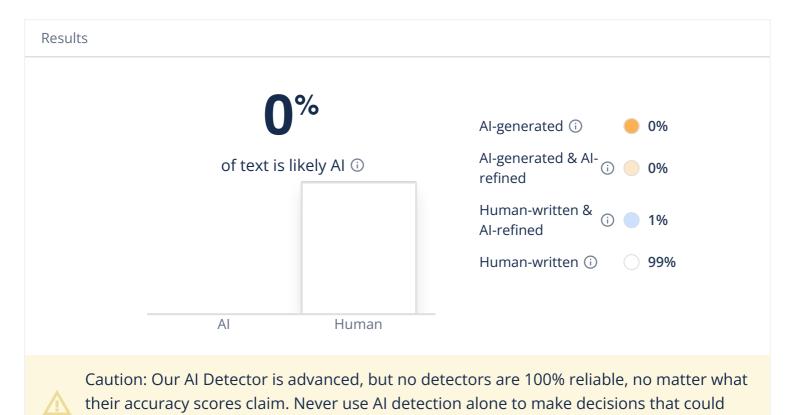
4,508 Words



DeepTumorNet: End-to-End Lung Cancer

Classification Using Deep Learning

Dharmika Tank School of Computer Science and Engineering,

impact a person's career or academic standing.

Dr. Vishwanath Karad MIT World Peace University, Pune

dharmikatank30@gmail.com

Jay MehtaSchool of Computer Science and Engineering,

Dr. Vishwanath Karad MIT World Peace University, Pune jaykalpeshmehta@gmail.com

Abstract—Using deep learning methods, the End-to-End Lung Cancer Classification project offers a thorough framework for identifying and categorizing lung cancer. The technology analyzes medical images using a Convolutional Neural Network (CNN) to accurately detect cancers. The project ensures a smooth and repeatable machine learning pipeline by integrating Flask for deployment, DVC for data and model versioning, and MLflow for experiment tracking. It is flexible for both research and practical application since it provides automated preprocessing, model training, and evaluation. The main objective is to develop an effective and scalable system that supports early

cancer diagnosis, cutting down on diagnostic delays and enhancing patient outcomes.

Keywords: CT scan images, CNNs, deep learning, lung cancer, transfer learning, and model deployment

Introduction

Lung cancer proves deadly for a major percentage of patients who succumb to cancer due to its status as the world's most dangerous disease. Research shows that lung cancer drives 18% of fatal cancer cases thus becoming one of the most damaging cancer types that resists treatment. Too late detection of lung cancer mostly occurs because early-stage cancer lacks noticeable symptoms while hiding under other respiratory conditions thus leading to increased mortality rates. Traditional diagnostic methods that base their analysis heavily on expert reading of biopsies and radiography scans and tissue analysis become an issue due to their time-intensive and unreliable processes which prompts the need for an automated early detection solution.

The implementation of advanced artificial intelligence produces medical imaging benefits by enabling automatic diagnostic systems that provide accurate results which can be applied to various cases. Medical image classification achieves significant success when utilizing Convolutional Neural Networks (CNNs) to detect tumors in addition to identifying diseases and medical anomalies in healthcare facilities. A deep learning system identifies complex medical image patterns in extensive medical datasets to produce superior analysis results than conventional approaches. Deep learning systems detect malignant tumors through their analysis of both benign tumors and different lung cancer subtypes to improve patient health results.

