An Adaptive Neuro-Fuzzy Inference System-Based Lung Cancer Detection System

M. Prema Kumar[1], K. Padma Vasavi[2], M. V. Subba Rao[3], G. Challa Ram[4], D Ramesh varma[5]

[1] Department of ECE, Shri Vishnu Engineering College for Women (A), Bhimavaram, Andhra Pradesh, India medapatipremakumar@gmail.com

[2,3,4,5] Department of ECE, Shri Vishnu Engineering College for Women(A), Bhimavaram, Andhra Pradesh, India

Abstract: In the last forty years, there have been dramatic growths in the area of medical and healthcare systems. During this period, the true causes of several illnesses were uncovered, fresh diagnostic techniques were created, and new medications were created. Despite these advances, illnesses like cancer continue to plague us because people remain susceptible to the system. Cancer is the second-biggest reason of mortality worldwide, accounting for around one in every six deaths. Therefore, early illness detection considerably increases the likelihood of survival. Among all cancers, lung cancer has the highest risk. Using the capabilities of Artificial Intelligence (AI), tumour diagnosis may be automated to analysis larger capacity in lesser time and at a lesser cost.

A Machine Learning based Lung Cancer Detection System (ML-LCDS) is suggested in this research. The automated identification and localization of tumour locations in lung imaging are increasingly crucial for saving patients' lives via prompt medical therapy. In this study, a lung tumour detection, categorization, and segmentation method based on machine learning are suggested. The tumour categorization stage first applies an adaptive median filtration to the original lung computerised tomography picture and then applies Discrete-Timing Complicated Wavelet Transformation (DT-CWT) to divide the whole picture into several sub-bands. In addition to the deconstructed sub-bands, Discrete Wavelet Transformation (DWT), and co-occurrence characteristics are calculated and identified using an ANFIS. The tumour segmentation stage detects tumour locations on this identified abnormal lung image using morphological features. The proposed system exhibits a precision of 93.4%, accuracy of 95.1%, specificity of 90.6%, sensitivity of 92.8%, False positive rate of 0.22%, false negative ratio of 0.18%, and classification accuracy of 98.2%. The outcomes of the simulation show that the proposed system for finding and predicting lung cancer is accurate and precise. The proposed method outperforms all methods and provides better lung cancer detection accuracy than others.

Keywords: Lung Cancer, Neuro-fuzzy, Tumour Classification, Tumour Prediction, Feature Selection, Feature Extraction.

1. INTRODUCTION

Over the past few decades, noteworthy progress has been achieved in medical and healthcare systems. This has resulted in the discovery of fundamental etiologies of diverse ailments, the emergence of novel diagnostic methodologies, and the inception of creative therapeutic interventions. The timely identification of cancer is paramount in enhancing the likelihood of survival among patients. Lung cancer is a type of cancer that presents a significant risk and requires effective and precise detection methods when compared to other forms of cancer.

Cancer is a group of illnesses characterised by the development of aberrant cells inside the human as an outcome of genetic changes [1]. When tumour cells are created, they proliferate uncontrollably and spreading to the entire organs. In the absence of treatment, the majority of cancer cases may end in death. Following cardiovascular illnesses, tumour is the primary reason of mortality globally [2]. More than 18M new tumour reasons and 9.55M mortality were recorded globally in 2018. There will be over 1.8M new cancer diagnoses and 608,000 mortalities in 2020. According to statistics gathered between 2015 and 2018, two out of five Americans will be treated for cancer throughout their lifespan. Cancer cells may form in any region of the body, with the lungs, ovaries, brains, intestines, rectum, kidney, abdomen, epidermis, and prostate is the most often afflicted organs [3]. Lung and colon tumours are responsible for the biggest number of fatalities. That was reasonable for approximately 2.5 million fatalities and over 2.9 million new

diagnoses in the United States alone. There are several causes of cancer, varying from behavioural factors, such as higher body mass indices, such as exposure to Ultraviolet (UV) and irradiation, specific biological and genetic carcinogens [4]. Nonetheless, the reason changes from patient to patient. Discomfort, exhaustion, nausea, a dry cough, clotting, and bruises are among the most common cancer complaints. Without a diagnostic technique such as a Computed Tomography (CT) scan [5], Magnetic Resonance Imaging (MRI) [6], Positron Emission Tomography (PET) view [7], ultrasonography, or a biopsy, it is complex to confirm the existence of cancer. In many situations, patients show small or no signs in the beginning phases, and by the time indications manifest, it is often too late.

A lung tumour is the most prevalent kind of cancer in the world, and it is caused by the development of aberrant patterns or abnormal cell proliferation in the lung area. In the US, lung tumours have a higher mortality rate than other cancers and tumours, such as lung tumours. The area on the lung is called a nodule, and it is visible on the CT scan. The aberrant development of cells is the primary cause of lung tumour formation. Common lung tumour symptoms include a strong cough with breathing difficulties and persistent chest discomfort. Lung tumours may be classified as cancerous or benign [8]. The benign lung tumours have begun to produce irregular patterns in the patches of the lung pictures. The primary cause of aggressive lung tumours is the intense creation of aberrant patterns surrounding the nodules of the lungs.

Machine Learning (ML) has several applications in pathology, varying from the diagnosis of illnesses to expert machines that, based on a patient's symptoms, may recommend traditional medications [9]. This field is still in its adolescence, and a great deal of study must be conducted before such applications may be used in clinical settings. Furthermore, it demonstrates the potential of Artificial Intelligence (AI) and suggests how it will be utilised in the medical industry in the next few years [10].

The need for the proposed research:

- To improve the likelihood of a successful course of therapy and patient survival rates, early identification of lung cancer is essential.
- To increase precision and efficiency, innovative computational algorithms must be developed due to the complexity and unpredictability of lung cancer imaging data.
- In the examination of medical imaging, AI and machine learning have shown great promise, allowing automatic and impartial tumor diagnosis.

The proposed adaptive neuro-fuzzy inference system addresses the shortcomings of conventional methods by combining the advantages of fuzzy logic, neural networks, and machine learning to improve the accuracy and resilience of lung cancer diagnosis.

Identifying and assessing pulmonary neoplasms present complex challenges owing to the heterogeneous and inconspicuous characteristics of tumor manifestations in lung radiographic information [38]. Conventional diagnostic techniques depend on human proficiency and are frequently associated with prolonged duration, high cost, and susceptibility to human fallibility. Consequently, an urgent requirement exists for automated systems that can effectively and precisely detect and locate lung tumors, facilitating prompt medical intervention and early diagnosis [39].

The motivation for the research:

- Clinical Implications: Early identification of lung cancer is essential for improving patient survival rates and the likelihood of effective therapy. The suggested research focuses on a significant clinical need, possibly saving lives, minimizing patient suffering, and advancing healthcare systems by creating an accurate and effective lung cancer detection system.
- Improve accuracy: The proposed system aims to boost lung cancer detection accuracy by utilizing AI and ML capabilities.
- Efficiency in terms of time and cost: By automating the identification of lung cancer, the suggested approach drastically cut down on analysis time, allowing for quick medical action and perhaps improving patient health. ML methods lower the cost of manually diagnosing and interpreting lung imaging data.

- Thorough analysis: Integrating ANFIS combines the advantages of fuzzy logic, neural networks, and ML to analyse lung imaging data. By capturing intricate patterns and changes in lung cancer photos, this technique improves the identification and classification of the disease.

An Adaptive Neuro-Fuzzy Inference System is employed in the proposed system [40]. The process involves a series of stages, namely tumor identification [41], classification [42], and segmentation [43], to achieve a thorough and dependable detection of lung cancer. The utilization of machine learning methodologies and artificial intelligence is employed to scrutinize vast quantities of lung imaging data, thereby augmenting the ability to identify tumors within a compressed time frame and at a diminished expense [44].

The ML-LCDS system under consideration integrates various technical features to augment its efficiency and precision. The proposed method encompasses:

- The adaptive mean filter is utilized to eliminate noise and unnecessary elements from lung CT images, enhancing overall image quality.
- The Gray-level Co-occurrence Matrix is utilized for attribute selection, thereby improving the identification of pertinent features for tumor detection.
- The Discrete-Time Complex Wavelet Transformation is employed to extract features from CT images. This method facilitates the production of clear lung data, which can be useful for tumor classification and segmentation.
- In the ultimate stage of the system, a Deep Belief Network is utilized to perform image categorization, differentiating between images of normal lung tumors and those that are malignant.

The ML-LCDS system suggested in this study seeks to attain elevated levels of precision, accuracy, specificity, sensitivity, and classification accuracy, surpassing current methodologies and yielding superior outcomes in the detection of lung cancer.

