PREFACE

Cancer detection has always been a major issue for the pathologists and medical practitioners for diagnosis and treatment planning. The manual identification of cancer from microscopic biopsy images is subjective in nature and may vary from expert to expert depending on their expertise and other factors which include lack of specific and accurate quantitative measures to classify the biopsy images as normal or cancerous one. The automated identification of cancerous cells from microscopic biopsy images help in alleviating the above mentioned issues and provides better results if the biologically interpretable and clinically significant feature based approaches are used for the identification of disease.

What is cancer: The cancer is nothing but an advance stage of tumor. Cancer is easier to treat and cure if it has diagnosed early. All cancers begin in cells, the body's basic unit of life. To understand cancer, it is helpful to know what happens when normal cells become cancer cells. Many types of cells make the body. These cells grow and divide in a controlled way to produce more cells as they needed to keep the body healthy. When cells are becoming old or damaged, are died and replaced by new cells. However, sometimes this orderly process goes wrong. The genetic material (DNA) of a cell can be damaged or changed, producing mutations that affect normal cell growth and division. When this happens, cells do not die when they should and new cells form when the body does not need them. The extra cells may form a mass of tissue called a tumor. Malignant tumors are cancerous. Cells in these tumors can invade nearby tissues and spread to other parts of the body. The spread of cancer from one part of the body to another part

of the body called metastasis. Some cancers do not form tumors. For example, leukemia is a cancer of the bone marrow and blood. Figure 1 describe the mutation of cancerous cell.

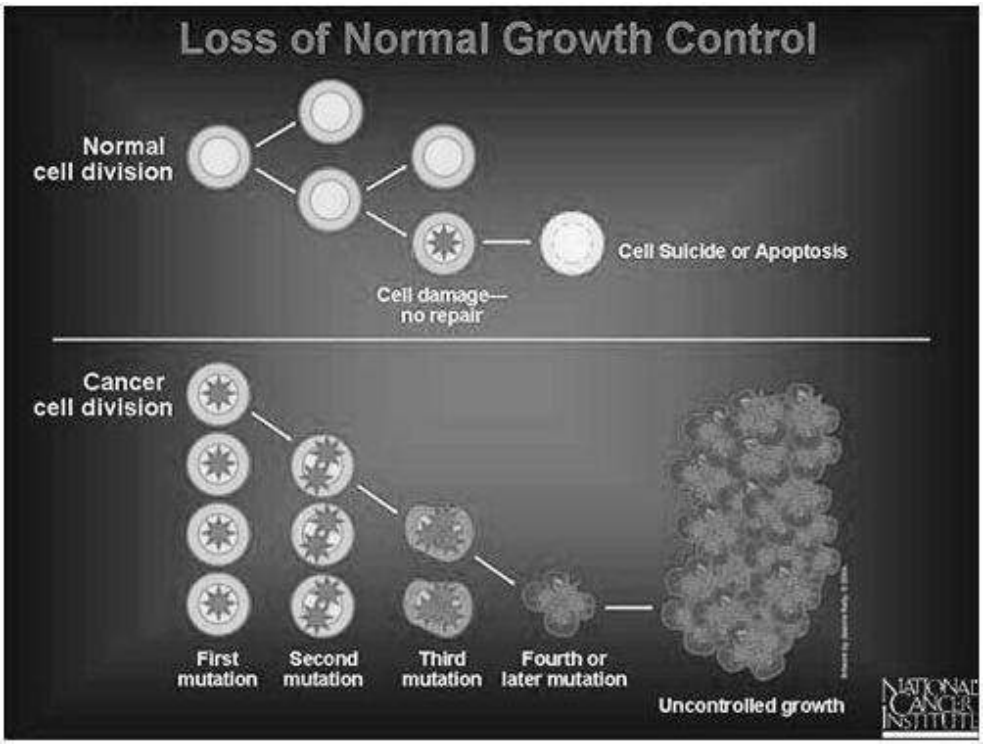

Figure 1: Understanding Cancer mutation of the cell (www.cancer.gov)

Types of cancer: There are about 200 type of cancer according to national cancer institute. Most of the cancer falls in following categories:

Carcinoma: The most common types of cancer arise from the cells that cover external and internal body surfaces. Lung, breast, and colon are the most frequent cancers of this type of cancer begin in the skin or tissues that line or cover internal organs.

Sarcoma: This type of cancer begins in the bones, Cartilage, fat, mussels and blood vessels or other connectives or supportive tissues.

