**Artificial Intelligence in Medical science: Early Lung Cancer Cell Detection Using Deep Learning**

 **Abstract:**

Artificial Intelligence is playing a vital and crucial role now a days in medical science. It is the evolution which is leading to drastically change the face of medical field. This chapter mainly contributes the work enhancement done in lung cancer. Lung carcinoma is the major cause of death throughout the world, nearly 18% deaths are due to cancer as the survival rate is very low. The rise of Artificial Intelligence (AI) and Machine Learning (ML) techniques and their applications in various fields have brought immense value in providing insights into advancement in support of cancer control. Deep learning is a significant and emerging AI strategy for the innovative change in healthcare domain. Deep learning is a type of machine learning that uses a layered algorithmic architecture for analyzing data. Deep learning algorithms are applied for detecting lung cancer cell in its early stages. This will be helpful for radiologist and doctors for proper treatment of the patients to cure. Advancement in the DL decreasing the mortality of the lung cancer. CNN is playing the vital role and enhancing the performance of the system. The recent development in Deep learning helps to early detect lung cancer finally beneficial to the society.

**Artificial intelligence**

Artificial intelligence (AI) has recently made significant strides in computer science and informatics, and it is now becoming a crucial component of contemporary healthcare. Medical practitioners are supported by AI algorithms and other applications powered by AI in clinical settings and current research. The medical sciences make substantial use of computer systems with artificial intelligence. Common uses include remote patient treatment, prescription transcription, increasing doctor-patient communication, medication discovery and development from beginning to end, and patient diagnosis. Modern computer algorithms have lately achieved accuracy levels that are comparable to those of human experts in the field of medical sciences, despite the fact that computer systems frequently perform jobs more quickly than humans do. Some believe that humans will soon be fully replaced by machines. The goal of this essay is to examine how artificial intelligence is altering the field of medical science and to distinguish marketing hype from fact. AI is a broad concept that lacks a clear definition. An AI programme uses historical data to forecast or categorise objects [8]. The dataset utilised for training, the pretreatment technique, the algorithm used to produce the prediction model, and the pre-trained model—which speeds up the modelling process and incorporates prior knowledge—are the key components. The study of developing algorithms to solve problems without being explicitly programmed, such as decision trees (DTs), support vector machines (SVMs), and Bayesian networks (BNs), is known as machine learning (ML), which is a subtype of artificial intelligence (AI). Deep learning is a further subclass of ML, featured with multiple layered ML, achieving feature selection and model fitting at the same time [[9](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8946647/#B9-cancers-14-01370)].

**Lung cancer:**

Cancer is still viewed as a hazardous disease with high mortality rates. Comparing all cancers, lung cancer has the greatest mortality rate or death rate [1, 2], and it is thought to be the deadliest carcinoma globally [3]. As a result, numerous researchers are concentrating on approaches for exploiting digital pictures, particularly Computed Tomography (CT), to detect lung cancer nodules. The use of X-rays in CT scans produces many pictures, making it difficult for radiologists to identify minute nodules from these images [4]. The fundamental task carried out by the radiologist for the diagnosis of lung cancer is the analysis and interpretation of nodules.

Numerous scientists and researchers are developing automated medical procedures that will save time and money [5–9]. The size and form of nodules, which can be classed as benign or malignant, usually provided the first indication of cancer in the majority of instances. Lung nodules less than 3 cm in diameter are typically regarded as benign nodules, but those greater than 3 cm are thought to be malignant or lung masses.

Genomic instability is a trait of lung cancer, a malignant tumour of the lungs. among 2020, 1.8 million individuals died from lung cancer worldwide [1], making it the second-most prevalent cause ocancer- related mortality among men

.

