A HYBRID IMAGE PROCESSING METHOD FOR LUNG CANCER DETECTION USING WATERSHED AND GENETIC ALGORITHM

[1]Navjot Kaur, [2]Guresh Pal Singh, [3]Ritika Sood
[1]M.Tech Student, [2]Associate Professor, [3]Assistant Professor
[1]Computer Science & Engineering
[1]Beant College of Engineering & Technology, Gurdaspur, India

Abstract: In today's world cancer is most common disease that causes death and lung cancer is one of them. Lung cancer detection at early stages will increase the survival rate of the patient and also provide more option for treatment. In recent years if detection of cancer at early stages is done than the survival rate of patient is increased from 10 to 40%. The lung cancer is most common cause of death and dangerous disease, so faster method for detection required. The abnormal growth of cells present within the lungs can be detected using image processing technique. Image processing technique classifies the MRI images to distinguish the affected areas. In this paper a hybrid algorithm that utilizes watershed and genetic algorithm is utilized to detect lung cancer. For classification the decision tree is used that will enhance the detection rate. It will also increase the linear accuracy and recognition rate.

Keyword: lung cancer, watershed algorithm, genetic algorithm.

1. INTRODUCTION

The main cause of cancer death is Lung cancer and detection of cancer at early stage is difficult that cause high mortality rate. In world the number of people died due to lung cancer is high as compare to any other cancer like breast, skin and colon. So the early detection of lung cancer will decrease rate of deaths. Recently according to WHO around 6.8 million deaths caused due to lung cancer worldwide and also it continue to rise.[1] Increase of abnormal cells of lung is caused lung cancer. The tumour will increase the cells and form new one, also abnormal one. This will be detect at early stage so that proper treatment will provided on time, but many will notice this disease when it is too late and it is impossible to surgery. The detection of lung cancer is important at early stage is important for better treatment. [2]The MRI images are used for diagnosis and the tumour starts at lungs part then it is called primary lung cancer. There are two types of lung cancer

- Small cell cancer
- Non-small cell cancer

In our work we focused on detecting tumour and its stages using hybrid algorithm that contains watershed and genetic algorithm.

2. RELATED WORK

In [3]proposed a region growing algorithm that segment the CT scan images of lung. It starts with seed pixel and then continues on checking other pixels of its neighbourhood. It utilizes a criterion for similarity index and it is similar then it include region. Then this region is examined further.

In [4]proposed algorithm to detect cancer cells from lung MRI image and it minimize the error rate in detection. It utilizes sobel edge for detection and matrix for label the edges. It uses gradient to find the sobel edges. The change in intensity of image is given by image gradient. [5]Suggests an approach the use CAD (computer aided design) to detect the lung cancer by using edges of MRI images. [6]Proposes threshold algorithm that detect sputum cell from the image. In[7] proposes water shed transformation algorithm that segment the images. In this to process the image morphological operations are used and eliminate the over segmented area. After that gradient is reconstruct and the shape of image gradient is maintained.[8] In this paper we mainly focused on detecting tumour affected regions in image and also evaluate accuracy of the system. [8]proposed lung cancer mechanism using the MRI images. The detection mechanism used convolution neural network.[9] Proposed a mechanism to predict lung cancer using the fuzzy neural network based approach. results are presented in terms of classification accuracy.[10]proposed diffusion based watershed algorithm for tumour cell detection at early stage. The problem with this approach is limited disease detection.

3. TECHNIQUES FOR LUNG CANCER DETECTION

There are number of techniques which are associated with proposed system that can be used to detect lung cancer. These techniques are as described below

Watershed algorithm

This algorithm is primary used to detect and extract features out of the image data present. Point spread function is used to determine the abnormal section of the image.

Genetic algorithm

This algorithm is used to optimise the result produced. The produced result is compared against the fitness function and objective function. Selection, mutation and crossover are the critical steps in the formation of this algorithm.

Decision Tree

This algorithm is primarily used for the classification purpose. This mechanism uses If-Then else rules for the classification of the result.

