International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Convolution Neural Network Approach for Lung Cancer Detection

Shobha Singh

Integrated M.Tech, Department of Computer Science and Engineering, Gautam Buddha University, Greater Noida, India

---

Abstract - *Medical Image Processing Modern three-dimensional (3-D) medical imaging offers the potential and promise for major advances in science and medicine as higher fidelity images are produced. Lung cancer are related to smoking or less often to exposure to radon or other environmental factors that's why this can be prevented. But still it is not yet clear if these cancers can be prevented or not. The technique of SVM is proposed in the previous research work. The SVM classification method give low accuracy for the lung cancer detection. The novel approach of classification is required which can improve accuracy for the lung cancer detection. In this research work, approach of segmentation, feature extraction and CNN will be applied for locating, characterizing cancer portion. It is analysed that proposed approach improve accuracy, precision and recall.*

Key Words: **CNN, Lung cancer, GLCM, K-means**

1. INTRODUCTION

Image processing is a branch of science using which information of the images is uncovered. Image processing is a sort of signal processing where the input is an image, (such as a photograph or video frame) and the output of image processing may be either an image or a set of features or metrics of the given input image. A number of image processing methods treat the image as a 2D (two-dimensional) signal and apply general signal-processing methods to this image [1].Medical Image Processing is a relatively a new research field. This field is based on the application of computer vision techniques to data sets acquired using medical imaging modalities. These modalities include ultrasound, MRI (Magnetic Resonance Imaging), CT (Computed Tomography), SPECT (Single Photon Emission Computed Tomography), PET (Positron Emission Tomography) and fundus photography [5]. In the last few years, the popularity of MIP (Medical Image Processing) has been increased a lot. Numerous applications are being developed for reducing the workload of doctors. These applications have the ability of medical image interpretation for the efficient diagnosis of irregularities. The recent research community is focussing on developing the automatic screening systems for different pathologies. Medical imaging can be described as a process of generating visible images of inner body parts for scientific and medicinal analysis and treatment. This process enables doctors to see the working of the interior body tissues. This process follows the disease detection and control. Lung cancer is a state in which cells are split in disorder manner into the lungs. The tumours start developing and the breathing ability of a person is diminishing. One of the major reasons of deaths from cancer and the second most detected disease in patients in US is lung cancer. The cancer detection in the early stages helps the patient in curing the disease. It is not easy to recognise the lung cancer in its former phases as the symptoms of cancer as well as of respiratory infection are same. There may be a condition in which no symptom is found. The leading reason of lung cancer is smoking with cigarette. The tobacco, second-hand smoking or a hereditary problem of lung cancer are the other main reasons of lung cancer. The lung cancer is classified into two kinds by doctors after observing the form of it under the microscope. The two general kinds of lung cancer are mentioned below:

- Small cell lung cancer: This cancer takes place in the person who smokes a lot. It is not as much common cancer as non-small cell cancer.
- Non-small cell lung cancer: This term is used for the numerous kinds of lung cancers that have the same behaviour. The squamous cell carcinoma, adeno-carcinoma and large cell carcinoma were comprised in the non-small cell cancers of lung.

The symptoms of lung cancer are not happened in the cancer patients always before the later phase of the disease. The symptoms of lung cancer may be observed in some of the patients but they consider them as symptoms of acute illness. The chest infections, headaches, less appetite, voice changes, breathing problem, cough that become worse [10]. Lung cancer is one of the most severe and prevailing cancer

all over the world as per the stage of discovery of the cancer cells in the lungs. Therefore, it is necessary to detect this disease in the early stage. The early detection of this disease plays an extremely vital and crucial role for avoiding the serious advanced stages of this disease so that the its percentage of distribution can be reduced. Figure shows a general structure of lung cancer detection using lung cancer detection. The task of lung cancer detection is performed in different stages. These stages are:

a. Image capture

b. Image enhancement

c. Image segmentation

d. Feature extraction

2. LITERATURE REVIEW

Nidhi S. Nadkarni, et.al (2019) proposed an image processing methodology to detect the lung tumor in its earlier stages [17]. This methodology was one of the most broadly used techniques in the medical field. An automated approach had illustrated in this paper for the lung cancer detection in the images of CT scan. The median filtering was utilized for the pre- processing of an image and mathematical morphological operations were implemented for the lung region of interest segmentation. These techniques were suggested in the detection of lung cancer algorithm. The calculation of geometrical features had done from the region of interest that was extracted. The support vector machine was employed for the classification of CT scan images using geometrical features.