The study offers a whole machine learning system for lung cancer detection using computed tomography (CT) scan pictures. When deep learning architecture systems link to workflow systems managing experimental tracking and dataset management and model deployment, they attain high

accuracy outcomes and system robustness. Four existing CNN networks with ResNet, InceptionV3, NASNetMobile integrated are included in the application. ResNet's exact implementation and new residual learning system help to produce better results than picture classification challenges. InceptionV3's enhanced feature extraction system lets diverse spatial patterns be detected to produce improved medical picture representations. Because of its mix of portable design and optimal computing efficiency, NASNetMobile becomes a diagnostic tool for real-time medical scenarios on mobile devices. Key parameters including accuracy, F1-score, specificity, and recall are used to analyze the performance of different models in order to complete a thorough review. Recall measures the model's sensitivity to positive examples, whereas specificity evaluates its capacity to accurately detect negative situations. The F1-score offers a balance between precision and recall, while accuracy shows how accurate forecasts are overall. The best model for detecting lung cancer can be found by utilizing several designs and taking into account trade-offs between computational efficiency and diagnostic accuracy. The relative efficacy of each model is demonstrated using visual aids such confusion matrices, ROC (Receiver Operating Characteristic) curves, and performance graphs.

The work employs several fundamental machine learning techniques to attain scalability, reproducibility, and transparency simultaneously. MLflow experiment tracking tools track every stage of the model parameterization process together with training run records and performance indicators. Strong dataset version management offered by the data version control system DVC guarantees consistent data and promotes team-based research between departments. Flask creates a web-based tool for model dissemination so doctors may use it in real-world medical environments. From these linked technologies, our team uses a simple Al diagnostic method to create a system that can progress through changes for next medical imaging operations.

We have incorporated normal and adenocarcinoma CT scan images into our study to give classification tasks visual understanding. Model interpretation and verification of deep learning effectiveness receives support from these images because they show tissue differences between cancerous lesions and healthy lung areas. Medical images from real-life clinical settings improve the study's practical applicability while facilitating the creation of Al-enhanced diagnostic systems meant for medical facilities.

The research supports future development of medical diagnosis through AI by establishing an advanced deep learning system approach that follows systematic guidelines and scaling principles.

The study presents a highly precise automatic lung cancer screening technology that will help medical experts and radiologists in their practice. Future advancements of this AI technology will lead to improved early detection rates along with fewer diagnostic errors resulting in better patient survival outcomes throughout the world.

Literature Survey

The healthcare sector encounters significant challenges in identifying lung cancer due to its persistently elevated death rate. The late detection of cancer results in adverse patient health outcomes, as it is one of the leading causes of global mortality from cancer. Timely detection is crucial for survival as it markedly enhances patient outcome success. The medical diagnosis of lung cancer utilizing Convolutional Neural Networks (CNNs) in deep learning (DL) methodologies has demonstrated significant efficacy in analyzing medical images. Medical practitioners today perceive radiological imaging differently because new approaches replace traditional diagnostic methods with enhanced precision and efficacy. Keerthi et al. (2025) [1] illustrated how deep neural networks discern complex patterns in medical images that are not readily perceptible to human observers in their examination of lung cancer categorization effectiveness. Research indicates that VGGNet and ResNet architectures from the CNN family exhibit significant potential for medical image processing and tumor detection, facilitating earlier cancer identification and minimizing diagnostic mistakes. These models serve as a formidable instrument in the fight against lung cancer, achieving exceptional accuracy when trained on extensive and diverse datasets. Hybrid models that integrate CNNs with transfer learning methodologies have emerged as a pivotal area of interest in recent years due to their enhancement of lung cancer diagnostic accuracy. The utilization of pre-trained models from public databases enables researchers to improve performance on domain-specific datasets via transfer learning. This technique significantly enhances medical picture analysis due to its requirement for minimal available labeled data. Marappan et al. (2024) [3] conducted research assessing the utilization of ResNet in conjunction with transfer learning to improve predictive accuracy, especially in scenarios of limited data availability. The approach addresses the challenge of limited labeled datasets by utilizing pre-trained weights from conventional picture datasets, hence enhancing model operational efficiency. The models exhibited exceptional precision in identifying malignant lesions by calibrating them with specific lung cancer data sets. The hybrid technique exhibits significant potential in clinical application by mitigating data scarcity and imbalance while improving generalization.

