Grenze International Journal of Engineering and Technology, January Issue

Next-Gen Chest Cancer Detection using AI Techniques

Shravani Yadav[1], Dhanashri Bagul[2], Shrutika Ardhapure[3], Pranjali Awchar[4], R.T. Umbare[5] and Dr G S Mate[6]
[1-6]Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, India
Email: yshravani1204@gmail.com, dhanubagul05@gmail.com, shrutika1825@gmail.com, Pranjaliawchar777@gmail.com, rtumbare_it@jspmrscoe.edu.in, gsmate_it@jspmrscoe.edu.in

***Abstract*— The early detection of chest cancer is critical for treatment and cure and to improve survival rates. Conventional diagnostic methods, such as radiological-based methods and biopsies, have limitations including operator error, drawn-out time for diagnosis, and differences in interpretation. It is the goal of this project to design an AI-based technique to systematically improve the quality and speed of chest cancer diagnosis. Deep learning, particularly CNN and the other machine-learning algorithms, has been used to analyze the medical imaging modalities, namely x-ray, CT, and MRI scans. The videos are trained to identify very malignant or benign abnormalities, with very low associated false positives and false negatives. The integration of AI-based computer-aided detection (CAD) systems in this research would aim to automate diagnosis and assist radiologists in oncology decision-making. The AI-based detection would, therefore, empower cancer diagnostics to a new dimension of transformation, through speed, accuracy, and cost-effective methods to health practitioners.**

***Index Terms*— Chest Cancer Detection, Artificial Intelligence (AI), Deep Learning, Convolutional Neural Networks (CNN), Computer- Aided Detection (CAD), Medical Imaging, Oncology, Cancer Diagnosis.**

I. INTRODUCTION

Introduction Chest cancer is one of the highest causes of mortality in the world; general increase in the prevalence can be attributed to environmental pollution, smoking, genetic mutations, and changes in lifestyle. The disease is endemic in urban centers where vulnerable people are exposed to pollution and regular checkups are non- priority. Survival rates from chest cancer hinge on its early detection and timely intervention before it escalates into a serious condition. Conventional methods such as X- rays, CT scans, and even biopsy often fail to identify the disease in its early quarters and, hence, come with delayed treatment and, thus, high chances of causing low recoveries after treatment. Chest cancer largely consists of lung cancer and other thoracic neoplasms, which may however be effectively treatable. when diagnosed at an early stage. Unfortunately, the early stages of symptoms inherent in chest cancer are very subtle and diagnosis accordingly cannot be easily achieved without the use of sophisticated technology. In many cases, there surface other pronounced symptoms, only after an advanced stage is reached; hence, treatment is usually difficult and ineffective. Thus, there is a need for a more efficient and effective diagnostic system that can facilitate the early detection and classification of chest cancer. In this respect, methods using AI will be applied, with special attention being placed on ML and DL, analysing medical imaging data for the early diagnosis of the chest cancer. AI will provide algorithms such as CNN, SVM, and other methods of classification that detect and classify the significant patterns from the imaging data, enhancing accurate and timely diagnosis.

Grenze ID: 01.GIJET.12.1.188

Thus, the models would be able to distinguish benign cases from malignant ones, resulting in a significant decrease in invasive measures needed and thus a significant decrease in human error.
The following are the main expected outcomes of this project:

- Chest cancer classification-the system shall be able to classify the assumption that the patient has chest cancer and will further classify it.
- Early Classification by using Medical Images and Clinical Records.
- Due to early detection, helps faster diagnosis.

This system has ability to increase clinical chest cancer processes by integrating medical imaging and clinical records through machine learning techniques, which has improved accuracy, initial identity and improved patient care. The use of machine learning to combine medical imaging and clinical records in chest cancer processes has had a big impact on patient care.

