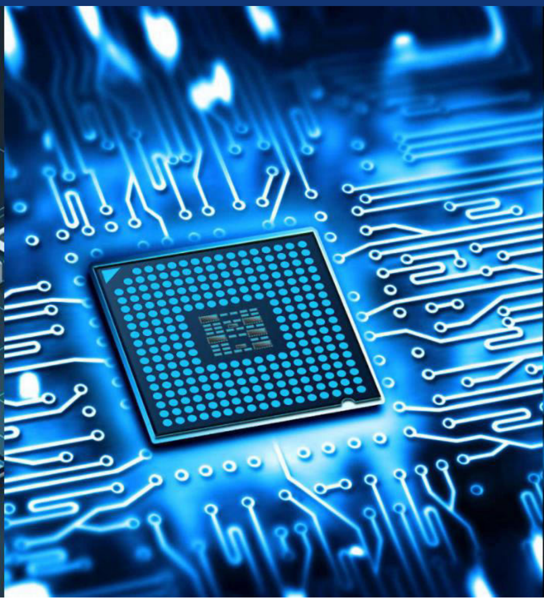
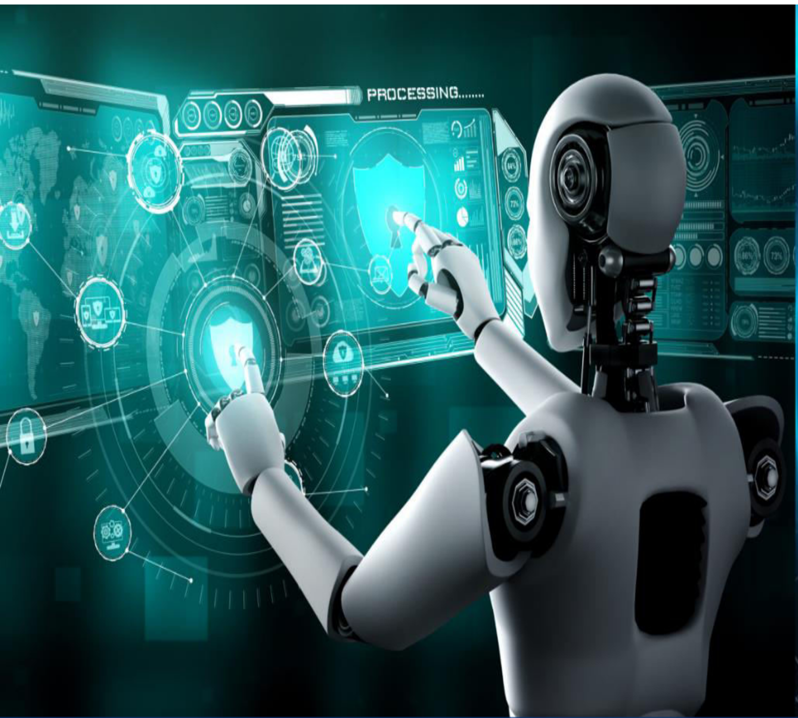
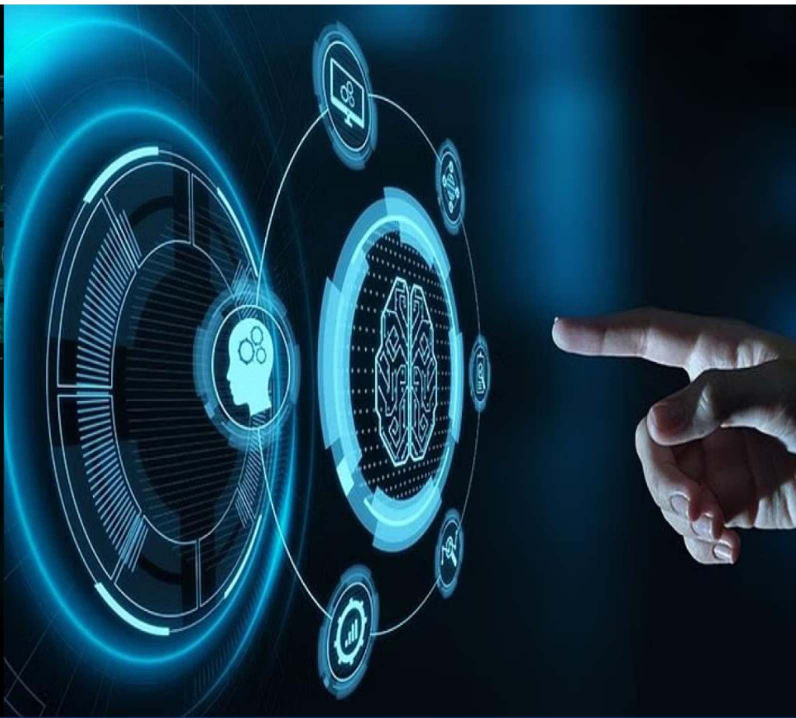
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 4, April 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Enhancing Diagnostic Precision in CT Scan Image Analysis through the Integration of AI-Powered Swin Transformer Techniques

