbf{(38)}$$

$$ImageContrast = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} |I(x, y - I(i,j))| \quad \textbf{(39)}$$

$$ImageCurvature = |c_0| + |c_1| + |c_2| + |c_3|$$

$$c_0 = M1 * I$$

$$c_1 = M_1^T * I$$

$$c_2 = \frac{3}{2} M_2^T * I - M_2^T * I$$

$$c_3 = \frac{3}{2} M_2^T * I - M_2^T * Irevx$$

$$M_1 = \frac{1}{6} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

$$M_2 = \frac{1}{5} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad \textbf{(40)}$$

GCRIS

$$Helmli\,\&\,Scherer'sMeanMethod = \begin{cases} \dfrac{\mu(x,y)}{I(x,y)}, \mu(x,y) \geq I(x,y); \dfrac{I(x,y)}{\mu(x,y)}, otherwise \end{cases} \quad \textbf{(41)}$$

$$SteerableFilters = \sum_{i,j \in \Omega(x,y)} I_f(i,j) \quad \textbf{(42)}$$

$$SpatialFrequencyMeasure = \sqrt{\sum_{i,j} I_x(i,j)^2 + \sum_{i,j} I_y(i,j)^2} \quad \textbf{(43)}$$

$$Vollath'sAutocorrelation = \sum_{(i,j) \in \Omega(x,y)} (I(i.j), I(i+1.j)) - \sum_{(i,j) \in \Omega(x,y)} (I(i.j), I(i+2.j)) \quad \textbf{(44)}$$

GCRIS

Table 2.4.2.1 The subfamilies of Focus Measure Operators (FMOs).

Method Types	***Used Features***	
DCT-BASED	Energy Ratio	Reduced Energy Ratio
WAVELET-BASED	Sum of Wavelet Coefficient	Variance of Wavelet Coeff.
	Ratio of Wavelet Coefficient	
LAPLACIAN-BASED	Modified Laplacian	Energy of Laplacian
	Variance of Laplacian	Diagonal Laplacian
STATISTICS-BASED	Gray Level Variance	Normalized Gray Level
		Histogram Entropy
	Gray Level Local Variance	Histogram Range
GRADIENT-BASED	Gaussian Derivative	Energy of Gradient
	Threshold Gradient	Squared Gradient
	Tenengrad	Tenengrad Variance
MISCELLANEOUS	Absolute Central Moment	Steerable Filters
	Brenner’s Method	Spatial Frequency
	Contrast	Vollath's Correlation
	Curvature	Helmli's Mean Method

2.4.3 Classification

Classification is a supervised learning technique. In the classification part of image processing, data should be divided into three parts which are training, testing and validation. This division could be 80%, 10%, and 10% or 70%, 15%, and 15% respectively. The majority of the data should be a part of the training [43]. In this phase, you give training data to the system and this system learns given data considering their labels. Thus, a model or system is trained. After that, using this trained model, label of validation set is predicted. Later, comparing predicted label and validation set label, performance metrics are obtained (accuracy, sensitivity, specificity etc.). These values are called validation performance metrics. Using this validation dataset, the best model is selected and recorded [44]. The percentage of validation set can be decided based on the number of hyperparameters. Hyperparameters play an important role in training model because model structure is created based on these hyperparameters. Using validation set, the best hyperparameters are selected and the best model is recorded [45]. If the number of hyperparameters is low one can decrease the percentage of validation set. Training part is used to create a suitable model by giving the system the associated classification labels (or classes), and the system learns which label belongs to which data type. After the system is trained and the model is formed, validation step is initiated. In that step dataset generally consists of 10 to 15% of all the available data and this dataset is used to improve and update the system hyperparameters. According to the validation accuracy or other metrics, model can be improved in order to get reliable results. After the model is trained and hyperparameters are updated/optimized, testing step is employed. In this step the images that the system has not seen before are used to predict their labels/classes using the available model and hyperparameters. These predicted labels are compared with the actual ones, and the performance metrics are calculated [43]. The most common used classifier types are; logistic regression, naive Bayes, support vector machines, decision trees, random forest and support vector machine. The explanation of classifiers is given below;

- The Naïve Bayes classifier is based on the ground of Bayes' theorem. It can work on unbalanced datasets. This algorithm calculates the probability of each state and classifies according to the highest probability value. It can reach high values with a small dataset. If a particular value is never observed for a discrete feature in training set but in test set the model assigns zero probability to test samples that contains this value. It cannot predict and this situation is known as zero frequency. In order to solve this problem some correction technique can be used such as Laplacian smoothing prediction. The usage areas of this method are real time prediction, multiclass prediction, and text classification [46].

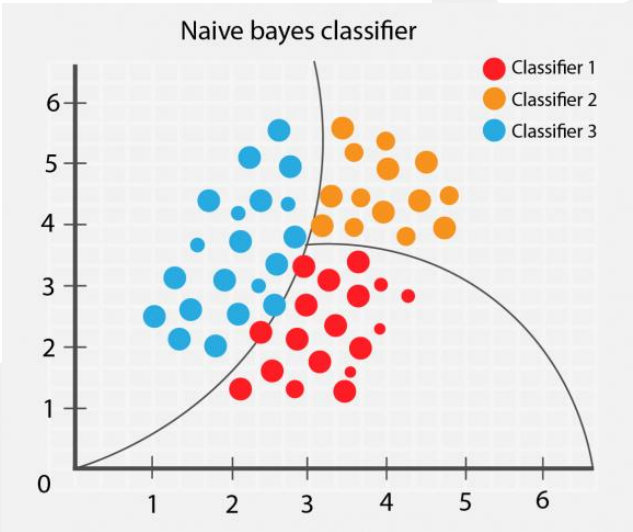


Figure 2.4.3.1 Naïve Bayes classifier [47].

- Logistic regression is a regression method which is used to perform classification on a dataset. It works if dependent variable or the results, can take 2 different values such as yes/no, man/woman. It is preferred for linear classification problems frequently thus it is similar to linear regression [48].

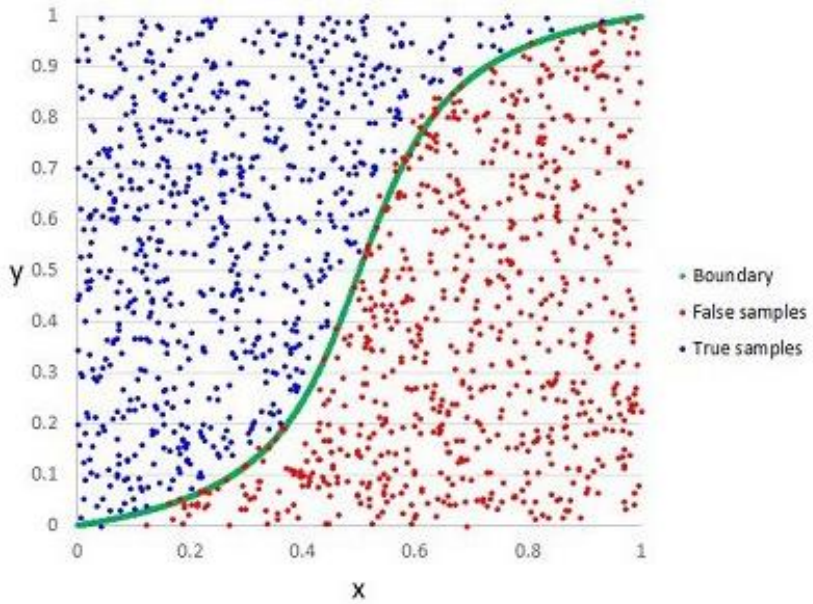


Figure 2.4.3.2 Logistic regression classifier [49].

- Decision tree which is used for both classification and regression problems is one of the most popular models in machine learning. The purpose is to create a model that estimates the value of a variable by extracting simple rules from data properties and learning these rules. Decision tree models can be modified to handle missing values [50].

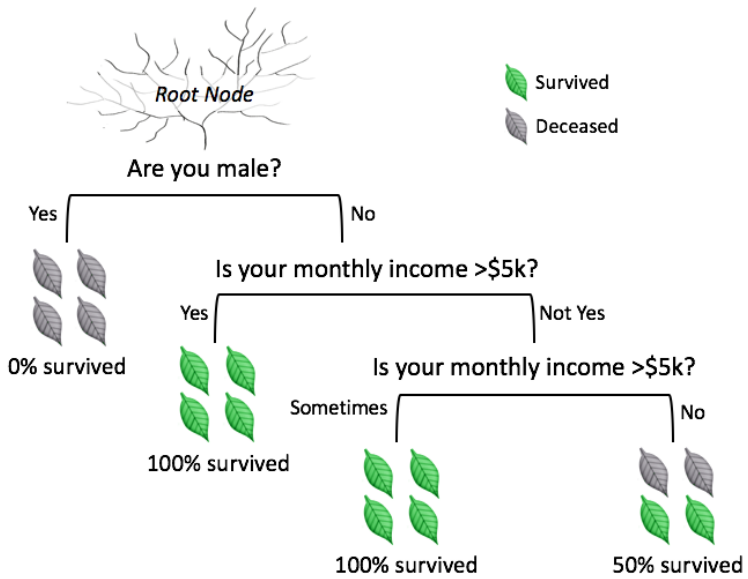


Figure 2.4.3.3 Decision tree classifier [51].

- Random forest classification method is similar to decision tree algorithms. It can be used for both classification and regression analysis like decision tree. The working principle of random forest is creating more than one decision tree and while producing a result, the estimates in the decision trees are averaged. In addition to this, random forest algorithm is an ensemble method [52].

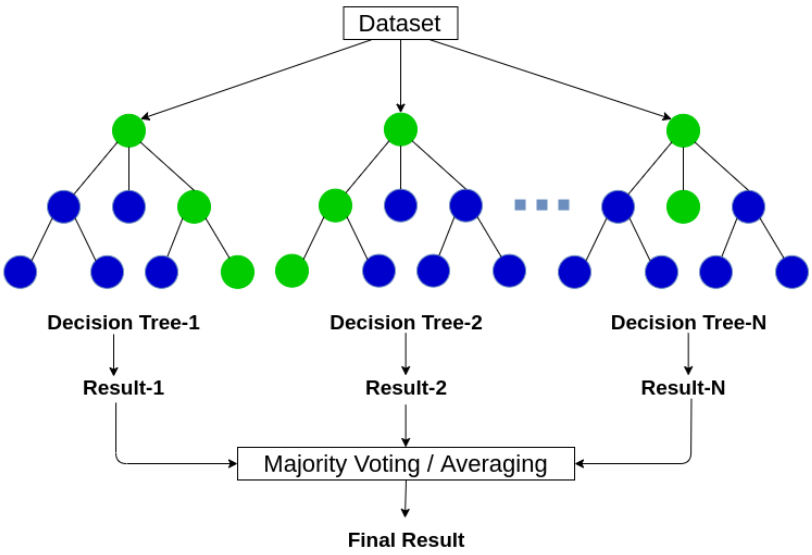


Figure 2.4.3.4 Random forest classifier [53].

- k-NN is the most simple and preferred classification method. It is a non-parametric and lazy learning algorithm. Lazy learning means that there is no learning step in the training phase. This system does not learn from the training set, it memorizes training set. When you need to predict, it searches the closest neighbor from all datasets. In this algorithm a K value is determined and the meaning of this value is the number of neighbors to search. When a value is determined, the distances between test example and training examples are computed and prediction is made based on the closest K samples in training set. Euclidean function is generally used in distance calculation. However, Manhattan, Minkowski and Hamming functions can also be used to measure distance. After calculation of the distance, it is ordered and majority voting is used to assign a suitable class [54].

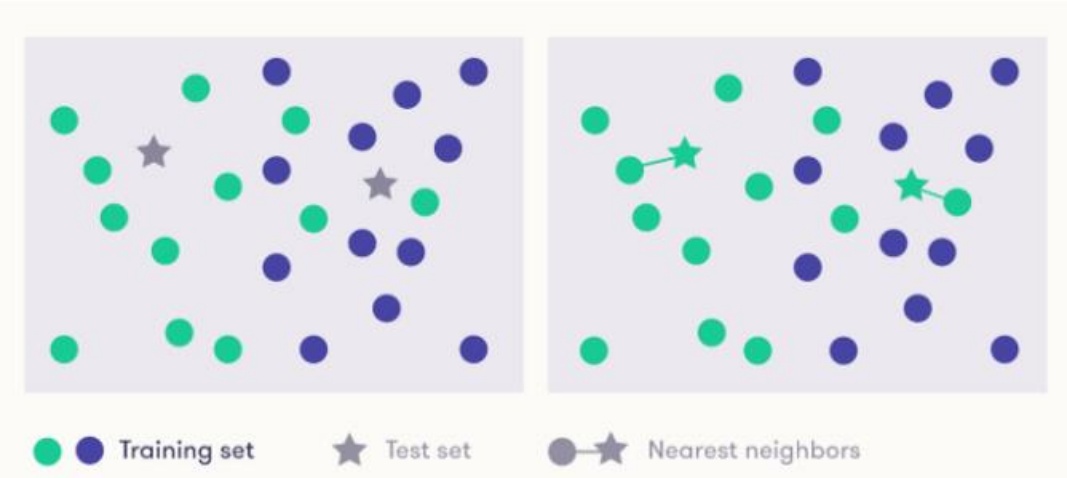


Figure 2.4.3.5 k-NN classifier [55].

- SVM is a similar classification algorithm with Logistic Regression. Both of them try to find the best line in order to separate two classes. This line passes from the furthest place of elements that belongs to two classes. It is a nonparametric classifier. Although SVM can classify both linear and non-linear datasets, it generally tries to classify data linearly but, in some cases, it is not possible. In order to get rid of this situation kernels are used. If a new dimension is created, linear classification can be possible [56].

