Lung Cancer Detection Model UsingConvolution Neural Network and Fuzzy Clustering Algorithms

P. Prakashbabu[1], D. Ashok Kumar[2], T.Vithyaa[3]
Department of Computer Science, Government Arts College, Kulithalai, Karur ,Tamilnadu - 639 120, India.

Abstract: *This paper discusses the formation of Lung cancer detection system by using the techniques of Image processing. The system formed can take any type of medical image within the three choices consisting of CT, MRI and Ultrasound images. Here the proposed model is developed using Fuzzy-C-Means and Convolution Neural Network (CNN) algorithm used for feature selection. This paper is an extension of image processing using lung cancer detection and produces the results of feature extraction and feature selection after segmentation. The system formed accepts any one of medical image within the three choices consisting of MRI, CT and Ultrasound image as input. After preprocessing of image, wiener filter is used for remove noise and unwanted region. This present work proposes a method to detect the cancerous cells effectively from the CT, MRI scan and Ultrasound images. Pixel Segmentation has been used for FCM segmentation and filter is used for De-noising the medical images. Simulation results are obtained for the cancer detection system using MATLAB and comparison is done between normal lung and abnormallung medical images.*

Keyword : *Lung cancer,lung segmentation,Fuzzy-C-Means,CNN, Feature extraction.*

I. INTRODUCTION

One of the major reasonsfor non-accidental death is cancer. It has been proved that lung cancer is the topmost cause of cancer death in men and women worldwide. The death rate can be reduced if people go for early diagnosis so that suitable treatment can be administered by the clinicians within specified time. Cancer is, when a group of cells go irregular growth uncontrollably and lose balance to form malignant tumors which invades surrounding tissues. Cancer can be classified as Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC).In this paper we confine to Non-small cell lung cancer (NSCLC) as it is more prevalent than small cell lung cancer (SCLC). There's a difference between the diagnosis and treatment of non-small cell and small cell lung cancer. The various ways to detect lung cancer is by the use of image processing, fuzzy c-means and convolutional neutral network to develop Computer aided diagnosis. In this paper, CT scan image, MRI scan image and ultrasound images are used. A CT scan or Computerized Axial Tomography (CAT) scan is the most sensitive and specific detection modality produces cross-sectional images of specific areas of scanned object by the use of computer processed combination of many X-ray images taken from different angle [1]. Radio waves and magnetic field is used to form images of a body in an imaging technique known as Nuclear Magnetic Resonance Imaging (NMRI) The aim of this paper is to design a system which can take any one of the three images as input and produces the desired output. The algorithms used are sensitivity,specificity and accuracy.The proposed model consists of following steps such as: Collection of lung image data set, preprocessing, wiener filter and FCM segmentation of CT and MRI images. Every step is described in further sections.

We apply an extensive preprocessing techniques to get the accurate nodules in order to enhance the accuracy of detection of lung cancer. Moreover, we perform an end-to-end training of CNN from scratch in order to realize the full potential of the neural network i.e. to learn discriminative features. Extensive experimental evaluations are performed on a dataset comprising lung nodules from more than 1390 low dose CT scans [2].

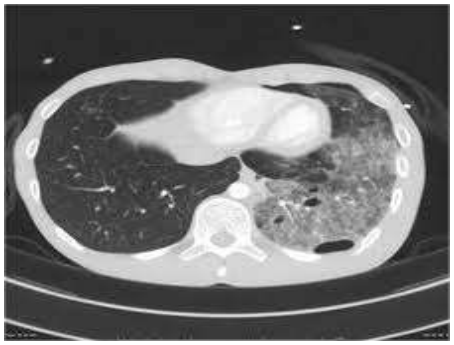

Figure 1: CT scan slice containing a small early stage lung cancer nodule.

II. REVIEW OF THE LITERATURE

Recently, deep artificial neural networks have been applied in many applications in pattern recognition and machine learning, especially, Convolutional neural networks (CNNs) which is one class of models [3]. Another approach of CNNs was applied on ImageNet Classification in 2012 is called an ensemble CNNs which outperformed the best results which were popular in the computer vision community [4]. There has also been popular latest research in the area of medical imaging using deep learning with promising results.

H. Suk et al. [5] suggested a new latent and shared feature representation of neuro-imaging data of brain using Deep Boltzmann Machine (DBM) for AD/MCI diagnosis. G. Wu et al. [6] developed deep feature learning for deformable registration of brain MR images to improve image registration by using deep features. Y. Xu et al. [7] presented the effectiveness of using deep neural networks (DNNs) for feature extraction in medical image analysis as a supervised approach. Kumar et al. [8] proposed a CAD system which uses deep features extracted from an autoencoder to classify lung nodules as either malignant or benign on LIDC database. In [9], Yaniv et al. presented a system for medical application of chest pathology detection in x-rays which uses convolutional neural networks that are learned from a non-medical archive. that work showed a combination of deep learning (Decaf) and PiCodes features achieves the best performance. The proposed combination presented the feasibility of detecting pathology in chest x-ray using deep learning approaches based on nonmedical learning. The used database was composed of 93 images. They obtained an area under curve (AUC) of 0.93 for Right Pleural Effusion detection, 0.89 for Enlarged heart detection and 0.79 for classification between healthy and abnormal chest x-ray.

