I J C T A, 9(2) 2016, pp. 1325-1331

Enhanced False Positive Reduction of Lung Nodule Using Pixel Based Machine Learning and Fuzzy Min-Max Neural Network

S. Sowmya[*] and G. Niranjana[**]

ABSTRACT

Cancer is a standout amongst the most genuine wellbeing issue on the planet. The death rate of lung cancer disease is the most astounding among every other kind of tumor. Survival from lung disease is specifically identified with its development at its discovery time. Early recognition of lung malignancy is the most encouraging approach to upgrade quiet risk for survival. In this paper, the Computerized tomography (CT) is taken as data since it is utilized to conclusion condition influencing the mid-section its substance and close-by structure. In first stage, Image Pre-preprocessing from CT pictures is performed to improve picture quality and evacuate commotion. In second stage Pixel Based Machine Learning (PML) is a procedure which utilizes "pixel" in pictures specifically rather than components figured from portioned article and it is mostly used to characterize the injuries whether it's benevolent or maligant.In third stage Fuzzy Min- max Neural Network is used to reduce False Positive Reduction Scheme.

Keywords: Lung Cancer, computerized tomography, pixel based machine learning, fuzzy min-max.

1. INTRODUCTION

Cancer is an essential base of grown-up death around the world. In India, the "International Agency for Research on Cancer" (IARC) figured in a roundabout way that around 635000 individuals lost their life from growth in 2008, approx. 8% of all evaluated overall tumor passings and around 6% of all passings in India. The genuine number of disease passings in India is liable to increment as expansion the populace and rates of tumor passings are required to rise. There are a few sorts of tumor; especially age-particular disease ascends because of utilizing tobacco smoke. Along these lines, social and topographical circulation identifying with lung growth is varying in India. It is important to control and analyze at early stage to spare from death all around [1]. The structure of the tumor cells are firmly associated and are a testing issue to distinguish lung growth at early stage since manifestations turn out just at later stages bringing about the death rate to be the most noteworthy among every single other sort of malignancy [2]. Thus, early recognition and early analysis can impressively enhance the survival rate [3]. Lung disease quiet bites the dust in extensive number than other malignancy like bosom, colon, and prostate growths. As showed by the most recent insights gave by whom, around 7.6 million passings overall occur every year. Besides, tumor passings are anticipate to continue expanding, to end up around 17 million worldwide in 2030.

Conclusion and Prognosis are therapeutic terms identified with clutters. Finding alludes to recognize a conceivable ailment includes grouping tests. Anticipation must be done after an analysis is made of particular state. The anticipation predicts the passing in a couple short months or weeks. There are a few methods to analyze lung growth, for example, "Chest Radiograph (X-Ray), Computed Tomography (CT), Magnetic Resonance Imaging (MRI output) and Sputum Cytology". Machine Learning and Data Mining are the instruments for effective grouping, i.e. "Simulated Neural Network, Decision tree, Bayesian Network,

[*] PG Scholar, Dept of CSE, SRM University, Chennai, *E-mail: sowmiyasenthil1@gmail.com*

[**] Assistant Professor(SG), Dept of CSE, SRM University, Chennai, *E-mail: niranjana.g@ktrsrmuniv.ac.in*

Association Rule Mining, Clustering, Classification. Artificial Neural Networks" (ANN) has been utilized to perform complex calculations in a few applications for effective discovery of tumors. Such calculations are finished by processing hubs (neurons) and weighted associations, which cooperate with all information and inputs. Neurons for the most part work in parallel and sorted out in layers, and criticism associations both inside the layer and toward contiguous layers.

