

# Independent Weighted Feature Set with Linked Feature Reduction Model for Lung Cancer Stage Detection using Machine Learning Model

Makineni.Siddardha Kumar<sup>1</sup>, Kasukurthi Venkata Rao<sup>2</sup>

<sup>1</sup>Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Andhra University, Visakhapatnam-530003, India

makinenisiddardha@gmail.com

<sup>2</sup>Department of Computer Science and Systems Engineering, Andhra University College of Engineering, Andhra University, Visakhapatnam-530003, India.

professor\_venkat@yahoo.com

**Abstract**—Lung cancer is a potentially fatal disease that is affected to 18% of population every year. Finding the exact location of a cancer and identification of lung cancer stage continues to be difficult for medical professionals. The true reason for cancer and a comprehensive cure is still unknown. Treatment for cancer is possible if detected at an early stage with accurate stage detection. Finding areas of the lung that have been impacted by cancer requires the use of image processing techniques like noise reduction, highlight filtration, recognizable proof of effected lung regions, and perhaps a comparison with data on the curative history of lung cancer. This research investigates whether or not technology enabled by machine learning algorithms and image processing can correctly classifies and predict lung cancer. For images, the dimensional feature channel is used in the preliminary processing stage. The proposed model considers Magnetic Resonance Imaging (MRI) images for detection of lung cancer. This research proposes an Independent Weighted Feature Set with Linked Feature Reduction (IWFS-LFR) model for accurate lung cancer stage detection based on the size of the tumour. The tumour stage can be accurately predicted using the feature attribute similarity calculation for accurate detection of lung cancer stage for proper diagnosis. The proposed model when contrasted with the traditional model exhibits better performance.

**Keywords**- Lung cancer, Classification, Computed Tomography (CT) scans, Machine Learning (ML), Artificial Neural Network (ANN).

## I. INTRODUCTION

Cancer of the lung is the most lethal form of the disease. As a result, many countries are developing diagnostic tools for lung infections that can be used early on. Reduced computed tomography (CT) screening administered three times annually to high-risk members significantly reduces success rates, as determined by the NLST study. Because of these methods, a radiologist will need to review a substantial quantity of MRI images [1]. Due to the fact that even for trained medical professionals, differentiating between injuries can be difficult, radiologists are under increasing pressure as the number of MRI scans they need to audit rises. Robotic arrangements are being developed to aid specialists in reducing their workload, advancing symptomatic exactness by minimizing subjectivity [2], accelerating analysis, and taking down restorative costs in light of the expected increase in the number of preventative/early-detection measures [3].

In order to identify cancerous cells and the stage in the lung area, it is necessary to detect and analyse a set of defining characteristics [4]. Cancer risk is determined by a mix of how often you watch highlights and how often you watch the same

ones [5]. Even for a therapeutic expert, this is difficult work because there is no simple correlation between the presence of a knob and a favourable cancer outcome. Common computer-aided demonstrative (CAD) [6] methods make use of quantity, structure, ambiguity, changeability, spiculation, homogeneity [7], and other now illustrated properties. The MRI Lung Images with Cancer Cell Size and Shapes are shown in Figure 1.

