International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 03 | June-2015 www.irjet.net p-ISSN: 2395-0072

REVIEW AND ANALYSIS OF FUSION MODEL FOR THE CLASSIFICATION OF LUNG CANCER DISEASE USING GENETIC ALGORITHM

Neha Sharma[1], Anil Kumar[2]

[1] *M.tech, CSE Department, GNDU Amritsar, Punjab, India*
[2] *Professor, Anil Kumar, GNDU Amritsar, Punjab, India*

Abstract - *Automatic categeorization of lung diseases in computed tomography (CT) images is an important diagnostic tool for computer-aided system of diagnosis . In this study, we define a new image based feature extraction technique for classification of lung CT images. A novel hybrid based method was obtained by combining the Gabor filter and Walsh Hadamard transform features using median absolute deviation (MAD) technique. Thus, it includes the advantages of combined models. The proposed system comprises of 3 stages. In the initial stage, the images are given as inputs and features are obtained by applying by novel fusion based feature wrenched method, followed by second stage, in which obtained features are selected by applying genetic algorithm which selects the top ranked features. In the final stage, categeorisers namely decision tree, memory based classification (KNN), Multi layer perceptron Neural Networks (MLP-NN) are employed to perform classification of the lung diseases. A total of 500 datasets for the diseases influenza, tuberculosis, pleural effusion and normal lung were used for training and testing. The classification precision of above 90% is accomplished by multilayer perceptron neural network classifier. The system has been trial outed with a number of real Computed Tomography lung images and has achieved satisfactory results in classifying the lung diseases.*

Key Words: Genetic algorithm, computer Tomography.

1. INTRODUCTION

From the recent years, medical CT Images have been used in medical field of diagnosis widely. In the recent years, medical CT Images have been applied in clinical diagnosis widely. It helps physicians to find and locate pathological changes with more precision. Computed tomography images can be distinguished for different tissues according to their different gray levels [1]. Lung diseases may be caused by infection, an outlet at the workplace, medications and various deffects. X-ray chest radiography and computer tomography (CT) are two common anatomic imaging models that are mostly used in the detection and diagnosis of a various lung diseases. Genetic algorithms (GAs) are among the most famous methods to do a feature selection. GAs has the ability to acquire either the exact or approximate solutions in very large search spaces within measureable time. GA carries out adaptive searching, succeeded by the standard concepts from natural genetics and evolution based on natural selection. Owing to their potential, GA is used in this work to perform selection. [3]

CT images of the lungs produced are with noise and no obvious changes in the grayscale boundary and other features, and are not susceptable to split. Therefore, the first thing to do is preprocessing of an image. The aim is to remove these places which are not productive to hew, which is to erase the noise of the lung CT images. Consequently, the boundary part of the CT image which changes smoothly should be sharpened in order to make possible the following work of segmentation.

The morphological smoothening and median filter techniques are used to separate the noise from the images and improve the image. The median filter will detach the salt and the pepper noises and produces the improved image. The erosion and dilation is called as Opening operation which vanquish the bright details diminish where dilation followed by erosion is closing operation which suppresses the dark details are calculated. In this work the darker details are overpower where the dilation is performed first and then it is followed by erosion. The preprocessed image is developed by using the median filter and morphological smoothening.

2. RELATED WORK

(Kalender)[1]An amazing overview of CT images and technology and applications .CT imaging had emerged far enough in respective of resolution and resolution to make it a liable tool in the trend of images of the lungs. A combination of pixel-based and knowledge-based algorithms for lung segmentation in CT images [4]. However, the maximum number of features actually describes a problem. It generates the " curse of dimensionality problem , where the indexing structures lowers and the significance of each feature diminishing, making the process of indexing and capturing and obtaining extremely time consuming. Pathologies are identified using automated CAD system. It assists the radiologist in study the digital images to bring out the possible outcomes of the diseases. The medical images

are acquired from different imaging systems such as CT scans etc.

