IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

LUNG CANCER DETECTION SYSTEM BASED ON DEEP LEARNING

[1]Ramya C,[2]Prof. Aruna P G,[3]Dr. Bhagya H K
[1]M.Tech, Student,[2]Assistant Professor,[3]Professor
[1]Department of Electronics & Communication Engineering,
[1]KVG College of Engineering, Kurunjibag, Sulllia, D.K-574327

Abstract : Lung cancer stands as one of the most pervasive and deadliest cancer types on a global scale. Detecting lung cancer at an early stage can significantly heighten the likelihood of successful treatment and improved patient outcomes. Over the past years, image processing techniques have emerged as promising tools for facilitating early lung cancer detection. This paper presents a comprehensive review of the image processing techniques employed in lung cancer detection. The study delves into the diverse modalities of medical imaging, including X-rays, CT scans, and MRI, and explores the image processing techniques implemented for feature extraction and classification. The review accentuates the importance of employing image processing techniques for lung cancer detection, as they enable the identification of subtle changes in lung tissue that may elude the human eye. Additionally, the review emphasizes the pressing need for developing more accurate and robust image processing techniques to enhance early detection and treatment of lung cancer. In conclusion, the utilization of image processing techniques for lung cancer detection has demonstrated promise in recent times. This review sheds light on the potential benefits of leveraging these techniques to facilitate early detection and treatment of lung cancer, and highlights the importance of continued research and advancement in this field. Lung cancer is a prevalent and highly lethal form of cancer that affects populations

Keywords: Cancer Detection; Image processing; Feature extraction; Enhancement Watershed Masking.

INTRODUCTION

Lung cancer is a disease of abnormal cells multiplying and growing into a tumor. Lung cancer is a disease of abnormal cells multiplying and growing into atumor. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and the center of the chest. Lung cancer often spreads toward the center of the chest because the natural flow of lymph out of the lungs is toward the center of the chest. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node orto another part of the body through the bloodstream. Cancer that starts in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adenocarcinoma, and Squamous cell carcinomas.

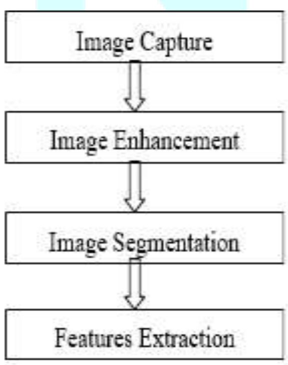

Figure 1. Lung cancer image processing stages

The rank order of cancers for both males and females among Jordanians in 2008 indicated that there were 356 cases of lung cancer accounting for (7.7 %) of all newly diagnosed cancer cases in 2008. Lung cancer affected 297 (13.1 %) males and 59 (2.5%) females with a male-to- female ratio of 5:1 Lung cancer ranked second among males and 10th among females. Figure 1 shows a general description of a lung cancer detection system that contains four basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access) [3]. The Second stage applies several techniques of image enhancement, to get the best level of quality and clearness.

The third stage applies image segmentation algorithms which playa n effective role in image processing stages, and the fourth stage obtains the general features from an enhanced segmented image which gives indicators of normality or abnormality of images.

Lung cancer is the most dangerous and widespread cancer in the world according to the stage of discovery of the cancer cells in the lungs, so the process of early detection of the disease plays a very important and essential role to avoid seriously advanced stages to reduce its percentage of distribution. This research aimed to detect features for accurate image comparison as pixels percentage and mask labeling.

Nomenclature

CT SCAN – Computed Tomography Scan

AHHMM – Adaptive Hierarchical Heuristic Mathematical Model

1. MATERIALS

Patient Recruitment

For the AHHMM challenge, 200 lung patients were recruited in this study at the Department of Pulmonary Oncology in the First Hospital of Changsha, from January 2016 to November 2017. According to the American Joint Committee on Cancer (AJCC) staging system, patients first diagnosed with lung/bronchus cancer (site: C34.1-C34.9; histology type: adenocarcinoma, squamous cell carcinoma, and small cell carcinoma) were recruited. Other inclusion criteria included: 1) pathologically confirmed patients with surgery biopsy maintained; 2) no radiotherapy before surgery; 3) aged between 30 and 90 yr. The exclusion criteria were: 1) multiple primary cancers 2) metastatic lung cancer; 3) patients with immune deficiency or organ-transplantation history; 4) patients who did not provide informed consent. This study was approved by the Ethics Committee of the First Hospital of Changsha. Informed consent was obtained from each patient before the examination. Necessary demographic and clinical information for each patient, such as age, gender, stage, pathology, etc. were collected.