The field of research has explored lung cancer detection model improvement with machine learning methods combined with deep learning approaches. The authors in Khouadja and Naceur (2023) [4] explored machine learning systems' functions to improve full detection solutions using deep learning methodologies. Research into deep neural networks has been conducted jointly with conventional machine learning algorithms by their investigatory team. The deep learning models handle complex classifications yet machine learning technologies work during the first detection phase by extracting features from images and performing preprocessing operations. By applying hybrid methods during real-time clinical operations the medical sector obtains speedier and more precise diagnoses with enhanced details of computer system resources. A detailed discussion about CNN-based system model optimization emerges from Sumithra et al. (2023) [6]. According to the research team deep learning models need to achieve both high accuracy levels and efficient operation for successful healthcare implementation since they must quickly generate results that support swift clinical choices. Effective and reliable medical diagnoses play an essential role in medicine since late medical diagnosis negatively impacts patient outcomes.

In order to improve the quality of the images that are used to train deep learning models, preprocessing medical images is another key component in the diagnosis of lung cancer. Preprocessing techniques such as image normalization, scaling, and segmentation that help with more precise cancer diagnosis are able to improve the clarity and accuracy of the characteristics that are present in medical images. Nawreen et al. (2021) [8] focused on use of image processing techniques, specifically for CT scan analysis, in order to improve the quality of the data before feeding it into deep learning models. This was done in order to increase the data's overall quality. In order to guarantee that the CNN models are able to extract relevant information in a more effective manner, these preprocessing processes improve the quality of the images. Rehman et al. (2021) [9] investigated machine learning models for classification tasks that were specifically created for the processing of CT images of the chest. This research contributed to the growing body of work that has been done on image preprocessing. In order to improve the effectiveness of the model, their research made it abundantly evident how important it is to have both a robust classification model and precise data pretreatment approaches. The usefulness of various image processing approaches, such as segmentation algorithms, that are utilized to differentiate lung tumors and other relevant characteristics was also investigated by Karthikeyan et al. (2021) [10]. In order to ensure that deep learning models are able to focus on the critical areas of the image that are suggestive of lung cancer, these methods are essential for ensuring that the accuracy of predictions can be increased.

Multiple obstacles stand in the way of the successful development of lung cancer detection systems for clinical applications, which inhibits their implementation. Through the utilization of extensive and extensive datasets that are also of high quality, the performance of deep learning models is able to attain its peak potential. Yadav and Badre (2020) [11] emphasize that there is a lack of complete datasets, which in turn prevents models from achieving general application. Despite the fact that deep learning models are extremely powerful, the decision-making processes they use continue to be difficult for people to comprehend. This is for the basic reason that their approaches of interpretation have drawn much criticism. Medical practitioners have great implementation difficulties since they must have a thorough awareness of the factors that models project. Models must be absolutely understandable to users if they are to inspire confidence in them. This is so because human decisions affecting people's life must be grounded on the predictable outcomes. Research in this topic should advance in two different directions: first, by raising the degree of accuracy reachable, and second, by creating transparent and honest explanation systems.

Methodology

The proposed approach of lung cancer classification is based on a disciplined pipeline. Apart from model training and evaluation and implementation stages, the whole system offers thorough coverage covering data collecting and preprocessing as well. By using this methodical technique, we can make sure the deep learning diagnostic system shows outstanding performance together with scalability and maintainability factors and repeatability. The design system conforms to all accepted industry guidelines concerning artificial intelligence methods and medical imaging technologies. The pipeline consists of these sequential phases, which form its basis:

1. Information Collecting

The process starts with the data collecting of premium medical photos. These datasets come from three respectable scientific organizations: The Cancer Imaging Archive (TCIA), the National Institutes of Health (NIH), and Kaggle. Regular lung images and scans showing cancer-infected tissue abound in medical images taken on CT scans. These datasets provide most of the data deep learning models need for both model assessment projects and training. Raw photos are first quality checked and filtered to guarantee both appropriate data quality and accurate data adherence before training.

