Journal of Scientific & Industrial Research
Vol. 79, February 2020, pp. 132–134

Pulmonary Tumor Detection by virtue of GLCM

G Kanagaraj[1]* and P Suresh Kumar[2]

[1]Department of Electronics & Communication Engineering, AVS Engineering College, Salem-003, India

[2]Department of Electrical& Electronics Engineering, Mahendra Engineering College (A), Namakkal, India

Received 13 September 2018; revised 19 August 2019; accepted 20 October 2019

As per the technical evolution and latest trend, Image processing techniques has become a boon in medical domain especially for tumor detection. Presence of tumor in Lungs which leads to lung cancer is a prominent and trivial disease at 18%. This is important to be detected at early stage thereby decreasing the mortality rate. The survival rate among people increased by early diagnosis of lung tumor. Detection of tumor cell will improve the survival rate from 14 to 49%. The aim of this research work is to design a lung tumor detection system based on analysis of microscopic image of biopsy using digital image processing. This can be done using Gray Level Co-Occurrence Matrix (GLCM) method and classified using back propagation neural network. This method is used for extracting texture features based on parameters such as contrast, correlation, energy, and homogeneity from the lung nodule. The microscopic lung biopsy images are classified into either cancer or non-cancer class using the artificial neural network algorithm. The proposed system has proven results in lung tumor detection and diagnosis.

Keywords: Gray Level Co-Occurrence Matrix, microscopic lung biopsy, artificial neural network algorithm

Introduction

The digital image processing involves texture features extraction using the Gray Level Co-Occurrence Matrix (GLCM) method and image classification using the back propagation neural network algorithm[1]. It uses artificial neural networks in the classification stage[2]. The system extracts hidden knowledge from a historical lung cancer disease database prediction system can be further enhanced and expanded[3].One of the techniques is content based image retrieval Computer Aided Diagnosis System (CAD) for early detection of lung nodules from the Chest Computer Tomography (CT) images[4]. For determining the cancer cells from medical images, various image processing and soft computing techniques can be used. Basically CAD system involves three steps. Those steps are as: Pre-processing, segmentation of the lung and classification of the nodule candidates[5]. Recently, machine learning methods on Computed Tomography (CT) images have been used in the diagnosis of lung cancer to accelerate the diagnosis process and assist physicians[6].

Proposed System

Texture is an important characteristic for the analysis of many types of images because it provides a rich source of information about the image. Also it provides a key to understand basic mechanisms that underlie human visual perception. In this research four statistical feature of texture (Contrast, Correlation, Homogeneity and Energy) was calculated from gray level Co-occurrence matrix (GLCM) of equal blocks (30×30) from both tumor tissue and normal tissue of three samples of CT-scan image of patients with lung cancer. It was found that the contrast feature is the best to differentiate between textures, while the correlation is not suitable for comparison, the energy and homogeneity features for tumor tissue always greater than its values for normal tissue. The proposed system uses image processing techniques such as optimal thresholding, total variation denoising, morphological operations for segmentation, Gray-level co-occurrence matrix feature extraction (GLCM). Basic representation of proposed system shown in figure 1.

Input Image

In our research, CT methodology has been used. Computed Tomography (CT) imaging also known as ―CAT scanning (Computed Axial Tomography), uses the digital computer together with a rotating X-ray device that generates detailed cross sectional images or slices of the different organs and various body parts such as the spine, lungs, liver, kidneys,

*Author for Correspondence
E-mail: raj.techstorm@gmail.com

Fig. 1 — Block Diagram of Proposed System

pelvis, pancreas, pelvis, extremities Brain and blood vessels. Unlike to MR imaging, CT is efficient as it has the ability to image a combination of soft tissue, bone, and blood vessels among the various imaging techniques such as MRI and X-ray. For example, conventional x-ray imaging of the head can only show the dense bone structures whereas this imaging does an excellent job of showing soft tissue and blood vessels, but MR is not that useful in detailing the bony structures.

The next step in GLCM based proposed method is Image Denoising. It is necessary to reduce the image noise for accurate analysis and diagnosis. Any noise reduction algorithm aims to improve the fidelity of an image which actually removes the uncorrelated and random structures and retains the resolution.

Total Variation

Usually, CT images have random noise. Here, we have chosen the non-linear total variation based on partially differential equation algorithm proposed by Leonid et at. As it effectively smoothens the image and removes random noise present. The 2D lung CT image is the input to the denoising module in JPEG format of size 512 x 512 shown in figure 2(a) & (b). Let $X_0(a,b)$ denote the pixel values of a noisy image. Let x(a,b) denote the desired clean image, and n be the additive noise, which is represented in

$$X_o(a,b)=x(a,b)+n(a,b).$$

The total variation of the image is minimized subject to a constraint that involves the statistics of the noise and the Lagrange multipliers are used to impose as in. With time as a parameter, the solution is obtained using the gradient projection method.

Segmentation

In proposed system, Segmentation is done by extracting the tumor from the area of interest (lung) and then analyses the separately obtained area for nodule detection. For normal lung, segmentation can be done by making use of excellent contrast between surrounding tissues and the air.