In [10], Suna W. et al., implemented three different deep learning algorithms, Convolutional Neural Network (CNN), Deep Belief Networks (DBNs), Stacked DenoisingAutoencoder (SDAE), and compared them with the traditional image feature based CAD system. The CNN architecture contains eight layers of convolutional and pooling layers, interchangeably. For the traditional compared to algorithm, there were about 35 extracted texture and morphological features. These features were fed to the kernel based support vector machine (SVM) for training and classification. The resulted accuracy for the CNN approach reached 0.7976 which was little higher than the traditional SVM, with 0.7940. They used the Lung Image Database Consortium and Image Database Resource Initiative (LIDC/IDRI) public databases, with about 1018 lung cases.

In [11], J. Tan et al. designed a framework that detected lung nodules, then reduced the false positive for the detected nodules based on Deep neural network and Convolutional Neural Network. The CNN has four convolutional layers and four pooling layers. The filter was of depth 32 and size 3,5. The used dataset was acquired from the LIDC-IDRI for about 85 patients. The resulted sensitivity was of 0.82. The False positive reduction gotten by DNN was 0.329.

In [12], R. Golan proposed a framework that train the weights of the CNN by a back propagation to detect lung nodules in the CT image sub-volumes. This system achieved sensitivity of 78.9% with 20 false positives, while 71.2% with 10 FPs per scan, on lung nodules that have been annotated by all four radiologists.Convolutional neural networks have achieved better than Deep Belief Networks in current studies on benchmark computer vision datasets. The CNNs have attracted considerable interest in machine learning since they have strong representation ability in learning useful features from input data in recent years.

Fuzzy k-c-means clustering algorithm used for medical image segmentation which was introduced in Ajala, 2012 [13]. Here fuzzy-c-means is a method of clustering algorithm which allows one piece of data belongs to two or more clusters and k-means is a simple clustering method in which we use low computational complexity as compared to fuzzy c-means. When both Clustering methods were combined to produce a more time efficient segmentation algorithm called as fuzzy-k-c-means clustering algorithm. They offered that thresholding which is the most elementary technique for medical image segmentation, in which this algorithm divides pixels in different classes depending upon their gray level. It is also said that it approaches division of scalar images by forming a binary partition of the intensity values of an image and lastly determines an intensity value. This intensity value is termed as threshold, which separates the desired classes. Classifier techniques which were used for pattern recognition, partitions a feature space derived from the image using data with known labels. A feature space is a set of N*M matrix where N relates to the number of observations and M relates to the number of attributes. Classifiers are known as supervised methods since they require training data which are manually segmented and then used it for automatically segmenting new data.

In [14], Fatma, 2012 two more segmentation methods were used which were Hopfield Neural Network (HNN), and Fuzzy C-Mean (FCM) clustering algorithm. In this they found that the HNN provides enhanced, accurate and reliable

segmentation results than FCM clustering in all cases. The HNN also divides the nuclei and cytoplasm regions while FCM failed in the detection of the nuclei. FCM only detected a part of the nucleus not the whole nucleus in a particular cell. Also FCM was not found subtle to intensity variations because the segmentation error at convergence was found larger with FCM in comparison to HNN. According to the utmost latest estimates of the statistics which are provided by world health organization indicates that there happened around 7.6 million deaths worldwide each year because of this type of cancer. Moreover, they also found that mortality from cancer are estimated to rise continuously, and will come near to 17 million deaths worldwide in 2020. So, better methods are required to extract the nucleus region for very early detection. A magazine in (IEEE, Pulse) provided us the knowledge about current trends in medical image analysis.

In [15] Mokhled, 2012 first images which were improved through Gabor filter. It has given better results than other enhancement techniques. They only worked on colored image enhancement and not extract the nucleus region and even not the cell region. In Features Extraction stage they acquire the general features of the enhanced and segmented image which later they used in Binarization. A refined Charged Fluid Model (CFM) along with improved Otsu's method was used for the automatic segmentation of MRI images.

In [16] Sajith, 2012 glandular cells were detected by using multiple color spaces and two clustering algorithms which were K-means and Fuzzy C-means.

A novel lung segmentation technique was proposed by Lin-Yu-Tseng et al to improve segmentation accuracy as well as to separate and eradicate the trachea from lungs [17]. Anita Choudhary et al used Digital Image Processing Techniques to achieve more quality and accuracy [18]. AzianAzamini Abdullah et al described the development of an algorithm that detects symptoms of lung cancer in X-ray films by CNN (Cellular Neural Network) templates simulation [19].