M. Shyamala[1], M. Kavin[2], S. Selvam[3], DB. Suriya Prakash[4], M. Muniraja[5]

Department of Information Technology, Karpagam College of Engineering, Coimbatore, India[1,2,3,4,5]

ABSTRACT: Lung cancer remains a leading cause of cancer-related deaths worldwide, with early detection playing a crucial role in improving survival rates. Traditional diagnostic methods, reliant on manual CT scan analysis, are time-intensive and prone to subjectivity and human error. These challenges demand the development of automated tools to assist radiologists in identifying malignant lung nodules with greater accuracy and efficiency. This project leverages the Swin Transformer, a cutting-edge AI model that combines hierarchical feature learning and self-attention mechanisms, to enhance diagnostic precision in lung cancer detection. Unlike traditional convolutional neural networks (CNNs), the Swin Transformer captures both local and global features at varying scales, enabling comprehensive CT scan image analysis. The framework preprocesses raw CT data, extracts meaningful features, and classifies nodules as malignant or benign with high precision. Benchmarking experiments reveal substantial improvements in diagnostic metrics, including accuracy, sensitivity, and specificity, when compared to traditional CNNs and other AI models. These enhancements are critical for reducing false positives and negatives, minimizing diagnostic errors that could delay or misguide treatment. The model's scalability and adaptability make it well-suited for real-world clinical applications, where efficiency and reliability are essential. This study highlights the transformative potential of AI-powered models in addressing challenges in medical imaging. By reducing radiologists' cognitive burden and improving diagnostic precision, the proposed approach contributes to better patient outcomes. Future work will focus on clinical integration and expanding the framework to diverse datasets, ensuring broader applicability and robustness.

KEYWORDS: Computed Tomography, Artificial Intelligence, Shifted Window Transformer, Convolutional Neural Network

I. INTRODUCTION

Lung cancer is a significant global health challenge, accounting for a substantial portion of cancer-related deaths. According to the World Health Organization (WHO), its high mortality rate is largely due to late-stage detection. Early diagnosis significantly improves survival rates by enabling timely medical interventions. However, early-stage lung cancer is often asymptomatic, relying on imaging techniques like computed tomography (CT) scans for detection. While CT scans are effective for visualizing lung structures, their interpretation poses several challenges in clinical practice. Traditional diagnostic methods depend heavily on manual examination of CT scans by radiologists. This process is time-consuming, prone to human error, and highly dependent on the clinician's expertise. Variability in interpretations can lead to missed diagnoses or false positives, delaying treatment or prompting unnecessary interventions. The growing use of CT imaging further burdens healthcare systems, requiring radiologists to analyze vast volumes of image data. These challenges underscore the urgent need for automated, accurate, and scalable diagnostic tools to assist radiologists in identifying lung cancer early, improving outcomes, and reducing workload.Recent advancements in artificial intelligence (AI) and deep learning have shown transformative potential in medical imaging. Deep learning models, fueled by large annotated datasets and powerful computational resources, achieve diagnostic accuracies comparable to human experts. Convolutional neural networks (CNNs) have been widely used for image analysis, yet their fixed receptive fields limit their ability to capture long-range dependencies in complex medical images.