(Hu and Brown)[2[[3], dynamic programming has been used to extract highest cost path, which matches to the junction line between the left and right lung. Hu etal. have used to generalize the threshold to segment lung regions instead of fixed threshold. Smoothing of the segmented lung areas is then accomplished in three separate stages to deal with three different issues: filling gaps caused by pulmonary vessels, deleting minute airgaps and removing large airgaps .

(Samuel) [4] has put up the process of tracking on the border to find boundaries at the regions. A fully automated approach for pulmonary segmentation is an important preprocessing step in a lung CAD system.

(Rushin Shojaii, Javad Alirezaie and Paul Babyn)[5] describes the different threshold adaptive values for the different slices of the ares of the disease and after that the pros and cons of the methods have been discussed in the detail.

(Hassan MR, Hossain MM, Bailey J, Ramamohanarao and Seiji Ishikawa)[6] describes proposed the various methods to discuss the complete study of classification of the diseases and all the affected areas and all the randomized techniques and their evolutionary strategies.

(Qiang Li, Shusuke Sone, Kunio Doi)[7] describes the slicing processing method of the lung CT images, using mathematical morphology, EM algorithm segmentation for nodules segmentation and extraction. Thus it provides great assistance for medical image practitioners in diagnosis to patients and improves the results of diagnosis.

(Moon, T.K.)[8]The growth rate of pulmonary nodules is a mark of differentiation between benign nodules and hypertentensed nodules.the strengths and weaknesses are discussed in detail. It can assist doctors by dividing the CT images nodules which are detected to find their volumes and to compute multiplication rate within a given frame of time.

3. GAPS IN EXISTING LITERATURE

The fusion method is studied in the classification of lung diseases using Genetic Algorithm. Various methods have been proposed and developed to find the solution in reasonable amount of time which are studied and discussed under literature review section. The gaps in study are identified and are discussed along with found drawbacks of existing methods as follows:

The major drawback of most of the existing heuristics is that they are evaluated with the use of modulated fusion method, less accuracy has been achieved in the process of developing the chromatographic images of the pulmonary lung disease.

During the classification of disease using decision trees and neural network method, scalability of the resultant images are not enhanced properly.

4. METHODOLOGY PROPOSAL

It defines how different gaps discussed above will be removed using the proposed algorithm using genetic operators.

The proposed genetic algorithm is given below:

PSUEDOCODE

[Step1] Begin: Generate random population of P solutions (chromosomes);

[Step2]For each individual i P: calculate fitness (i);

[Step3]For i=1 to number of generations; Randomly select an operation (crossover or mutation);

[step5]If crossover; Select two parents at random ia and ib; Generate on offspring ic = crossover (ia and ib);

[step6]Else If mutation; Select one chromosome i at random; Generate an offspring ic = mutate (i); End if;

[step7]Calculate the fitness of the offspring ic; If ic is better than the worst chromosome then replace the worst chromosome by ic;

[step8]Next i;

[step 9]Check if termination = true;

End;

5. CONCLUSIONS AND FUTURE SCOPE

In this work a novel fusion based feature extraction is proposed and feature selection is done by genetic algorithm that selects the high ranked characterstics and categeorization is performed through J48 , decision trees and neural network classifiers to classify the lung CT dataset. The algorithm has been developed based on the concept of texture and pixel co-efficient characterstics. This method greatly works well for the identification of lung diseases with high sensitivity, specificity and accuracy. Studies show that MLP NN classifier with median absolute deviation techniques and genetic algorithm for feature selection produces better results. This makes a point on the choice of using fusion with the genetic algorithm with the classifier. This approach has potential for future development because of this simplicity that will help to classify the types of lung diseases. The developed classification system is made to provide valuable diagnosis for the medical theory. The work

can further be carried by involving more feature extraction and selection methods for the classification of more lung diseases.

