iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71901

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Enhancing Lung Cancer Diagnosis: A Classification Based Machine Learning Approach

Janakiraman S[1], Priyadharshini K[2]
[1]Assistant Professor, [2]MCA, Department of Master of Computer Applications, Er.Perumal Manimekalai College of Engineering, Hosur

Abstract: Out of all types of cancer in the world, lung cancer is one of the biggest killers and is classified as either non-small cell (NSCLC) or small cell (SCLC). Making sure lung cancer is properly staged is necessary for making a good treatment plan and for understanding prognosis. Most lung cancer staging today involves seeing and reading CT images by hand which requires much time, may lead to errors and strongly depends on how skilled the radiologist is. Such limitations frequently mean that patients are diagnosed later and treated in suboptimal ways. The project here suggests an innovative framework to automatically detect and stage lung cancer. Radiomics is used in the system to measure features like intensity, shape and texture featured inside a tumor from CT scans. PCA is introduced to lighten the computational stress by selecting and maintaining the most important features. Input from the reduced features feeds a custom neural network, the TNMClassifier which is developed to categorize Tumor (T), Node (N) and Metastasis (M) stages using the TNM staging system. The design employs fully connected layers along with dropout layers, so it isn't likely to overfit and performs well. Classifying the TNM stages of lung cancer is accurate at 98% with the proposed system which helps address the weaknesses of existing solutions. This new method may help simplify work for clinicians, reduce chances of wrong diagnoses and provide better patient results
Keywords: Non-Small Cell Lung Cancer (NSCLC),Small Cell Lung Cancer (SCLC),Lung Cancer Staging, CTImaging, Radiomics, Feature Extraction, Principal Component Analysis (PCA) Dimensionality Reduction, TNM, Staging System, TNM Classifier, Neural Network, Tumor (T) Stage, Node (N) Stage, Metastasis (M)

I. INTRODUCTION

In the chest, the lungs (which help with breathing) are cone-shaped organs that sit separated by the heart and structures in the mediastinum. The organ has a medial surface and dividing lines known as anterior, posterior and inferior edges. Broadly, the coast of the lungs touches the rib cage, but their medial face surfaces are smaller. The lungs are attached to the bronchus, blood vessels, lymphatic vessels and nerves which all enter through a crack on the surface of the mediastinum called the helium, making up the lungs root. The right lung is both bigger and heavier than the left lung. Since we are right-side up, the heart is on the left side, making the left lung slight and causing an indentation called the cardiac impression to let the heart fit. Because of this indentation, the upper front part or anterior inferior, of the superior lobe forms a thin strip called the lingual.Finding Out if Someone Has Lung CancerWith imaging tests, an X-ray picture can expose an abnormal mass in the body. You should undergo a CT/PET-CT scan if the X-ray scan gives no significant result. You can clearly see small lesions on a CT scan.

The CT scan plays a major role in spotting symptoms of Cancer and estimating its risk. In addition to what is seen on a chest X-ray, a CT scan of the chest aids in seeing the details of the Cancer.

II. PROPOSED WORK

Through deep learning and radiomics, the proposed system can spot and assign stages to lung cancer cases, without much support from a manual radiologist assessment. Thanks to advanced machine learning, the system can help doctors make quick and accurate diagnoses for better planning of treatment.

Oncology Diagnostics Web App.Web-based technology forms the main tool used for diagnosing and grading lung cancer. The website is easy for anyone to use, so clinicians can easily add CT scan pictures, analyze them and print out reports. Because the workflow is simplified, the application makes cancer diagnosis faster and lets clinicians know instantly about the cancer's development. It also ensures safety of health records and simple integration with the hospital network.Accurate prediction and stage of Lung Cancer

Precise and efficient TNM staging is possible thanks to the use of radiomics and PCA by the system. Radiomics uses CT images to take data that human eyes may miss during a regular scan.

By applying PCA, we can discard some features and well-maintain those that matter most for classification. Using the deep learning model TNMClassifier, features from the data are examined and the stages of Tumor (T), Node (N) and Metastasis (M) are classified accurately. Automating the process of test actions improves consistency and reduces chances of mistakes.

