🌐 *United International Journal of Engineering and Sciences* 🌐
*(UIJES – A Peer-Reviewed Journal); ISSN:2582-5887 | Impact Factor:8.075(SJIF)*
📓 *Volume 5 | Special Issue 1 | 2025 Edition*
*National Level Conference on "Advanced Trends in Engineering Science & Technology" – Organized by RKCE*

# A Review On Early Detection Of Cancer Using AI

**O. Teja[1],K.Pavani[2],K. Monika[3], P. Rajasri[4]**

*UG Scholars, Department of Computer Science And Engineering, RK College Of Engineering Vijayawada, India.*

obillaneniteja@gmail.com[1],pavanikonakalla303@gmail.com[2],raji999@gmail.com[3] ,mounikakatariadv@gmail.com[4]

***Abstract*- The identification of lung cancer at the early stage is very demanding and difficult task due to the construction of the cell. The cancer grows in the body when cancerous cells start to develop uncontrollably. The image processing plays vital role in the prediction of lung cancer at early stage which is also helpful in treatment to avoid the lung cancer. This proposed system is developed to detect lung cancer at early stage with the help of image processing techniques and artificial neural network classifier to design computer based diagnosis system. In this system, during the preprocessing step, several image enhancing techniques, masks are applied using morphological operations and thresholding technique, which eliminates background and surrounding tissue. Region of interest (ROI) is calculated using region based segmentation algorithm. Circle fit algorithm is used to extract the desired nodule. Radius, Mean Intensity, Area, Euler Number and ECD features are extracted in feature extracting step. Finally, Back propagation algorithm is used to train Artificial Neural Network (ANN) in categorization stage.v**

**Key Words-Cancer, Diabetics, Diagnosis, patients.**

## I. INTRODUCTION

The main cause of lung cancer is growth of cells in lung tissue which is irregular and out of control. One of the reasons is smoking. If it is detected earlier, then there will be a good chance of curing. Screening is the one of the important step for lung cancer detection. Screening is the process used to detect and identify the nodule. A nodule appears as round and white in color on a Computed Tomography scans images or an chest X-ray.

[1] There are two types of nodules one is a benign and second one is a malignant. A nodule with diameter 3 cm or less is called a Pulmonary or non-cancerous nodule. These nodules are also called as benign. A nodule whose diameter is larger than 3 cm is poisonous and called as malignant nodule. Malignant nodule should be identified as early possible because it is likely to be cancerous nodule. To check whether these nodules are expanding, they are needed to be observed over the time. If there is a change in the size of nodule and it is growing, then there is a probability of getting cancer. So, a nodule should be observed.

[2] As compared with other types of cancer, the long-term term endurance rate of lung cancer patient is very lo w. So, the identification of lung cancer at an early stage is very important and it provides vital research platform in medical image processing field.

## II. LITERATURE SURVEY

The detection and segmentation of lung nodules based on computer tomography images (CT) is a basic and significant step to achieve the robotic needle biopsy[3]. In this paper, we reviewed some typical segmentation algorithms, including thresholding, active contour, differential operator, region growing and watershed. To analyse their performance on lung nodule detection, we applied them to four CT images of different kinds of lung nodules [4]. The results show that thresholding, active contour and differential operator do well in the segmentation of solitary nodules, while region growing has an advantage over the others on segmenting nodules adhere to vessels. For segmentation of semi-transparent nodules, differential operator is an especially suitable choice [5]. Watershed can segment nodules adhere to vessels and semi-transparent nodules well, but it has low sensitivity in solitary nodulesOur aim is to review and explain the capabilities and performance of currently available approaches for segmentation of lungs with pathologic conditions on chest CT images, with illustrations to give radiologists a better understanding of potential choices for decision support in everyday practice. The computer-based process of identifying the boundaries of lung from surrounding thoracic tissue on computed tomographic (CT) images, which is called segmentation, is a vital first step in radiologic pulmonary image analysis. Many algorithms and software platforms provide image segmentation routines for quantification of lung abnormalities; however, nearly all of the current image segmentation approaches apply well only if the lungs exhibit minimal or no pathologic conditions. When moderate to high amounts of disease or abnormalities with a challenging shape or appearance exist in the lungs, computer-aided detection systems may be highly likely to fail to depict those abnormal regions because of inaccurate segmentation methods. In particular, abnormalities such as pleural effusions, consolidations, and masses often cause inaccurate lung segmentation, which greatly limits the use of image processing methods in clinical and research contexts. In this review, a critical summary of the current methods for lung segmentation on CT images is provided, with special emphasis on the accuracy and performance of the methods in cases with abnormalities and cases with exemplary pathologic findings. The currently available segmentation methods can be divided into five major classes: (a) thresholding-based, (b) region-based, (c) shape-based, (d) neighboring anatomy–guided, and (e) machine learning–based methods. The feasibility of each class and its shortcomings are explained and illustrated with the most common lung abnormalities observed on CT images. In an overview, practical applications and evolving technologies combine the presented approaches for the practising.