JaneeAlam, et.al (2018) recommended an efficient algorithm for detection and prediction of lung cancer. In this algorithm multi-class classifier named SVM was carried out [18]. The cancer was detected using multi-stage classification. The possibility of lung cancer was also forecasted in this system. The improvement and segmentation of an image at each classification stage was performed independently. In the image enhancement, various processes were implemented such as scaling of image, transformation of color space and improvement in contrast. The segmentation was completed using threshold and marker-controlled segmentation that based on watershed. The binary classifier of support vector machine was employed for the classification. The higher accuracy had been achieved in the detection and predication of lung cancer using recommended method.

Sayali Satish Kanitkar, et.al (2015) studied that the lung cancer was the main reason of death in cancer. In the early stage, there were more options for treatment, less invasive surgery and the survival rate was high for the cancer detection [19]. If the cancer was detected in early stages in lung cancer, the survival rate of person who suffered from cancer increased from14 to 49%. The cancer was one of the most dangerous diseases in the whole world. The lung cancer mainly occurred due to the presence of cancer cells in the lungs. The chances of an effective treatment were increased when the detection of cancer started at its early phase. The images that were more efficient as compared to X-ray had utilized in it. To study the detection of lung cancer from the images of CT scan, MATLAB software was publicly available. The pre-processing and segmentation of an image, extraction of attributes and method for categorization were comprised in this process. A technique for cancerous cells detection from lungs had suggested in this paper that helped to decrease the detection error.

Anita Chaudhary, et.al (2012) investigated that the common reason for patient's death was lung cancer [20]. The chances to survive in the lung cancer were maximized if it was detected at its early stages. The survival rate of patients of lung cancer was increased from 14% to 49% in recent five years of this disease. The CT method was better than the X-ray. Various techniques were utilized for the present of lung cancer detection due to the time restraint. The image processing carried out for classification of lung cancer present in CT scan images. The implementation of MATLAB software was in every process. All the procedures that had used during image processing were discussed in it. To acquire the more accurate outcomes was the major objective of this paper. For this, several improvement and segmentation methods were employed in it.

S. Kalaivani, Pramit Chatterjee, et.al (2017) examined that the trained professionals were implemented the cancer detection methodology to detect the cancer in the early stage [21]. The tedious process included in this methodology. During the detection procedure, the probability of human error was very high so at that time, an automated procedure had required. An automated procedure had utilized in the advance cancer detection so that human error was decreased. That automated procedure made more accurate and was designed using ANN and algorithms of image processing to detect the lung cancer in its advanced phase.

N. Werghi, et.al (2012) proposed that the major cause of deaths of cancer patients was lung cancer in the whole

world. The survival rate in overall five years was only 15%. The chances of successful treatment were maximized by the earlier detection of cancer. A striking approach for cancer detection was computer-aided diagnosis system in which sputum stained smears images were utilized. This was an invasive, practical and less expensive approach. A Bayesian classification and segmentation of main shift were employed to detect and segment the sputum cells that were present in the images of sputum. The comparison of this suggested method with other competitive methods was performed and series of experiments were carried out with data set in which 88 images included.

LilikAnifah, et.al (2017) suggested that the common reason of cancer death worldwide was lung cancer. The patient was saved if the lung cancer detected earlier [23]. The GLCM feature that based on Neural Network Back-propagation was employed to detect the lung cancer. The lung data used had originated from the database named Cancer imaging archive database in which 50 CT images were comprised. The normal and the lung cancer were two clusters of CT image. The images were preprocessed, the features were extracted and the Neural Network Back-propagation was utilized for the lung cancer detection in this research. The accuracy of normal lung and lung cancer detection in CT image using this system was calculated as 80%.

Qing Wu, et.al (2017) recommended a new algorithm based on neural-network for the detection of SCLC from the CT images. This algorithm was also considered as EDM [24]. The lung cancer detection in early stages was provided in this work. The National Cancer Institute had presented lung Computed Tomography scans with high-resolution such as the training data and testing data. The 12 lung CT scans were taken in this work from the library. Six images were of healthy lungs and six images were scanned from patients using SCLC. The model was trained by selecting five scans from every group at random. Two scans were employed for testing. There accuracy achieved from these algorithms had evaluated as 77.8%.

