International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 9, November- 2012

Survey on Computer Aided Diagnosis of Cancer

Shanti Kadachha[1], Asst. Prof. Pinal Patel[2]

[1]Department of CSE, Government Engineering College, Sector-28, Gandhinagar Gujarat, India.

[2]Department of CSE, Government Engineering College, Sector-28, Gandhinagar Gujarat, India.

Abstract

With doctors increasingly relying on digital images, like those from MRIs, to diagnose cancer, image analysis often plays a key role in crucial decisions for patients. The more accurate the analysis, the more likely the patient will receive appropriate—and timely—treatment. Computer-aided diagnosis or detection (CAD) is a concept established by taking into account equally the roles of physicians and computers. With CAD, the performance by computers does not have to be comparable to or better than that by physicians, but needs to be complementary to that by physicians. In fact, a large number of CAD systems have been employed for assisting physicians in the early detection of various types of cancers [1].

This paper presents a survey of Computer Aided Cancer Detection and presents general computational steps which are needed in diagnosis of cancer.

1. Introduction

Today, cancer constitutes a major health problem. In the United States, it is the second leading cause of death. Approximately one out of every two men and one out of every three women get cancer at some point during their lifetime. Furthermore, the risk of getting cancer has been further increasing due to the change in the habits of people in our century such as the increase in tobacco use, deterioration of dietary habits, and lack of activity [2]. Cancer is an enormous global health burden, touching every region and socioeconomic level. Today, cancer accounts for one in every eight deaths worldwide – more than HIV/AIDS, tuberculosis, and malaria combined. In 2008, there were an estimated 12.7 million cases of cancer diagnosed and 7.6 million deaths from cancer around the world. More than 60 percent of all cancer deaths occur in low- and middle-income countries, many of which lack the medical resources and health systems to support the disease burden. Moreover, the global cancer burden is growing at an alarming pace; in 2030 alone, about 21.4 million new cancer cases and 13.2 million cancer deaths are expected to occur, simply due to the growth and aging of the population [3].

In the 1980s, the concept of automated diagnosis or automated computer diagnosis was already known from studies in the 1960s, but these early attempts were not successful. Thus, it appeared to be extremely difficult to carry out a computer analysis on lesions involved in medical images. It was, therefore, not easy to predict whether the development of CAD schemes would be successful or fail. Therefore, we thought that we should select research subjects which had the potential to have a major impact in medicine, if CAD could be developed successfully. Some of the most important subjects in medicine at that time were related to cardiovascular diseases, lung cancer, and breast cancer [1].

The main idea of CAD is to assist radiologists in interpreting medical images by using dedicated computer systems to provide ‘second opinions’. The final medical decision is made by the radiologists. Studies on CAD systems and technology show that CAD can help to improve diagnostic accuracy of radiologists, lighten the burden of increasing workload, reduce cancer missed due to fatigue, overlooked or data overloaded and improve inter- and intra-reader variability [4]. The primary goal of CAD is to increase the detection of disease by reducing the false negative rate due to observational oversights [5]. CAD is used in the diagnosis of breast cancer, lung cancer, colon cancer, prostate cancer, bone metastases, coronary artery disease and congenital heart defect [6].

2. Overview

The automated cancer diagnosis consists of three main computational steps: preprocessing, processing, feature extraction, and diagnosis. The aim of the preprocessing step is to eliminate the background noise and improve the image quality for the purpose of determining the focal areas in the image. The next step processing comprises nucleus/cell segmentation in the case of extracting cellular-level information. The preprocessing and preprocessing becomes the most important yet difficult step for a successful feature extraction and diagnosis.

After processing the image, features are extracted either at the cellular or at the tissue-level. The cellular-level feature extraction focuses on quantifying the properties of individual cells without considering spatial dependency between them. For a single cell, the morphological, textural, fractal, and/or intensity-based features can be extracted. The tissue-level feature extraction quantifies the distribution of the cells across the tissue; for that, it primarily makes use of either the spatial dependency of the cells or the gray-level dependency of the pixels. For a tissue, the textural, fractal, and/or topological features can be extracted.

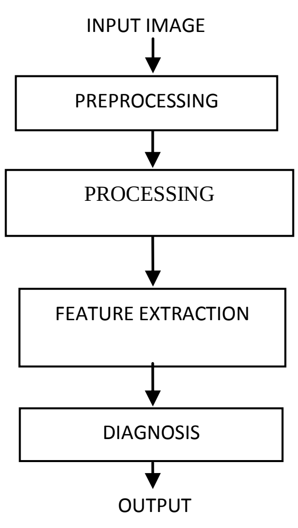

Figure 1. General steps in Computer Aided Cancer Detection

The aim of the diagnosis step is (i) to distinguish benignity and malignancy or (ii) to classify different malignancy levels by making use of extracted features. This step uses statistical analysis of the features and machine learning algorithms to reach a decision. An overview of these three steps is given in Figure 1 [2].