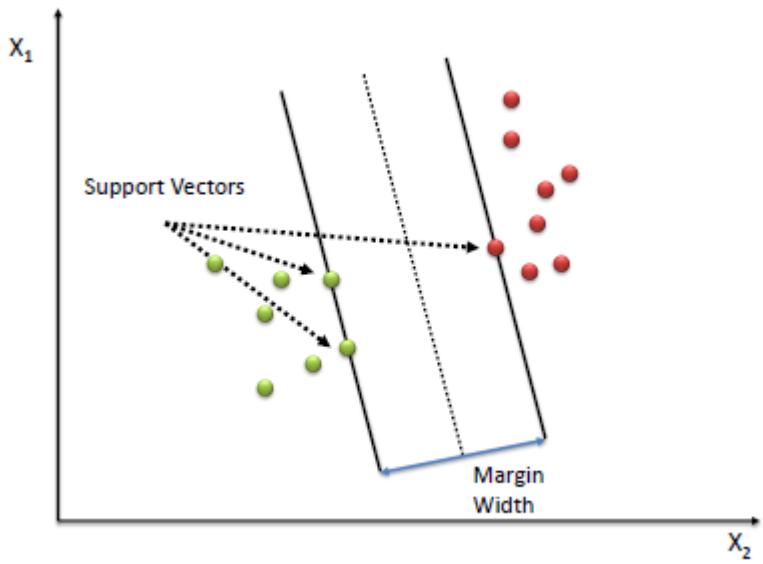
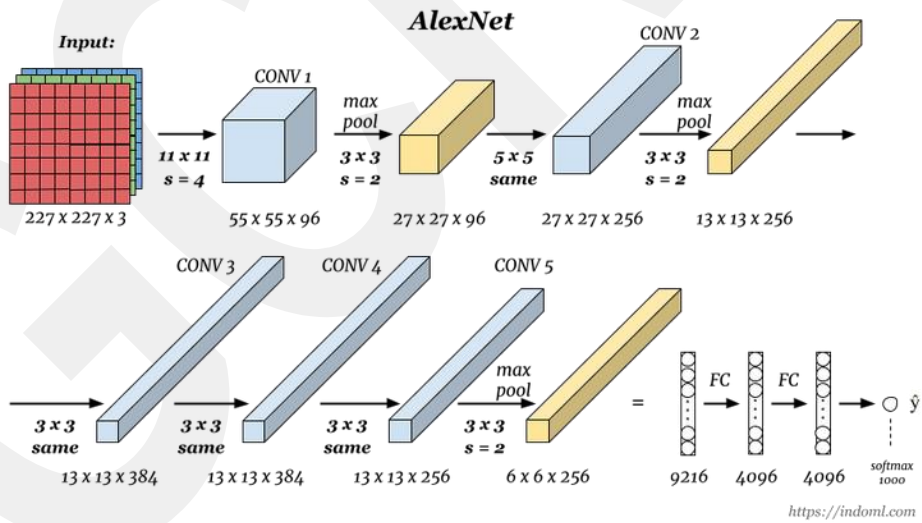


Figure 2.4.3.6 SVM classifier [57].

2.4.4 Transfer Learning

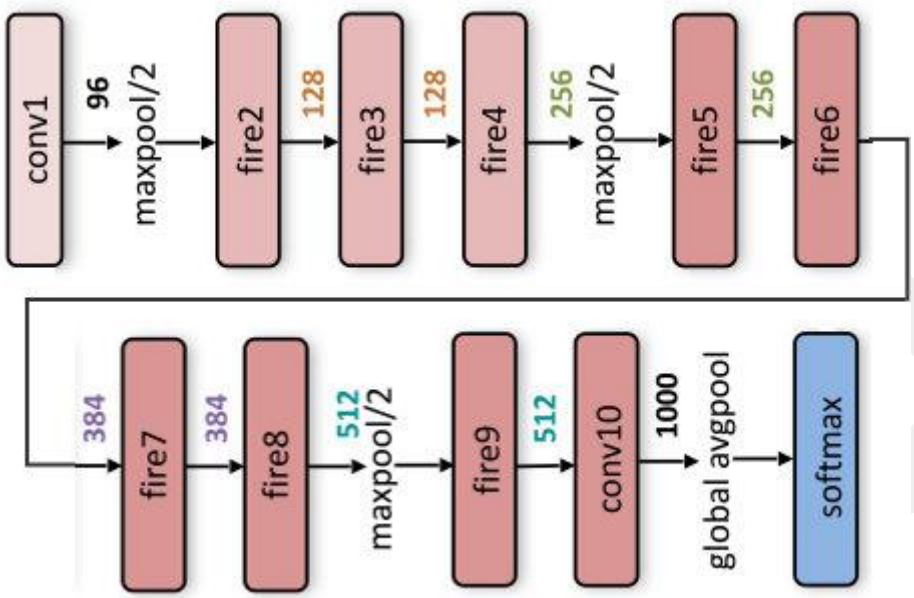
Training of some models is almost impossible to implement on standard computer processors due to model complexity or the size of the dataset. Therefore, graphics processing units (GPU) are needed. These trained models can be used in various ways to solve different problems as a result of trainings that take days and sometimes weeks. This is exactly what is called 'Transfer Learning'. For example, if the dataset is not large enough, transfer learning has advantages because this system is trained previously on huge amount of dataset such as ImageNet dataset (15 million). Transfer learning provides faster solutions to many problems in artificial intelligence studies. Different types of transfer learning methods are explained below [58].

AlexNet is a transfer learning method that includes 8 layers. This network is pretrained with more than one million images from ImageNet database. Using this model, images can be classified into 1000 categories using 60 million parameters. Input size of images for this network should be 227×227 [59].



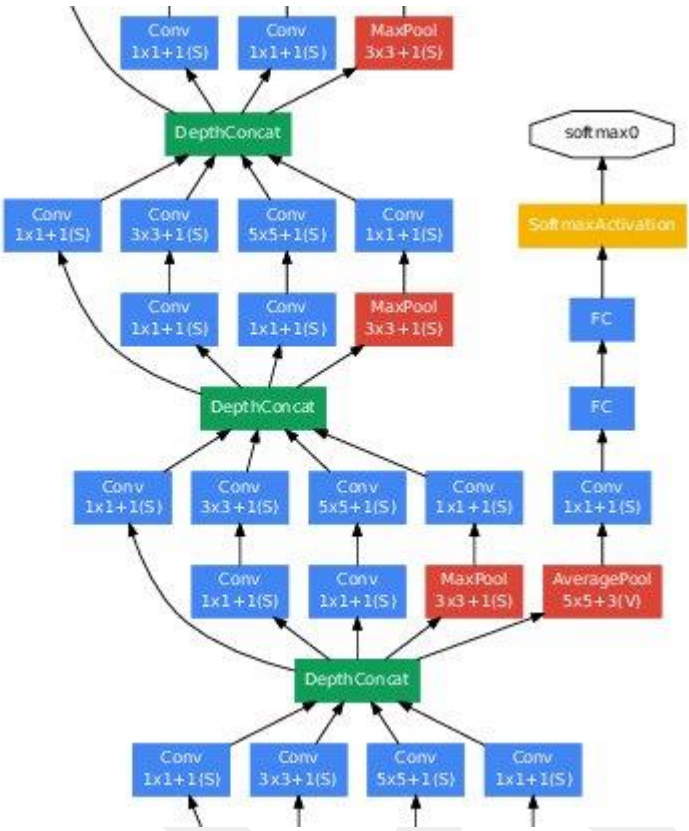
2.4.4.1 AlexNet architecture [59].

SqueezeNet is another transfer learning model which is pretrained with more than one million images from ImageNet database as AlexNet, however, this model has 18 layers. In addition to this, SqueezeNet is 3 times faster than AlexNet. In order to train the model using SqueezeNet, size of images should be 227×227 [60].



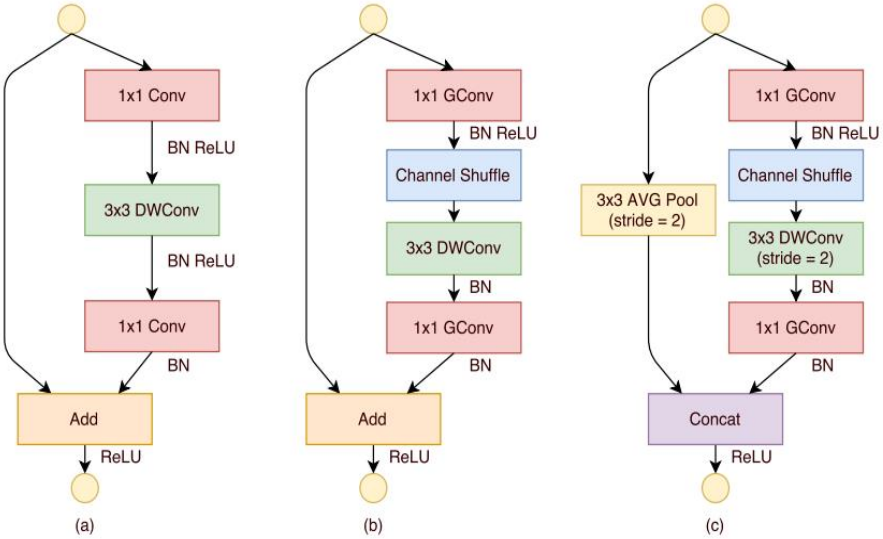
2.4.4.2 SqueezeNet architecture [60].

GoogLeNet network is trained with two different databases. One of these databases is ImageNet, the other one is Place365. When model is trained with ImageNet, it can classify images into 1000 different categories, however model with trained with Place365 can classify images into 365 categories with 4 million parameters. The number of layers is 22 and the size of input images is 224×224 [61].



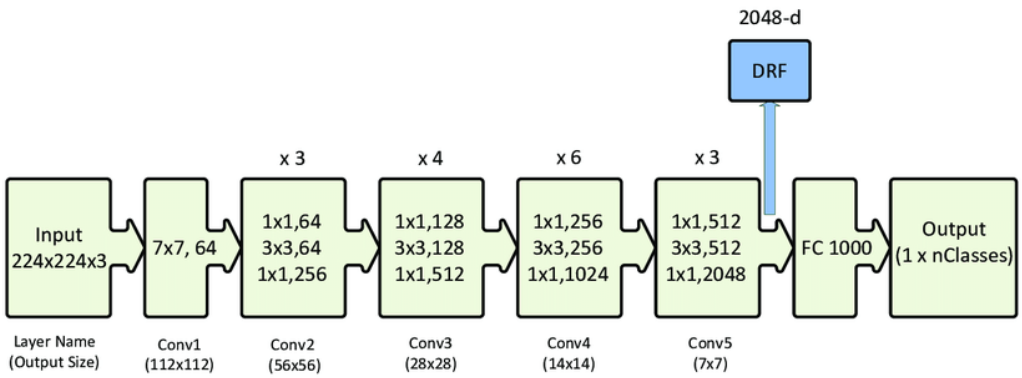
2.4.4.3 GoogLeNet architecture [61].

ShuffleNet is another kind of convolutional neural network type. This model is improved especially for mobile devices. This network uses two proposed operations that are pointwise group convolution and channel shuffle. This model decreases computation cost significantly. This model is trained with ImageNet dataset and size of input image is same as with GoogLeNet (224×224) [62].



2.4.4.4 ShuffleNet architecture [62].

ResNet stands for residual network. This network is trained with ImageNet and size of input images is 224×224. However, ResNet can be used as ResNet-18, ResNet-50 or ResNet 101. The numbers show the number of layers of used in the models [63].



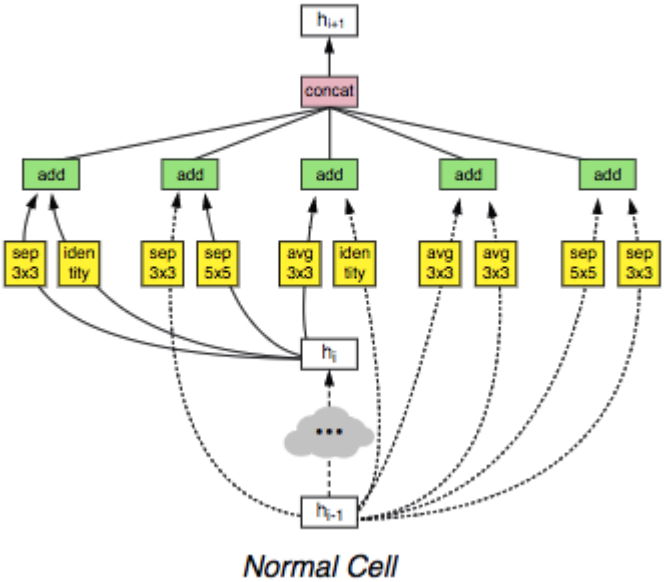
2.4.4.5 ResNet architecture [63].

MobileNet is created by using 53 layers. ImageNet database is used to train this model previously. Computation power of MobileNet is less than other transfer learning models that is why this model is suitable for embedded systems, mobile devices or computers without GPU [64].

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5\times$ Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
$5\times$ Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

2.4.4.6 MobileNet architecture [64].

NasNetMobile is another type of artificial neural network. NasNet stands for Neural Architecture Search. This model includes 913 layers, and the size of input image is 224×224. This model is pretrained using ImageNet database, thus this model can be useful for multiclass problems [65].



2.4.4.7 NasNetMobile architecture [65].

Chapter 3

3 Study-1

3.1 Effect of Interpolation on Specular Reflection in Texture Based Automatic Colonic Polyp Detection

Colon cancer develops in large intestine which is the last part of the digestive system. Many colon cancer cases begin with adenomatous polyps called small and benign cells [66]. Adenomatous polyps refer to abnormal growth of the tissue. Studies show that CC is usually caused by pre-existing adenomatous polyps [67]. According to the study of the American Cancer Society, CC is the third most common type of deadly cancer. Most CCs are composed of 'silent' tumors. These are slow-growing tumors that generally do not give any indication until they reach a large size. CC can be prevented and treated if it is diagnosed early [68]. Therefore, the polyps that the specialists detect are removed from the patient regardless of whether they are malignant or benign so as not to create future risk. The most common imaging methodology used for this procedure is called ‘colonoscopy.’ Colonoscopy allows examination of the intestines with the light and camera on the tip, but also the removal of polyp with the help of the catheter at the end. Colonoscopy is an operator-dependent process; thus, attention deficiency or fatigue of the operator can cause missed polyps. The polyps which are not detected early enough can turn into a cancerous structure, and after many years the disease can be diagnosed as an advanced cancer, in which the survival rate is lower than 10% [69]. In addition to the colonoscopy systems, in recent years, wireless capsule endoscopy (WCE) has been developed for human-independent computer-aided screening. In this method, the camera is placed in a vitamin-sized capsule. As the capsule travels along the digestive tract, thousands of images are captured by the recorder connected to the patient's waist.