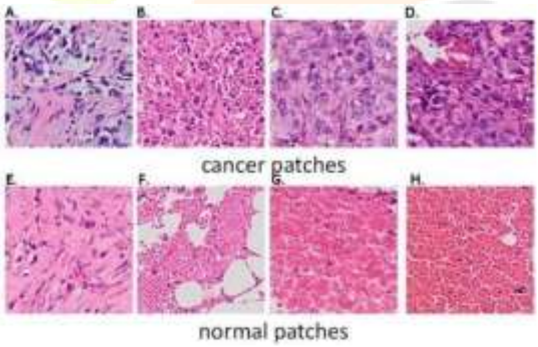

Fig. 2. Example of tumor patches and normal patches.

Data Preparation

Histological slides were stained with H&E scanned by a digital slide scanner (3DHISTECH Panoramic 250) with objective magnifications of 20x. A close look at different tissues in the slides can be seen in Fig. 2. One can see that the patch colors were quite different even among the patches from normal tissue due to the staining variability. The appearance of the cancer regions was also quite different because of the different cancer types. For instance, Fig. 2. (A) and (B) represent small-cell lung cancer, and Fig. 2. (C) and (D) represent squamous cell lung cancer and adenoid cell lung cancer. Fig. 2.(E)-(H) are normal patches. In total, 200 H&E stained slides were scanned and digitized. We randomly split those 200 slides into training and test sets. 150slides with annotations were released as the training set. 50 slides were held as the test set.

Fig. 3. The distribution of participants

The main types of cancer were included in our data: squamous cell carcinoma, small cell carcinoma, and adenocarcinoma. The ratio of them was approximately 6:3:1. One pathologist with 30 years of experience (the director of the pathology department) annotated the cancer regions for all 200 slides (See Fig. 1). We also asked the second pathologist (with 20 years of experience) to annotate the test set only. The annotation of the second pathologist was only used for accessing the inter-observer variability. Participants were allowed to use their training data for pre-training.

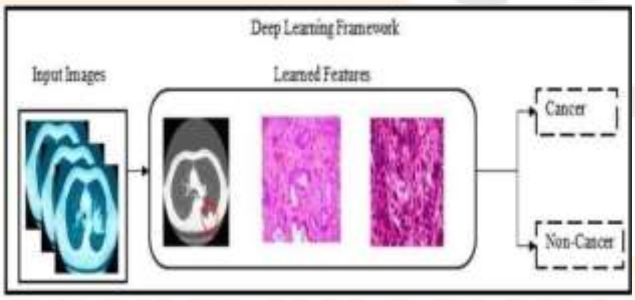

Fig 4. Training procedure of a deep learning method

All data were uploaded to Microsoft One Drive, Google Drive, and Baidu Pan for participants from different regions. Whole-Slide images were released inthe TIFF format. Manual annotations were in XML format. In clinical practice, more than one sample from the same biopsy was scanned. If samples had a similar shape, the pathologist only annotated one sample in the WSI. Participants were suggested to use ASAP4 to make a bounding box themselves to exclude the unused samples

2. *RESEARCH METHODOLOGY*

Image-Preprocessing

Medical image processing methods have been used to predict changes present in the lungs of the recorded CT pulmonary images. The lung cancer processing structure is shown in Figure 5. Lung tumor has successfully been detected by the medical image processing technique according to the processing structure of the lung pictures. The following section provides a detailed explanation of each processing phase. The images captured are examined in pixel noise prediction and contrasted details to improve the quality of the Computed Tomography pulmonary image since the image captured contains several incoherent details, and low pixel quality which decreases the accuracy of detected lung cancer. The pixel-intensive testing process has been utilized that essentially changes the pixel picture interpretation, and the accuracy of the Computed Tomography lung image is increased. The current pixel shift removes the inconsistent pixel and the noise pixel. Image histogram methods are used for image quality enhancement because it works better and simply on different images.