III. METHODOLOGY

In this method the lung cancer is detected and forcast from CT image using dynamic particle swarm optimization method. The processing techniques of proposed method are shown in fig 2.

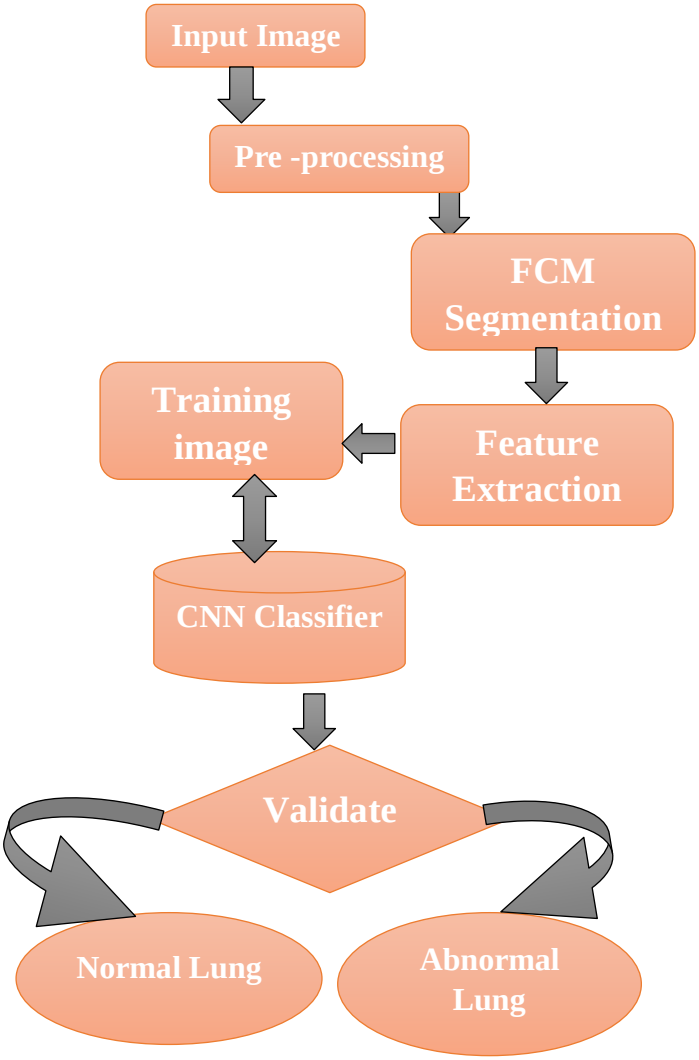

Figure 2: Proposed Model.

At first the CT image of lung cancer is read from the data base. Usually the acquired image contains low noise and if the noise is removed directly then is a chance of losing clarity so the included noise is removed by using processing technique. Median filter which is a nonlinear digital filter where it is used to reduce the noise. The enhanced image is them processed through the segmentation by FCM algorithm.The extracted image is then given to the classifier known as CNN which classify whether the lung nodule.

A. Input Image.

Image acquisition is a process of acquiring a digital image from data base. Generally the images are acquired by different types of scanners like MRI and CT. CT image is acquired from CT scanner. Computed Tomography (CT) is an imaging procedure that generates cross sectional images signifying in each pixel. This scan is a Non-Invasive and painless diagnostic tool. It also referred as CAT (computerized axial tomography).

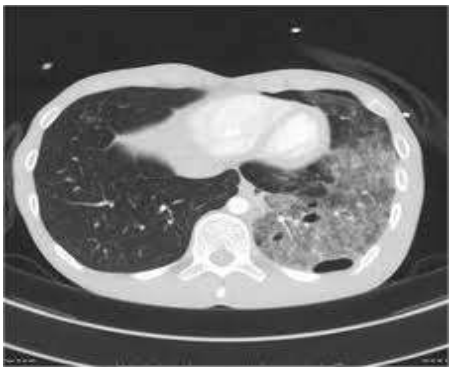

Figure 3: Input lung cancer image.

The input CT scan lung cancer image as shown in figure 3.

B. Preprocessing.

The images are subjected to pre-processing steps to remove noise and unwanted region. First, get the input image. Resize the image to the size acceptable to the processing system. Convert resized image into gray image in order to use only one color channel. Gray- scale comparison involves simple algebraic scalar operators. Gray scale image is enough to distinguish peaks of intensity. After converting the gray scale into binary image. That binary image is a digital image for each pixel with two possible values. The nextthing after acquiring an image is to redimension it. Because each image has different sizes so we can resize it with the same size. They convert it to a gray scale image after resizing an input image.

