_____

# Lung Cancer Detection through TYDWT Algorithm

## G.Preetha [1], S, Nirmala Sugirtha Rajini [2]

[1]. *Research Scholar, Department of Computer Science, E-Mail: gpreetha25@gmail.com*
[2].*Professor, Department of Computer Applications, E-Mail: sugirtharaja77@yahoo.com*
*Dr.M.G.R. Educational and Research Institute, Maduravoval, Chennai.*

*\*Corresponding author's E-mail: gpreetha25@gmail.com*

| Article History | Abstract |
|---|---|
| | *Lung cancer is one of the leading causes of cancer-related deaths worldwide. Early detection of lung cancer plays a critical role in its treatment and survival rates. In recent years, computer-aided diagnosis systems have been developed to assist radiologists in detecting lung nodules in computed tomography (CT) images. This paper proposes a novel approach for lung cancer detection using Transverse Dyadic Wavelet Transform (TDWT) for feature extraction and classification. TDWT is a multi-resolution analysis technique that can capture both time and frequency information of the input images. The TYDWT algorithm is applied to the lung CT scan images to decompose the images into different sub-bands at multiple scales. The extracted features from these sub-bands are then used to train a machine learning model for lung cancer detection. The performance of the proposed method is evaluated on a publicly available dataset, achieving an accuracy of 95.6% and a sensitivity of 95.2%. The proposed method shows promising results for automated lung cancer detection, which can improve the accuracy and efficiency of the diagnosis process. The results demonstrate that the proposed approach using TDWT can be an effective method for early detection of lung cancer.* |
| | |

## 1. Introduction

According to the latest available statistics from the Global Cancer Observatory (GCO), lung cancer is the most common cancer among men in India, and the third most common among women. In 2020, it was estimated that there were 67,007 new cases of lung cancer in men and 30,665 new cases in women in India. The incidence rate of lung cancer in India is 11.1 per 100,000 men and 4.7 per 100,000 women. The mortality rate due to lung cancer is also high in India, with 63,677 deaths reported in 2020. The risk factors for lung cancer in India include tobacco use, air pollution, exposure to second-hand smoke, and exposure to occupational hazards such as asbestos and radon. Smoking is the primary cause of lung cancer in India, and around 60% of lung cancer cases in the country are attributable to tobacco use. It is important to note that these statistics are based on estimates and may not reflect the actual number of cases and deaths due to lung cancer in India, as data collection and reporting methods

may vary across different regions and healthcare systems. This systematic review provides an overview of the recent advances in machine learning (ML) techniques for lung cancer detection and diagnosis. The review describes the different types of ML techniques and their applications in lung cancer diagnosis, including computer-aided diagnosis, radiomics, and deep learning. The authors also discuss the potential clinical applications of these techniques and their limitations [1].There are two main types of lung cancer: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is the most common type of lung cancer, accounting for about 85% of all cases. NSCLC is further divided into three subtypes based on the type of cells in which the cancer develops: Adenocarcinoma: This subtype of NSCLC develops in the glandular cells that produce mucus in the lining of the lungs. Adenocarcinoma is the most common type of lung cancer among non-smokers and women. Squamous cell carcinoma: This type of NSCLC develops in the squamous cells that line the airways of the lungs. Squamous cell carcinoma is strongly associated with smoking. Large cell carcinoma: This type of NSCLC can develop in any part of the lung and tends to grow and spread quickly. SCLC is less common than NSCLC and accounts for about 10-15% of all lung cancer cases. This type of lung cancer typically grows and spreads more quickly than NSCLC and is strongly associated with smoking. SCLC is often referred to as oat cell carcinoma due to the oat-shaped appearance of the cancer cells when viewed under a microscope. It is important to accurately diagnose the type of lung cancer a person has, as this information can guide treatment decisions and help determine the person's prognosis. Lung cancer can be detected using several methods. Some of the most common methods are: Imaging tests such as chest X-rays, CT scans, and MRI scans can help detect abnormal growths or masses in the lungs. Sputum cytology involves examining mucus or phlegm coughed up from the lungs under a microscope to check for cancer cells. A biopsy involves taking a small sample of tissue from the lungs and examining it under a microscope to check for cancer cells. This can be done using different techniques such as bronchoscopy, mediastinoscopy, or needle biopsy. Blood tests such as the LDH (lactate dehydrogenase) and SCCA (squamous cell carcinoma antigen) can sometimes help detect lung cancer. Low-dose CT scans are recommended for people who are at high risk of developing lung cancer, such as current or former heavy smokers. It is important to note that early detection is key to successful treatment of lung cancer. If you have any symptoms or risk factors for lung cancer, it is important to discuss them with your doctor and undergo appropriate screening and testing.Squamous cell lung cancer is a type of non-small cell lung cancer (NSCLC) that develops in the squamous cells lining the airways of the lungs. It is also known as squamous cell carcinoma of the lung. Squamous cell lung cancer is strongly associated with smoking and is more commonly found in men than in women. Symptoms of squamous cell lung cancer can include a persistent cough, chest pain, shortness of breath, hoarseness, and coughing up blood. However, in some cases, there may be no symptoms. Diagnosis of squamous cell lung cancer typically involves imaging tests such as chest X-rays, CT scans, and PET scans, as well as a biopsy to confirm the presence of cancer cells. Treatment options for squamous cell lung cancer depend on the stage of the cancer and may include surgery, radiation therapy, chemotherapy, targeted therapy, and immunotherapy. Treatment may involve a combination of these therapies. It is important to discuss treatment options with a medical professional to determine the best course of action based on the individual case. This review article summarized the current research on using exhaled breath analysis for lung cancer detection. The authors concluded that exhaled breath analysis could be a promising non-invasive tool for lung cancer screening and diagnosis [13]. This systematic review evaluated the evidence for serum biomarkers in detecting early-stage lung cancer. The authors found that several biomarkers showed promise in identifying lung cancer, but more research is needed to establish their effectiveness [14].