REFERENCES

[1]Kalender WA. Computed tomography: fundamentals, system technology, image quality, applications. 3rd ed. Erlangen: Publicis Publishing; 2011.
[2] S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images", IEEE Transactions on Medical Imaging Vol 20, No 6, June 2001.
[3] M. S. Brown, M. F. Mcnittgray, N. J. Mankovich, J. G. Goldin, J. Hiller, L. S. Wilson and D.R. Aberle, "Method for segmenting chest CT image data using an anatomical model – preliminary results", IEEE Trans Med Imaging. Vol. 16, pp. 828-839, Dec 1997.
[4]Samuel. G. Armato III and W. F. Sensakovic, "Automated lung segmentation for thoracic CT: impact on computer-aided diagnosis," In Academic Radiology, vol. 11, pp. 1011-1021, 2004.
[5]Rushin Shojaii, Javad Alirezaie, and Paul Babyn, "Automatic lung segmentation in ct images using wa-tershed transform," In Proceedings of the 8th IEEE In-ternational symposium on Computers and Communica-tion, vol. 2, pp. II-1270-3, 2005. Hassan MR, Hossain MM, Bailey J, Ramamohanarao
[6] Hyoungseop Kim, Tooru Nakashime, Yoshinori Itai, Shinya Maeda, Joo kooi Tan, and Seiji Ishikawa, "Automatic detection of ground glass opacity from the thoracic mdct images by using density features," In International conference on control, automation and sys-tems, pp. 1274-1277, 2007.
[7] K. Improving k-nearest neighbour classification with distance functions based on receiver operating characteristics. In: Machine learning and knowledge discovery in databases LNAI. Berlin Heidelberg: Springer; 2008. p. 489–504.
[8] Moon, T.K. The expectation-maximization algorithm. Signal Processing Magazine, IEEE, 1996/11, 13(6):47-60.
[9] Pugatch RD, Faling LJ. Computed tomography of the thorax: a status report. Chest 1981;80:618–26.
[10]Chiou YS, Lure YF, Freedman MT, Fritz S. Application of neural network based hybrid system for lung nodule detection. In: Computer-based medical systems, proceedings of sixth annual IEEE symposium, IEEE; 1993. p. 211–216.
[11] Golberg DE. Genetic algorithms in search, optimization and machine learning. Newyork: Addison Wesley;1989.
[12] Melanie M. An introduction to genetic algorithm. Bradford Book. London, England: The MIT Press Cambridge, Massachusetts; 1999, Fifth printing.
[13] Malcok M, Aslandogan Y, Yesildirek A. Fractal dimension and similarity search in high-dimensional spatial databases. In: IEEE international conference on information reuse and integration, Waikoloa, Hawaii, USA; 2006. p. 380–384.
[14] Speckmann EJ, Elger CE. Introduction to the neurophysiological basis of the EEG and DC potentials. In: Neidermeyer E, Lopes da Silva F, editors, Electroencephalography, 5th ed., Lippincott Williams & Wilkins, New York, 2005 [Chapter 2].
[15] Kanazawa K, Kawata Y, Niki N, Satoh H, Ohmatsu H, Kakinuma R, Kaneko M, Moriyama N, Eguchi K. Computeraided diagnosis for pulmonary nodules based on helical CT images. Comput Med Image Graph 2008;22(2):157–67.
[16] National Cancer Imaging Archive. Available at https://imaging.nci.nih.gov/. [17]Food and Drug Administration. Available at: http://www.fda.gov/oc/initiatives/criticalpath/reports/opp list.pdf,2006. [18]Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing(Second Edition)[M]. China Publishing House ofElectronics Industry, 2003: 420- 453.
[19]J. W. Turkey. Exploratory Data Analysis. Reading. MA: Addision-Wesley.1970.
[20]Dougherty, R. A. Lotufo. Hands-on Morphological Image Processing[M]. SPIE PRESS Bellingham, WA, 2003.
[21] Wang L.Enhancement of medical ultrasonic image based on gray level histogram equalization[J].J Sichuan Univ,2002,34(1):105-108.
[22]Haupt SE. Practical genetic algorithms. second ed. New Jersey, United States: John Wiley & Sons; 2004.