Although this procedure seems more advantageous for the patient, it is not very efficient in terms of examining about 60,000 images. The examination and processing of these images is time consuming; it has not yet become a preferred method [66].

3.2 Literature review

Computer-aided polyp detection for colonoscopic and WCE images can be performed according to the polyp shape or texture. In the literature, there are many studies focusing on shape-based [70,71,72,73,74] and texture-based polyp detection. Shape-based polyp detection generally uses geometrical shape, appearance or boundaries of polyps [70,75,76]. Texture is one of the most important characteristics used to describe the region of interest on the image. They are the measures of intensity variation of a surface that determine properties such as smoothness, roughness, and regularity. Previous works on colonoscopic images show that texture-based methods are more popular. For example, Wang et al. used local binary pattern (LBP) approach [77], Tjoa et al. employed color histogram and texture spectral features [78], and Alexandre et al. proposed a method using color space features [79]. In addition to these studies, the texture analysis approach like gray level co-occurence matrix (GLCM) was preferred in numerous publications [80, 81, 82]. These approaches use pixel intensities on the images. In order to get reliable results, quality of the image should also be high. The quality can be deteriorated by the light reflections, date or patient name information on the colonoscopic images. Especially, reflections of light which occur due to the LED light source used to illuminate the scene to visually examine the colon cause unwanted noise effects called ‘specular reflection.’ These undesired reflections occur on the polyp or other colon tissue, and affect the texture features obtained from the image. In order to perform texture-based polyp detection successfully the effect of specular reflection should be minimized. Previously, several researchers proposed approaches to eliminate reflection on images. For example, Guo et al. suggested a method to eliminate specular reflection for endoscopic images. In that study, they proposed two steps: (1) Detection of reflection using thresholding and (2) elimination of reflection using an inpainting algorithm [83]. Among other methods proposed in the literature, Stehle used spectral deconvolution algorithm to remove reflections [84], Arnold et al. [85] and Karapetyan et al. [86]

preferred inpainting algorithms which can also be called as the image interpolation technique. In general, image interpolation approaches compute new values for pixels whose intensity values are saturated due to specular reflection using adjacent/surrounding pixels that have normal intensity values. As it has many types, in this study we used three different interpolation methods: bilinear, nearest neighbor and bicubic interpolation.

In machine learning literature, texture-based automatic detection of colonic polyps involves extracting texture features from the images or parts of images (subimages) and applying classification techniques on these features to discriminate whether it includes a polyp structure or not. The aim of this study was to seek answers to the following two questions. (1) How are the texture features used in automatic detection of polyps affected by the interpolation on specular reflections? (2) If they are affected by the interpolation approach does it really affect the classification performance? In order to answer these questions, we obtained texture features from colonoscopic images with no specular reflections and the same features obtained from the same images with synthetically added reflections with various sizes, and interpolation applied. The interpolation approaches we used were the nearest neighbors, bilinear and bicubic interpolation. For answering the second question, we used t-test to investigate whether these interpolation methods caused any difference in terms of classification performance to discriminate polyps from the colon background. We performed automatic classification of polyps and background using random forest and k nearest neighbors (k-NN) approaches.

3.3 Method

In this study, we employed colonoscopy images coming from the “CVC-ClinicDB” database prepared in the Hospital Clinic of Barcelona, Spain [5]. The images of this database were acquired using white light conventional colonoscopy system. It is the official database used in the training stages of MICCAI 2015 Sub-Challenge on Automatic Polyp Detection Challenge in Colonoscopy Videos. The dataset includes 610 colonoscopic images which were obtained from 29 patients. Approximately 20-25 images were obtained from different angles from each patient. Each image contains polyps in different shapes, size, and numbers. In addition, ground truth images showing the location

of the polyps were provided. Sample polyp and ground truth images are given in Figure 3.3.1 [5].

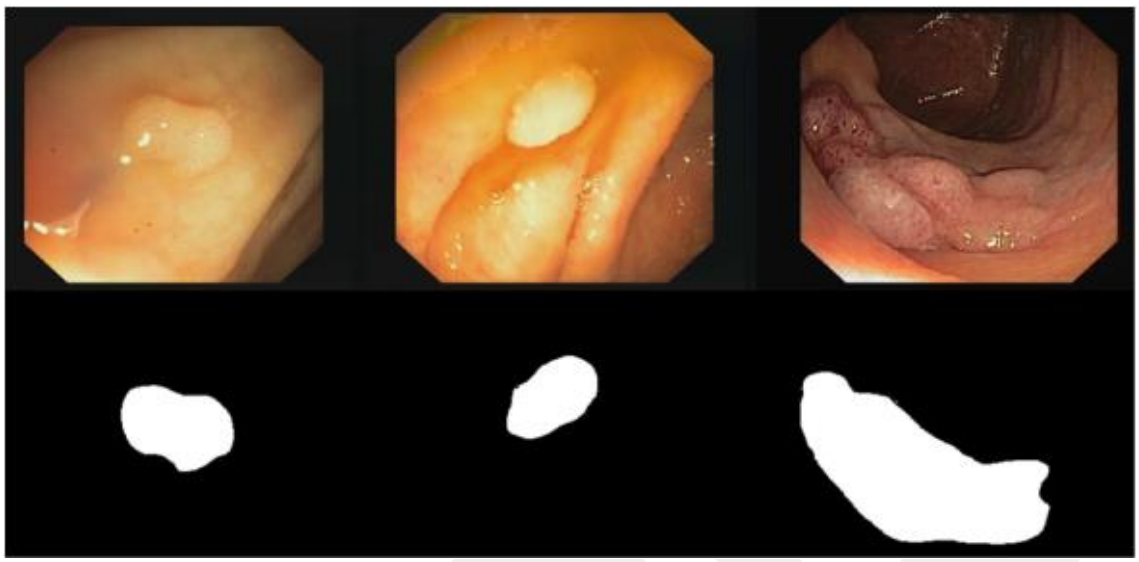


Figure 3.3.1. Three different polyps and their ground truth counterparts from CVC-Clinic DB database [5].

The size of the images used in this study was 288x384 (Figure 3.3.2). Because, the images include black parts on the edges that do not contain any information, we selected a rectangle from the initial images to result in 224x192 images (Figure 3.3.3) to be used in the subsequent phases of our study. First of all, we have converted these images into gray scale. For analysis and labeling purposes, we, then, divided each image into 32x32 squares which we referred to as "**tiles**." We obtained a total of 42 tiles from each image (Figure 3.3.4). In the final phase of this study, which we discussed later in this thesis, we automatically classified/discriminated healthy and polyp tissues. In that phase, we only needed tiles that contained pure heathy and pure polyp tissues. However, as shown in Figure 3.3.4, several tiles included both healthy and polyp tissues, such as tiles numbered as 10, 19, and 23. If tiles were pure healthy, i.e., there were no pixels corresponding to polyps or specular reflections, these tiles were labeled as healthy. For tiles with specular reflections reader should refer to Figure 3.3.5. It shows that there are spots that correspond to the specular reflection on some of the tiles like tile 21 and 27. Likewise, if a tile was labeled as pure polyp, that means there were no pixels corresponding to specular reflections or healthy tissues. For the determination of polyp tissue, we used ground truth images and logic operations, and for the specular reflections, we used a simple

thresholding approach. As a result of this process, we obtained a total of 15674 tiles from 610 images. The number of tiles that were labeled as pure polyp and pure healthy was 1426 and 14248, respectively.

Figure 3.3.2 Original image

Figure 3.3.3 Cropped image.

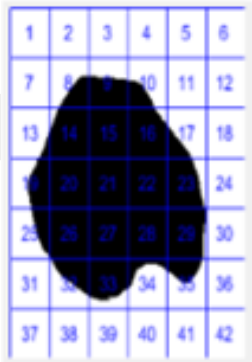


Figure 3.3.4 Ground truth tiles.

Figure 3.3.5 Tiles with reflection.

In order to investigate the effect of specular reflections on polyp detection we added real reflections on tiles labeled as pure polyp and pure healthy. Real reflections are added to answer effect of different size of reflection. In order to investigate the effect of real reflections, we mean that we selected four different sized “actual reflections” from tiles that contained specular reflections (which were not pure tiles). Figure 3.3.6 shows four real reflections we used in our subsequent analysis which corresponded to the reflections covering 2%, 10%, 20%, and 30% of the tiles. For example, 2 and 10 percent reflections mean that out of 1024 pixels (32x32 pixels in one tile) approximately 20 and 100 pixels, respectively, were replaced with pixels containing specular reflection.

Figure 3.3.6 Four different size “real” reflections added on the tiles corresponding to the reflections covering 2%, 10%, 20%, and 30% of the pixels in each tile (32x32 pixels).

In this study, we tested the effect of different image interpolation approaches on texture features which were used in the automatic classification of polyps. The image interpolation refers to interpolating (by fitting a function) pixel values corresponding to specular reflections using the pixel values without reflection. The concept we are referring here is also called inpainting.

In this part, we applied two-dimensional (2D) nearest neighbor, bilinear, and bicubic interpolation approaches on healthy and polyp tiles (a total of 15674 tiles) on which specular reflections were previously added as explained above. We used a built-in function called ‘fillmissing’ in MATLAB for image interpolation. We were able to choose and apply each interpolation methodology separately using this function. In several different studies [82,83,84,85] researchers reported results from various interpolation approaches to remove specular reflection from endoscopic images. In the nearest neighbor interpolation, the empty spaces will be filled in with the nearest neighboring pixel value. In the bilinear interpolation, the main idea is to perform linear interpolation first in one direction, and then in the other direction. Although each step is

linear in the sampled values and in the position, the interpolation as a whole is quadratic in the sample location [86]. The bicubic interpolation is a method that uses four neighboring pixels to fill the missing part of the image. For each of the neighboring four data points, we need to know its intensity value, its partial derivatives along both axes, and its cross-derivatives [87]. In this study, we have not used interpolation to fill missing values but to update pixels with reflections.

Image texture is one of the most important characteristics used to describe region of interest in an image. They are the measures of intensity variations of a surface that determine properties such as smoothness, roughness, and regularity. Texture based feature extraction methods are categorized into first order and second order techniques. First order statistics does not use neighborhood relationships, but second order statistics use neighborhood relationships.

After preprocessing, addition of specular reflections and interpolation of images with reflections, the next step was to extract image texture features. Actually, we have extracted features from the tiles with no specular reflections and from the same tiles but reflection added. By this way we wanted to investigate *the effect of interpolation on the tiles with specular reflections in terms of texture features* and the *discrimination of healthy tiles and tiles with polyps using the extracted features*. A total of 116 texture features were extracted using the approaches above for each tile with no reflections, 2%, 10%, 20% and 30% reflections added and interpolated. Totally, 116 features which are obtained by GLCM, GLRLM, NGTDM were used for texture analysis.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad \textbf{(45)}$$

In order to investigate the effect of image interpolation on texture features, we performed a t-test using SPSS after performing a W-test for normality check. The statistical significance level (p) was set to 0.05. The aim of this step was to detect any significant differences between texture features obtained from reflection free healthy tiles and texture features obtained from reflection added and interpolated healthy tiles. This analysis was performed four times. Firstly, we started statistical analysis by comparing reflection free

healthy tiles and 2% reflection added and interpolated healthy tiles. Next, this step was repeated for reflection with different sizes (10, 20 and 30% of the tile was covered with reflection). The same procedure was applied for tiles with polyps. This analysis was useful to determine robust features which were affected minimally from the image interpolation approaches we have implemented in this study.

In the final phase of this study, we have tested two different classification methods to automatically discriminate the healthy tiles from the tiles with polyp using texture features. To remind again we had a total of 15674 tiles from 610 images. The number of tiles that were labeled as polyp and healthy was 1426 and 14248, respectively.

The classification methods used here were k nearest neighbors (k-NN) and random forest. The aim of this part was to investigate the effect of interpolation on the texture based classification process. For this purpose, we used both reflection free tiles and reflection added and interpolated tiles. First, using the WEKA software the random forest and k-NN classification methods were applied to discriminate healthy tiles with no reflection from tiles with polyp with no reflection. The number of trees for the random forest approach was 100 and k value was set to 1 for k-NN classification. Later, we continued with the classification of reflection free healthy tiles and 2, 10, 20 and 30% specular reflection plus interpolation on polyp tiles. The same procedure was followed to perform classification between reflection free polyp tiles and 2, 10, 20, and 30% specular reflections plus interpolation on healthy tiles. We performed a 10-fold cross-validation for reducing the bias.

In the performance analysis of classification methods, we used overall accuracy and f-measure. f-measure is selected as the performance metric, because it is useful for class-imbalance problems like this one. The definitions of metrics known as precision, recall, and F-measure are given below in Eqs. 46-48. Precision and recall are measures for correctness and completeness, respectively. The f-measure takes the precision and the recall into account when computing the score. It can be interpreted as a harmonic mean of precision and recall.

$$Precision = \frac{TP}{TP+FP} \quad \textbf{(46)}$$

$$Recall = \frac{TP}{P} \quad \textbf{(47)}$$

$$f - measure = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} \quad \textbf{(48)}$$

3.4 Results

As mentioned above, the t-test was used to investigate the effect of interpolation on texture features. For this purpose, features obtained from reflection free polyp tiles and interpolated polyp tiles that included different size reflections were statistically compared. This process was carried out for bilinear, nearest neighbor, and bicubic interpolation techniques separately. In addition, the same procedure was followed for heathy tiles. When the t-test was applied on reflection free (healthy and polyp) tiles and tiles that were interpolated after 2% specular reflection was added, we observed that the number of features affected significantly ($p<0.05$) was between 0 to 8 out of 116 features for all interpolation techniques. In addition, the number of features affected significantly by interpolation was between 15 and 86 (12-70% of all features) for 10% reflection added and interpolated tiles. As expected, the number of features that was affected by interpolation increased when the size of reflection increased. Table 3.4.1 depicts the results of this part. In summary, if an image includes 2% specular reflection, any interpolation technique can be effective to eliminate reflection without changing texture features significantly. However, if the reflection percentage is over 10%, using interpolation can cause deformation on texture structure. We should note that there were 4 features which were same for both polyps and healthy tiles, which did not change significantly from added reflection and interpolation. Three of these robust features came from the gray level co-occurrence matrix as autocorrelation, sum of squares (variance), sum of average, and energy features obtained using the first order method.

