NUCLEI DETECTION AND SEGMENTATION IN DIGITAL HISTOPATHOLOGY

K.Mari Sheela[1], Dr.R.Velayutham[2]

[1]M.E. Scholar, Dept. Of CSE, Einstein college of Engineering, Tirunelveli, sheela.seenu03@gmail.com
[2] Head Of Department, Dept. Of CSE, Einstein College of Engineering, Tirunelveli.

Abstract: Cancer is a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body. There are more than 18 types of cancer, including breast cancer, skin cancer, lung cancer, blood cancer, mouth cancer and so on. Cancer is the second most common cause of death after heart disease. So the earlier cancer is diagnosed and treated, the better the chance of its being cured. Screening can help doctors find and treat several types of cancer early. Checking for cancer in people who have no symptoms is called screening. Early detection is important because when abnormal tissue or cancer is found early, it may be easier to treat. But in screening separate process is needed to find out the each cancer type. This is the main problem. To overcome this problem this project proposed one common screening system for all types of cancer. This is based on the cell because cancer is a disease of the cells in the body. So first take the cancer tissue as in the cell form. And then these cells are treated in same way. The main aim of this project is to segment the cancer affected cell from the given cell image and then find the cancer type and its severity such as begin, moderate, severe.

Keywords: Medical image processing, Screening System, segmentation,Histopathalogy

I.   INTRODUCTION

Medical Image Processing is one of the subdivision concepts of digital image processing. The digital image processing techniques have been used to analyze the scan reports of the patients to identify the diseases. There have been increases in the use of digital image processing techniques for the screening of disease after it was recommended as one of the method for screening methods. This system is going to implement the early diagnosis of cancer cell who have a no symptoms.cancer is a class of diseases characterized by out of control cell growth.there are over than 100 different types of cancer,and each is classified by the type of cell that is initially affected.the majority of cancers,some 90-95% of cases,are due to environmental factores. It is not generally possible to prove what caused a particular cancer because the various causes do not have specific fingerprints.for example,if a person who uses tobacco heavily develops ling cancer,then it was probably caused by the tobacco use,but since everyone has a small chance of developing ling cancer as a result of air pollution or radiation,the cancer may have developed for one of those reasons.cancer is generally not a transmissible disease.cancer types and its severity can be identify by,

1. Segment the cancer affected cell
2. After segmentation the features are extracted from the cancer affected cell to find the type of the cancer. This project uses the color, texture and shape features.
3. After features extraction ,cancer types and its severity can be find.

The aim of this project is to develop a system that have one common screening system is developed for all types of cancer. Using this project the cancer affected cell is segmented from the given cell image and then the cancer type and its severity is found.And also it includes that the system will suggest the appropriate treatment for the patient.

II. BACKGROUND

2.1 Medical image processing

Medical Imaging is a boon given by science and technology to humanity. While on the one hand medical diagnostics has advanced in leaps and bounds, on the other hand advancement in pattern recognition and machine intelligence techniques has intensified medical imaging. With these advancements it is now possible to make diagnostics in a non-invasive manner. An easy and correct diagnostic facilitates appropriate treatment earlier and thereby enhance chances of full recovery, while reducing the pain and suffering of patients.

2.2 Cancer

Cancer cells can also spread to other parts of the body. For example, cancer cells in the lung can travel to the bones and grow there. When cancer cells spread, it is called metastasis (muh-TAS-tuh-sis). When lung cancer spreads to the bones, it is still called lung cancer because that is where it started. It is not called bone cancer unless it started in the bones. Cancers can sometimes recur (or come back) after treatment,mostly at the site where they started but sometimes at a distant site like lungs, liver,brain, or bone.

Some cancers tend to grow and spread very quickly. Others grow more slowly. Cancers also respond to treatment in different ways. Some types are best treated with surgery.Others do better with drugs called chemotherapy (key-mo-THER-uh-pee). Often 2 or more treatments are used to get the best results.When someone has cancer, the doctor will want to find out what kind of cancer it is.

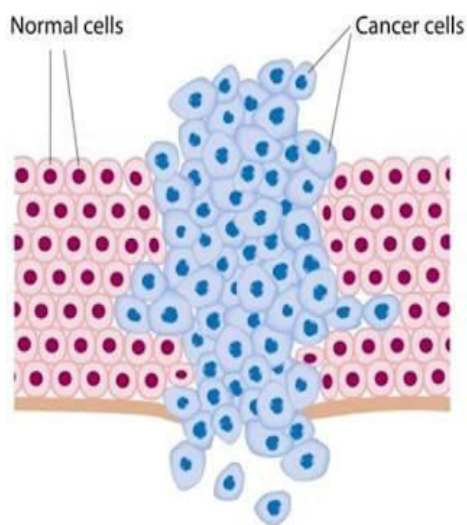People with cancer need treatment that is aimed at their type of cancer.