3. Proposed Methodology

The image set is required to be operated upon by the modified watershed algorithm with classification. Feature extraction is performed by the use of Genetic algorithm. Rules are implemented using decision tree for classification purpose. The detailed description of proposed system is listed as under

4. DATASET DESCRIPTION

The dataset which is used is derived from internet. LUNG CANCER DATASET is used for extraction. LUNG CANCER DATASET contains 10 images of 40 to 44 KB in size. The images are grey scale in nature. Dimension of images is 1024×1024.

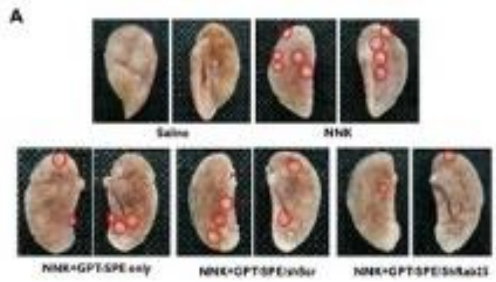

Figure 1: dataset

LUNG CANCER DATASET contains group of images with and without cancerous cell. Through the proposed system classification is generated to detect the disease and compared against the original image set.

APPLYING WATERSHED ALGORITHM FOR SEGMENTATION

Watershed is a transformation commonly applied on grey scale images. During the flooding operation adjacent catchment are constructed. The flooding process is performed on gradient images. The basins are constructed near the edges. It leads to over segmentation of images especially on noisy images. The algorithm is as follows

- Set of markers from where flooding should begin are selected.
- Priority is assigned to the neighbourhood of each marked pixel on the basis of gradient magnitude value selected. Label is assigned to each pixel inserted into priority queue.
- The pixel with least priority is extracted from priority queue and labels are checked.
- If labels are assigned and are same then same label pixels accept one are removed.
- The other pixels which are unlabelled are pushed into priority queue and steps are performed again.

After performing watershed algorithm, result is as follows

Figure 2: resulted image after applying watershed algorithm

The watershed algorithm is followed by genetic algorithm for feature extraction.

GENETIC ALGORITHM

Genetic algorithm is multi heuristic algorithm having multiple objectives associated with it. Genetic algorithm is associated with different phases. The pixels correspond to chromosomes. To perform feature extraction, selection operation takes place. The proposed system uses random selection operation. The fitness function evaluation is used to generate next population for feature extraction. The extracted features are compared against the threshold value. The threshold value is assumed to be base value above which optimality is achieved.

$Threshold = Optimal_{base-value}$

The fitness function evaluation takes place in order to obtain optimal results for classification.

$F(t) = Pixel_i(Features_j > Threshold)$

The mutation and crossover is performed only if threshold value is invalidated. Mutation and crossover is accomplished by identifying pixels having intensity values lower than desired levels. After which selection operation is performed again.

DECISION TREE IMPLEMENTATION

Decision tree is used for classification purpose. Classification of results required certain rules to be created. The Training rules correspond to features which are identified through genetic algorithm. The disease is detected if rules are violated. Decision tree rules are listed as under

If GM>0.5
If STD>0.5
If Kurtosis>0.5
If Moment>0.5
If Mean>0.5
Disease detected
Else
Disease not detected

Membership rules.

Maximum value 1 and minimum

A feature extracted from the segmented image is compared against these rules. Membership is decided only if feature lies within the range of 0 and 1. In order to detect the disease membership should be greater than 0.5.

5. PERFORMANCE ANALYSIS AND RESULTS

The performance of proposed system is analysed in terms of recognition rate and sigmoid function. Sigmoid function predicts the activation of training images. Training images matches against the testing images. Threshold must be satisfied in order to determine whether given image lie within the desired levels of features. If they do then they are classified as having given disease.

Recognition rate

Recognition rate indicates the success of simulation in the detection of problems within the given image set. For the success of proposed system recognition rate must be high.