Automated Recommendation

There is a recommendation module available to help clinicians choose the best treatments. The brain-inspired system recommends particular treatments according to the classified lung cancer stage such as surgery, chemotherapy, radiation therapy or using targeted drugs. With this module, information about the individual's previous condition, the latest medical guidelines and AI predictions are analyzed to tailor suggestions

III. METHODS

A. Oncology Diagnostics Online Application

The Web App for Oncology Diagnostics is built with the user in mind, so users can easily identify and stage lung cancer cases. Bootstrap was used to construct the Front-End, creating a responsive and simple interface for both users to use. The Back-End uses Python and Flask to handle CT images and make appropriate diagnostics. Secure storage of patient records and results is made possible thanks to MySQL's use for structured data storage. You can upload images, get an automated prediction for lung cancer, find out the cancer stage and get personal advice for your treatment. The result is improved accuracy, efficiency and more suitable choices for oncology treatment and care.

B. Customer Dashboard

1) Admin Panel

- Login: Provides safe document management for admins.
- Upload your training data by using the Import Dataset button.
- Start the model training and watch its progress.
- You can manage the accounts of both patients and doctors.
- Access doctors and hospital records: You can add new details or update already existing records.
- Suggestions & Recommendations: You can create and change disease-specific suggestions for your patients.

2) User/Patient/Doctor Dashboard

Registering and logging in are secured and user profile information can be managed through this section.
Users can choose to upload CT images for the detection of lung cancer using the tool.
View the report to see where the cancer is and what stage it is at.
You will be offered personal recommendations and suggestions for your care.

C. Dataset Description

Information for the LunCan model was derived from public lung CT scan images in medical imaging datasets. There are examples of each TNM stage for lung cancer and they have been annotated appropriately. Each photo is tagged as Tumor (T), Node (N) and Metastasis (M) which allows the use of a structured collection for training and testing models. Also included in the dataset are patient history and clinical reports which improve the accuracy of the model.

D. Import Dataset

Step four is to import the dataset into the system for more steps. Images are taken out of the database and their labels are also identified. Breaking the imported dataset into training, validation and testing helps make sure the model works evenly. By using data augmentation, you can deal with class disparities and improve your model's overall performance.

E. Preprocessing

You should always process the raw CT images before giving them to the deep learning algorithm. Among other things, you should:

- Resizing: Making all the images have the same dimensions so that the images in the dataset are uniform.
- Color Representation: Changing the appearance of a colored CT scan to gray levels.
- Removing noise from an image using filters to make it more visible.

- Binarization is a process that changes an image into binary code to pick out meaningful details.
- Marking-off areas: Isolating the lung regions using image processing so that analysis can be done with high accuracy.

F. Feature Extraction

Unique features of the tumors are found by examining their texture, outline and shades in lung CT pictures with radiomics-based methods. With these features, we can tell what stage cancer is in better. Radiomics measures features of a tumor that often go unnoticed by hand evaluation which increases the accuracy of TNM staging.

G. Classification

In classification, the TNMClassifier, a designed deep learning tool for lung staging, is put into practice. The CT scans are assigned by the classifier to Tumor (T), Node (N) or Metastasis (M) stages. Using this process, exact classification is achieved through deep neural networks working with the extracted features from images.

H. Make and shape your athletes with proper practice.

This model is developed by using various deep learning architectures available in TensorFlow and Keras. The process starts with preprocessing the data and then this process of training trains the weights with backpropagation. Such techniques are used to make the model work better and to prevent the model from overfitting the data. Running the model repeatedly (managing to use the full data repeatedly) helps it achieve higher accuracy at predicting TNM stages.

I. Deploy Model

When the model is declared working as intended, it becomes part of the Oncology Diagnostics Web App. During deployment, the trained model is combined with a Flask backend to detect and stage lung cancer in real-time. Users may upload CT scans through the web and the system responds by presenting TNM staging and proposing the best next treatment options.

IV. THE LUNG CANCER PREDICTOR

A. Upload Lung CT Image

To get started, the Lung Cancer Predictor module welcomes a lung CT scan image from the user which the user can send via the Oncology Diagnostics Web App. At the beginning, we test the image's size and quality to decide if it can be used. After validation, the image is processed which helps it be adjusted in size, become less noisy and focused on lung areas.

B. Cancer Localization

Once data is preprocessed, the system uses both radiomics features and deep learning methods to accurately identify and outline cancer in the lung CT image. The LunCan model, with TNMClassifier included, detects areas that might be tumors by evaluating how pixels, texture and other radiomic features look. Clinicians can see exactly where the detected cancer tissue is on the tumor by highlighting it with both heat maps and bounding boxes.