3. Preprocessing

The pre-processing stage prepares the image for further processing, analysis and interpretation. It consists of two important steps before processing of the image.

The first step involves the removal of artefact and unwanted parts in the background of the image. Adaptive Median filtering has been found to be very powerful in removing noise from two-dimensional signals without blurring edges. This makes it particularly suitable for enhancing images [7]. Then, an enhancement process is applied to the digital images. Image enhancement operations can be used to improve the appearance of images, to eliminate noise or error, or to accentuate certain features in an image. The contrast limited adaptive histogram equalization (CLAHE) method seeks to reduce the noise produced in homogeneous areas and was originally developed for medical imaging. This method has been used for enhancement to remove the noise in the pre-processing of digital images [8]. Selective median filtering with CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm will remove the noise and enhances the image for better segmentation [9].

4. Processing

After preprocessing of an image, processing step takes place and it is an important step in cancer detection. Processing of an image mainly include the segmentation and morphological operations.

4.1. Segmentation

The first process in this stage is segmentation which is the fundamental step in image analysis or image understanding. Segmentation divides image into its constituent regions or objects. The level to which segmentation is carried out depends upon the problem being solved [10]. It aids in extracting regions of interest in an image. A good segmentation must be able to separate objects from the background to obtain the region of interest [11]. For the segmentation of intensity images like digital mammograms, there are four main approaches, namely, watershed, threshold techniques, boundary-based methods and region-based methods [10].

- Watershed Transformation-:
  The watershed transformation considers the gradient magnitude of an image as a topographic surface. Pixels having the highest gradient magnitude intensities (GMIs) correspond to watershed lines, which represent the region boundaries. Water placed on any pixel enclosed by a common watershed line flows downhill to a common local intensity minimum (LIM). Pixels draining to a common minimum form a catch basin, which represents a segment [12].

- Threshold Technique-:
  The simplest method of image segmentation is called the thresholding method. This method is based on a clip-level (or a threshold value) to turn a gray-scale image into a binary image. The key of this method is to select the threshold value (or values when multiple-levels are selected) [12].

- Boundary/Edge Based Method-:
  Edge based segmentation is the location of pixels in the image that correspond to the boundaries of the objects seen in the image. It is then assumed that since it is a boundary of a region or an object then it is closed and that the number of objects of interest is equal to the number of boundaries in an image [13].

- Region-based methods-:
  The region based segmentation is partitioning of an image into similar/homogenous areas of connected pixels through the application of homogeneity/similarity criteria among candidate sets of pixels. Each of the pixels in a region is similar with respect to some characteristics or computed property such as colour, intensity and/or texture. The assumption in these techniques is that the partitions that are formed correspond to objects or meaningful parts of the image [13].

4.2. Morphological Operation

The resultant image has region along with small spurious objects which can be removed by performing morphological operations. Morphology is an operation of image processing based on shapes. Input and the output images are of same size. Output image is generated by applying a structuring element to the input image. Some of the basic morphological operations are dilation, erosion, top-hat transform, bottom-hat transform etc. Dilation and erosion operations are the basis of all other operations [14].

Let E be a Euclidean space or an integer grid, and A a binary image in E.

4.2.1. Erosion [15]

The erosion of the binary image A by the structuring element B is defined by:

$$A \ominus B = \{z \in E | B_z \subseteq A\} \qquad (1)$$

Where B_z is the translation of B by the vector z, i.e.,

$$B_z = \{b + z | b \in B\}, \forall z \in E \qquad (2)$$

The erosion of A by B is also given by the expression:

$$A \ominus B = \bigcap_{b \in B} A_{-b} \qquad (3)$$

4.2.2. Dilation [15]

The dilation of A by B is defined by:

$$A \oplus B = \bigcup_{b \in B} A_b \qquad (4)$$

The dilation can also be obtained by:

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \varnothing\} \qquad (5)$$

where B^s denotes the symmetric of B, that is,

$$B^s = \{x \in E | -x \in B\} \qquad (6)$$

5. Feature extraction

Feature extraction is a method of capturing visual content of an image. The objective of feature extraction process is to represent raw image in its reduced form to facilitate decision making process such as pattern classification. Various types of Features are:

- The **morphological** features provide information about the size and the shape of a nucleus/cell.
- The **textural** features provide information about the variation in the intensity of a surface and quantify properties such as smoothness, coarseness, and regularity.
- The **fractal-based** features provide information on the regularity and complexity of a cell/tissue by quantifying its self-similarity level.
- The **topological** features provide information on the cellular structure of a tissue by quantifying the spatial distribution of its cells.
- The **intensity-based** features provide information on the intensity (gray-level or colour) histogram of the pixels located in a nucleus/cell.