Sigmoid function

A **sigmoid function** is a mathematical function having a characteristic "S"-shaped curve or **sigmoid curve**. Often, *sigmoid function* refers to the special case of the logistic function shown in the figure and defined by the formula

$$S = \frac{1}{1+e^{-x}}$$

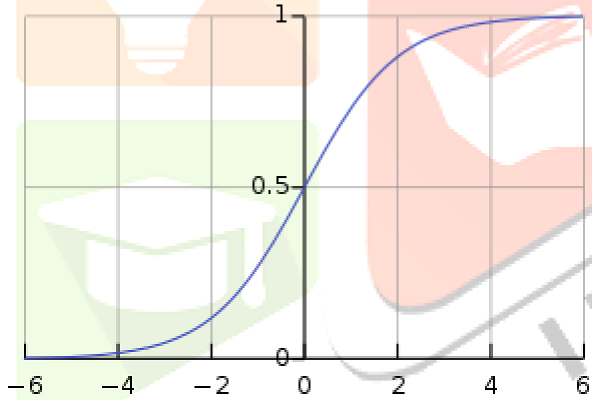

Figure 3: showing sigmoid function

Other examples of similar shapes include the Gompertz curve (used in modelling systems that saturate at large values of x) and the ogee curve (used in the spillway of some dams). Sigmoid functions have domain of all real numbers, with return value monotonically increasing most often from 0 to 1 or alternatively from −1 to 1, depending on convention.

A wide variety of sigmoid functions have been used as the activation function of artificial neurons, including the logistic and hyperbolic tangent functions. Sigmoid curves are also common in statistics as cumulative distribution functions (which go from 0 to 1), such as the integrals of the logistic distribution, the normal distribution, and Student's *t* probability density functions.

Snapshots

Snapshots corresponding to proposed system are as under

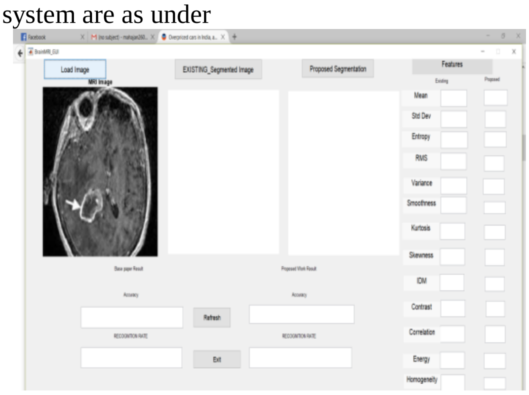

Figure 4: snapshot corresponding to load image

This screen demonstrates the selection of image when user clicks on the load button. The next screen demonstrates the existing system segmentation.

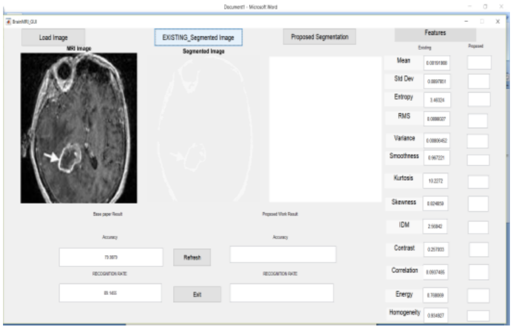

Figure 5: snapshot corresponding to existing segmented image

In this case benign cancer is detected. In other words classification will be of benign cancer

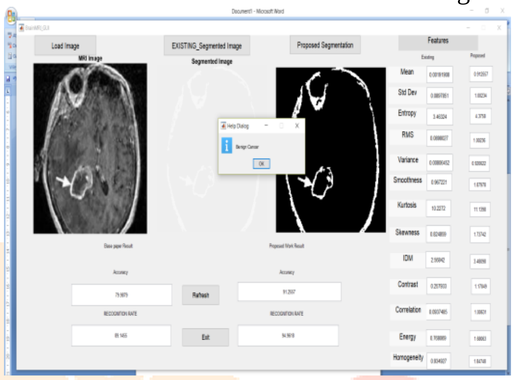

Figure 6: snapshot corresponding to proposed segmentation

In this case classification result gives benign cancer. The accuracy of proposed system along with recognition rate is high as compared to existing techniques.