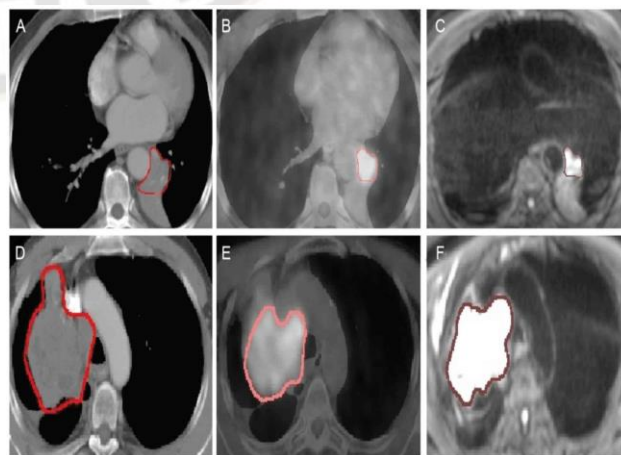


Fig 1: Numerous MRI Lung Images with Cancer Cell Size and Shape

To lessen the number of fatalities caused by lung cancer, early diagnosis and stage recognition is crucial. When monitoring those who are at a higher risk for lung cancer, MRI chest screening is the best standard. The computer-aided diagnosis (CAD) system was created to aid doctors in their assessment of medical imaging data, and it has since been proven to be an effective second opinion for doctors. Nodule feature extraction and clinical judgement inference are the original CAD task's sub-tasks [9]. Some methods for estimating malignancy risk use the measured texture features of selected nodules in MRI images in conjunction with the patient's clinical factors as input features for a machine learning classifier for accurate size detection to estimate the grade [10]. Clinical variables include age, gender, timing of specimen collection, family history of lung cancer, smoking history, and more; imaging measurements typically include nodule shape, nodule type, nodule region, nodule count and nodular boundary in MRI scans [11]. However, because to their subjective nature and lack of standardisation, these characteristics rarely provide a sufficient complete characterization of malignant lesion appearances [12].

More research has been done to implement machine learning based models in the CAD system to boost accuracy, decrease false positive rate, and speed up execution time for lung tumour grade detection, with the rise of machine learning techniques [13]. Standard procedures for these models include nodule recognition and segmentation, lesion feature extraction, and clinical judgement inference [14]. The machine learning based CAD system is superior to conventional methods because it can automatically model the structure of a nodule and retrieve distinctive features of a problematic nodule. Nodule shape descriptions were obtained using a spherical harmonic model for nodule segmentation, while nodule texture and intensity features were extracted using a machine learning model [15]. This research proposes an Independent Weighted Feature Set with Linked Feature Reduction model for accurate lung cancer stage detection based on the size of the tumour [16]. The tumour stage can be accurately predicted using the feature attribute similarity calculation for accurate detection of lung cancer stage for proper diagnosis [17].

## II. LITERATURE SURVEY

Predictive modelling of lung cancer diseases using continuous monitoring has been provided by Deepa et al. [3]. They utilised a categorization linked to a fuzzy clustering to accomplish this. Accurate image segmentation can only be achieved through the use of the fuzzy clustering method. Instead, the author used a fuzzy C-means cluster analysis approach to separate the features of the transition zone from those of the tumour picture. The Otsu thresholding technique was used in this study to separate the transition zone from the

lung cancer image. To further enhance the segmentation's presentation, the right corner picture is used in tandem with the morphometric, thinning procedure. In order to achieve classification in an incremental manner, a novel accumulative classification technique is combined with the state-of-the-art Association Rule Mining (ARM), the standard decision tree (DT), and the CNN. Operative procedures were performed using both standard images extracted from the database and the most up-to-date information on the patient's condition obtained from IoT devices worn by the patient. The study's conclusion is that the predictive modelling system has improved in precision.

Jasti et al. [5] used deep residual learning to create a technique for identifying lung cancer in CT scans. Using the UNet and ResNet models, the researchers have developed a pre-processing pipeline. The goal of this model is to identify cancerous regions of the lung and extract relevant features from those regions. In order to predict the likelihood that a CT scan shows cancer, a collective of XGBoost and random forest classifiers is used. The likeliness that a Computed tomography is malicious is then calculated as the sum of all the results predicted by the classifiers. There is an 84% improvement in accuracy using the LIDC-IRDI compared to the norm.

Lung cancer was segmented by De Potter et al. [6] using an energetic spline model. This method has been used to successfully acquire X-ray visuals of the lung. It is suggested to first use a median filter for noise detection during the pre-processing stage. Additional K-means and fuzzy C-means grouping are used for feature capture during the segmentation stage. After the X-ray image is segmented, the final result of this study's feature retrieval process is reached. The suggested model was created using the SVM technique for categorization. MATLAB is used to model the results of the cancer protection system through simulation. This research aimed to improve methods for identifying and classifying lung cancer by comparing normal and abnormal images.

An active contouring developed model by Li et al.[8] has been implemented. Lungs were separated using a variability level set function application. In order to make an accurate diagnosis of lung disease, parenchyma must be properly segmented. At present, we are comparing the under-consideration algorithm to four different kinds of active contour models. The test outcomes show that the provided method is trustworthy and can be calculated quickly. After reducing the number of features in lung CT scans, Zamani et al. [10] developed OODN (Optimal Deep Neural Network) and compared it to other classification techniques. Consequently, they were able to devise a more precise technique as a result. Lung cancer labelling has been sped up and human error has been reduced with the introduction of an

automated categorisation. Researchers found that machine learning algorithms were significantly more accurate and precise in distinguishing between normal and abnormal lung images. The study was able to correctly categorise images of the lungs with a peer selectivity of 94.56 percent, a correctness of 96.2 percent, and a sensitivity of 94.2 percent, as shown by the results. To improve CAT scans' ability to detect cancer, it is possible to implement a number of techniques that have already been demonstrated to work. The studies confirm this to be true.