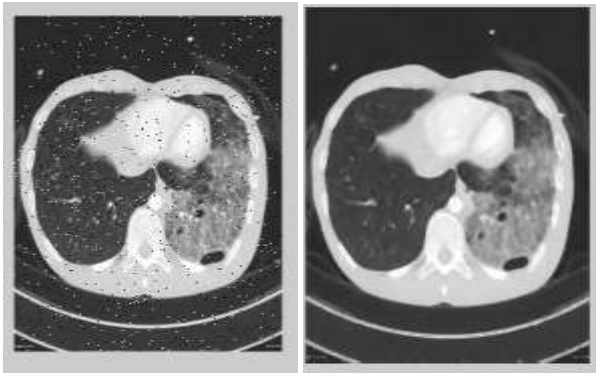

Figure 4: Noisy image.

Figure5: Wiener Filter Image.

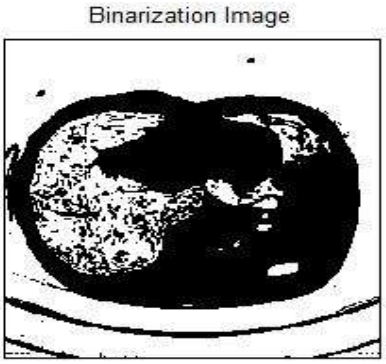

Figure6: Binarization image.

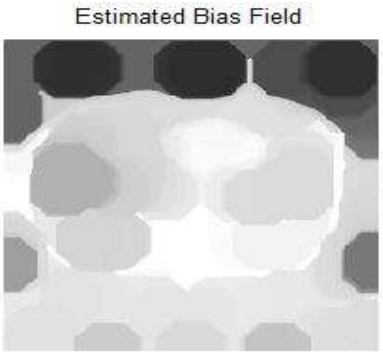

Figure 7: Estimated Bias Field.

Input image noise and unwanted region as show in figure 4, To remove noise and unwanted region, the input image is processed though filtered as show in figure 5, Binarization image as shown in figure 6 and Estimate bias fild image as show in figure 7.

C. Segmentation

It presents an automatic graph cut-based segmentation framework that uses a distance-constrained energy function to produce topologically restricted solutions. This term ensures that labels are assigned only to the lung pixels even in the presence of other anatomical regions with similar lung-like patterns. The Euclidean distance was specified to make it clear that the distance referred to in this work is the distance between two points, not the distance as a measure of the difference between two regions. Any metric can therefore be used to measure the distance between points or regions. The contribution of this work is to create an automatic method of lung segmentation using Graph Cut that produces topographically restricted solutions to accurately identify the lungs in a CT image.

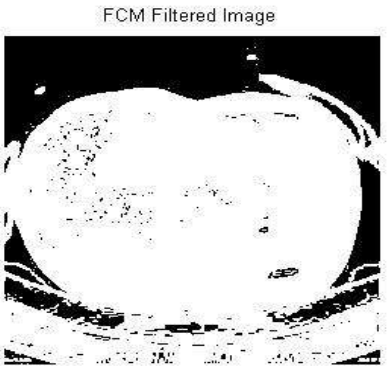

Figure8: FCM filter.

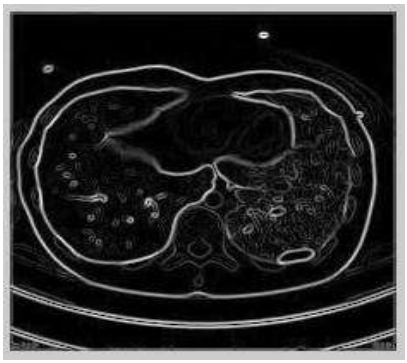

Figure9: Gradient magnitude.

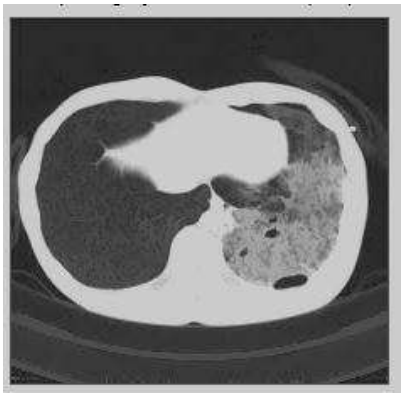

Figure 10: Opening – closing by reconstruction.

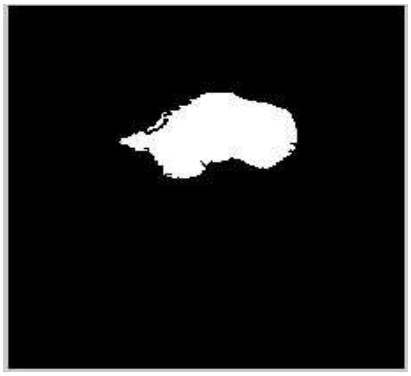

Figure11: Regional maxima of opening-closing by reconstruction.