Figure 1 Diagram differentiate normal cell and cancer affected cell

2.3 Databases

This system uses database,which is used to store images. Cell image which is affected by cancer disease are collected from hospital, Internet and scanned from biomedical books. The shape of the cancer affected cell will differ according to the geographical locations. So, the different type of cancer image are stored. The analyzed image is compared with the stored images to know the stage of disease. For the process of comparing and retrieving the appropriate matched image, the SVM algorithm is used.

III.   EXISTING SYSTEM

Once a patient has doubt of having cancer, doctors study the angle of cancer cell, lenghtand diameter of normal cell and visual fields. All these term should be analyzed at together to know a complete picture of the type and nature of cancer.

In existing system, they used techniques like standardness to address the colour problem.The segmentation of nuclei on histopathological images is more difficult. Since most of the nuclei are often part of histopathological structures presenting complex and irregular visual aspects.

To overcome the above problems the proposed method is used to detect and to segment the nucleus from the histopathological images.

IV.   PROPOSED SYSTEM

The treatment success is based on how early we detect the disease. In the existing system, there may be a chance of late detection of the disease identification. If the patient is suffered from diabetes or heart disease the existing system may take longer time to diagnosis. In order to fasten the diagnosis process the proposed system makes use of medical image processing techniques. One of the advantages in proposed system is that it does not take much time to identify the disease. In the proposed system the reports are used to analyze the disease cancer is present or not. Analysis is done through the comparison between the processed image and stored sample image in the database. Then the treatment for that disease is also suggested.

This project proposed the techniques to implement the screening system for detecting all types of cancer using histopathology images to get an idea about the caused cancer and its severity. In pre-processing step, colour normalization is performed.Nuclei Detection is to perform the blob detection.Nucleus Segmentation process is to detect the cancer affected cells using mean shift and outlier removal algorithms.Nucleus Separation process is done by using LRRS.SVM is used to classify the nucleus with the trained image from the dataset
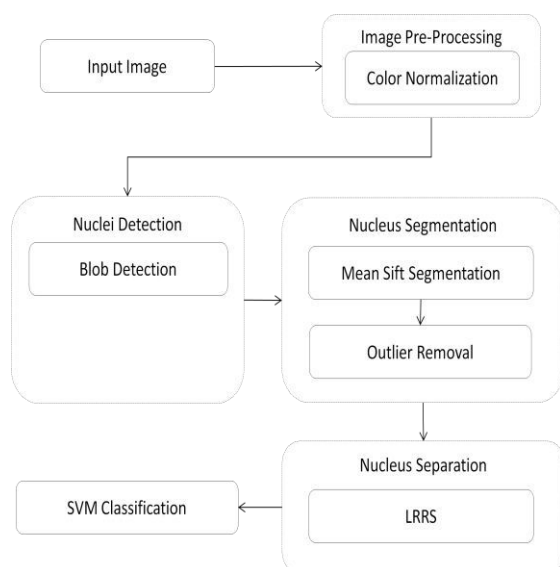
Figure 2 Block Diagram of detect and segment the cancer affected cell in histopathalogy image

The output of the preprocessing stage is passed to the next stage. In that stage the background pixel is segmented using K-mean clustering algorithm with two cluster class centre. After segmentation process, disease identification will be carried out. In this stage the above processed image is compared with the sample images in the database by using pattern recognition and pattern matching algorithms. After identification of the disease, the appropriate treatment methodology is also suggested.

V METHODOLOGY

5.1 PREPROCESSING STAGE

Image Pre-processing is used to enhance the visual appearance of images and Improve the manipulation of datasets. This project using the color normalization for image Pre-processing. Color normalization is concerned with artificial color vision and object recognition. In general, the distribution of color values of histopathology image depends on the staining process, different lighting conditions and microscope feature. Color also very important factor and input for segmentation algorithm. Color normalization allows for object recognition. Techniques based on color, to compensate for these variations. Algorithm that works for every pixel in the image is given by

$$\text{Normalized Red} = \frac{r}{\sqrt{r^2 + g^2 + b^2}},$$

$$\text{Normalized Green} = \frac{g}{\sqrt{r^2 + g^2 + b^2}},$$

$$\text{Normalized Blue} = \frac{b}{\sqrt{r^2 + g^2 + b^2}}$$

5.2 SEGMENTATION

The main objective of segmentation is to group the image into regions with same property or characteristics. It plays a major role in image analysis system by facilitating the description of anatomical structures and other regions of interest. Method of image segmentation include: mean shift segmentation and Outlier Removal.

5.2.1 Nuclei Detection using Mean Shift Segmentation

Initial segmentation means decompose the original image in to small parts. It reduce the variations of the images. Variation means texture, shape etc. In this project initial segmentation use the mean shift segmentation algorithm.