Fig 1:- Lung cancer cell

Cancer risk can be calculated based on nodule classification and other data. The primary level identification and categorization of distinct cancer kinds is where AI approaches play a significant role [10–14]. In recent years, deep learning (DL) models have been applied in a variety of industries, such as games, agriculture, and medical [15]. DL models perform admirably in each of these areas, especially in particular applications like picture segmentation, object recognition, and image classification [16–17[DL] is an area of artificial intelligence that uses interconnected nodes to carry out challenging tasks. DL algorithms are able to learn from training data rather than using preprogrammed instructions [13]. Deep learning has previously been used by many researchers to identify cancer [18–20].

|  |  |  |
| --- | --- | --- |
| S. No.  | Type of Cancer | Total cases (in millions) |
| 1 | Breast | 2.26 |
| 2 | Lung | 2.21 |
| 3 | Colon and rectum | 1.93 |
| 4 | prostate | 1.41 |
| 5 | skin (non-melanoma) | 1.20 |
| 6 | stomach | 1.0 |

Table 1:- Cancer cases

Due to evolution and different treatments lung cancer is classified in 2 categories:

• Non small cell carcinoma (NSCC): Most of the people who are having lung cancer have NSCC. Peoples having more than 65 years age and who smoke or who breathe a lot of smoke are most likely to get non small cell carcinoma.

• Small cell carcinoma (SCC): In small cell carcinoma cells of the lung start growing rapidly in an uncontrolled manner

A multi-stage method called radiomics makes use of standard medical images to help with clinical diagnosis and prognosis. Gillies et al. [21] described high-throughput quantitative feature extraction as "the conversion of images into mineable data and the subsequent analysis of these data for decision support." A tool called computer-aided diagnosis (CAD) helps doctors and radiologists evaluate diagnostic image data in order to identify or monitor a condition. Both radiomics and CAD use imaging methods such as computed tomography (CT) scans, magnetic resonance imaging (MRI), positron emission tomography (PET), X-rays, and others. Radiomics, a subfield of computer-aided design (CAD), on the other hand, consists of high-dimensional features extracted from medical images and used to interact with clinical outcomes, create models for hypothesis testing and decision support, identify, monitor, and diagnose disease, among other things [21–22]. Radiomics traits can be integrated with clinical data (sex, smoking status, age, etc.) or gene expression to enhance decision support analysis.

The National Cancer Institute (NCI) and the Quantitative Imaging Network (QIN) have supported radiomics' progress in the field of cancer, even though it can be used for any condition. The Radiomics workflow involves gathering patient images, locating and recovering regions of interest (Segmentation), creating high-dimensional characteristics from the segmented area, and then using machine learning models for classification or prediction. Medical photos that meet a patient's standard of care are obtained. A single or a number of regions of interest (ROI) are selected from the patient imaging. The larger tumour or a smaller tumour within it (referred to as a habitat) may very well represent the ROI. The ROI can then be utilised to extract deep radiomics, semantic, or high-dimensional agnostic properties.

Shape, size, margin, speculation, boundary, and external characteristics, such as pleural attachment and fissure attachment, all appear to be characterised in terms of semantic elements; shape characteristics include lobulation, concavity, and the like; size characteristics include long-axis diameter and short-axis diameter. An expert radiologist described semantic features of the ROI that provide important details about tumour characteristics [23].

**IDENTIFICATION OF LUNG CARCINOMA**

The identification of lung cancer is made in the following ways:

 • History and Physical examination

• Diagnostic tests

• Staging tests

Among the diagnostic procedures is sputum cytology. During sputum cytology, sputum samples are collected and examined under a microscope to check for any cancer cells. A biopsy (bronchoscopy, needle biopsy, surgery) is the removal of a tiny sample of tissue from a suspected site for examination and analysis under a microscope.

* chest X-ray It is the first imaging technique to look for lung cancer. Nodules greater than 1 cm are frequently visible on X-ray images. In the event that the chest X-ray exhibits any irregularities, patients may be advised to undergo staging testing.
* X-rays X-rays are produced by the radiation emitted by electromagnetic waves. As the images are being created, X-rays are used to highlight interior body regions. Radiation varies in different places of the body because different tissues absorb radiation in different ways.
* Computed tomography, or CT Computed tomography, commonly known as CAT or computerised axial tomography (CT), is frequently used in image processing techniques. The benefits of CT are numerous.
* Radiography Similar to X-rays, radiography is a highly general term. The two radiographic image types employed in medical imaging are fluoroscopy and projection radiography. This imaging technique, the first one offered in modern medicine, takes pictures using a wide x-ray beam.

The amount of metastases is evaluated via a PET scan, commonly referred to as positron emission tomography. The term "metastasis" refers to the initial malignancy spreading from one organ to another. PET scanning is a recent technological advancement.