The primary contributions are listed below:

- The adaptive mean filter eliminates the noise or unwanted components in the CT lung images.
- Gray-level Co-occurrence Matrix (GLCM) is utilised to enhance the available features for the attribute selection process.
- Discrete-Time Complex Wavelet Transformation (DT-CWT) is used for feature extraction from the CT images and for producing crip lung data.
- Deep Belief Network (DBN) is used for image classification to find the malicious and normal lung tumour images at the last phase.

2. BACKGROUND TO THE CANCER DETECTION AND CLASSIFICATION METHODS

A precise examination of lung nodules is required to evaluate malignancy and the likelihood of lung cancer. Several attempts are being made to construct a comprehensive and scalable tumour segmentation algorithm to aid radiologists. The attempts may be divided into traditional image analysis approaches and machine learning methods. This section gives an overview of newly suggested approaches in each area.

Malik et al. introduced a novel approach, which is a Convolutional Neural Network (CNN) model with multi-classification capabilities [11]. The model is specifically tailored to classify chest radiographs into three categories: COVID-19, pneumonia, and lung cancer. The experimental findings indicate a classification accuracy of 94.5% for COVID-19, 92.1% for pneumonia, and 91.3% for lung cancer.

Shim et al. introduced a novel approach that employs machine learning and radiomic texture features to forecast the occurrence of lung cancer [12]. The approach centers on prioritizing the significance of features using multimodal radiomic texture features. The novelty of the methodology is attributed to its capacity to extract pertinent features and prioritize their significance in forecasting lung cancer. The method's efficacy was demonstrated through experimental results, yielding a prediction accuracy of 85.6% for lung cancer.

Nanglia et al. introduced a hybridized approach utilizing Support Vector Machines (SVM) and neural networks to classify lung cancer [13]. The proposed approach amalgamates the advantageous features of support vector machines and neural networks to enhance the precision of lung cancer categorization. The experimental findings indicate a classification accuracy of 92.6% and a sensitivity of 92.4% for lung cancer classification.

Khan et al. devised a comprehensive framework for the classification of lung cancer based on Computer Tomography (CT) imaging [14]. The methodology entailed the amalgamation and curation of classical features based on contrast. The innovation of the methodology is rooted in the amalgamation of diverse feature fusion methodologies. The experimental findings indicate a classification accuracy of 92.5% for identifying lung cancer.

The integration of CNN in a lung cancer classification pipeline was proposed by Bonavita et al. [15]. The methodology was centred on the evaluation of malignancy in pulmonary nodules. The experimental findings indicate that the assessment of pulmonary nodule malignancy exhibited a sensitivity of 89.6% and a specificity of 89.2%.

Asuntha et al. introduced a deep-learning approach for identifying and categorizing lung cancer [16]. The method's innovation stems from deep learning to analyze lung images and classify lung cancer effectively. The experimental findings indicate a general precision rate of 94.2% in identifying and categorizing lung cancer.

Leader et al. performed a single-cell analysis on humans' Non-Small Cell Lung Cancer (NSCLC) lesions [17]. The objective of the investigation was to enhance the categorization of tumors and the grouping of patients through the utilization of single-cell analysis. The method's novelty is rooted in its single-cell analysis utilization to provide a more comprehensive comprehension of NSCLC. The study's results yielded advancements in tumor classification and patient stratification.

Bishnoi et al. introduced a transfer learning model for lung cancer classification based on Tensor- Regression Tree (RT) [18]. The employed approach involved utilizing transfer learning and optimization through Tensor-RT to enhance the effectiveness and precision of lung cancer categorization. Based on the experimental results, the proposed model achieved a classification accuracy of 96.7% for lung cancer.

Marentakis et al. devised a technique for classifying lung cancer histology utilizing radiomics and deep learning models [19]. The study employed a hybrid approach that integrated radiomics features and deep learning to classify lung cancer histology based on CT images precisely. The obtained experimental outcomes revealed a classification accuracy of 92.3% in the context of histological classification of lung cancer.

Toda et al. introduced a technique utilizing a semi-conditional Generative Adversarial Network (GAN) to produce synthetic CT images of lung cancer with controlled shapes [20]. The objective of the methodology was to produce authentic artificial CT images of lung cancer with regulated shape attributes. The results indicate that the method utilizing synthetic CT images is effective for classifying lung cancer types, achieving an accuracy rate of 87.2%.

Maleki et al. introduced a k-Nearest Neighbor (k-NN) approach to predict lung cancer prognosis using a genetic algorithm to select relevant features [31]. The technique's novelty is attributed to the amalgamation of k-NN categorization and genetic algorithm-based feature selection, which enhances the prognosis of lung cancer. The suggested approach yielded an accuracy of 87.3% for lung cancer prognosis, as evidenced by the experimental results.

Alsadoon et al. introduced a framework that serves as a tool for assessing the efficacy of deep learning techniques in the timely identification and categorization of lung cancer [32]. The findings of this research make a valuable contribution to the advancement of deep learning-based methodologies that can be utilized for the timely identification and categorization of lung cancer.

Meraj et al. devised a technique for identifying lung nodules through semantic segmentation and classification, incorporating optimal features [33]. The methodology employed semantic segmentation methodologies to identify lung nodules, followed by their classification using optimal features. The proposed method yielded a detection accuracy of 94.6% and a classification accuracy of 91.2% for lung nodules, as evidenced by the experimental results.

Shakeel et al. introduced a novel approach for detecting lung cancer automatically [34]. Their method involved the utilization of discrete AdaBoost-optimized ensemble learning generalized neural networks. The proposed approach entailed the integration of the AdaBoost algorithm for feature selection and ensemble learning technique with

generalized neural networks to detect lung cancer. The suggested approach yielded a lung cancer detection accuracy of 92.5%, as evidenced by the experimental results.

Shakeel et al. introduced a novel approach for detecting lung cancer in CT images through an enhanced deep neural network and ensemble classifier [35]. The approach utilized an enhanced deep neural network framework and an ensemble classifier to identify lung cancer precisely. The proposed method yielded a classification accuracy of 94.7% for the detection of lung cancer, as evidenced by the experimental results.

Traditional approaches have the following limitations:

- Tumour separation serves a crucial role in the identification of tumour severity. Numerous studies have achieved suboptimal tumour segmentation results, which are unsuitable for further tumour severity diagnosis.
- Most research has identified the inner limit of tumour areas in aberrant lung images, which is the primary reason for diagnosing tumour severity levels.
- Conventional approaches evaluated in the research survey focused on detecting tumour areas in aberrant lung pictures using their suggested techniques but needed to expand their technique to diagnose the risk level of segmenting tumour regions in faulty lung pictures.

Research gaps are given below:

- There has been little investigation into combining several imaging techniques for thorough and precise lung cancer diagnosis.
- There hasn't been enough study of the temporal aspect of lung cancer diagnosis, such as using longitudinal imaging information to monitor tumor development over time.
- Insufficient attention is paid to uncertainty estimates and confidence intervals in lung cancer detection designs, despite their importance for clinical and decision-making applications.

3. MATERIALS AND METHODS

The study employed a Chest CT-Scan image dataset to detect and classify lung cancer through experimentation [30]. The dataset was compiled from multiple sources and encompassed a heterogeneous assortment of pulmonary images.

The dataset of Chest CT-Scan images consists of high-resolution CT scans taken of the chest area. The study encompassed a range of lung tumor cases, comprising both malignant and benign instances, thereby offering a comprehensive portrayal of lung cancer scenarios. The images played a crucial role in developing and assessing the proposed technique for identifying lung cancer.

The methodology for identifying and categorizing lung cancer cases integrates diverse techniques from machine learning and computer vision domains. The present study employed various techniques in the development of the suggested system.

- Preprocessing was conducted on the CT-Scan images to eliminate noise and amplify the pertinent features [45]. Adaptive mean filtering and histogram equalization methods were applied to enhance the quality of images and enable the precise identification of lung tumors.
- The methodology suggested in this study utilized sophisticated feature extraction techniques to capture specific information from the CT-Scan images [46]. The present study employed the DT-CWT and GLCM techniques to extract pertinent features that effectively delineate lung tumors.
- A classification model was utilized to differentiate between malignant and benign lung tumors by classifying the extracted features. Several models, such as CNNs [47], DBNs [48], and ensemble classifiers [49], were investigated to attain precise and resilient classification outcomes.