Machine learning (ML) assumes a vital part in the medicinal imaging ûeld, including restorative picture investigation and computer-aided diagnosis (CAD). A standout amongst the most well known employments of ML in therapeutic picture examination is the classiûcation of articles, for example, sores into specific classes (e.g., strange or typical, sores or non-injuries, and threatening or favorable) in light of info components (e.g., differentiation, range, and circularity) got from sectioned item hopefuls (This class of ML is alluded to include based ML). The undertaking of ML here is to decide "ideal" limits for isolating classes in the multidimensional element space which is shaped by the info highlights [4]. ML calculations for classiûcation incorporate direct discriminant investigation, quadratic discriminant examination, multilayer perceptron, and bolster vector machines. Such ML calculations were connected to lung knob discovery in mid-section radiography and thoracic CT, classiûcation of lung knobs into benevolent or harmful in mid-section radiography and thoracic CT, location of microcalciûcations in mammography, recognition of masses in mammography, classiûcation of masses into considerate or threatening in mammography, polyp identification in CT colonography, deciding subjective comparability measure of mammographic pictures, and identification of aneurysms in cerebrum MRI.

2. LITERATURE REVIEW

Elizabeth *et al.*, [5] in their work have introduced a PC helped determination framework fit for selecting a critical cut for the examination of every knob from an arrangement of cuts of a computed tomography (CT) filter in digital imaging and communications in medicine (DICOM) group has been created for the conclusion of lung growth. Here RBFNN (Radial Basis Function Neural Network) is utilized. An exactness of 94.44% has been accomplished in arranging the knobs as dangerous and non-harmful. Lessens the computational unpredictability of the CAD framework. The seriousness and rate of development of lung malignancy is not recognized.

Pu *et al.*, [6] utilized a marked separation field as a part of the entire arrangement of pictures and identified the maximums found as potential knobs, which were then positioned by separation to the average hub, got by method for the grouping procedure and the use of walking blocks calculation. An aggregate of 52 tests were performed on a restrictive base that contained 184 lung knobs and whose outcomes were 95.1% affectability and a mean of 1200 suspect voxels per exam. Fiebich et.al, [7] utilized low dosage and dynamic shape pictures too, consolidated with a multi-thresholding calculation, beginning at focuses that had greatest similarity to a circle and got from the auto values in each voxel. The specialists figured out how to distinguish 58% of the knobs in this base, with a rate of 1.38 false positives.

Lee *et al.*, [8] built up another methodology with respect to programmed location of kind knobs. They utilized hereditary calculation construct layout coordinating procedure with respect to CT pictures. Kanazawa *et al.*, [9] proposed a fluffy bunch based CAD framework for the ID of pneumonic knobs. Biradar *et al.*, [10] planned a CAD framework to recognize kind lung knobs by utilizing CT pictures. They utilized the extraction of districts of hobby and essential picture preparing strategies. Choi *et al.*, [11] proposed a CAD framework to naturally arrange lung knobs. Besides, there are different ANN-based CAD in writing. Suzuki *et al.*, [12] proposed an example acknowledgment system taking into account ANN utilizing lowdose CT pictures for diminishment of false encouraging points in mechanized location of lung knobs. Coppini et al. [13] exhibited a neural-system based framework for the PC supported identification of lung knobs in mid-section radiogram.

3. PROPOSED SYSTEM MODEL

The proposed system architecture is given in the figure 3.1. It consists of the following parts:1) Image preprocessing, 2) Feature Extraction and Selection, 3) Pixel based Machine Learning, 4) Fuzzy min-max neural Network.

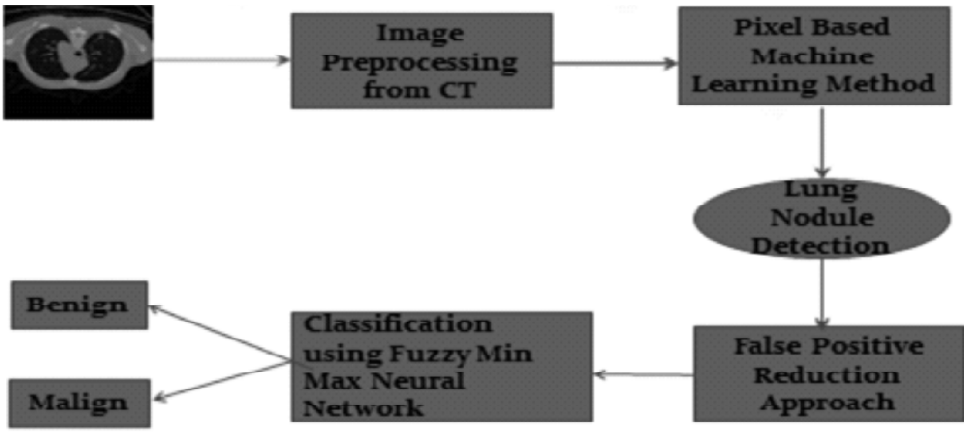


Figure 3.1: Proposed System Architecture

3.1. Image Preprocessing

It is utilized to enhance the picture information that stifles undesirable mutilation or upgrades some picture highlight vital for further handling. Preprocessing from CT pictures is performed to improve picture quality and expel commotion utilizing channel.