C. Staging

Once the tumor is clearly found, the system stages the cancer using the TNM system.

- Tumor (T) – The size of the original cancer and how farit has grown or spread.
- Node (N) – Found in nearby lymph nodes.
- M (short for Metastasis) – Cancer has reached other organs in the body.

The TNMClassifier uses extracted information to determine the stage of lung cancer and help doctors decide on the best management. In the web application, you can see your staging results and visual images, plus advice on potential treatments for your stage.

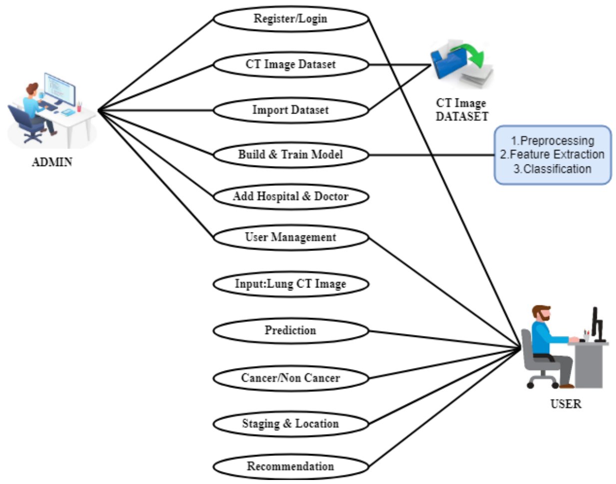
V. RECOMMENDATION

Based on what stage of lung cancer is found, the Recommendation Module suggests the right treatment options for that person. It presents both traditional and medical therapy options, including surgery, chemotherapy, radiation, immunotherapy or targeted therapy.

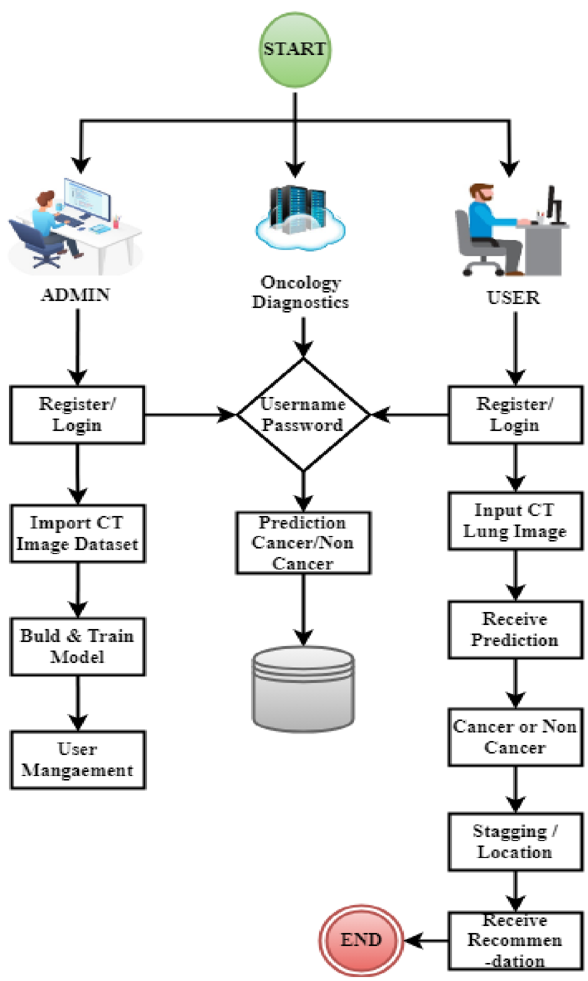
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

The system also suggests places to receive care and treatment based on where the patient is and how far the cancer has advanced. It gives patients advice about their day-to-day activities, as well as ongoing care suggestions for things like stopping smoking, diet changes and following regular screening rules. There are psychological support resources for patients dealing with the disease. Recommendations are all viewable on the Oncology Diagnostics Web App for quick use by both medical professionals and patients.