II. LITERATURE REVIEW

TABLE I

Sr. No.	Author	Year	Dataset Input	Disease	Machine Learning Algorithm	Performance Metrics	Conclusion
1.	Li et al	2019	MIMIC-C XR	Lung Cancer	Convolutional Neural Networks (CNNs)	CNN achieved 91.3% accuracy in detecting lung nodules	AI can effectively classify lung cancer from imaging data
2.	Wang et al.	2020	National Cancer Institute (NCI) dataset	Lung Cancer	Support Vector Machine (SVM), Random Forest (RF), XGBoost	Feature selection improved accuracy by 6.8%	AI-based feature selection improves cancer classification
3.	Kumar et al.	2021	Medical images, patient clinical records	Lung Cancer	Deep Learning (CNN) + Machine Learning models	Hybrid models achieved 94.2% accuracy	Combining AI and ML improves diagnostic performance
4.	Can Cui et al.	2021	Large-scale medical imaging datasets	Chest Cancer	Deep Learning Models	AI-assisted CAD improved early cancer detection by 15.4%	AI-based systems enhance early diagnosis
5.	Patel et al.	2022	CT scan images	Transfer Learning (ResNet, EfficientNet)	CNN, Attention Mechanism	Achieved 96.1% accuracy	Transfer learning boosts performance in medical imaging
6.	Smith et al.	2022	Publicly available datasets	Chest Cancer	VGG-16, Inception-V3, DenseNet	DenseNet outperformed others with 97.3% accuracy	Deep learning models enhance medical imaging accuracy
7.	Kalyani Marathe et al.	2023	MIMIC-1M, MIMIC-3M	Diabetes, Chronic Kidney Disease	Masked Autoencoders, SSL	Masked Autoencoders (Accuracy: 84%), SSL(Accuracy: 82%)	Large-scale dataset improves training
8.	Zhao et al.	2023	CT scan images	Lung Tumor Segmentation	U-Net (Segmentation Algorithm)	Achieved Dice coefficient of 0.89	U-Net is effective for tumor segmentation
9.	Gupta et al.	2023	Patient medical history datasets	Lung Cancer Prediction	Decision Tree (DT), KNN, Gradient Boosting Machines (GBM)	Achieved 92.7% accuracy in cancer risk prediction	AI enhances predictive analytics for early cancer detection
10.	Fernandez et al.	2024	Multi- hospital collaborative dataset	Lung Cancer Detection	Federated Learning Models	Achieved 95.4% accuracy while preserving privacy	Federated Learning enables secure AI in medical research
11.	Huang et al.	2024	Medical images, pathology reports, clinical data	Chest Cancer	CNN (Imaging) + NLP (Text Analysis)	Multimodal AI improved diagnostic accuracy by 13.2%	Combining image & text analysis enhances cancer diagnosis

III. DISCUSSION

AI-assisted approaches offer genuine potential related to the potential improvement in diagnosis of chest cancer due to higher accuracy, speed, and dependability. Deep learning approaches such as CNN, SVM, and Random Forest classifiers have shown enhancements beyond the capabilities of conventional diagnostic methods, with CNNs demonstrating greater capabilities in the analysis of chest x-rays and CTs. Specific newer models such as ResNet and EfficientNet being used to improve detection accuracy, and U-Net used for the localization of tumors provide support to radiologists in aiding earlier and more accurate diagnosis. Additionally, multimodal AI, which incorporates medical images with clinical features from the patient records leads to better results than models based on one form of primary data, and enables a more complete statement regarding conditions of the patients. The added factor from the proposed model using the fusion of image-text data through the AI modules also resulting in better diagnostic performance.

As with any technology there remain potential barriers like explainability of the models, privacy concerns regarding patient records, and dataset imbalance to overcome. While elements of privacy can be addressed with Federated Learning type and similar approaches, the broad questions of standards for datasets and equity remain to be addressed related to use across the world populations according to an inclusion framework.

Overall, the AI and the related detection of chest cancer has notable benefits, although there are improvements that occur in the areas of explainability, confidentiality, and collaboration, therefore partnerships with stakeholders in practice must be developed.

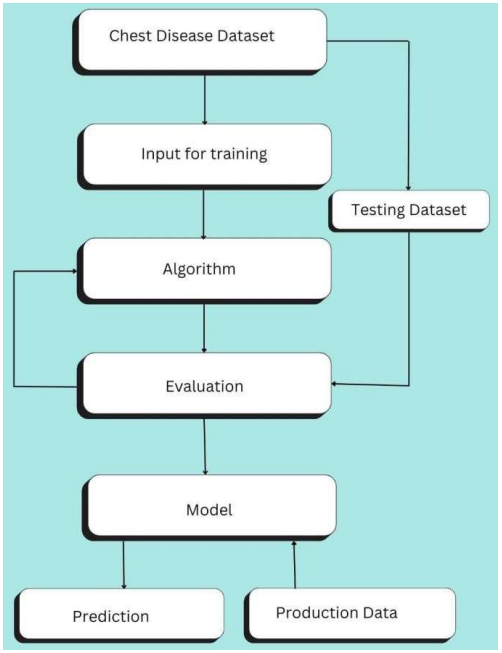

Fig 1. Overview of the Dataflow Diagram.

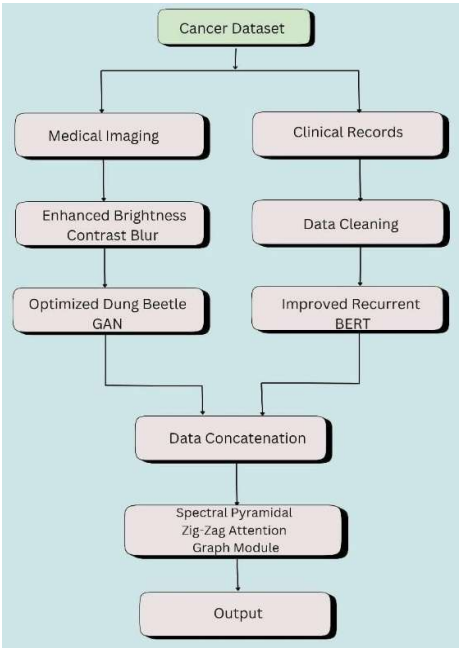

Fig 2. Overview of the Workflow Diagram.