Kyamelia Roy, et.al (2019) proposed that the development of precision and to determine the particular value to detect the lung carcinoma earlier was the main objective of this study. Themethod of biomedical image processing and knowledge discovery in data were combined together and employed in it [25]. The CT scan image of lungs was pre-processed. The segmentation of these CT scanned images had performed in ROI. The classification of the dissimilar features was done using Random Forest method. The SURF algorithm utilized for the Saliency enhancement. The extraction of entropy co-relation and variance from the Saliency Enhanced images was done with classifier of SVM. The obtained image was healthy or carcinomic had been recognized in classification. The performance was analyzed when minimum objective used with the number of function evaluation plot. The Random forest algorithm and Support Vector Machine classifier were implemented to perform the whole procedure. The Support Vector Machine provided the efficient outcome. The results demonstrated that the efficiency was achieved 94.5%, the sensitivity and specificity had evaluated as 74.2% and 77.6% respectively.

K. Punithavathy, et.al (2015) presented that the main goal of this paper was the development of technique to detect the lung cancer automatically from the PET/CT images [26]. The CLAHE) and Wiener filtering employed so that the artifacts were removed due to noise and variation in contrast. The morphological operators utilized for the extraction of lung ROI. The more information related to texture was extracted from Haralick statistical texture features than visual assessment. Thus, Haralick features were implemented in this technique. The classification of regions was done using FCM that described that the regions were normal or abnormal. For the implementation of suggested technique lung PET/CT images of patients were used and MATLAB software was used. With the ROC curve, the evaluation of suggested technique was acquired. The accuracy obtained for the classification and detection of cancer was evaluated as 92.6%.

3. RESEARCH GAPS

The image processing is the approach which can process digital information stored in the form of pixels. The medical image processing is the field of image processing in which medical information like MRI images, CT images are processed. This research work is related to lung cancer detection from MRI images. The lung cancer detection techniques have various phases which are pre-processing, segmentation, features extraction and classification. In the phase of pre-processing noise from the input image is removed and in the segmentation phase, the image will be segmented into certain parts from further processing. In the third phase, the approach of features extraction will be applied which extract features of input image. In the last phase, the technique of classification will be applied which will classify tumor and non tumor portions. In the previous years, many techniques are designed for the lung cancer detection which are based on neural network approach. In the designed techniques, the problem exits of tumor localization and characterization. The technique for lung

cancer detection designed by Alexis Arnaud solves the problem of tumor localization and characterization. The designed technique has high complexity due which execution time is very high. The techniques need to designed which can localize and characterize tumor portion accurately and in least amount of time.

4. RESEARCH METHODOLOGY

This research work is related to lung cancer detection from the CT scan image using image processing techniques. The proposed methodology has the four phases for the lung cancer localization and characterization. Following are the various phases of the lung cancer detection:-

This research work is related to lung cancer detection from the CT scan image using image processing techniques. The proposed methodology has the four phases for the lung cancer localization and characterization. Following are the various phases of the lung cancer detection:-

1. Pre-processing:- The pre-processing is the first phase in which CT scan image is taken as input. The technique of image de-noising will be applied which will remove noise from the input image.

2. Segmentation:- In the second phase, the approach of region based segmentation will be applied which will segment the similar and dissimilar regions from the CT scan image The outh's segmentation technique is applied for the segmentation. The sectioned picture attained from thresholding comprises several benefits like lesser storage space, speedy dispensation velocity and easiness in exploitation in comparison with gray level picture that generally includes 256 steps. In the presented work, a gray scale picture is utilized for thresholding process. In this process, rgb picture is converted into binary picture. The obtained picture is in the form of black and white.

3. Feature Extraction:- The feature extraction is the third phase, in which GLCM algorithm will be applied for the feature extraction of the CT scan image In this step, the GLCM algorithm is applied for the feature extraction. The GLCM algorithm will extract the textural features of the input image. The GLCM algorithm extracts 13 features of the image for the tumor detection

$$\text{Energy} = \sqrt{\sum_{i,j=0}^{N-1} p_{i,j}^2} \text{----------(1)}$$

$$\text{Entropy} = \sum_i p_i log_x i \text{----------(2)}$$

$$\text{Contrast} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \text{----------(3)}$$

