QATAR UNIVERSITY

COLLEGE OF ENGINEERING

CANCER DETECTION AND IDENTIFICATION ON SCARCE AND LOW-RESOLUTION DATA

BY

SUCHITHRA KUNHOTH

A Thesis Submitted to

the Faculty of the College of

Engineering

in Partial Fulfillment

of the Requirements

for the Degree of

Masters of Science in Electrical Engineering

June 2018

COMMITTEE PAGE

The members of the Committee approve the Thesis of Suchithra Kunhoth

defended on 21/05/2018.

_____

Dr. Mustafa Serkan Kiranyaz, Dr. Somaya Al Maadeed
Thesis Supervisor

_____

Dr. Nader Meskin
Committee Member

_____

Dr. Uvais Qidwai
Committee Member

_____

Dr. Selim Aksoy
Committee Member

Approved:

_____

Khalifa Al-Khalifa, Dean, College of Engineering

ABSTRACT

KUNHOTH SUCHITHRA., Masters : June : 2018, Masters of Science in Electrical Engineering

Title: Cancer Detection and Identification on Scarce and Low- Resolution Data

Supervisor ofThesis: Mustafa Serkan Kiranyaz, Somaya Al Maadeed

Machine learning algorithms have been contributing immensely in the biomedical sector with the innovation of several automatic and semi-automatic diagnostic devices. They are well suited for applications such as the diagnosis of cancer, which is a prevalant and devastating disease nowadays. A computer-aided diagnostic system can detect and classify various tumor tissues and thereby ensures a reliable and rapid screening procedure. They serve as an additional confirmatory tool which is independent of pathologist expertise and experience. In this thesis study, we perform comparative evaluations among several recent approaches for cancer detection and identification on a scarce and low-resolution biopsy image dataset. The biopsy samples comprise of normal as well as cancerous colorectal tissues, collected from the Al-Ahli hospital, Qatar. We have built two separate image datasets, multispectral and optical from the collected samples. Using our multispectral image acquistion system, images are acquired in various wavelength bands spanning from visible to near infrared to build the first dataset. The second dataset is composed of optical images (in RGB raw format) of the same samples.

A Multispectral image based tumor identification system was developed using rotation invariant Local Phase Quantization technique and Support vector Machine (SVM) classifier. The comparative evaluations demonstrate that it could outperform Local Binary Pattern for the feature extraction and Random Forests (RF) in classification of the

colorectal tumor tissues. We compared the classification accuracies yielded with the two image modalities- multispectral and RGB and the former one exhibited higher accuracy. Furthermore, we have presented a band selection strategy to eliminate the redundant bands from the multispectral imagery. This could reduce the computation time along with improving the classification accuracies.

As the main contribution of the thesis, we propose a compact and adaptive CNN approach for the detection and identification of the tumor tissues on the RGB image dataset. This approach is fully automatic with the absence of any manual pre-processing, tuning or prior segmentation phase to aid the classification algorithm. Its performance is compared against the SVM classifier with three different kernel types and five state-of-the-art texture feature extraction methods including rotation invariant Local Binary Pattern, rotation invariant Local Phase Quantization, and Haralick features. The proposed systematic approach with adaptive and compact CNNs and the top performing conventional method with the best texture feature have achieved the highest identification accuracies with respect to the task of discriminating four classes of colorectal tissues. However, the proposed method has achieved the highest cancer detection performance, around 94.5%, as compared to the best detection score of 87% achieved by the best conventional method. This is despite the fact that the proposed method used low-resolution image data (64x64 pixels) in contrast to the original patch resolution (300x300) used by the conventional methods. Finally, the proposed approach can further exhibit a superior computational complexity and minimal false alarms. The promising results throw light on the competence of adaptive CNNs for cancer detection in low-resolution images from a limited dataset.

To my dear parents….

ACKNOWLEDGMENTS

Firstly, I would like to thank Qatar National Research Fund for the grant provided through National Priority Research Programme (NPRP) No. 6-249-1-053. I am gratefully indebted to their financial support, which helped me to accomplish the whole thesis work.

I would like to thank my supervisor, Prof. Mustafa Serkan kiranyaz for the successful completion of my thesis. I am extremely grateful for the valuable comments, which greatly helped to foster my technical writing skills. I am thankful for the time and effort he put to review my thesis. I must express my profound gratitude to my Co-Supervisor Dr. Somaya Al Maadeed for her valuable support she offered me during my entire Master's course. She steered me in the right direction to achieve my goals. I recognize her support, which helped me a lot to balance my job and studies. This accomplishment would not have been possible without her who gave me the motivation to join the course.

I would also like to acknowledge all my friends in Qatar University for their help and assistance throughout my years of study. Special Thanks to my friend Christofer Fellicious, who did the proof reading for my thesis, which greatly helped to improve my writing.

Above all, I offer my deepest gratitude to my family members for their love and prayers. I thank extremely to the continuous encouragement of my father and mother. Thanks to my husband and mother in law as well. Finally to my daughter, Baby Nishika who gave the utmost support to finish my course and thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Cancer has been found to be the second leading cause of death all over the world. It is responsible for 8.8 million deaths in 2015[1]. This statistic indicates that nearly one out of six deaths worldwide is attributed to cancer. Based on the incidence rates, lung cancer holds the first position followed by breast and colorectal cancer. According to the worldwide statistics for mortality rates from cancer in the year 2012, lung cancer (1.6 million deaths), liver cancer (745,000 deaths), and stomach cancer (723,000 deaths) come in the first three positions [2]. It has turned out to be a dangerous disease because of its increasing probability and high mortality rate. Medical researches have provided only specific treatment approaches that can increase the life expectancy to a small extent. It is evident that timely screening reduces the impact on patient's life, but appropriate attention is not paid to the matter [3]. Although several diagnostic methodologies were proposed in the past, histopathological techniques or biopsy analysis is considered to be the major screening tool for majority of the cancer types.

Cost-effective diagnostic methodologies have been introduced with the rapid technological advancements in the field of image processing and machine learning [4]. Automatic inspection of the biopsy sample serves to screen a larger mass of population within a limited time while maintaining or improving the diagnostic results. Computer aided analyses of the biopsy images serve to minimize the dependency of diagnostic accuracy on any specific individual. As far as a pathologist is concerned, it is a tedious task to inspect and analyse several tissue samples. The complexity of the task elevates it further when it comes to the case of categorizing the cancerous tissues into specific classes based on their malignancies. This monotonous task of investigation under the microscope for

continuous hours can produce erroneous results which can be either false positives or false negatives. Moreover, the inter-personal variation in the analysis has a prominent role in the diagnostic outcome. The expertise and experience of the pathologist will be counted towards the accuracy of detection and identification of cancer tissues [5]. Automated analysis can improve the detection accuracy along with the inter operator variation problem being solved. Automated cancer screening tools aid to deliver accurate results in a short period of time.

Histopathological analysis, which is the major diagnostic approach followed in cancer screening can be accomplished by a pattern recognition system. In those approaches, features are extracted from the digitized biopsy samples and fed to a suitable classifier which demarcates the cancerous lesions from the non-cancerous ones. Having an efficient algorithm, it can even discriminate between the different tissue types or tumor grades. The image acquisition can follow two main modalities such as optical, e.g., color (RGB) or grayscale as well as multispectral imaging techniques. Since the overall object characterization with RGB imaging is limited by the use of only three color channels, multispectral methods seem to be more promising. Several pre-processing steps commonly incorporated in these algorithms include image enhancement, segmentation of the prospective tumor region, etc. The image enhancement is meant to rectify some issues that may arise as a result of non- uniform illumination conditions during the imaging process. Stain normalization has also been suggested for histopathology images in order to cope with problems such as color differences in the tissue images due to varying conditions of tissue preparation, stain reactivity variations when using the stains from different manufacturers, individual dependencies and the use of microscopes from different

manufacturers or different modes of lighting [6]. Segmentation phase is incorporated to emphasize the glandular regions alone from the tissues for extracting relevant features. Automated methods for segmentation have proved its efficiency in certain instances. Commonly, the segmentation is accomplished manually or with certain semi-automatic techniques. Different tissue types can be distinguished with the extracted features such as color, size and shape of gland and nuclei, statistical measures, and texture characteristics etc. As explained here, the general schematic of pattern recognition based cancer tissue classification system follow the processing stages as shown in the Figure 1. Feature extraction is followed by the classification stage. State-of-the-art classifiers such as Support Vector Machines (SVM), Random Forests (RF), Artificial Neural Networks (ANN) are usually deployed to realize the supervised learning strategies. In recent years, deep convolutional neural networks (CNNs) have become a major tool for image classification and object recognition tasks including several medical applications [7-13]. Compared to conventional classification approaches, CNNs have proved to yield state-of-the-art accuracies with significant performance gaps.

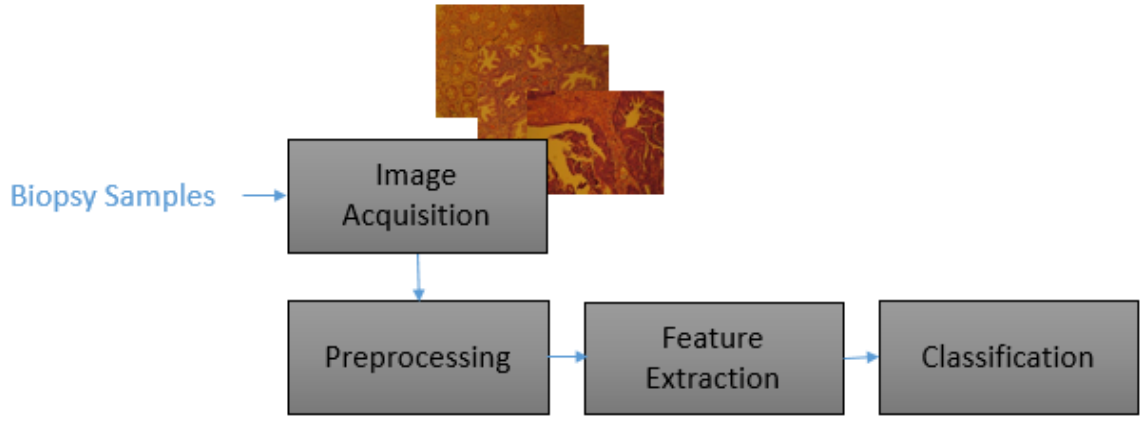


Figure 1. Basic schematic for cancer detection/identification from biopsy images

1.1 Objectives

This thesis study aims to perform comparative evaluations among several recent approaches for cancer detection and identification using our colorectal biopsy image datasets. The set of biopsy samples comprise specimens from 4 major categories of colorectal tissues including one normal and three abnormal (cancer) classes. The samples were acquired using both multispectral and RGB modalities. The prospective cancerous regions were identified by the pathologist to serve as the ROIs for our classification and identification algorithms. The main objectives of the thesis work can be summarised as follows:

1) Investigation of the performance of texture based feature extraction methods including LBP, LPQ and their variants along with classifiers such as SVM and RF

for cancer tissue classification. Experiments should be conducted on the multispectral images

2) Comparison of the two different imaging modalities: RGB and multispectral imaging in the context of cancer identification
3) The development of a band selection methodology to select the relevant multispectral image bands alone for further processing
4) The development of a compact adaptive CNN approach for the cancer detection and identification task, which makes use of RGB image dataset for the validation. The results are further compared with the conventional methods. We have considered 5 different feature extraction techniques: rotation invariant LPQ (rLPQ), rotation invariant LBP (rLBP), uniform rotation invariant LBP, Haralick and the combined rotation invariant LPQ- rotation invariant LBP feature and SVM classifier with 3 different kernel types, Radial Basis Function (RBF), Linear and Polynomial. Texture feature based classification needs to be carried out on comparatively higher resolution images than that used for CNN. Both the cancer detection (normal- abnormal tissue discrimination) and the identification (4- class classification) should be performed with the conventional and CNN approaches.

The major highlights of the proposed adaptive CNN scheme include:

- Dataset with limited number of biopsy images; unlike very large number of images as with deep CNNs
- Low- resolution image patches utilized in CNN experiments to validate the algorithm efficiency on low- resolution images

- No manual pre-processing and segmentation phase
- No additional hardware requirements for our compact adaptive CNNs with limited complexity

1.2 Contributions

We propose a systematic approach using a compact and adaptive CNN for the detection and identification of the tumor tissues. Unlike the deep CNNs, the proposed network configuration could work effectively with low-resolution patches generated from a limited dataset. The adaptive property eliminates the restrictions imposed on the input image size as well as the number of hidden layers. Being compact, the computational complexity can be reduced. We have developed a biopsy image database with both multispectral and RGB image counterparts. It is comprised of 200 images, which is reasonably good compared to the existing datasets. The variability within the dataset is guaranteed well with our biopsy samples of 151 patients, which is the highest number of patients ever used for this challenging task. The images are distributed equally in each of the 4 classes. Our algorithm is devoid of any prior manual or semi-automatic image enhancement or segmentation stages. In this study, results of the proposed CNN based approach are further compared against five conventional classification methods each based on a state-of-the-art texture feature.

1.3 Thesis Outline

We commence with a review of the related works in Chapter 2, where different methodologies available for cancer detection and identification till date will be discussed. The contributions of our work are also introduced with a view to address the limitations of

existing methodologies. Chapter 3 introduces the datasets, medical aspects of the concerned samples as well as the data acquisition and augmentation techniques that we followed. The image acquisition techniques for both the RGB and multispectral imaging are detailed in this chapter. All the experiments on our multispectral image dataset are summarised in the Chapter 4. The comparative study on multispectral and RGB techniques, and the proposed band selection strategy is also included in this chapter. The proposed approach with the adaptive and compact CNNs is presented in Chapter 5 with relevant technical explanations and implementation details. The conventional methods used for the comparative evaluations of the proposed approach are also discussed here. Chapter 6 presents the experimental results along with the comparative evaluations. Chapter 7 summarizes the work with the relevant conclusions we reached. Finally, the scope of some potential future work is discussed.

CHAPTER 2: BACKGROUND AND RELATED WORK

Several computer vision approaches have attempted to detect and classify the tissues related to different types of cancers. Based on the feature extraction techniques adopted, they can be classified as belonging to texture, morphological, and approaches based on Deep Learning. The related work in the literature is organized in the succeeding section based on this categorization.