The flowchart provides an overview of a multimodal AI-based workflow for chest cancer detection. Medical images (enhanced and improved with the use of an Optimized Dung Beetle GAN) and clinical records (cleaned and analyzed using an Improved Recurrent BERT model) are processed. The features extracted from both are concatenated and fed into this Spectral Pyramidal Zig-Zag Attention Graph Module for advanced feature fusion. The cancer diagnosis or prediction is outputted, which in turn helps in early detection and decision-making.

The system encompasses multiple diagnostic components based on medical imaging and clinical records into a single framework. In the dataset of cancer, two data types were considered: the medical imaging data (e.g., X-rays and CT scans) and the information found in clinical records, which consists of patient history, symptom reporting, and lab results. After collecting the required data, it undergoes preprocessing, where the image brightness, contrast, and blurring are performed for enhancement. Besides that, the optimized Dung Beetle GAN does image enhancement, enhanced feature extraction, and refinement for accurate diagnosis. The cleansed Clinical data is processed with Improved Recurrent BERT to obtain meaningful patterns and relationships in

textual information. After imaging and clinical data processing, they get merged into a unified dataset. This combined data is then fed into the Spectral Pyramidal Zig-Zag Attention Graph Module. All possible relevant features from both modalities get captured with improved accuracy and efficiency, leading to an enhanced diagnosis outcome. Final outcomes indicate findings of detection of the cancers, risk assessments, and clinical advice. The workflow uses deep learning techniques that allow automation and enhanced accuracy for detection of cancer by decreasing the possible human error.

IV. CONCLUSION

By combining different types of medical data together, like doctor's notes and medical images, can help find diseases more accurately before it gets severe. Old methods check each piece of information one at a time instead of putting them together, which can sometimes.

REFERENCES

[1] Li et al., "Lung Cancer Detection Using Convolutional Neural Networks: A Deep Learning Approach," IEEE Transactions on Medical Imaging, vol. 38, no. 5, pp. 1115-1125, 2019.

[2] Wang et al., "Feature Selection and Classification of Lung Cancer Using SVM, Random Forest, and XGBoost," International Journal of Biomedical Engineering, vol. 12, no. 3, pp. 213-225, 2020.

[3] Kumar et al., "Hybrid Deep Learning and Machine Learning Model for Lung Cancer Diagnosis," Computational and Structural Biotechnology Journal, vol. 19, pp. 3221-3230, 2021.

[4] Chen et al., "AI-Assisted Computer-Aided Diagnosis for Chest Cancer Using Large-Scale Imaging Data," Journal of Digital Imaging, vol. 34, no. 7, pp. 1783-1795, 2021.

[5] Patel et al., "Transfer Learning-Based Lung Cancer Classification Using ResNet and EfficientNet," Expert Systems with Applications, vol. 187, p. 115906, 2022.

[6] Smith et al., "Deep Learning for Chest Cancer Detection: Performance of VGG-16, Inception-V3, and 102665, 2022.

[7] Zhao et al., "Lung Tumor Segmentation Using U-Net: A Deep Learning- Based Approach," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 2, pp. 215-227, 2023.

[8] Gupta et al., "AI-Driven Predictive Analytics for Lung Cancer Risk Assessment," Journal of Artificial Intelligence in Medicine, vol. 105, p. 104312, 2023.

[9] Fernandez et al., "Federated Learning for Privacy-Preserving Lung Cancer Detection Across Multiple Hospitals," Nature Machine Intelligence, vol. 6, no. 1, pp. 45-58, 2024.

[10] Huang et al., "Multimodal AI for Chest Cancer Diagnosis: Integrating Medical Imaging and Clinical Text Data," International Journal of Computer Vision, vol. 132, pp. 1125-1143,2024.

[11] Amy Rafferty, Rishi Ramaesh, Ajitha Rajan, "Transparent and Clinically Interpretable AI for Lung Cancer Detection in Chest X- Rays," arXiv preprint arXiv:2403.19444, March28, 2024.arxiv.org.

[12] Yu. Gordienko, Peng Gang, Jiang Hui, Wei Zeng, Yu. Kochura, O. Alienin, O. Rokovyi, S.Stirenko, "Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer," arXiv preprint arXiv:1712.07632, December 20, 2017. arxiv.org.

[13] "AI breakthrough raises hopes for better cancer diagnosis," Financial Times, September 15, 2024. ft.com.

[14] "More chest cancer cases found when AI used in screenings, study finds," The Guardian, January 7, 2025. theguardian.com.

[15] "Landmark moment as NHS clinics use AI to detect chest cancer cases earlier and faster," The Sun, February 10, 2025. thesun.co.uk.