Table 1: recognition rate

Result	Threshold=0.4
EXISTING	PROPOSED
79	91
81	92
79.879	90.23323
85	95.878
83	93.4343

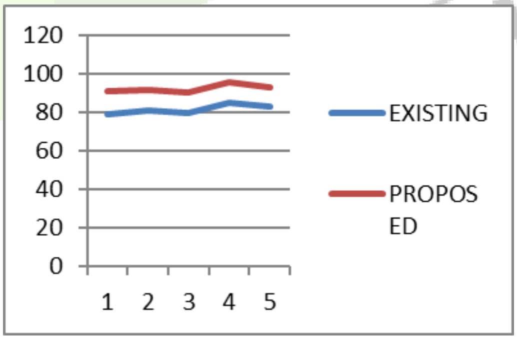

Figure 7: false matching rate (FMR)

Table 2: Sigmoid Function

EXISTING	Proposed
89	94
89.433	94.4334
90.2323	95.343
90.2324	95.787
91.1223	96.233434

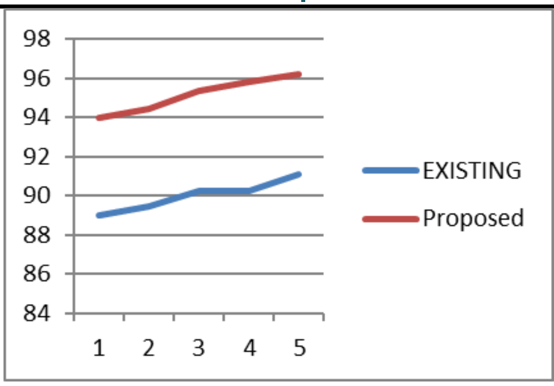

Figure 8: showing accuracy

6. CONCLUSION AND FUTURE SCOPE

The proposed scheme of thing produces efficient mechanism in order to determine the accuracy of disease detection using hybrid approach of Watershed and FSVM. Analysis process detects disease efficiently and handles the complex image with least complexity.

In future, detection process can be accomplished using LBP (linear binary pattern) with genetic algorithm.

References

[1] “SELECTED AUTOMOBILE COMPANIES AND ITS,” no. 10, 2014.

[2] N. Kureshi, S. S. R. Abidi, and C. Blouin, “A predictive model for personalized therapeutic interventions in non-small cell lung cancer,” *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 1, pp. 424–431, 2016.

[3] M. E. S. Processing, “Detection of Lung Cancer Using Marker-Controlled Watershed Transform,” vol. 00, no. c, 2015.

[4] F. Taher, N. Werghi, H. Al-ahmad, and R. Sammouda, “Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods,” vol. 2, no. 3, pp. 136–142, 2012.

[5] B. Abdillah, A. Bustamam, and D. Sarwinda, “Image processing based detection of lung cancer on CT scan images Image processing based detection of lung cancer on CT scan images,” 2017.

[6] I. Engineering, “Lung Cancer Recognition in CT Image Using,” pp. 3405–3409, 2016.

[7] C. Engineering, “Region of Interest Based Medical Image Compression,” pp. 1–12, 2014.

[8] P. Rao, N. A. Pereira, and R. Srinivasan, “Convolutional Neural Networks for Lung Cancer Screening in Computed Tomography (CT) Scans,” *IEEE Access*, pp. 489–493, 2016.

[9] S. H. Hawkins, J. N. Korecki, Y. Balagurunathan, Yuhua Gu, V. Kumar, S. Basu, L. O. Hall, D. B. Goldgof, R. A. Gatenby, and R. J. Gillies, “Predicting Outcomes of Nonsmall Cell Lung Cancer Using CT Image Features,” *IEEE Access*, vol. 2, pp. 1418–1426, 2014.

[10] D. D. Yi Yin, Oliver Sedlaczek, Benedikt Müller, Arne Warth, Margarita González-Vallinas, Bernd Lahrmann, Niels Grabe, Hans-Ulrich Kauczor, Kai Breuhahn, Irene E. Vignon-Clementel and ?, “Tumor cell load and heterogeneity estimation from diffusion-weighted MRI calibrated with histological data: an example from lung cancer,” *IEEE Trans. Med. Imaging*, vol. 0062, no. c, 2017.