The results of the second part of this study were summarized in Table 3.4.2. The f-measure values were computed and given for the comparison of interpolation methods

and classification approaches to automatically discriminate tiles with polyps from the healthy background tiles with and without specular reflection of different sizes. For both classification approaches f-measure values were higher for bicubic interpolation when compared to other two methods. As the percentage of the specular reflection increased, f-measure values increased except for two cases (bicubic interpolation increasing from 20 to 30% reflection). Random forest was found to be performing better than k-NN in all cases. Table 3.4.3 shows the accuracy of the best f-measure results for bilinear, nearest neighbor and bicubic interpolation methods and random forest and k-NN classification approaches. The accuracy levels were approximately 90% for the discrimination of polyp tiles from the healthy tiles without any specular reflection. The accuracies went up to ~99% as we increased the reflection percentages for both classification approaches.

Table 3.4.3 shows the accuracy of the best f-measure results for bilinear, nearest neighbor and bicubic interpolation methods and random forest and k-NN classification approaches. The accuracy levels were approximately 90% for the discrimination of polyp tiles from the healthy tiles without any specular reflection. The accuracies went up to ~99% as we increased the reflection percentages for both classification approaches.

GCRIS

Table 3.4.1 Summary of t-test results for the comparison of features obtained from reflection free tiles and interpolated tiles that included different size reflections.

Reflection Free Tiles	**Tiles with Specular Reflection plus Interpolation**	**Number of Significantly Different Features ($p<0.05$)**		
		Bilinear	**NN**	**Bicubic**
Reflection Free Polyp	Polyp 2% reflection	0	0	1
Reflection Free Polyp	Polyp 10% reflection	15	21	82
Reflection Free Polyp	Polyp 20% reflection	29	28	83
Reflection Free Polyp	Polyp 30% reflection	30	66	82
Reflection Free Healthy	Healthy 2% reflection	4	6	8
Reflection Free Healthy	Healthy 10% reflection	31	30	86
Reflection Free Healthy	Healthy 20% reflection	44	64	95
Reflection Free Healthy	Healthy 30% reflection	51	79	93

Table 3.4.2 Summary of f-measure values for the comparison of interpolation methods and classification approaches to automatically discriminate tiles with polyps from the healthy background tiles with and without specular reflection of different sizes.

Reflection Added on **Healthy** Tiles (%)	Reflection Added on **Polyp** Tiles (%)	Random Forest			k-NN		
		f-Measure					
		BL	NN	BC	BL	NN	BC
0	0	0.376	0.376	0.376	0.389	0.360	0.360
0	2	0.376	0.383	0.384	0.352	0.357	0.352
0	10	0.376	0.391	0.835	0.358	0.374	0.734
0	20	0.406	0.474	0.956	0.407	0.420	0.931
0	30	0.450	0.675	0.872	0.420	0.576	0.786
2	0	0.389	0.390	0.379	0.370	0.366	0.365
10	0	0.411	0.429	0.829	0.377	0.389	0.731
20	0	0.490	0.571	0.942	0.427	0.464	0.893
30	0	0.585	0.758	0.864	0.478	0.612	0.736

Table 3.4.3. Accuracy of the best f-measure results for bilinear, nearest neighbor, and bicubic interpolation methods.

Interpolation Types	Reflection Added on **Healthy** Tiles (%)	Reflection Added on **Polyp** Tiles (%)	**Overall Accuracy %**	
			Random Forest	**k-NN**
No Interpolation	0	0	92.56	87.88
Bilinear	0	30	93.02	88.82
	30	0	94.21	90.67
Nearest Neighbor	0	30	95.13	92.13
	30	0	96.24	93.36
Bicubic	0	20	99.21	98.75
	20	0	98.94	98.05

GCRIS

Chapter 4

4 Study-2

4.1 Comparison of Transfer Learning and Conventional Machine Learning Methods for Non-informative Frame Elimination from Colonoscopic Images

Detection of colon abnormalities is one of the most challenging tasks for gastroenterologists. Colonoscopy is the most common method to record videos and frames from colon to monitor any abnormality. However, the frames or videos obtained during the procedure are exposed to significant amount of unwanted artifacts such as motion artifact due to the fast movement of the colonoscopy probe or the capsule, specular reflection due to the light source used at the probe or in the capsule, improper contrast levels due to insufficient or excessive illumination inside the colon, gastric juice and bubbles, or residuals. Disease detection process should be conducted using clear frames, which are called informative. The main aim of this study was to investigate the effectiveness of conventional machine learning and transfer learning methodologies in detecting non-informative colonoscopy frames automatically.

4.2 Literature review

Automatic frame elimination technique is needed for both CC and WCE. In the literature, there are many studies focusing on non-informative frame elimination in which researchers have investigated different types of non-informative frames and feature extraction methodologies. For example, Ballesteros et al. proposed a method based on edge detection to separate informative images from non-informative images. Even though the implementation of this approach was simple, threshold determination was subjective and not adaptive [88]. A similar study was carried out by Oh et al. using Canny edge detection and thresholding [89]. In another study by Oh et al., they studied the separation of in-focus and out-of-focus frames [90]. Sun et al. suggested a method to remove non-informative frames, which included gastric juice and bubbles, from WCE videos. The local histogram, local binary pattern (LBP), and discrete cosine transform (DCT) were the feature extraction approaches used in that study [91]. Another study conducted by van Dongen et al. aimed at automatic detection of informative frames for early detection of oesophageal cancer. They used the color histogram, and the DCT coefficients as the features to be employed in the classification [92]. Tajbakhsh et al. divided the images into tiles and obtained two-dimensional (2D) DCT dominant coefficients from each tile. Later, they reconstructed the image and used a difference map in order to detect if that frame was non-informative or not. According to this study, non-informative frames included bubbles, blurring due to motion, and reflection artifacts [93]. Tong et al. supposed that Harr wavelet transform was an efficient method to detect a blurry image, and even to quantify the blurriness level of the image [94]. Arnold et al. conducted a similar study by using the discrete wavelet transform (DWT) coefficients as features obtained from the colonoscopy videos. They considered only the luminance channel of the image [95]. An et al. proposed a method to detect out-of-focus frames using discrete Fourier transform and texture analysis [96]. Cho et al. suggested using non-informative frame elimination to identify bleeding, polypectomy, residue or stool on the colonoscopy videos [97].

The focus measure operators (FMOs) include a set of feature extraction methods that were grouped into six categories/families. These categories are the gradient operator, Laplacian, DWT, DCT, image statistics, and a category that included the miscellaneous feature extraction methods. Using FMOs one can compute the focus level of an image for

every pixel. FMOs have been proposed for measuring the image quality [98] or sharpness of an image [99]. In addition, they were employed on the microscopic images [100]-[96]. In this particular problem, wavelet transform, [88]-[89], and discrete cosine transform based [92]-[93] features were used in previous studies.

In this study, we proposed using an individual methodology or combination of methodologies to extract features from the colonoscopy images to determine the non-informative frames with different artifact types, such as the motion artifact, specular reflection, improper contrast levels, gastric juice and bubbles, and residuals. Based on our knowledge, FMOs as a complete set, and texture analysis approaches like GLRLM, and NGTDM have never been used to eliminate non-informative colonoscopy frames. In addition to this, our database included six different artifact types as opposed to one or two types of artifacts used in several previous studies. However, increasing the number of features would cause increased computation time which would hamper the applicability of this system in real-time. The best f-measure and accuracy values were obtained using FMOs, that is why we also studied the performance and computation time of each FMO family separately.

In recent years, different image processing problems have been solved by using deep learning (DL) approaches, which requires large datasets. Convolutional neural networks are among the most popular DL approach in this area. However, for a specific problem in medical field, accessibility to huge amount of data is not always possible. Transfer learning has been developed to solve this kind of a problem [101]. In addition to feature extraction and classification applications, deep learning (DL) is a relatively new technique that is used to detect non-informative frames. In a recent study, Yao et al. worked on non-informative frames only with blur and reflection artifacts and used GLCM and convolutional neural networks (CNN) for the feature extraction and classification. Their database included 12,830 informative and 3,829 non-informative frames [102]. In another study, Islam et al. employed transfer learning approaches such as AlexNet, GoogleNet, ResNet and SimpleNet. In their study non-informative frame set included images with artifacts like blur, water and bubble [103]. Putten et al. preferred using ResNet and Hidden Markov Model in their study, and their database included a total of 3883 frames [104]. In these studies, there was no comparison between conventional machine learning and transfer learning approaches, and the artifact types were limited.

In our study we used AlexNet, SqueezeNet, GoogleNet, ShuffleNet, ResNet50, ResNet18, NasNetMobile, and MobileNet architectures to identify non-informative frames. All the architectures used in our study are pretrained using ImageNet dataset. These architectures were trained with over than one million images. We used the pre-trained network to train the system with our own dataset in MATLAB. All of the pretrained models have different number of layers, thus number of convolution layers, fully connected layers, pooling layers and parameters varies from model to model [105]. Information about used architectures in our study given table 4.2.1.

Table 4.2.1. The depth, number of parameters and input size of image for different type of transfer learning models.

Network	**Depth**	**Parameters (Million)**	**Image Input Size**
AlexNet	8	60	227
GoogLeNet	22	7	224
ResNet18	18	11,7	224
ResNet50	50	25.6	224
ShuffleNet	50	5	224
SqueezeNet	18	1.2	227
NasNetMobile	913	5	224
MobileNet	53	4.2	224

We tested the use of mentioned architectures as the transfer learning methodologies on the same database in order to compare these approaches. A comprehensive study like ours has not been published until now in this context. Elimination of non-informative frames using the methods mentioned above is the first step in our long-term goal to detect colon

diseases automatically which requires working on clear images. Automatic detection of non-informative frames requires the training, validation and test processes that include the labelled images to be used in the classification. In this study, we propose using both, as an individual category or in combination of categories, to extract features from the images to determine the non-informative frames with different artifact types, such as the motion artifact, specular reflection, improper contrast levels, gastric juice and bubbles, and residuals.

4.3 Method

In this study, conventional colonoscopy videos that are publicly available in https://www.gastrointestinalatlas.com/index.html were used. A total of 11,491 frames were extracted from 43 videos. The videos included the images from both healthy and diseased colons. Thus, our database consisted of images with colons that were healthy or had ulcerative colitis, Crohn's, cancer, or polyps. In the first phase of the study, we only focused on the differentiation of the images in terms of informativeness and did not care about whether they were coming from a healthy or a diseased colon. In this study, we divided our database into training, validation and test sets, and the number of patients were 37, 3, 3 respectively. The number of frames for training, validation and test sets were 10,113, 714 and 664 respectively.

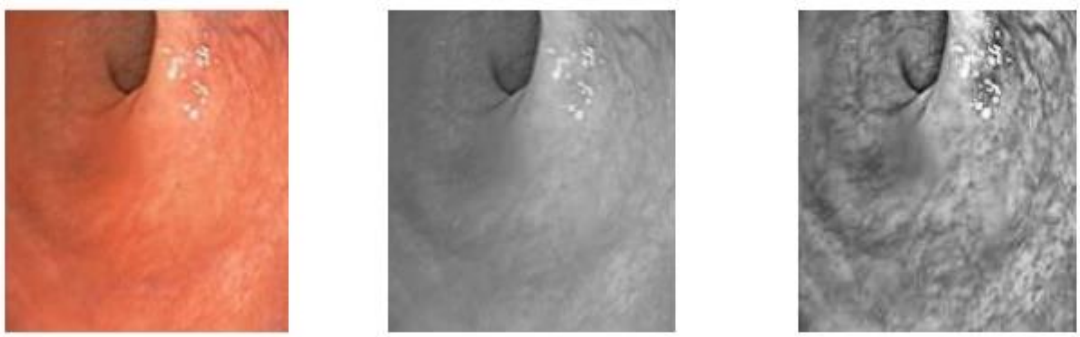
Ten-fold cross validation and hyperparameter optimization were used. While 5,064 images were manually labelled as "non-informative" due to the motion artifact, specular reflection, improper contrast levels, gastric juice and bubbles, and residuals, 5,049 images were labelled as "informative" for training set. We note that our database was split using stratified sampling. Figure 4.3.1 shows four images as examples of informative and non-informative frames from our image database. Sizes of the images extracted from the videos were different. In order to perform a reliable comparison, we automatically cropped all images to adjust the new size to become 176-by156 pixels. This process also

eliminated the black region that included the date and the patient name from the colonoscopy images.

a) Clear b) Bubbles c) Specular Reflection d) Motion-artifact

Figure 4.3.1. Four images as examples of informative and non-informative frames from our image database [106].

In order to enhance the contrast in the images, we employed adaptive histogram equalization (AdaptHistEq) approach [107] which is different from the regular histogram equalization approach. In this method histogram equalization was applied on small tiles (8-by-8 pixel squares), not on the whole image. This approach outperformed the regular histogram equalization because it had contrast-width limit that led to the prevention of the noise amplification (see Figure 4.3.2 for sample results) and made frames more visible. After the pre-processing step was completed, the features were extracted from the images using the texture features described below.



(a) (b) (c)

Figure 4.3.2. (a) Original frame, (b) Gray-scale, (c) Adaptive histogram equalization output [106].

In this study, in order to perform feature extraction, we used four popular second order texture analysis approaches called gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), neighborhood gray tone difference matrix (NGTDM), focus

measure operators (FMOs) and three first order statistics, such as kurtosis, standard deviation, and skewness to determine the non-informative frames with different artifact types, such as the motion artifact, specular reflection, improper contrast levels, gastric juice and bubbles, and residuals.