* Magnetic-resonance imaging A big magnet is used in MRI scans to create 3-D images. Any region in particular that could not be well interpreted on a CT scan can be studied using MRI. Examining the participation of cancer cells with MRI is helpful. Brain tumour identification often involves the use of a magnetic resonance imaging technique. When a nuclear magnetic resonance (NMR)
* Sonographic images Medical imaging creates images by using high-frequency, wide-band sound waves in the Megahertz (MHz) range that are reflected differently by tissue.

Identification of the Lung Carcinoma Problem Radiologists retrospectively review lung scans for indications of cancer. The region of interest that is not suspicious enough to recall can be used to spot an anomaly in photos. When radiologists handle lung pictures, there are two basic issues that arise. The radiologist may have overlooked the symptom and not given it any attention, which is the first potential issue.

The second scenario involves looking at the indicator but deciding it is innocuous, typical, or not worrisome enough to warrant further investigation. In addition to these issues, there are further difficulties with mass detection.

They are listed as follows:

• The intensity levels in different lung areas vary substantially, making it challenging to define characteristics for segmentation.

• The grey level segmentation is challenging due to subtle variations in grey level across the image. • Tumours are sometimes difficult to see.

• The image's poor lighting and significant noise levels, which can range from 10-15% of the maximum pixel entity.

following breast cancer, the leading cause of cancer-related death in women [24]. Annual low-dose computed tomography (CT) screening decreased lung cancer mortality by 20% compared to chest radiography, according to the National Lung Screening experiment (NLST), a randomised control experiment including more than 50,000 high-risk patients in the United States [1]. As a result, low-dose CT scanning initiatives have started to be implemented in the US and other countries. It's one of the most challenging features of CT for radiologists to evaluate so many images. There could be up to 600 slices in a CT scan. The analysis of such massive amounts of data presents a considerable problem for radiologists.

**Role of Deep Learning in cancer**

Artificial intelligence (AI) is the process of using computers and technology to mimic human-like intelligent behaviour and critical thought. The term was first used by John McCarthy in 1956 [1]. Additionally, it gives us the ability to assess and grasp complex medical data, aiding in diagnosis, management, and the prediction of treatment outcomes across a range of clinical presentations. The medical sector could undergo a fundamental transformation thanks to artificial intelligence. Because of the accessibility of digital data, machine learning, and computer infrastructure, AI applications have been able to expand into fields that were previously believed to be unachievable with

Radiography is a key area for the early use of AI techniques in the medical industry, claims Tang. It is anticipated that the use of AI will considerably increase the breadth, quality, and value of radiology's contribution to patient care and public health over the coming ten years. Workflow for radiologists is anticipated to alter and evolve significantly. One class of representation learning methods called deep learning uses picture data to learn hierarchical feature representation. Deep learning has the benefit of being able to produce high level feature representations straight from the original image data.

Additionally, parallel architectural deep learning approaches have improved accuracy in many sectors recently, including image identification, object detection, and speech recognition, with the use of Graphic Processing Units (GPUs). For instance, current research indicates that CNNs deliver guaranteed performance in cancer diagnosis and detection [18].

**Deep learning Approaches**

According to reports, a deep learning-based CAD system holds great promise for the accurate automatic identification of lung illness in medical imaging [34,35,36]. A neural network model with numerous levels of data representation makes up the deep learning model. Unsupervised, reinforcement-based, and supervised learning are the three subcategories of deep learning methodologies.

Unsupervised learning, which analyses the data and then organises intrinsic similarities between the input data, does not require user instruction. As a result, semi-supervised learning is a mixed model that, despite various difficulties, can produce a win-win situation. Techniques for semi-supervised learning employ both labelled and unlabeled data. Both labelled and unlabeled data are used to increase the decision boundary's accuracy. In terms of clustering and nonlinear dimensionality reduction, Auto-Encoders (AE), Restricted Boltzmann Machines (RBM), and Generative Adversarial Networks (GAN) excel. Training typically requires a lot of labelled data, which increases cost, effort, and difficulty. To eliminate labelling and create a more reliable model, researchers have used deep clustering [39,40].