4. PROPOSED MACHINE LEARNING BASED LUNG CANCER DETECTION SYSTEM (ML-LCDS)

In medical picture classification, identifying a patient's ailment from a medical database is a computationally promising task. The present study analysed lung picture classification databases with the primary objective of achieving maximal illness prediction precision. The flowchart diagram of the proposed work is given in Fig. 1:

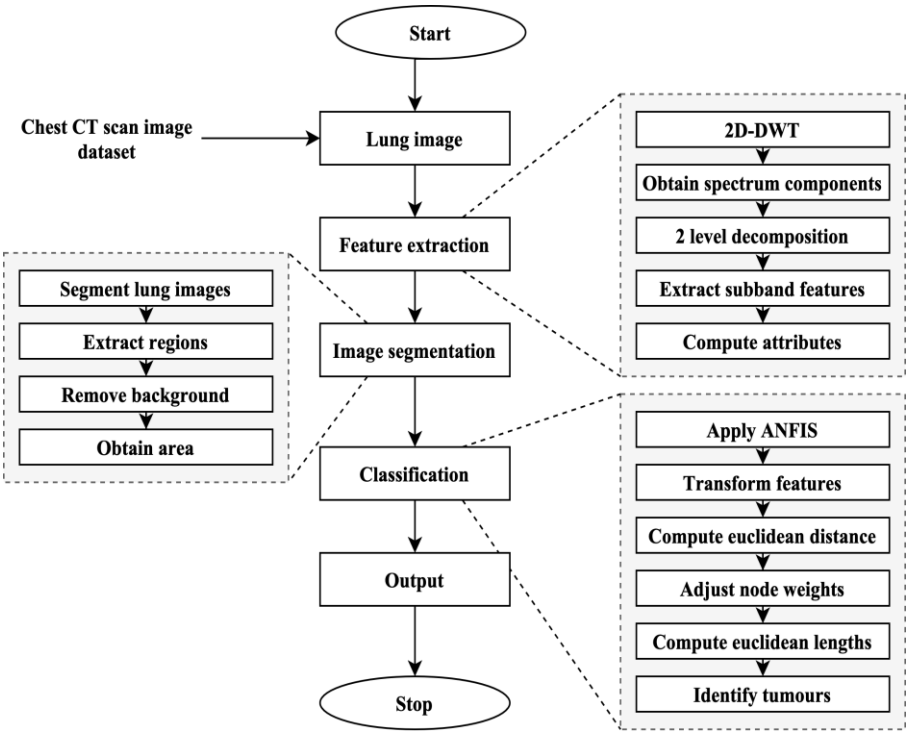

Figure 1. Flowchart of the proposed method

Proposed methodology

- System Model

The methodology known as ML-LCDS can be delineated in the following manner:

- Feature Extraction:

The task is to identify pertinent characteristics from the lung CT images provided, represented as a set of features denoted as $F = \{f_1, f_2, \cdots, f_n\}$. Each feature, f_i, corresponds to a distinct attribute extracted from the pictures.

- ANFIS Categorization Method:
  - Initialization:

Assign a predetermined value, such as 0.8, to the learning rate (η).

Commence by assigning a predetermined value, such as 0.5, to the internal nodes' weights (cap W).

  - Euclidean Distance Calculation:

Compute the Euclidean distance for every source feature vector and record the results in the distance matrix, denoted as $D(i, j)$.

  - Minimal and Maximal Euclidean Distances:

Compute the minimum Euclidean distance (E_1) and maximum Euclidean distance (E_2) based on the provided distance matrix.

  - Calculation of E_1 and E_2:

The updated value of E_1, denoted as E'_1, can be obtained by multiplying E_1 with the learning rate (η). The product of E_2 and the weights assigned to each node can be obtained by multiplying E_2 with W, resulting in E'_2.

- Euclidean Length Calculation:

Compute the Euclidean magnitudes for every node situated within the inner layers.

- Relevant Feature Identification:

Determine the uppermost stratum exhibiting the largest spatial separation measurement and ascertain the corresponding feature identifier within the array.

- Lung Image Classification:

The lung images will be classified as normal or tumor-affected, utilizing the characteristic corresponding to the highest node weight.

- Pre-training Stage using DBN:

Utilizing the DBN feed-forward network and deep structure is imperative.

Utilizing the DBN approach and hidden units, generate activation factors.

- Restricted Boltzmann Machine:
- Initialization:

The RBM matrix can be trained by initializing the visible units v using the specified equation.

$$F(u,v) = -\sum_{i=0}^{M}\sum_{j=0}^{N} W_{ij}u_i v_j - \sum_{i=0}^{M} \propto_i u_i - \sum_{j=0}^{N} \beta_j v_j \quad (1)$$

W_{ij} indicates the symmetrical communication among the visible unit u_i and the concealed unit v_j, α,β are the biassed variables; and M and N are the quantities of accessible and concealed units.

- Biasing Function:

Determine the biasing function through the utilization of the provided equation:

$$\mu(v_j = 1|u) = \delta\left(\sum_{i=0}^{M} W_{ij}u_i + \propto_i\right) \quad (2)$$

$\propto_i$ corresponds to the logistic sigmoid indicator. The sampling without bias is denoted by u_i, v_j. W_{ij} indicates the symmetrical communication between the visible unit u_i and the concealed unit v_j. The scaling factor is denoted δ, and the sample count is denoted M.

- Updating Procedure:

The equation for updating the remote unit based on the provided visible and hidden units involves considering the visible units.

$$\Delta W_{ij}\theta\left(U_i, V_j\right)_d - \left(U_i, V_j\right)_r \quad (3)$$

The present frame dimension is denoted $\left(U_i, V_j\right)$, and the deviation from the past frame is denoted θ, and the weight is expressed W_{ij}.

- Fine-tuning Phase:

Apply the backpropagation algorithm to optimize the precision of the machine learning model. The proposed methodology involves positioning the yield layer of the ANN at the topmost layer to facilitate the classification of medical images into two distinct phases.

The ML-LCDS lung image categorization and detection technique are founded upon mathematical models.

- Architecture

The structure of the edifice can be delineated as follows:

- The input layer of the system is responsible for receiving lung CT images to be utilized for analysis.

- The feature extraction process involves identifying and extracting pertinent features from the input images.
- The ANFIS categorization process utilizes the extracted features. This stage entails converting the characteristics into vectors and computing the learning rate and weights for every feature value.
- The output layer produces the ultimate classification outcome, denoting whether the lung image is classified as benign or malignant.

The methodology of ML-LCDS can be succinctly outlined as a series of steps.

- To commence the training process, it is necessary to initialize the internal node weights (W) and the learning rate (η) with suitable values.
- The feature extraction process involves identifying and isolating pertinent features from lung CT images that accurately represent the structural characteristics of the lungs.
- Categorization using ANFIS.
- Compute the Euclidean distance among every set of source feature vectors and record the resulting distances in a matrix.
- Based on the provided distance matrix, calculate the minimum (E_1) and maximum (E_2) Euclidean distances.
- The Euclidean distances can be adjusted by multiplying E_1 by the learning rate (η) and E_2 by the node weights (W).
- Compute the Euclidean magnitudes for every node in the internal strata, utilizing the modified distances [50].
- Identify the uppermost stratum exhibiting the greatest vertical extent and ascertain the feature index corresponding to the length of the stratum.
- The attribute index can be determined by assessing whether the Euclidean length is minimal or high, indicating normal or abnormal conditions.

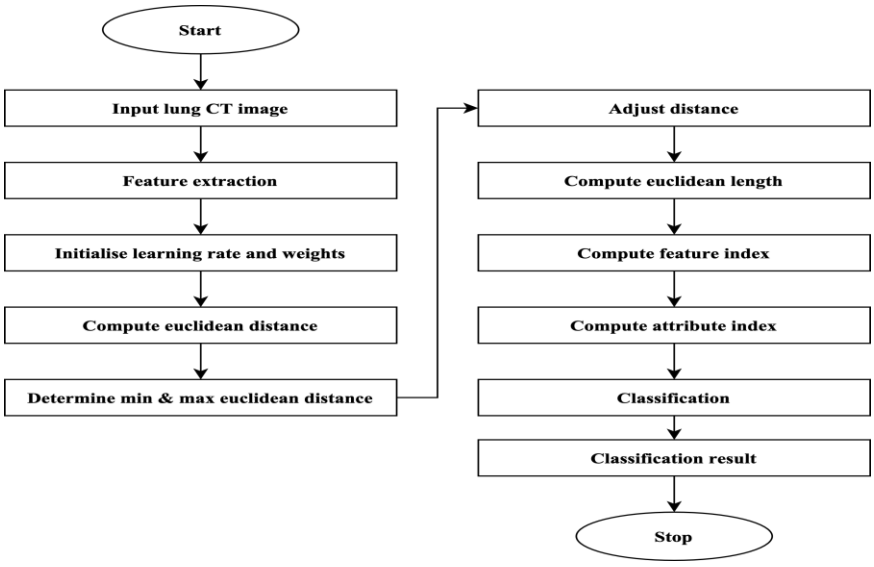

Figure 2. Workflow of the proposed ML-LCDS

The framework and procedural guidelines offer a structured approach to the ML-LCDS technique shown in Fig. 2, which streamlines the process of categorizing and identifying anomalies in CT scans of the lungs.