Automatic classification of frames from a video as informative and non-informative used features extracted from the frames employing different methodologies explained above. The min-max normalization was applied to the extracted features, i.e., the maximum and minimum values were equalized to 1 and 0, respectively. In the classification phase, different types of classifiers were employed such as the random forest, support vector machines and decision tree approaches. However, we decided to use the decision tree classifier because it was relatively faster than the other classifiers without compromising the accuracy. In addition, decision tree classification has many advantages such as being comprehensive and user-friendly and having high specificity [108]. It is also a preferred technique for image classification, pattern recognition and character recognition [109]. Decision tree is a classification method that creates a model in the form of a tree structure consisting of decision nodes and leaf nodes according to the classification, feature and target. The decision tree algorithm is developed by dividing the data set into smaller and even smaller pieces. A decision node can contain one or more branches. The first node is called a root node. A decision tree can consist of both categorical and numerical data [101]. In this study, we included all 143 features to train and test the system, then focused on FMOs and studied the performances of different feature families listed under FMOs using decision tree classification approach.

4.4 Results

The main goal of this study was to compare the performances of conventional machine learning and transfer learning-based classification approaches for informative and non-informative frame discrimination. In that context, machine learning based approach was investigated using different number of features using different subcategories of feature extraction methodologies. The performances of machine learning based classification are

shown in Table 4.4.1 for each type of feature extraction method. As shown in Table 4.4.1, when we compared GLCM, GLRLM, NGTDM, FMOs and image statistics-based feature extraction approaches, the highest accuracy and f-measure values were obtained using FMOs, ~89% and 0.79 respectively. The accuracy and f-measure values obtained using the combination of all features were not as high as when only FMOs, as a complete set, were employed. In addition, we examined the classification performances based on the feature families of FMOs such as the DCT, gradient, Laplacian, wavelet, miscellaneous, and image statistics using decision tree classifier (Table 4.4.2). The classification process using miscellaneous (MISC), only 8 features, and DCT based features (only 2 features) yielded accuracies of 79 and 83% respectively, which were not significantly worse than the accuracies obtained with all FMOs features. The other feature families did not yield promising results when they were tested individually. Table 4.4.3 shows transfer learning results of informative and non-informative frame discrimination.

In this study, feasibility of 8 transfer learning methods have been investigated. These methods were trained with different type and number of images after that we used this model for our database. In the first step, we initialized this model using training set and after that we used validation and test set. The names of models we preferred are AlexNet, SqueezeNet, GoogLeNet, Shuffle, ResNet50, ResNet18, NasNetMobile, and MobileNet. When Table 4.4.3 is considered the best performance metric results belong to AlexNet and the lowest performance metric results belong to MobileNet. However, even the lowest performance obtained using a transfer learning method is higher than the best result of conventional machine learning approaches. When other transfer learning results are examined, it is seen that the accuracy values vary between 90 and 99%. Table 4.4.4 depicts the performance comparison of conventional machine learning and transfer learning (AlexNet and fully connected layer) indicating that we obtained better results using transfer learning with respect to almost all performance metrics. Although the highest accuracy level was 88% in machine learning part, the transfer learning accuracy was 99%. We also observed that the computational cost was more advantageous in transfer learning than the fastest machine learning based classification approach.

Table 4.4.1 Machine learning results of 5 different texture features using decision tree algorithm.

Extracted Feature Type	**# of Features**	**Accuracy**	**f-measure**	**Precision**	**Recall**
GLCM	100	0.8509	0.6551	0.7231	0.8987
NGTDM	5	0.7410	0.6126	0.4739	0.8662
GLRLM	7	0.6491	0.4466	0.3561	0.5987
FMOs	28	0.8886	0.7874	0.7173	0.8726
Statistics	3	0.6220	0.4990	0.3634	0.7962
Combination of 5 features	143	0.8419	0.7042	0.6313	0.7962

Table 4.4.2 Machine learning results of FMO subcategories using decision tree algorithm.

FMO feature type	***# of features***	***Accuracy***	***f-measure***	***Precision***	***Recall***
Gradient-based	6	0.7575	0.5903	0.4915	0.7389
Laplacian-based	4	0.5858	0.5201	0.3582	0.9490
Wavelet-based	3	0.6084	0.4758	0.3481	0.7516
Statistics	5	0.7425	0.5799	0.4720	0.7516
DCT-based	2	0.8358	0.7212	0.6326	0.8981
Miscellaneous	8	0.7963	0.6800	0.6120	0.7650

Table 4.4.3 Transfer learning results of informative and non-informative discrimination.

Transfer Learning Models	***Accuracy***	***f-measure***
AlexNet	**0.9985**	**0.9968**
SqueezeNet	**0.9880**	**0.9748**
GoogLeNet	0.9849	0.9690
ShuffleNet	0.9639	0.9250
Resnet-18	0.9623	0.9191
NasNetMobile	0.9428	0.8774
Resnet-50	0.9367	0.8591
MobileNet	0.9051	0.8153

Table 4.4.4. The best results of both machine learning (ML) and transfer learning (TL) algorithm.

Comparison Methods		**Accuracy**	**f-measure**	**Precision**	**Recall**
ML	FMOs	0.8886	0.7874	0.7173	0.8726
TL	AlexNet	0.9985	0.9968	1	0.9936

In this study, we investigated the answers to the following two questions: (1) Is it possible to automatically discriminate informative frames from the non-informative ones successfully using different texture features such as GLCM, GLRLM, NGDTM, FMOs, and image statistics and typical classifiers? (2) Does transfer learning give better results when compared to machine learning? According to our results, among the feature extraction methods focus measure operators yielded the best performance. We found here that the non-informative frame elimination was possible using FMOs. In order to perform the elimination, we used 5 different feature extraction methods that included a total of 143 different features. We may argue that all subcategories of FMOs would be employed to eliminate non-informative frames. When subcategories were analyzed separately, the

miscellaneous family which have not been used in the literature so far gave the second-best results for our dataset. These miscellaneous family included 8 features; absolute central moment, Brenner's measure, image contrast, image curvature, Hemli's and Scherer's mean, steerable filters-based features, spatial frequency measure, and Vollath's autocorrelation. However, using transfer learning, specifically AlexNet, we obtained the best performance when compared to machine learning.

This study, being a preliminary work for our future studies, paved a road for important developments such as (1) the selection of informative frames automatically and decreasing the computational time while doing this, (2) the automatic classification of colon disease types using only the informative frames using transfer learning and (3) the formation of a pipeline that will help us to label the frames of a video in a setting that is close to the real time processing.

GCRIS

Chapter 5

5 Study-3

5.1 Automatic Abnormality Detection on Colonoscopic Images

Removing non-informative frames from a video or collection of images and proceeding with informative ones is critical in the automatic detection of diseases. In order to detect abnormalities in a colon automatically, we need to determine informative frames first. The artefacts inherent in the colonoscopy videos are not only a serious problem for the experts performing the procedure on site but also for a possible computer-aided diagnosis system. This was the focus of the previous chapter. In the third study of this dissertation work, the aim was to develop automatic colon abnormality detection. The abnormalities we took into account included Crohn's and ulcerative colitis (UC) diseases, colon cancer and polyp. We have divided this study into two parts. The first part aimed at determining whether any frame contained any of the diseases mentioned above (frame with one of the colon diseases we mentioned above) or a healthy colon tissue (healthy frame) regardless of the type of diseases. This was a binary classification problem. The goal of the second part was to perform a multi-category classification in which any frame was discriminated into one of the five classes as four types of diseases and healthy.

5.2 Literature review

Automatic abnormality detection on colonoscopic images is an ongoing effort, and there is no perfect solution yet. Currently, the experience of the gastroenterologist plays a huge role in diagnosing diseases. However, we believe that computer-aided diagnosis (CAD) systems would help them in stressful colonoscopy procedures. Time limitations and fatigue may cause incorrect diagnosis or missing an existing disease. Such a CAD system would inform the physician and lessen such unwanted outcomes. In the literature, one can find several studies on colonoscopy images for automatic abnormality detection in which the definition of abnormality varies. Polyps, inflammation, tumour, diverticulosis and bleeding can be listed as some of the abnormalities that were studied so far. The majority of previous studies have used conventional machine learning approaches. For example, Charfi and Ansari proposed a method to discriminate normal and abnormal WCE frames to detect polyp, inflammation and bleeding using local binary pattern (LBP) variance and discrete wavelet transform [110]. In another study on colonoscopy images, frames with polyp, tumour and blurriness were defined as abnormal [111]. They used fuzzy color and texture histogram and color and edge directivity descriptor as two different feature extraction methods, and preferred Inception-v3 transfer learning model. For automatic detection of diseases Sindhu et al. conducted a study on WCE frames employing machine learning methods. The gray level co-occurrence matrix (GLCM) and scale invariant Fourier transform (SIFT) were used to extract features in this study, and polyp, tumour, bleeding and Crohn's disease were defined as abnormal [112]. Pogorelov et al. suggested using fuzzy color and histogram texture method for real-time GPU-accelerated implementation. The abnormal frames included polyp, ulcerative colitis (UC), Crohn's, diverticulosis and cancer [113]. Another study aimed at detecting WCE frames with inflammation and polyp using only SIFT approach [114]. In the study by Iakovidis et al., color features were investigated for the detection of WCE frames with polyp, bleeding and ulcers [115]. In 2013, automatic classification was performed by combining GLCM, LBP and color features. In this study, bleeding and lesions were assumed as abnormal classes in the dataset [116]. Detecting abnormalities using texture features and neural network was conducted by Karkanis et al. whose dataset included only healthy frames and frames with polyps [117]. Almost two decades ago, Li et al. investigated a different

approach in which patch-based classification using CIE-lab color texture feature to detect frames with tumour, polyp, lipoma and inflammation was used [118].

In this dissertation work, our aim was to automatically detect the most common colon abnormalities. We should note that non-informative frames will be eliminated during the pre-processing stage, thus we selected and used the most appropriate/informative frames in this part of the study.

5.3 Method

In this part of the dissertation work, our colonoscopic image database was constructed with 2787 manually selected informative frames. The frames included both healthy colon tissue and colon with different types of diseases. In the first phase of this study, we focused on differentiation of frames with healthy and diseased colon (binary classification), and in the second phase we investigated the detection of disease types. The data set used in the first phase included 1512 frames with healthy colon and 1275 frames with four types of diseases. First, the frames in the healthy colon dataset was divided into training, validation and test sets, which included 1203, 171 and 138 frames respectively. Secondly, the dataset that included the frames with diseases was also divided into training, validation and test sets, with 990, 135 and 150 frames respectively. Ten-fold cross validation and hyperparameter optimization were performed. In addition to this, we split our dataset into training, validation and test sets using stratified sampling method. The ratio of samples in training, validation and test sets were 80, 10 and 10%, of all dataset respectively. Therefore, total number of frames in the training, validation and test sets were 2193, 306 and 288 respectively. Figure 5.3.1 depicts five images as examples of frames with healthy and diseased colons from our image database. For the second part of our study, which aimed at detecting individual disease types, again we used stratified sampling preserving the ratios of disease types. Starting from a total of 1275 frames we divided our database into training, validation and test sets. The number of frames allocated for training, validation and test sets were 990, 135 and 150 respectively. The number of frames for cancer, Crohn’s, polyp and UC were 375, 375,

154 and 371 respectively. Training, validation and test sets for disease-based dataset is explained in table 5.3.1.

Table 5.3.1 Number of frames in the database which are coming from healthy and diseased colons.

		TRAINING	**VALIDATION**	**TEST**	**TOTAL**	**TOTAL**
HEALTHY	*HEALTHY*	1203	171	138	1512	1512
DISEASE	*CANCER*	313	31	31	375	1275
	CROHN'S	257	32	86	375	
	POLYP	120	13	21	154	
	ULCERATIVE COLITIS	300	59	12	371	

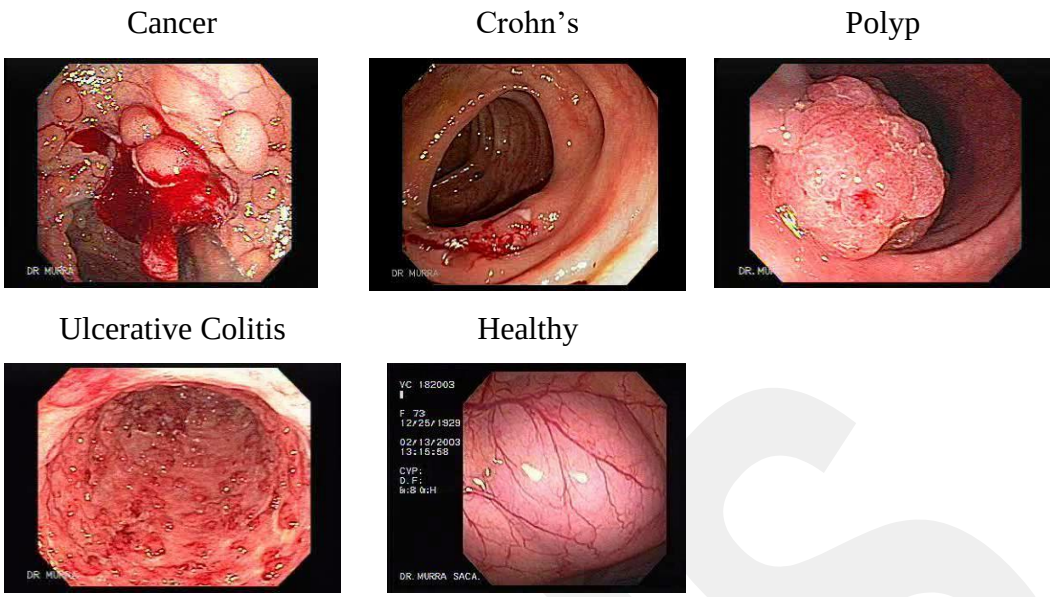


Figure 5.3.1 Sample images from our dataset showing diseased and healthy colons [106].

The aim of this study was to detect the presence of disease and disease types using both conventional machine learning and transfer learning based algorithms. In that context, machine learning based approaches were investigated using different number of features from different subcategories of feature extraction methodologies. The feature extraction methods we have employed were the same as the ones we have used in the previous chapters, such as GLCM, GLRLM, NGTDM, FMOs and image statistics. A total of 143 features were obtained using these approaches. In addition to this, we deployed conventional machine learning methods which are decision tree, SVM and k-NN. The default number of split for decision tree was 15, sigma value for SVM was 1 and number of neigborhood for k-NN classifier was 5. Apart from machine learning application, we used the same transfer learning architectures as we have used in the previous chapter to detect the presence of disease and to discriminate the disease types.