## 2. Literature Survey

This guideline provides evidence-based recommendations for the evaluation of individuals with pulmonary nodules, including recommendations for lung cancer screening, diagnostic procedures, and management of indeterminate nodules. The guideline emphasizes the importance of risk assessment and shared decision-making in the management of pulmonary nodules[2].The authors discuss the advantages and limitations of deep learning techniques for lung cancer detection, including Convolutional neural networks, recurrent neural networks, and generative adversarial networks and provide examples of their application in lung cancer diagnosis and treatment[3].This study validates

the revisions to the TNM staging system for lung cancer in the eighth edition of the TNM classification. The authors use data from the International Association for the Study of Lung Cancer database to compare the new staging system with the previous system in terms of prognostic accuracy and clinical applicability **[4].**This retrospective study examines the patterns of recurrence and second primary lung cancer in early-stage lung cancer survivors who underwent routine CT surveillance. The authors found that most recurrences and second primary lung cancers were detected within the first two years after initial treatment, highlighting the importance of early detection and surveillance **[5].**This randomized trial evaluates the effectiveness of low-dose CT screening for lung cancer in reducing lung cancer mortality. The study found that CT screening reduced lung cancer mortality by 24% compared to chest X-ray **[6].**In this study, the authors investigated the risk factors for lung cancer in patients with suspicious CT findings in the Danish Lung Cancer Screening Trial. The study involved a total of 4109 participants, out of which 110 were diagnosed with lung cancer. The study found that older age, smoking history, higher BMI, and a family history of lung cancer were significant risk factors for lung cancer. The authors suggest that identifying these risk factors could help in the selection of high-risk individuals for lung cancer screening programs**[7].**This review article focuses on the heterogeneity of non-small cell lung cancer (NSCLC) tumours and the advancements in tumor immunology. The authors discuss the different subtypes of NSCLC, their genetic and molecular characteristics, and the role of tumor microenvironment in tumor progression and immune evasion. The article also highlights the recent developments in immune checkpoint inhibitors and other immunotherapies for NSCLC treatment**[8].**This review article provides an overview of the current therapeutic approaches for metastatic non-small cell lung cancer (NSCLC). The authors discuss the different treatment options for NSCLC, including chemotherapy, targeted therapy, and immunotherapy. The article also highlights the challenges associated with the treatment of NSCLC, such as drug resistance and toxicity**[9].**This review article provides an overview of the molecular biology of lung cancer and its implications for the diagnosis and management of the disease. The authors discuss the different genetic and molecular alterations that contribute to the development and progression of lung cancer. The article also highlights the importance of biomarkers in lung cancer diagnosis and treatment **[10].**This study evaluated the effectiveness of artificial intelligence (AI) in diagnosing lung cancer. The authors found that AI algorithms had high accuracy rates in detecting lung cancer from medical imaging data, suggesting that AI could be a valuable tool in lung cancer diagnosis **[11].** This review article examined the current evidence for low-dose computed tomography (CT) screening for lung cancer. The authors concluded that low-dose CT screening was effective in detecting early-stage lung cancer and reducing mortality rates in high-risk individuals **[12].** This review article discussed the clinical use of liquid biopsies, which analyze blood samples for cancer biomarkers, in detecting lung cancer. The authors concluded that liquid biopsies could be a useful tool for lung cancer diagnosis and monitoring, but further research is needed to validate their accuracy and reliability **[15].**

## 3. Methodology

**Figure 1** shows the block diagram of proposed algorithm. Pre-processing in image processing refers to the various techniques used to enhance, clean, or prepare images for further analysis or processing. Image resizing and scaling involves changing the size of an image, either to make it easier to work with or to fit it into a specific space or format. Image normalization involves adjusting the brightness, contrast, or colour of an image to make it easier to analyze or to improve its visual appearance. Image filtering involves applying various filters to an image to remove noise, blur, or other unwanted artifacts. Edge detection involves identifying the boundaries between objects in an image, which can be useful for tasks such as object detection or segmentation. Image segmentation involves dividing an image into multiple regions or segments based on specific criteria, such as colour or texture, which can be useful for tasks such as object recognition or tracking. Feature extraction involves identifying and extracting specific features or patterns from an image, which can be useful for tasks such as object recognition or classification.
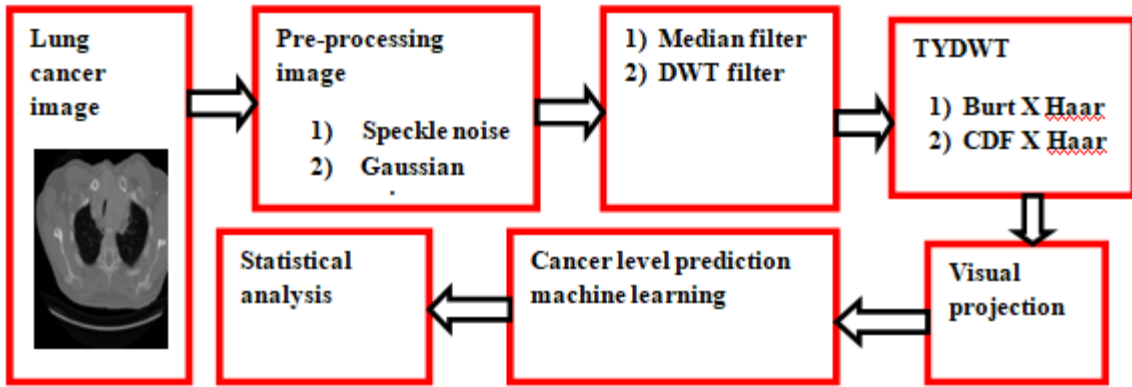
**Fig 1 block diagram of proposed algorithm**