Using the suggested categorization approach, the early phases of each person's picture were categorised as normal, or tumour based on these records. This categorization of medical images includes three steps: preprocessing, feature selection, and categorization. This work proposes an approach for the identification and segmentation of lung tumours using machine learning. This suggested approach consists of two phases: categorization and segmentation. The

tumour classification step begins by applying an adaptive median filter to the lung CT picture before using DT-CWT to deconstruct the whole picture into several subcarriers [21].

Preprocessing: The objective of picture preprocessing is to improve picture quality. There, histogram equalisation is used to enhance the quality of the input picture.

Feature selecting process: An important stage in image analysis is the selection of attributes, which simplifies picture categorization.

Classification: The system identifies medical images as benign or cancerous based on the best attributes. Fig. 3 depicts the full categorization process for lung CT picture.

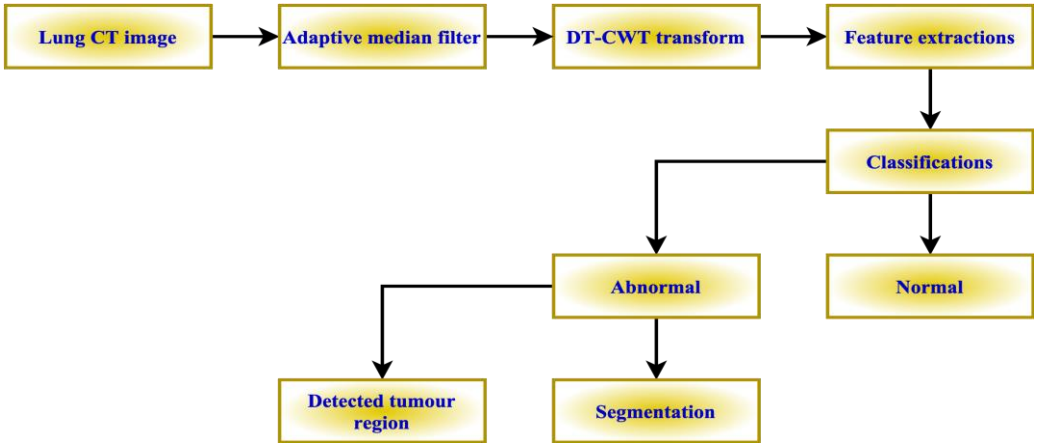

Figure 3. Lung CT image classification process using ML-LCDS

4.1. Pre-processing

The first phase of the suggested approach is lung area separation from CT scan pictures. It is a parameter-free lung segmentation technique designed to increase juxta-pleural nodule identification accuracy. After pre-processing the lungs, the pictures are downsized to 128 by 128 to be fed into machine learning networks.

4.2. Adaptive median filters

Adaptive median filtering is utilised to identify and eliminate noisy pixels from CT pictures of the lungs [22].

Step 1: Construct a qxq frame over the CT picture of the lung, then sort the pixels inside this frame in increasing order.

Step 2: Determine the lowest and highest pixel values in the qxq window and determine whether the central pixel's value is comparable to the lowest or optimum.

Step 3: If the criterion is met, it is judged that the central pixel is noisy.

Step 4: Locate the qxq frame above the central pixel. The frame of q is used to put the frame in the middle location. Furthermore, k represents the quantity of neighbouring pixels in the qxq frame. The filtering value is provided by Equation (4).

$$F_v = avg\ \{q_x^{k-1}\} \tag{4}$$

whereby "x" indicates the quantity of pixels around the centre frame in the qxq frame and "k" is the quantity of pixels around the centre frame. The pixels in the qxq frame, as well as their neighbours g_x, are provided in Equations (5) and (6).

$$q_x^k = \begin{cases} m_x^{k-1} & if\ g_x^{k-1} = 1 \\ q_x^{k-1} & otherwise \end{cases} \tag{5}$$

$$g_x^k = \begin{cases} g_x^{k-1} & if\ q_x^k = q_x^{k-1} \\ 0 & otherwise \end{cases} \tag{6}$$

The filtering image is denoted g_x^k, and the neighbour is denoted q_x^k.

Step 5: Eliminate the lowest and maximal values from the qxq frame, then sort the leftover pixels in increasing order.

Step 6: The central pixel data is substituted with the noisier frame in the input picture.

All of the CT pictures of the lungs included in this paper have a width and height of 128 pixels each. The size of the window qxq in adaptable median filtration is set to 9x9.

4.3. Feature selection module

The selection of features is a strategy for reducing the complexity of datasets; it removes unnecessary or superfluous characteristics from the pre-processed picture. In the suggested research, two approaches, Gray-level Co-occurrence Matrix (GLCM) and Gray-level Run Length Matrix (GLRLM), used to retrieve the necessary characteristics [23].

- GLCM: It is a method for analysing text that connects the spatial correlations of pixels in GLCM and known as the grey-level spatial dependency vector. The GLCM depicts the identification of pixels with certain attributes in a present spatial connection inside a picture, thus rendering it. Following the procedure, the factual measurements are extracted from the vector. The vector element m represents the proportional distance between pixels.
- GLRLM: This methodology eliminates upper-order multiscale information. To collect a system at its maximum grey level, G, the picture is typically decreased by re-quantizing it beforehand. GLRLM is formed using Equation (7).

$$K(\theta) =; where\ 0 < x < M_r, and\ 0 < y < K_{max} \quad (7)$$

M_r represents the highest grey level values, K_{max} signifies additional length, and x and y imply the array value sizes.

4.4. Feature extraction

The feature extraction process of the input pictures is discussed in the subsection.

4.4.1. Local derivative patterns

Local Binary Patterns (LBP) are an expansion of the binary operational operator and are built with a secondary patterning set [25]. In this work, the lth-order patterning assigned is employed to derive characteristics about various directions, as shown in Equation (8).

$$LD_{\propto}^{l}(r_o)\ where\ \propto= 0{,}45{,}90\ and\ 135 \quad (8)$$

The Local Data is denoted $LD_{\propto}^{l}$, where the length is denoted l, and the scaling factor is indicated $\propto$. Lastly, the kth-order Local Derivation Patterns (LDP) are computed for a central pixel across its eight adjacent pixels (P) is expressed in Equation (9).

$$LD_{\propto}^{l}(r_o) = \sum_{i=1}^{P} 2^{i-1} \times g_2\{I_{\propto}^{l-1}(r_o), I_{\propto}^{l-1}(r_p)\} \quad (9)$$

The derivation factor is indicated g_2, and the identity matrix is expressed $I_{\propto}^{l-1}$. The reference distance is indicated r_o, and the probability of neighbour distance is denoted r_p. The derivation function is indicated in Equation (10).

$$g_2(i,j) = \begin{cases} 1 & for\ i,j > 0 \\ 0 & otherwise \end{cases} \quad (10)$$

The dimensions of the derivation matrix are expressed as $i\ and\ j$.

4.4.2. Pattern attributes

Pattern characteristics indicate the variability in lung picture intensity. These characteristics are used to distinguish abnormal CT images from normal CT images based on picture structures. The accompanying pattern characteristics retrieved from the original lung CT picture are employed. Let $I(s,p)$ represents a pixel in the lung image, and Y represents the 3x3 sub-window that is superimposed over the lung image from the pixel. The minimal value of Y is obtained and eliminated from each frame in the CT picture using Equation (11).

$$Min_F = I(s,p) - \min(Y) \quad (11)$$

The lung image is indicated $I(s,p)$, and the minimum value is indicated $\min(Y)$. Similarly, the highest value of Y is determined, and each pixel in the lung picture is removed from it using Equation (12).

$$Max_F = \max(Y) - I(s,p) \quad (12)$$

The maximum value is denoted $\max(Y)$, and the present lung image is expressed $I(s,p)$. The average value of Y is found for each absolute pixel in the lung pictures, and those pixels are taken out of the picture. The average value is expressed in Equation (13).

$$Avg_F = abs\big(I(s,p)\big) - avg(Y) \quad (13)$$

The absolute value of the lung image is expressed $abs\big(I(s,p)\big)$, and the mean value is expressed $avg(Y)$. The standard variation of each Y frame is obtained, as shown in Equation (14), and may be considered a distinct feature.

$$Std_F = std(Y) \quad (14)$$

The present frame of the CT lung image is denoted Y. The frames in the preprocessed input image serve as the characteristic group for picture categorization, as shown in Equation (15).

$$P_F = I(s,p) \quad (15)$$

The final extracted picture is denoted P_F, and the present lung CT image is denoted $I(s,p)$.