2.1 Texture features

The whole slide histopathologic analysis in [14] is based on the well-known grey level co-occurrence (GLCM) textural features. The classification system was built upon a support vector machine (SVM) algorithm to differentiate normal and cancerous colorectal tissue samples. The study in [15] aims to compare the performances of various texture features including local binary patterns, Haralick features and local intensity order patterns for multispectral image based colorectal tissue classification. It was the local binary patterns when combined with Support Vector Machine classifier that gave the best classification accuracy of 91.3%. A range of texture features including Markovian co-occurrence matrices, run-length analysis, Laws features, wavelet decomposition, and Fourier analysis was evaluated in [16] to predict the risk of breast cancer. The automated breast cancer identification system in [17] relies on textural and architectural image features. Both cancer detection and identification were carried out on a dataset of 48 breast tissue biopsy slide images. Gabor features turned out to exhibit the highest discriminatory capability for the task. Chaddad et al. [18] proposed an improved version of the snake algorithm for segmentation followed by Haralick texture feature extraction from the segmented images. A four-class identification schema for the colorectal tissue types such

as Carcinoma (Ca), Intraepithelial Neoplasia (IN) and Benign Hyperplasia (BH) was experimented further in the work. Melanoma skin cancer detection using GLCM features and multilayer perceptron [19] could be performed with 92% accuracy when validated on a dataset with 102 dermoscopy images. The methodology in [20] follows an approach involving active contour segmentation of pathologic ROI followed by the extraction of texture features such as discrete wavelets, GLCM etc. A three-class identification system was implemented on a database of 480 images generated from 30 patients. Texture characterization with textons enabled to build the automated Gleason grading system in [21]. The framework for texture classification relied on two simple stages: filtering and clustering. Further, Gleason grades 3 and 4 were discriminated with the support of an SVM classifier. In [22], the concept of gray level co-occurrence matrix is extended for the texture characterization in multiband images and hence to discriminate the healthy and pathological prostatic tissues. In order to reduce the multispectral band representation, a band selection technique was employed which selects the best relevant bands of multispectral prostate cancer database. The system in [23] was intended to analyze the significance of multispectral imaging in order to discriminate cancer and non-cancer nuclei in liver tissues. Textural features were extracted using Gabor descriptors, which include 30 Gabor patterns at different scales and orientations. The results have indicated approximately 99% classification accuracy. It was the images taken in wavelength bands 418-467 nm, 481-513 nm and 548-641 nm that proved efficient to classify normal and hepatocellular carcinoma in high-magnification. Qi et al. [24] have conducted a performance study to compare the classification accuracy using both multispectral imaging (MSI) and standard bright-field imaging (RGB) to characterize breast tissue microarrays.

Feature extraction comprised texton library training and histogram construction. Experimental results showed that the multispectral images delivered a higher classification rate than RGB images. Meaningful features offered for classification by multispectral imaging technology is well exhibited in the two latter works.

2.2 Morphological features

Apart from the texture features, color and morphometric characteristics were also incorporated for the Gleason grading of prostatic biopsy images in [25]. Performance of several classification algorithms such as Bayesian, nearest neighbor, and SVMs was also investigated. Detection of cancer tissues in [26] follows a series of stages: microscopic image enhancement, segmentation of background cells, feature extraction, and finally the classification. Shape and color-based features are utilized along with several texture descriptors including Tamura's feature and Law’s texture energy features. The study in [27] has proposed a method for colorectal cancer tissue classification using shape features and nearest neighbor classifier. Area, Xor Cell-Convex, and Solidity features were proved out to efficiently differentiate the carcinoma tissues from the other cancer types. Color and shape geometry features were utilized for the skin cancer detection and identification in [28]. The algorithm involved an automatic skin lesion segmentation stage as well to separate the unwanted background prior to feature extraction. A hybrid feature set encompassed of geometric characteristics, morphological features, traditional SIFT, etc., were applied on 174 colon biopsy RGB images to implement a cancer detection system [29]. An accuracy of 98% was achieved with the help of a Radial basis function (RBF) kernel SVM classifier. Structural features such as roundness and shape distribution were considered for the classification of prostatic biopsy images into five Gleason grades [30].

A prior segmentation stage is incorporated in the algorithm to identify the glandular regions. The classification follows a tree structure approach with each stage performing a binary classification using intensity and shape features. Evaluation on two different datasets with 91 and 199 images produced 95% and 85% accuracy respectively. The methodology in [31] also commences with the glandular segmentation using the color space information. Fifteen structural features were extracted from the glandular regions in order to classify them as benign, grade 3 and grade 4 tissue types. Various state of art classifiers like nearest neighbor, decision tree, SVM etc., were experimented in the work. A segmentation and identification schema for hepatocellular carcinoma (HCC) biopsy images was proposed in [32]. In order to remove noise and highlight the nuclear shape, a dual morphological grayscale reconstruction method was adopted. Watershed transform and snake models were applied for the nuclei segmentation. A set of 14 features including but not limited to nuclear size, nucleocytoplasmic ratio, irregularity, and haralick were extracted from the segmented nuclei to perform the classification task. With the aid of an SVM based decision graph classifier, the proposed approach could yield an accuracy of 94.54% on a dataset of 804 images with 5 different grades of tissues. Another work for prostate cancer grading have been done with morphological features from individual glands [33]. Automated gland boundary segmentation was done with level set method followed by the extraction of seven boundary features. Utilizing the SVM classifier, the algorithm yielded an accuracy of 75% in distinguishing benign epithelium and grade 3, 85.71% between benign epithelium and grade 4, and 72.73% between grade 3 and grade 4. The dataset comprised of 29 tissue patches. Similar research has been carried out for 4 class grading on 54 tissue image patches [34]. The system extracts architectural, nuclear density,

gland morphology as well as texture based features from the tissue patterns to classify the biopsy images according to the 4 Gleason grades. The SVM classifier could result in an accuracy of 92.8% when discriminating between Gleason grade 3 and stroma, 92.4% between stroma and epithelium, and 76.9% between Gleason grades 3 and 4. The methodology in [35] relies on the concept of statistical shape model of manifolds (SSMM). They make use of the morphologic features of glands from prostatic specimens to automatically distinguish between Gleason grades 3 and 4. Khalid et al [36] presented a binary grading system for prostate cancer tissue types. The algorithm commences with a segmentation stage based on textural features. Architectural changes in nuclei and stroma, luminal characteristics etc. were captured in the feature extraction phase. The classification results on a dataset of 88 ROIs from 30 whole slide images exhibited an accuracy of 97.6%.

2.3 Approaches based on Deep Learning

In recent years, Deep Learning with CNNs has become the de-facto standard for image classification and object recognition tasks including medical applications. But the major drawback lies in the necessity of a massive size dataset for the training of such deep networks usually with 100M or above parameters. This turns out to be a prominent issue since the biomedical datasets are often available with a limited number of samples. A four-class identification strategy for breast biopsy images using deep CNNs was proposed in [37]. Features extracted by CNNs were used to differentiate normal tissue, benign lesion, in situ carcinoma and invasive carcinoma. An augmented dataset of 70000 different patches was generated from the original 250 training images by applying arbitrary rotations and mirroring. Then each of the patches was considered to have the same label as the original image. Algorithm evaluation of 20 images yielded 77% accuracy for four class

identification and 80% accuracy for non-carcinoma vs carcinoma discrimination. When the features extracted by CNN were trained on an SVM classifier, the two class discrimination accuracy could be improved up to 83%. A benign vs malignant classification of biopsy images of the breast was performed in [38]. The dataset was composed of 7909 images with nearly 2000 image belonging to each magnification level from 40, 100, 200 and 400. Four different patch generation approaches were applied to yield a minimum of 54 and maximum 1000 patch per original image. The training was done patchwise and testing involved the accumulation of results over the individual patches of the original image. The maximum accuracy of 90% was reported which is obtained for 40x magnification level. Detection and classification of cell nuclei in cancerous tissue is also a prominent task, which was accomplished using a spatially constrained convolutional neural network in [39]. Multispectral colorectal images were classified in [40] into one of the three classes using CNN approach. The highest accuracy of 99.17% was achieved in the segmented image regions. However, the dataset was very small comprising 30 images in total and the segmentation was accomplished using a semi-automatic procedure. The database of [38] was deployed for the realization of a multi-class breast cancer classification system in [41]. A class structure based deep CNN was proposed in the work to perform the classification among 4 different grades each of benign and malignant classes. Data augmentation was carried out for building the training partition of the database. The average accuracy of the patient level testing was 93.2%, and image level testing was 93.8% among all the magnification factors. Another deep learning method was proposed in [42] to perform a magnification independent classification of breast cancer tissues. The same database of [38] was utilized here also. An average recognition rate of 82.13% was achieved for

benign/malignant classification problem. The multi-task CNN architecture introduced here was able to predict both the image magnification level and its malignancy status simultaneously.

Many works in the literature lack proper validation because of the insufficient dataset. Especially the deep CNNs require datasets with massive sizes, e.g., in the “Big Data” scale. In fact, this requirement alone makes deep 2D CNNs inapplicable to many practical problems that have limited datasets including the problem addressed in this study [40]. The variability of the dataset is also another aspect which depends on the number of patients from which the biopsy samples are collected. Even though [20] and [29] indicate a dataset of more than 100 images, they are taken from only 30 and 68 different patients, respectively. A slightly larger dataset of [38] is generated from the biopsy samples of only 82 patients. Several works [18,20,28,40] have incorporated a segmentation stage in order to remove the background efficiently. This is basically needed to localize the tumor regions prior to feature extraction. But some of them rely on semi-automated or even manual procedures, which obviously reduces the practical usability of the proposed approach.

CHAPTER 3: COLORECTAL TUMOR DATASETS

We have used a colorectal tissue dataset comprising four different types of tissues for the validation of the considered algorithms. The medical aspects ([43-45]) of colorectal cancer tissues is briefly explained in the Section 3.1.The image acquisition and data augmentation procedures are detailed next.

3.1 Colorectal Tumor: Medical Aspects

This disease affects the colon or rectum, the parts of the large intestine. Initially, most colon cancers appear as colorectal polyps, which refer to the abnormal growths inside the colon or rectum. Although they are considered to occur without a malignant property, neoplastic polyps found in the colon can be the precursor of the colorectal carcinoma. The different types of colonic polyps based on their malignancy potential are described in the succeeding sections.

3.1.1 Non-neoplastic polyps

An abnormal mucosal maturation or inflammation can lead to the formation of the non- neoplastic polyps. These make up about 90% of the epithelial polyps in the large intestine of persons above the age of 60. The different types of non- neoplastic polyps are as presented below:

1) Hyperplastic polyps

These are the small nipple-like, hemispherical, smooth protrusions of the mucosa which may appear individually or in multiple numbers. Approximately half of such polyps occur in the recto sigmoid region. Although they are not dangerous, some hyperplastic polyps on

the right side of the colon can be the precursors of the colorectal carcinoma. This transformation occurs due to a series of genetic mutations in these lesions. Most hyperplastic polyps are less than 5mm in size and are characterized by cells having a hypermature appearance.

2) Hamartomatous polyps

These are the benign malformations which have the same growth rate as that of a normal tissue.

a) Juvenile Polyps:- Hamartomatous proliferations of the lamina propria are typically found in younger children. It is characterized by large lesions of 1-3 cm in diameter which are rounded, smooth or lobulated and having a stalk up to 2 cm long. It commonly occurs individually in the rectum. Based on several factors such as inheritance method, severity and degree of involvement of GI tract, the Juvenile Polyposis Syndrome (JPS) is subdivided into four subgroups such as Infantile Polyposis Syndrome, Juvenile Polyposis Coli, Generalized JPS and Gastric JPS .

b) Retention Polyps:- These refers to the Juvenile Polyps in adults, but are comparatively smaller lesions.

c) Peutz- Jeghers Polyps:- These polyps develop as a result of an autosomal dominant Puetz- Jeghers syndrome, which in turn occurs by germ- line mutations in the LKB1 gene that encodes a serine threonine kinase.

3) Inflammatory Polyps

These refer to the non- dysplastic crypts formed as a result of some injury to the mucosa. The number and the appearance of such polyps may be dependent on the level of severity of the causative factor. Mucosal prolapse and the inflammatory bowel disease are the

commonly described reasons for the arrival of such inflammatory polyps.

3.1.2 Neoplastic polyps

In contrast to the non-neoplastic type, these have a malignancy nature. Neoplastic polyps are those with a higher probability to develop cancer.

1) Adenomas

Epithelial proliferation and dysplasia may lead to the formation of the malignant neoplastic polyps, which are the adenomas. There is an equal chance of occurrence in both males and females and the hereditary aspect also prevails in case of sporadic adenoma. The malignancy with the adenomatous polyp may depend on the polyp size, histologic architecture and the severity of epithelial dysplasia. Adenomas may have a larger size ranging from few millimeters to several centimeters compared to the hyperplastic polyps. Among these, the maximum diameter is a prominent feature that affects the adenoma to carcinoma transformation. The three subtypes of adenomatous polyps are:

a) Tubular Adenoma:- It is the most common type among the three adenomas. It appears small and pedunculated but the smallest tubular adenomas are sessile. The larger ones will have slender stalks and raspberry like heads.

b) Villous adenoma:- They are the dangerous ones and mostly found in older persons. A large and sessile structure is the characteristic feature of villous adenoma. They appear as villiform extensions of the mucosa covered by dysplastic columnar epithelium and 40% of these lesions tend to have invasive carcinoma

c) Tubulovillous adenoma:- The appearance, features and also the risk of these adenoma lies intermediate between the tubular and villous adenoma.

Other than having the pedunculated or sessile structures, another non- polypoid group called flat adenomas were discovered later. They are sometimes almost depressed and are found to exhibit a high grade dysplasia.

2) Familial Polyposis Syndrome

Familial polyposis syndrome is an autosomal dominant disorder that may later turn out to a malignant situation. A number of 500 to 2500 colonic adenomas, both tubular and villous may develop over the mucosal surface in this Famililal Adenomatous Polyposis disease (FAP). The genetic defect due to APC gene on chromosome 5q21 was found out to be the causative factor of FAP.

3.1.3 Colorectal carcinoma

The adenomatous polyps may transform into the malignant adenocarcinoma if not cured at any early stage. Males are 20% more affected when compared to females. The development of carcinoma from adenomatous lesions is directly related to the number of adenomas. Thus the probability of cancer is higher in patients with the Familial Polyposis Syndrome.

The whole classification scheme described here for the colorectal polyps/tumor can be illustrated in the Figure 2.

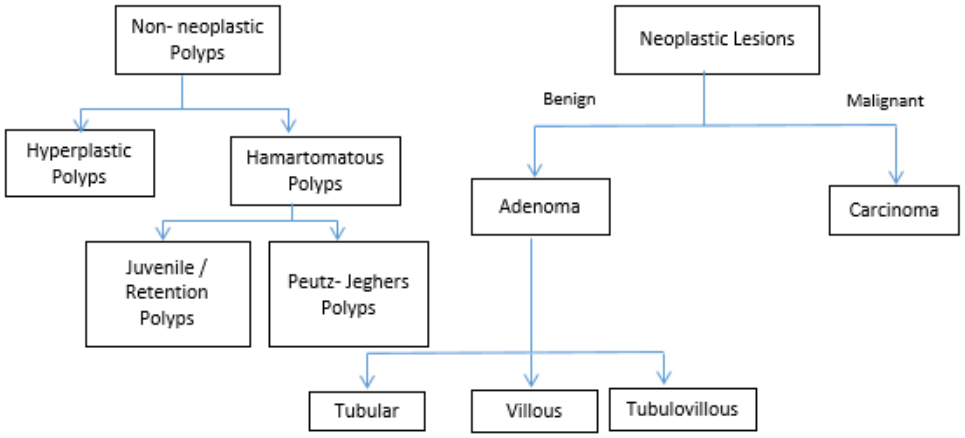


Figure 2. Colonic tumor/abnormalities classification

3.2 Colorectal Biopsy Samples

The colorectal tissue slides and the concerned medical details were collected from the Pathology and Laboratory Medicine lab at Al-Ahli hospital, Qatar. This study was reviewed and approved by Qatar University's Institutional Review Board (QU-IRB). A total of 164 tissue samples collected between the years 2007 and 2016 were provided by the hospital from 151 different patients. Each sample belongs to one of the four classes of colorectal tissue: Normal, Hyperplastic polyp (HP), Tubular Adenoma with low-grade dysplasia (TA_LG) and Carcinoma (CA). The distribution of images with respect to the

number of patients as well as the number of biopsy slides can be deduced from Table 1.