When it comes to diagnosing lung cancer, Palani et al. [12] had put a premium on the application of image processing techniques. The investigation of lung cancer is getting the deep learning treatment. Lung cancer, by far the most common type of the disease, is responsible for an alarming rise number of deaths. A tomography scan was used to assess a person's risk for developing lung cancer. Nodules are a common indicator of cancer because they represent the growth of premalignant tissue. Radiologists with the proper training can identify nodules and reliably speculate on whether or not they are cancerous. False positive and false negative results can also be generated by these radiologists. Under constant duress, the patient undergoes a massive data evaluation, and a timely decision that is best for the client is made. As a result, the solution is likely to lie in creating a computer-aided detection method that can rapidly detect features using the input of radiology.

Whole-slide haematoxylin-and-eosin-stained slides of lung disease and squamous-cell carcinoma was obtained by Nithila et al. [15]. Both TCGA (The Cancer Genome Atlas) and Stanford TMA (Tissue Microarray Database) were mined for images of patients, with an additional 294 images added. Human pathology evaluations, no matter how thorough and careful, are still unable to reliably predict the outcome of a patient's treatment. The TMA cohort was used to verify the success rate of the proposed model. The results of this investigation suggest that automatically generated characteristics may aid in prognosticating lung cancer patients and, as a result, in the creation of personalised treatments.

As the leading cause of cancer-related deaths worldwide, lung cancer accounts for approximately 1.4 million deaths annually and is often not diagnosed until it has progressed significantly. Therapeutic imaging techniques such as PET and CT are used to precisely pinpoint the affected area. The current structure designed by Talukdar et al. [17] reveals how PET/CT was used to detect lung cancer and how various methods of therapeutically monitoring breathing were implemented. Imaging workflow is also explained more clearly. Here, within the confines of the current setup, the authors are not employing any particular calculations. PET/CT guided

radiation treatment and motion overseen radiotherapy are two ideas that can be integrated into patient-specific workflows. PET/CT imaging has become standard practise in lung cancer care, but it is not yet fully coordinated with treatment planning. PET and CT are being applied for the first time in lung cancer chronic care, and they are being used for decision making and disease staging respectively.

### III. PROPOSED MODEL

The current framework shows how the respiratory process was regulated and elucidates the causes of lung disease. Filtering, Histogram equalisation, Picture upgrade, clamour expulsion, etc. are all pre-processing techniques that must be applied to the images before the affected area can be removed cleanly. The goal of image pre-processing is to remove the redundant content from checked images without changing the focal points of the demonstration. Pre-processing improves the quality of every image [18]. By switching out the pixel value for a intermediate pixel esteem, a moo recurrence image was created in the proposed model.

It has been claimed that machine learning-based lung cancer grade prediction models can aid doctors in the management of ambiguous pulmonary nodules found incidentally or during screening [19]. A potential benefit of using such a system would be a reduction in the amount of benign lesions that are followed up on or worked up unnecessarily by doctors. The segmentation of the nodule is necessary because the next process, feature extraction [20], will be applied to a ROI centred on the nodule. A circular region of interest (ROI) was drawn around each nodule, and then a threshold was applied to that region. The user may then tweak the threshold to fine-tune the segmentation, and then utilise manual editing tools to get rid of any extraneous voxels that the categorization had included along with the nodule of interest. Usually, this is the method used to exclude nearby vessels [21]. In total, 12 texture features were taken from the nodule itself and the surrounding area, both of which were defined automatically. Insight is gained from realising that one set of textural features wouldn't be able to represent the patterns because of the vastly varied ranges of Hounsfield units in the nodule and the surrounding parenchyma.

We used an entirely automated feature selection approach to zero in on the bare minimum of characteristics needed to achieve optimal classification performance on training dataset. As it would be computationally prohibitive to try all possible permutations of the available set of features, we employed a algorithm that, beginning with the best possible pair of features discovered through an exhaustive search of all possible pairs, chose features in order to maximise

achievement over training data at each step. The proposed model framework is shown in Figure 2.