VI. UML DIAGRAM



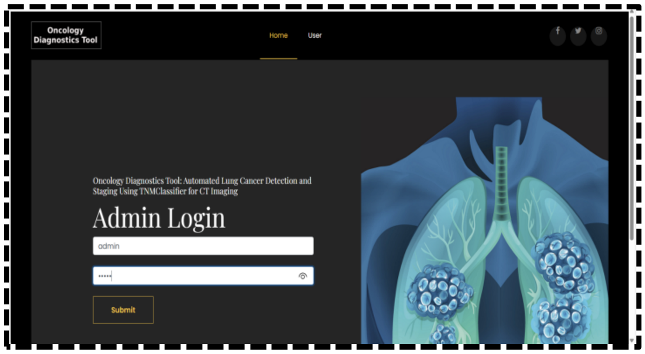
VII. ACTIVITY DIAGRAM

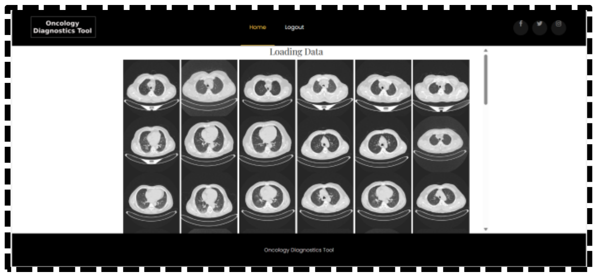


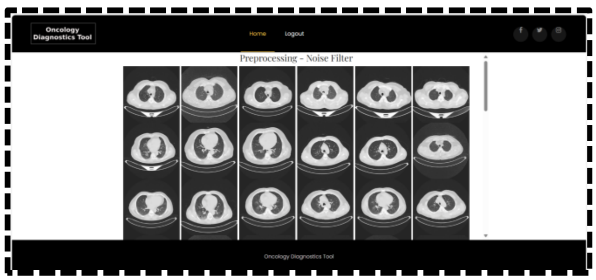
VIII. RESULT

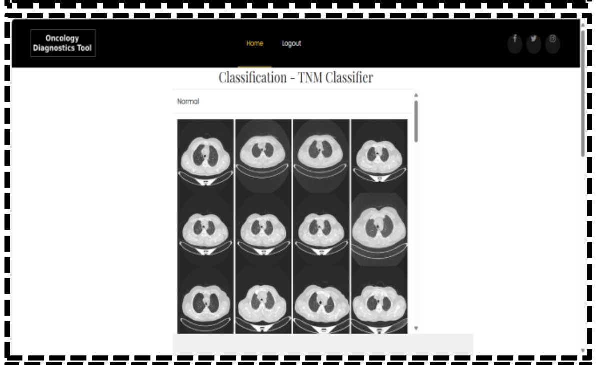
TC ID: TC001
User loads a valid CT scan image onto the websiteDone: Image is successfully uploaded and processed by the system
Solution: Smoothly uploaded and processed the images with no mistakesStatus: Pass
TC ID: TC002
Input: The user uploads a document that software cannot read.The result should be an error message showing on the screen.The error message appeared as expected.
Status: Pass
TC ID :TC003
Input: User enters correct login credentials.
End Result: User account accesses the dashboard successfully.User saw their dashboard after they were logged in successfully
Status: Pass
TC ID: TC004
The user provides wrong login information.
Result Expected: System prevents entry and displays an error messageWhat actually happened: A prompt appears, asking for a new login.
Status: Pass
TC ID : TC005
Before analysis, the CT scan image is preprocessed.The outcome is the image is resized, any noise reduced and the process is completed as intendedPreprocessing was done without errors.Status: Pass
TC ID : TC006
System pulls out the radiomics features from the CT scan.Once completed, your data should have successful feature extraction.Features were successfully and error-free extracted.
Status: Pass
TC ID : TC007
Method used: System uses PCA to reduce the number of features.The goal is to successfully use dimensionality reduction.The actual result was that PCA ran smoothly.
Status: Pass
TC ID: TC008
We use TNMClassifier to classify images in this example.
Outcome: The system can accurately classify the stage of lung cancer.
Actual Result: Classification results are shown accurately
Status: Pass
TC ID : TC009.
User asks for cancer staging data.The results of TNM classification appear as expected.Cancer staging was successfully presented to everyone in attendance.
Status: Pass
TC ID: TC010.
The system sends treatment recommendations after the user finishes the assessment.
Result: The system recommends appropriate treatment according to the detected stage.
The recommendations for treatment were shown as expected.
Status: Pass