Table 1. Image distribution with respect to number of patients and slides

	Normal	HP	TA_LG	CA	Total
No. of images	50	50	50	50	200
No. of patients	31	36	38	49	151
No. of biopsy slides	38	38	38	50	164

The tissues were initially fixed overnight in 10% neutral buffered formalin. It is followed by placing them in an auto processor where it will be subjected to different concentrations of ethyl alcohol, xylene, and paraffin. After that, the tissue is embedded in paraffin and cut by microtom into 5 micron thickness. Finally, they are mounted on glass slides and stained by hematoxylin and eosin stain.

Except for the normal class, specific marking was done by the consultant histopathologist in order to indicate the ROI belonging to the concerned class. Normal class stands for the non-cancerous, non-malignant class without any disease symptoms [43]. Hyperplastic polyp belongs to the non-neoplastic polyp category. Although they are harmless, some hyperplastic polyps on the right side of the colon can be the precursors of

the colorectal carcinoma. Tubular Adenoma falls under the neoplastic polyp category. The malignancy with the adenomatous polyp may depend on the polyp size, histologic architecture and the severity of epithelial dysplasia. The adenomatous polyps may transform to the malignant carcinoma if not cured at an early stage.

3.3 Image Acquisition

We have built two different setups in our Lab in order to prepare multispectral and RGB databases of the samples. Specific sensors and microscope were used for the image acquisition. Both types of images are captured under transmission microscopy mode. ROIs captured in the multispectral imaging case were recorded manually and exactly the same region was acquired with the RGB imaging setup also. Both techniques are detailed in the following subsections.

3.3.1 Multispectral Imaging (MSI)

Our work [46] presents the multispectral image acquisition system to develop a database for the colorectal biopsy slides. Each biopsy sample is acquired in different wavelength bands corresponding to both visible and near infrared spectrum. With the acquired partial database, a preliminary experiment was carried out which involved an LBP algorithm with RF classifier.

We followed the tunable filter approach to build our multispectral imaging system. Since a single tunable filter covering the whole range from visible to near infrared is not available, we had to setup two different arrangements for acquiring images in both ranges. Our detector is the Xeva-1.7-320 VisNIR digital camera with a thermoelectrically cooled InGaAs detector head. The model XEN000165 we are using offers a frame rate of 60Hz,

with a camera link digital output interface. It supports spectral imaging in the range of 400 to 1700 nm. The Varispec liquid crystal tunable filter (LCTF) VIS-20-20, LNIR-20-20 enables the hyperspectral acquisition in the visible and near infrared bands respectively. The tuning of wavelength in the desired range is enabled by the Varispec filter software. Image acquisition is controlled by the Xeneth Software delivered with the camera.

The ROI to be acquired is decided based on the proper recommendation from the histopathologist. Once the bands in the visible range are acquired, the co-ordinate marking scheme in the zeiss microscope allows capturing the same region over the NIR bands also, after replacing the VIS LCTF (visible range) with the LNIR LCTF (near infrared range)

1) Visible Range Acquisition

Acquisition of hyperspectral images in the visible range involves the deployment of Xenics camera, VIS Filter, VIS Relay lens as well as the microscope. Zeiss axioscope a1 microscope with halogen illumination was employed for the 10x magnification of specimens. The imaging setup used for multispectral image acquisition in the visible range is shown in Figure 3

The usage of the microscope (objective lens) at the end implies that the lens is not used just in front of the camera. This particular setup demands the usage of an additional optical component, a relay lens so as to reduce the vignette. The relay lens also supports easy focus adjustment. The VIS Filter software allows the tuning of filter's wavelength from 400 to 720 nm so that corresponding images in any of this desired wavelength range can be acquired.

Thirteen multispectral visible bands were captured for each specimen at wavelength intervals of 20nm starting from 470nm.

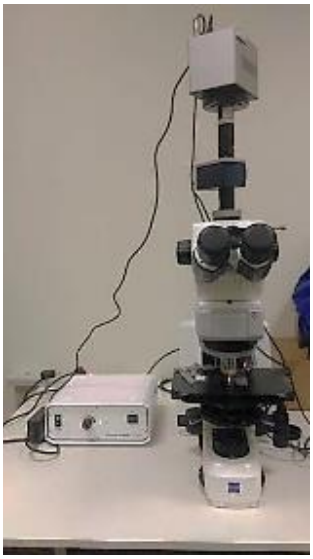


Figure 3. Multispectral image acquisition in the visible band

2) Near Infrared Acquisition

Acquisition of hyperspectral images in the near infrared range involves the deployment of Xenics camera, LNIR Filter, LNIR Relay lens as well as the microscope. The same objective lens of 10x specification was used here too. During the infrared imaging, the heat protection filter arranged in the light path of the microscope for visible range acquisition had to be removed. We captured 26 multispectral near infrared bands for each specimen at wavelength intervals of 20nm ranging from 1150nm to 1650 nm.

The resolution of our Xenics camera is 320x256 pixels, which resulted in a 320x256x39 multispectral cube for each specimen. A total of 200 images with 50 images per class contributed to our MSI database. Some of the bands captured in VIS as well as

NIR range for a ‘normal’ colorectal tissue are shown in Figure 4.

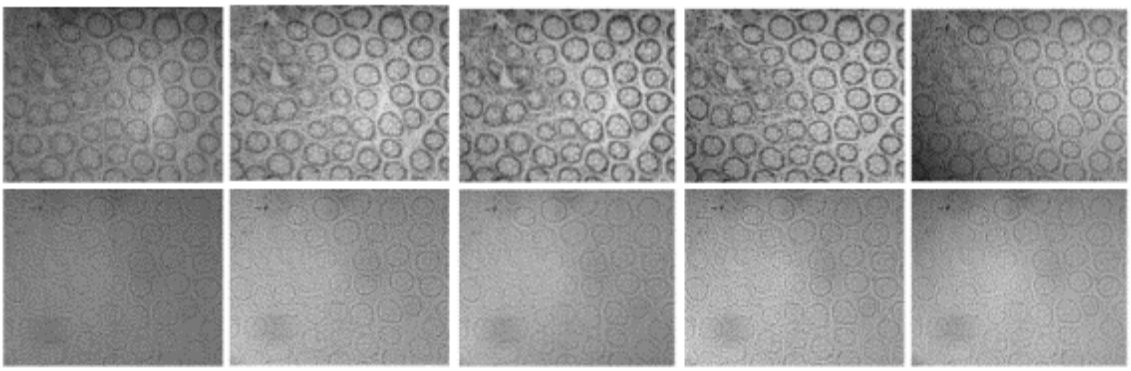


Figure 4. Multispectral image bands of a normal tissue: VIS Range (Top), NIR range (Bottom)

3.3.2 Standard RGB Imaging

The biopsy slides were digitized using a Canon Power Shot A650IS digital camera mounted over the halogen illuminated zeiss microscope (Figure 5). The resolution of the digital camera was 640x480 pixels. The images were acquired with the setting of a 10x objective lens (Figure 6). The same set of 200 images in MSI database was captured with the RGB setup.

Figure 5. RGB Image Acquisition setup

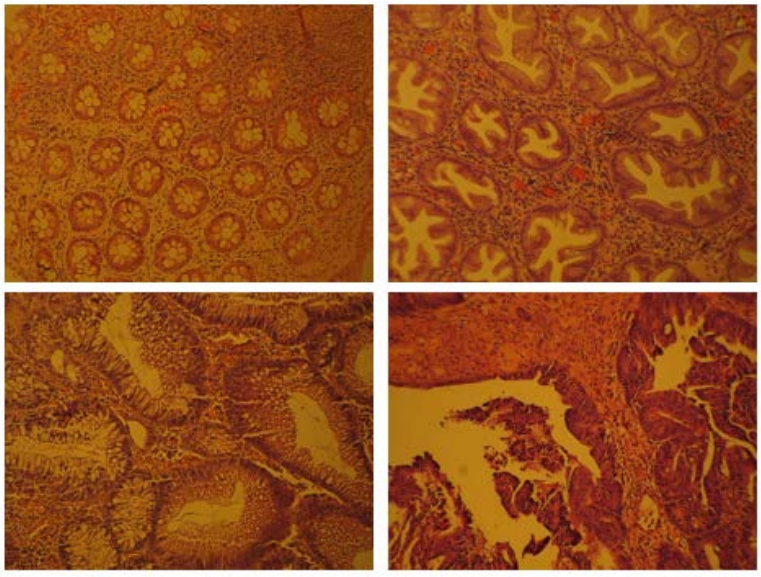


Figure 6. Normal Tissue (Top Left), Hyperplastic polyp (Top Right), Tubular Adenoma with low-grade dysplasia (Bottom Left), Carcinoma (Bottom Right)

3.4 Visual Inspection of biopsy images

Having a look at Figure 6, it seems that the visual characteristics of glandular structures differ greatly between the classes. The glands in the normal tissue often appear circular in shape. Depending upon whether transverse or longitudinal cross section was considered during biopsy, it may change to slightly elliptical shape. The structure will be well organized within the tissue. The hyperplastic polyp tissues appear to be a transformed version of the normal tissues with elongated glands. It is also characterized by a star shaped lumen. In tubular adenoma with low grade dysplasia, there will be thickly populated cells on the epithelial lining. The size and number of these cells may vary. The carcinoma tissue looks much dissimilar to other classes. We cannot identify individual glands within quite a disorganized structure.

Now we can have a closer look on a series of images belonging to each of these classes (Figure 7). We can infer the high complexity underlying the task of automatic discrimination between the four types of tissues. It is obvious that there are several instances of inter-class similarities and intra-class variations. Among the classes TA_LG and CA there tend to be a great deal of confusion. It can be inferred from the images below.

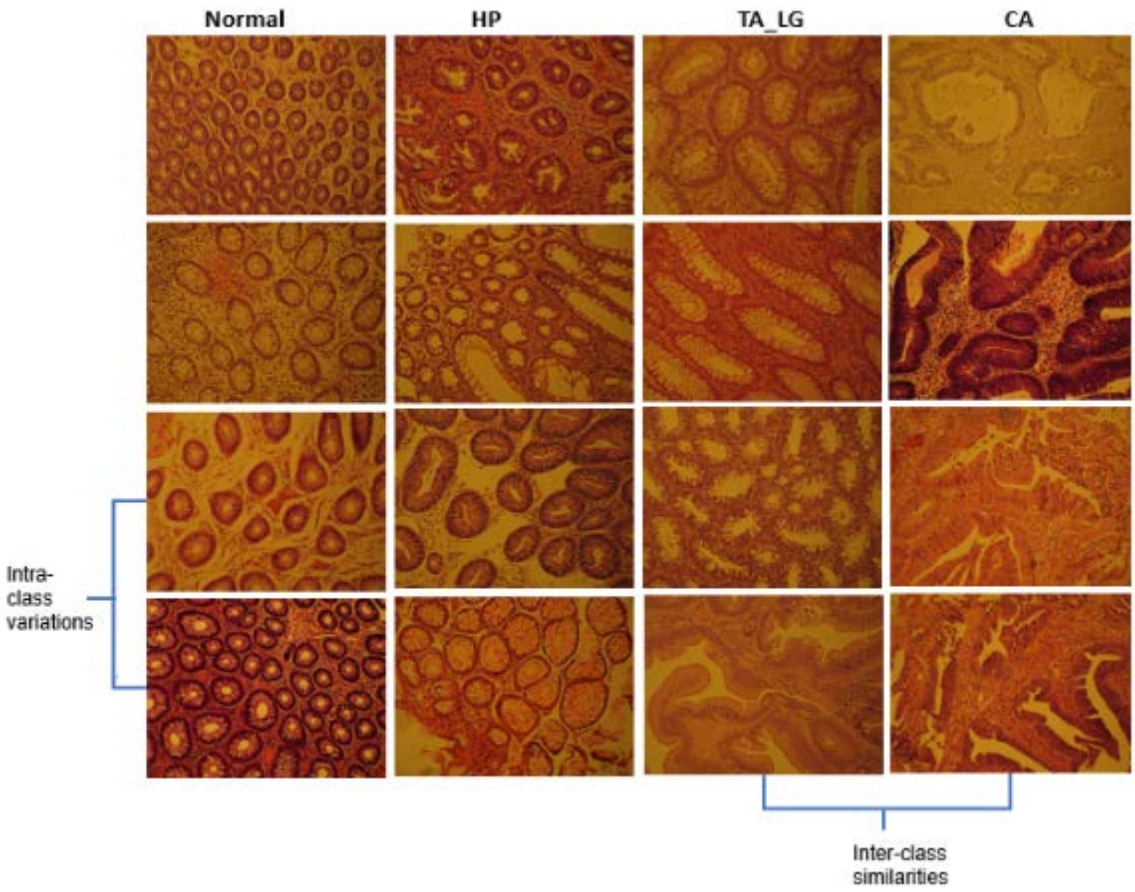


Figure 7. Sample images from the dataset: Normal tissue (column 1), Hyperplastic polyp (column 2), Tubular Adenoma with low grade dysplasia (column 3), Carcinoma (column 4)

3.5 Data Augmentation

The dataset should be sufficiently large in order to have enough samples for the training and testing phases of supervised learning algorithms. We applied specific augmentations to our MSI and RGB datasets with an intention to build a larger database. The augmentation

procedure on the latter dataset was meant to provide a greater number of images than the former since they were being used for the CNN experiments.

3.5.1 MSI Dataset

Each multispectral image was split into 4 equal sized non-overlapping patches. This resulted in the enlarged database comprising 800 images (patches), with 200 images per class. The images of decreased resolution, 128x160 was used for the evaluation of algorithms.