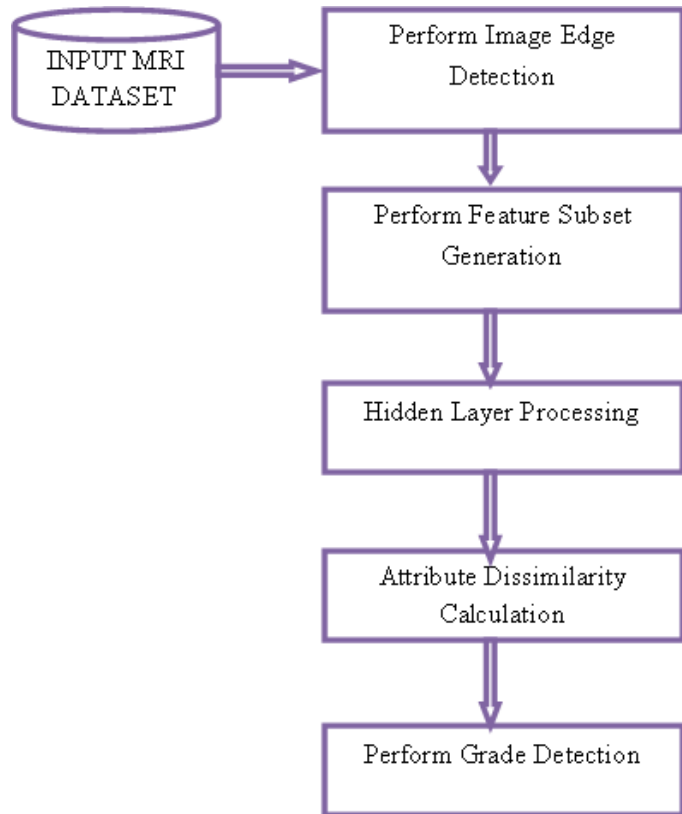


Fig 2: Proposed Model Framework

Feature extraction refers to the technique of precisely elucidating the quantity of resources needed from a large data set. When the highlights have been chosen, they need to be removed. In an essential system, the extraction structure uses mathematical and methodical methods to isolate the specific parts or forms that are sought after. The highlighted parts need to be extracted. This research proposes an Independent Weighted Feature Set with Linked Feature Reduction (IWFS-LFR) model for accurate lung cancer stage detection based on the size of the tumour.

The images in the dataset is loaded and the image analysis is performed for extracting the pixels for processing. The image analysis is performed as

$$Image(p)_{i,j}^M = \left( \sum_{i \in LDset} (IM) + \sum_{i=1}^M getIMG(i) + getIntensity(i) \right)$$

All the images in the LDset are analyzed and the current image is loaded as i, j is the next image to be loaded for analysis.

The images edge detection is performed using canny edge model for accurate edge detection so that the dissimilar regions are identified for accurate size and shape detection. The image edge detection is performed as

$$Edge(Img[p, q]) = \sqrt{\sum_{l=1}^M (getImg(p, q) - minIntensity(p, q)) + \frac{getattr(maxIntensity(p, p + 1))}{minIntensity(q, q + 1)}}$$

Here p,q are pixel coordinates of the current image loaded and the exact edges will be identified based on the pixel to pixel checking.

The features in the segments are considered and these features are used for training the model for tumour and grade detection. The feature extraction is performed as

$$FExtract(Edge(i)) = \sum_{i=1}^M \sum_{y=1}^x \max(Edge_{ij}^m | |p_j - q_i|^2 + p * q + F(getattr(p, q)))$$

Here F is the function used for considering all the feature attributes in the image.

The feature subset is generated for training the model and the features are selected based on the correlation factor. The feature subset generation and training is performed as

$$FSset(Img(p, q)) = \sum_i \sum_j \maxattr(FExtract(i, j)) + \sum_{i \in LDset} \frac{\maxattr(FExtract(p + 1, q + 1))}{len(FExtract)}$$

$\lambda$  indicates the comparative value of tumour pixels. The hidden layers are generated and considers the feature subset for processing. The hidden layers will consider the inputs and analyze the minute attributes for accurate prediction. The hidden layer processing is performed as

$$\begin{aligned}
 Hlayer(Img(p, q)) &= Link(FSset(p, q)) \\
 &+ \sum_{i=1}^{getmaxrange(FSset(i))} \lambda_{p,q}(FSset(p + 1)) \\
 &+ \min(FSset(q + 1))
 \end{aligned}$$