Algorithm 1 Modified K-Means clustering algorithm

```
K-means
(attributes (m), priority (no of a cluster))
{
Sort (attributes.Priority);For (j=1 to k)
{
Centroid [j]=findmean (Cj);
}
For (i=1 to k (k-1)/2)
{
Dis [i] =dis (centroid [1tok]);Temp [i]=dis [i]/n;
}
Until (–m!=0)
{
Apply m to plotted points;
Shift +or -temp towards other centroid;For (j=1 to k)
{
Centroid [j]=findMean (Cj);
}
For (i=1 to k (k-1)/2)
{
Dis [i]=dis (centroid [1 to k]);Temp [i]=dis [i]/m;
    }
}
}//end
```

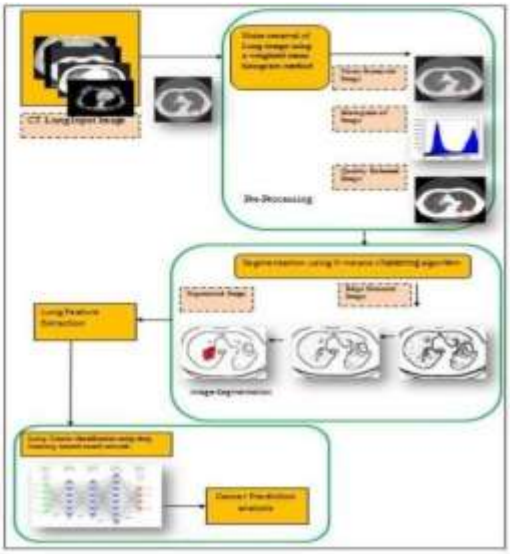

Fig 5. The architecture of the proposed AHHMM method

Image-Segmentation

The next major step is the segmentation of the region injured by cancer by the use of the K-mean algorithm using the enhanced lung CT image. The implemented segmentation approach inspects the pixel similarity in the lung CT images and divides the images into several sub-settings to predict the area concerned. Analysis of the image pixels in the image and group of the same super pixels into a similar group for abrupt pixels detection.

The pixel or information is subsequently tested during the image fragmentation process to predict the comparability of data using the same value of spectrum or pixels. The pixel similarity attribute involves the quantitative image evaluation used to accurately shape the cluster. The enhanced quality of the Lung Computed Tomography image pixels is tested and every pixel is considered an element in the app. For parts of the affected region, an undirected graph is created.

Image-Classification

The final stage is the identification of lung cancer through an explosion-trained deep learning neural network (DITNN). The image training should be carried out utilizing a deep- learning NN before the classification process is done since it does not need manual features. Alternatively, the deep learning process uses a captured segmented image or lung CT the image that utilizes a huge number of hidden layers to Identify the boundaries of the image and the corresponding features utilized for the networks creation with a large amount f information.

1. RESULTS AND DISCUSSION

The findings of segmentation included nodules connecting to the inner wall and precise segmentation of the lung parenchyma. In the classification and detection of objects, efficient segmentation is very important. An efficient method of segmentation is developed based on the distribution of lung CT images by spatial pixel intensity. CT images will simply remove pulmonary parenchyma in morphological operations; however, tissue around the lung can be included. Figure 6 shows the segmentation results from comparison 6(a) shows the CT images 6(b) shows the segmentation established by region growing

6(c) shows the Lung segmentation established by the Modified K-means clustering algorithm 6(d) shows the Lung segmentation established by AHHMM approach.

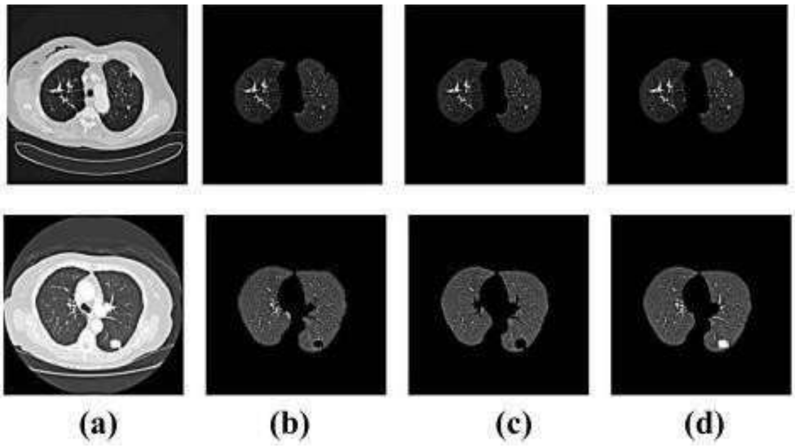

Fig 6. Comparison of segmentation outcomes for all the grades

(a) Computed Tomography images
(b) segmentation established by region growing
(c) Lung segmentation established by Modified K-means clustering algorithm
(d) Lung segmentation established by AHHMM approach