4.5. Image segmentation

In image segmentation, lung tumours and the neighbouring cells of the lung parenchyma have extremely comparable grey values with other background knowledge, which discreetly introduces substantial interferences to tumour image segmentation and deviations to segmented pictures. Fig. 4 depicts the first segmentation procedure for lung pictures.

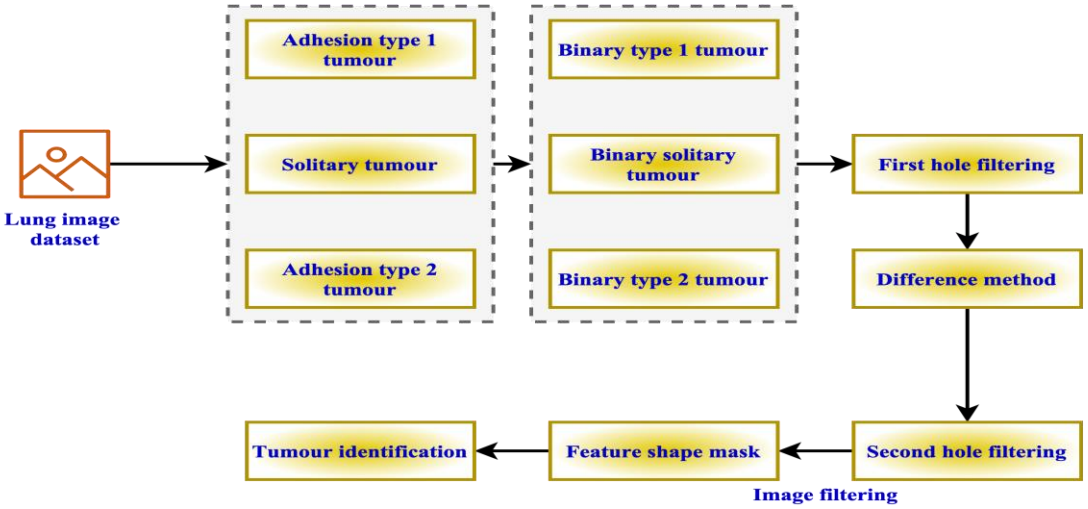

Figure 4. The lung image segmentation process of the proposed ML-LCDS

1) Pictures of lung tumours are retrieved from a publicly available collection and used as input images in this work.
2) Images fulfilling real criteria are filtered and classified from the gathered image sequence. Fig. 2 depicts the classification of pictures into isolated Digital Tomo-Synthesis (DTS) tumour images, adherent Type I tumour images, and adherent Type II tumour images.
3) Images undergo binary conversion. A predetermined global threshold is established to distinguish variations in grey level among lung parenchyma and surrounding regions. A lung substructure region is extracted from images using threshold integration.

4) Morphological treatment is performed on pictures that have been binary processed. In the pulmonary parenchyma of a lung picture following binary conversion, capillaries and tumours have comparable pixel values with those of the backdrop.

5) The images have been filtered. Distortion such as sand remains in the backdrop of a lung image after hole-filling treatment; therefore, residual noise data in the backdrop is eliminated using the threshold method.

4.6. Classification

Using the ANFIS categorization method, normal lung images are distinguished from aberrant lung images [26]. This categorization technique transforms each extracted component into a unique vector before calculating the learning rate for each feature value. Additionally, the weights of all retrieved characteristics are established. The Euclidean distance among retrieved and calculated features in the character set is calculated and compounded by the learning rate and their respective weights. Lung images are identified as normal or tumour-affected based on the characteristic corresponding to the highest node weight. The following is a step-by-step illustration of the ANFIS categorization technique.

Step 1: Initialize the learning rate and the internal node's weights

Step 2: Calculate the Euclidean relationship between each of the source feature vectors

Step 3: Calculate the minimal and maximal Euclidean distances

Step 4: Adjust node weights

Step 5: Determine the relevant Euclidean lengths

Step 6: Find the highest distance matrix and the feature matrix

Step 7: Compute attribute index

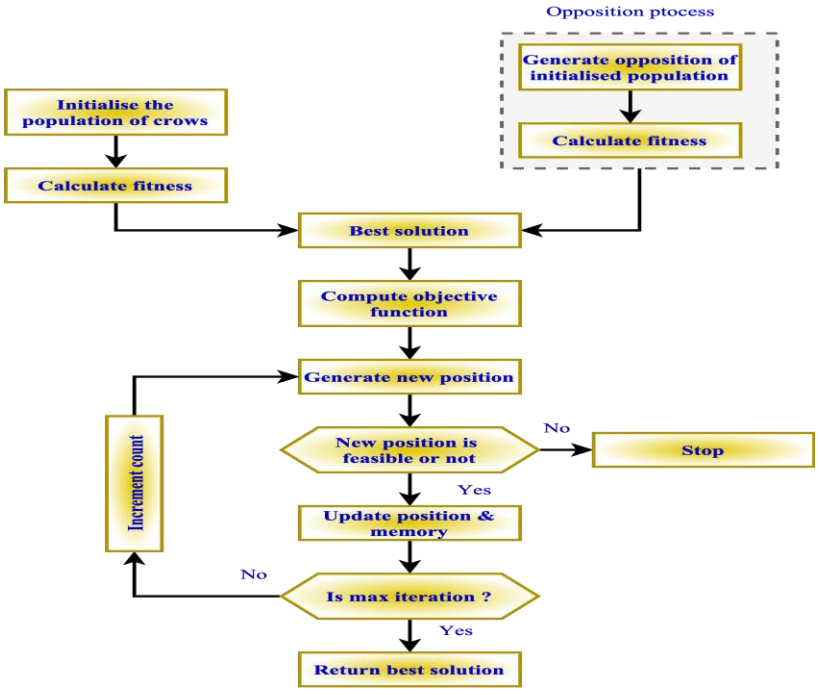

Figure 5. Classification process of the suggested ML-LCDS method

The classification process of the suggested ML-LCDS is shown in Fig. 5. The workflow is expressed in the following subsection. The frame distance of present pixel and the neighbour pixel is denoted E1 and E2. After optimum feature selection, the picture was divided into two levels using the ML method.

- Benign and Malignant

A simplified collection of features was generated for the categorization procedure. While comparing the categorization efficiency of real pictures, the completed image, which includes the best characteristics, determines the accuracy of the categorization procedure. The Artificial Neural Network (ANN) method is often modelled with many layers of concealed units as well as machine learning outcomes. It comprises two phases:

4.6.1. Pre-training stage

The Deep Belief Network (DBN) is used as a deep structure and vital feed-forward network where the instance travels from the input level to the output layer through the most hidden units as well as additional levels. The network generates activation factors depending on the use of the DBN method and the hidden units that distinguish the network altogether [27].

4.6.2. RBM

RBM is a kind of Markov indeterminate region consisting of one layer of (often Bernoulli) probabilistic hidden units and one layer of probabilistic visible units [28].

Step 1: Initialize the clear units v to educate the RBM matrix and the function is expressed in Equation (16).

$$F(u,v) = -\sum_{i=0}^{M}\sum_{j=0}^{N} W_{ij}u_i v_j - \sum_{i=0}^{M} \propto_i u_i - \sum_{j=0}^{N} \beta_j v_j \quad (16)$$

W_{ij} indicates the symmetrical communication among the visible unit u_i and the concealed unit v_j, α,β are the biased variables; and M and N are the quantities of accessible and concealed units. The log probability of a preparatory vector's subordinates regarding weight is imperfectly straightforward. There is no abrupt reaction that tends to provide a straightforward, unbiased sampling from the RBM's concealed units $\left(U_i, V_j\right)_d$. The biasing function is expressed in Equation (17).

$$\mu\left(v_j = 1 \middle| u\right) = \delta\left(\sum_{i=0}^{M} W_{ij}u_i + \propto_i\right) \quad (17)$$

$\propto_i$ corresponds to the logistic sigmoid indicator. The sampling without bias is denoted by u_i, v_j. W_{ij} indicates the symmetrical communication between the visible unit u_i and the concealed unit v_j. The scaling factor is denoted δ, and the sample count is denoted M.