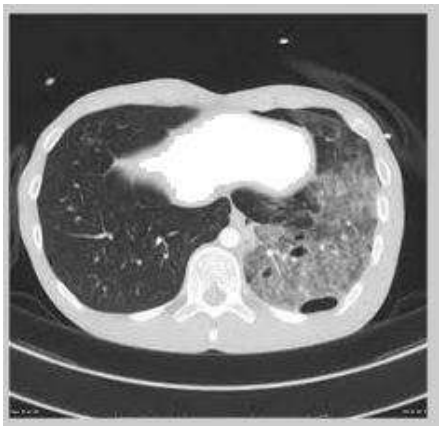

Figure12: Modified regional maxima superimposed on original image.

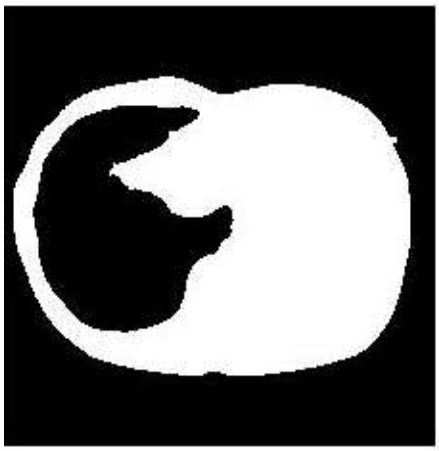

Figure 13: Threshold opening – closing by reconstruction.

D. Training.

Back-propagation algorithm is used to train the CNN to detect lung tumors in CT image of size 512 × 512 pixel. It consists of two phases. In the first phase, a CNN consists of multiple volumetric convolution, rectified linear units (ReLU) and max pooling layers is used to extract valuable volumetric features from input data. The second phase is the classifier. It has multipleFC and threshold layers, followed by a SoftMax layer to perform the high-level reasoning of the neural network. No scaling was applied to the CT images of the dataset to preserve the original values of the DICOM images as much as possible. During training, the randomsub-volumes extracted from the CT images of the training set and are normalized according to an estimate of the normal distribution of the voxel values in the dataset.

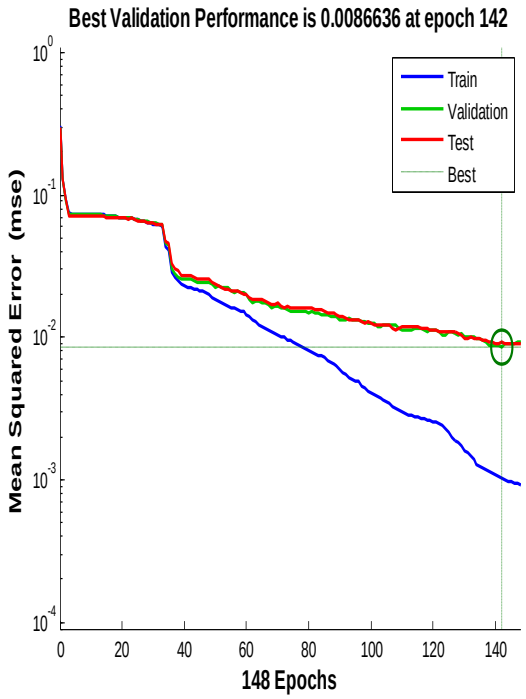

Figure 14: CNN Performance Validation.

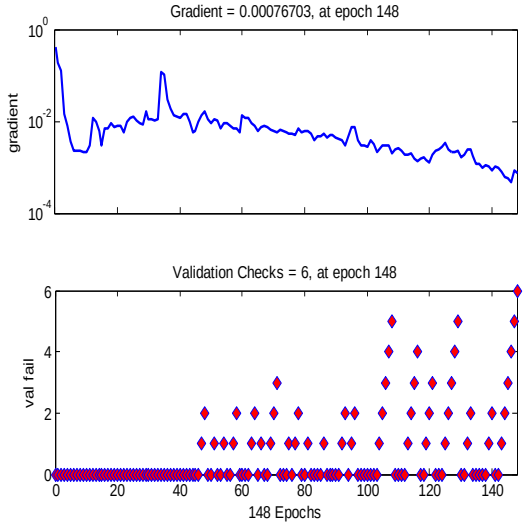

Figure 15: CNN Training state.

Mean squared errorforvalidation, training, andtest isdescribed in figure14.Validation check and training stateas show in figure15.

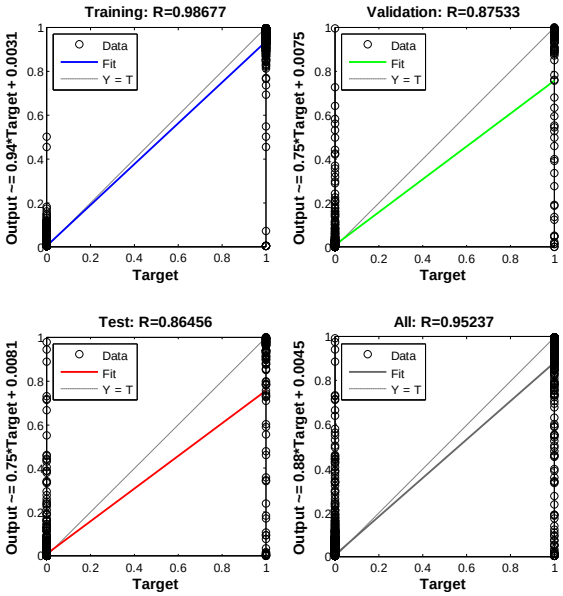

Figure 16: Regression.