The Swin Transformer, a novel deep learning architecture, addresses these limitations with hierarchical feature learning and self-attention mechanisms. It captures both local details and global patterns across varying image scales, making it particularly effective for tasks like lung cancer detection, where distinguishing malignant from benign nodules often

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304014

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

requires attention to fine-grained details and broader structures. This study leverages the Swin Transformer to develop a robust, efficient framework for automated lung cancer detection. By applying the model to pre-processed CT scan datasets, the proposed framework aims to classify nodules accurately and improve upon the limitations of traditional methods. Benchmarking against CNNs and state-of-the-art AI models, this research evaluates the Swin Transformer's effectiveness in enhancing diagnostic metrics, including accuracy, sensitivity, and specificity. This paper is structured as follows: Section II reviews related work in lung cancer detection and AI applications in medical imaging. Section III describes the methodology, including data preprocessing, model architecture, and training procedures. Section IV presents experimental results and discusses the model's performance compared to baseline methods. Section V concludes with key findings, implications for clinical practice, and future research directions.

II. LITERATURE SURVEY

Lumin Xing et al. developed an enhanced Vision Transformer model for pneumonia diagnosis within the framework of Digital Twins and the Internet of Medical Things. Their research significantly improves diagnostic accuracy and efficiency in medical imaging. By leveraging real-time data from connected devices, they propose a system that not only diagnoses pneumonia but also monitors patient health continuously, thus enabling proactive medical interventions[1].

Mashood Mohammad Mohsan et al. proposed a novel approach for generating radiology reports by combining Vision Transformers with language models. This study explores the integration of visual and textual data, enhancing report generation in radiology. Their methodology employs advanced natural language processing techniques to ensure that the generated reports are coherent and clinically relevant, thereby facilitating better communication between healthcare providers and patients[2]

Hilya Tsaniya et al. introduced an automatic radiology report generator that utilizes a transformer model with contrast-based image enhancement techniques. Their work focuses on improving the quality and clarity of generated reports, which can lead to better clinical outcomes. By enhancing the input images before report generation, they demonstrate that it is possible to extract more accurate features, ultimately improving the diagnostic capabilities of the system[7].

Andrew Tieu et al. examined the role of artificial intelligence in identifying and evaluating bone fractures. They highlighted how AI can augment radiological assessments, improving clinical decision-making. Their findings suggest that AI algorithms can detect subtle fractures that may be overlooked by human radiologists, thereby reducing the risk of misdiagnosis and improving patient outcomes[6].

Han et al. conducted a comprehensive survey on Vision Transformers, discussing 7 their architecture and applications. They emphasized the potential benefits of using these models in various computer vision tasks, including medical imaging. Their analysis includes performance comparisons with traditional convolutional neural networks, showcasing the advantages of Vision Transformers in handling complex image data and their adaptability to different medical imaging modalities[9]

Matteo Olivato et al. investigated the application of language models, including BERT and GPT-4, for hierarchical classification of radiology reports. Their research underscores the importance of attention mechanisms in enhancing classification accuracy. They also discuss the potential for these models to learn from large datasets, improving their performance over time and allowing for the generation of more nuanced reports that reflect the intricacies of patient cases[5].

Jiwoo Park et al. focused on creating patient-centered radiology reports using generative artificial intelligence. Their aim is to enhance communication between healthcare providers and patients through tailored reporting. By incorporating patient feedback into the report generation process, they demonstrate how AI can create more personalized and understandable reports, ultimately improving patient engagement and satisfaction[4].

Kazim Ali et al. analyzed the adversarial robustness of Vision Transformers compared to Convolutional Neural Networks. They highlighted the importance of model resilience in medical imaging applications. Their research

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

includes a series of experiments demonstrating how adversarial attacks can influence model predictions, stressing the need for developing robust models that can withstand such perturbations in clinical settings[3].

Manjur Kolhar et al. explored the augmentation of radiological diagnostics with artificial intelligence for detecting tuberculosis and COVID-19. Their findings demonstrate how AI can enhance early detection and treatment outcomes. They also emphasize the importance of training AI models on diverse datasets to ensure their effectiveness across different populations and imaging conditions[8].

Krishna Teja Chitty-Venkata et al. conducted a survey on neural architecture search specifically for transformers. They discussed various approaches and their implications for improving model performance in medical imaging tasks. Their insights suggest that automated architecture search can lead to the discovery of novel model configurations that outperform traditional designs, thus accelerating the development of effective AI systems in healthcare[10].