The algorithms used for preprocessing are: (i) Median filter and (ii) Sobel edge detector.

(i) Median filter: It is basic and capable non –linear channel which is based request measurements. It is anything but difficult to execute strategy for smoothing pictures.

It is utilized for decreasing the measure of power variety between one pixel and the other pixel. In this channel, we don't supplant the pixel estimation of picture with the mean of all neighboring pixel values; we supplant it with the middle worth. At that point the middle is figured by first sorting every one of the pixels values into rising request and afterward supplant the pixel being ascertained with the center pixel esteem. The median filter gives the best result of removing noise.

(ii) Sobel Edge detector: It is the procedure of finding the edge of a picture which is vital in finding the rough total slope at every point I of an information dark scale picture. The Sobel edge finder utilizes a couple of 3*3 convolution veil, one assessing angle in X course and other evaluating inclination in y bearing. The sobel finder is unimaginably delicate to commotion in pictures if successfully highlight them as edges. Subsequently, sobel administrator is suggested in enormous information correspondence found in information exchange.

3.2. Feature Extraction and Selection

The locales of hobby were extricated to separate kind/defame tumors in lung CT pictures. The separation of tumors can be performed by the assistance of measurable and shape components of tumors. At the point when the info information to a calculation is too vast to ever be prepared and it is suspected to be repetitive (e.g. the dullness of pictures exhibited as pixels), then it can be changed into a decreased arrangement of components. This process is called feature extraction.

The following algorithm is used for feature extraction and selection.

(i) Gray-Level Co-occurrence matrix: The Gray-Level Co-event Matrix (GLCM) relies on upon the extraction of a dim scale picture. It considers the relationship between two neighboring pixels, the primary pixel is known as a wellspring of point of view and the second is known as a neighbor pixel. The GLCM is a square grid with Ng measurement, where Ng approaches the quantity of dim levels in the picture. Every component of the grid is the quantity of event of the pair of pixel with worth i and a pixels with quality j which are at separation d.

3.3. Pixel based Machine Learning

It utilizes "Pixel" in pictures specifically rather than elements ascertained from the sectioned picture.The main purpose of PML is used to classify the nodules whether it is normal lesion or abnormal lesion.It obtains the image "Pixel by Pixel" from the given input information.

The following algorithm used for pixel based machine learning:

(i) Massive Training Artificial Neural Network (MTANN)

We are developing the neural channels and the neural edge locator to oblige different picture handling and example acknowledgment errands, and we should call this procedure a massive training artificial neural network (MTANN). The engineering and the preparation strategy for the MTANN are appeared in Fig. 3.2. The MTANN comprises of a changed multilayer ANN, which can straightforwardly handle information dark levels and yield dim levels. In the MTANN, picture preparing or design acknowledgment is performed by checking of a picture with the adjusted ANN in which the enactment elements of the units in the info, covered up, and yield layers are a straight, a sigmoid, and a direct capacity, separately. The MTANN utilizes a direct capacity as the initiation capacity of the unit in the yield layer on the grounds that the qualities of an ANN were altogether enhanced with a straight capacity when connected to the consistent mapping of qualities in picture preparing, 35, 36, 32.