The cancer affects level fixed using training, validation, and test as regression in figure 16.

E. Validate.

The neural network based on convolutional and segmentation has been implemented in MATLAB and the system is trained with sample data sets for the model to understand and familiarize the lung cancer. A sample image has been fed as an input to the trained model and the model at this stage is able to tell the presence of cancer and locate the cancer spot in the sample image of a lung cancer. The process involves the feeding the input image, preprocessing, feature extraction, identifying the cancer spot and indicate the results to the user. In case of the malignancy is present, a message indicating the presence of will be displayed on the screen .

Figure17: Cancer detected part.

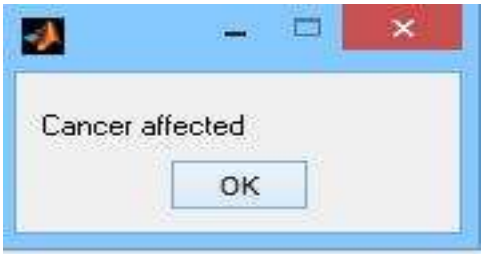

Figure 18: Result.

The cancer affected part as show in figure 17. The result for figure 18.

IV. CLUSTERING ALGORITHMS

A. Fuzzy- C-Means Clustering Algorithm.

Clustering is the method of separating the data into homogenous units by considering the relationship of objects. The clustering method is the allocation of the feature vectors into N clusters. Every n^{th}cluster has C_n as its center. Fuzzy Clustering is employed in numerous areas such as pattern recognition and fuzzy detection. Among various kinds of fuzzy clustering methods, Fuzzy C-Mean clustering (FCM) is the extensively used one. FCM utilizes reciprocal distance to determine fuzzy weights. The input of this process is a pre known number of clusters, N. The mean position of every the members of a cluster is identified. The output is the segregating of N clusters on a class of objects. Thegoal of the FCM cluster is to reduce the total weighted mean square error, (MSE). The FCM consents each feature vector to match with several clusters of different fuzzy membership values. The final segmentation is based on the optimum weight of the feature vector over all clusters. The steps involved in the FCM algorithm are given below.

Input: feature vectors (image voxels) v= $\{v_1, v_2, \ldots\ldots v_n\}$N=number of cluster.

Output: A group of clusters that lessning the sum of error of distance.

Steps:

1. Set random weight for every pixel using fuzzy weighting with positive weights $\{W_v^n\}$ ranging from 0-1.
2. Normalize the starting weights for eachv^nvoxel on all N clusters by using the below equation.
   $W_v^n / \sum_{i=1N} W_v^i$ (1)
3. Normalize the weights on n=1 1,….,N for each v to get W_v^n as given below

$$W_v^n = \frac{W_v^n}{\sum_{i=1}^{V} W_v^n}, v=1,2,\ldots\ldots\ldots,V \quad (2)$$

4. Estimate new centroidsC_n,n=1,…….,n from

$$C_n=\sum_{r=1}^{V} W_v^n(v), n=1,2,\ldots\ldots N \quad (3)$$

5. Update the weights $\{W_v^n\}$by using(4)

$$W_v^n = \frac{(1/||Xp - Cn||^2)^{1/(p-1)}}{\sum_{i=1}^{N}(1//||Xp - Cn||^2)^{1/(p+1)}}, \quad n=1,2,\ldots.N, v=1,2,..V \quad (4)$$

6. If the input is altered, do again from step 3, else stop the process.
7. Set each pixel to a cluster according to the maximum weight.

B. Convolution Neural Network.

Architecture of one hidden layer is depicted in Figure 19. It is examined for its skill to classify theNodules. This network consists of three layers namely, one input layer, one hidden layer, and one output layer. The input layer has P neurons that represent the P x P pixel of the image obtained from segmentation process. The hidden layer contains groups of N x N neurons organized as a sovereign N x N feature map (where N=P-r+1) and the r x r area is represented as the interested area. Each hidden neuron selects input from a r x r adjacent section on the input image section. If the neurons in the similar feature map are one neuron distant, then their interested areas in the input layer are one pixel distant. Each neuron of the similar feature map is reserved to take the identical group of R weights and accomplish the equal action on the resultant fragments of the input image.