Leukemia: This type of cancer starts in forming blood forming tissue, such as bone marrow and causes large number of abnormal blood cells.

Lymphoma and Myeloma: This type of Cancer begins in cell of immune systems.

Central nervous system cancers: This type of cancer begins in the tissues of the brain and spinal cord.

Recent Trends for cancer detection

Cancer screening: Cancer screening detects cancers when they are at an early stage or in the case of cervical cancer screening before they have developed. We know that screening saves thousands of lives each year.

Biopsy: A biopsy is the removal of a small amount of tissue for examination under a microscope. Other tests can suggest that cancer is present, but only a biopsy can make a definite diagnosis. A pathologist (a doctor who specializes in interpreting laboratory tests and evaluating cells, tissues, and organs to diagnose disease) analyzes the sample removed from the biopsy. There are different type of biopsies, classified by the technique and/or size of needle used to collect the tissue sample.

Needle biopsy: A biopsy uses a small needle to remove a small sample of cells.

Core needle biopsy: A core needle biopsy uses a larger needle to remove a larger sample of tissue. This is usually the preferred biopsy technique for finding out whether an abnormality on a physical examination or an imaging test is cancer. A vacuum-assisted biopsy removes more than one large core of tissue.

Surgical biopsy: A surgical biopsy removes the largest amount of tissue. This biopsy may be incision (removal of part of the lump) or decisional (removal of the entire lump). Because definitive surgery is best done after a cancer diagnosis has been made, a surgical biopsy is usually not the recommended way to diagnose

breast cancer. Most often, non-surgical core biopsies are recommended to diagnose breast cancer to keep surgery to one operation to remove the tumor if it is cancerous.

Image-guided biopsy: Image-guided biopsy is used when a distinct lump cannot be felt, but an abnormality is seen with an imaging test, such as a mammogram. During this procedure, a needle is guided to the location with the help of an imaging technique, such as mammography, ultrasound, or MRI. A stereotactic biopsy is done using mammography to help guide the needle. A small metal clip may be put into the breast to mark where the biopsy sample was taken, in case the tissue is cancerous and more surgery is needed. An image-guided biopsy can be done using a fine needle, core, or vacuum-assisted biopsy depending on the amount of tissue being removed.

About 32% of Indian population gets cancer at some point during their life time. Cancer is one of the common disease in India which has responsibility to maximum mortality with about 0.3 million death per year. The chances of getting affected by this disease are accelerated due to change in habits in the people such as increase in use of tobacco, deterioration of dietary habits, lack of activities, and many more. The possibility of cure from cancer is increased due to recent combined advancement in medicine and engineering. The chances of curing from cancer are primarily in its detection and diagnosis. The selection of the treatment of cancer totally depends on its level of malignancy. Medical professionals use several techniques for detection of cancer. These techniques may include various imaging modalities such as X-Ray, Computer Tomography (CT), Positron Emission Tomography (PET), Ultrasound, Magnetic Resonance Imaging (MRI); and pathological tests such as urine test, blood test etc.

For better detection of cancer pathologists' uses histopathology biopsy images that is the examination of microscopic tissue structure of the patient. Thus biopsy image analysis is a vital technique for cancer detection. Histopathology is the study of symptoms and indications of the disease using the microscopic biopsy images. To visualize various parts of the tissue under a microscope, the sections are dyed with one or more staining components. The main goal of staining is to reveal the components at cellular level and counter-stains are used to provide color, visibility and contrast. Hematoxylin-Eosin (H&E) is staining components that have been used by pathologists for over few decades. Hematoxylin stains cell nuclei as a blue in color while Eosin stains cytoplasm and connective tissues are of pink color. The histology is related to the study of cells in terms of structure, function and interpretations of the tissue and cells. Microscopic biopsies are most commonly used for both disease screening because of the less invasive natures. The characteristic of microscopic biopsy images has presence of isolated cells and cell clusters. The microscopic biopsy images are easier to analyze specimens compared to histopathology due to absence of non- complicated structures. The accurate manual identification of cancer from microscopic biopsy images has always been a major issue by the pathologists and medical practitioners observing at cell or tissue structure under the microscope.