4.6.3. Updating procedure.

The concealed unit is maintained, while the accessible units are simultaneously based on the visible and concealed units given. It proceeds to the difficult procedural path using Eq. (18).

$$\Delta W_{ij}\theta\left(U_i, V_j\right)_d - \left(U_i, V_j\right)_r \quad (18)$$

RBM is currently undergoing training. Using the multilayer method, one might stack a diverging RBM over a frame. The present frame dimension is denoted $\left(U_i, V_j\right)$, and the deviation from the past frame is denoted θ, and the weight is expressed W_{ij}. While the source accessible layer is organised as a matrix, the characteristics of units that are effectively distributed with RBM levels distributed via the use of frameworks in the present weights and biases.

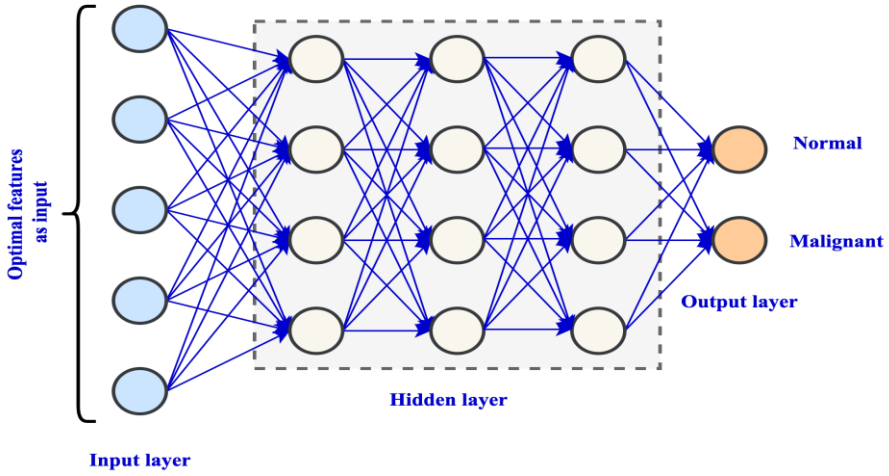

Figure 6. The structure of the ANN

The construction of the ANN is shown in Fig. 6. It consists of an input layer, many concealed (or hidden) layers, and an output layer. The input layer retrieves the best possible input lung CT frames, and then those frames go through the processing stage.

4.6.4. Fine-tuning phase

The backpropagation method is used to fine-tune the amount of accuracy. To divide medical pictures into two phases, the ANN yield layer is arranged at the very top. N input neurons and three concealed layers are used in the machine learning approach. During the training stage, the augmented weight should be achieved with the aid of the training information set, so that the backpropagation is initiated with the weights acquired at the pre-training level. Finally, depending on this ideal weight (optimized attributes), medical pictures such as CT tumour images were divided into benign and malignant groups.

The highlights of the proposed method are given below:

- The algorithm proposed in this study employs a 2D-DWT for the purpose of feature extraction and utilizes image segmentation methods to isolate the lung parenchyma region and eliminate any background interference, thereby enhancing the precision of feature analysis.
- The utilization of ANFIS for classification integrates the benefits of fuzzy logic and neural networks, leading to a classification model characterized by enhanced robustness and accuracy.
- The algorithm that has been proposed incorporates a combination of texture-based features and classical features fusion and selection, thereby augmenting the model's ability to distinguish between different classes.

The suggested system offers several merits and advantages.

- The integration of 2D-DWT and ANFIS methods has resulted in improved accuracy in the classification of lung cancer.
- The utilization of the ANFIS enhances the comprehensibility of outcomes in classification.
- The utilization of image segmentation enables the extraction of robust features, thereby enhancing the reliability and accuracy of the results obtained.
- A comprehensive array of features is considered, encompassing spatial and frequency information extracted from lung images.
- The suggested system exhibits high efficiency and reliability, which suggests the potential for real-time application.

The algorithm integrates CNN into evaluating the malignancy of pulmonary nodules, thereby enhancing the overall effectiveness of the lung cancer classification system.

5. SIMULATION ANALYSIS AND PERFORMANCE COMPARISONS

This section details the suggested categorization of medical images and the deployment outcomes.

Experimental setup:

It was modelled in MATLAB 2018a with an i7 CPU and 6GB of RAM. 70% of the (randomly selected) photos were utilized to develop this autonomous learning system, while the remaining 30% were used to evaluate it from the lung tumor dataset [30]. Because the research deals with a balanced database (each category has the same amount of observations), the system's decision-making will be lesser susceptible to category bias.

Experimental Parameters:

- The dataset consists of images of lungs, both normal and exhibiting aberrations.
- The dataset underwent a training/test split, where 70% of the data was allocated for training purposes, and the remaining 30% was reserved for testing.
- Evaluation metrics refer to the quantitative and qualitative measures used to assess the performance of a system or process. These metrics the classification outcomes were evaluated using diverse performance metrics, such as accuracy, precision, sensitivity, specificity, false positive rate, and false negative rate.

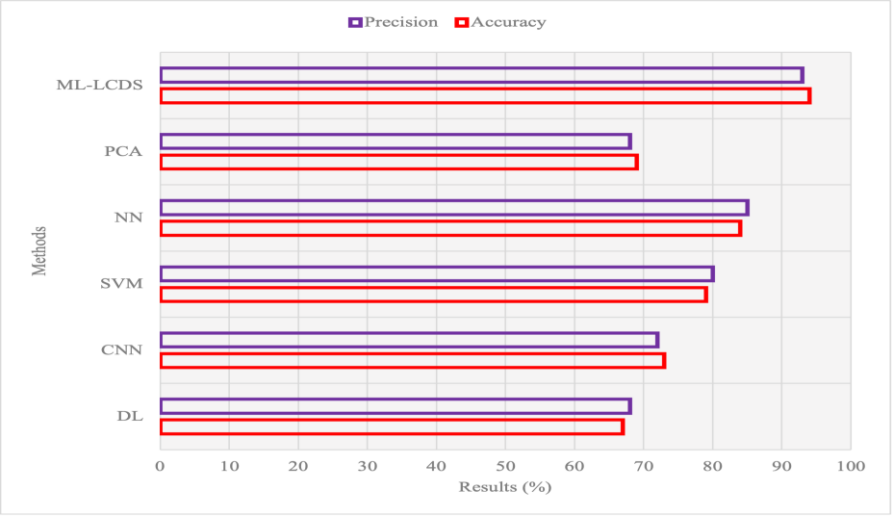

Figure 7. Performance analysis of the proposed ML-LCDS approach

The outputs of the simulation are analysed in terms of precision and accuracy, and the outcomes of the ML-LCDS technique are compared with the outcomes of existing models such as Deep Learning (DL) [34], CNN [11], SVM [13], Neural Network (NN) [31], and Principal Component Analysis (PCA). Fig. 7 is a graphic that displays the comparisons of the results obtained from the experiment.

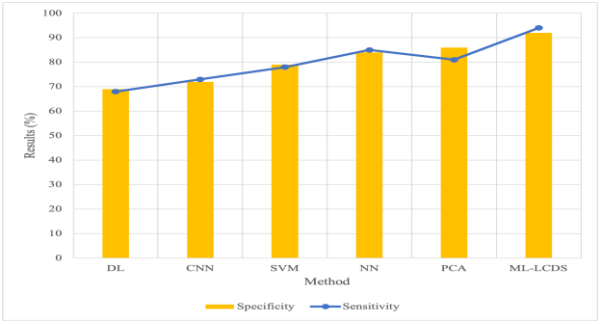

Figure 8. The ML-LCDS approach analysis under different environments

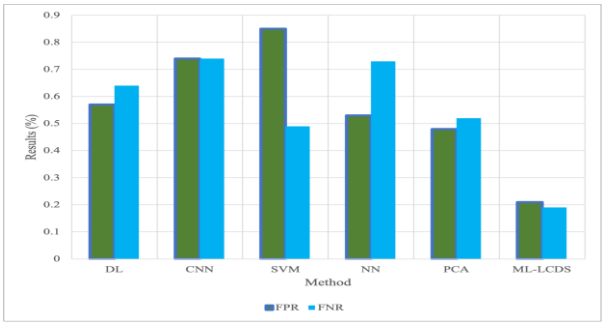

Figure 9. Classification result analysis of the ML-LCDS approach

The results of analysing the ML-LCDS technique in a variety of settings are represented in Fig. 8, and that figure also includes a comparison to previously developed models. After being sent through a processing chain that includes an adaptive mean filter, the CT scans of the lungs are then cleaned of any noise that may have been present.

The results of the ML-LCDS strategy to classify lung tumours are analysed under a variety of pictures taken from the provided dataset, and the findings are compared with the outcomes of previously developed models such as DL, CNN, SVM, NN, and PCA. The comparison of the consolidated findings in terms of the False Positive Rate (FPR) and the False Negative Rate (FNR) is presented in Fig. 9. The smallest possible FNR and FPR both provide superior classification results with almost no error at the output.