Mean Accuracy Analysis

The suggested k-means lung segmentation, which offers 96% lung cancer segmentation, only maintains this mean accuracy in low-dose Computed Tomography images, and the same approaches are recommended for separating the lungs into high-resolution Computed Tomography images. The classifier accuracy cannot be determined by the complete set of data set attributes. The measured value is compared with the extracted and trained features for cancer classification based on deep learning. Double-time matching enhances the accuracy of prediction and decreases the failure rate efficiently. The mean accuracy of the classifier specifies how many samples have been correctly predicted by the total number of samples and is shown as follows. Figure 8 shows the mean accuracy ratio analysis of the proposed AHHMM method.

$$\text{Mean Accuracy Ratio} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total no of predictions}} * 100$$

TABLE 1. Mean accuracy analysis.

Methods	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image 10
Double convolutional deep neural network (DCDNN)	78.5	88.4	89.3	90.2	91.3	92.8	93.4	94.6	94.7	94.9
Artificial Neural Network (ANN)	92.8	93.6	92.54	92.31	93.4	92.97	92.5	93.78	94.2	94.4
Deep Convolutional neural network (DCNN)	96.5	97	97.4	96.1	97.6	96.2	97.4	97.5	97.6	97.8
Improved profuse clustering technique and deep learning instantaneously trained neural network (IPCT-DLITNN)	97.2	97.5	97.1	97.8	97.9	97.45	97.9	98.01	97.9	98.2
Adaptive Hierarchical Heuristic Mathematical model (AHHMM)	98.09	98.4	98.2	98.4	98.6	98.1	98.5	98.2	98.8	98.9

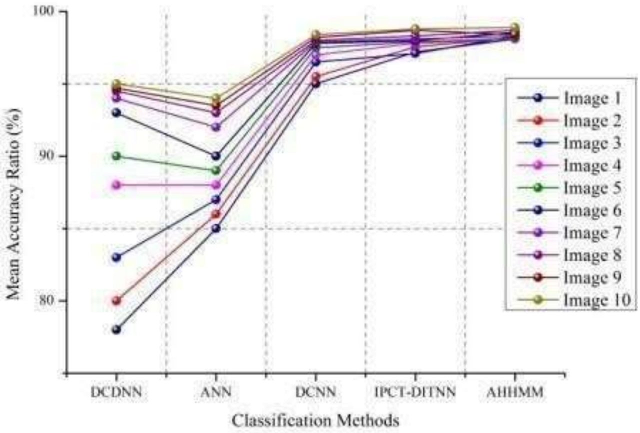

Table 1 shows the mean accuracy analysis of the proposed AHHMM method. Biopsy predicts lung cancer, it is rather difficult to maintain precision and

accuracy. The CT scan is therefore done by the transmission of X-rays for the analysis of changes in the body. The technique of screening has helped to predict cancer of the lung, but it is difficult to maintain an earlier recognition of large cell carcinoma and cancer detection.

2. CONCLUSION AND FUTURE SCOPE

Lung cancer is a dangerous disease, and early-stage detection is therefore necessary. This paper presents a deep learning-assisted Adaptive Hierarchical Heuristic Mathematical Model (AHHMM) to predict lung cancer on computed tomography images. This paper uses the Modified K-means algorithm to pre-classify pictures into slices of images in the same image, where the DNN will concentrate on the image classification of images in similar images. The next thing is the convolution layer with filtered edges to scan for lung cancer thoroughly. Then, estimating the weighted mean function which replaces the pixel utilizing the cumulative distribution and likelihood distribution method improved the image quality. The injured portion has been segmented by a pixel-like value measurement after the image has been improved. Based on the similarity calculation, spectral- related features have been extracted. The proposed AHHMM system predicts computed tomography scanning images of lung cancer successfully. At the end of the system, you can say that the system is satisfying its desires. The findings of the evaluation showed that around 90% of the images have been correctly identified. Hybridized Heuristic Mathematical Model will be implemented in the future for predicting lung cancer at an earlier stage.