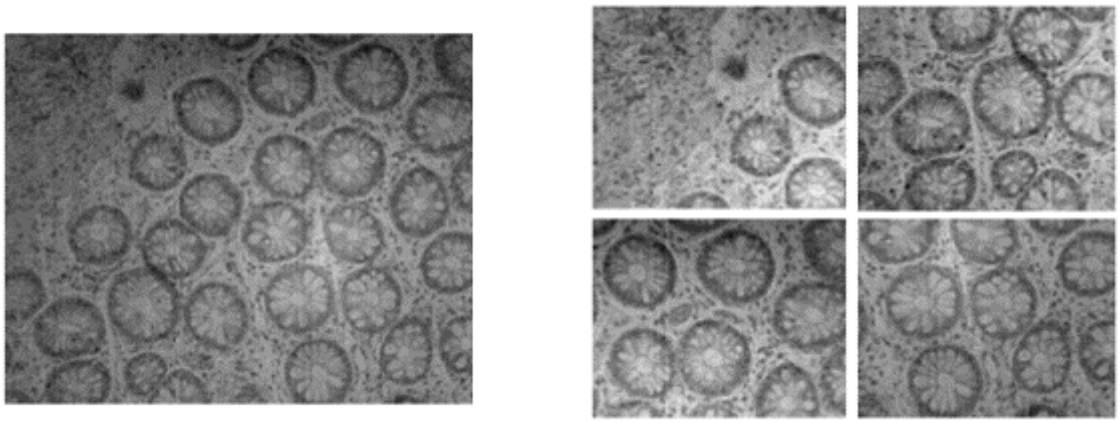


Figure 8. (Left) Original Image (Right) 4 patches generated from the image

3.5.2 RGB Dataset

The 640x480 image was initially divided into four 300x300 sized patches; two from

the top half and two from the bottom half. Each of the 4 patches was subjected to 3 major rotations (90°, 180°, and 270°) and a transpose (Figure 9). This enabled us to build a database of 4000 images (patches) with 1000 patches per class.

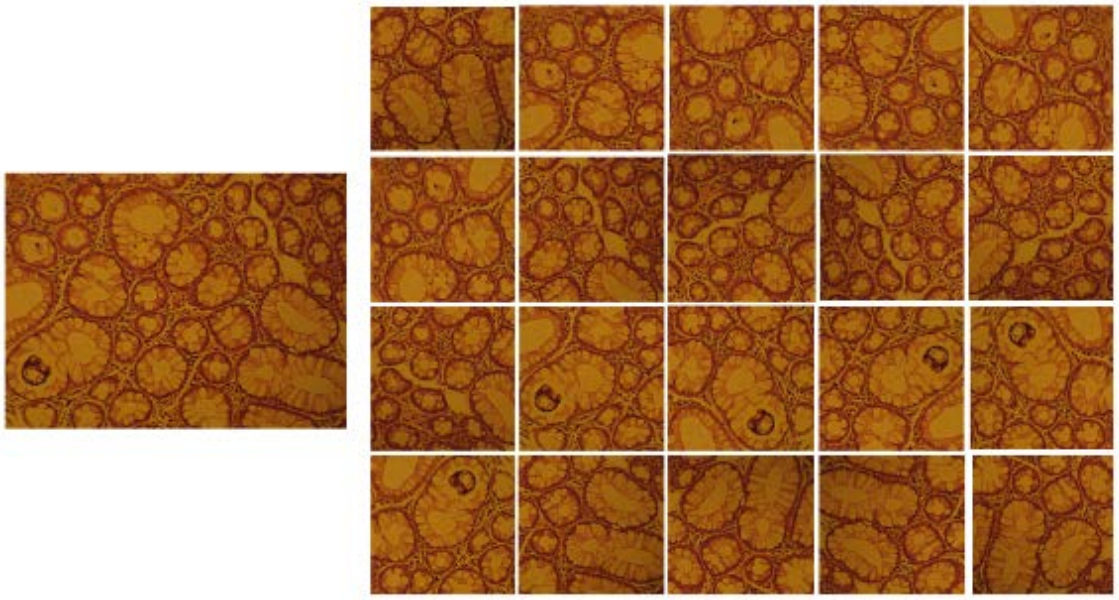


Figure 9. (Left) Original Image (Right) 20 patches generated from the image

CHAPTER 4: CANCER IDENTIFICATION ON MULTISPECTRAL IMAGES

We have provided a review of tumor detection methodologies utilizing multispectral images in [47]. Some prominent researches carried out to accomplish the detection of cancers affecting cervix, breast, colon, lungs, and prostate are discussed in the work. The ultimate aim of hyper/ multispectral imaging of biopsy slides is to achieve an efficient demarcation between the normal and abnormal tissues in automated diagnostic systems. Since a high computational load is accompanied with the multispectral image processing, it is expected to surpass the efficacy of RGB imaging.

The most challenging problem with multispectral data processing is the phenomenon known as, “Curse of Dimensionality”. Acquisition of medical datasets with high spatial & spectral resolution can also be a complex task. Another aspect we have to deal with is the high expenses incurred in the procurement of concerned equipments. Capturing the extremely fine variations both spectrally and spatially paves the development of an ideal automated system for the classification of the various tissue types. Extraction of relevant information from the high dimensional data cubes can be a challenging phase when dealing with multispectral images. The time constraint due to the increased complexity of the multispectral image processing should be overcome by the selection of the appropriate spectral bands from the redundant data.

It can be seen that most of the works in the literature involve experiments on a small number of images. This can be due to the unavailability of the multispectral dataset, which demands complex acquisition. A small dataset is not sufficient to quantify the success rate of specific algorithms which may need extensive training for learning purpose. Multi/Hyper spectral imaging provides imaging in invisible bands such as infrared, thus

offering the capability to reveal information that cannot be seen by the naked eye. But the full spectrum ranging from ultraviolet to near and the mid infrared have not been completely exploited with the current spectral imaging systems in the cancer diagnostic field. All multispectral and hyperspectral imaging methods till date have utilized only the visible part of light spectrum. This may not be much adequate to reveal the complete tissue characteristics, as near-infrared / mid-infrared spectrum remains unutilized.

4.1 Tumor identification system based on multispectral imaging

Our research work [48] presents a multispectral image based colorectal tumor identification system. The algorithm validation is performed on our augmented MSI database comprising 800 images. ie, 200 images each from the 4 classes, viz. normal, hyperplastic polyp, tubular adenoma with low grade dysplasia and carcinoma tissues. All the texture feature based methodologies in our work extract corresponding feature independently for each band and concatenate together to form the feature vector for classification. Rotation invariant Local phase quantization (rLPQ) feature extraction on our multispectral images have yielded a classification accuracy of 86.05% with an SVM classifier. Moreover, the experiments were carried out on another small multispectral image dataset which had 3 categories of tissues. The obtained results were also satisfactory to demonstrate the effectiveness of rLPQ in the context of tumor identification.

Feature Extraction includes LBP, Uniform Rotation invariant LBP, LPQ, and rLPQ. For all type of features extracted from the multispectral images, the classification performance was compared for two supervised learning techniques: SVM and RF. The whole dataset was split into 70% for training and the remaining for validation. These partitions was done randomly from the 800 patches and hence irrespective of the patient,

biopsy sample and image. The classification performances are presented in Table 2. All classification accuracies are based on a 50 fold data shuffling method of holdout validation.

Table 2. Classification accuracies (%) on our multispectral image dataset

Method	SVM	RF
LBP	77.86	62.10
Uniform Rotation invariant LBP	83.61	72.18
LPQ	67.52	60.14
Rotation invariant LPQ	**86.05**	72.04

It is obvious that rLPQ outperforms all other methods studied. Regarding the performance of classifiers, SVM could provide higher classification accuracies when compared to the RF method.

In the same work, we conducted experiments on another multispectral image dataset (Dataset II) [15]. But it was composed of image bands from VIS spectrum alone. The 29 images of the database were split into 16 patches of dimensions 128*128*16, which allowed us to have a larger dataset. These patches were then labeled with the same label as the image from which they were extracted. Therefore, the resultant database on which the tests were conducted consists of 160 CA, 160 BH and 144 IN images of size 128*128*16. The classification results of LPQ, rLPQ as well as Uniform rLBP on the dataset are shown in Table 3.

Table 3. Classification accuracies (%) on Dataset II

Method	SVM	RF
Uniform rotation invariant LBP	**91.49**	88.81
LPQ	80.92	79.91
Rotation invariant LPQ	90.39	84.53

It can be concluded that the rLPQ and uniform rLBP methods yielded comparable accuracies for Dataset II, where it was required to discriminate between three classes. It is worth noticeable that SVM classifier outperformed the RF classifier here also, with an insignificant margin.

4.2 Comparative evaluations between RGB and Multispectral Imaging based approaches

The four textural features; LBP, Uniform rLBP, LPQ, and rLPQ are used in the comparative study to evaluate the classification performances achieved with the two image modalities in the scenario of cancer tissue identification. Since the resolution of images in the MSI and RGB datasets were different, we performed a resizing of the images as in the latter one. Moreover, RGB dataset was augmented on the basis of the method followed for the MSI dataset (section 3.5.1). Eventually, all experiments were carried out on the generated 800 patches of both the databases (Table 4).

Table 4. Comparison of classification accuracies with Multispectral and RGB Images

Method	Multispectral Images	RGB Images
LBP	**77.86**	65.32
Uniform Rotation invariant LBP	**83.61**	66.99
LPQ	**67.52**	65.29
Rotation invariant LPQ	**86.05**	80.71

Table 4 reveals that the multispectral imaging approach is delivering the best results for all the methods, with the greatest accuracy improvement on Uniform rotation invariant LBP. The higher accuracy has to comprise with the increased computational complexity arising from the 39 bands in MSI compared to the 3 in RGB images. The processing time for each algorithm is depicted in Figure 10. The rLPQ technique using multispectral imagery provides the highest accuracy and consumes a greater time in comparison to its RGB counterpart.

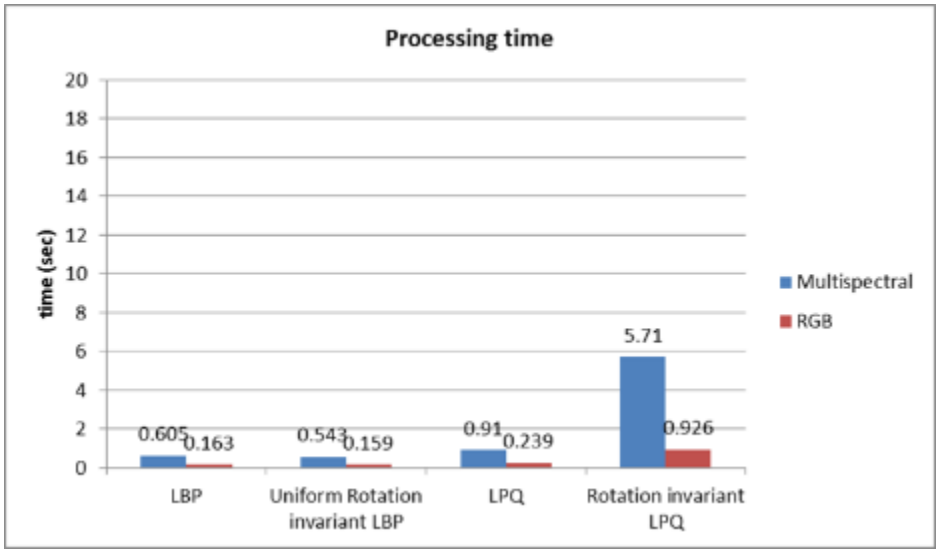


Figure 10. Comparison of computational complexity

4.3 Band Selection in Multispectral Imagery

Selection of the relevant bands from the multispectral image data can be the solution to reduce the computational complexity without significantly suffering from the classification performance. Our idea was based on mutual information. In information theory, the mutual information measures the statistical dependence between two random variables [49]. In other words, it can indicate the amount of information one variable contains regarding another variable. It has been utilized as a tool to realize image registration, where it serves to measure the similarity between two images.

The redundancy in the image bands can be captured with the calculation of mutual information between the neighboring bands. Higher values of MI come with the bands which are highly similar. Initially, mutual information (MI) is computed between the

individual neighboring bands. The bands which have the MI value less than a certain threshold alone are retained. If there are multiple neighboring bands which are similar (MI above the threshold), then the band with the highest information should be identified and selected. It is based on the entropy values of that particular band. The band which has the highest entropy corresponds to the one with maximum information and is retained. When the threshold is varied from high to low, the number of bands that are selected will be reduced. The whole process is represented in Figure 11.

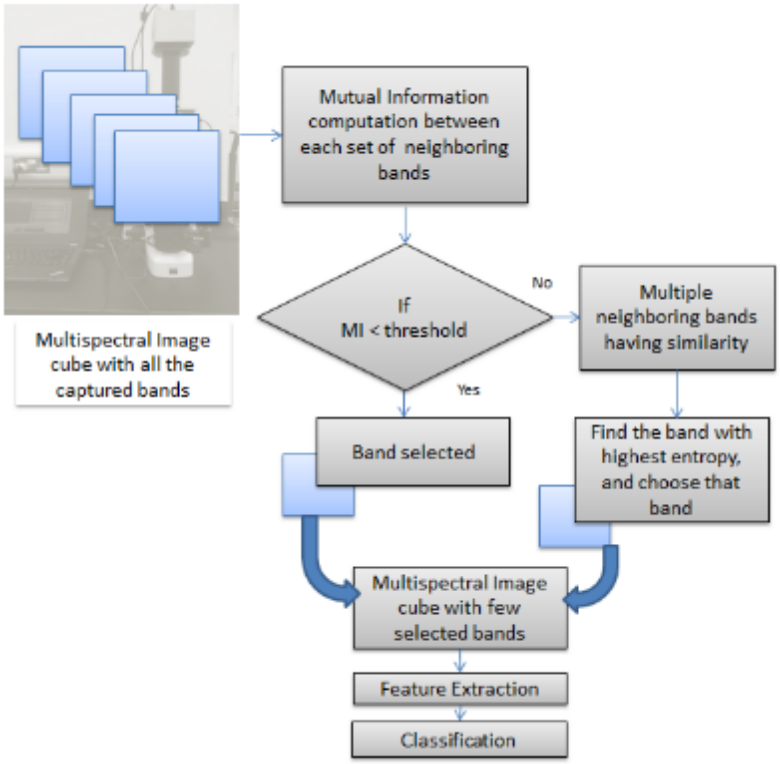


Figure 11. Proposed methodology for band selection

Figure 12 shows a plot of mutual information of the respective bands for several sample images from the database. It is obvious that all curves follow a similar pattern, which indicates the bands that should be retained for a dataset can be generalized from this finding. This process insists on performing the band selection algorithm as an offline procedure and should not be repeated for new images that are collected from the same setup. In contrast to existing methods [50], this approach enables to reduce the computation time with a much simpler algorithm.