Presently cancer detection is done by pathologists through evaluation of microscopic biopsy of cancerous tissues by examining the tissue structures, distributions of cells in tissue, regularities of cell shape to determine the level of abnormalities that may be present in the sample under investigation. The outcomes of these examinations may be normal, benign and malignant tissues. The manual evaluation of microscopic biopsy for cancer detection leads to

subjective, time consuming and varies with perceptions and level of expertise of pathologists. To overcome these challenges automated cancer diagnosis is needed for objective, fast, accurate and quantitative results. These steps involve image pre-processing, enhancement and restoration, segmentation, feature extraction to quantify properties of local area, and classification of sample image into normal and abnormal categories e.g. benign and malignant ones.

What do medical practitioners look under a microscopic biopsy image for cancer detection?

- Shape and size of cells
- Shape and size of nuclei
- Distribution of cell in the tissue

Challenges on cancer detection from microscopic biopsy images

- The first challenge is microscopic biopsy itself, because it has very complex structure and interpretation of microscopic biopsy images is very typical.
- Accurate segmentation of cell and nuclei is another challenge for cancer detection from microscopic biopsy images.
- A quantitative assessment of microscopic biopsy is required for prediction of level of abnormalities.
- Accurate grading is need for better survival, treatment and planning.

Objective of the thesis

The main objective of the present work is to design and implement a framework for cancer detection from microscopic biopsy images. The success of design and development of cancer detection model for microscopic biopsy images

depends on the development of an appropriate enhancement, segmentation, feature extraction and classification approaches. Among these steps the segmentation is most important steps of cancer detection. Segmentation process has ability to separate cell and nuclei from microscopic biopsy images to extract the local property shape and morphology of the cell.

In this thesis, following important problems of cancer detection from miscopy biopsy images are investigated

- Brief descriptions of the model for cancer detection techniques from microscopic biopsy images are presented.
- Over all CAD tools for detection and classification of cancer from microscopic biopsy images on four fundamental tissues type are implemented and investigated by selecting the best available approached in the literature in each stages.
- Design and development of incremental and hybrid model for segmentation of microscopic biopsy images to improve the overall results of cancer detection.
- Design and development of an adaptive fourth order PDE based FCM segmentation approach for segmentation of microscopic biopsy images corrupted with poison noise.

Contributions:

The main contributions of this thesis are as follows:

- A detailed literature review on computer aided diagnosis (CAD) tools for detection of cancer from microscopic biopsy images are investigated and analyzed.

- Clinically significant set of hybrid features are extracted from microscopic biopsy images. Based on above mentioned features, the best suited classification approaches are tested on four fundamental tissues of microscopic biopsy images.
- A hybrid color k-means segmentation approached is proposed, implemented and compared with other popular approaches for segmentation of breast tissues of microscopic biopsy images and its ROI segmented ground truth images. To improve the cancer detection accuracy.
- Fourth order partial differential equation (FPDE) based nonlinear filter adapted to Poisson noise (AFPDE) with fuzzy c-means (FCM) segmentation method is proposed. This method is capable of handling restoration, noise removal and segmentation of the microscopic biopsy images. The proposed approach is tested on breast cancer microscopic data set with ROI segmented ground truth images. The microscopic biopsy data set contains 31 benign and 27 malignant images of 896×768.

This thesis may aid researchers and medical practitioners of the area of cancer detection to go through the state-of-the-art methods for recent development in cancer detection and segmentation from microscopic biopsy images. The results come from the above approaches are very helpful to the doctors and pathologists. The proposed approach automatically determines whether the microscopic biopsy images are cancerous or not. If the cells are cancerous, the biopsy results can tell your doctor where the cancer originated and the type of cancer. This information may help guide treatment options. Other special tests on the cancer cells also can help to guide treatment choices.

Throughout this thesis the main objective is to design and develop an effective and efficient framework for cancer detection from microscopic biopsy images using image processing and pattern recognition techniques. The overall thesis is organized into six chapters. The abstract of each chapter are given as follows:

Chapter 1 provides the introduction, motivation and problem description for the present work including thesis scope/objectives, and contributions. Finally, this chapter concludes with the organization that illustrates the coverage of chapter in the thesis.

Chapter 2 provides a systematic survey on computer aided diagnosis (CAD) tools for detection of cancer from biopsy images using image processing and pattern recognition techniques. These steps involve image pre-processing, enhancement and restoration, segmentation, feature extraction to quantify properties of image, and classification of sample image into normal and abnormal categories e.g. benign and malignant ones.