The dissimilarity verification is performed and the dissimilar regions are detected to identify the tumor size and shape. The dissimilarity calculation among the features and the final grade prediction is performed as

$$DisC_{i,j} = \exp\left(\frac{Img(Hlayer(p + 1) - Img(Hlayer(p)))}{\left(\frac{\lambda^2}{2}\right) * len(Hlayer)}\right)$$

$$\begin{aligned}
 GradDetset[M] &= DisC^{(i)}(p, q) \\
 &+ \sum_{i=1}^M \max(FSset\left(\left(\left(p_i - q_i^{(j)}\right)\right)\right) \\
 &+ \maxattr(FExtract(p + 1, q + 1))
 \end{aligned}$$

#### IV. RESULTS

Lung cancer consistently ranks among the deadliest human diseases. In addition, among the various types of cancer, it is one of the most frequently seen and major causes of death. Lung cancer diagnoses are on the rise. Each year, India sees around 70,000 new cases. Unfortunately, the disease is notoriously difficult to diagnose in its early stages because patients rarely exhibit any symptoms. That is why finding cancer early can be crucial in preventing deaths. Patients have a better chance of recovery and cure if their condition is identified early. Successful cancer detection relies heavily on modern technology. Based on their findings, many researchers have put forth various approaches. Several CAD techniques and systems have been proposed, developed, and introduced in recent years in an effort to leverage computer technology to address this issue. Systematic approaches of predicting cancer malignancy have been developed, and they typically make use of a combination of Machine Learning and deep learning approaches.

The difficulty in reproducing the better outcomes is a result of the large number of parameters that need to be hand-crafted in order to assess the optimal performance using these methods. Computing relies heavily on classification, the process of categorising data. Most cancer cells are stacked on top of one another in their structure. As a result, finding cancer early is harder. The development of high-performance gradient systems, phased-array receiver coils, and improved imaging sequences has made MR imaging of the lung a realistic option

in recent years. Assuming optimal conditions of effective breath-holds with consistent gating or triggering, the minimum detectable size of lung nodules using MRI is presently thought to be 3-4 mm. Under these circumstances, the diagnosis rate for nodules larger than 3 mm is 90% and the detection rate for nodules 5 mm and larger is 100%. Using the unique sensitivity profiles of individual coil components in multi-channel electromagnetic fields receive coil arrays or transmit/receive coil arrays, parallel imaging can drastically cut down on the time it takes to acquire an image. By taking use of the sparsity of the pictures in the right transform domain, the compressed sensing technique speeds up imaging acquisition from a severely undersampled data set. This research proposes an Independent Weighted Feature Set with Linked Feature Reduction (IWFS-LFR) model for accurate lung cancer stage detection based on the size of the tumour.

In this study, we use an MRI image dataset of the lungs to test our proposed EKNN approach. The dataset consists of more than 300 MRI scans from several hospitals in Tamil Nadu. The 256x256 regions were cut out of the several source photos. On the basis of the physician's recommendation, the complete dataset has already been labelled as benign or malignant. The image detection is performed to identify the accurate lung edge so that feature extraction can be done only in the edge. The Image Edge Detection Time Levels of the proposed and traditional models are shown in Figure 3.

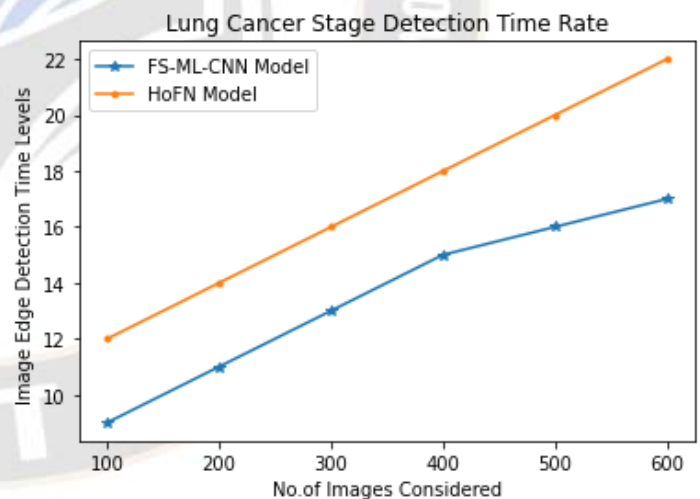


Fig 3: Image Edge Detection Time Levels

The feature subset is generated from the extracted features that are used to train the model for the accurate grade detection. The feature subset generation accuracy levels of the proposed and existing models are shown in Figure 4.