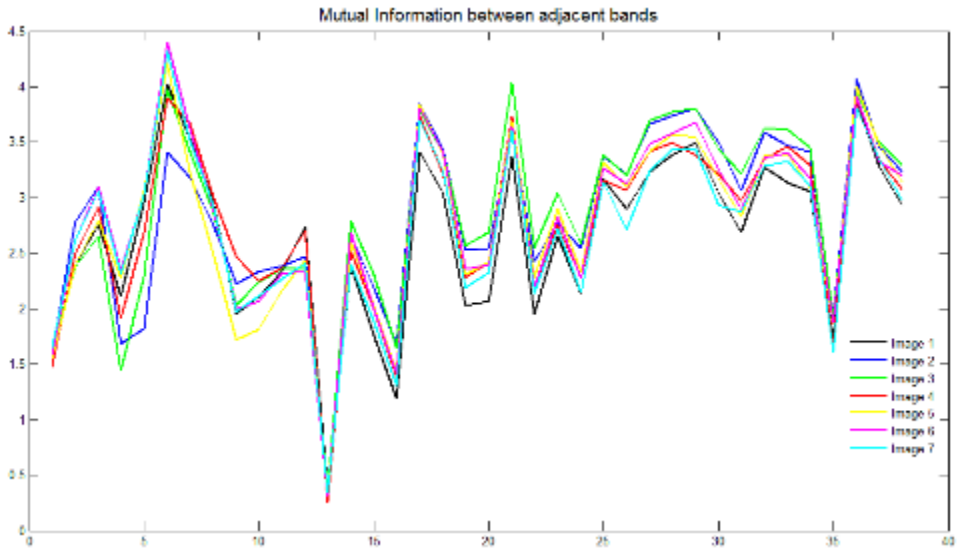


Figure 12. Mutual information among adjacent thirty nine bands: The plot is given here for randomly selected seven images in the database

We applied this band selection algorithm for tumour identification using the Rotation invariant LPQ features. The LPQ is selected here because it is the one with the highest accuracy as presented in Table 4. The local window size used in the precomputed LPQ filters applied with the rotation invariant LPQ descriptors was the default value of 9. Here, we are setting the window size values to 3, 5, 7 and 9 and the corresponding band selection results are investigated. The classification methods that use all the 39 bands are represented in Table 5. In order to decide the acceptable MI, we have chosen 4 different thresholds for which a different number of bands could be selected. Table 6 shows the classification accuracies with different number of bands with the variation in the filter size as well.

Table 5. Classification accuracies without band selection (For filter window size 3,5, 7, and 9)

Filter size	3	5	7	9
Rotation Invariant LPQ	86.11	86.39	**87.70**	86.05

Table 6. Classification accuracies with band selection: An MI threshold of 3.25 leads to the selection of 22 image bands from the 39 dimensional multispectral image cube

No. of bands	22	19	17	10
Filter size	(MI< 3.25)	(MI<3)	(MI<2.75)	(MI<2.5)
3	86.17	85.30	83.62	78.86
5	86.83	87.31	86.82	85.12
7	87.87	**87.91**	87.50	87.15
9	86.52	86.77	86.19	85.37

A quick scan through the above 2 tables conveys that classification accuracies are increasing or remaining largely the same compared to the results obtained with the full set of 39 bands. It is evident that the accuracy has improved from 87.70% to 87.91%, when only 19 bands are used. This finding indicates that many redundant bands could be eliminated with the band selection. Consequently, the processing time could be greatly reduced. Moreover, the feature extraction from mere relevant bands resulted in increased accuracy. Even with the 10 bands, an accuracy of 87.15% is obtained, which is not significantly less than the result obtained without band selection.

In order to evaluate the significance of band selection algorithm, we carried out our band selection approach on another dataset (Dataset II)[15]. The images in this database is augmented in a similar manner as that of our MSI dataset (Section 3.5.1), which generated 464 images.

Table 7 shows the classification accuracies obtained when using the original set of 16 bands. The results are displayed for the different window sizes of 3, 5, 7 and 9 of the LPQ filters. The band selection algorithm produces the classification results indicated in Table 8.

Table 7. Classification accuracies without band selection (For filter window size 3, 5, 7, and 9)

Filter size	3	5	7	9
Rotation Invariant LPQ	**92.21**	91.96	91.68	90.32

Table 8. Classification accuracies with band selection: An MI threshold of 3 leads to the selection of 13 image bands from 16 dimensional multispectral image cube

No. of Bands	13	10	8	5
Filter size	(MI<3)	(MI<2.75)	(MI<2.5)	(MI<2.25)
3	92.57	93.83	93.77	90.11
5	91.44	**94.09**	93.96	92.75
7	91.22	92.24	92.60	92.86
9	90.73	92.39	92.55	91.84

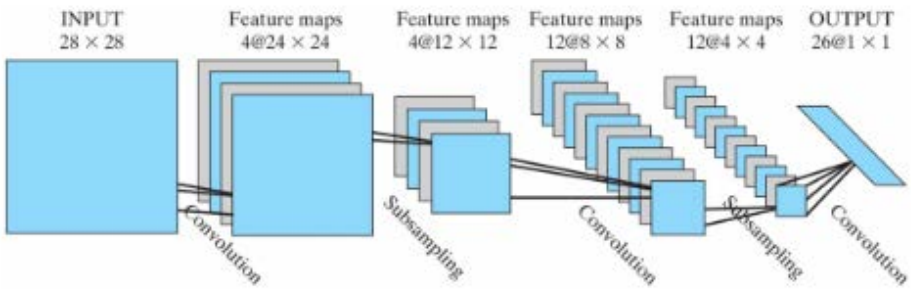
Similar to our MSI dataset, the classification accuracy has improved with the band selection. The previous accuracy of 92.21% with the 16 bands was improved to 94.09% when 6 bands were removed. This finding demonstrates the band selection benefits for better classification and reduced computational complexity. Feature extraction on mere 5 bands from the total 16 could deliver a higher accuracy than those obtained from the full set of bands. This signifies the ability of our algorithm to eliminate the redundant bands by mutual information and identify the prominent bands from the image entropies.

CHAPTER 5: THE PROPOSED APPROACH

5.1 CNN Overview

CNNs are the basic feed-forward artificial neural networks which were initially developed as crude models of mammalian visual cortex. A typical CNN structure can be visualized in Figure 13. The input layer is fed with our original images which will be comprised of specific number of channels according to the modality of images. There are sub-sampling layers between each convolutional layers which is intended to decimate the feature maps of the neurons of previous layers. After the different subsampling layers, a scalar (1-D) neuron should be reached in the last subsampling layer. Finally, CNNs contain fully-connected layers that have the structure similar to Multi-Layer Perceptrons.

We should consider and specify certain parameters prior to the CNN training phase. This includes width and height of the input image, filter dimensions at each level, number of CNN hidden layers, number of MLP hidden layers, number of neurons in each layer, and subsampling factor for each level. Basically, the input image dimension should comply with the number of CNN layers, filter size and subsampling factors so that at the final layer, we can reach to a scalar feature map.



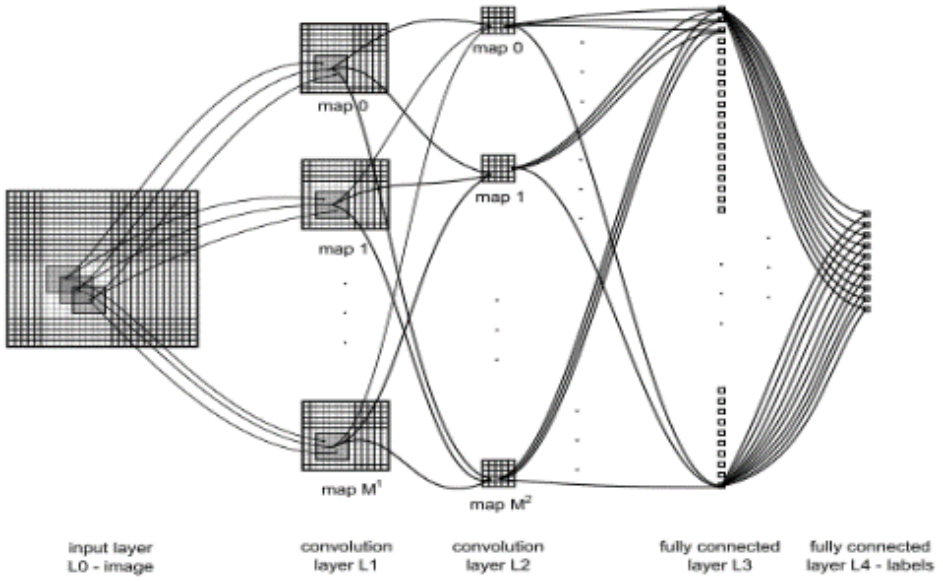


Figure 13. Structure of the conventional CNN (top), Convolutions with only convolutional layers shown (bottom) [51]

5.2 Adaptive 2D CNNs

In the proposed adaptive CNN implementation, there are two types of hidden layers: CNN layers into which conventional “convolutional” and “subsampling” layers are merged, and fully-connected (or MLP) layers. Neurons of the hidden CNN layers are, therefore, modified in such a way that each neuron is capable of both convolution and down-sampling. The intermediate outputs of each neuron are sub-sampled to obtain the final output of that particular neuron. The final output maps are then convolved with their individual kernels and further cumulated to form the input of the next layer neuron. In order to simplify the CNN analogy and to have the freedom of any input layer image dimension

independent of the CNN parameters, neurons of the hidden "CNN layers" are modified as shown in Figure 14. Here, each neuron is capable of convolution and down-sampling. The final output of the k[th] neuron at layer l, s_k^l, is, therefore, the sub-sampled version of the intermediate output, y_k^l. The input map of the next layer neuron will be obtained by the cumulation of the final output maps of the previous layer neurons convolved with their individual kernels, as follows:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv2D\left(w_{ik}^{l-1}, s_i^{l-1}, 'NoZeroPad'\right) \quad (1)$$

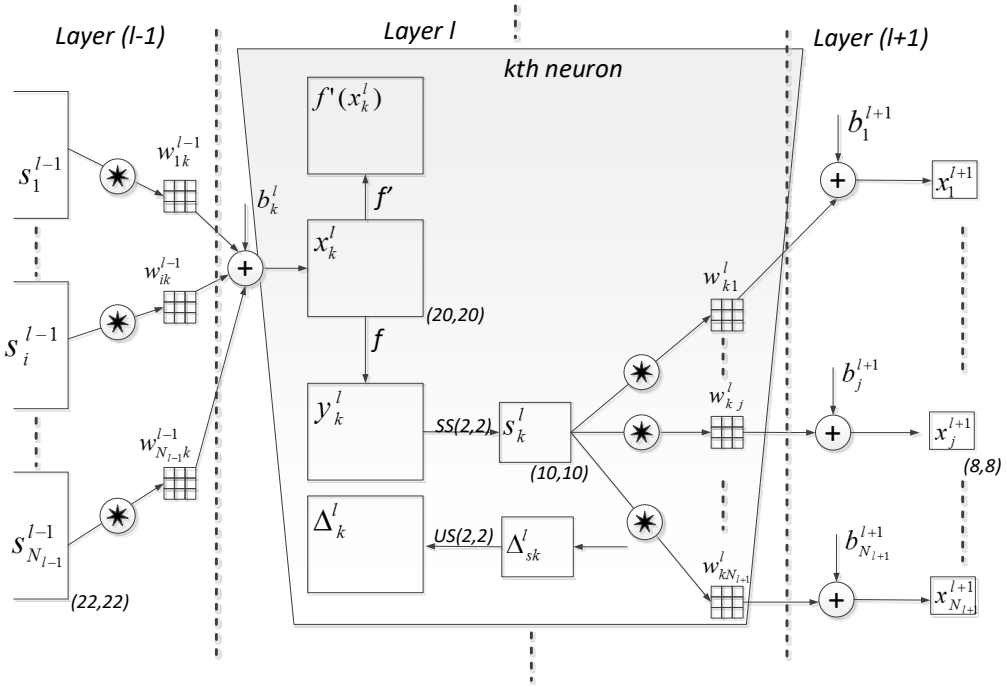


Figure 14. Three consecutive CNN layers in the adaptive CNN implementation

The number of hidden CNN layers can be set to any number. This ability is possible in this implementation because the sub- sampling factor of the output CNN layer (the hidden CNN layer just before the first MLP layer) is set to the dimensions of its input map, e.g., if the layer l+1 would be the output CNN layer, then the sub-sampling factors for that layer is automatically set to *ssx* = *ssy* = 8 since the input map dimension is 8x8. Besides the sub-sampling, note that the dimension of the input maps are gradually decreasing due to the convolution without zero padding. As a result of this, the dimension of the input maps of the current layer is reduced by (K_x-1, K_y-1) where Kx and Ky are the width and height of the kernel, respectively. The input layer is fed with the down-sampled patches each of which has three color channels (R, G, B). We performed down-sampling to 64x64 pixels in order to mimic the low-resolution data.

For an N-class problem, for each patch with its corresponding target and output vectors, $[t_1, \dots, t_N]$ and $[y_l^L, \dots, y_N^L]$ respectively we are interested to find out the derivative of this error with respect to each individual network parameter (weights and biases). Let *l*=1 and *l*=*L* be the input and output layers, respectively. The error (MSE) in the output (MLP) layer can be expressed as:

$$E = E(y_1^L, \dots\dots, y_{N_L}^L) = \sum_{i=1}^{N_L} (y_i^L - t_i)^2 \qquad (2)$$

The Back Propagation (BP) training of the MLP layers is identical to the conventional BP for MLPs and hence skipped here. The BP training of the CNN layers is composed of 4 distinct operations as detailed below.

5.2.1 Inter BP among CNN layers: $\Delta s_k^l \overset{\Sigma}{\leftarrow} \Delta_l^{l+1}$

The basic rule of BP states: If the output of the k[th] neuron at layer *l* contribute a neuron *i* with weight w_{ki}^l in the next layer *l*+1, the next layer neuron's delta Δ_l^{l+1} will contribute with the same weight to form $\Delta_k^l$of the neuron in the previous layer *l*. This means:

$$\frac{\partial E}{\partial s_k^l} = \Delta s_k^l \overset{\Sigma}{\leftarrow} \Delta_l^{l+1}, \forall i \in \{1, N_{l+1}\} \tag{3}$$

Where *E* is the total error (MSE)

Specifically:

$$s_k^l = \sum_{i=1}^{N_{l+1}} \frac{\partial E}{\partial x_i^{l+1}} \frac{\partial x_i^{l+1}}{\partial s_k^l} = \sum_{i=1}^{N_{l+1}} \Delta_l^{l+1} \frac{\partial x_i^{l+1}}{\partial s_k^l} \tag{4}$$

Where;

$$x_i^{l+1} = \cdots + s_k^l * w_{ki}^l + \cdots \tag{5}$$

It is obviously hard to compute the derivative directly from the convolution. Instead, let us focus on a single pixel's contribution of the output, $s_k^l(m,n)$, to the pixels of the $x_i^{l+1}(m,n)$ with the assumption of a 3x3 kernel.

$$\begin{aligned} x_i^{l+1}(m-1,n-1) &= \cdots + s_k^l(m,n).w_{ki}^l(2,2) + \cdots \\ \boldsymbol{x_i^{l+1}(m-1,n)} &= \boldsymbol{\cdots + s_k^l(m,n).w_{ki}^l(2,1) + \cdots} \end{aligned} \tag{6}$$

$$x_i^{l+1}(m+1, n+1) = \cdots + s_k^l(m,n).w_{ki}^l(0,0) + \cdots$$

This is illustrated in Figure 15 where the role of an output pixel, $s_k^l(m,n)$, over two pixels of the next layer's input neuron's pixels, $x_i^{l+1}(m-1, n-1)$ and $x_i^{l+1}(m+1, n+1)$ is indicated.