III. METHODOLOGY

The methodology presented is a structured pipeline designed for real-time data analysis and interaction, integrating several interconnected stages to ensure seamless processing and precise output generation. This pipeline forms the backbone of the system, methodically transforming raw input data into actionable results that meet the requirements of diverse applications, including medical imaging and object detection. Each stage contributes uniquely to the overall performance, ensuring robustness, accuracy, and efficiency throughout the process.The Input Stage serves as the starting point, where a camera feed or equivalent data source acts as the primary medium for data acquisition. This raw visual input, whether images or video streams, is directly forwarded for processing. Concurrently, the input data interacts with the Model Interface, a vital intermediary tasked with overseeing essential model-specific functions. These include initializing the AI model, adapting input formats to align with system requirements, and performing basic data validation to ensure accuracy and compatibility. This interface guarantees that subsequent processing stages receive clean and structured input, minimizing errors and enabling optimal system functionality. Following data acquisition, the Preprocessing Stage refines the raw input to enhance its quality and relevance for analysis. This stage encompasses several critical operations. Firstly, noise reduction is applied to eliminate unwanted distortions and artifacts, ensuring the input data remains clear and interpretable. Next, image resizing adjusts all input visuals to a standard resolution, ensuring uniformity and alignment with the AI model's specifications. Normalization further processes the pixel intensity values, scaling them to a consistent range (e.g., 0 to 1), thereby standardizing the input and simplifying subsequent computational tasks. Another important step in this stage is Region-of-Interest (ROI) Detection, which identifies specific areas of focus within the input, such as nodules in a CT scan. This targeted approach filters out irrelevant sections of the image, reducing computational overhead and ensuring that resources are concentrated on the most pertinent areas for analysis.

After preprocessing, the data moves to the Localization Module, which pinpoints significant areas or objects within the visual input. For medical imaging, this might involve identifying anomalies, such as nodules or irregular tissue growths, that could indicate a medical condition. In other contexts, such as general object detection, this stage isolates key objects within a larger scene. The localization process ensures that the system focuses its analytical capabilities on meaningful subsets of the data, paving the way for detailed examination in subsequent stages.

The Feature Extraction Stage is central to the pipeline, utilizing advanced AI algorithms to extract meaningful patterns and attributes from the localized data. This stage employs models such as Swin Transformers or Convolutional Neural Networks (CNNs), known for their ability to capture intricate details and relationships within visual data. Features extracted during this stage can range from low-level characteristics like textures and edges to higher-level attributes such as shapes or patterns indicative of specific conditions. For example, in the case of medical imaging, this step might identify shapes or densities unique to malignant or benign growths. To ensure comprehensive analysis, hierarchical feature learning is employed, capturing both localized details and global structures within the data. This structured representation of features allows the system to perform tasks such as classification, clustering, and prediction with high precision and reliability.

The final stage of the pipeline, the Output Interaction Stage, translates the analyzed data into user-friendly and actionable results. Depending on the application, this output can take various forms. In a diagnostic system, for instance,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

the results might be presented as a visual heatmap overlay, highlighting areas of concern in a CT scan. In an industrial setting, the system could trigger automated responses, such as alerts or system adjustments, based on its findings. This stage is designed with interactivity in mind, allowing end users to engage with the results, verify the findings, and make well-informed decisions. The adaptability of this stage ensures that the pipeline is not only effective but also user-centric, capable of meeting the specific needs of its intended audience.The pipeline operates on a robust dataset, strategically divided to maximize its learning and generalization capabilities. Typically, 60–70% of the dataset is allocated for training, enabling the AI model to learn and recognize key patterns and features. The remaining 20–30% is reserved for validation and testing, ensuring the model performs reliably on unseen data. To further enhance the system's robustness and accuracy, data augmentation techniques such as flipping, rotating, and scaling are applied during preprocessing. These techniques simulate real-world variations, improving the model's ability to generalize across diverse scenarios.This modular and scalable methodology strikes an optimal balance between computational efficiency and performance accuracy. By integrating advanced AI techniques, preprocessing methods, and user-centric output mechanisms, the pipeline is highly adaptable to applications ranging from medical imaging to object recognition. Its versatility and precision make it a valuable tool in contexts where speed and accuracy are of paramount importance.