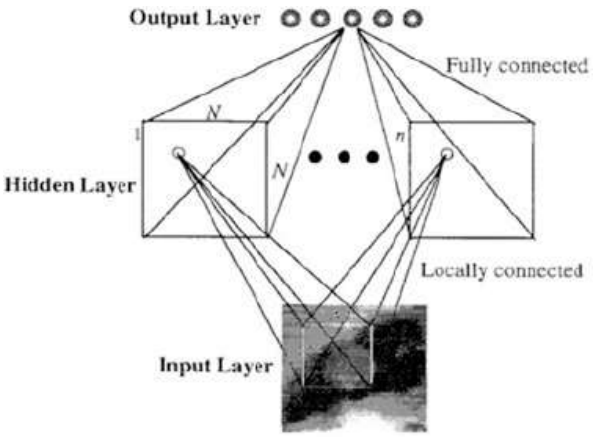

Figure 19: Architecture of One Hidden Layer CNN.

The advantage of hindering the weights permits the network to achieve shift-invariant pattern recognition. Hence, the total action is represented as the r x r convolution kernel. The feature map is the output obtained from the convolution of the input with the r x r convolution kernel. Each hidden neuron yj creates its output by means of an activation function represented. The minimum and maximum activation functions are zero and one, correspondingly.

W_{ij} - the weight between the hidden neuron, j and the pixel, iof the input image.

X_i- gray value of the input pixel i.
A_j - the bias of the hidden neuron j.
$x_1,x_2,\ldots,x_r$ Pixels on input image and they are connected to the neuron, j.
The output layer is entirely linked to the hidden layer. The sigmoid activation function, zo of the output neuron is represented by,
W_{0j} - Weight between the output neuron and neuron, j in the hidden layer
nN^2 - total number of neurons in the hidden layer
g_o - bias of the output neuron.
Hence, the network contains $(O+P^2+nN^2)$ number of neurons and $(nN^2 (R^2 +O+ 1) +O)$ number of links. These numbers include the input neurons and bias links also. The number of independent links is given by $nN^2(O+ 1) + nk^2 + O$. O represents the number of output neurons.

The network weights as well as the bias weights are altered by the application of the Back Propagation (BP) algorithm. The BP algorithm iteratively alters the weights with theintention of reducing the total error of the actual output vector from the target vector. The error functionto be reduced is called as the Sum-of-Squared Error (SSE).During training, the interested areas within one hidden class are restricted to consume the equal form of weights. The weights between hidden and output layers and the weights of every interested area, are altered by means of stochastic mode. In this method, the weight difference for each training sample is obtained from each back-propagated error and are altered instantaneously for every neuron.

V. DATA SECTION

Our primary dataset is the patient lung CT scan dataset from KagglesData Science Bowl (DSB) 2017 [20]. The dataset contains labeled data for 1387 patients, which we divide into training set of size 968, and test set of size 419. For each patient, the data consists of CT scan data and a label (0 for no cancer, 1 for cancer). Note that the Kaggle dataset does not have labeled nodules. For each patient, the CT scan data consists of a variable number of images (typically around 100- 400, each image is an axial slice) of 512 × 512 pixels. The slices are provided in DICOM format. Around 75% of the provided labels in the Kaggle dataset are 0, so we used a weighted loss function in our malignancy classifier to address this imbalance.

Because the Kaggle dataset alone proved to be inadequate to accurately classify the validation set, we also used the patient lung CT scan dataset with labeled nodules from the Lung Nodule Analysis 2016 (LUNA16) Challenge [21] to train a U-Net for lung nodule detection. The LUNA16 dataset contains labeled data for 888 patients, which we divided into a training set of size 710 and a validation set of size 178. For each patient, the data consists of CT scan data and a nodule label (list of nodule center coordinates and diameter). For each patient, the CT scan data consists of a variable number of images (typically around 100-400, each image is an axial slice) of 512 × 512 pixels.

LUNA16 data was used to train a U-Net for nodule detection, one of the phases in our classification pipeline. The problem is to accurately predict a patient's label ('cancer' or 'no cancer') based on the patient's Kaggle lung CT scan.

VI. EXPERIMENTAL RESULTS.

The enactment of the study proposed is valuedby benchmark metrics: Sensitivity, Specificity, and Accuracy.The description of these metrics and how their values are estimated. They are valued using confusion matrix which includes true and false positive and true and false negative. The true negative and positive envisage that the cases are diseased and non-diseased in which they are in fact diseased and non-diseased. The false negative and positive are simply contradictory to the true negative and positive.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (5)$$

Sensitivity was truthful positive estimates divided by the entire positives.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (6)$$

Specificity was truthful non- positive estimates divided by the entire negatives.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

Where TP - True Positive, TN – True Negative, FP- False Positive, FN- False Negative.

Table 1: Kaggles Dataset Accuracy Calculate.

	Sample Dataset	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Training dataset	968	842	10	12	104	89%	46%	88%
Testing/ Validation	419	376	5	7	31	93%	42%	91%
Average								**89%**

Table 2: LUNA16 Dataset Accuracy Calculate.