Considering the pixel as an MLP neuron that is connected to other MLP neurons in the next layer, according to the basic rule of BP one can then easily write the delta of $s_k^l(m,n)$ as follows:

$$\frac{\partial E}{\partial s_k^l}(m,n) = \Delta s_k^l(m,n)$$
$$= \sum_{i=1}^{N_{l+1}} \left(\sum_{r=-1}^{1} \sum_{t=-1}^{1} \Delta_l^{l+1}(m+r, n+t).w_{ki}^l(1-r, 1-t) \right) \quad (7)$$

If we generalize it for all pixels of Δs_k^l,

$$\Delta s_k^l = \sum_{i=1}^{N_{l+1}} conv2D(\Delta_l^{l+1}, rot180(w_{ki}^l), 'ZeroPad') \quad (8)$$

Note that this is a full convolution with zero padding by (K_x-1, K_y-1) zeros to each boundary of the Δ_l^{l+1}in order to achieve an equal dimensions (width and height) for Δs_k^l and Δ_l^{l+1} with the s_k^l.

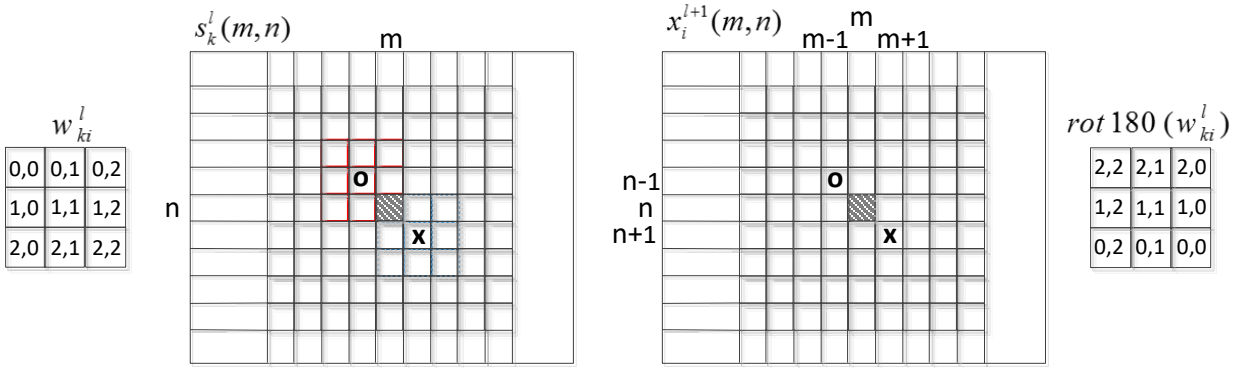


Figure 15. A single pixel's contribution of the output, $\boldsymbol{s_k^l(m,n)}$, to the two pixels of the $\boldsymbol{x_i^{l+1}}$ using a 3x3 kernel.

5.2.2 Intra BP within a CNN neuron: $\Delta_k^l \leftarrow \Delta s_k^l$

Once the first BP is performed from the next layer, *l*+1, to the current layer, *l*, then we can further back-propagate it to the input delta. Let zero order up-sampled map be: $us_k^l = up_{ssx,ssy}(s_k^l)$, then one can write:

$$\Delta_k^l = \frac{\partial E}{\partial x_k^l} = \frac{\partial E}{\partial y_k^l}\frac{\partial y_k^l}{\partial x_k^l} = \frac{\partial E}{\partial us_k^l}\frac{\partial us_k^l}{\partial y_k^l} f'(x_k^l) = up(\Delta s_k^l)\beta\, f'(x_k^l) \quad (9)$$

Where $\beta = (ssx, ssy)^{-1}$ since each pixel of s_k^l was obtained by averaging ssx.ssy number of pixels of the intermediate output, y_k^l. Instead of averaging if maximum pooling is used, then Eq. **(9)** should be adapted accordingly.

5.2.3 BP from first MLP layer to output CNN layer

As illustrated in Figure 16, the output layer of CNN can be connected to the 1st MLP layer and hence the outputs of this layer CNN neurons, s_k^l, are scalars. In other words, s_k^l and of course, Δs_k^l are now all scalars and to achieve this recall that the subsampling factors, ssx and ssy, in this particular layer are all set to the dimensions of the input map (ssx= 8, ssy= 8) as in the figure.

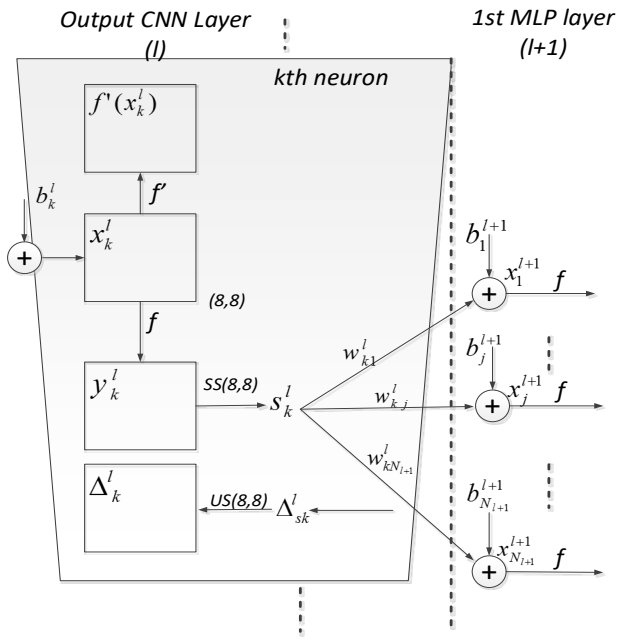


Figure 16. Output CNN layer- 1st MLP layer

Similarly the weights of this CNN layer neurons, w_{ki}^l, are also all scalar and instead of convolution, scalar multiplication is performed as in a regular MLP. So from MLP layer to

the CNN layer, the regular (scalar) BP is simply performed as in Eq. (10)

$$\frac{\partial E}{\partial s_k^l} = \Delta s_k^l = \sum_{i=1}^{N_{l+1}} \frac{\partial E}{\partial x_i^{l+1}} \frac{\partial x_i^{l+1}}{\partial s_k^l} = \sum_{i=1}^{N_{l+1}} \Delta_i^{l+1} w_{ki}^l \quad (10)$$

And intra BP to get: $\Delta_k^l \overset{BP}{\leftarrow} \Delta s_k^l$ is identical as in Eq. (9)

$$\Delta_k^l = \frac{\partial E}{\partial x_k^l} = up\left(\Delta s_k^l\right) \beta\, f'(x_k^l) \quad (11)$$

Finally, the weight and bias sensitivities, too, are identical to a regular MLP

$$\frac{\partial E}{\partial w_{kj}^l} = \frac{\partial E}{\partial x_j^{l+1}} \frac{\partial x_j^{l+1}}{\partial w_{kj}^l} = \Delta_j^{l+1} s_k^l \quad (12)$$

and $$\boldsymbol{\frac{\partial E}{\partial b_k^{l+1}} = \frac{\partial E}{\partial x_k^{l+1}} \frac{\partial x_k^{l+1}}{\partial b_k^{l+1}} = \Delta_k^{l+1}}$$

5.2.4 Computation of weight and bias sensitivities

As in the regular BP on MLPs, the delta of the i[th] neuron at layer l+1, Δ_i^{l+1} will be used to update the bias of that neuron and all weights of the neurons in the previous layer connected to that neuron.

$$x_i^{l+1} = b_i^{l+1} + \cdots + y_k^l w_{ki}^l + \cdots \rightarrow \frac{\partial E}{\partial w_{ki}^l} = y_k^l \Delta_i^{l+1} \quad (13)$$

and $$\boldsymbol{\frac{\partial E}{\partial b_l^{l+1}} = \Delta_i^{l+1}}$$

The update rule of conventional BP states: The sensitivity of the weight connecting the k[th] neuron in the current layer to the i[th] neuron in the next layer depends on the output of the current layer neuron, and the delta of the next layer neuron. For CNN layer neurons

we need to follow a similar approach to find out weight and bias sensitivities. Figure 17 illustrates the convolution of the output of the current layer neuron, s_k^l, and kernel, w_{ki}^l, to form the input of the i[th] neuron, x_i^{l+1}, at the next layer l+1.

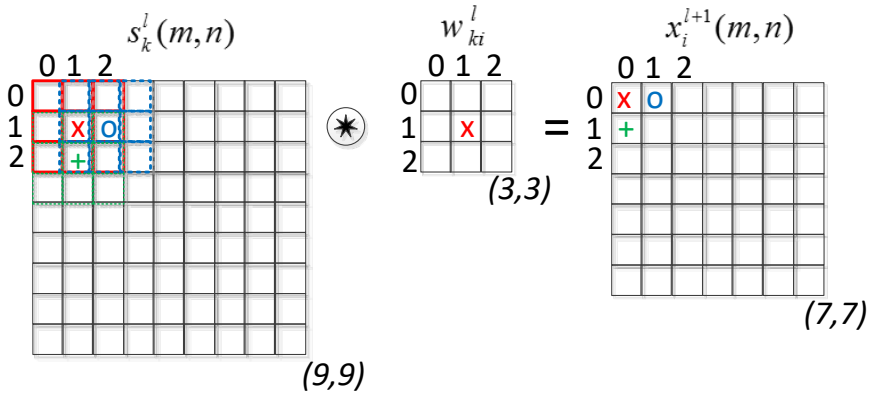


Figure 17. Convolution of the output of the current layer neuron, $\boldsymbol{s_k^l}$, and kernel, $\boldsymbol{w_{ki}^l}$, to form the input of the [ith] neuron, $\boldsymbol{x_i^{l+1}}$, at the next layer l+1

So if we focus on the contribution of each kernel element over the output, in analytical form one can write:

$$x_i^{l+1}(0,0) = .. + w_{ki}^l(0,0)s_k^l(0,0) + w_{ki}^l(0,1)s_k^l(0,1) + w_{ki}^l(1,0)s_k^l(1,0) + ..$$

$$x_i^{l+1}(0,1) = .. + w_{ki}^l(0,0)s_k^l(0,1) + w_{ki}^l(0,1)s_k^l(0,2) + w_{ki}^l(1,0)s_k^l(1,1) + .. \quad (14)$$

$$\boldsymbol{x_i^{l+1}(1,0) = .. + w_{ki}^l(0,0)s_k^l(1,0) + w_{ki}^l(0,1)s_k^l(1,1) + w_{ki}^l(1,0)s_k^l(2,0) + ..}$$

….

$$x_i^{l+1}(m,n) = .. + w_{ki}^l(0,0)s_k^l(m,n) + w_{ki}^l(0,1)s_k^l(0,n+1)$$
$$+ w_{ki}^l(1,0)s_k^l(m+1,n) + ..$$
$$x_i^{l+1}(m,n) = \sum_{r=-1}^{1}\sum_{t=-1}^{1} w_{ki}^l(r+1,t+1)s_k^l(m+r,n+t) + ..$$

Since each weight (kernel) element is used in common to form each neuron input, $x_i^{l+1}(m,n)$, the derivative will be the cumulation of delta- output product for all pixels. ie,

$$\frac{\partial E}{\partial w_{ki}^l(r,t)} = \sum_m \sum_n \Delta_i^{l+1}(m,n)s_k^l(m+r,n+t) \quad (15)$$

$$\Longrightarrow \frac{\partial E}{\partial w_{ki}^l} = conv2D(s_k^l, \Delta_i^{l+1}, 'NoZeroPad')$$

Similarly, the bias for this neuron, b_k^l , contributes to all pixels in the image (same bias shared among all pixels), so its sensitivity will be the cumulation of individual pixel sensitivities as expressed in Eq. (16)

$$\frac{\partial E}{\partial b_k^l} = \sum_m \sum_n \frac{\partial E}{\partial x_k^l(m,n)} \frac{\partial x_k^l(m,n)}{\partial b_k^l} = \sum_m \sum_n \Delta_k^l(m,n) \quad (16)$$

As a result, the iterative flow of the BP for each patch in the training set can be stated as follows:

1) Initialize weights (kernels) and biases (e.g., randomly, U(-0.1, 0.1)) of the CNN.
2) For each BP iteration (*t=1:iterNo*) DO:
    a. For each patch, *p*, in the train set, DO:

i. FP: Forward propagate from the input layer to the output layer to find the output of each neuron at each layer, $y_i^l, \forall i \in [1, N_l]\ and\ \forall l \in [1,L]$

ii. BP: Compute delta error at the output (MLP) layer and back-propagate it to first hidden CNN layer to compute the delta errors, $\Delta_k^l, \forall k \in [1, N_l] and\ \forall l \in [2, L-1]$

iii. PP: Post-process to compute the weight and bias sensitivities using Eqs. (15) and (16)

iv. Update: Update the weights and biases with the (cumulation of) sensitivities found in (c) scaled with the learning factor, ε:

$$w_{ik}^{l-1}(t+1) = w_{ik}^{l-1}(t) - \varepsilon \frac{\partial E}{\partial w_{ik}^{l-1}} \qquad (17)$$

$$b_k^l(t+1) = b_k^l(t) - \varepsilon \frac{\partial E}{\partial b_k^l}$$

5.3 Methodology

We have proposed an adaptive CNN algorithm to implement the cancer detection and identification system. As illustrated in Figure 18, the proposed systematic approach for cancer detection and identification has four main blocks: Image acquisition from biopsy samples, patch generation for data augmentation, classification by an adaptive CNN and final decision by majority ruling. Biopsy samples are digitized using a Zeiss microscope with the 10x objective lens. Over each sample, an image is acquired with 640x480 pixel

resolution and then augmented to deliver 20 300x300 pixel patches as explained in the section 3.5.2. These patches are further resampled to a size of 64x64 to mimic low-resolution data which is then used as the input to the proposed adaptive 2D CNN that was trained by using Back-Propagation (BP) in advance. Once the patches of each image are classified, the final decision is composed by the majority rule as follows: the values that represent the probabilities for belonging to each class obtained from the classifier are averaged over the patches and the final decision for that specific image is taken based on the highest probability obtained.