5.4 Results

5.4.1 Binary (Healthy Frames vs. Frames with Disease) Classification

- The machine learning results are shown in table 5.4.1. Decision tree yielded the highest accuracy and f-measure values for disease detection problem when compared to SVM and k-NN classification models.

Table 5.4.1.1 Performance metrics of machine learning results for binary classification.

Models	Accuracy (%)	f-measure
Decision Tree	**84.64**	**0.8792**
SVM	71.24	0.7179
k-NN	63.73	0.6873

- The performance of transfer learning based methods is given below in table 5.4.2. When accuracies were compared, the best performance was obtained using SqueezeNet architecture. The accuracy and f-measure results were ~90% and ~0.92 respectively. In addition, AlexNet architecture yielded the second-best result in terms of accuracy and f-measure metrics. The accuracy value was 86% and f- measure was 0.86.

Table 5.4.1.2 Performance metrics of transfer learning results of binary classification.

Models	Accuracy (%)	f-measure
SqueezeNet	**90.63**	**0.9153**
AlexNet	**85.76**	**0.8600**
NasNetMobile	76.73	0.8109
MobileNet	76.38	0.7975
GoogleNet	76.38	0.8033
ResNet-18	79.51	0.8279
ShuffleNet	69.09	0.7342

- The above two tables show that the detection of disease presence can be performed using transfer learning algorithms. When conventional machine learning and transfer learning results are compared in terms of accuracy and f-measure values, it is clear that SqueezeNet and AlexNet architectures yield better results than the best performing machine learning method.

5.4.2 Multi-label Classification

- In the second part of the study, we investigated the performances of conventional machine learning and transfer learning methodologies on a dataset that included frames with healthy colon and frames with polyp, cancer, ulcerative colitis, and Crohn's disease. Table 5.4.3 shows the performance metrics of machine learning algorithms; decision tree, SVM, and k-NN. While the highest value for accuracy was obtained using decision tree classifier (69%), the highest f-measure was obtained using SVM (0.88). Disease type detection was not as easy as disease detection because in this part the most important problem was the number of frames. The number of frames was limited in this part which led to inefficient training and test phases.

Table 5.4.2.1 Performance metrics of conventional machine learning results for multi-label classification.

METHODS	**Accuracy (%)**	**f-measure**
Decision Tree	**69.28**	**0.630**
SVM	**68.63**	**0.880**
k-NN	**50.00**	**0.504**

- Transfer learning results of disease detection are given below in table 5.4.4. According to these results, GoogLeNet gave the best accuracy with 80%. In addition to this, AlexNet and ResNet-18 have the second-best performances. Their accuracy values were very close to each other. In addition to accuracy results, f-measure values had similar characteristics with the accuracy results.

- Other transfer learning approaches did not yield higher performances than the machine learning methods did. However, implementation of machine learning algorithms takes more time than the transfer learning applications.

Table 5.4.2.2 Performance metrics of transfer learning results of disease type detection.

Models	**Accuracy (%)**	**f-measure**
GoogLeNet	**80.39**	**0.7308**
AlexNet	**78.76**	**0.7087**
ResNet-18	**78.10**	**0.7600**
ResNet-50	69.93	0.7147
SqueezeNet	65.69	0.5393
ShuffleNet	61.44	0.7600
MobileNet	55.56	0.7122
NasNetMobile	46.88	0.3762

Chapter 6

6 Conclusions and Future Prospects

6.1 Conclusions

In this thesis work, three different studies have been conducted. Even though these three studies were different from each other, the main aim was to reduce experts' workload and to enable automatic detection of non-informative frames and frames with diseases. These studies were conducted to create preliminary step for future real-time based studies.

In the first study, we tried to answer the following question: Does the interpolation of specular reflections encountered in colonoscopic images affect the texture features to be used in the automatic detection of polyps? First, we made comparisons between texture features obtained from an image with no specular reflections and the same features obtained with synthetically added reflections with various sizes plus the interpolation. Secondly, we performed automatic classification between polyp and background colon based on texture features obtained from the interpolated images. In the classification phase, we did not use the image as a whole, rather 32x32 sub-images extracted from the gray-scale colonoscopic images. We can summarize our key findings as follows:

- The size of the specular reflection is an important factor. When the images include 2% reflection, interpolation could be effective. At most 0.9% and 7% of the texture features were affected significantly from the interpolation on subimages with polyps and background, respectively. However, when the reflection size was greater than 10%, the interpolation was not effective. It is worth noting that bicubic interpolation restored the texture features the most.

- In the classification of polyp and the background, the random forest approach performed better than the *k*-NN algorithm. Using WEKA software, we obtained classification results

and stated the f-measure and accuracy values. The overall accuracy level was more than 92% for all interpolation methods. The results showed that there was an improvement when BC was used over other interpolation techniques for 10, 20, 30% reflection.

The previous work on this problem was limited to the interpolation (or removal) of specular reflections on colonoscopic, endoscopic or other medical images. For endoscopic images, Guo et al. suggested a method to eliminate specular reflection [83], Stehle used spectral deconvolution algorithm to remove reflections [84], Arnold et al. [85] and Karapetyan et al. [86] investigated different inpainting algorithms. In addition, Tchoulack et al. proposed using a real-time inpainting algorithm to eliminate specular reflections from endoscopic images [119]. Aydi et al. showed the use of interpolation methods to remove reflections on iris images [120]. None of the abovementioned studies focused on the effect of these methods on the texture features and accuracy of automatic classification of polyp and the background.

In the second study, we investigated the answers to the following two questions: (1) Is it possible to automatically discriminate informative frames from non-informative ones successfully using different texture features such as GLCM, GLRLM, NGDTM, FMOs, and image statistics and typical classifiers? (2) Does transfer learning give better results when compared to machine learning? According to our results, among the feature extraction methods FMOs (focus measure operators) yielded the best performance. We found here that the non-informative frame elimination was possible using FMOs. In order to perform the elimination, we used 5 different feature extraction methods that included a total of 143 different features. We may argue that all subcategories of FMOs would be employed to eliminate non-informative frames. When subcategories were analyzed separately, the miscellaneous family which have not been used in the literature so far gave the second-best results for our dataset. These miscellaneous family included 8 features; absolute central moment, Brenner's measure, image contrast, image curvature, Hemli's and Scherer's mean, steerable filters-based features, spatial frequency measure, and Vollath's autocorrelation. However, using transfer learning, specifically AlexNet, we obtained the best performance when compared to machine learning.

In numerous studies the researchers have focused on the elimination of non-informative frames from videos. Non-informative frame elimination is a tough task not only for

conventional colonoscopy videos but also for different types of endoscopic and wireless capsule endoscopy videos. Fan et al. studied on WCE frames to select informative images using lumen depth of lumen perception and motility assessment [121]. In another study, WCE images were segmented according to having bubbles and turbid, or being clear using color and texture features [122]. During the bronchoscopy process detecting non-informative frames was the focus of another study in which images with reflection, loss of focus, impurity and motion blur were included. They employed zero cross edge detection, color transformation (HSV), MPEG-7 edge and DCT spectrum [123]. Moreover, Rangseekajee and Phongsuphap suggested a method to classify thoracoscopic images as informative or non-informative using edge-based techniques [124]. Another study was conducted by Hwang et al. to discriminate non-informative images on colonoscopy images using discrete Fourier transform and texture analysis (gray level co-occurrence matrix) [125]. The methods that were and their results are shown in table 6.1.2. It is clear to see that in previous works the comparison of conventional machine learning and transfer learning results were limited. However, our results were obtained using not only one texture feature family or transfer learning method. Our transfer learning results for non-informative frame elimination were promising and the performances were higher than the previous works. In addition to these outputs, our non-informative database included 6 different type of artifacts that are motion artifact, specular reflection, improper contrast, gastric juice, bubbles and residuals not only two or three types of artifacts.

Table 6.1.1 Results of previous works related with non-informative frames.

Reference Number	***# of Dataset***	***Method***	***Classification***	***Results (P.metrics)***
Fan et al. [121]	500 frames	• Histogram • Mean Shift	• Mean Shift	87.5 (Specificity)
Arivazhagan et al. [122]	50 videos	• RGB Color • Surf Feature	• Support Vector Machine	85.2 (Accuracy)

Grega et al. [123]	768 frames	• Edge Detection	• Neural Network	93 (Accuracy)
		• HSV Analysis		89 (Accuracy)
		• MPEG-7edge Histogram		93 (Accuracy)
		• DCT Spectrum		90 (Accuracy)
Ranksekajee et al. [124]	387 frames	• Isolated Pixel Value	• Thresholding	95.1 (Accuracy)
An et al. [125]	5,971 frames	• DFT • GLCM	• k-Mean Clustering	98.1 (Accuracy)
Oh et al. [89]	323,000 frames	• Edge Detection	• Thresholding	0.96 (Recall)
Tong et al. [38]	2,355 frames	• Haar Wavelet	• Thresholding	98.6 (Accuracy)
Oh et al. [90]	923 frames	• Edge Detection	• k-Mean Clustering	95 (Accuracy)
Arnold et al. [95]	15,000 frames	• 2D Dwt	• Bayesian	92.3 (Accuracy)
Sun et al. [40]	180,000 frames	• LBP • DCT	• k-NN	99.3 (Accuracy)
Dongen et al. [92]	2,172 frames	• DCT • Color	• Regression Tree	97 (Accuracy)
Ballesteros et al. [88]	2,000 frames	• Edge Detection • Threshold	• Support vector Machine	95 (Accuracy)
Yao et al. [102]	16,659 frames	• CNN • GLCM	• Convolutional Neural Network	0.77 (f-measure)

Islam et al. [103]	6,805 frames	• AlexNet • GoogLeNet • ResNet • SimpleNet	• Convolutional Neural Network	0.94 (f-measure)
Putten et al. [104]	3,883 frames	• ResNet • HMM	• Convolutional Neural Network	0.91 (f-measure)

In the third study, we investigated the answers to the following questions: (1) Is it possible to detect abnormalities automatically from informative frames using both machine learning and transfer learning method? (2) Is it possible to detect disease types using machine learning and transfer learning? (3) Does transfer learning give better results when compared to the machine learning? In this study, for machine learning part we did not obtain results according to feature extraction types because the main aim of this study was to compare transfer learning and conventional machine learning results. In machine learning part, we used 143 different features which are GLCM, GLRLM, NGTDM, FMOs and statistics. In the transfer learning part, we used 9 different methods which are AlexNet, GoogleNet, ResNet18, ResNet50, SqueezeNet, ShuffleNet, MobileNet and NasNet. While SqueezeNet and AlexNet gave the best two results for abnormality detection, GoogleNet, Alexnet and ResNet-18 gave the best three results for abnormality type detection. We understand from the fact that the performances were higher with transfer learning than conventional machine learning methods for both abnormality and abnormality type detection, transfer learning can be used more efficiently and successfully.

Table 6.1.3 shows the previous studies related with abnormality detection in colonoscopy or WCE images. According to this table, previous studies generally preferred to use texture features or color features to detect abnormalities. Studies on transfer learning algorithms to detect colon abnormalities are still limited. However, our last study has been conducted using both conventional machine learning and transfer learning algorithms. We preferred both using texture features and transfer learning models. We used 7 different types of transfer learning models and 143 texture features in our study.

Table 6.1.2 Results of previous works related with abnormality detection.

Reference Number	***# of Dataset***	***Method***	***Classification***	***Results (P.metrics)***
Yoshida et al. [75]	100 frames	• DWT • LBPV	• Linear Discriminant Analysis	93 (Accuracy)
Pogorelov et al. [111]	300 frames	• Fuzzy Color & Texture • Histogram • Color & Edge Directivity Descriptor • Inception-v3	• Decision Tree • Random Forest • k-NN	96 (Accuracy)
Sindhu et al. [112]	1,385 frames	• GLCM • SIFT	• Multilayer Perceptron NN	97 (Accuracy)
Gueve et al. [114]	600 frames	• SIFT	• Support Vector Machines	89 (Accuracy)
Iakovidis et al. [115]	137 frames	• Color • HSV • YCbCr	• Support Vector Machines	0.89 (AUC)
Manivannan et al. [116]	2,100 frames	• GLCM • LBP • Color	• Support Vector Machines	90 (Accuracy)
Karkanis et al. [117]	100 frames	• GLCM • Multilayer feed forward NN	• Artifical Neural Network	93 (Accuracy)
Peng et al. [118]	58 frames	• CIE-lab	• Support Vector Machines	82 (Accuracy)

6.2 Contribution to Global Sustainability

According to WHO statistics in 2018, colon cancer is the second most common type of cancer among woman and the third in man. Other type of colon diseases also can turn into colon cancer in the long term, thus, early diagnosis has a huge impact on saving human life. According to UNDP sustainable development goals, our purpose meets fifth goal which is related to health among all 17 goals. In this study, the main goal was to reduce experts' workload and misdiagnosis rate because colonoscopy procedure has many disadvantages in terms of patients and experts. However, conventional colonoscopy (CC) is still the most preferred technique because when compared to WCE and CT colonoscopy, it is more user friendly than the other two methods. When CC procedure is examined from the perspective of patients, it is painful and scary thus patients do not have a colonoscopy unless they have to. In addition to this, when CC is examined in terms of experts, it is clear to say that this method is expert-dependent. Expert-dependent procedures are affected by the experience or fatigue of the experts, which can affect the results. Also, colon has folded structure and this structure makes experts' disease detection harder. When these situations are considered, enhancing the disease detection process in CC plays an important role. In this thesis work, we tried to make non-informative frame elimination and disease detection automatically. This automatic detection of non-informative frames and diseases will guide the experts. In the next step, our studies will be transformed into a real-time system in which automatic elimination and detection will be possible using a software to be embedded in colonoscopy devices during the procedures in real time. During the operation, experts will get help from this software and perform decisions more efficiently and accurately. Thus, misdiagnosis rate will be reduced and non-informative frame elimination will gain some time for the experts.

6.3 Future Prospects

We can summarize our future prospects for the three studies performed in this thesis work separately as follows:

.