2. Versioning and Preprocessing of Data

A dependable data preprocessing system is developed to achieve performance consistency and training round reproducibility. The procedure consists of:

The training process requires data cleaning to eliminate scans that are of poor quality since these images might mislead the model.

A standard normalization process transforms pixel values to the same level of brightness and creates uniform image contrast between all photos.

Data Version Control (DVC) functions as a tool to maintain technical reproducibility while handling modifications in dataset content.

3. Management of Data and Repositories

Along with codebases and model checkpoints, the platform DagsHub acts as a management system for datasets guarantees both experiment collaboration and maintenance of openness. By means of version handling elements that enable simple tracking and repetition operations for machine learning applications, the system offers seamless experimentation control.

4. Model Training and Selection

A classification model implements Convolutional Neural Network (CNN) architecture to execute its operations. Multiple deep learning models are selected as baseline candidates for this project including ResNet and InceptionV3 as well as NASNetMobile. The CT scans of lungs are used to refine the running generation of models which began with weight transfer learning methods. The third step requires the selection of suitable architecture design and involves parameter setup followed by optimization through Adam or RMSprop methods.

5. Metrics computation and performance evaluation

Following model training, a thorough assessment is conducted utilizing a number of metrics:
Accuracy: Indicates whether predictions were accurate.
Assess the accuracy and recall of the classification of cases that are malignant and those that are not.
F1-Score: Strikes a balance between recall and precision.
Sensitivity and Specificity: Calculate the rates of true positives and true negatives, respectively.
Performance is tracked and compared across models and settings using MLflow.
Fig. 1. System Design of Lung cancer
Classification Using Deep Learning
6. Model Retraining and Hyperparameter Adjustment
Hyperparameter tweaking is carried out using Grid Search, Random Search, and Bayesian Optimization if the first results are insufficient.

The model is retrained for better performance once the ideal parameters have been determined.
7. Evaluation and Prediction of the Final Model
A designated hold-out test set is used to assess the improved model. In order to verify clinical readiness, predictions are compared to ground truth labels, ensuring an objective evaluation.
8. Model Implementation
Flask is used to deploy the finished model as a REST API. Among the crucial actions are:
-Model Serialization
-Both batch and real-time inference
This guarantees dependable deployment in a variety of settings.
Constant observation and model updates
The system is constantly checked for performance drift after deployment. To maintain accuracy and flexibility, the model is updated and retrained on a regular basis using fresh data and user feedback.
Analysis and Discusion

Fig. 2. ResNet Performance Parameters

Chest CT scan images from the dataset were used for training the ResNet model to test its classification performance using multiple evaluation criteria. Recording 0.0734 loss throughout the final results phase, the study shows the model effectively separates healthy from malignant lung tissues with 97.06% accuracy. With 97.21% good precision, the model shows minimum false positives and recalls real cancer cases with 97.06%. The model keeps great dependability since its computed F1-score is 97.05% and it achieves exact accuracy while efficiently extracting diagnostic data. The model shown perfect specificity since it precisely identified every instance of negative situations, thereby helping to reduce unnecessary medical treatments. ResNet demonstrates itself as a strong deep learning architecture with remarkable performance measures supporting its practical clinical application for lung cancer diagnosis. ResNet shows outstanding performance traits that qualify for discussion in operational medical diagnosis systems supported by artificial intelligence technology.