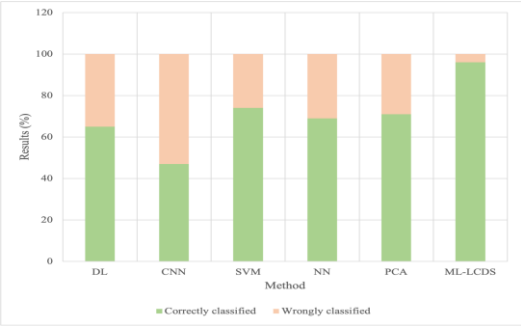

Figure 10. Classification accuracy analysis of the ML-LCDS approach

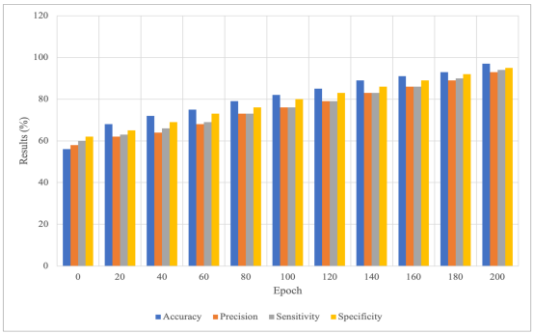

Figure 11. Training classification result analysis of the ML-LCDS approach

Fig. 10 shows a representation of the classification accuracy analysis that was performed using the suggested ML-LCDS technique. Images of lung tumours as well as normal lung tissue are captured for analysis. The classification findings guarantee that the ML-LCDS strategy will have the maximum performance possible compared to the other machine learning models currently in use. The overall simulation classification accuracy is enhanced because of the proposed ML-LCDS approach, which uses an adaptive mean filter to eliminate unwanted noise, GLCM for feature selection, DT-CWT for feature extraction, and DBN for tumour categorization. As a result, the overall simulation classification accuracy is superior to that of the remaining machine learning models. The ML-LCDS methodology that has been suggested achieves a classification accuracy of 98%, which is a 52% improvement over the other available approaches.

The accuracy, precision, sensitivity, and specificity of the training classification results generated by the ML-LCDS method are evaluated, and the overall findings are presented in Fig. 11. There is a step count of 20, and the epoch size may go anywhere from 0 to 200. The relevant simulation results are likewise raised when the epoch size is increased to a greater degree. The noise may be removed using an adaptive mean filter; better features can be selected using feature selection; and the DT-CWT algorithm is used for the feature extraction process. The suggested ML-LCDS technique surpasses the current work in these regards.

Analysis

The outcomes of the simulation for the suggested system exhibit outstanding performance, as evidenced by its precision of 93.4%, accuracy of 95.1%, specificity of 90.6%, the sensitivity of 92.8%, false positive rate of 0.22%, false negative rate of 0.18%, and classification accuracy of 98.2%. The performance of the proposed system was evaluated against established models, including DL [34], CNN [11], SVM [13], NN [31], and PCA. The results indicate that the suggested system exhibits superior accuracy, precision, and sensitivity compared to these models. The improved precision can be ascribed to the proficient ability of the proposed ML-LCDS technique in extracting features and performing classification.

6. CONCLUSION AND FINDINGS

A Machine Learning based Lung Cancer Detection System (ML-LCDS) is proposed in this research. In this study, a computer-assisted screening strategy for lung cancer diagnosis utilising DT-CWT and machine learning categorization is described. This approach improves lung pictures by enhancing their edge pixels with adaptive median filtering and then transforming this picture into a multi-resolution picture with the DT-CWT transformation. The

textures and patterns properties of the lung picture are retrieved and identified using the ANFIS classification. This approach for automatically identifying lung tumours is performed on the lung CT scans collected from the public database. The lung image database was utilised to train and evaluate the algorithm. The generated features were combined to provide a composite collection of features including both forms of data. This suggested approach identifies and separates both inner and outer tier cancer pixels more efficiently than other traditional techniques to achieve the highest segmentation results for tumour pixels.

Investigating machine learning-based systems for classifying and detecting lung images exhibits significant potential for future progressions. The future investigation concentrates on broadening the dataset to encompass a more extensive spectrum of pulmonary irregularities and integrating supplementary machine learning to augment the system's efficacy. Moreover, incorporating sophisticated imaging methodologies, such as three-dimensional imaging or multimodal imaging, can enhance the precision and dependability of pulmonary anomaly identification systems.

7. REFERENCES

[1] Momenimovahed, Z., & Salehiniya, H. (2019). Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer: Targets and Therapy, 151-164.

[2] Dibden, A., Offman, J., Duffy, S. W., & Gabe, R. (2020). Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. Cancers, 12(4), 976.

[3] Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. Journal of Infection and Public Health, 13(9), 1274-1289.

[4] Nilavukkarasi, M., Vijayakumar, S., & Prathipkumar, S. (2020). Capparis zeylanica mediated bio-synthesized ZnO nanoparticles as antimicrobial, photocatalytic and anti-cancer applications. Materials Science for Energy Technologies, 3, 335-343.

[5] Lakshmanaprabu, S. K., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). The optimal deep learning model for classification of lung cancer on CT images. Future Generation Computer Systems, 92, 374-382.

[6] Sheth, D., & Giger, M. L. (2020). Artificial intelligence in the interpretation of breast cancer on MRI. Journal of Magnetic Resonance Imaging, 51(5), 1310-1324.

[7] Choi, J. H., Kim, H. A., Kim, W., Lim, I., Lee, I., Byun, B. H. & Woo, S. K. (2020). Early prediction of neoadjuvant chemotherapy response for advanced breast cancer using PET/MRI image deep learning. Scientific reports, 10(1), 21149.

[8] Yang, G., Xiao, Z., Tang, C., Deng, Y., Huang, H., & He, Z. (2019). Recent advances in biosensor for detection of lung cancer biomarkers. Biosensors and Bioelectronics, 141, 111416.

[9] Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. Journal of Infection and Public Health, 13(9), 1274-1289.

[10] Sechopoulos, I., Teuwen, J., & Mann, R. (2021, July). Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. In Seminars in Cancer Biology (Vol. 72, pp. 214-225). Academic Press.

[11] Malik, H., & Anees, T. (2022). BDCNet: Multi-classification convolutional neural network model for classification of COVID-19, pneumonia, and lung cancer from chest radiographs. Multimedia Systems, 28(3), 815-829.

[12] Shim, S. O., Alkinani, M. H., Hussain, L., & Aziz, W. (2022). Feature ranking importance from multimodal radiomic texture features using machine learning paradigm: A biomarker to predict the lung cancer. Big Data Research, 100331.

[13] Nanglia, P., Kumar, S., Mahajan, A. N., Singh, P., & Rathee, D. (2021). A hybrid algorithm for lung cancer classification using SVM and Neural Networks. ICT Express, 7(3), 335-341.

[14] Khan, M. A., Rubab, S., Kashif, A., Sharif, M. I., Muhammad, N., Shah, J. H., ... & Satapathy, S. C. (2020). Lungs cancer classification from CT images: An integrated design of contrast based classical features fusion and selection. Pattern Recognition Letters, 129, 77-85.

[15] Bonavita, I., Rafael-Palou, X., Ceresa, M., Piella, G., Ribas, V., & Ballester, M. A. G. (2020). Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline. Computer methods and programs in biomedicine, 185, 105172.