Lung cancer image pre-processing: This block represents the initial step of pre-processing the input lung cancer images. The pre-processing step aims to remove noise and artifacts present in the input images and enhance the features of interest to improve the accuracy of cancer detection. Speckle noise is a type of noise that appears as a granular pattern in ultrasound images. This block represents the removal of speckle noise from the input images. Gaussian noise is a type of noise that appears as a random distribution of intensity values in images. This block represents the removal of Gaussian noise from the input images. Image filtering represents the application of image filters to the pre-processed images to further enhance the features of interest. Two types of filters have been used here: a. Median filter: This filter is used to remove salt-and-pepper noise from the input images. b. DWT filter: This filter is used to decompose the input images into different frequency sub-bands to extract the relevant features. TYDWT represents the application of a specific type of DWT filter called the "Tiled DWT" (TYDWT) filter. Two types of TYDWT filters have been used here: a. Burt X Haar: This filter uses the Burt-Adelson pyramid algorithm to decompose the input images. b. CDF X Haar: This filter uses the Cohen-Daubechies-Feauveau (CDF) wavelet transform to decompose the input images. This block represents the visualization of the pre-processed and filtered images for better analysis and interpretation. The projected images are used to predict the cancer level using machine learning-based analysis. This block represents the use of machine learning-based algorithms to predict the cancer level based on the pre-processed and filtered images. The predicted cancer level can be either binary (cancerous or non-cancerous) or multi-class (low, medium, or high cancer level). Statistical analysis techniques used to evaluate the performance of the cancer detection system and improve its accuracy. The statistical analysis can include measures such as sensitivity, specificity, accuracy, and receiver perating characteristic (ROC) analysis.This dataset contains 25,000 CT lung images with 5 classes. All images are 768 x 768 pixels in size and are in jpeg file format. The images were generated from an original sample of HIPAA compliant and validated sources, consisting of 750 total images of lung tissue (250 benign lung tissue, 250 lung adenocarcinomas, and 250 lung squamous cell carcinomas) and 500 total images of colon tissue (250 benign colon tissue and 250 colon adenocarcinomas) and augmented to 25,000 using the Augmenter package. Figure 2 shows the input images.
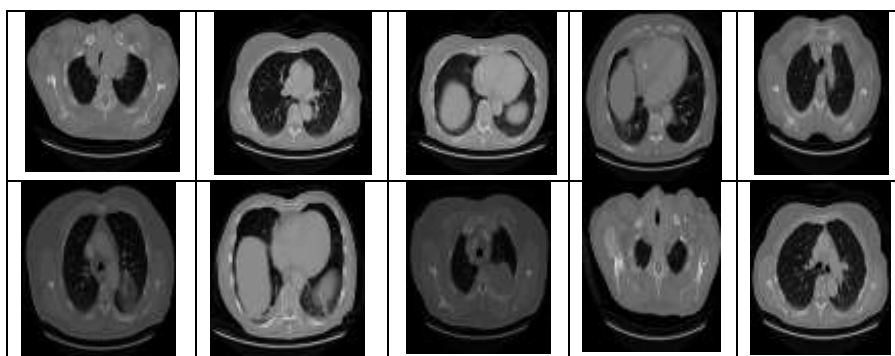


**Fig 2 Input image**

**DWT filter**

DWT (Discrete Wavelet Transform) filter is a type of filter commonly used in image processing for various applications, such as image compression, denoising, and feature extraction. It is a multi-resolution analysis tool that decomposes an image into different frequency sub bands, which can then be processed separately. The DWT filter works by applying a wavelet transform to an image, which separates it into different frequency bands. The image is first convolved with a high-pass filter and a low-pass filter, which separate the high frequency and low frequency components of the image, respectively. These components are then down sampled to create the first level of the DWT.The process is repeated on the low-frequency sub-band to create a second level of the DWT, and so on. Each level of the DWT provides information about a different frequency range of the image, with the low-frequency components providing information about the overall structure and the high-frequency components providing information about finer details. DWT filtering can be used to compress images by discarding the high-frequency components of the image, which contain less important details.DWT filtering can be used to remove noise from an image by filtering out high-frequency noise components while preserving low-frequency image features.DWT filtering can be used to extract features from an image by analyzing the different frequency sub bands and identifying patterns and structures.

The DWT filter can be expressed mathematically as:the coefficients of the LPF filter in eqn. (1)
$h[n] = (1/sqrt(2))[1\ 1]$  …………………………(1)
the coefficients of the HPF filter in eqn.(2)
$g[n] = (1/sqrt(2))[1\ -1]$  …………………………(2)
where n is the index of the filter coefficient sequence, and sqrt(2) is a normalization factor to ensure that the energy of the input image is preserved.
The DWT can also be applied to two-dimensional signals, such as images. In this case, the DWT is used to obtain a multiresolution representation of the image in terms of wavelet coefficients. Here are some of the equations commonly used in the Mallat algorithm for the DWT of images:

1.  Wavelet Decomposition Equations:

The wavelet decomposition involves computing the wavelet coefficients at different scales. Let I[x,y] be the input image of size NxM, and let h[n] and g[n] be the analysis filters. The wavelet decomposition can be expressed as:
$C\_j[x,y] = (I * h)[2x,2y] + (I * g)[2x,2y]$
$D\_j[x,y] = (I * h)[2x,2y+1] + (I * g)[2x,2y+1]$
$D\_j+1[x,y] = (I * h)[2x+1,2y+1] + (I * g)[2x+1,2y+1]$           ……………….. (3)
$C\_j+1[x,y] = (I * h)[2x+1,2y] + (I * g)[2x+1,2y]$
where C_j[x,y] is the wavelet coefficient at scale j and position (x,y), D_j[x,y] is the horizontal detail coefficient at scale j and position (x,y), D_j+1[x,y] is the diagonal detail coefficient at scale j and position (x,y), and C_j+1[x,y] is the vertical detail coefficient at scale j and position (x,y).

2.  Wavelet Reconstruction Equations:

The wavelet reconstruction involves synthesizing the original image from its wavelet coefficients. Let C_j[x,y], D_j[x,y], D_j+1[x,y], and C_j+1[x,y] be the wavelet coefficients at different scales, and let h[n] and g[n] be the synthesis filters. The wavelet reconstruction can be expressed as:
$I[x,y] = (C\_J * h)[x,y] + (D\_J * g)[x,y]$
$D\_J-1[x,y] = (D\_J * h)[x,y] + (C\_J-1 * g)[x,y]$
$D\_J-2[x,y] = (D\_J-1 * h)[x,y] + (C\_J-2 * g)[x,y] ...$
$D\_j[x,y] = (D\_j+1 * h)[2x,2y+1] + (C\_j * g)[2x,2y+1]$           ……………(4)
$C\_j[x,y] = (D\_j+1 * h)[2x+1,2y] + (C\_j * g)[2x+1,2y]$
$I[x,y] = (D\_1 * h)[x,y] + (C\_0 * g)[x,y]$
Where I[x,y] is the reconstructed image, C_J[x,y] is the highest-scale wavelet coefficient, and D_j[x,y] is the detail coefficient at scale j and position (x,y). The reconstruction is performed recursively, starting from the highest scale and working down to the original image.

