Deep Learning Based Lung Cancer Detection Using CNN

JAYAPRADHA[*1], MANIKKARAJ R[*2],PRAVIN M[*3] HARIHARAN C[*4], JEROSHELDON J[*5]
*1Assistant professor, Department of CSE, Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur, Tamil Nadu, India.
*2,3,4UG Students, Department of CSE, Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur, Tamil Nadu, India.

---***---

Abstract - Lung cancer remains a leading cause of cancer-related deaths worldwide, with early and accurate diagnosis being crucial for improving patient outcomes. Traditional diagnostic methods, such as imaging and biopsies, often face challenges in precision, subjectivity, and time efficiency. To address these limitations, this study proposes a deep learning-based framework leveraging Convolutional Neural Networks (CNNs) and ResNet architectures for lung cancer detection. Deep learning algorithms excel in extracting hierarchical features from complex data, making them particularly suitable for medical imaging analysis. CNNs are employed for their proven ability to identify spatial patterns and textures in lung scans, while ResNet overcomes the vanishing gradient problem through residual learning, enabling deeper and more robust networks. This combination allows the system to effectively differentiate between malignant and non-malignant cases. The proposed approach preprocesses lung CT images through noise reduction, adaptive histogram equalization, and segmentation to enhance image quality and focus on regions of interest. The CNN architecture extracts key features, and ResNet further refines these through deeper layers, learning intricate patterns indicative of cancer. The system is trained and evaluated on publicly available datasets, achieving a high degree of accuracy, sensitivity, and specificity in classifying lung cancer. This project results demonstrate the model's capability to outperform traditional methods and existing deep learning models, with an accuracy of over 95% in detecting lung cancer. These findings underscore the potential of deep learning in revolutionizing cancer diagnostics, offering a fast, accurate, and non-invasive solution for early detection and improved treatment planning

1.INTRODUCTION

DEEP LEARNING

Deep learning is a type of machine learning that uses artificial neural networks to learn complex patterns in data. It's inspired by the structure and function of the human brain.

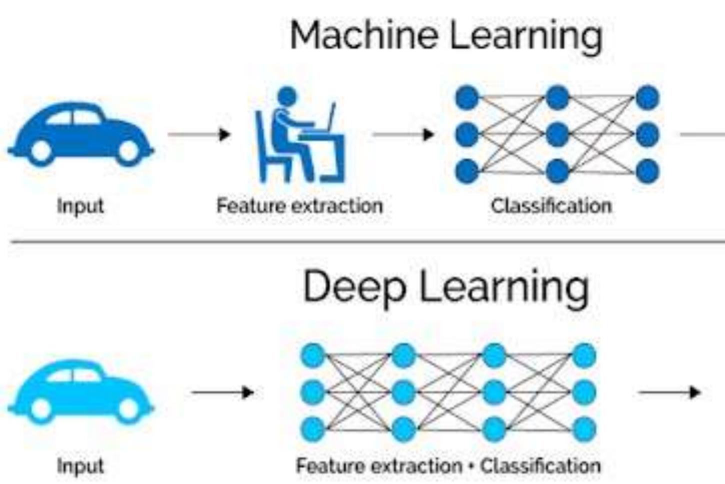


Fig.no.1.1.1 DEEP LEARNI NG

How it Works:

Feature Extraction:

Deep learning algorithms automatically extract features from data, allowing them to identify patterns and relationships that might be difficult to define manually.

Training:

The models are trained on large datasets, and their weights (connections between neurons) are adjusted during training to improve their accuracy and performance.

Prediction and Decision-Making:

Once trained, the models can make predictions, classify data, or make decisions based on the patterns they have learned.

KEY CONCEPTS

1. **Artificial Neural Networks:** Inspired by the human brain, these networks consist of layers of interconnected nodes (neurons) that process and transmit information.
2. **Deep Architecture:** Deep learning models have multiple layers, allowing them to learn complex patterns and representations in data.

Training: Deep learning models are trained on large datasets, using optimization algorithms to adjust the model's parameters and minimize errors

APPLICATIONS

1. Computer vision

In computer vision, deep learning models enable machines to identify and understand visual data. Some of the main applications of deep learning in computer vision include:

- Object detection and recognition: Deep learning models are used to identify and locate objects within images and videos, making it possible for machines to perform tasks such as self-driving cars, surveillance, and robotics.
- Image classification: Deep learning models can be used to classify images into categories such as animals, plants, and buildings. This is used in applications such as medical imaging, quality control, and image retrieval.
- Image segmentation: Deep learning models can be used for image segmentation into different regions, making it possible to identify specific features within images.