Feature extraction step is important step to get high classification rate. A set of features are extracted in order to allow a classifier to distinguish between normal and abnormal pattern. The abnormality can be identified on the basis of textural appearance. Extracted features are used in neural classifier to train it for the recognition of particular class either normal or abnormal. The ability of the classifier to assign the unknown object to the correct class is dependent on the extracted features [16].

6. Diagnosis

After determining an appropriate set of features, the next step is to distinguish the malignant structures from their counterparts. In this step, a cell or a tissue is assigned to one of the classes of cancerous, benign, or healthy. As a part of diagnosis, it is also possible to classify the malignancy level of the tissues (i.e., grading). In this case, the classes are the possible grades of the cancer of interest.

For diagnosis, one group of studies employs a statistical test on the features. Another group of studies uses machine learning algorithms to learn (from data) how to distinguish the different classes from each other. Among those algorithms are the neural networks, k-nearest neighbourhood algorithm, logistic regression method, fuzzy systems, linear discriminate functions, and decision trees [2].

7. Conclusion

This paper presents the survey of Computer Aided Detection/Diagnosis (CAD) of Cancer. General steps in CAD based cancer detection are Preprocessing, Processing which include segmentation and morphological operations, Feature Extraction and diagnosis which determines the abnormalities.

In the preprocessing step, the focal areas determined. This comprises to eliminate the noise and improve the image quality. Segmentation step is used for extracting the regions of interest in an image. Morphological Operations are used for removing the unwanted clusters. Next is the feature extraction step. This step quantifies the properties of the biological structures of interest, extracting features. After determining an appropriate set of features, Diagnosis step determines the suspected malignant in an image.

8. Acknowledgment

The authors are grateful to Department, Faculties and friends who are directly or indirectly involved in this work.

9. References

[1] Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential Kunio Doi, Ph.D, 2007.

[2] Cigdem Demir and B¨ulent Yener, "Automated cancer diagnosis based on histopathological images: a systematic survey", Technical Report, Rensselaer Polytechnic Institute, Department Of Computer Science, Tr-05-09.

[3] Cancer Facts & Figures 2012, American Cancer Society, Atlanta, Georgia.

[4] Hiroshi Fujitaa, Yoshikazu Uchiyamaa, Toshiaki Nakagawaa, Daisuke Fukuokab,Yuji Hatanakac, Takeshi Haraa, Gobert N. Leea, Yoshinori Hayashia,Yuji Ikedoa, Xin Gaoa, Xiangrong Zhoua" Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs" computer methods and programs in biomedicine 9 2 (2 0 0 8) 238–248

[5] Ronald A Castellino "Computer aided detection (CAD): an overview" Cancer Imaging. 2005; 5(1): 17–19.

[6] Computes Aided Diagnosis from Wikipedia, the free encyclopedia

[7] S. Vijaya Kumar, M.Naveen

[8] Lazarus, and C. Nagaraju" A Novel Method for the Detection of Morphological Gradient Watershed Segmentation Algorithm" International Journal of Engineering Science and Technology Vol. 2(7), 2010, 2616-2622

[9] Prof. Samir Kumar Bandyopadhyay "Pre-processing of Mammogram Images" International Journal of Engineering Science and Technology, Vol. 2(11), 2010, 6753-6758

[10] B. Senthilkumar" A Review on Computer Aided Detection and Diagnosis - Towards the Treatment of Breast Cancer" European Journal of Scientific Research, ISSN 1450-216X Vol.52 No.4 (2011), pp.437-452

[11] Ganesh Choudhari, Debabrata Swain, Dipali Thakur ,and Kiran Somase" Colorography: an Adaptive Approach to classify and detect the Breast Cancer using Image Processing" International Journal of Computer Applications (0975 – 8887) Volume 45– No.17, May 2012

[12] Rafael C. Gonzalez, Richard E. Woods,"Digital Image Processing", Pearson Education, Inc. and Dorling Kindersley Publications, Inc.

[13] Segmentation (image processing) From Wikipedia, the free encyclopedia

[14] Tranos Zuva, Oludayo O. Olugbara, Sunday O. Ojo and Seleman M. Ngwira" Image Segmentation, Available Techniques, Developments and Open Issues" anadian Journal on Image Processing and Computer Vision Vol. 2, No. 3, March 2011

[15] D. Narain Ponraj, M. Evangelin Jenifer, P. Poongodi, and J. Samuel Manoharan" Morphological Operations for the Mammogram Image to Increase the Contrast for the Efficient Detection of Breast Cancer" European Journal of Scientific Research ISSN 1450-216X Vol.68 No.4 (2012), pp. 494-505 M

[16] Morphological Operations From Wikipedia, the free encyclopedia

[17] R.nithya and B.santhi" Comparative Study On Feature Extraction Method For Breast Cancer Classification" Journal of Theoretical and Applied Information Technology 30th November 2011. Vol. 33 No.2