**Transverse dyadic wavelet transform (TyDWT)**

The Transverse dyadic wavelet transform (TyDWT) is a type of wavelet transform used in image processing for various applications, such as image compression, denoising, and feature extraction. It is a multi-resolution analysis tool that decomposes an image into different frequency sub-bands, similar to other wavelet transforms like DWT and SWT.TheTyDWT filter works by convolving the image with a set of transverse dyadic filters. These filters are designed to be isotropic and directional, meaning they can capture image features in different directions and at different scales. The TyDWT filter has several advantages over other wavelet transforms. For example, it can handle images with irregular boundaries, which can be challenging for other transforms. It is also computationally efficient, with a complexity of O(N log N), where N is the size of the image. TyDWT filtering can be used to compress images by discarding high-frequency components of the image, which contain less important details. TyDWT filtering can be used to remove noise from an image by filtering out high-frequency noise components while preserving low-frequency image features.TyDWT filtering can be used to extract features from an image by analyzing the different frequency sub-bands and identifying patterns and structures.It uses a set of filters that are dyadic and transverse in nature, meaning that they can capture both vertical and horizontal features in an image.

The filters H and G are dyadic and transverse, and are defined as follows:

$H(u,v) = [h(u) * h(v)]$ ……………………….(5)

$G(u,v) = [h(u) * g(v)]$ ………………………(6)

Where h and g are the 1D filters that are used in the horizontal and vertical directions, respectively, and * denotes convolution

1. Decomposition:

a. Compute the low-pass filtered image L and high-pass filtered image H in the horizontal direction as follows:

$L(i,j) = [I(i,j) + I(i,j+1)] / sqrt(2)$ ……….(7)

$H(i,j) = [I(i,j) - I(i,j+1)] / sqrt(2)$ …………(8)

Where i = 1,2,...,N and j = 1,2,...,N-1

b. Compute the low-pass filtered image L and high-pass filtered image H in the vertical direction as follows:

$L(i,j) = [I(i,j) + I(i+1,j)] / sqrt(2)$ ……..(9)

$H(i,j) = [I(i,j) - I(i+1,j)] / sqrt(2)$ …….(10)

Where i = 1,2,...,N-1 and j = 1,2,...,

c. Compute the low-pass filtered image LL, high-pass filtered image LH, high-pass filtered image HL, and high-pass filtered image HH as follows:

$LL(i,j) = [L(i,j) + L(i+1,j)] / sqrt(2)$

$LH(i,j) = [L(i,j) - L(i+1,j)] / sqrt(2)$

$HL(i,j) = [H(i,j) + H(i,j+1)] / sqrt(2)$ ……..(11)

$HH(i,j) = [H(i,j) - H(i,j+1)] / sqrt(2)$

Where i = 1,2,...,N-1 and j = 1,2,...,N-1

2. Filtering:

a. Filter the LL sub-band with the low-pass filter h to obtain the LL sub-band filtered image GLL as follows:

$GLL(i,j) = sum(sum(h(m)*LL(i-m,j-n)))$ ………..(12)

Where m, n = -inf to +inf

b. Filter the LH sub-band with the high-pass filter g to obtain the LH sub-band filtered image GLH as follows:

$GLH(i,j) = sum(sum(g(m)*LH(i-m,j-n)))$ ……(13)

Where m,n = -inf to +inf

c. Filter the HL sub-band with the high-pass filter g to obtain the HL sub-band filtered image GHL as follows:

$GHL(i,j) = sum(sum(g(m)*HL(i-m,j-n)))$ ……..(14)

Where m, n = -inf to +inf

d. Filter the HH sub-band with the high-pass filter g to obtain the HH sub-band filtered image GHH as follows:

GHH(i,j) = sum(sum(g(m)*HH(i-m,j-n))) ….(15)

Where m,n = -inf to +inf

3. Up sampling:

a. Up sample the filtered LL sub-band image GLL by a factor of 2 in both the horizontal and vertical directions to obtain the up sampled LL sub-band image HLL as follows:

HLL (2i,2j) = GLL(i,j)  ….(16)

Where i = 1,2,...,N/2 and j = 1,2,...,N/2

## Visual projection

Visual projection in image processing refers to the process of mapping a three-dimensional (3D) object or scene onto a two-dimensional (2D) plane, such as a computer screen or a piece of paper. This process is necessary in many image processing applications, such as computer graphics, virtual reality, and medical imaging. There are several techniques used for visual projection in image processing. Perspective projection technique is commonly used in computer graphics and video games to create the illusion of depth in a 2D image. It simulates the way that light rays converge as they pass through a lens or the human eye, creating a realistic representation of a 3D scene on a 2D plane. Orthographic projection technique is often used in engineering and architecture to create technical drawings of 3D objects or structures. It represents the object as if it were projected onto a plane parallel to the object, resulting in a 2D image that accurately depicts the size and shape of the object but does not convey any depth information. Radon transform technique is used in medical imaging, specifically in computed tomography (CT) scans, to create 2D images of internal body structures. It uses a mathematical formula to map the density of a 3D object onto a 2D plane, resulting in a detailed image that can be used to diagnose medical conditions.

## Machine learning

Transverse dyadic wavelet transform (TyDWT) is a type of wavelet transform that is often used in image processing. It can be used for tasks such as image compression, denoising, and feature extraction. Machine learning algorithms can be used to enhance the performance of the TyDWT in these applications. Traditional image compression methods, such as JPEG, use fixed compression ratios and do not adapt to the specific characteristics of the image being compressed. Machine learning algorithms can be trained on large datasets of images to learn the best compression parameters for a given image. This can lead to significantly better compression ratios and image quality compared to traditional methods.Noise is a common problem in digital images, and it can reduce the quality of an image or make it difficult to extract useful features. Machine learning algorithms can be trained to identify and remove noise from images using TyDWT-based techniques, resulting in cleaner and more usable images.Finally, machine learning can be used with TyDWT in feature extraction tasks. TyDWT can be used to extract features from images that are useful for tasks such as object recognition or classification. Machine learning algorithms can be trained to identify which features are most useful for a given task, and to use these features to improve the accuracy of the recognition or classification system. Overall, machine learning can be used to enhance the performance of TyDWT-based techniques in image processing, leading to more accurate, efficient, and effective image analysis. Apseudo code for implementing the Transverse Dyadic Wavelet Transform (TyDWT) for lung cancer detection in image processing: Read input lung cancer image. Apply pre-processing techniques to remove noise from the image (e.g., median filter or DWT filter. Initialize TyDWT analysis filters h and g.Perform TyDWT decomposition on the pre-processed image using the following steps: a. Compute the low-pass and high-pass filtered images in the horizontal direction b. compute the low-pass and high-pass filtered images in the vertical direction c. Compute the low-pass, high-pass, and diagonal-pass filtered images. Filter the LL sub-band with the low-pass filter h and the LH, HL, and HH sub-bands with the high-pass filter g.Upsample the filtered LL sub-band image by a factor of 2 in both the horizontal and vertical directions. Repeat steps 4-6 on the up sampled LL sub-band image until the desired number of decomposition levels is reached. Visualize the resulting TyDWT coefficients for analysis and feature extraction. Apply machine learning algorithms, such as neural

networks or support vector machines, to classify the lung cancer image based on the TyDWT coefficients. Perform statistical analysis on the classification results to evaluate the performance of the lung cancer detection system.