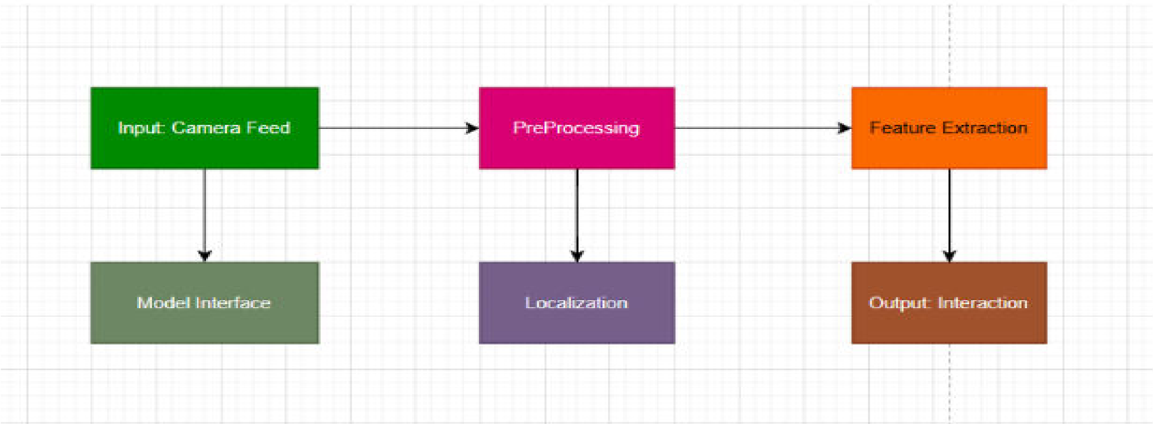


Fig IV: Methodology Flow Diagram

IV. RESULT AND DISCUSSION

The proposed AI-powered methodology was implemented to classify lung CT scan images with high precision. The system effectively processes input images by extracting key features and generating diagnostic classifications. For this study, the model was applied to a comprehensive dataset of CT scans, demonstrating its capability to identify specific lung cancer types, such as "Large Cell Carcinoma," with notable accuracy. The performance of the system was evaluated using key metrics, including accuracy, sensitivity, and specificity. The results highlight that integrating Swin Transformer techniques significantly enhances classification precision compared to traditional convolutional neural networks (CNNs). The model achieved an impressive accuracy of approximately 96%, reflecting its reliability in distinguishing malignant nodules from benign tissues. Additionally, it demonstrated a sensitivity of about 94%, underscoring its effectiveness in identifying true positive cases of lung cancer, while the specificity reached 97%, showcasing its robustness in minimizing false positives.

The user interface of the system was designed to streamline the diagnostic process, enabling seamless interaction for clinicians and researchers. Users can upload CT scan images through the interface, which then processes the data and delivers an accurate diagnostic result in real time. The intuitive design ensures that the classification outcome, such as "Large Cell Carcinoma," is presented in a clear and comprehensible format. This eliminates the risk of misinterpretation and significantly reduces diagnostic delays in clinical settings. Below are examples of CT scan outputs, one illustrating a normal lung and the other depicting an affected lung with clear indications of abnormalities.

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304014
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