	Sample Dataset	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Training dataset	710	601	9	11	89	86%	45%	86%
Testing/ Validation	178	148	4	5	21	88%	44%	85%
Average								**85%**

VII. CONCLUSION

In this paper we developed a convolutional neural network (CNN) architecture to detect nodules in patients of lung cancer and detect. This step is a preprocessing step for CNN. While we perform well considering that we use less labeled data than most state-of-the-art CAD systems. As an interesting observation, the first layer is a preprocessing layer for segmentation using different techniques. Threshold, FCM and CNN are used to identify the nodules of patients.

The network can be trained end-to-end from image patches. Its main requirement is the availability of training database, but otherwise no assumptions are made about the objects of interest or underlying image modality.

In the future, it could be possible to extend our current model to not only determine whether or not the patient has cancer, but also determine the exact location of the cancerous nodules. The most immediate future work is to use FCM segmentation as the initial lung segmentation. Also, we saved our model at accuracy, but perhaps we could have saved at other metrics. Other future work include extending our models to images for other cancers. The advantage of not requiring too much labeled data specific to our cancer is it could make it generalizable to other cancers.

VIII. REFERENCES

[1] W.J. Choi and T.S. Choi, "Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach," Entropy, vol. 15, no. 2, pp. 507–523, 2013.

[2] A. Chon, N. Balachandar, and P. Lu, "Deep convolutional neural networks for lung cancer detection," tech. rep., Stanford University, 2017.

[3] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision.," in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), pp. 253–256, IEEE, 2010.

[4] K. Alex, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25 (NIPS 2012) (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, 2012.

[5] H. Suk, S. Lee, and D. Shen, "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis," NeuroImage, vol. 101, pp. 569–582, 2014.

[6] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of mr brain images.," Medical Image Computing and Computer-Assisted Intervention, vol. 16, no. Pt 2, pp. 649–656, 2013.

[7] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 1626–1630, 2014.

[8] D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in ct images," in 2015 12th Conference on Computer and Robot Vision, pp. 133–138, June 2015.

[9] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," Proceedings - International Symposium on Biomedical Imaging, vol. 2015-July, pp. 294–297, 2015.

[10] W. Sun, B. Zheng, and W. Qian, “Computer aided lung cancer diagnosis with deep learning algorithms,” in SPIE Medical Imaging, vol. 9785, pp. 97850Z–97850Z, International Society for Optics and Photonics, 2016.

[11] J. Tan, Y. Huo, Z. Liang, and L. Li, “A comparison study on the effect of false positive reduction in deep learning based detection for juxtapleural lung nodules: Cnnvsdnn,” in Proceedings of the Symposium on Modeling and Simulation in Medicine, MSM ’17, (San Diego, CA, USA), pp. 8:1–8:8, Society for Computer Simulation International, 2017.

[12] R. Golan, C. Jacob, and J. Denzinger, “Lung nodule detection in ct images using deep convolutional neural networks,” in 2016 International Joint Conference on Neural Networks (IJCNN), pp. 243–250, July 2016.

[13] AjalaFunmilola A, Oke O.A, Adedeji T.O, Alade O.M, Oyo Adewusi E.A, ―Fuzzy k-c-means Clustering Algorithm for Medical Image Segmentation‖, Journal of Information Engineering and Applications, ISSN 2224-5782 (print) ISSN 2225-0506 (online), Vol 2, No.6, 2012.

[14] Christian D., Naoufel W., Fatma T., Hussain, "Cell Extraction from Sputum Images for Early lung Cancer Detection", IEEE 978-1-4673-0784-0/12, 2012 .

[15] Mokhled S. AL-TARAWNEH, ―Lung Cancer Detection Using Image Processing Techniques‖, Leonardo Electronic Journal of Practices and Technologies, ISSN 1583-1078, Issue 20, January-June 2012.

[16] SajithKecheril S, D Venkataraman, J Suganthi and K Sujathan, "Segmentation of Lung Glandular Cells using Multiple Color Spaces", International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.2, No.3, June 2012 .

[17] Tseng L, Huang L, “An Adaptive Thresholding Method for Automatic Lung Segmentation in CT Images”, IEEE AFRICON,pp 1-5, September 23-25, 2009.

[18] Chaudhary A, Singh S S, “Lung Cancer Detection Using Digital Image Processing”, IJREAS vol 2, 1351-1359, Issue 2, 2012.

[19] Abdullah A A and Mohamaddiah H, “Development of Cellular Neural Network Algorithm for Detecting Lung Cancer Symptoms”, IEEE EMBS Conference on Biomedical Engineering & Sciences, 138- 143, 2010.

[20] Kaggle, “Data science bowl 2017.” https://www.kaggle.com/c/datascience-bowl-2017/data, 2017.

[21] LUNA16, “Lung nodule analysis 2016.” https://luna16.grandchallenge.org/, 2017.