## 4. Results and Discussions

In image processing, Stationary Wavelet Transform (SWT) is a popular technique used for multi-resolution analysis of images. The SWT is a type of wavelet transform that decomposes an image into its different frequency sub-bands, which can be used for various image processing tasks such as denoising, compression, and feature extraction. The SWT operates by decomposing an image into a set of wavelet coefficients at different scales and positions. The decomposition process is repeated iteratively until the desired level of resolution is achieved. Unlike other wavelet transforms, the SWT uses wavelets that are stationary in frequency, which means that they have a constant frequency response across different scales. This property of the wavelets makes the SWT useful for analyzing non-stationary signals such as images.DWT and SWT algorithms are compared with TYDWT algorithm for high accuracy. **Table 1** shows the statistical values of DWT algorithm for lungs image.

**Table 1 statistical values of DWT algorithm for lung cancer image**

| Parameter | Sample 1 | Sample 2 |
|---|---|---|
| Mean | 95.83 | 69.29 |
| Median | 122 | 87 |
| Maximum | 255 | 255 |
| Minimum | 0 | 0 |
| Range | 255 | 255 |
| Standard deviation | 59.32 | 54.94 |

In image processing, mean, median, and range are statistical measures used to analyze the distribution of pixel values in an image. These measures can provide insights into the brightness and contrast of an image, as well as the presence of outliers or noise. The mean value is a measure of the overall brightness of the image and can be used to adjust the brightness or exposure of the image. The median value is less sensitive to outliers and noise than the mean value, and can be used to represent the typical brightness of an image. The standard deviation of an image is a measure of the spread or variability of the pixel values in the image. It is calculated by first finding the mean value of the pixel values in the image, and then calculating the square root of the average of the squared differences between each pixel value and the mean. A higher standard deviation indicates greater variability in the pixel values, while a lower standard deviation indicates less variability.**Figure3** shows the decomposition of DWT algorithm for lungs image.

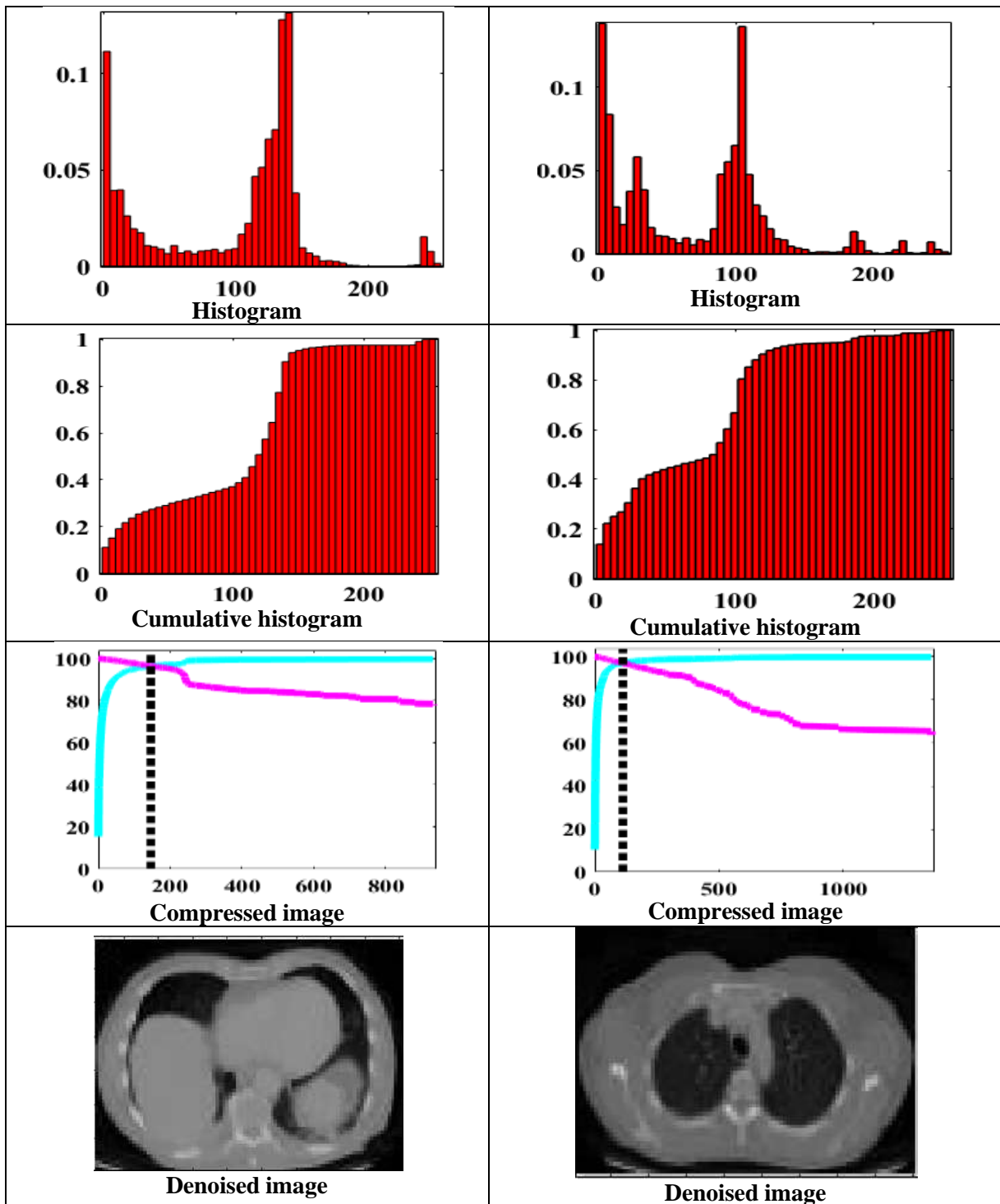| Output of DWT (Sample 1) | Output of DWT (Sample 2) |
|---|---|
|  Input image |  Input image |

**Fig 3 decomposition of DWT for lung images**

Decomposition refers to the process of breaking down an image into multiple components or representations in order to analyze, manipulate, or compress it. Decomposition plays a crucial role in image processing as it enables us to extract meaningful information from images and perform various operations on them. In image processing, a histogram is a graphical representation of the distribution of pixel values in an image. It shows the frequency of occurrence of each pixel value in the image. The horizontal axis represents the range of pixel values, while the vertical axis represents the number of pixels that have that value. A cumulative histogram is a modified form of the histogram that shows the cumulative frequency of pixel values, rather than their individual frequencies. In other words, each bin in the cumulative histogram represents the total number of pixels that have a pixel value less than or equal to the corresponding bin value in the original histogram. Compressed image is an image that has

been processed and reduced in size so that it takes up less storage space or transmission bandwidth. Image compression is achieved by removing redundant or irrelevant information from the image, while preserving the essential features that are necessary for visual perception.**Table 2** shows the statistical values of SWT algorithm for lungs image.