Acknowledgements

I would like to extend my heartfelt thanks to everyone who has contributed to this journal research paper. I am deeply grateful to my mentor, Prof. Aruna P G, for their guidance, encouragement, and expert feedback throughout the research process. Their invaluable insights and advice have been essential in shaping the direction and scope of this paper.

I would like to extend my heartfelt thanks to everyone who has contributed to this journal research paper. I am deeply grateful to my PG coordinator, Dr. Bhagya H K, for their guidance, encouragement, and expert feedback throughout the research process. Their invaluable insights and advice have been essential in shaping the direction and scope of this paper.

I would also like to express my appreciation to the reviewers of this paper, whose constructive comments and suggestions have greatly improved the quality of our work. Their meticulous and insightful review process has helped me to refine my research and present it in a more effective and compelling manner.

Furthermore, I would like to thank ,my classmates and collaborators who have provided their valuable feedback, ideas, and support during the course of this research. Their contributions have been instrumental in shaping my approach and refining my results.

Finally, I would like to acknowledge the support of my friends and family, who have encouraged me in every step of the way. Their unwavering supportand belief in my abilities have been a constant source of motivation and inspiration.

Thank you all for your contributions to this research paper.

REFERENCES

[1] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, „„Classification and mutation prediction from non- small cell lung cancer histopathology images using deep learning,"" Nature Med., vol. 24, no. 10, pp. 1559–1567, 2018.

[2] P. Gupta and A. K. Malhi, „„Using deep learning to enhance head and neck cancer diagnosis and classification,"" in Proc. IEEE Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN), New York, NY, USA: ACM, vol. 28, Jul. 2018, pp. 1–7.

[3] B. A. Skourt, A. El Hassani, and A. Majda, „„Lung CT image segmentation using deep neural networks,"" Procedia Comput. Sci., vol. 127, pp. 109– 113, Jan. 2018.

[4] D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher,

L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P.Naidich, and S. Shetty, „„End-to-end lung cancer screening with three-dimensional deep learning onlow-dose chest computed tomography,"" Nature Med., vol. 25, no. 6, pp. 954–961, Jun. 2019.

[5] N. Nasrullah, J. Sang, M. S. Alam, and H. Xiang, „„Automated detection and classification for early stage lung cancer on CT images using deep learning,"" in Proc. 30th Pattern Recognit. Tracking, May 2019, Art. no. 109950

[6] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, „„Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks,"" Measurement, vol. 145, pp. 702–712, Oct. 2019.

[7] S. K. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, „„Optimal deep learning model for classification of lung cancer on CT images,"" Future Gener. Comput. Syst., vol. 92, pp. 374–382, Mar 2019.

[8] S. Trajanovski, D. Mavroeidis, C. Leon Swisher, B. G. Gebre, B. S. Veeling, R. Wiemker, T. Klinder, A. Tahmasebi,

S. M. Regis, C. Wald, B. J. McKee, S. Flacke, H. MacMahon,and H. Pien, „„Towards radiologist-level cancer risk assessment in CT lung screening using deep learning,"" 2018.

[9] S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, „„An interportable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification,"" Expert Syst. Appl., vol. 128, pp. 84–95, Aug. 2019.

[10] G. Jakimovski and D. Davcev, „„Using double convolution neural network for lung cancer stage detection,"" Appl. Sci., vol. 9, no. 3, p. 427, 2019.

[11] I. M. Nasser and S. S. Abu-Naser, „„Lung cancer detection using artificial neural network,"" Int. J. Eng. Inf. Syst., vol. 3, no. 3, pp. 17–23, 2019.

[12] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa,„„Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks,"" Measurement, vol. 145, pp. 702–712, Oct. 2019.

G. Litjens et al., "A survey on deep learning in medical image analysis," Med. Image. Anal., vol. 42, pp. 60–88, 2017

[13] L. C. Chen et al., "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell., vol, 40, no. 4, pp. 834–848, Apr. 2018