## III. Methodology

The median filter is generally used to diminish noise in an image. In the image, the median filter checks its nearby pixels to decide whether that neighbouring pixel is similar or not. In this filter it replaces pixel value with its neighboring median pixel values. Histogram equalization technique is used to adjust image intensity to enhance contrast. It is the graphical interpretation of the image's pixel intensity values. It can be interpreted as the data structure that stores the frequencies of all the pixel intensity levels in the image.

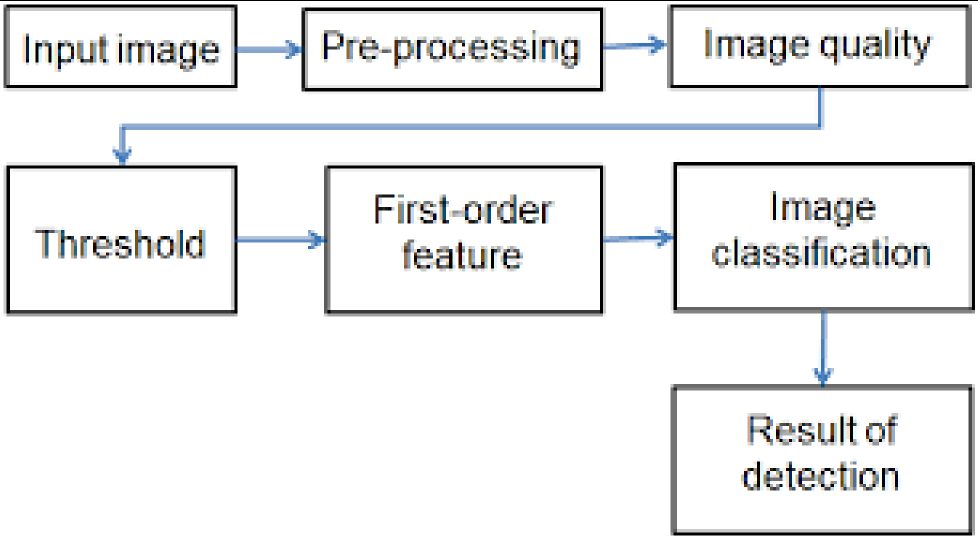



Fig 1. Methodology of early cancer detection using AI

The Fig 1 methodology of early cancer detection using AI involves a combination of data collection, data preprocessing, feature extraction, model training, and model evaluation.

**Step 1: Data Collection**

**1. Medical Images**: Collecting medical images such as MRI, CT scans, mammography images, and histopathology slides.

**2. Clinical Data:** Collecting clinical data such as electronic health records (EHRs), medical history, and demographic information.

**3. Genomic Data:** Collecting genomic data such as genetic mutations, gene expression data, and copy number variations.

**Step 2: Data Preprocessing**

1. Data Cleaning: Removing missing or erroneous data.

2. Data Normalization: Normalizing data to a common scale.

3. Data Transformation: Transforming data into a suitable format for analysis.

**Step 3: Feature Extraction**

1. Image Features: Extracting features from medical images such as texture, shape, and intensity.

2. Clinical Features: Extracting features from clinical data such as age, sex, and medical history.

3. Genomic Features: Extracting features from genomic data such as genetic mutations and gene expression levels [5].

**Step 4: Model Training**

Machine Learning Algorithms: Training machine learning models using algorithms such as convolutional neural networks (CNNs), support vector machines (SVMs), and random forests.

**3.2 Dataset Description**

Data is a fundamental stage in data mining to improve data efficiency.

Datasets for early cancer detection using AI vary depending on the type of cancer and detection method. Here are some notable ones:

National Lung Screening Trial (NLST) Dataset: This dataset contains 42,290 CT scans from 14,851 patients and has been used to train AI models for lung cancer detection, achieving 95.5% accuracy in malignancy risk prediction.

Breast Cancer Datasets: Researchers have used datasets containing over 1 million mammograms from 500,000 women to develop AI models for breast cancer detection, achieving comparable accuracy to radiologists.

Cervical Cancer Datasets: Studies have utilised datasets with images and clinical data to develop deep learning models for cervical cancer screening, achieving reproducible and clinically translatable resultsFig.2. Accuracy Comparison Graph.