Fig. 3. InceptionV3 Performance Parameters

Training the InceptionV3 model with the given dataset then testing it across several assessment criteria tested on the lung cancer classification performance. The acquired final results showed a loss of 0.0875 in addition to 97.06% accuracy, thereby confirming the great predictive potential of the model. Because it correctly identifies malignant patients, the model shows 97.06% recall and achieves 97.21% precision, therefore helping to eliminate erroneous positive predictions. The strong classification of the model comes from its F1-score value of 97.05%, which in a balanced manner corresponds to recall and precision measurements. Such high degrees of sensitivity (100%)

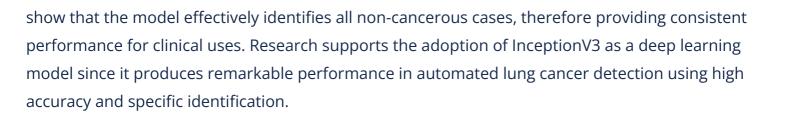
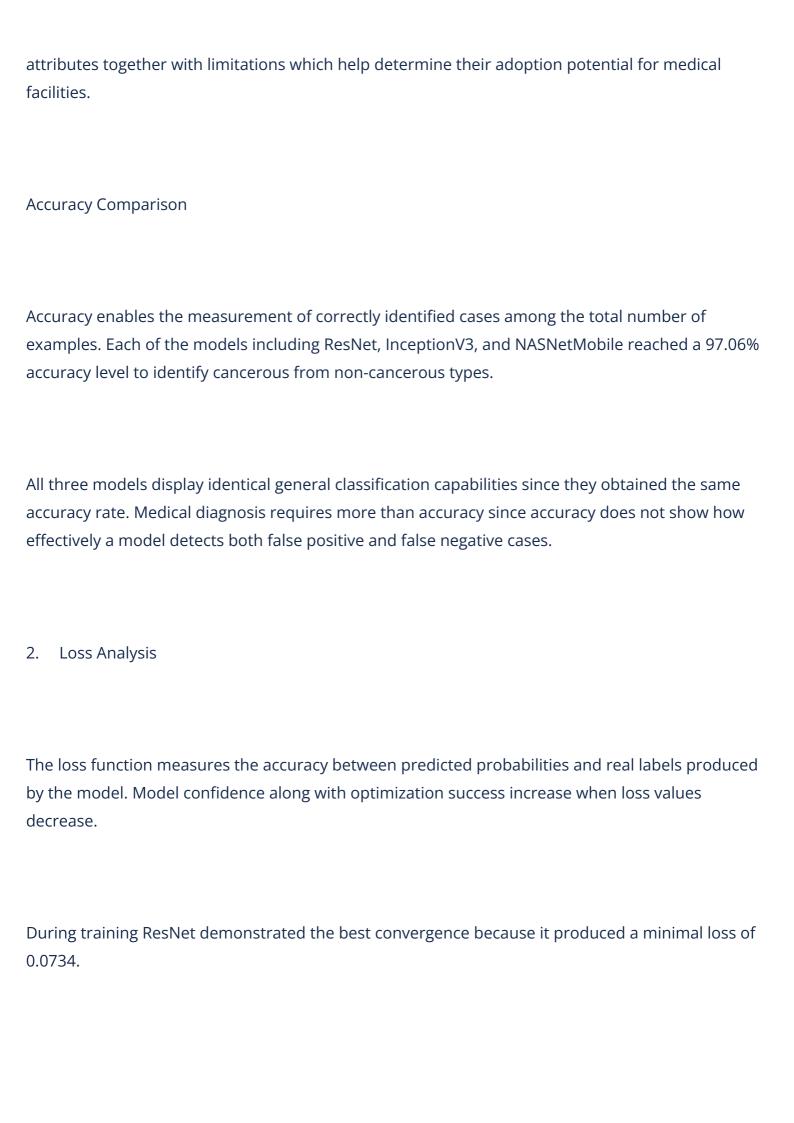


Fig. 4. ResNet Performance Parameters

During its testing phase, the NASNetMobile produced good results in lung cancer categorization under several performance standards. Since laboratory results show 0.0987 loss rate with 97.06% accuracy, they confirm the capacity of the model to identify malignant from benign diseases. The model maintained 97.06% recall, so guaranteeing the detection of most real cancer cases, and attained 97.21% precision, so reducing false discoveries. An F1-score value at 97.05% showing balanced precision-recall ratios provides the model confirmation. The model is suitable for medical environments that have to avoid possible human damage from erroneous results since it can detect all non-cancerous situations with 100% accuracy.