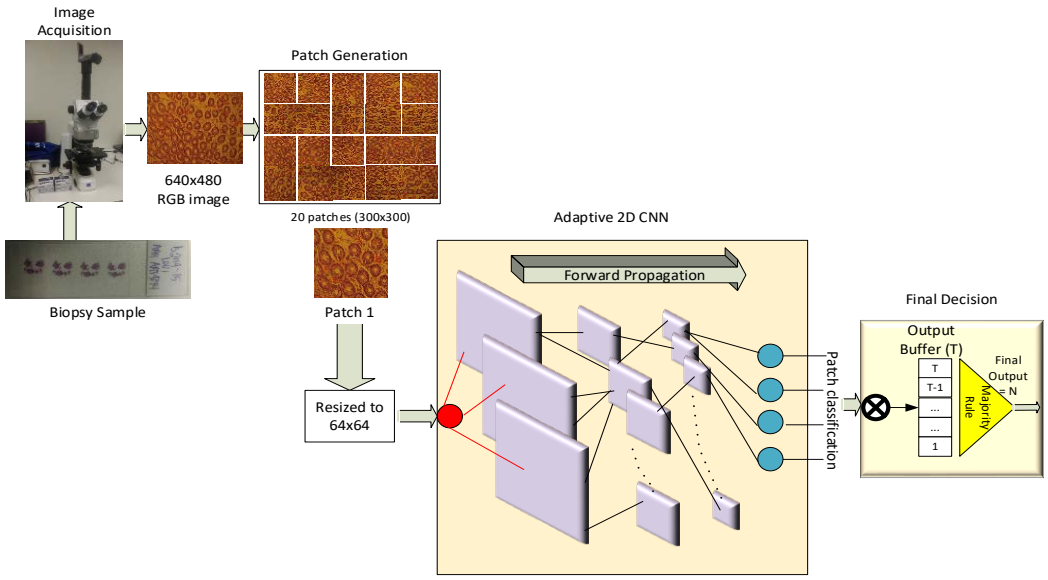


Figure 18. Proposed system architecture for cancer detection and identification

5.4 Investigated Conventional Methods

Conventional classification methods consist of two consecutive processes: feature extraction and classification. For the former, 5 state of art texture-based feature extraction algorithms are evaluated. For the latter, we used the SVM classifier with 3 different kernels. Both of them are detailed in the following sections.

5.4.1 Feature Extraction Methods

(1) Rotation Invariant Local Phase Quantization (rLPQ)

Local Phase Quantization is a feature descriptor proposed recently by Ville Ojansivu [52]. It is based on the quantized phase of the discrete Fourier transform (DFT) computed over local image windows. Local computation of DFT phase at every pixel location followed by concatenation of the resulting codes as a histogram generates the LPQ feature vector. Several applications including recognition of blurred faces, fingerprint liveness detection etc. have exploited the LPQ features successfully. The codes produced by the LPQ operator are usually insensitive to centrally symmetric blur. The invariance to uniform illumination changes arising from the utility of mere phase information is a distinctive characteristic of the technique. The LPQ feature extraction procedure generates a 256-dimensional feature vector.

A blur and rotation insensitive local phase quantization texture descriptor was later proposed by the same authors in [53]. It consists of mainly two stages:- (1) estimation of local characteristic orientation (2) extraction of directed binary descriptor vector. Both stages apply the phase of the locally computed Fourier transform coefficients. These coefficients are insensitive to centrally symmetric image blurring. The characteristic orientation is estimated from the quantized STFT (Short time Fourier transform)

coefficients. The second stage of descriptor extraction follows the same principle as that of ordinary LPQ feature extraction. But the neighborhood at each location will be rotated to the direction of the estimated characteristic orientation.

(2) Rotation Invariant Local Binary Pattern (rLBP)

Local Binary Pattern (LBP) [54], proposed by Ojala et al. is a texture descriptor which is a special case of texture spectrum model. The key idea behind LBP is that it is capable of encoding the relationship between a specific pixel and its neighboring pixels. Accordingly, a code word is generated for each pixel. Each bits '1' or '0' in the code word represents whether the pixel's value is greater than or less than the neighboring pixel. When 8 neighborhood points are considered, the resulting code word will be of length 8. Decimal conversion of these code words followed by a histogram generation finally yields the feature vector.

There are several variants for LBP including rotation invariant LBP, uniform rotation invariant LBP [55] etc. The $LBP_{P,R}$ operator produces 2^P distinct output values, corresponding to the different binary patterns that can be formed by the total P pixels in the neighborhood. When the image is rotated, the gray values will correspondingly move along the perimeter of the circle around a certain pixel. Rotating a particular binary pattern obviously result in a different LBP value. This does not happen with patterns consisting of mere 0s (or 1s). Those will remain constant at all rotation angles. To remove the effect of rotation, a unique identifier is assigned to each rotation invariant local binary pattern. 36 unique rotation invariant local binary patterns can exist in the case of P= 8. A rotation invariant feature descriptor can hence reduce the feature size from 256 to 36 along with achieving rotation invariance.

(3) Uniform Rotation Invariant Local Binary Pattern (Uniform rLBP)

Uniform rLBP tends to reduce the original LBP feature size by such a representation that involves same bin for several similar patterns. The concept of uniform patterns arises from the fact that some binary patterns occur more frequently in texture images than others. A specific LBP is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. So in uniform LBP, all the uniform patterns from the whole 256 patterns are assigned separate labels and all the non-uniform patterns are accumulated to a single label. Since a circularly symmetric neighbor set of P pixels comprise only P+1 uniform binary patterns, the ultimate feature vector has a small length of 10 in the case of an 8 neighborhood.

(4) Haralick

Haralick features are calculated from the GLCM matrices extracted from the images [56]. These matrices represent an average spatial relationship which the gray levels in the image have between one another. We can deduce a measure of the frequency with which an intensity value 'i' occurs at a predefined spatial relationship with another value 'j'. These spatial relationships can be indicated by the 4 different orientations such as 0, 45, 90 and 135 degrees. Finally, the following Haralick features are computed from the normalized GLCM matrices p(i,j) for each of the orientations. The four features obtained for each orientation can be concatenated together to yield the Haralick descriptor.

Energy: $\sum_{i,j} p(i,j)^2$

Contrast: $\sum_{i,j} |\boldsymbol{i-j}|^2 \boldsymbol{p(i,j)}$ (18)

Homogenity: $\sum_{i,j} \frac{p(i,j)}{1+|i-j|}$

Correlation: $\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$

We calculated the above four features: Energy, Contrast, homogeneity, and correlation for each of the 4 different orientations resulting in a 16 dimensional feature vector. 'Energy' is basically the sum of squared elements in the GLCM matrix. The attribute 'Contrast' will measure the intensity contrast between each pixel and its neighbor. Homogeneity accounts for the closeness of distribution of elements in the GLCM to diagonal elements of GLCM. 'Correlation' indicates the level of correlation between each pixel and its neighbor for the whole image.

(5) Concatenated feature: rLPQ+ rLBP

This represents the concatenated feature vector generated by combining the feature descriptors from both rLPQ and rLBP methods. We have chosen these two because of the better results produced by both techniques compared to the other texture features studied. The resultant feature vector had a length of 292.

5.4.2 Classification

Support vector machine (SVM) was introduced in 1963 by Vapnik. It is a discriminative classifier which is defined by a separating hyperplane. Given a labeled training data input, the algorithm outputs an optimal hyperplane which categorizes new data, test dataset [57]. It would not be a difficult task in the case of linearly separable data. But in the case of patterns that are not linearly separable, the original data should be mapped to a new space by means of a kernel function [58]. The main objective of SVM is

to maximize the margin around the separating hyperplane. This decision function is characterized by a small subset of the training samples, called as support vectors that lie closest to the hyperplane. The problem of finding the optimal hyper plane is basically an optimization problem and can be solved by Lagrange multipliers.

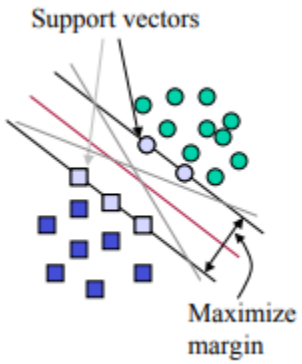


Figure 19. SVM classification [58]

Suppose we have a set of (input, output) training pair samples with input features (x_1, $x_2 \ldots x_n$), and the output result y. The training process generates a set of weights whose linear combination predicts the value of y, which in turn decide the output class. The equation defining the decision surface separating the classes is a hyperplane of the form: $f(x)= w^T x + b = 0$; w: weight vector, x: input vector; b: bias. When an unknown point measured on features x_i is given, the classification can be done based on the sign obtained

for the expression f(x).

Non-linear SVMs perform a mapping of the data to a higher dimensional space in order to achieve separability. Kernel functions are exploited here to achieve this transformation. They are based on calculating the inner products of two vectors. The three types of kernel we have used in our experiments are defined as follows.

Linear kernel: $k(x_i, x_j) = x_i^T x_j$

RBF kernel: $\boldsymbol{k(x_i, x_j) = e^{-(\frac{||x_i - x_j||^2}{2\sigma^2})}}$ (19)

Polynomial kernel: $k(x_i, x_j) = (x_i x_j + 1)^b$

CHAPTER 6: EXPERIMENTAL RESULTS

The experimental results of the proposed approach, as well as the conventional methods, are presented in the subsequent sections. The system implementation details, as well as the evaluation matrices used for the comparative study, are also explained briefly.

6.1 Experimental Setup

Texture feature extraction algorithms were tested and evaluated in MATLAB. Libsvm toolbox [59] of MATLAB was utilized for SVM training and testing. The parameter estimation (cost and gamma) was done using grid search method. An 8-neighborhood schema was used for the LBP implementations. For the LPQ, a local window size of 3 was used with frequency estimation based on STFT with Gaussian window. A precomputed filter of window size 7 was applied to the rotation invariant LPQ descriptor. The implementation of the adaptive CNN is performed using C++ over MS Visual Studio 2015 in 64bit. This is a non-GPU implementation; however, Intel ® OpenMP API is used to obtain multiprocessing with a shared memory. We used a computer with I7-4700MQ at 2.4GHz (8 CPUs) and 16Gb memory for the training and testing purpose. The CNN configuration parameters are presented in Table 9.

Table 9. CNN Parameters

Parameter Name	Value
Convolution filter size	5
Subsampling factor	2
No. of CNN Layers	3
No. of MLP Layers	2

For all experiments, we employed an early stopping training procedure with setting the maximum number of BP iterations to 50 and the minimum training error to 8% to prevent over-fitting. We initially set the learning factor, $\varepsilon = 10^{-3}$ and then a global adaptation has been applied during each BP iteration in order to dynamically set a proper learning factor for the iteration, i.e., if the train MSE decreases in the current iteration we slightly increase ε by 5%; otherwise, we reduce it by 30%, for the next iteration.

6.2 Evaluation Metrics

We performed 5-fold cross validation over the dataset with the train rate set to 80%. Therefore, for each fold, 160 images and corresponding 3200 patches were used for training and the rest was used for the test. This allows us to test the proposed system over all the records. The patches are divided in such a way that patches from same image will not come simultaneously in training and testing. Once the final decision is performed by the majority rule over the patch classification results, the confusion matrices (CMs) of the classification results of the test images are then cumulated to compute the final CM. We

used the classification accuracy (Acc) computed over the final CM to evaluate the overall cancer identification performance. For the cancer detection performance, the final CM that consists of classification results of 4 categories was first deducted to a binary CM that has the numbers for normal and cancer categories. In other words, the classification results of the three cancer categories are merged into a single "cancer" class. Over the binary CM, we computed several standard metrics, i.e., cancer detection accuracy (Acc), sensitivity (Sen), specificity (Spe), and positive predictivity (Ppr). The definitions of these standard performance metrics using the hit/miss counters obtained from the binary CM elements such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN), are as follows: Accuracy is the ratio of the number of correctly classified images to the total number of images, Acc = (TP+TN)/(TP+TN+FP+FN); Sensitivity (Recall) is the rate of correctly classified cancer images among all cancer images, Sen = TP/(TP+FN); Specificity is the rate of correctly classified normal images among all normal images, Spe = TN/(TN+FP); and Positive Predictivity (Precision) is the rate of correctly classified cancer images in all images classified as cancer, Ppr = TP/(TP+FP).

6.3 Results and Comparative Evaluations

Cancer detection is a binary classification where the images are categorized either cancerous or normal. Cancer identification extends this to a multi-class classification where the cancer types are further categorized. The classification accuracies obtained by the conventional methods are given in Table 10. In the table, accuracies with respect to the three experimented SVM kernels are presented for the two tasks.

Among the investigated texture features, rLPQ exhibited the highest accuracy for both cancer identification and detection tasks. Except for the combined feature, all the other

methods exhibited comparatively lower classification accuracies. But when discrimination to be carried out was between just two classes, both rLBP and Uniform rLBP could compete with rLPQ and the combined feature to yield comparable results. This demonstrates the efficiency of rLPQ in the scenarios where there are multiple numbers of classes. The Linear kernel yielded the best classification performance using rLPQ while the Polynomial kernel has the best performance when using the other texture features. No significant performance improvement was observed when features are concatenated. The classification performance of the Haralick features was the poorest.

Table 10. Classification Accuracy (%) for Conventional methods with Multi- class and Binary classification

Method	Cancer Identification			Cancer Detection		
	RBF	**Linear**	**Poly.**	**RBF**	**Linear**	**Poly.**
rLPQ	67.5	**74**	71	84.5	**87**	83.5
rLBP	58	57.5	**65.5**	81	82.5	**86**
Uniform rLBP	56.5	54	**59.5**	80	79.5	**83**
Haralick	38	40	**42.5**	71	65	**73**
rLPQ+ rLBP	64	73.5	**74**	84	86	**87**

Since the combined rLPQ and rLBP features achieved the highest accuracy, it is used for the comparative evaluations of the proposed approach. In order to have a more detailed evaluation, the 4-class confusion matrices (CMs) obtained for cancer identification are presented in Table 11 and Table 12 respectively. Classes 1 - 4 represent the normal, hyperplastic polyp, tubular adenoma with low-grade dysplasia and carcinoma, respectively.

Table 11. Confusion matrix of the Proposed approach

		Truth			
		1	**2**	**3**	**4**
Real	**1**	41	1	0	1
	2	5	29	6	3
	3	1	5	31	0
	4	3	15	13	46

Table 12. Confusion matrix of the Top performing conventional method

		Truth			
		1	**2**	**3**	**4**
Real	**1**	33	8	1	0
	2	14	32	3	6
	3	3	6	43	4
	4	0	4	3	40

According to the results, it is clear that the proposed systematic approach with an adaptive and compact CNN and the top performing conventional method with the best texture feature achieved the highest cancer identification performances within a close margin. This is despite the fact that the proposed method used low-resolution image data (64x64 pixels) while the competing methods used the original patch resolution (300x300). This demonstrates the high robustness of the compact CNNs against low image resolution. Another observation worth mentioning is that the false alarm rate (9/50 = 18%) of the proposed approach is minimal while the best performing conventional method almost doubles it (17/50 = 34%).

Using the CMs of cancer identification presented in Table 11 and Table 12, the corresponding binary CMs for cancer detection can be computed as presented in Table 13 and Table 14. From the binary CM the standard evaluation metrics, cancer detection accuracy (*Acc*), sensitivity (*Sen*), specificity (*Spe*), and positive predictivity (*Ppr*) are computed and presented in Table 15.