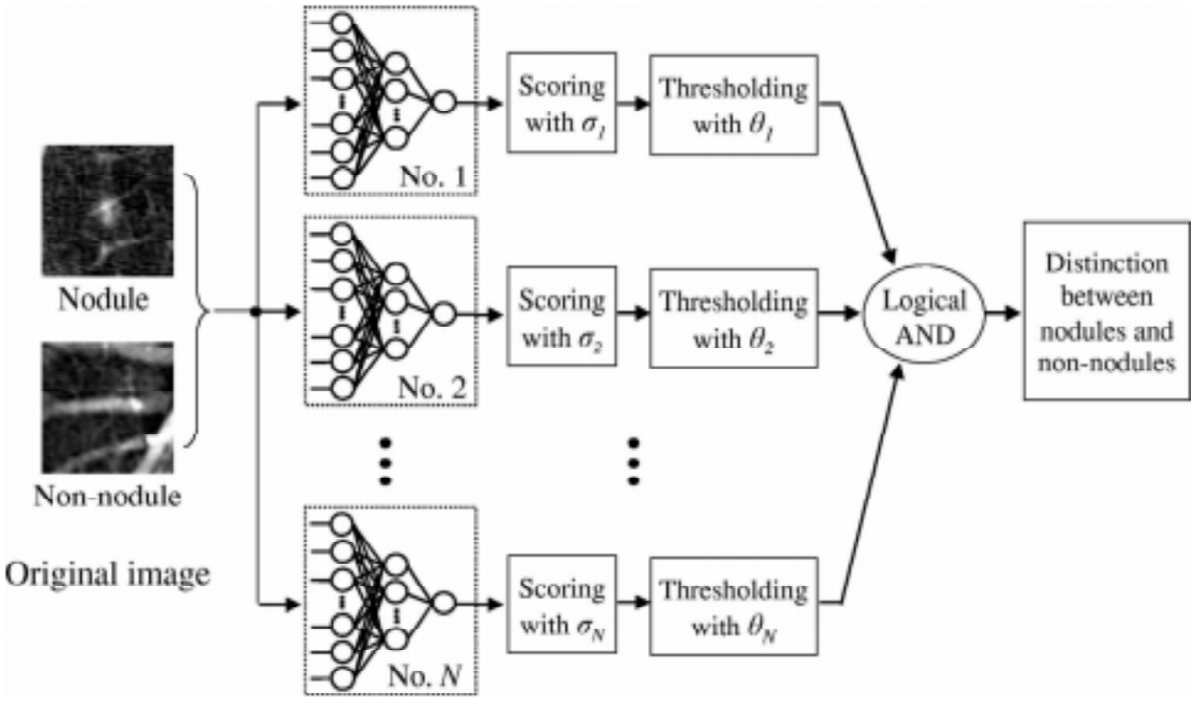


Figure 3.2: MTANN Architecture

3.4. Fuzzy min-max neural Network

It is mainly used to reduce False Positive Reduction scheme. The Fuzzy min-max neural network was advanced by fuzzy K-means clustering for false positive reduction.

The algorithm used for fuzzy min-max neural network is: Fuzzy k-Means algorithm.

(i) Fuzzy K means algorithm: The k-means calculation is a calculation to bunch n objects in view of properties into k segments, where k < n. It is like the desire boost calculation for blends of Gaussians in that they both endeavor to discover the focuses of common groups in the information. It accept that the article qualities shape a vector space.

A calculation for apportioning (or bunching) N information focuses into K disjoint subsets Sj containing information guides so as toward minimize the aggregate of-squares paradigm.

$$J = \sum_{j=1}^{k} \sum_{n \in S_j} \left| x_n - \mu_j \right|^2$$

Where,

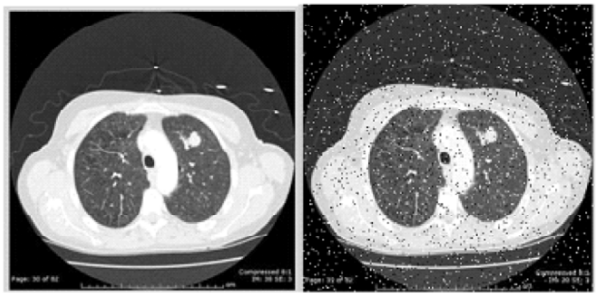
x_n is a vector representing the nth data point and

μ_j is the geometric centroid of the data points in Sj.

4. EXPERIMENTAL RESULTS

The execution assessment of our proposed framework was performed with MATLAB programming. All analyses were performed by utilizing a PC with 3.4 GHz i7 processor, 8 GB memory and Windows 7 working framework.