Fig. 5. Comparative Analysis of Deep Learning Models for Lung Cancer Classification

Important performance values such as accuracy, precision, recall, F1-score and loss were utilized to evaluate the ResNet and InceptionV3 and NASNetMobile models for lung cancer classification. Each metric contributes to model evaluations through its ability to decrease misidentification errors and accurately measure their capacity to detect malignant cases. This analysis reveals practical



InceptionV3 required better optimization steps or a different learning rate to achieve its loss level of 0.0875 since its loss outpaced ResNet's lower figure of 0.0734.
NASNetMobile struggled with difficulty in processing complex patterns found in the dataset since it demonstrated a loss of 0.0987.
Generalization performance together with prediction certainty improves when loss rates decrease. ResNet achieved the lowest loss rate so it offers more stability in predictions than InceptionV3 and NASNetMobile.
3. Precision, Recall, and F1-score Comparison
Precision and recall deliver higher importance in medical applications because they assess how effectively the model detects both true and false samples.
The precision level stands at 97.21% which indicates what percentage of accurate cancer predictions exist in the system. The high precision of the model decreases wrong positive predictions thus maintaining essential patient safety by avoiding unwanted medical procedures for people without lung cancer. The precision values from each model were equivalent which demonstrates their identical capability in managing false positive outcomes.
A high model recall performance reaches 97.06% because it reflects the ability to detect genuine lung cancer cases. In medical imaging the detection of cancerous cases (recall) assumes high importance because the mistreatment of cancerous abnormalities has serious patient

consequences. Each model demonstrated the same recall performance which makes them suitable for identifying lung cancer patients in CT scans.
The F1-score reaches 97.05% through its method of calculating the harmonic mean between precision and recall to find an optimal trade-off between the two measurement points. When an F1-score remains high it indicates the model successfully detects lung cancer with minimum impact on precision and recall metrics. Again, all three models performed equally well, proving their robustness in classification tasks.
Computational Efficiency and Suitability for Deployment
Although all three models performed equally well in terms of accuracy and key performance metrics, their computational efficiency and suitability for deployment differ significantly.
ResNet:
Strengths:
Achieves the lowest loss, indicating superior learning and convergence.
Well-suited for deep feature extraction, making it highly effective in medical imaging.
Performs well in large-scale datasets due to its deep residual connections.

Limitations:
Computationally intensive, requiring more processing power and memory.
Longer training time compared to NASNetMobile.
Best Use Case:
High-performance hospital diagnostic systems
Research environments requiring extensive feature extraction
Cloud-based AI services with sufficient computational resources.
InceptionV3
Strengths:
Although factorized convolutions optimize performance between prediction precision and calculation speed.
The model achieves results similar to ResNet but needs fewer computing resources thus making it an optimal choice for systems that operate on constrained power supplies.