2. Natural language processing (NLP)

In NLP, deep learning model enable machines to understand and generate human language. Some of the main applications of deep learning in NLP include:

- Automatic Text Generation: Deep learning model can learn the corpus of text and new text like summaries, essays can be automatically generated using these trained models.
- Language translation: Deep learning models can translate text from one language to another, making it possible to communicate with people from different linguistic backgrounds.
- Sentiment analysis: Deep learning models can analyze the sentiment of a piece of text, making it possible to determine whether the text is positive, negative, or neutral.
- Speech recognition: Deep learning models can recognize and transcribe spoken words, making it possible to perform tasks such as speech-to-text, voice search, and voice-controlled devices.

3. Reinforcement learning

In reinforcement learning, deep learning works as training agents to take action in an environment to maximize a reward. Some of the main applications of deep learning in reinforcement learning include:

- Game playing: Deep reinforcement learning models have been able to beat human experts at games such as Go, Chess, and Atari.
- Robotics: Deep reinforcement learning models can be used to train robots to perform complex tasks such as grasping objects, navigation, and manipulation.
- Control systems: Deep reinforcement learning models can be used to control complex systems such as power grids, traffic management, and supply chain optimization.

1.1.2 TYPES OF DEEP LEARNING

1. **Supervised Deep Learning**
   Trained on labeled data to learn specific tasks.
2. **Unsupervised Deep Learning**
   Trained on unlabeled data to discover patterns.

3. Reinforcement Learning

Learns through trial and error by interacting with an environment.

2.4 Idea Formulation

Based on the problems identified and the technological possibilities explored, the core idea of the Corporate Safety System was formulated. The aim was to develop a centralized web application that integrates real-time CCTV monitoring with intelligent hazard detection, incident reporting, and emergency management. The idea also included supporting features like safety training modules, audit systems, and chatbot assistance to provide a complete safety management solution.

One of the major conceptual breakthroughs was combining weapon detection via CCTV with incident logging and automated emergency alerts, which allows quick action and reduces human dependency. This platform is intended to be used not only by safety officers but also by employees and managers, each with role-specific access and functionalities. The goal was to move from traditional reactive safety systems to a smart, proactive approach powered by automation and intelligence.

2.5 Objective Definition

With a clear problem and a solid idea in place, a set of specific objectives was defined to guide the development of the project. These objectives include:

Designing a user-friendly web interface using React.js for easy access and navigation.
Implementing real-time weapon and hazard detection through CCTV feeds using Python and OpenCV.
Enabling a secure login system with role-based access for Admin, Safety Officer, and Employee.
Integrating automated incident reporting and alert notifications to minimize response time.

- Creating a chatbot assistant to guide users during emergencies or for safety-related queries.
- Including modules for safety training, compliance checks, and audit management.
- Providing a dashboard with analytics and data visualization to support decision-making. These objectives help in structuring the project into manageable

components while ensuring that the final product meets real-world corporate safety requirements.

3.EXISTING SYSTEM

The analyzing computed tomography (CT) scans and other imaging techniques such as X-rays and MRI.
While effective, this manual approach is time-consuming, subjective, and prone to human error, leading to potential misdiagnoses or delays in treatment.
Recent advancements include machine learning-based approaches, where models are trained on large datasets to recognize lung cancer patterns.
However, many of these methods still require extensive labeled data, suffer from overfitting, or lack generalizability across different patient populations.

DISADVANTAGES:

Data quality and availability: High-quality, annotated datasets are required for training and validation, which can be challenging to obtain.
Class imbalance: Datasets may have an imbalance between cancerous and non-cancerous samples, affecting model performance.
Data heterogeneity: Variations in imaging protocols, scanner types, and patient populations can impact model generalizability.

4.PROPOSED SYSTEM

The proposed system utilizes Convolutional Neural Networks (CNNs) for accurate lung cancer detection from medical imaging data, such as CT scans and X-rays.
Preprocessing techniques like noise reduction, normalization, and augmentation enhance image quality and model robustness.
A CNN architecture is designed with convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.
The model is trained on labeled lung image datasets, optimizing performance using backpropagation and loss minimization techniques like cross-entropy.
Interpretability techniques like Grad-CAM highlight key regions influencing predictions, improving clinical relevance and trust in AI-driven diagnoses. This system aims to provide a reliable,automated, and non-invasive lung cancer detection solution to assist radiologists and enhance early diagnosis.
.