Table 13. Confusion matrix for cancer detection using the proposed approach

		Truth	
		1	**2**
Real	**1**	41	2
	2	9	148

Table 14. Confusion matrix for cancer detection using the Top performing conventional method

		Truth	
		1	**2**
Real	**1**	33	9
	2	17	141

Table 15. Performance Comparison of Proposed method and Conventional methods for Cancer Detection

Method	Sensitivity	Precision	Specificity	Accuracy
Adaptive CNN	**94.27**	**98.67**	**95.35**	**94.50**
rLPQ	89.24	94.00	78.57	87.00
rLBP	90.67	90.67	72.00	86.00
Uniform rLBP	87.66	90.00	67.39	83.00
Haralick	80.38	84.67	45.24	73.00
rLPQ+ rLBP	89.74	93.33	77.27	87.00

The cancer detection results clearly indicate that the proposed systematic approach with the compact and adaptive CNNs at its core achieved a superior detection performance with an elegant Precision and Recall (Sensitivity) levels around 95%. This indicates such a level that is more than 4% higher than the top-performing method. However, the most distinctive performance gap occurred on the Specificity level (>95%) which is around 17% higher. This shows the robustness of the proposed method in terms of the minimal false alarm rate achieved. Cancer detection can serve as a methodology for finding the target ROIs and aid the pathologists to proceed with further investigations.

6.4 Computational Complexity Analysis

We implemented the proposed system using C++ over MS Visual Studio 2015 in 64bit. This is a non-GPU implementation; however, Intel ® OpenMP API is used to obtain multiprocessing with a shared memory. The experiments are performed on a computer with I7-4700MQ at 2.4GHz (8 CPUs) and 16Gb memory. In theory, this should yield 8x speed improvement but in practice, the observed speed improvement was between 4.8x to 5x. As mentioned earlier, one of the crucial advantages of the proposed approach is its significantly low computational cost that makes real-time cancer detection and identification feasible. Specifically, with the aforementioned computer implementation, the total time for forward-propagation of a 64x64 patch over the adaptive CNN is about 88.35ms.

We used the MATLAB R2015a on a computer with i7-5600U CPU at 2.6GHz and 8 Gb RAM for the training and testing of conventional methods. It took around 57.00 ms for performing the classification on a single patch.

CHAPTER 7: CONCLUSIONS AND FUTURE WORK

7.1 Summary

With this work, we aimed to perform comparative evaluations among several recent approaches for the detection and identification of tumor tissues. In order to accomplish this objective, we have developed an image database for colorectal tumor biopsies in two modalities: multispectral and RGB. The variability of biopsy samples in terms of the number of patients is comparatively higher than those used in the existing researches. Our multispectral image database also remains unique in the sense we attempted to acquire images in the near infrared bands also.

The classification performances of texture based algorithms with state-of-the-art classifiers were experimented on our multispectral image dataset. A comparative study was also carried out in order to demonstrate the significance of multispectral images over RGB images in the realization of an automated cancer identification system. Furthermore, we have proposed a band selection algorithm to remove the redundant bands in the multispectral imagery. The significance of our algorithm is demonstrated well with the reduced computational time and improved classification accuracies.

As the main contribution of this thesis, we proposed a novel cancer detection and identification approach based on CNNs. The relevance of the algorithm can be realized when we have a scarce and low-resolution dataset. The proposed system is based on compact and adaptive CNNs and in contrast to the deep counterparts which require a massive size dataset, they can learn and generalize well on such limited and low-resolution data. Furthermore, we also performed comparative evaluations against the conventional

approaches which used state-of-the-art texture features. The proposed approach was able to achieve a comparable cancer identification accuracy despite the resolution was significantly low. However, it achieved the best cancer detection performance with a significant margin compared to conventional methods. Besides the elegant detection accuracy, it also provides very low false alarm rate with no manual feedback/tuning. Finally, it is a fully automatic method and can function without any additional hardware.

7.2 Future Work

Our study revealed that the images in the MSI database could perform better than those in the RGB database. This prompts us to experiment further with the multispectral images. Moreover, the proposed band selection algorithm can be utilized in order to lessen the computational complexity. As part of the future work, we intend to experiment and compare the CNN- texture-based techniques on the MSI dataset. The database can also be improved further by adding more biopsy images. We also aim to diversify our dataset by adding biopsy samples from more number of patients. However, collection of data from patients have to comply with the ethical aspects which should be considered carefully.

Currently, we have used the raw RGB images in the classification algorithms. Since our images are captured under halogen illumination mode, we need to verify whether a preprocessing stage such as stain normalization could improve our results. All of these will be the topic of our future work.

REFERENCES

[1] World Health Organisation, Fact Sheets, Available: http://www.who.int/mediacentre/factsheets/fs297/en/. Last Accessed: March 2018 .

[2] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, International journal of cancer. 136 (2015).

[3] L.C. Kobayashi, C. von Wagner, J. Wardle. Perceived Life Expectancy Is Associated with Colorectal Cancer Screening in England, Annals of Behavioral Medicine. 51 (2017) 327-336.

[4] M. Asiedu, S. Guillermo, N. Ramanujam. Low-cost, Speculum-free, Automated Cervical Cancer Screening: Bringing Expert Colposcopy Assessment to Community Health, Annals of Global Health. 83 (2017) 199.

[5] J.K. Turner, G.T. Williams, M. Morgan, M. Wright, S. Dolwani. Interobserver agreement in the reporting of colorectal polyp pathology among bowel cancer screening pathologists in Wales, Histopathology. 62 (2013) 916-924.

[6] A.M. Khan, N. Rajpoot, D. Treanor, D. Magee. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution, IEEE Transactions on Biomedical Engineering. 61 (2014) 1729-1738.

[7] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, Inf.Sci. 415 (2017) 190-198.

[8] T. Ince, M. Zabihi, S. Kiranyaz, M. Gabbouj, Learned vs. Hand-Designed Features for ECG Beat Classification: A Comprehensive Study, EMBEC & NBC 2017, Springer, 2017, pp. 551-554.

[9] S. Kiranyaz, T. Ince, R. Hamila, M. Gabbouj, Convolutional neural networks for patient-specific ecg classification, 37th Annual International Conference of the IEEE Engineering in Medicine Biology and Socitey(2015) 2608-2611.

[10] S. Kiranyaz, T. Ince, M. Gabbouj. Real-time patient-specific ECG classification by 1-D convolutional neural networks, IEEE Transactions on Biomedical Engineering. 63 (2016) 664-675.

[11] S. Kiranyaz, T. Ince, M. Gabbouj. Personalized Monitoring and Advance Warning System for Cardiac Arrhythmias, Scientific Reports. 7 (2017) 9270.

[12] A. Janowczyk, A. Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases, J.Pathol.Inform. 7 (2016) 29-3539.186902. eCollection 2016.

[13] G. Litjens, C.I. Sánchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis, Scientific reports. 6 (2016) 26286.

[14] L. Jiao, Q. Chen, S. Li, Y. Xu, Colon cancer detection using whole slide histopathological images, World Congress on Medical Physics and Biomedical Engineering (2013) 1283-1286.

[15] R. Peyret, A. Bouridane, S.A. Al-Maadeed, S. Kunhoth, F. Khelifi, Texture analysis for colorectal tumour biopsies using multispectral imagery, 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society(2015) 7218-7221.

[16] A. Manduca, M.J. Carston, J.J. Heine, C.G. Scott, V.S. Pankratz, K.R. Brandt, et al. Texture features from mammographic images and risk of breast cancer, Cancer Epidemiol.Biomarkers Prev. 18 (2009) 837-845.

[17] S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features, 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2008) 496-499.

[18] A. Chaddad, C. Tanougast, A. Dandache, A. Al Houseini, A. Bouridane, Improving of colon cancer cells detection based on Haralick's features on segmented histopathological images, IEEE International Conference on Computer Applications and Industrial Electronics (2011) 87-90.

[19] M.A. Sheha, M.S. Mabrouk, A. Sharawy. Automatic detection of melanoma skin cancer using texture analysis, International Journal of Computer Applications. 42 (2012) 22-26.

[20] A. Chaddad, C. Desrosiers, A. Bouridane, M. Toews, L. Hassan, C. Tanougast. Multi Texture Analysis of Colorectal Cancer Continuum Using Multispectral Imagery, PloS one. 11 (2016) e0149893.

[21] P. Khurd, C. Bahlmann, P. Maday, A. Kamen, S. Gibbs-Strauss, E.M. Genega, J.V.Frangioni, Computer-aided Gleason grading of prostate cancer histopathological images using texton forests, IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2010) 636-639.

[22] S. Bouatmane, M.A. Roula, A. Bouridane, S. Al-Maadeed. Round-Robin sequential forward selection algorithm for prostate cancer classification and diagnosis using multispectral imagery, Mach Vision Appl. 22 (2011) 865-878.

[23] B. Oranit, A. Chamidu, N. Hiroshi, A. Kota, K. Fumikazu, Y. Masahiro. Multispectral Band Analysis: Application on the Classification of Hepatocellular Carcinoma Cells in High-Magnification Histopathological Images, Journal of Cytology & Histology. 3 (2015) 1.

[24] X. Qi, F. Xing, D.J. Foran, L. Yang, Comparative performance analysis of stained histopathology specimens using RGB and multispectral imaging, SPIE Medical Imaging (2011) 79633B-79633B-9.

[25] A. Tabesh, M. Teverovskiy, H. Pang, V.P. Kumar, D. Verbel, A. Kotsianti, et al. Multifeature prostate cancer diagnosis and Gleason grading of histological images, IEEE Trans.Med.Imaging. 26 (2007) 1366-1378.

[26] R. Kumar, R. Srivastava, S. Srivastava. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features, Journal of medical engineering. 2015 (2015).

[27] A. Chaddad, C. Tanougast, A. Golato, A. Dandache. Carcinoma cell identification via optical microscopy and shape feature analysis, Journal of Biomedical Science and Engineering. 6 (2013) 1029.

[28] O. Abuzaghleh, B.D. Barkana, M. Faezipour, Automated skin lesion analysis based on color and shape geometry feature set for melanoma early detection and prevention, IEEE Long Island Systems, Applications and Technology Conference (2014) 1-6.

[29] S. Rathore, M. Hussain, A. Khan. Automated colon cancer detection using hybrid of novel geometric features and some traditional features, Comput.Biol.Med. 65 (2015) 279-296.

[30] R. Farjam, H. Soltanian-Zadeh, R.A. Zoroofi, K. Jafari-Khouzani, Tree-structured grading of pathological images of prostate, 5747 (2005) 840-852.

[31] K. Nguyen, B. Sabata, A.K. Jain. Prostate cancer grading: Gland segmentation and structural features, Pattern Recog.Lett. 33 (2012) 951-961.

[32] P. Huang, Y. Lai. Effective segmentation and classification for HCC biopsy images, Pattern Recognit. 43 (2010) 1550-1563.

[33] S. Naik, A. Madabhushi, J. Tomaszeweski, M.D. Feldman, A quantitative exploration of efficacy of gland morphology in prostate cancer grading, IEEE 33rd Annual Northeast Bioengineering Conference (2007) 58-59.

[34] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, J. Tomaszeweski, Automated grading of prostate cancer using architectural and textural image features, 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2007) 1284-1287.

[35] R. Sparks, A. Madabhushi. Statistical shape model for manifold regularization: Gleason grading of prostate histology, Comput.Vision Image Understanding. 117 (2013) 1138-1146.

[36] M.K.K. Niazi, K. Yao, D.L. Zynger, S.K. Clinton, J. Chen, M. Koyutürk, et al. Visually meaningful histopathological features for automatic grading of prostate cancer, IEEE journal of biomedical and health informatics. 21 (2017) 1027-1038.

[37] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, et al. Classification of breast cancer histology images using Convolutional Neural Networks, PloS one. 12 (2017) e0177544.

[38] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks, International Joint Conference on Neural Networks (IJCNN) (2016) 2560-2567.

[39] K. Sirinukunwattana, S.E.A. Raza, Y. Tsang, D.R. Snead, I.A. Cree, N.M. Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images, IEEE Trans.Med.Imaging. 35 (2016) 1196-1206.

[40] H. Haj-Hassan, A. Chaddad, Y. Harkouss, C. Desrosiers, M. Toews, C. Tanougast. Classifications of multispectral colorectal cancer tissues using convolution neural network, Journal of pathology informatics. 8 (2017).

[41] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, S. Li. Breast cancer multi-classification from histopathological images with structured deep learning model, Scientific reports. 7 (2017) 4172.

[42] N. Bayramoglu, J. Kannala, J. Heikkilä, Deep learning for magnification independent breast cancer histopathology image classification, International Conference on Pattern Recognition(2016) 2440-2445.

[43] Vinay Kumar, Ramzi S.Cotran, Stanley L., Robbins, Robbins. Basic Pathology, 7th Edition.

[44] Neil A. Shepherd, Bryan F. Warren, Geraint T. Williams, Joel K. Greenson, Gregory Y. Lauwers, Marco R. Novelli. Morson and Dawson's Gastrointestinal Pathology, 5th Edition.

[45] Steven G. Silverberg, Ronald A. DeLellis, William J. Frable, Virginia A. LiVolsi, Mark R. Wick. Silverberg's Principles and Practice of Surgical Pathology and Cytopathology, 2-Volume Set, 4th edition.

[46] S. Kunhoth, S. Al Maadeed, Building a multispectral image dataset for colorectal tumor biopsy, 13th International Wireless Communications and Mobile Computing Conference (IWCMC)(2017) 1745-1750.

[47] S. Al Maadeed, S. Kunhoth, A. Bouridane, R. Peyret, Multispectral imaging and machine learning for automated cancer diagnosis, 13th International Wireless Communications and Mobile Computing Conference (IWCMC) (2017) 1740-1744.

[48] S. Kunhoth, S. Al Maadeed, Multispectral Biopsy Image Based Colorectal Tumor Grader, Annual Conference on Medical Image Understanding and Analysis (2017) 330-341.

[49] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens. Multimodality image registration by maximization of mutual information, IEEE Trans.Med.Imaging. 16 (1997) 187-198.

[50] A. MartÍnez-UsÓMartinez-Uso, F. Pla, J.M. Sotoca, P. García-Sevilla. Clustering-based hyperspectral band selection using information measures, IEEE Trans.Geosci.Remote Sens. 45 (2007) 4158-4171.

[51] D.C. Cireşan, U. Meier, J. Masci, L.M. Gambardella, J. Schmidhuber. High-performance neural networks for visual object classification, arXiv preprint arXiv:1102.0183. (2011).

[52] V. Ojansivu, J. Heikkilä, Blur insensitive texture classification using local phase quantization, International Conference on Image and Signal Processing (2008) 236-243.

[53] V. Ojansivu, E. Rahtu, J. Heikkila, Rotation invariant local phase quantization for blur insensitive texture analysis, 19th International Conference on Pattern Recognition(2008) 1-4.

[54] T. Ojala, M. Pietikäinen, D. Harwood. A comparative study of texture measures with classification based on featured distributions, Pattern Recognit. 29 (1996) 51-59.

[55] T. Ojala, M. Pietikainen, T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans.Pattern Anal.Mach.Intell. 24 (2002) 971-987.

[56] R.M. Haralick, K. Shanmugam. Textural features for image classification, IEEE Trans.Syst.Man Cybern. (1973) 610-621.

[57] B. Scholkopf, A.J. Smola, Learning with kernels: support vector machines, regularization, optimization, and beyond, MIT press 2001.

[58] R. Berwick. An Idiot's guide to Support vector machines (SVMs), Retrieved on October. 21 (2003) 2011.

[59] C. Chang, C. Lin. LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (TIST). 2 (2011) 27.