Pancreatic Cancer Datasets: Researchers have developed machine learning models using datasets with clinical and genomic data to predict pancreatic cancer risk.

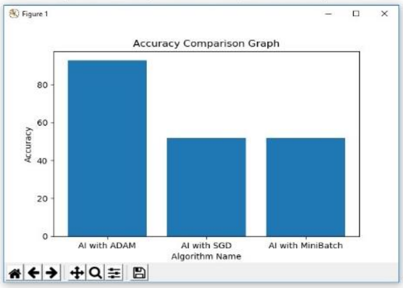
Skin Cancer Datasets: AI models have been trained on datasets with images of skin lesions to detect skin cancer, achieving high accuracy and sensitivity

## Iv. Results

### Improved Accuracy

**1. High Sensitivity:** AI models can detect cancer at an early stage with high sensitivity.

| Algorithm Name | Accuracy | Precision | Recall | FSCORE | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| AI with ADAM | 92.85714285714286 | 92.7560944802324 | 93.25670498084291 | 92.91954022988506 | 96.42857142857143 | 100.0 |
| AI with SGD | 51.78571428571429 | 17.261904761904763 | 33.33333333333333 | 22.74509803921569 | 96.42857142857143 | 100.0 |
| AI with MiniBatch | 51.78571428571429 | 17.261904761904763 | 33.33333333333333 | 22.74509803921569 | 100.0 | 0.0 |




**2. High Specificity:** AI models can reduce false positives and improve specificity.

**3. Accurate Diagnosis:** AI models can provide accurate diagnoses, reducing errors.

Early Detection

**1. Detection at Early Stage:** AI models can detect cancer at an early stage, improving treatment outcomes.

**2. Improved Survival Rates:** Early detection can lead to improved survival rates.

**3. Reduced Mortality:** Early detection can reduce mortality rates.

**4. Personalised Treatment:** AI models can provide personalised treatment recommendations.

**5. Targeted Therapy:** AI models can help identify targeted therapy options.

**6. Improved Patient Outcomes:** Personalised medicine can lead to improved patient outcomes.

## V. Conclusion and Future Scope

### 5.1. Conclusion

The CA D Systems are beneficial to detect cancerous nodules& have a lot to offer to modern medicine. A nodule is identified with required area by using a circle fit algorithm with a maximum radius, which eliminates the unnecessary selection of wrong nodules. After every iteration, we get more accurate results. This led the system to provide an Accuracy of 95.6%. The Sensitivity & Specificity of the system are 93.1% &100% respectively. Based on CT images, this system will give accurate and effective results of lung nodule detections, including benign or malignant lung nodules. In Future work, this system will help to diagnose cancer in different organs of the human body. Techniques used in this system can be implemented in reducing the growth of abnormal cells or spreading to other parts of the body. This system can be enhanced for MRI and Ultrasound images. The results obtained from A NN classifier are more precise and accurate but it requires more number of data inputs as compared with SVM classifier.

### 5.2. Future Scope

Artificial intelligence (AI) is revolutionising early cancer detection, offering promising advancements in diagnostic accuracy and accessibility.

Future directions: Multicenter Detection Tools:  AI systems like Grail's Gallery test can detect over 50 types of cancer from a single blood sample.

Real-Time Assistance: AI provides real-time support to doctors, flagging potential cancerous areas and speeding up the diagnostic process.

Digital Pathology: AI analyses tissue samples, classifying and grading cancer cells with higher accuracy and consistency than human pathologists[31].

## VI. **REFERENCES**

[1]World Health Organisation.Cancer: fact Sheet no. 297. 2015 July 8. http://www.who.int/mediacentre/factsheets/fs297/en/.Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2015. CA Cancer J Clin.2015; 60(5):277–300.

[2] The diagnosis of lung cancer (update) Published by the National Collaborating Centre for Cancer (2nd Floor, Front Suite, Park House, Greyfriars Road, Cardiff, CF10 3AF) at Velindre NHS T rust, Cardiff, Wales [2011] Database from: The Cancer Imaging Archive.

[3] Nanusha, "Lung Nodule Detection Using Image Segmentation Methods", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 6, Issue 7, July 2017

4] Awais Mansoor Ph.D. et al, "Segmentation and Image Analysis of Abnormal Lungs at CT: Current Approaches, Challenges, and Future Trends", Radio Graphics 2015; 35:1056–1076 Published online 10.1148/rg.2015140232

[5] Ozge Gunaydin et al "Comparison of Lung Cancer Detection Algorithm" 2019, 978-1-7281-1013-4/19/ $31.00 © 2019 IEEE