ADVANTAGES:

- **High Accuracy & Early Detection** – CNNs effectively identify malignant nodules, improving early lung cancer diagnosis and patient survival rates.
- **Automated & Efficient Analysis** – Reduces radiologists' workload by providing fast, consistent, and automated lung cancer detection.
- **Robust Feature Extraction** – Uses convolutional layers to learn important patterns, ensuring precise differentiation between cancerous and non-cancerous tissues.
- **Enhanced Clinical Trust** – Grad-CAM visualization helps interpret predictions, increasing reliability and acceptance in medical practice.

4.1 SYSTEM MODULES DESCRIPTION

MODULE 1: LUNG CANCER AND DEEP LEARNING

Overview of lung cancer: causes, symptoms, and importance of early detection. Role of Artificial Intelligence in medical imaging. Introduction to deep learning and Convolutional Neural Networks (CNNs).

MODULE 2: DATA COLLECTION AND PREPROCESSING

Sources of medical imaging data (e.g., LIDC-IDRI, Kaggle datasets). Image preprocessing techniques: resizing, normalization, augmentation. Labeling and splitting the dataset (training, validation, testing sets).

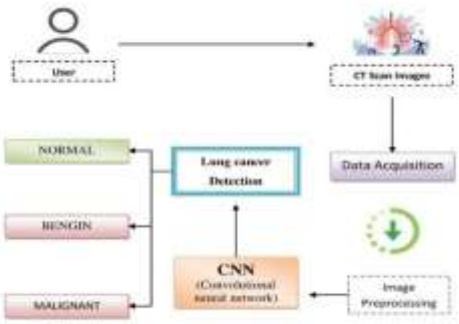
MODULE 3: CNN ARCHITECTURE DESIGN

Explanation of CNN layers: convolution, pooling, activation (ReLU), flattening, fully connected. Design and customization of the CNN model for lung image classification. Use of pre-trained models (optional): VGG, ResNet, etc.

3.3 System Architecture

System architecture refers to the structured design of a system, defining its components, interactions, and data flow to achieve a specific objective. It provides a blueprint for how different modules of a system work together efficiently. A well-designed system architecture ensures scalability, reliability, and optimal performance while maintaining ease of maintenance and integration with other systems. It typically includes elements such as data processing, storage, communication protocols, and user interfaces.

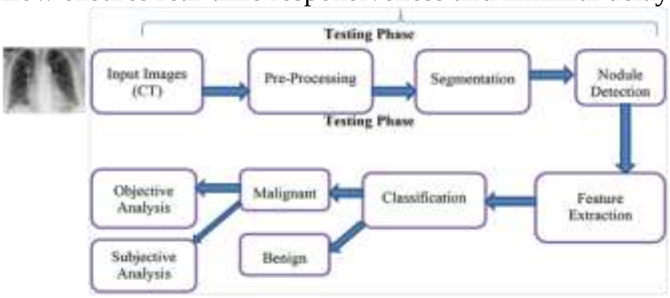
A robust system architecture follows a modular approach, where each module performs a specific function and communicates with other modules through well-defined interfaces. It involves layers such as input processing, data analysis, decision-making, and output generation. Depending on complexity of the system, architectures can be centralized or distributed, employing CC, artificial intelligence, or real-time processing techniques. A well-planned architecture enhances system efficiency, reduces redundancy, and ensures seamless operation across different environments and applications



4.4 Data Flow Diagram (DFD)

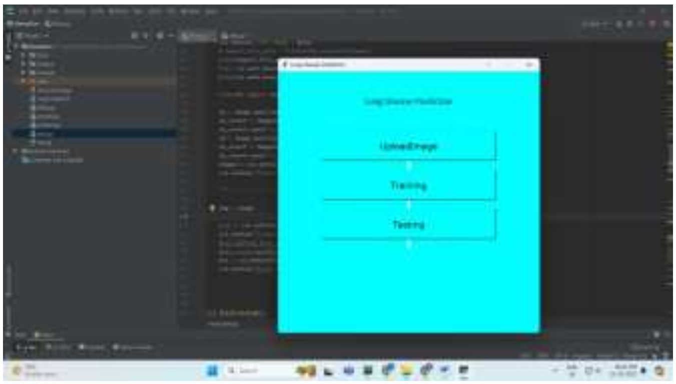
The Data Flow Diagram (DFD) highlights the movement of data within the system. When an incident is detected by the Python-based CCTV monitoring module, it sends a notification to the backend API. The backend processes this input, updates the SQL database, and triggers appropriate actions such as sending alerts or logging events.