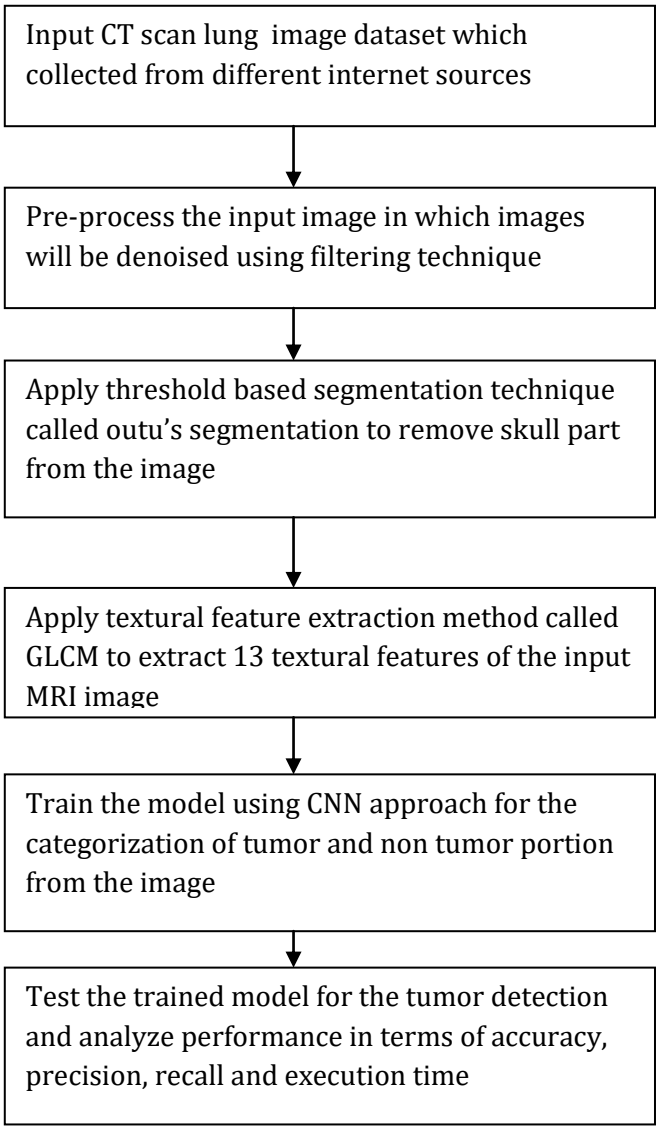

Fig -1: Proposed Methodology Flow chart

4. Classification:- In the last phase, the approach of CNN will be applied which can categorize and localize the cancer part. All the data points of an individual class are separated by the best hyperplane, this can be identified through the classification provided by CNN. In the CNN the largest the best hyperplane is described by the largest margin between the two classes. There are no interior data points when there is maximum width between the slabs parallel to the hyperplane

which is also known as margin. The maximum margin in hyperplane is separated by the CNN algorithm.

5. RESULT AND DISCUSSION

The proposed approach is implemented in MATLAB. The proposed algorithm use the computer vision and neural network tool box to implementation. The results are analyzed in terms of accuracy. The dataset is collected from authentic data source which is ADNT [26].

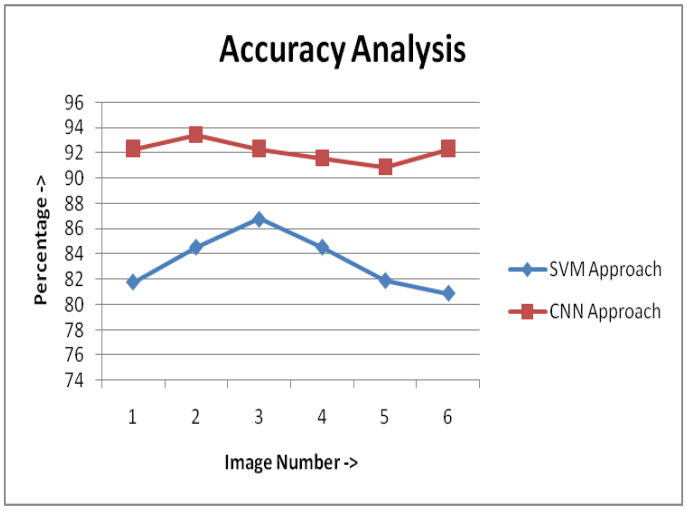

Fig -2: Accuracy Analysis

As shown in figure 2, the accuracy of the existing system which is SVM approach is compared with the proposed approach which is CNN approach. The system is tested on different number of images and it is analysed that CNN give best results as compared to SVM approach.