The most popular unsupervised learning algorithms in medical pictures are convolutional neural networks (CNN), deep convolutional neural networks (DCNN), and recurrent neural networks (RNN). Because less pre-processing is needed, CNN architecture is one of the most used supervised deep learning methods for lesion segmentation and classification. Recently, CNN architectures have been used to classify and segment medical pictures, such as Mask R-CNN [41] and AlexNet [42] and VGGNet [43]. More layers with complicated nonlinear interactions are typically found in DCNN designs, which have shown to perform regression and classification with respectable accuracy [44,45,46]. A higher-order neural network with RNN design can accommodate network output to re-input [47]. Elman network with feedback links from the buried layer is applied by RNN to reinforcement learning technique was first applied in Google Deep Mind in 2013 [[**48**](https://www.mdpi.com/2072-6694/14/22/5569#B48-cancers-14-05569)]. Since then, numerous studies have been conducted to enhance the accuracy, sensitivity, and specificity of lung cancer detection using reinforcement learning techniques. Labelled datasets are used in semi-supervised learning techniques like deep reinforcement learning and generative adversarial networks.

A learning algorithm is typically used in supervised learning, and labels are applied to the input data in accordance with the labelling data collected during training. On CT images, a number of supervised deep learning techniques have been used to detect anomalies with anatomical localization. These methods have some shortcomings, including the need for a significant amount of labelled data during training, the assumption of fixed network weights after training, and the inability to improve upon results after training. As a result, a few-shot learning (FSL) model is created to minimise the amount of data needed for training.

Due to their exceptional performance, convolutional neural networks (CNNs) have lately grown in prominence in the machine learning sector. Neurons with programmable weights and biases make up CNNs. This method is based on the artificial neural network (ANN) structure, which was modelled after the biological neuron. The fact that CNNs learn filters on their own gives them an advantage over conventional neural networks. The learnable filters that make up the CNN layer parameters enable the system to adapt to new problems. These filters use a convolutional method to extract spatial information from the incoming data. Because of this, CNNs excel at a variety of tasks, including object detection, video analysis, speech recognition, natural language processing, and medical image analysis [34–36]. But in order to prevent overfitting,

A CNN has been arranged in the form of the layers.

• ReLU layers

• Convolutional layers

 • Pooling layers

• a Fully connected layer



Fig 2:- Convolutional Neural Networks

* Convolutional layer The primary goal of this convolution phase is to concentrate the highlights in the data image. Consistently, CNN starts with the convolutionary layer. The features are extracted from the input image in this procedure, and a feature map is created.
* ReLU The straight redressed unit layer comes after the convolution layer. Here, the feature maps are used to implement the enactment function, which increases the non-linearity in the network and makes it simple to remove negative values.
* Pooling layer The size of the input can gradually be decreased by the pooling process. The pooling stage can reduce overfitting. By increasing the amount of parameters required, it will quickly identify the necessary ones.
* Fully connected layer Here are all the traits and their matched attributes. It is possible to use the classification approach with significant percentile imprecision. The primary methods of measuring and recording the inaccuracy
* Softmax layer:-To convert the abnormal network activity over expected performance groups to a probability distribution in neural grids, softmax is also used. The Softmax has been used to solve numerous problems in numerous academic domains. The decimal's probability will imply 1.0. Consider the related Softmax versions, such as Full Softmax, which has the ability to determine the chance of each potential class. To categorise the images, CNN employs successive convolutional and pooling layers.

The CNN's pooling layer decreases the dimension and categorises the object without respect to its spatial information, or where the object is actually situated in the image [40]. The benefit and the disadvantage of CNN's pooling feature are both present. Some crucial information that is crucial for object detection and image segmentation is lost during the pooling operation.

CNN is having pretrained networks and transfer learning which again improves the system performance in terms of accuracy, precision, recall, F1 score.

Chapter summary- This chapter covers the importance of AI in medical science related to detecting lung nodule in early stages. The traditional methods for lung cancer cell detection and their pitfalls .The advancement of AI like Machine learning, Deep Learning how these approaches helping the radiologist, doctors and finally to the patients to survive. Deep Learning approaches like Convolutional Neaural Networks , Recurrent Neural Network, LSTM has a great impact in medical field. CNN is preferred most for lung cancer cell detection and classification process, whether the cell is benignant or malignant.