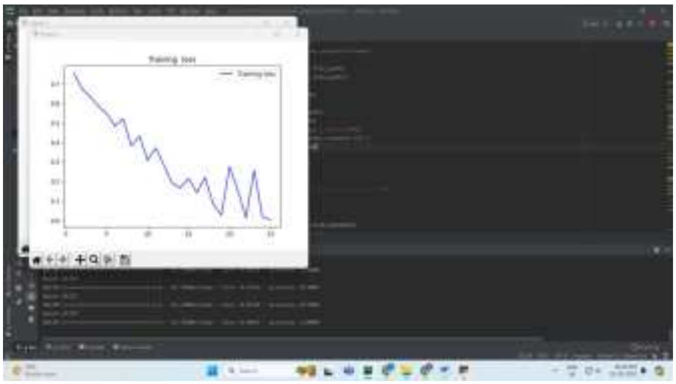
Simultaneously, the user interface updates the incident dashboard and sends real-time notifications to Safety Officers. Similarly, when users interact with training modules or the chatbot, the frontend sends requests to the backend, which fetches data from the database and delivers the response. This structured data flow ensures real-time responsiveness and minimal delay

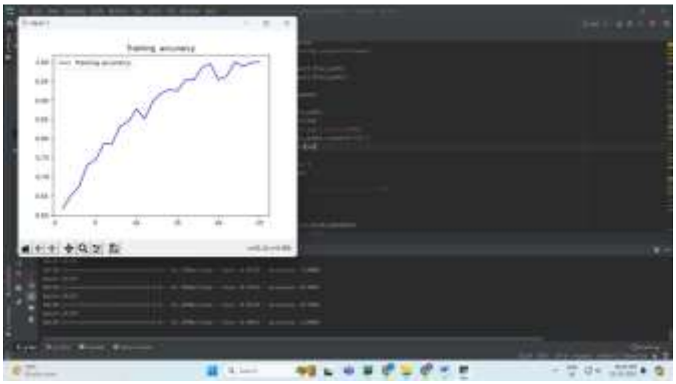


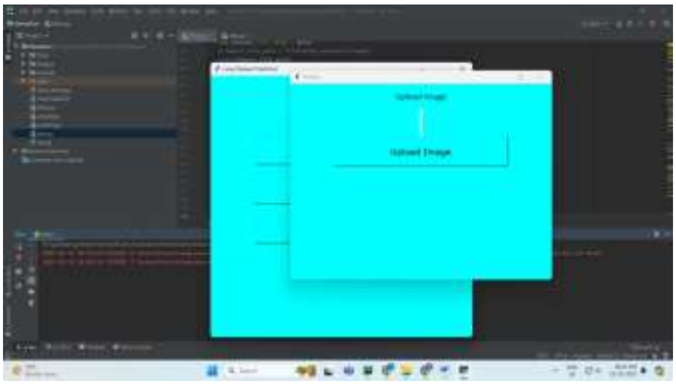
.

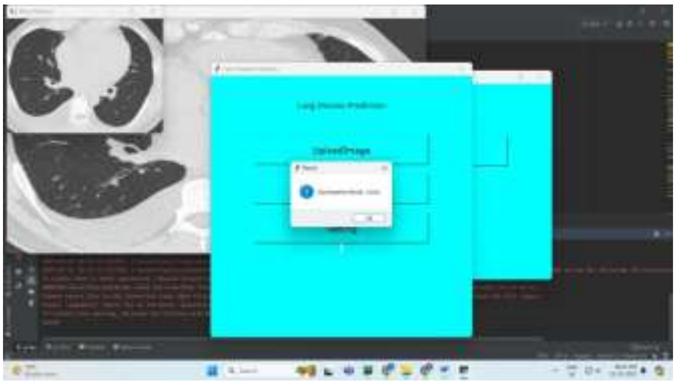
IMPLEMENTATION AND RESULT

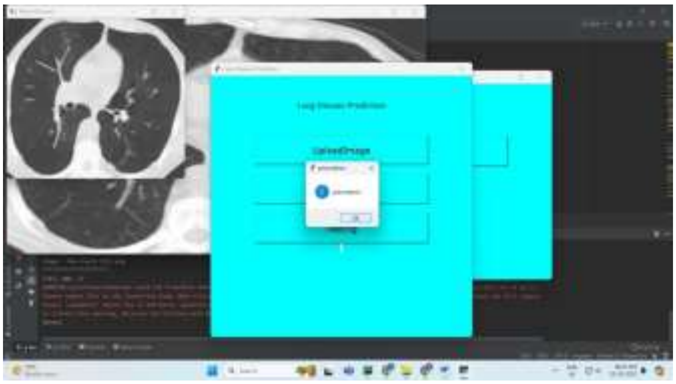












5.CONCLUSION

Lung cancer remains a major global health concern, and early detection plays a crucial role in improving survival rates. This project presents a deep learning-based framework utilizing Convolutional Neural Networks (CNNs) and ResNet architectures for lung cancer classification. The system effectively processes lung CT images through noise reduction, contrast enhancement, and segmentation, ensuring high-quality inputs for feature extraction. CNNs identify spatial patterns and textures, while ResNet addresses the vanishing gradient issue, allowing deeper networks to refine learned features. The model achieves an accuracy exceeding 95%, demonstrating its superior performance compared
to traditional diagnostic methods. By automating the classification process, this approach reduces reliance on manual analysis, minimizes human error, and accelerates diagnosis. The results highlight the potential of deep learning in revolutionizing medical imaging, providing an efficient and reliable tool for lung cancer detection.