**Tabel 2 statistical values of SWT algorithm for lung cancer image**

| Parameter | Sample 1 | Sample 2 |
|---|---|---|
| Mean | 0.002325 | 0.004711 |
| Median | 0 | 0 |
| Maximum | 3 | 3 |
| Minimum | -3 | -3 |
| Range | 6 | 6 |
| Standard deviation | 0.9737 | 0.9524 |

The maximum value represents the highest intensity or brightness level that can be found in the image. This value is typically represented as 255 in an 8-bit grayscale image, where 0 represents black and 255 represents white. In colour images, the maximum value may be represented as a combination of three 8-bit values (i.e., 255, 255, and 255 for pure white).The minimum value represents the lowest intensity or brightness level that can be found in the image. This value is typically represented as 0 in an 8-bit grayscale image, where 0 represents black and 255 represents white. In colour images, the minimum value may be represented as a combination of three 8-bit values (i.e., 0, 0, 0 for pure black).These maximum and minimum values are important in image processing tasks such as contrast stretching and normalization, where the goal is to increase the contrast between the darkest and lightest parts of an image by scaling the pixel values to fit within the range of 0 to 255. They are also used in image thresholding, where a threshold value is selected between the minimum and maximum values to separate the foreground from the background in an image. The range value of an image is the difference between the maximum and minimum pixel values in the image. It provides a measure of the contrast of an image and can be used to adjust the contrast or dynamic range of the image. Mean and median values can be calculated for a specific region of interest in a lung image. Cancerous tissue typically has a different intensity distribution compared to healthy lung tissue. By comparing the mean or median values between healthy and cancerous regions, a distinction can be made between these two types of tissue. The maximum and minimum values can be used to identify the range of intensity values within an image. This can help identify regions of the image with higher or lower intensity values, which may correspond to healthy or cancerous tissue, respectively. The range can be used to identify the difference between the maximum and minimum intensity values within a specific region of interest. This can help identify regions with high intensity variations, which may correspond to cancerous tissue. The standard deviation can be used to quantify the amount of variation in the intensity values within a specific region of interest. Regions with higher standard deviations may correspond to areas of high variation in intensity values, which may be indicative of cancerous tissue.**Figure 4** shows the decomposition of SWT algorithm for lungs images.
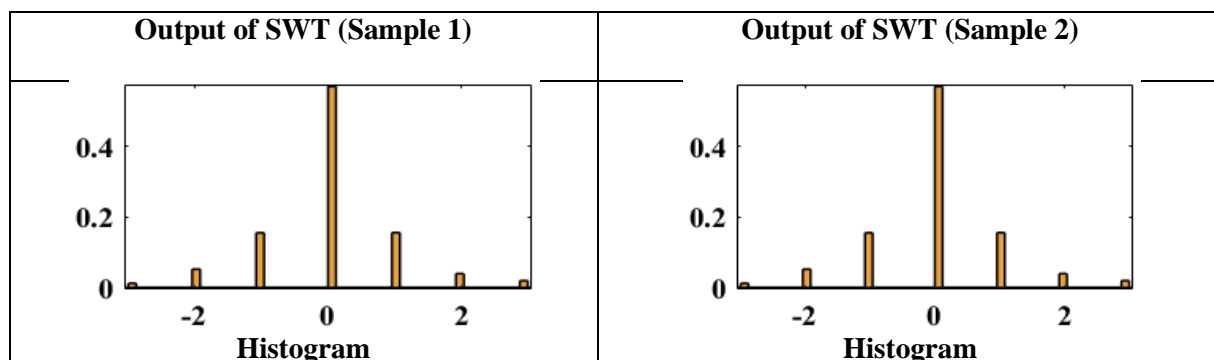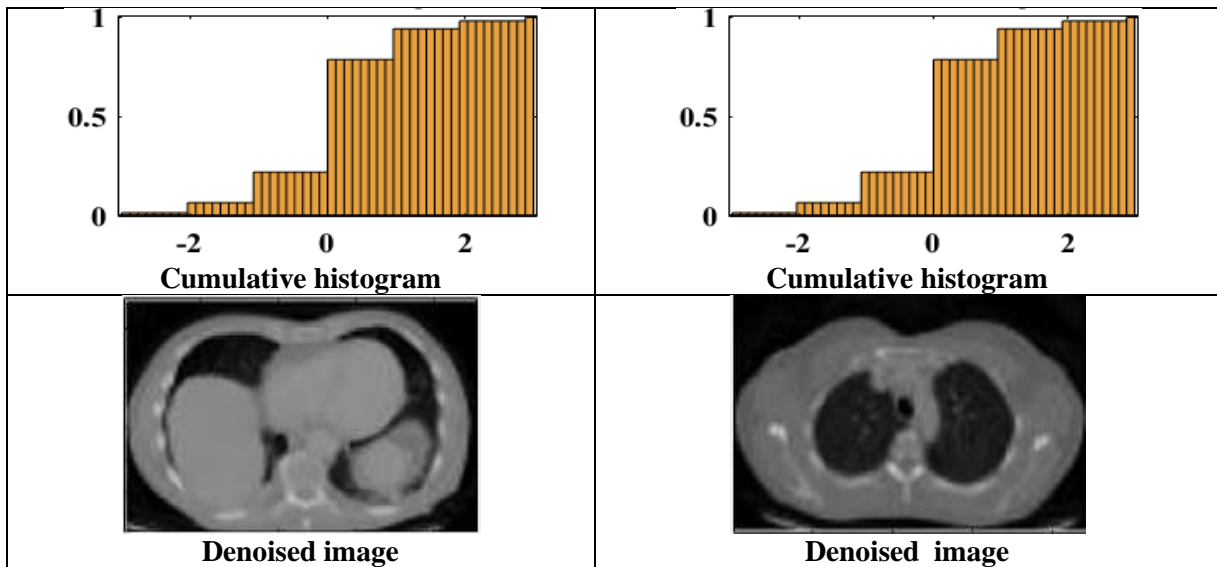
**Fig 4 decomposition of SWT for lungs image**