- Future prospects of study 1:

  Although we believe that our results are promising, we must acknowledge that these findings need to be validated or augmented by future studies in which more images are included in the database. One of the limitations of this study is that we had to study with an unbalanced database. In this study, only 10% of the sub-images (tiles) came from the polyps. In addition, the effect of interpolation in the detection of polyps was investigated on the sub-images not the whole image. A deep learning-based approach will be developed to compensate for the specular reflections. Furthermore, a future study will tackle a real-time specular reflection removal approach. We will also investigate the use of other feature extraction and selection methods, and other classifiers such as SVM and ensemble techniques.

- Future prospects of study 2:

  This study, being a preliminary work for our future studies, paved a road for important developments such as (1) the selection of informative frames automatically and decreasing the computational time while doing this, (2) the automatic classification of colon disease types using only the informative frames using transfer learning and (3) the formation of a pipeline that will help us to label the frames of a video in a setting that is close to the real time processing. We should note that informative and non-informative frame discrimination is still a hot topic for endoscopic images. Our study proposed that non-informative image elimination could be performed by focus measure operators on colonoscopic images, and could be extended for use in other endoscopic image types. In addition to this, our study claimed that our dataset split was stratified and the test frames were never used in the training and validation phases. Although our results were promising for conventional colonoscopy images, similar approaches could be applied for WCE images.

- Future prospects of study 3:

  Abnormality detection in colonoscopic images is still a hot topic. In order to perform abnormality detection automatically, the diversity of images in the database should be increased. Obtaining huge amount of data for different types of abnormalities is not as easy as obtaining non-informative frames because non-informative frames are obtained frequently and depend on camera angle. However, diseases should be obtained from different patients, in order to train the system properly. After increasing the number of frames, a hyperparameter optimization procedure can be applied and more reliable and better results can be obtained for disease type detection. In addition to this, the number of diseases diversity can be increased. Real-time application can be embedded into the colonoscopy devices; this system will be useful guide for the experts. Thus, they will be able to make reliable decisions on the disease types relatively more easily.

  Figure 6.3.1. shows our future prospect of developing a pipeline of real-time application of the methodologies developed during this thesis work. This pipeline will be developed and tested on colonoscopy videos in collaboration with gastroentereologists.

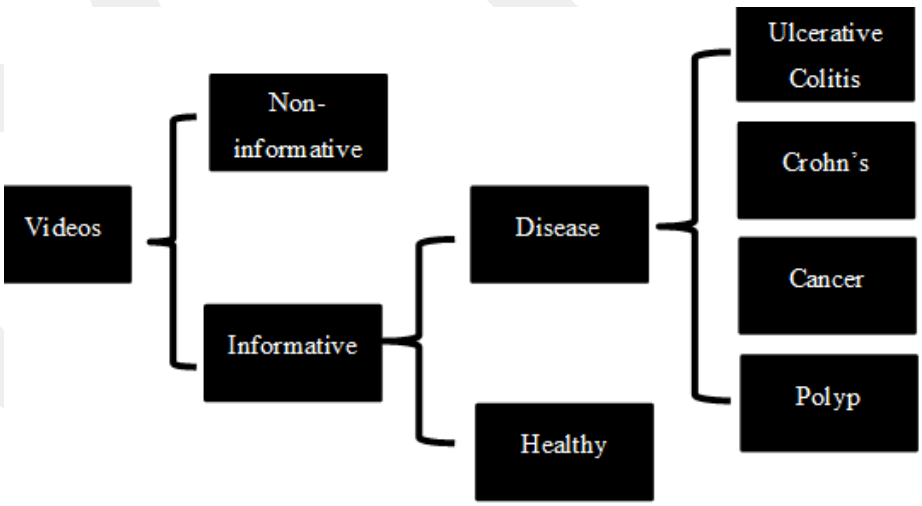


Figure 6.3.1 Real-time processing pipeline of automatic disease detection on colonoscopy videos.

BIBLIOGRAPHY

[1]American Cancer Society. "Cancer Facts and Statistics Https://Www.Cancer.Org/Research/Cancer-Facts-Statistics.Html, [Online], July (2020).

[2] 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, " *IEEE Transactions on Medical Imaging*, 37, 2196-2210, (2018).

[3] A. Sieg, "Capsule Endoscopy Compared with Conventional Colonoscopy for Detection of Colorectal Neoplasms," *World Journal Gastrointestinal Endoscopy*, 3, 81-85, (2011).

[4] Z. Liang and R. Richards, "Virtual Colonoscopy vs. Optical Colonoscopy," *Expert Opinion on Medical Diagnostics*, 4, 159-169, (2010).

[5] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F.Vilariño, "WM-DOVA Maps for Accurate Polyp Highlighting in Colonoscopy: Validation vs. Saliency Maps from Physicians," *Computerized Medical Imaging and Graphics*, 43, 99–111, (2015).

[6] R. N. Kaçmaz, B. Yılmaz, and Z. Aydın, "Effect Of Interpolation on Specular Reflections in Texture-Based Automatic Colonic Polyp Detection," *International Journal of Imaging Systems and Technology*, 30, 1-9, (2020).

[7] Y. Nigam, J. Knight, N. Williams, "The Anatomy and Functions of the Large Intestine, " *Nursing Times*, 115, 50-53, (2019).

[8] https://teachmeanatomy.info/abdomen/gi-tract/colon/, [Online], July (2020).

[9]https://www.medtronic.com/us-en/patients/treatments-therapies/colon-disease.html, [Online], July (2020).

[10]https://www.aboutkidshealth.ca/Article?contentid=924&language=English, [Online], July (2020).

[11]https://www.medicinenet.com/colon_polyps/article.htm, [Online], July (2020).

[12] A. Gonzalez et al., "Radiation Related Cancer Risks from CT Colonography, Screening: Risk Benefit Analysis," *American Journal of Roentgenolog*, 4, 816-823, (2011).

[13]http://ehyadarman.com/products/30/NeuViz-16-Classic-CT?lang=en, [Online], July (2020).

[14]https://www.dreamstime.com/stock-illustration-ct-scan-d-virtual-colonoscopy-colorful-there-example-anatomical-elements-which-can-be-very-useful-diagnostic-image88863864, [Online], July (2020).

[15] H. Song and K.Shim "Current Status and Future Perspectives of Capsule Endoscopy," *Intestinal Research*, 14, 21-29, (2016).

[16]https://www.researchgate.net/publication/221916265_Multiscale_Texture_Descriptors_for_Automatic_Small_Bowel_Tumors_Detection_in_Capsule_Endoscopy, [Online], July (2020).

[17]https://www.dicardiology.com/article/capsule-endoscopy-systems-safety-patients-cardiovascular-implants, [Online], July (2020).

[18] M.Hafner, "Conventional Colonoscopy: Technique, Indications, Limit," *European Journal of Radiology*, 61, 409-414, (2007).

[19]https://www.itnonline.com/content/eliminating-looping-colonoscopy-procedures, [Online], July (2020).

[20]https://www.gastroenterologyadvisor.com/general-gastroenterology/factors-influencing-adequacy-of-colonoscopy-preparation-in-hospitalized-patients/, [Online], July (2020).

[21]https://freecontent.manning.com/the-computer-vision-pipeline-part-3-image-preprocessing/, [Online], July (2020).

[22] G. Kumar and P. K. Bhatia, "A Detailed Review of Feature Extraction in Image Processing Systems," *Fourth International Conference on Advanced Computing and Communication Technologies*, 5-12, (2014).

[23] D. Tian, "A review on image feature extraction and representation techniques," *International Journal of Multimedia and Ubiquitous Engineering*, 8, 385-395, (2013).

[24] S. Kale, "Color, Shape and Texture Feature Extraction for Content Based Image Retrieval System," *International Journal of Advanced Research in Computer and Communication Engineering*, 5, 303-306, (2016).

[25] N. Tajbakhsh, S.R. Gurudu, and J. Liang, “A Classification-Enhanced Vote Accumulation Scheme For Detecting Colonic Polyps,” *Abdominal Imaging*, 53-62, (2013).

[26] K. Geetha and C. Rajan, “Automatic Colorectal Polyp Detection in Colonoscopy Video Frames,” *Asian Pacific Journal of Cancer Prev*ention, 17, 4869–4873, (2016).

[27] J. Bernal, J. Sánchez, and F. Vilariño, “Towards Automatic Polyp Detection with a Polyp Appearance Model, ” *Pattern Recognit*ion, 45, 3166–3182, (2012).

[28] R. Bhotika, P. R. S. Mendonça, S. A. Sirohey, W. D. Turner, Y. Lee, J. M. McCoy, R. E. B. Brown, and J. V. Miller, “Part-Based Local Shape Models for Colon Polyp Detection, ” *Medical Image Computing and Computer-Assisted Invention*, 4191, 479-486, (2006).

[29] J. Bernal, J. Sanchez and F. Vilarino, “Impact of Image Preprocessing Methods on Polyp Localization in Colonoscopy Frames,” *35th Annual International Conference of IEEE EMBS*, 7350-7354, (2013).

[30] S. Ameling, S. Wirth, D. Paulus, G. Lacey, and F. Vilarino, “Texture-Based Polyp Detection in Colonoscopy,” *Bildverarbeitung für die Medizin*, 346-350, (2009).

[31] A. Dahal, J. Oh, W. Tavanapong, J. Wong, and P. C. De Groen, “Detection of Ulcerative Colitis Severity in Colonoscopy Video Frames,” *International Workshop Content Based Multimedia Index*, (2015).

[32] S. M. Krishnan, X. Yang, K. L. Chan, S. Kumar, and P. M. Y. Goh, “Intestinal Abnormality Detection From Endoscopic Images,” *Proceeding 20th Annual International. Conferences IEEE Engineering in Medicine and Biology Society*, 2, 895–898, (1998).

[33] L.Gueye, S. Yildirim-Yayilgan, F. A. Cheikh, and I. Balasingham, “Automatic Detection of Colonoscopic Anomalies Using Capsule Endoscopy,” *International Conference on Image Process*, 1061–1064, (2015).

[34] B. Ergen and M. Bayram., “İstatistiksel Uzaysal Alan Metotlarının Içerik Tabanlı Tıbbi Görüntü Erişimi İçin Bir Uygulama,” *Fırat Üniv. Mühendislik Bilimleri Dergisi*, 2, 87-93, (2011).

[35] M. Sonka, V. Hlavac, R. Boyle, “Image Processing, Analysis, and Machine Vision,” *International Thomsom Publishing*, 4, 752-754, (2015).

[36] M. Amadasun and R. King, “Textural Features Corresponding to Textural Properties,” *IEEE Transaction on Systems, Man and Cybernetics*, 5, 1264–1274, (1989).

[37] X. Marichal, W. Ma, and H. Zhang, "Blur Determination in the Compressed Domain Using DCT Information," *IEEE International Conf Image Process*, 2, 386-390, (1999).

[38] H. Tong., M. Li., H. Zhang, and C. Zhang, "Blur Detection for Digital Images Using Wavelet Transform," *IEEE Int. Conference Multimedia. Expo*, 17–20, (2004)

[39] A. Thelen, S. Frey, S. Hirsch, and P. Hering, "Improvements In Shape-From Focus for Holographic Reconstructions with Regard to Focus Operators, Neighborhood-Size, and Height Value Interpolation," *IEEE Transactions on Image Processing*, 18, 151–157, (2009).

[40] Z. Sun, B. Li, R. Zhou, H. Zheng, and H. Meng, "Removal of Noninformative Frames for Wireless Capsule Endoscopy Video Segmentation," *IEEE International Conference on Automation and Logistics*, 294–299, (2012).

[41] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of Focus Measure Operators for Shape-From-Focus," *Pattern Recognition*, 46, 1415–1432, (2013).

[42] A. Santos, et al., "Evaluation of Autofocus Functions in Molecular Cytogenetic Analysis," *Journal of Microscopy*, 188, 264–272, (1997).

[43] K.Manjula, K.Vijayarekha and P.Vimaladevi, "Review on Classification Algorithms in Image Processing," *International Journal of Innovative Trends in Engineering & Research*, 2, 1-5, (2017)

[44] K. Karjus, M. Hebart and R. Vicente, "An Efficient Data Partitioning to Improve Classification Performance While Keeping Parameters Interpretable," *Plos One*, 11, 1-16, (2016).

[45] https://medium.com/@jorgesleonel/hyperparameters-in-machine-deep-%20%20%20learning-ca69ad10b981, [Online], August (2020).

[46] W. Zhang and F. Gao, "An Improvement to Naïve Bayes for Text Classification," *Advanced in Control Engineering and Information Science*, 15, 2160-2164, (2011).

[47] https://towardsdatascience.com/introduction-to-na%C3%AFve-bayes-classifier-fa59e3e24aaf, [Online], August (2020).

[48] S. Dreiseitl and L. Machado, "Logistic regression and artificial neural network classification models: A methodology review," *Journal of Biomedical Informatics*, 35, 352-359, (2002).

[49] https://medium.com/@ekrem.hatipoglu/machine-learning-classification-logistic-regression-part-8-b77d2a61aae1 , [Online], August (2020).

[50] C. Kingsford and S.L Salzberg, "What are the decision trees?," *National Biotechnology*, 26, 1011-1013, (2009).

[51] https://www.kdnuggets.com/2016/09/decision-trees-disastrous-overview.html, [Online], August (2020).

[52] A.Sarica, A.Cerasa and A.Quattrone, "Random Forest Algortihm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review, " *Frontiers in Aging Neuroscience*, 9, 1-12, (2017).

[53] https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/**,** [Online], August (2020).

[54] Z. Zhang, "Introduction to Machine Learning: k-Nearest Neighbors", *Annals of Translational Medicine*, 4, 1-7, (2016).

[55]https://Medium.Com/Sifium/Machine-Learning-Types-Of-Classification-9497bd4f2e14, [Online], August (2020).

[56] T. Evgeniou and M. Pontil, "Workshop on Support Vector Machines: Theory and Application, " *Machine Learning and Its Applications, Advanced Lectures*, 7-12, (2001).

[57]https://medium.com/@ekrem.hatipoglu/machine-learning-classification-support-vector-machine-kernel-trick-part-10-7ab928333158, [Online], August (2020).

[58] K.Weiss, T.Khoshgoftaar and D.Wang, "A Survey of Transfer Learning", *Journal of Big Data*, 3, 1-40,(2016).