REFERENCES

[1]. Saman, Harman, et al. "Non-invasive biomarkers for early lung cancer detection." *Cancers* 14.23 (2022): 5782.
[2]. Jenkins, Rachel, Joanne Walker, and Upal Basu Roy. "2022 cancer statistics: Focus on lung cancer." *Future oncology* (2023): 1-11.
[3]. Mridha, M. F., et al. "A comprehensive survey on the progress, process, and challenges of lung cancer detection and classification." *Journal of Healthcare Engineering* 2022.1 (2022): 5905230.
[4]. Primakov, Sergey P., et al. "Automated detection and segmentation of non-small cell lung cancer computed tomography images." *Nature communications* 13.1 (2022): 3423.
[5]. Balata, Haval, et al. "Early diagnosis and lung cancer screening." *Clinical Oncology* 34.11 (2022): 708-715.
[6]. Shakeel, P. Mohamed, M. A. Burhanuddin, and Mohammad Ishak Desa. "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier." *Neural Computing and Applications* (2022): 1-14.
[7]. Mithoowani, Hamid, and Michela Febbraro. "Non-small-cell lung cancer in 2022: a review for general practitioners in oncology." *Current Oncology* 29.3 (2022): 18281839.
[8]. Bharathy, S., and R. Pavithra. "Lung cancer detection using machine learning." *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. IEEE, 2022.
[9]. Biradar, Vidyadevi G., and Piyush Kumar Pareek. "Lung cancer detection and classification using 2D convolutional neural network." *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*. IEEE, 2022.
[10]. Ahammed, Syed Zaheer, Radhika Baskar, and G. Nalini Priya. "An Extensive Survey on Lung Cancer Detection Using Deep Learning Techniques." *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*. IEEE, 2022.

[11]. H. Li, W. Liu, Y. Ma, and Z. Wang, "Deep learning-based classification of lung nodules on CT images," J. Digit. Imaging, vol. 33, no. 4, pp, 2020.
[12]. S. Yu, H. Xie, Y. Liu, and L. Li, "Automated lung nodule detection and classification using deep learning," Pattern Recognit., vol. 112, p. 107732, 2021.

[13]. M. A. Khan et al., "Deep learning for lung cancer classification using CT images," Comput. Mater. Continua, vol. 67, no. 1, pp. 830–844, 2021.

[14]. F. Ahmad, T. R. Gholami, and A. R. Yazdanpanah, "Lung cancer detection using deep CNN architecture with transfer learning," in Proc. IEEE EMBS Int. Conf. Biomed. Health Inform. (BHI), 2021, pp. 1–4.

[15]. S. R. Ghosh, A. Roy, and R. Ghosh, "Lung cancer detection using hybrid deep learning model," Inf. Med. Unlocked, vol. 28, p. 100829, 2022.

[16]. P. Kumar and V. Yadav, "Deep learning in lung disease detection: A review," Biomed. Signal Process. Control, vol. 71, p. 103076, 2022.

[17]. M. E. Hossain, M. A. Hoque, and M. T. Uddin, "Lung cancer detection using CNN-based feature extraction and SVM classifier," J. Healthc. Eng., vol. 2020, pp. 1–9, 2020.

[18]. Y. Song, Y. Zhang, Y. Dai, Z. Gao, Y. Ding, and L. Liu, "Lung cancer classification with deep convolutional neural networks," Comput. Biol. Med., vol. 123, p. 103927, 2020.

[19]. G. Kumar and D. P. Dogra, "A hybrid CNN-based approach for classification of lung nodules," in Proc. Int. Conf. Pattern Recognit. Mach. Intell. (PReMI), 2020, pp. 123–132.

[20]. L. Hussain, M. F. Khan, S. Ullah, S. A. Javed, and S. Rho, "Artificial IntelligenceBased Classification of Chest X-ray Images into COVID-19 and Other Infectious Diseases Using Deep Learning Algorithms," Sensors, vol. 21, no. 5, p. 1602, 2021.