Table 1: Accuracy Analysis

Image Number	SVM Approach	CNN Approach
1	82	92
2	84.5	93.5
3	86.52	92
4	84	92
5	82	91.5
6	81	92

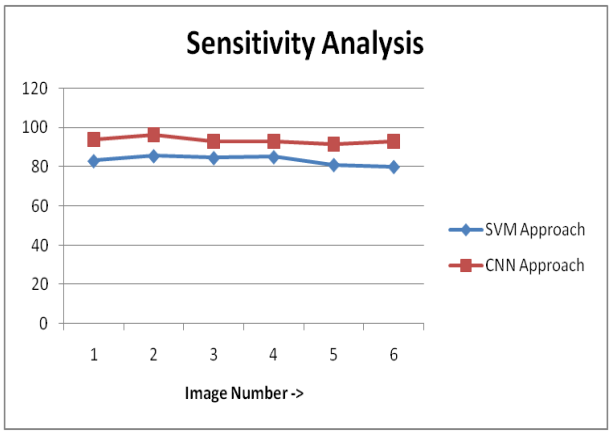

Fig -3: Sensitivity Analysis

As shown in figure 3, the sensitivity of the existing system which is SVM approach is compared with the proposed approach which is CNN approach. The system is tested on different number of images and it is analyzed that CNN give best results as compared to SVM approach.

Table 2: Sensitivity Analysis

Image Number	SVM Approach	CNN Approach
1	81	92
2	83.5	94.5
3	84.52	93
4	85	93
5	81	90.5
6	81	93

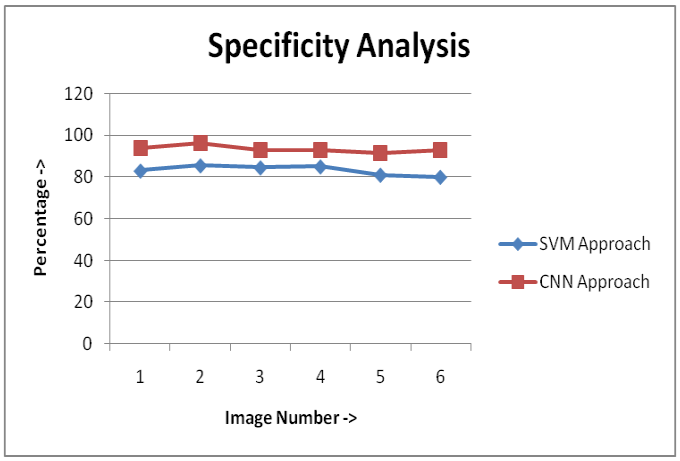

Fig -4: Specificity Analysis

As shown in figure 4, the specificity of the existing system which is SVM approach is compared with the proposed approach which is CNN approach. The system is tested on different number of images and it is analysed that CNN give best results as compared to SVM approach.

Table 3: Specificity Analysis

Image Number	SVM Approach	CNN Approach
1	81	92
2	83.5	94.5
3	84.52	93
4	85	93
5	81	90.5
6	81	93

6. CONCLUSIONS

To detect the presence of cancer nodule CT scan images are used. Further the pre-processing composed of two processes. Image enhancement and image segmentation are that two processes. The lung cancer detection have various steps which include pre-processing, segmentation, feature extraction and classification. In the previous research work, SVM classifier is used for the lung cancer detection.

The SVM classifier give low accuracy for the lung cancer detection. The dataset of the lung cancer detection is collected from the Kaggle. In the proposed method CNN is used for the classification with the GLCM algorithm for the feature extraction. The simulation is done in MATLAB and performance is analyzed in terms of specificity, sensitivity and accuracy. It is analyzed that proposed approach improve accuracy up to 8 percent as compared to existing method.

REFERENCES

[1] Shradha fule, "Lung Cancer Detection Using Image Processing Techniques", 2017, International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 12

[2] Ayushi Shukla, Chinmay Parab, Pratik Patil, Prof. Savita Sangam, "Lung Cancer Detection using Image Processing Techniques", 2018, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04

[3] Bariqi Abdillah, Alhadi Bustamam, and Devvi Sarwinda, "Image processing-based detection of lung cancer on CT scan images", 2016, The Asian Mathematical Conference

[4] Febr Mokhled S. AL-TARAWNEH, "Lung Cancer DetectionUsing Image Processing Techniques", 2012, Leonardo ElectronicJournal of Practices and Technologies.