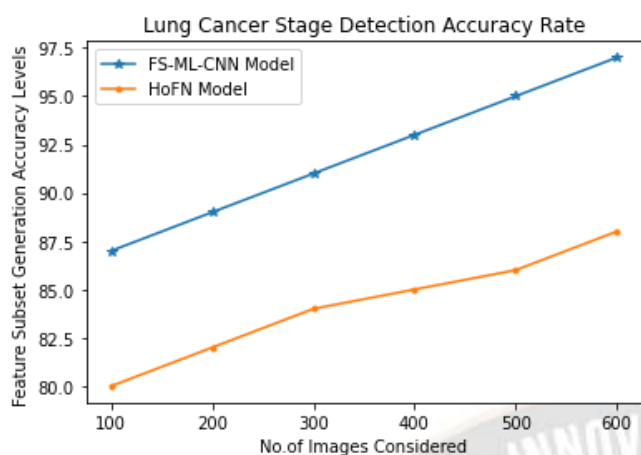


Fig 4: Feature Subset Generation Accuracy Levels

The extracted features will be processed in the hidden layers for enhancing the grade detection accuracy rate. The Hidden Layer Feature Processing Time Levels of the proposed and existing models are shown in Figure 5.

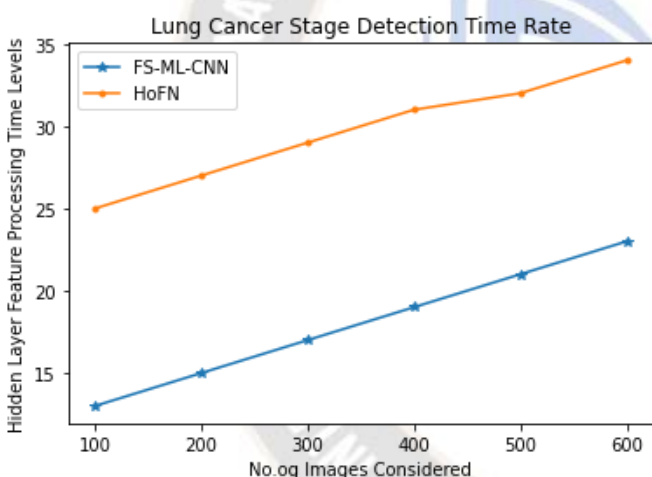


Fig 5: Hidden Layer Feature Processing Time Levels

The Attribute Dissimilarity Detection Accuracy Levels are identified for grade detection by analyzing the tumour size. The Attribute Dissimilarity Detection Accuracy Levels of the proposed and traditional models are shown in Figure 6.

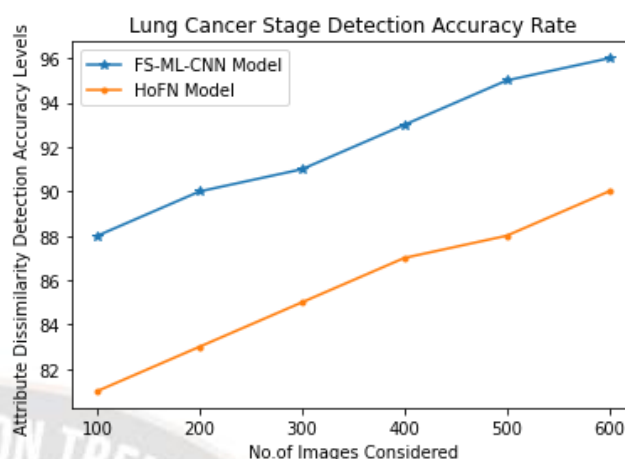


Fig 6: Attribute Dissimilarity Detection Accuracy Levels

The proposed model grade detection accuracy is high when compared to traditional model. The Grade Detection Accuracy Levels of the proposed and traditional models are shown in Figure 7.