In Chapter 3, a framework for automated detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features is proposed and examined. The various stages involved in the proposed methodology include enhancement of microscopic images, segmentation of background cells, features extraction, and finally the classification. An appropriate and efficient method is employed in each of the design step of the proposed framework after making a comparative analysis of commonly used method in each category. For the enhancement of the microscopic biopsy images, the contrast limited adaptive histogram equalization approach is used to highlight the details of the tissue and structures. For the segmentation of

background cells, k-means segmentation algorithm is used because it performs better in comparison to other commonly used segmentation methods. In feature extraction phase, it is proposed to extract various biologically interpretable and clinically significant shape as well as morphology based features from the segmented images. These include gray level texture features, color based features, color gray level texture features, Law's Texture Energy (LTE) based features, Tamura's features, and wavelet features. Finally, the K-Nearest Neighborhood (KNN) based method is used for classification of images into normal and cancerous categories because it is performing better in comparison to other commonly used methods for this application such as fuzzy KNN and Support Vector Machine (SVM) based classifiers. The performance of the proposed framework is evaluated using well known parameters for four fundamental tissues (connective, epithelial, muscular and nervous) of randomly selected 1000 microscopic biopsy images.

In Chapter 4, a hybrid combination of color k-means and marker control watershed based segmentation approach is proposed to be applied for the segmentation of cell and nuclei of microscopic biopsy images. The proposed approach is tested on breast cancer microscopic data set with ROI segmented ground truth images. Finally, the results obtained from proposed framework are compared with the results of popular segmentation algorithms; Fuzzy c-means, color k-means, texture based segmentation as well as adaptive thresholding approaches. The experimental analysis shows that the proposed approach is providing better results in terms of accuracy, sensitivity, specificity, false positive rate (FPR), Global consistency error (GCE), Random index (PRI), and variance of information (VOI) as compared to other segmentation approaches. The proposed

framework is associated with larger value of accuracy, sensitivity, specificity, random index (RI), and smaller value of FPR, FNR, GCE, and VOI in comparison to other methods. The performance measured for microscopic biopsy of breast tissues dataset in terms of the accuracy, sensitivity, specificity, FPR, FNR, PRI, GCE, and VOI were 0.9904, 0.8430, 0.9921, 0.0078, 0.1570, 0.9810, 0.0103 and 0.1030 respectively. Hence, we are in position to con-clude that, the proposed approach is performing better in terms of all parameters and it is suitable for the segmentation of microscopic biopsy images.

Chapter 5, present fourth order partial differential equation (FPDE) based nonlinear filter adapted to Poisson noise (AFPDE) with fuzzy c-means (FCM) segmentation method. This method is capable of effectively reducing the blocky artifacts while achieving good tradeoff between Poisson noise removals with edge preservation of the microscopic biopsy images. The proposed approach is tested on breast cancer microscopic data set with ROI segmented ground truth images. The microscopic biopsy data set contains 31 benign and 27 malignant images of 896×768. The region of interest (ROI) selected ground truth of all 58 images are also available for this data set Finally, the result obtained from proposed approach is compared with the results of popular segmentation algorithms; fuzzy c-means, color k-means, texture based segmentation, and total variation fuzzy c-means (TVFCM) approaches. The experimental results shows the proposed approach is providing better results in terms of various performance measures such as Jaccard Index, dice index, Tanimoto coefficient, area under curve (AUC), accuracy, true positive rate (TPR), true negative rate (TNR), false positive rate (FPR), false negative rate (FNR), random index (RI), global consistency error (GCE), and variance of information (VoI) as compared to other segmentation approaches used

for cancer detection. Experimental analysis and results it has been observed that the proposed framework is associated with larger value of accuracy, sensitivity, specificity, random index (RI), and smaller value of FPR, FNR, GCE, and VOI in comparison to other methods. The performance measured for microscopic biopsy of breast tissues dataset in terms of the Jaccard, dice, Tanimoto, accuracy, TPR, TNR, FPR, FNR, RI, GCE and VI are 0.5668, 0.7167, 0.5668, 0.9884, 0.6267, 0.9976, 0.0024, 0.3733, 0.9771, 0.0149 and 0.1262 respectively. From experimental analysis the proposed AFPDEFCM approach is performing better in terms of all parameters and it is suitable for the segmentation of microscopic biopsy images for cancer detection.

Finally in Chapter 6, we summarize main findings of this thesis and provides future perspectives of the research in this thesis. It describes the benefits of the presented work for the society, pathologist and the medical practitioners.