Figure 4 represents CT pictures for the yields of methodology in the proposed framework. Figure 4(a) displays the original lung image, Figure 4(b) is the noisy image (salt and pepper noise) and Figure 4(c) shows the filtered image.



4(a) 4(b)

4(c)

Figure 4: (a) Input Image, 4(b) Noisy Image, 4(c) Filtered image

The following figure 4.2 shows the sobel edge detection of the lung image. Here, the Nodule part is detected and separated by using edge detection technique.

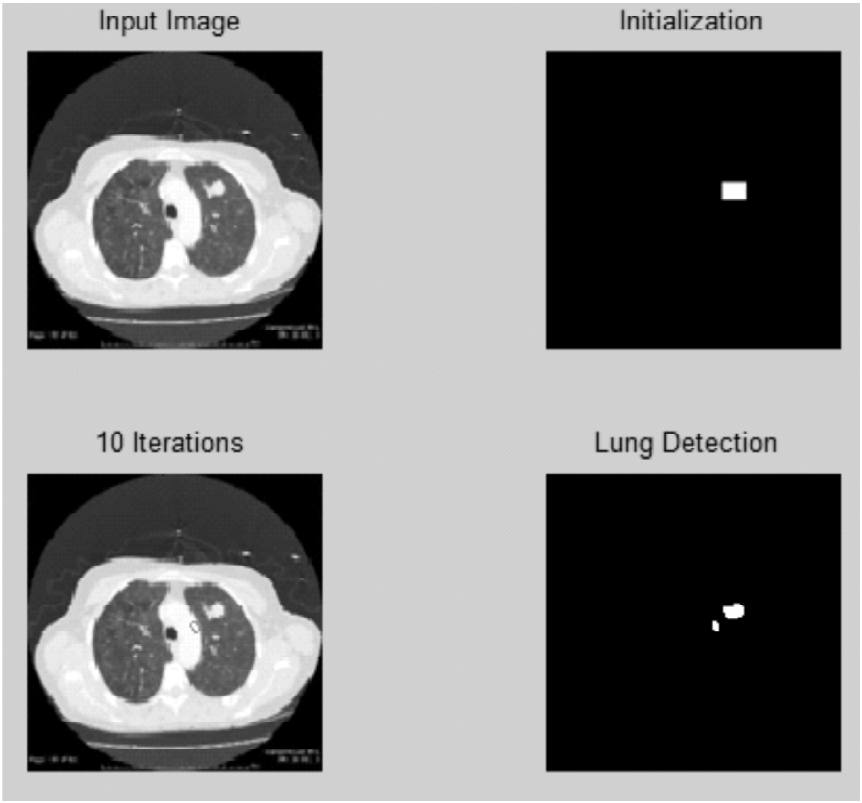


Figure 4.2: Edge detection Processes

The following figure 4.3 shows the regression of our system.