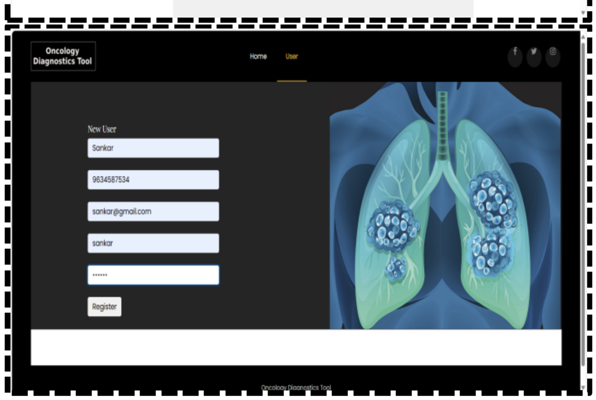
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

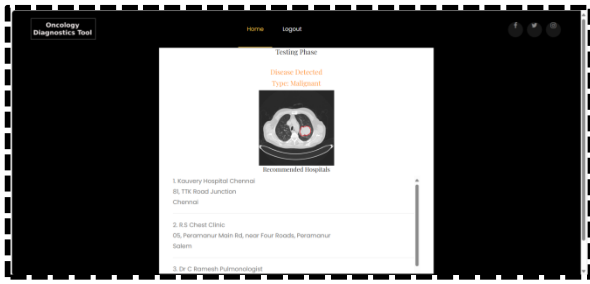












IX. CONCLUSION

Finally, the project has been built to help with automated lung cancer detection and staging in easy-to-use form. Combining Python, Flask, MySQL, TensorFlow, OpenCV, Pillow, Pandas, NumPy, Seaborn and Bootstrap, the system allows for fast processing and correct identification of lung cancer stages. With the End User Dashboard, admins, doctors and patients can each manage their tasks, transfer images, analyze studies instantly and receive specific guidance tailored to their needs. This system relies on the LunCan Model which handles the pipeline using three steps and a classifier: first, the data is preprocessed; second, features are extracted using Radiomics; third, PCA reduces these features; and finally, the TNMClassifier identify types of lung cancer in the data. Thanks to this model, doctors can reliably stage cancers because it handles every type of cancer earlier in the staging process. Using CT scan images, the Cancer Predictor module identifies lung cancer regions, marks tumor areas, recognizes stage of the cancer and helps diagnose and plan trea quickly. Furthermore, the Recommendation System helps medical staff design proper treatments with reference to the cancer stage that was found. Using machine learning and medical imaging helps connect AI and clinical oncology by making the system much more efficient and intelligent for diagnosis. The project is set to revolutionize existing lung cancer diagnosis, so more cases are found early, patients have better outcomes and healthcare processes are streamlined.

X. ACKNOWLEDGMENT

The authors declare that they have no reports of acknowledgments for this

REFERENCES

[1] S. Albahli, H. T. Rauf, M. Arif, M. T. Nafis, and A. Algosaibi, ''Identification of thoracic diseases by exploiting deep neural networks,'' Comput., Mater. Continua, vol. 66, no.3, pp. 3139–3149, 2021.
[2] A. Naik and D. R. Edla, ''Lung Cancer classification on computed tomography images using deep learning,'' Wireless Pers. Commun., vol. 116, pp. 655–690, Jan. 2021.
[3] S. Albahli, ''AI-driven deep CNN approach for multi-label pathology classification using chest X-rays,'' PeerJ Comput. Sci., vol. 7, p. e495, Apr. 2021.
[4] T. Meraj et al., ''Lung Cancers detection using semantic segmentation and classification with optimal features,'' Neural Comput. Appl., vol. 33, pp. 10737–10750, 2020.
[5] W. Zuo, F. Zhou, and Y. He, ''An embedded multi-branch 3D convolution neural network for false positive reduction in lung Cancer detection,'' J. Digit. Imag., vol. 33, no. 4, pp. 846–857, Aug. 2020.
[6] C.-J. Lin and Y.-C. Li, ''Lung Cancer classification using Taguchi-based convolutional neural networks for computer tomography images,'' Electronics, vol. 9, no. 7, p. 1066, Jun. 2020.
[7] P. Zhai, Y. Tao, H. Chen, T. Cai, and J. Li, ''Multi-task learning for lung Cancer classification on chest CT,'' IEEE Access, vol. 8, pp. 180317–180327, 2020.