Mean shift Segmentation Algorithm

1. Consider the given image.
2. The image is divided into pixels.
3. n be the total number of pixels.
4. Find the feature vector set and initialized
5. The mean shift segmentation is defined as

$$m_G(y_i) = \frac{\sum_{j=1}^{n} x_j g\left(\left\|\frac{y_i - x_j}{h}\right\|^2\right)}{\sum_{j=1}^{n} g\left(\left\|\frac{y_i - x_j}{h}\right\|^2\right)} - y_i$$

5.2.2 Outlier Removal

The detected blobs without outliers will now serve as the internal markers marking the rough position of potential nuclei. External markers roughly delineating cell boundaries will help form a complete marker map for the watershed transform with the marker-controlled strategy. Since the external markers must not be too close to the boundaries of the regions that are nuclei to be segmented.

5.3 NUCLECI SEPARATION

After the segmentation stage disease separation will be performed. This segmentation

algorithm is used to segment the nuclei regions. It is also called as split and merge algorithm. After mean shift segmentation calculate the intensity for each pixels. The mean shift segmented image contains many segmented regions. This region set Is given to the input of LRRS segmentation.

5.3.1 Local Region Recursive Segmentation (LRRS)

LRRS Algorithm

To get the input region set from the mean shift segmented image.

- Calculate mean intensity for each region.
- Find the global threshold using otsu method.
- Then truncate the region whose mean intensity is greater than global threshold value.
- After truncation remove the cytoplasm for most regions.
- The remaining regions are merged and form a new region set.
- These regions are contains several nuclei and noisy elements.
- For each merged region estimate the number of intensity values and the area.
- The local region which satisfies the following conditions.

$$A(R'_p) > T_{area} \quad \text{and} \quad V > 2$$

- Finally create the candidate nuclei region set.
- Then apply the split and merge algorithm until the regions to satisfy the above condition.

5.4 NECLECI CLASSIFICATION

In classification use SVM is used to classify the nucleus. It classifies the original image with the trained image from the dataset.

5.4.1 SVM Classification

In machine learning support vector machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis.

Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible.

Finding the Closest Pair of Points

- First of all, we observe that, finding the closest pair of points in kernel space requires n2 kernel computations where n represents the total number of data points. But, in case we use a distance preserving kernel like the exponential kernel the nearest neighbors in the feature space are the same as the nearest neighbors in the kernel space. Hence we need not perform any costly kernel evaluations for the initialization step.

Adding a Point to the Support Vector Set

- Given a set S which contains only Support Vectors, we wish to add another Support Vector c to S.

Pruning

- When we say that a point p 2 S is blocking the addition of c to S what we mean is that p of that point may become negative due to the addition of c to S. What it physically implies is that p is making a transition from S to the well classified set R.

```
Algorithm 1 Simple SVM
candidateSV = { closest pair from opposite classes }
while there are violating points do
    Find a violator
    candidateSV = candidateSV ∪ violator
    if any αₚ < 0 due to addition of c to S then
        candidateSV = candidateSV \ p
        repeat till all such points are pruned
    end if
end while
```

5.5 TREATMENT SUGGESTION

Treatment suggestion is the final stage of this system. The type of disease and its stage is identified in the previous phase. Based on the severeness of the disease the treatment is suggested to the patient. Depending upon the percentage of matching with disease affected image the appropriate treatment will be suggested from the data collection.

VI CONCLUSION

This system is very helpful in the medical field at the hospital environment. It is more comfortable way for the patient in diagnosing and early detection of the diseases. On the other hand, it is the flexible way for the histopothologists to analyze the reports. The proposed automatic cancer disease analyzer would be useful at remote locations where medical experts are not available. The cost required to this system is affordable. The existing system would take much amount, especially for diabetes and heart patients more report should be analyzed like body pressure, urea, etc. But in the proposed system only the scan report of the patient considered, not other factors like whether the person is already affected by any other disease or disorders.

REFERENCES

[1] B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos- Hoff, "Iterative voting for inference of structural saliency and characterization of subcellular events," IEEE Trans. Image Process., vol. 16, no. 3,pp. 615–623, Mar. 2007.

[2] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," IEEE Trans. Pattern Anal.Mach. Intell., vol. 25, no. 8, pp. 959–973, Aug. 2003.

[3] L. Yang, P. Meer, and D. J. Foran, "Unsupervised segmentation based on robust estimation and color active contour models," IEEE Trans. Inf. Technol. Biomed., vol. 9, no. 3, pp. 475–486, Sep. 2005.

[4] X. Qi, F. Xing, D. J. Foran, and L. Yang, "Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set," IEEE Trans. Biomed. Eng., vol. 59, no. 3, pp. 754–765, Mar. 2012.

[5] A. D. Belsare1 and M. M. Mushrif "Histopathological Image Analysis Using Image Processing Techniques: An Overview" Signal &