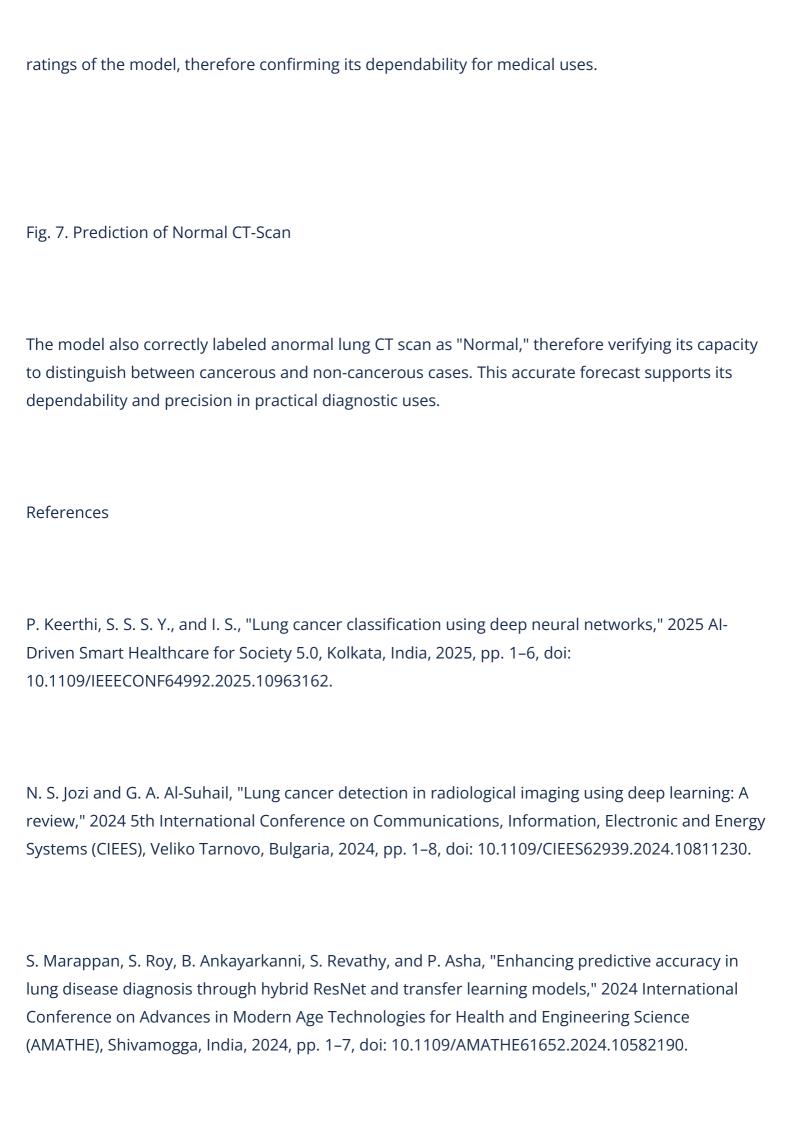
The system allows flexible modifications that speed up inference operations making it desirable for medical use with moderate hardware platforms.
Limitations:
NAS presents margins of loss closer to those of ResNet which indicates possible instability in training and optimization processes.
Users must thoroughly adjust their hyperparameters to obtain the highest possible performance output.
Best Fit For:
The healthcare facilities which use average computing systems operate within this category.
Medical AI startups working on performance optimization and stability will find NASNetMobile beneficial.
Al solutions serving healthcare in the cloud environment need scalability and reasonable resource allocation capability.

NASNetMobile
Strengths:
The model has been specifically developed for embedded devices and mobile platforms to operate with maximum efficiency in resource utilization.
The model achieves performance near ResNet along with InceptionV3 without using similar processing power.
The system suits mobile applications perfectly because of its simplified deployment process thus serving well for portable real-time testing needs.
Limitations:
According to loss values NASNetMobile shows the most uncertain predictions of the three networks.
The model has difficulty when it comes to generalizing across larger or complex datasets compared to ResNet models.

Best Fit For:
Real-time diagnostics on mobile devices.
Telemedicine systems need device-based processing capabilities.
Facilities that operate with restrictions on their availability of powerful computing systems.
Key Insights and Discussion
The comparative analysis reveals that all three models—ResNet, InceptionV3, and NASNetMobile—are highly effective in lung cancer classification, with identical accuracy and strong precision-recall balances. However, the best model depends on the deployment scenario:
ResNet is best for high-accuracy applications where computational power is not a constraint.
InceptionV3 provides a balance between accuracy and efficiency, making it suitable for scalable medical AI applications.