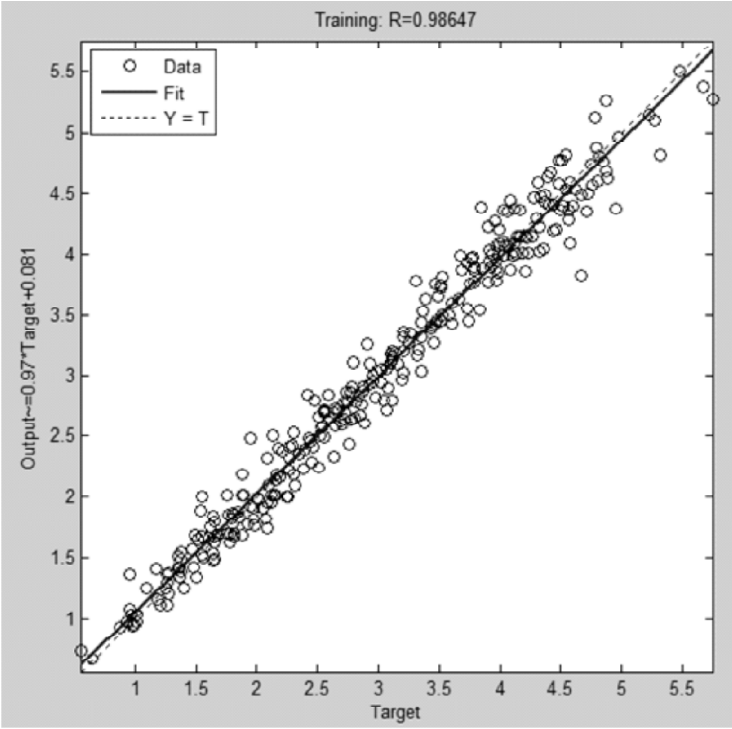


Figure 4.3: Regression of our system

5. CONCLUSION

In this paper, we clarified about the Lung Cancer discovery after connected a few techniques and a give an inventive path for Lung Cancer identification utilizing "Fuzzy Min-Max Neural Network and Fuzzy C Mean". The order techniques are connected to both FMN and FCM on the X-beam 130 carcinogenic and noncancerous datasets accessible. Lung Cancer X-beam Image Datasets can be utilized as a part without bounds to recognize the best results. This work exhibited a technique for programmed discovery of lung knobs. The stages included, from acquiring exams to showing arranged knobs, and additionally the outcomes got, exhibit this is a promising system.

REFERENCES

[1] Ankit Agrawal and AlokChoudhary, (2011), "Identifying HotSpots in Lung Cancer Data Using Association Rule Mining", 11th IEEE International Conference on Data Mining Workshops, 2011.

[2] Vinod Kumar, Dr. Kanwal Garg, (2012), "Pattern Identification by Image Processing for Early Detection of Lung Cancer", Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra, June-2012.

[3] Vinod Kumar, Dr. Kanwal Garg, (2012), "Neural Network Based Approach for Detection of Abnormal Regions of Lung Cancer in X-Ray Image", International Journal of Engineering Research & Technology, ISSN: 2278-0181, IJERT Vol. 1 Issue 5, July 2012.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, (2001),"Pattern Recognition", Wiley Interscience, Hoboken, NJ, USA, 2nd edition, 2001.

[5] D. S. Elizabeth, H.K.Nehemiah, (2011), "Computer-aided diagnosis of lung cancer based on analysis of the significant slice of chest tomography image", IET Dec. 2011.

[6] J. Pu, B. Zheng, J.K. Leader, X.-H. Wang, D. Gur, (2008), "An automated CT based lung nodule detection scheme using geometric analysis of signed distance field", Medical Physics, Vol. 35, No. 8,pp. 3453–3461.

[7] C. Schneider, A. Amjadi, A. Richter, M. Fiebich, (2009),"Automated lung nodule detection and segmentation," in Proceeding of the SPIE, vol. 7260, 2009.

[8] Y. Lee, T. Hara, H. Fujita, S. Itoh, T. Ishigaki, (2001), "Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," IEEE Transactions on Medical Imaging, 20(7), July 2001.

[9] K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyama ve K. Eguchi, (1998), "Computeraided diagnosis for pulmonary nodules based on helical CT images", Computer Medical Image Graph., Vol. 22, No. 2, pp. 157-167, 1998.

[10] V. Biradar, U. Patil, (2013), "Computer Aided Detection (CAD) System for Automatic Pulmonary Nodule Detection in Lungs in CT Scans", The International Journal of Engineering and Science (IJES), Vol. 2, No.1,pp. 18-21, 2013.

[11] W.-J. Choi ve T.-S. Choi, (2013), "Automated Pulmonary Nodule Detection System in Computed Tomography Images: A Hierarchical Block Classification Approach", entropy, Vol. 15, pp. 507-523, 2013.

[12] K. Suzuki, S. G. Armato, F. Li, S. Sone, K. Doi, (2003), "Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography", Medical Physics, Vol. 30, 1602, 2003.

[13] G. Coppini, S. Diciotti, M. Falchini, N. Villari, G. Valli, (2003),"Neural Networks for Computer-Aided Diagnosis: Detection of Lung Nodules in Chest Radiograms", IEEE Transactions on Information Technology in Biomedicine, vol. 7, No.4, pp. 344-357, 2003.