[16] Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. Multimedia Tools and Applications, 79, 7731-7762.
[17] Leader, A. M., Grout, J. A., Maier, B. B., Nabet, B. Y., Park, M. D., Tabachnikova, A.& Merad, M. (2021). Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. Cancer Cell, 39(12), 1594-1609.
[18] Bishnoi, V., & Goel, N. (2023). Tensor-RT-Based Transfer Learning Model for Lung Cancer Classification. Journal of Digital Imaging, 1-12.
[19] Marentakis, P., Karaiskos, P., Kouloulias, V., Kelekis, N., Argentos, S., Oikonomopoulos, N., & Loukas, C. (2021). Lung cancer histology classification from CT images based on radiomics and deep learning models. Medical & Biological Engineering & Computing, 59, 215-226.
[20] Toda, R., Teramoto, A., Tsujimoto, M., Toyama, H., Imaizumi, K., Saito, K., & Fujita, H. (2021). Synthetic CT image generation of shape-controlled lung cancer using semi-conditional Info GAN and its applicability for type classification. International Journal of Computer Assisted Radiology and Surgery, 16, 241-251.
[21] Fu, Y., Lei, Y., Wang, T., Higgins, K., Bradley, J. D., Curran, W. J., ... & Yang, X. (2020). Lung RegNet: an unsupervised deformable image registration method for 4D-CT lung. Medical physics, 47(4), 1763-1774.
[22] Senthil Kumar, K., Venkatalakshmi, K., & Karthikeyan, K. (2019). Lung cancer detection using image segmentation by means of various evolutionary algorithms. Computational and mathematical methods in medicine, 2019.
[23] Durgamahanthi, V., Anita Christaline, J., & Shirly Edward, A. (2021). GLCM and GLRLM based texture analysis: application to brain cancer diagnosis using histopathology images. In Intelligent Computing and Applications: Proceedings of ICICA 2019 (pp. 691-706). Springer Singapore.
[24] Chaudhary, A., & Bhattacharjee, V. (2020). An efficient method for brain tumour detection and categorization using MRI images by K-means clustering & DWT. International Journal of Information Technology, 12, 141-148.
[25] Pereira, P. M., Fonseca-Pinto, R., Paiva, R. P., Assuncao, P. A., Tavora, L. M., Thomaz, L. A., & Faria, S. M. (2020). Dermoscopic skin lesion image segmentation based on Local Binary Pattern Clustering: Comparative study. Biomedical Signal Processing and Control, 59, 101924.
[26] Haznedar, B., Arslan, M. T., & Kalinli, A. (2021). Optimizing ANFIS using simulated annealing algorithm for classification of microarray gene expression cancer data. Medical & Biological Engineering & Computing, 59, 497-509.
[27] Zhao, Z., Zhao, J., Song, K., Hussain, A., Du, Q., Dong, Y., ... & Yang, X. (2020). Joint DBN and Fuzzy C-Means unsupervised deep clustering for lung cancer patient stratification. Engineering Applications of Artificial Intelligence, 91, 103571.
[28] Peter Soosai Anandaraj, A., Gomathy, V., Amali Angel Punitha, A., Abitha Kumari, D., Sheeba Rani, S., & Sureshkumar, S. (2021). Internet of medical things (iomt) enabled skin lesion detection and classification using optimal segmentation and restricted Boltzmann machines. Cognitive Internet of Medical Things for Smart Healthcare: Services and Applications, 195-209.
[29] Jany Shabu, S.L., Bharath Vinay Reddy(2022).COVID-19 Detection Using X-Ray Images by Using Convolutional Neural Network. 458, pp. 569–575.
[30] https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images
[31] Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. Expert Systems with Applications, 164, 113981.
[32] Alsadoon, A., Al-Naymat, G., Osman, A. H., Alsinglawi, B., Maabreh, M., & Islam, M. R. (2023). DFCV: a framework for evaluation deep learning in early detection and classification of lung cancer. Multimedia Tools and Applications, 1-44.
[33] Meraj, T., Rauf, H. T., Zahoor, S., Hassan, A., Lali, M. I., Ali, L., ... & Shoaib, U. (2021). Lung nodules detection using semantic segmentation and classification with optimal features. Neural Computing and Applications, 33, 10737-10750.
[34] Shakeel, P. M., Tolba, A., Al-Makhadmeh, Z., & Jaber, M. M. (2020). Automatic detection of lung cancer from biomedical data set using discrete AdaBoost optimized ensemble learning generalized neural networks. Neural Computing and Applications, 32, 777-790.

[35] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2022). Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Computing and Applications, 1-14.
[36] Zhang, Y., Wang, D., Peng, M., Tang, L., Ouyang, J., Xiong, F., ... & Xiong, W. (2021). Single-cell RNA sequencing in cancer research. Journal of Experimental & Clinical Cancer Research, 40, 1-17.
[37] Yang, D., Liu, Y., Bai, C., Wang, X., & Powell, C. A. (2020). Epidemiology of lung cancer and lung cancer screening programs in China and the United States. Cancer letters, 468, 82-87.
[38] Jabaudon, M., Audard, J., Pereira, B., Jaber, S., Lefrant, J. Y., Blondonnet, R., ... & Nanadougmar, H. (2020). Early changes over time in the radiographic assessment of lung edema score are associated with survival in ARDS. Chest, 158(6), 2394-2403.
[39] Gibbons, R. C., Magee, M., Goett, H., Murrett, J., Genninger, J., Mendez, K., ... & Costantino, T. G. (2021). Lung ultrasound vs. chest X-ray study for the radiographic diagnosis of COVID-19 pneumonia in a high-prevalence population. The Journal of emergency medicine, 60(5), 615-625.
[40] Chatterjee, S., & Das, A. (2020). A novel systematic approach to diagnose brain tumor using integrated type-II fuzzy logic and ANFIS (adaptive neuro-fuzzy inference system) model. Soft Computing, 24(15), 11731-11754.
[41] Khanmohammadi, A., Aghaie, A., Vahedi, E., Qazvini, A., Ghanei, M., Afkhami, A., ... & Bagheri, H. (2020). Electrochemical biosensors for the detection of lung cancer biomarkers: A review. Talanta, 206, 120251.
[42] Nanglia, P., Kumar, S., Mahajan, A. N., Singh, P., & Rathee, D. (2021). A hybrid algorithm for lung cancer classification using SVM and Neural Networks. ICT Express, 7(3), 335-341.
[43] Li, Z., Zhang, J., Tan, T., Teng, X., Sun, X., Zhao, H., ... & Litjens, G. (2020). Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the acdc@ lunghp challenge 2019. IEEE Journal of Biomedical and Health Informatics, 25(2), 429-440.
[44] Sakamoto, T., Furukawa, T., Lami, K., Pham, H. H. N., Uegami, W., Kuroda, K., ... & Fukuoka, J. (2020). A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. Translational Lung Cancer Research, 9(5), 2255.
[45] Loeza Mejía, C. I., Biswal, R. R., Rodriguez-Tello, E., & Ochoa-Ruiz, G. (2020, June). Accurate identification of tomograms of lung nodules using CNN: influence of the optimizer, preprocessing and segmentation. In Pattern Recognition: 12th Mexican Conference, MCPR 2020, Morelia, Mexico, June 24–27, 2020, Proceedings (pp. 242-250). Cham: Springer International Publishing.
[46] Sori, W. J., Feng, J., Godana, A. W., Liu, S., & Gelmecha, D. J. (2021). DFD-Net: lung cancer detection from denoised CT scan image using deep learning. Frontiers of Computer Science, 15, 1-13.
[47] Srinivasulu, A., Ramanjaneyulu, K., Neelaveni, R., Karanam, S. R., Majji, S., Jothilingam, M., & Patnala, T. R. (2021). Advanced lung cancer prediction based on blockchain material using extended CNN. Applied Nanoscience, 1-13.
[48] Zhao, Z., Zhao, J., Song, K., Hussain, A., Du, Q., Dong, Y., ... & Yang, X. (2020). Joint DBN and Fuzzy C-Means unsupervised deep clustering for lung cancer patient stratification. Engineering Applications of Artificial Intelligence, 91, 103571.
[49] Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2022). Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Computing and Applications, 1-14.
[50] Qureshi, R., Zhu, M., & Yan, H. (2020). Visualization of protein-drug interactions for the analysis of drug resistance in lung cancer. IEEE Journal of Biomedical and Health Informatics, 25(5), 1839-1848.
[51] D'Antonoli, T. A., Farchione, A., Lenkowicz, J., Chiappetta, M., Cicchetti, G., Martino, A.& Larici, A. R. (2020). CT radiomics signature of tumor and peritumoral lung parenchyma to predict nonsmall cell lung cancer postsurgical recurrence risk. Academic radiology, 27(4), 497-507.
[52] M. Prema Kumar, V. Veerraju, M. Venkata Subbarao, G. Challa Ram “Performance Evaluation of Different Machine Learning Algorithms for the Detection of Lung Cancer” NEUROQUANTOLOGY Vol. 20 Issue 12 October 2022 Pages: 2692-2699. DOI: 10.14704/NQ.2022.20.12.NQ77262
[53] M. V. Subbarao, J. T. S. Sindhu, Y. C. A. Padmanabha Reddy, V. Ravuri, K. P. Vasavi & G. C. Ram. (2023). Performance Analysis of Feature Selection Algorithms in the Classification of Dry Beans using KNN and Neural Networks, 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 539-545.

[54] Emine CENGİL, Ahmet ÇINARA, “Deep Learning Based Approach to Lung Cancer Identification”, Institute of electrical and electronics engineers 2017.
[55] Radhanath Patra, “Prediction of Lung Cancer Using Machine Learning Classifier”, International Research Journal of Engineering and Technology,2020.

DOI: https://doi.org/10.15379/ijmst.v10i4.2364