[5] Muhammad Usman, Muhammad Shoaib and Mohamad Rahal, "Lung Cancer Detection Using Digital Image Processing", 2013, PIERS Proceedings, Stockholm, Sweden

[6] Anita Chaudhary and Sonit Sukhraj Singh, "Multiresolution Analysis Technique for Lung Cancer Detection in Computed Tomographic Images", 2012, International Journal of Research in Engineering & Applied Sciences, IJREAS Volume 2, Issue 2

[7] Disha Sharma, Gagandeep Jindal, "Computer Aided Diagnosis System for Detection of Lung Cancer in CT

Scan Images", 2011, International Journal of Computer and Electrical Engineering, Vol. 3, No. 5

[8] Samir Kumar Bandyopadhyay, "Edge Detection from CT Images of Lung", 2012, International Journal of Engineering Science & Advanced Technology Volume 2, Issue 1, pp 34 – 37

[9] Retz Mahima Devarapalli1, Hemantha Kumar Kalluri, Venkatesulu Dondeti, "Lung Cancer Detection of CT Lung Images", 2019, International Journal of Recent Technology and Engineering (IJRTE), Volume-7, Issue-5S4

[10] Nisar Ahmed Memon, Anwar Majid Mirza, and S.A.M. Gilani, "Segmentation of Lungs from CT Scan Images for Early Diagnosis of Lung Cancer", 2006, World Academy of Science, Engineering and Technology

[11] Disha Sharma, Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", 2011, International Conference on Computational Techniques and Artificial Intelligence

[12] Al-Tarawneh, M.S., "Lung Cancer Detection Using ImageProcessing Techniques", 2012, Leonardo Electronic Journalof Practices and Technologies, Volume 11, No. 21,pp. 147-58

[13] Malik, B. Singh, J.P., Singh, V.B.P., and Naresh, P.,"Lung Cancer Detection at Initial Stage by Using ImageProcessing and Classification Techniques", 2016, Lung Cancer, Volume 3, No. 11

[14] Gajdhane, M.V.A., and Deshpande, L., "Detection ofLung Cancer Stages on CT scan Images by Using VariousImage Processing Techniques", 2014, InternationalOrganization of Scientific Research Journal of ComputerEngineering, pp. 2278-0661

[15] Patil, B.G., and Jain, S.N., "Cancer Cells Detection usingDigital Image Processing Methods", International Journalof Latest Trends in Engineering & Technology", 2014, Volume 3

[16] Moffy Vas, Amita Dessai, "Lung cancer detection system using lung CT image processing", 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)

[17] Nidhi S. Nadkarni, Sangam Borkar, "Detection of Lung Cancer in CT Images using Image Processing", 2019, 3rd International Conference on Trends in Electronics and Informatics (ICOEI)

[18] Janee Alam, Sabrina Alam, Alamgir Hossan, "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifier", 2018, International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)

[19] Sayali Satish Kanitkar, N. D. Thombare, S. S. Lokhande, "Detection of lung cancer using marker-controlled watershed transform", 2015, International Conference on Pervasive Computing (ICPC)

[20] Anita Chaudhary, Sonit Sukhraj Singh, "Lung Cancer Detection on CT Images by Using Image Processing", 2012, International Conference on Computing Sciences

[21] S. Kalaivani, Pramit Chatterjee, Shikhar Juyal, Rishi Gupta, "Lung cancer detection using digital image processing and artificial neural networks", 2017, International conference of Electronics, Communication and Aerospace Technology (ICECA), Volume: 2

[22] N. Werghi, C. Donne, F. Taher, H. Alahmad, "Segmentation of sputum cell image for early lung cancer detection", 2012, IET Conference on Image Processing (IPR 2012)

[23] Lilik Anifah, Haryanto, Rina Harimurti, Zaimah Permatasari, Puput Wanarti Rusimamto, Adam Ridiantho Muhamad, "Cancer lungs detection on CT scan image using artificial neural network backpropagation based gray level co-occurrence matrices feature", 2017, International Conference on Advanced Computer Science and Information Systems (ICACSIS)

[24] Qing Wu, Wenbing Zhao, "Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm", 2017, International Symposium on Computer Science and Intelligent Controls (ISCSIC)

[25] Kyamelia Roy, Sheli Sinha Chaudhury, Madhurima Burman, Ahana Ganguly, Chandrima Dutta, Sayani Banik, Rayna Banik, "A Comparative study of Lung Cancer detection using supervised neural network", 2019, International Conference on Opto-Electronics and Applied Optics (Optronix)

[26] K. Punithavathy, M.M. Ramya, Sumathi Poobal, "Analysis of statistical texture features for automatic lung cancer detection in PET/CT images", 2015, International Conference on Robotics, Automation, Control and Embedded Systems (RACE)

[26] https://www.kaggle.com/jesyfax/lung-cancer-detection