Thresholding

Thresholding is the most widely used method for image segmentation. The foregrounds discriminated from the background in this process. The Otsu's

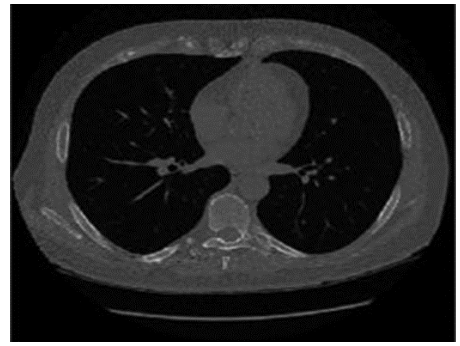

Fig. 2(a) — Original Noisy Image

Fig. 2(b) — Tumor Image

method is one of the best global thresholding methods. It performs unsatisfactorily for those poor quality images that have low contrast and non-uniform illumination but works well with dearly scanned images. The pre-processed lung CT consists of high intensity pixels in the body and low intensity pixels in the lung and surrounding air. To segregate the body and non-body pixels through an iterative procedure, optimal thresholding proposed and applied on the pre-processed lung image. Ta is the segmentation threshold. The initial threshold To is average Segmentation of Lung Tumor Using GLCM Techniques. The mean gray-level of the body pixels and non-body pixels after segmentation are denoted as µb and µn. The new threshold Ta+1is determined using equation. Ta+1 = µx+µy. Repeat steps until there is no significant difference between threshold values in successive iterations (Ta+1 = T). The lung image after segmentation with optimal threshold value contains non-body pixels and using morphological operation the low-density regions within the image is removed.

Morphology

Lung nodules are the small masses of tissue in the lung. They appear approximately round and white in CT scan or X-ray images. In the proposed method,

Table 1 — Features of Extracted Image

FEATURES	VALUES
Tumor Area	426
Tumor convex area	515
Equiv diameter	22.314
Tumor Convex Eccentricity	0.423
Tumor Convex Solidity	0.736
Image contrast	315.514
Image correlation	0.823
Image Energy	0.0633
Image Homogeneity	0.3576

lung nodule is our region of interest. 4- Connected labelling algorithm is used in the proposed method. Overview of 4- connected labelling algorithm is explained as taken a binary image, then find negative of the image. Check the north and west pixel for every pixels. The pixel belongs to the same region if the pixel to the west or north has the same intensity value. Assign the same label to the current pixel. If the pixel to the north has the same value and the pixel to the west has a different value, assign the north pixel's label to current pixel. If the north and west neighbours of pixel have different pixel values, create a new label and assign that label to the current pixel.

Feature Extraction

The process of extracting the textural features from the spatial distribution that can be used to characterize images is Feature extraction. GLCM is one the best method that describes the image texture. GLCM is calculated as how often a pixel with the intensity an occurs in a specific spatial relationship to pixel with value bits defined as a matrix of relative frequencies p (a, b). Following parameter has been calculated and presented in table 1.

- Area: It is a scalar value that gives the actual number of pixels in the Region Of Interest (ROI).
- Convex Area: It is a scalar value that gives the number of pixels in convex image of the ROI which is a binary image with all pixels within the hull filled in.
- Equivalent Diameter: It is the diameter of a circle with the same area as the ROI as defined in below. Equiv Diameter = $4*Area\pi$.
- Solidity: It is the proportion of the pixels in the convex hull that are also in the ROI. Solidity = Area Convex Area.
- Energy: It is the summation of squared elements in the GLCM and its value range between 0 and.
- Contrast: It is the measure of contrast between an intensity of pixel and its neighboring pixels over the whole ROI, where N is the number of different gray levels
- Homogeneity: It is the measure of closeness of the distribution of elements in the GLCM to the GLCM of each ROI and its Value ranges between 0 and 1.
- Correlation: It is the measure correlation of pixel to its neighbor over the ROI. Segmentation of Lung Tumor Using GLCM Techniques.
- Eccentricity: The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. After feature extraction, it then applies classification technique to predict the lung tumor as either benign or malignant.

Conclusion

The system proposed here has a promising results compared to other techniques. The proposed system is very efficient in segmentation principles for obtaining the feature extraction. In our research, the denoising of the image is done as part of the first phase and in the second phase; the lung region is differentiated from surrounding anatomy. This newly developed system is capable of classifying images with 96% accuracy on the training stage, and 85.5% accuracy on the testing stage. In next phase, once the Region of Interest (ROI) is extracted, features extraction is performed by GLCM.

References

1 Kusworo A, Catur E W, Arispuji W, Rahmat G, Adi P & Rizky, Detection of Lung Cancer Using Gray Level Co-Occurrence Matrix (GLCM) and Back Propagation Neural Network Classification, *J of Engg Sci & Tech Rev* 11 (**2**) (2018)

2 Kalaivani S, Pramit C, Shikhar J & Rishi G, Lung cancer detection using digital image processing and artificial neural networks, *Int Conf of Elect, Comm & Aero Tech Dec* (2017).

3 Pooja R & Anuradha T, Detection of Lung Cancer Stages using Image Processing and Data Classification Techniques, *Int Conf of Conv In Tech* (2017).

4 Munimanda P C, Venkateshwara M & Rajinikanth T, Detection of Lung Cancer Using Digital Image Processing Techniques: A Comparative Study, *Int J of Med Imag*, **5**(5) (2017).

5 Snehal D, Shubhangi C, Sneha J & Arjun N, A review paper on computer aided system for lung cancer detection, *Int Conf of Big Data, IoT & Data Sci* **2**(1) (2017).

6 Badrul A M & Mohammad A Y, Detection of lung cancer from CT image using image processing and neural network, *Int Conf On Elec. Engg & Info Comm Tech* (2015)