[59] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, (2012).

[60] I. Forrest, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-Level Accuracy with 50x Fewer Parameters and <0.5 MB Model Size, " *Preprint, Submitted November*, 4, (2016).

[61] S. Christian, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper With Convolutions," In *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, 1-9, (2015).

[62] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6848-6856, (2018).

[63] H. Kaiming, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," In *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, 770-778, (2016).

[64] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. C. Chen, "Mobilenetv2: Inverted Residuals and Linear Bottlenecks," *Conference on Computer Vision and Pattern Recognition* , 4510-4520, (2018).

[65] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8697-8710, (2018).

[66] https://www.mayoclinic.org/diseases-conditions/colon-cancer/symptoms-causes/syc-20353669, [Online], August (2020).

[67] J. Ong, A. Seghouane, and K. Osborn, "Polyp Detection in CT Colonography Based on Shape Characteristics and Kullback-Leibler Divergence," *Biomedical Imaging: From Nano to Macro*, 636-639, (2008).

[68] A. Wolf et al., "Colerectal Cancer Screening for Average- Risk Adults" *2018 Guideline Update from the American Cancer Society*, 68, 250-281, (2018).

[69] N. Tajbakhsh, S. R. Gurudu, and J. Liang, "Automatic Polyp Detection Using Global Geometric Constraints and Local Intensity Variation Patterns," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 8674, 179- 187, (2014).

[70] S. Hwang, J. Oh, W. Tavanapong, J. Wong, and P. C. De Groen, "Polyp Detection in Colonoscopy Video Using Elliptical Shape Feature," *International Conference on Image Processing*, 2, 465–468, (2007).

[71] M. Park, S. J. Jin, R. Hofstetter, M. Xu, and B. H. Kang, "Automatic Colonic Polyp Detection by the Mapping Using Regional Unit Sphere," *International Conferences on Multimedia Ubiquitous Engineering*, 144–149, (2008).

[72] J. Bernal, J. Sánchez, and F. Vilariño, "Towards Automatic Polyp Detection with a Polyp Appearance Model," *Pattern Recognition*, 45, 3166–3182, (2012).

[73] J. Lynn Ong and A. Seghouane, "Feature Selection Using Mutual Information in CT Colonography," *Pattern Recognition Letters*, *32*, 337-341, (2011).

[74] J. Lynn Ong and A. Seghouane, "From Point to Local Neighborhood: Polyp Detection in CT Colonography Using Geodesic Ring Neighborhoods," *IEEE Transactions on Image Processing*, 20, 1000-1010, (2011).

[75] H. Yoshida and J. Nappi, "Three-Dimensional Computer-Aided Diagnosis Scheme for Detection of Colonic Polyps," *IEEE Transactions Medical Imaging*, 20, 1261–1274, (2001).

[76] N. Tajbakhsh, C. Chi, S. R. Gurudu, and J. Liang, “Automatic Polyp Detection From Learned Boundaries,” *IEEE 11th International Symposium Biomedical Imaging*, 97–100, (2014).

[77] P. Wang, S. M. Krishnan, C. Kugean, and M. P. Tjoa, “Classification of Endoscopic Images Based on Texture and Neural Network,” *Annual Reports Research Reactor Institute, Kyoto University*, 4, 3691–3695, (2001).

[78] M. P. Tjoa and S. M. Krishnan, “Feature Extraction for the Analysis of Colon Status from the Endoscopic Images,” *Biomedical. Engineering Online*, 2, 9, (2003).

[79] L. A. Alexandre, J. Casteleiro, and N. Nobre, “Polyp Detection in Endoscopic Video Using Systems,” *11th European Conference on Principles Practice Knowledge Discovery. Databases*, 4702, 358–365, (2007).

[80] T. Ghosh, S. A. Fattah, C. Shahnaz, A. K. Kundu, and M. N. Rizve, “Block Based Histogram Feature Extraction Method for Bleeding Detection in Wireless Capsule Endoscopy,” *IEEE Region 10 Conference*, 1–4, (2015).

[81] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, and D. A. Karras, “Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features,” *IEEE Trans. Inf. Technol. Biomed*., 7, 141-152, (2003).

[82] M. Benäœo And R. Hudec, “Novel Method for Color Textures Features Extraction Based on GLCM,” *Radioengineering*, 64–67, (2007).

[83] J. J. Guo, D. F. Shen, G. S. Lin, J. C. Huang, K. C. Liu, and W. N. Lie, “A Specular Reflection Suppression Method for Endoscopic Images,” *IEEE 2nd International Conference Multimedia Big Data*, 125–128, (2016).

[84] T. H. Stehle, “Specular Reflection Removal in Endoscopic Images,” *Proc. 10th International Student Conference. Electrical. Engineering. Poster*, 6, (2006).

[85] A. Ghosh, M. Arnold, S. Ameling, and G. Lacey, “Automatic Segmentation and Inpainting of Specular Highlights for Endoscopic Imaging,” *Eurasip Journal Image Video Processing*, 2010, (2010).

[86] G.Karapetyan and H.Sarukhanyan, “Automatic Detection and Concealment of Specular Reflections for Endoscopic Images,” *9th International Conference on Computer Science Information Technology*, (2013).

[87] R. Olivier and C. Hanqiang, “Nearest Neighbor Value Interpolation,” *International Journal of Advanced Computer Scence Application*, 3, 4, (2012).

[88] C. Ballesteros, M. Trujillo, C. Mazo, D. Chayes, and J. Hoyos, “Automatic Classification of Non-Informative Frames in Colonoscopy Videos,” *Lecture Notes in Computer Science*, 10125, 401–408, (2017).

[89] J. Oh, S. Hwang, W. Tavanapong, P. C. De Groen, and J. Wong, “Blurry Frame Detection and Shot Segmentation in Colonoscopy Videos,” *International Society for Optical Engineering*, 5307, 531–542, (2004).

[90] J. Oh, S. Hwang, J. Lee, W. Tayanapong, J. Wong, and. P. C. De Groen, “Informative Frame Classification for Endoscopy Video,” *Medical Image Analysis*, 11, 110–112, (2007).

[91] Z. Sun, B. Li, R. Zhou, H. Zheng, and M.Q. Meng, “Removal of Noninformative Frames for Wireless Capsule Endoscopy Video Segmentation,” *IEEE International Conference on Automation Logistics*, 294–299, (2012).

[92] N. C. van Dongen, F. van der Sommen, S. Zinger, E. J. Sekoon, and P. H. N. de With, “Automatic assessment of informative frames in endoscopic video,” *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 119-122, (2016).

[93] N. Tajbakhsh, H. Sharma, Q. Wu, S. R. Gurudu, and J. Liang, “Automatic Assessment of Image Informativeness in Colonoscopy, Abdominal Imaging, Computational and Clinical Applications,” *Lecture Notes in Computer Science*, 8676, 51-158, (2014).

[94] H, Tong, M. Li, H. Zhang, and C. Zhang, “Blur Detection for Digital Images Using Wavelet Transform,” IEEE *International Conference on Multimedia Expo* 17–20, (2004).

[95] M. Arnold, A. Ghosh, G. Lacey, S. Patchett, and H. Mulcahy, “Indistinct Frame Detection in Colonoscopy Videos,” *13th International. Machine Vision. Image Processing*, 47–52, (2009).

[96] Y. An, G. Kang, I.J. Kim, H.S. Chung, and H. Park, “Computer-Aided Diagnosis System for Colon Abnormalities Detection in WCE Shape from Focus Through Laplacian Using 3D Window,” *Second International Conference on Future Generation Communication and Networking*, 2, 46 –50, (2008).

[97] M. Cho, J.H. Kim, H.J Kong, K.S. Hong, and S. Kim, “A Novel Summary Report of Colonoscopy: Timeline Visualization Providing Meaningful Colonoscopy Video Information,” *International. Journal Colorectal Disease*, 33, 549–559, (2018).

[98] A. M. Eskicioğlu, P. S. Fischer, “Image Quality Measures and Their Performance,” *IEEE Transactions on Communications*, 43, 2959–2965, (1995).

[99] C.Y. Wee and R. Paramesran, “Measure of Image Sharpness Using Eigenvalues,” *Information Sciences*, 177, 2533-2552, (2007).

[100] H. Xie, W. Rong, and L. Sun, “Wavelet-Based Focus Measure and 3-D Surface Reconstruction Method for Microscopy Images,” *International Conference on Intelligent Robots and Systems*, 229–234, (2006).

[101] M. Carbonneaou, V. Cheplygina, E. Granger, and G. Gagnon, “Multiple Instance Learning,” *Pattern Recognition*, 77, 329-353, (2018).

[102] H. Yao, R. W. Stidham, R. Soroushmehr, J. Gryak,, K. Najarian, “Automated Detection of Non-Informative Frames for Colonoscopy Through a Combination of Deep Learning and Feature Extraction,” *41st Annual International Conference of The IEEE Engineering in Medicine and Biology* Society, (2019).

[103] A. B. Islam, M. R., Alammari, J. Oh, W. Tavanapong, J. Wong, and P. C. De Groen, “Non-Informative Frame Classification in Colonoscopy Videos Using CNNs,” *Proceedings of The 3rd International Conference on Biomedical Imaging, Signal Processing*, 2402-2406, (2018).

[104] J. Van Der Putten, J. De Groof, F. Van Der Sommen, M. Struyvenberg, S. Zinger, W. Curvers,P. H. N. De With, “Informative Frame Classification of Endoscopic Videos Using Convolutional Neural Networks and Hidden Markov Models,” *IEEE International Conference on Image Processing*, 380-384, (2019).

[105] A. Krizhevsky, Sutskever I, and Hinton G. E, “ImageNet Classification With Deep Convolutional Neural Networks,” *Neural Information Processing Systems*, 1097–1105, (2012).

[106] https://www.gastrointestinalatlas.com/, [Online], August, (2020).

[107] B. Kurt and V. V. Nabiyev, “Dijital Mamografi Görüntülerinin Kontrast Sınırlı Adaptif Histogram Eşitleme ile Iyileştirilmesi,” *VII. Ulusal Tıp Bilişimi Kongresi*, 67-79, (2010).

[108] H.Patel and P.Prajavati, “Study and Analysis of Decision Tree Based Classification Algorithms,” *International Journal of Computer Science and Engineering*, 6, 74-78, (2018).

[109] R. Safavian and D. Landgrebe, “A Survey of Decision Tree Classifier Methodology,” *Man and Cybernetics*, 3, 660-674, (1991).

[110] S. Charfi and M. E. Ansari, “Computer-Aided Diagnosis System for Colon Abnormalities Detection in Wireless Capsule Endoscopy Images,” *Multimedia Tools and Applications*, 77, 4047-4064, (2017).

[111] K. Pogorelov *et al.*, “Efficient Disease Detection in Gastrointestinal Videos –Global Features Versus Neural Networks,” *Multimedia Tools and Applications*,76, 22493-22525, (2017).

[112] C. P. Sindhu and V. Valsan, “A Novel Method for Automatic Detection of Inflammatory Bowel Disease in WCE,” *International Conference on Signal Processing, Communications and Networking*, (2017).

[113] K.Pogorelov *et al.*, “GPU-Accelerated Real-Time Gastrointestinal Diseases Detection,” *IEEE 29th International Symposium on Computer-Based Medical Systems*, (2016).

[114] L. Gueye, S. Yildirim-Yayilgan, F. A. Cheikh, and I. Balasingham, “Automatic Detection of Colonoscopic Anomalies Using Capsule Endoscopy,” *IEEE International Conference on Image Processing*, 1061-1064, (2015).

[115] D. K. Iakovidis and A. Koulaouzidis, “Automatic Lesion Detection in Wireless Capsule Endoscopy; A Simple Solution for a Complex Problem,” *IEEE International Conference on Image Processing*, (2014).

[116] S. Manivannan, R. Wang, E. Trucco, and A. Hood, “Automatic Normal-Abnormal Video Frame Classification for Colonoscopy,” *10th International Symposium on Biomedical Imaging*, 644-647, (2013).

[117] S. A. Karkanis et al., “Detecting Abnormalities in Colonoscopic Images by Texture Descriptors and Neural Networks,” *The Workshop Machine Learning in Medical Application*, 59-62, (1999).

[118] L. Peng et al., “Detecting Abnormal Regions in Colonoscopic Images By Patchbased Classifier Ensemble,” *Proceedings of the 17th International Conference on Pattern Recognition*, 3, 774-777, (2004).

[119] S. Tchoulack, J. M. P. Langlois, and F. Cheriet, “A Video Stream Processor for Real-Time Detection and Correction of Specular Reflections in Endoscopic Images,” *North East Workshop on Circuit and Systems and TAISA Conferences*, 49–52, (2008).

[120] W. Aydi, N. Masmoudi, and L. Kamoun, “New Corneal Reflection Removal Method Used in Iris Recognition System,” *International Journal of Electronics and Communication Eng*ineering, 5, 697-700, (2011).

[121] Y. Fan, M. Meng, and Li Baopu, “A Novel Method for Informative Frame Selection in Wireless Capsule Endoscopy Video,” *International Conference IEEE Engineering in Medicine and Biology Society*, (2011).

[122] S. Arivazhagan, L. Sylvia, W. Jebarani, and V. J. Daisy, “Categorization and Segmentation of Intestinal Content and Pathological Frames in Wireless Capsule Endoscopy Images,” *International Journal of Imaging Robot*, 13, 134–147, (2014).

[123] M. Grega, M. Leszczuk, M. Duplaga, and R. Fraczek, “Algorithms for Automatic Recognition of Non-Informative Frames in Video Recordings of Bronchoscopic Procedures,” *Advances in Intelligent and Soft Computing*, 535-545, (2010).

[124] N. Rangseekajee and S. Phongsuphap, “Endoscopy Video Frame Classification Using Edge-Based Information Analysis Pre-Processing,” *Computers in Cardiology*, 38, 549–552, (2011).

[125] Y. H. An, S. Hwang, J. H. Oh, J. K. Lee, W. Tavanapong, P. C. De Groen, and J. Wong, "Informative-Frame Filtering in Endoscopy Videos," *Medical Imaging, Image Processing*, 5747, 291–302, (2005).

GCRIS