NASNetMobile is ideal for lightweight applications where speed and low computational cost are priorities.
The slight variations in loss values suggest differences in optimization efficiency, with ResNet showing the best convergence, followed by InceptionV3 and then NASNetMobile. This makes ResNet the most stable model in terms of learning capability.
From a deployment perspective, NASNetMobile is the most efficient for real-time and mobile Albased healthcare applications, while ResNet is more appropriate for hospital-grade Al systems that require extensive computation and deep feature extraction. For medium-scale installations in hospitals or cloud-based artificial intelligence systems, InceptionV3 offers a balanced solution.
Final Model Recommendation for Different Scenarios
Scenario
Recommended Model
High-accuracy hospital Al systems
ResNet
Al-assisted radiology (cloud-based)
ResNet/InceptionV3
Mid-range hospital AI deployment

InceptionV3
Research and large-scale datasets
ResNet
Telemedicine and remote diagnosis
NASNetMobile
Mobile-based AI diagnosis apps
NASNetMobile
Real-time embedded AI in devices
NASNetMobile
This grouping guarantees that every model is used depending on its advantages, therefore offering the ideal mix of deployment practicality, performance, and efficiency.
Results
Fig. 6. Prediction of Cancerous CT-Scan
The model proved useful in real-world diagnosis situations by effectively spotting adenocarcinoma cancer from the given lung CT imaging. This accurate forecast supports the high accuracy and recall



O. Khouadja and M. S. Naceur, "Lung cancer detection with machine learning and deep learning: A narrative review," 2023 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET), Hammamet, Tunisia, 2023, pp. 1–8, doi: 10.1109/IC_ASET58101.2023.10150913.
[5] I.V. and D. Menaka, "Real-time detection of lung cancer using CNN," 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), Vellore, India, 2023, pp. 1–6, doi: 10.1109/ViTECoN58111.2023.10157316.

[6] B. Sumithra, G. Vallathan, M. Raman Kumar, and K. Govindharaju, "Deep learning for accurate

chest disease classification: A CNN-based approach for lung cancer subtypes and normal cells,"

2023 International Conference on System, Computation, Automation and Networking (ICSCAN),

[7] M. Mamun, A. Farjana, M. Al Mamun, and M. S. Ahammed, "Lung cancer prediction model using

ensemble learning techniques and a systematic review analysis," 2022 IEEE World AI IoT Congress

[8] N. Nawreen, U. Hany, and T. Islam, "Lung cancer detection and classification using CT scan image

processing," 2021 International Conference on Automation, Control and Mechatronics for Industry

4.0 (ACMI), Rajshahi, Bangladesh, 2021, pp. 1–6, doi: 10.1109/ACMI53878.2021.9528297.

(AlloT), Seattle, WA, USA, 2022, pp. 187–193, doi: 10.1109/AlloT54504.2022.9817326.

Puducherry, India, 2023, pp. 1-7, doi: 10.1109/ICSCAN58655.2023.10394855.

[9] A. Rehman, M. Kashif, I. Abunadi, and N. Ayesha, "Lung cancer detection and classification from chest CT scans using machine learning techniques," 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), Riyadh, Saudi Arabia, 2021, pp. 101–104, doi: 10.1109/CAIDA51941.2021.9425269.

[10] R. D. Karthikeyan, R. G., V. V., G. B. C., and K. M., "A review of lung cancer detection using image processing," 2021 Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, 2021, pp. 1–4, doi: 10.1109/STCR51658.2021.9588835.

[11] A. Yadav and R. Badre, "Lung carcinoma detection techniques: A survey," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020

[10] R. D. Karthikeyan, R. G., V. V., G. B. C., and K. M., "A review of lung cancer detection using image processing," 2021 Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, 2021, pp. 1–4, doi: 10.1109/STCR51658.2021.9588835.

[11] A. Yadav and R. Badre, "Lung carcinoma detection techniques: A survey," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020

4, doi: 10.1109/STCR51658.2021.9588835.