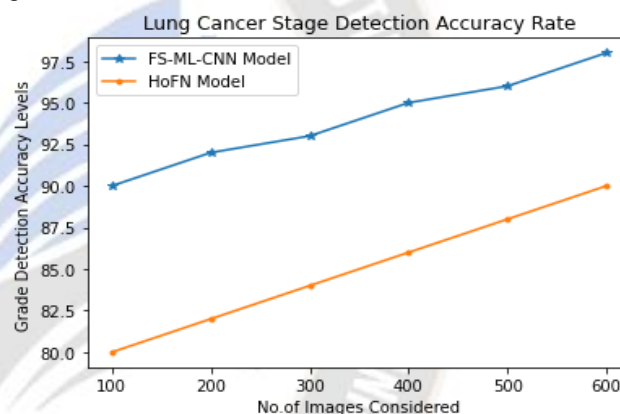


Fig 7: Grade Detection Accuracy Levels

## V. CONCLUSION

About one million people per year lose their lives due to lung cancer, making it one of the most lethal forms of the disease. Lung nodule identification on MRI scans is urgently needed in the current medical climate. Therefore, CAD systems are essential for detection of lung cancer stage. With its widespread practical application, image processing is an indispensable skill. Cancerous tumours in the body can be detected through MRI imaging of a lungs. Locating areas of the lung that are impacted by cancer requires the use of image processing techniques like noise removal, extraction of features, identification of defected regions, and possibly comparing with data on the health information of lung cancer. This research demonstrates the feasibility of using machine learning and image processing to reliably classify and predict lung cancer stage. The first step is to gather a collection of photographs. After that, a calculated filter is applied to the images as a form of pre-processing. As a result, picture quality

improves over time. The images are then divided into regions. The cancer lesion is detected and the stage identification is performed. This research proposes an Independent Weighted Feature Set with Linked Feature Reduction model for accurate lung cancer stage detection based on the size of the tumour. The proposed model achieves 98% grade detection accuracy when compared to traditional models. In future, minimal feature set can be applied for minor grade detections also to enhance the detection rate.