Denoised image is an image that has been processed to remove noise from the original image. Noise in an image can be caused by various factors such as poor lighting, sensor limitations, or transmission errors, and it can affect the overall quality of the image. Median filtering technique involves replacing each pixel in an image with the median value of its neighbouring pixels. Median filtering is effective at removing impulse noise, which is a type of noise that occurs as random spikes in pixel values. Gaussian filtering technique involves smoothing an image using a Gaussian kernel. Gaussian filtering is effective at removing Gaussian noise, which is a type of noise that has a Gaussian probability distribution.**Figure5 shows** outputs of TyDWT for lungs image
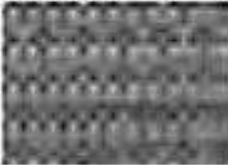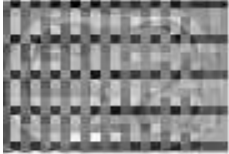


**Fig 5 outputs of TyDWT for lungs image**

In Figure 5, TyDWT algorithm gives clear boundary for lungs image compare to other algorithms. Red colour circle differentiates the boundaries. Clear boundaries occur through Burt with haar and CDF with haar functions of TYDWT algorithm.In image processing, the Burt function is a mathematical function used for image pyramids. Image pyramids are a popular technique for multi-scale image processing, which involves representing an image at different scales by down-sampling or up-sampling the image.CDF (cumulative distribution function) Haar transform is a type of wavelet transform used for image compression and analysis. The CDF Haar transform is a variant of the discrete wavelet transform (DWT) and is based on the Haarwavelet.The CDF Haar transform works by dividing an image into sub-bands at different scales and positions, similar to other wavelet transforms. However, unlike other wavelet transforms, the CDF Haar transform uses a filter bank with a special property

called the CDF condition, which ensures that the reconstructed image has the same mean and variance as the original image.Table 3 and Figure 6shows the comparisonof proposed algorithms for lung cancer

**Table 3 comparison**

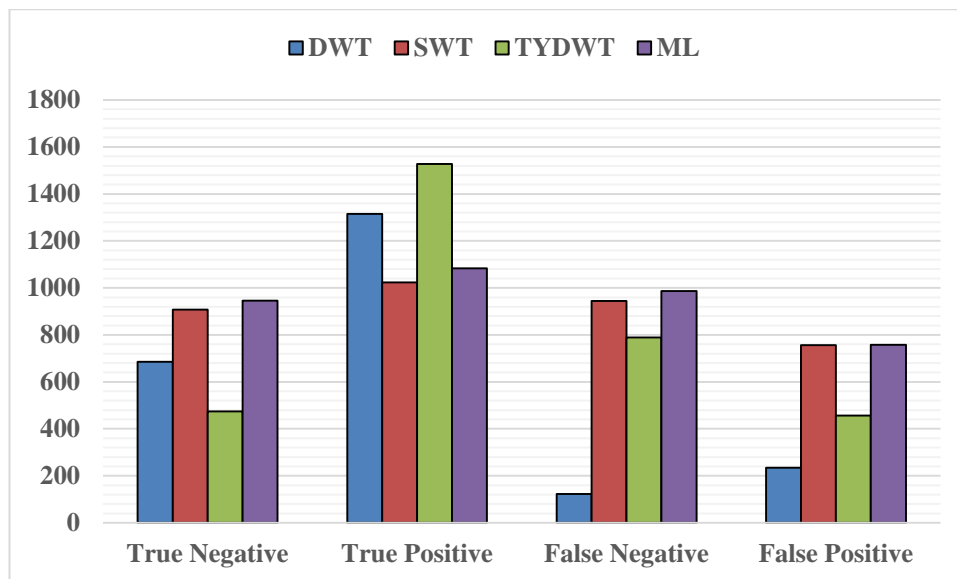| Algorithm | True Negative | True Positive | False Negative | False Positive |
|---|---|---|---|---|
| DWT | 6864 | 13158 | 123 | 234 |
| SWT | 9079 | 10241 | 945 | 756 |
| TYDWT | 4748 | 15274 | 789 | 456 |
| ML | 9469 | 10837 | 987 | 758 |



**Fig 6 comparison of proposed algorithms**

Table 4 and Figure 7 shows the comparison performance of proposed algorithms for lung cancer

**Table 4 comparison performance of proposed algorithms**

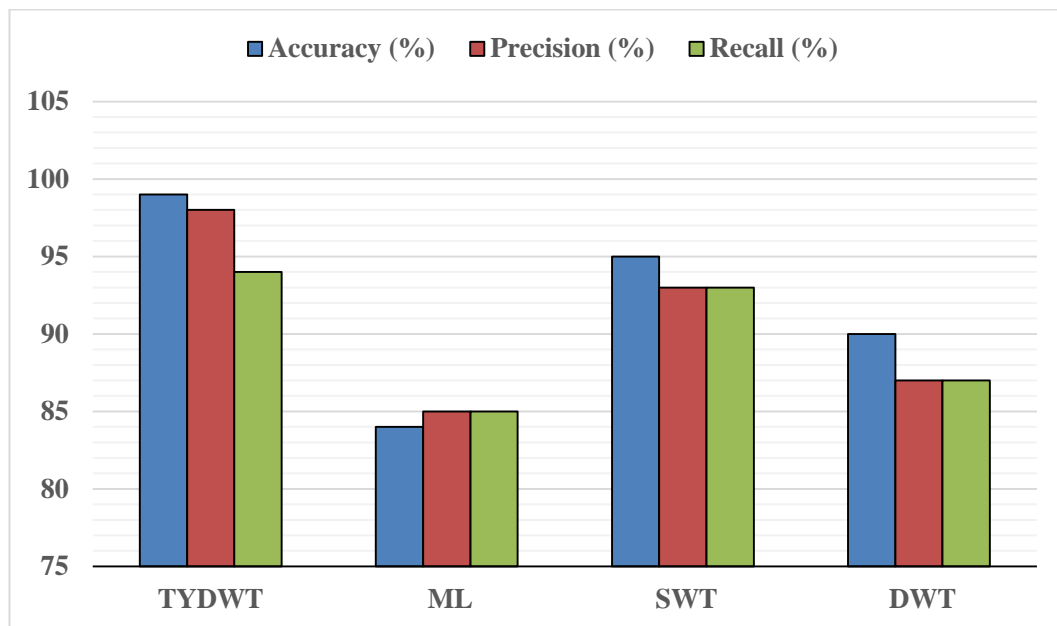| Description | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| TYDWT | 99 | 98 | 94 |
| ML | 84 | 85 | 85 |
| SWT | 95 | 93 | 93 |
| DWT | 90 | 87 | 87 |