## References

- [1]. J. Zhou et al., "Cohesive Multi-Modality Feature Learning and Fusion for COVID-19 Patient Severity Prediction," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 5, pp. 2535-2549, May 2022, doi: 10.1109/TCSVT.2021.3063952.
- [2]. Ahmed Medjahed S., AitSaadi T., Benyettou A., Ouali M. Kernel-based learning and feature selection analysis for cancer diagnosis. *Applied Soft Computing* .2017;51:39–48. doi: 10.1016/j.asoc.2016.12.010.
- [3]. Deepa N., Prabadevi B., Maddikunta P. K., et al. An AI-based intelligent system for healthcare analysis using Ridge-Adaline Stochastic Gradient Descent Classifier. *The Journal of Supercomputing*. 2021;77(2):1998–2017. doi: 10.1007/s11227-020-03347-2.
- [4]. Liu C., Hu S. C., Wang C., Lafata K., Yin F. F. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. *Quantitative Imaging in Medicine and Surgery*. 2020;10(10):1917–1929. doi: 10.21037/qims-19-883.
- [5]. Jasti V., Zamani A. S., Arumugam K., et al. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. *Security and communication networks*. 2022;2022:7. doi: 10.1155/2022/1918379.1918379.
- [6]. De Potter B., Huyskens J., Hiddinga B., et al. Imaging of Urgencies and Emergencies in the Lung Cancer Patient. *Insights into imaging*. 2018;9(4):463–476. doi: 10.1007/s13244-018-0605-6.
- [7]. Chaudhury S., Krishna A. N., Gupta S., et al. Effective image processing and segmentation-based machine learning techniques for diagnosis of breast cancer. *Computational and Mathematical Methods in Medicine*.2022;2022:6. doi: 10.1155/2022/6841334.6841334.
- [8]. Anthony Thompson, Ian Martin, Alejandro Perez, Luis Rodriguez, Diego Rodríguez. Utilizing Machine Learning for Educational Game Design. *Kuwait Journal of Machine Learning*, 2(2). Retrieved from <http://kuwaitjournals.com/index.php/kjml/article/view/183>
- [9]. Li J., Wang Y., Song X., Xiao H. Adaptive multinomial regression with overlapping groups for multi-class classification of lung cancer. *Computers in Biology and Medicine* .2018;100:1–9. doi: 10.1016/j.compbiomed.2018.06.014.
- [10]. Halder A., Kumar A. Active learning using Fuzzy-Rough Nearest Neighbor classifier for cancer prediction from microarray gene expression data. *Journal of Biomedical Informatics* . 2020;34(1):p. 2057001. doi: 10.1142/S0218001420570013.
- [11]. Zamani A. S., Anand L., Rane K. P., et al. Performance of machine learning and image processing in plant leaf disease detection. *Journal of Food Quality* .2022;2022:7. doi: 10.1155/2022/1598796.1598796.
- [12]. Sandhiya S., Kalpana Y. An artificial neural network (ANN) based lung nodule identification and verification module. *Medico-Legal Update*. 2019;19(1):p. 193. doi: 10.5958/0974-1283.2019.00039.2.
- [13]. Palani D., Venkatalakshmi K. An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification. *Journal of Medical Systems*. 2019;43(2):p. 21. doi: 10.1007/s10916-018-1139-7.
- [14]. Bhatia S., Sinha Y., Goel L. *Soft Computing for Problem Solving*. Singapore: Springer; 2019. Lung cancer detection: a deep learning approach; pp. 699–705.
- [15]. Joon P., Bajaj S. B., Jatain A. *Progress in Advanced Computing and Intelligent Engineering*. Singapore: Springer; 2019. Segmentation and detection of lung cancer using image processing and clustering techniques; pp. 13–23.
- [16]. Nithila E. E., Kumar S. S. Segmentation of lung from CT using various active contour models. *Biomedical Signal Processing and Control*.2019;47:57–62. doi: 10.1016/j.bspc.2018.08.008.
- [17]. Lakshmanaprabu S. K., Mohanty S. N., Shankar K., Arunkumar N., Ramirez G. Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems* .2019;92:374–382. doi: 10.1016/j.future.2018.10.009.
- [18]. J. N. S. S., J. N. ., & E. N., G. . (2023). A Novel Blockchain-Based Lightweight Encryption Technique in Fog Based IoT for Personal Healthcare Data Application. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3s), 119 -. Retrieved from <https://ijisae.org/index.php/IJISAE/article/view/2549>
- [19]. Talukdar J., Sarma P. A survey on lung cancer detection in CT scans images using image processing techniques. *International Journal of Current Trends in Science and Technology*. 2018;8(3):20181–20186.
- [20]. Yu K. H., Zhang C., Berry G. J., et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications*. 2016;7(1):p. 12474. doi: 10.1038/ncomms12474.
- [21]. Cirujeda P., Cid Y. D., Muller H., et al. A 3-D Riesz-covariance texture model for prediction of nodule recurrence in lung CT. *IEEE transactions on medical imaging* . 2016;35(12):2620–2630. doi: 10.1109/TMI.2016.2591921.
- [22]. Sangamithraa P. B., Govindaraju S. Lung tumour detection and classification using EK-mean clustering. *Proceedings of the 2016 IEEE International Conference on Wireless Communications, Signal Processing and Networking, WISPNET; 2016; Chennai, India.*

- 
- [23]. Kurkure M., Thakare A. Lung cancer detection using genetic approach. Proceedings -2nd International Conference on Computing, Communication, Control and Automation, ICCUBEA; 2017; Pune, India.
- [24]. Kureshi N., Abidi S. S. R., Blouin C. A predictive model for personalized therapeutic interventions in non-small cell lung cancer. *IEEE journal of biomedical and health informatics*. 2016;20(1):424–431. doi: 10.1109/JBHI.2014.2377517.
- [25]. Kumar A., Gautam B., Dubey C., Tripathi P. K. A review: role of doxorubicin in treatment of cancer. *International Journal of Pharmaceutical Sciences and Research*. 2014;5(10):4117–4128.
- [26]. Degadwala, D. S. ., & Vyas, D. . (2021). Data Mining Approach for Amino Acid Sequence Classification . *International Journal of New Practices in Management and Engineering*, 10(04), 01–08. <https://doi.org/10.17762/ijnpme.v10i04.124>
- [27]. Kulkarni A., Panditrao A. Classification of lung cancer stages on CT scan images using image processing. *IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT*; 2014; Ramanathapuram, India. 2014. pp. 1384–1388.
- [28]. Westaway D. D., Toon C. W., Farzin M., et al. The International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society grading system has limited prognostic significance in advanced resected pulmonary adenocarcinoma. *Pathology* 2013;45(6):553558. doi:10.1097/PAT.0b013e32836532ae.
- [29]. Chaudhary A., Singh S. S. Lung cancer detection on CT images by using image processing. Proceedings: Turing 100 - International Conference on Computing Sciences, ICCS; 2012; Phagwara, India.