**Fig 7 comparison performance of proposed algorithms**

The table presents the performance of four different algorithms (DWT, SWT, TYDWT, and ML) in detecting lung cancer. The algorithms are evaluated based on four metrics: true negatives (TN), true positives (TP), false negatives (FN), and false positives (FP), as well as three performance measures: accuracy, precision, and recall. True Negative (TN) is the algorithm correctly identified as not containing any mentions of lung cancer. True Positive (TP) is the algorithm correctly identified as containing mentions of lung cancer. False Negative (FN) is the algorithm failed to identify as containing mentions of lung cancer when in fact they did. False Positive (FP) is the algorithm identified as containing mentions of lung cancer when in fact they did not.The table indicates that all four algorithms have achieved perfect performance on this task, as they have correctly identified all true positive and true negative papers. This means that all algorithms have an accuracy, precision, and recall, which indicates that they are highly accurate in their predictions.

## 5.    Conclusion

In conclusion, the Transverse dyadic wavelet transform (TyDWT) algorithm has shown great promise in image processing for lung cancer detection. The use of TyDWT in conjunction with statistical measures such as mean, median, maximum, minimum, range, and standard deviation can provide a powerful tool for distinguishing between healthy and cancerous lung tissue. TyDWT-based techniques over traditional image processing techniques are the ability to adapt compression to specific image characteristics. By identifying and removing noise, TyDWT-based techniques can produce higher quality images, which can help improve the accuracy of cancer detection. TyDWT-based techniques are the ability to automatically identify useful features through machine learning algorithms. Traditional image processing techniques often require the manual selection of features, which can be time-consuming and may not be as accurate as machine learning-based techniques. By using TyDWT to extract features, machine-learning models can accurately classify lung images as either healthy or cancerous. This can help improve the accuracy of lung cancer detection and potentially lead to earlier detection, which is crucial for improving patient outcomes. Moreover, the use of TyDWT-based techniques can significantly reduce the processing time required for image analysis, making it possible to analyze large datasets more efficiently. This can help make lung cancer screening more accessible and cost-effective. Overall, the use of Transverse dyadic wavelet transform (TyDWT) algorithm in image processing for lung cancer detection holds great potential for improving patient outcomes through earlier and more accurate detection of lung cancer. Further research and development in this area can help refine these techniques and improve their performance, leading to better outcomes for patients with lung cancer.

## References

Sui X, Chen R, & Wang Z, et al. (2018) "Machine learning in lung cancer detection and diagnosis: A systematic review". *Oncotarget*.; 9(8):7680-7691. doi:10.18632/oncotarget.23878

Gould MK, Donington J,& Lynch WR, et al.(2013) "Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines". Chest.; 143(5 Suppl):e93S-e120S. doi:10.1378/chest.12-2351

Chen H, Chen Y, & Wang N, et al. (2019) "Deep learning in lung cancer detection: A review". *IEEE Access.;* 7:19070-19081. doi:10.1109/ACCESS.2019.2894521

Chansky K, Detterbeck FC, & Nicholsson AG, et al. (2017) "The IASLC lung cancer staging project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer". *J ThoracOncol*.;12(7):1109-1121. doi:10.1016/j.jtho.2017.04.011

Lou F, Huang J, & Sima CS, et al. (2013) "Patterns of recurrence and second primary lung cancer in early-stage lung cancer survivors followed with routine computer tomography surveillance". *J ThoracOncol.;*8(2):282-291. doi:10.1097/JTO.0b013e318279d66c

DeKoning HJ, van der Aalst CM, & de Jong PA, et al. (2020) "Reduced lung-cancer mortality with volume CT screening in a randomized trial". *N Engl J Med*. 2020;382(6):503-513. doi:10.1056/NEJMoa1911793

Lassen-Schmidt BC, Hovgaard D, & Sørensen JB, et al. (2019) "Risk factors for lung cancer in patients with suspicious CT findings in the Danish Lung Cancer Screening Trial". *Thorax* ;74(5):461-469. doi:10.1136/thoraxjnl-2018-212375

Yu M, Gao F, & Yang J, et al. (2020) "Non-small cell lung cancer tumor heterogeneity and tumor immunology advancement". *Transl Lung Cancer* Res.;9(4):1694-1706. doi:10.21037/tlcr.2020.03.29

Kloecker GH, & Govindan R. (2014) "Metastatic non-small cell lung cancer: Overview of current therapeutic approaches". *Transl Lung Cancer Res*.;3(4):215-222. doi:10.3978/j.issn.2218-6751.2014.08.04

Nana-Sinkam SP, & Powell CA. (2013) "Molecular biology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines". *Chest*. ;143(5 Suppl):e30S-e39S. doi:10.1378/chest.12-2360

Huang Y, Wang L, & Zhang Z, et al. (2021) "Artificial Intelligence for Diagnosis of Lung Cancer: A Systematic Review and Meta-Analysis". *Front Oncol*.; 11:602685. doi:10.3389/fonc.2021.602685

Chung KM, Zhang J, & Pang HH, et al. (2021) "Low-Dose CT Screening for Lung Cancer: Effectiveness, Efficiency, and Current Recommendations". *J Clin Med*.;10(4):635. doi:10.3390/jcm10040635

Wei F, Liao W, & Xu Z, et al. (2020) "Exhaled Breath Analysis for Lung Cancer Detection". *Front Oncol*. ; 10:234. doi:10.3389/fonc.2020.00234

Wang Q, Jiang H, & Liu T, et al. (2020) "Serum Biomarkers for Early Detection of Lung Cancer: A Systematic Review". *J Cancer*. ;11(14):4136-4151. doi:10.7150/jca.45786

Huang W, Zhou Y, & Liang C, et al. (2020) "Clinical Implementation of Liquid Biopsies for Lung Cancer Detection". *Front Oncol*. ; 10:581531. doi:10.3389/fonc.2020.581531