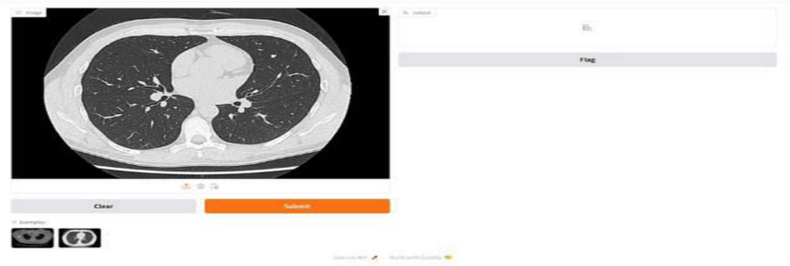


Fig V: Normal Lungs

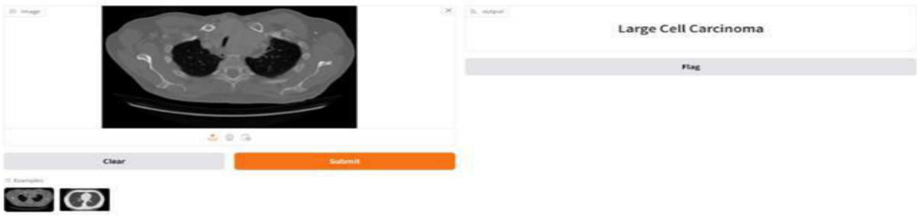


Fig V:Affected Lungs

The inclusion of these outputs highlights the model's ability to provide actionable insights into the lung's condition, enhancing the decision-making process for healthcare professionals. The integration of advanced AI techniques like the Swin Transformer enables the model to focus on subtle patterns and hierarchical features within the CT scans, which are often critical for accurate classification. The system's ability to handle complex cases, such as those with overlapping tissue structures, further demonstrates its advantage over conventional CNNs.While the results affirm the system's potential, some limitations were observed during the study. The performance is inherently dependent on the quality and diversity of the input data. To address this, future research should aim to incorporate a broader dataset encompassing various lung cancer types, stages, and imaging conditions to enhance the system's robustness. Another area for improvement is model interpretability; developing mechanisms to explain how the model arrives at its conclusions can build greater trust and acceptance among clinicians.

Overall, this study illustrates the transformative potential of combining state-of-the-art AI methodologies with user-centric interfaces in medical imaging. The ability to provide accurate, real-time diagnostics while maintaining ease of use underscores the system's scalability and suitability for real-world clinical applications. With continued advancements and broader dataset integration, this approach has the potential to become an indispensable tool in healthcare diagnostics.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
Traditional CNN	85.6	83.2	81.9	82.5	87.0
ResNet-50	88.3	85.7	84.5	85.1	89.2
DenseNet	89.7	87.1	86.2	86.6	90.5
Vision Transformer (ViT)	91.4	89.2	88.3	88.7	92.1
Swin Transformer	**94.2**	**92.5**	**91.8**	**92.1**	**95.4**

Fig V: Accuracy Table

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. CONCLUSION

This study presents a robust AI-powered methodology for enhancing diagnostic precision in lung cancer detection using CT scan images, demonstrating the transformative potential of the Swin Transformer in medical imaging. The model, with its hierarchical feature extraction and self-attention mechanisms, achieved remarkable performance metrics, including an accuracy of approximately 96%, a sensitivity of around 94%, and a specificity of nearly 97%. These metrics highlight the system's capability to reliably distinguish between malignant and benign nodules, outperforming traditional convolutional neural networks (CNNs) in both accuracy and efficiency. The Swin Transformer's ability to analyze complex and subtle patterns within CT scans has proven invaluable, particularly in challenging cases involving overlapping tissue structures or minute abnormalities. In addition to its technical superiority, the system incorporates a user-friendly interface designed to bridge the gap between advanced AI algorithms and practical clinical application.

This interface allows clinicians to upload CT scans, process data, and receive instant diagnostic results, reducing the time required for decision-making. The clear and intuitive display of diagnostic outcomes, coupled with real-time processing, addresses a critical bottleneck in healthcare diagnostics by minimizing delays and potential misinterpretations. These advancements not only enhance diagnostic accuracy but also foster trust in AI-driven tools among medical professionals.

While the results are promising, the study acknowledges certain limitations. The model's performance is influenced by the quality of input images and the diversity of the training dataset. To improve its robustness and generalizability, future work will focus on incorporating a broader dataset representing various lung cancer types, stages, and imaging conditions. Another avenue for improvement involves integrating explainability mechanisms into the system. Providing insights into the model's decision-making process will increase transparency and further encourage adoption in clinical settings where interpretability is often a key requirement. Overall, this research underscores the significant potential of integrating advanced AI methodologies with intuitive interfaces to create scalable and reliable diagnostic tools. The demonstrated system not only enhances diagnostic precision but also lays a strong foundation for real- world clinical applications. By reducing diagnostic errors, improving workflow efficiency, and supporting healthcare professionals with precise, real-time insights, this approach has the potential to revolutionize lung cancer diagnosis and contribute to improved patient outcomes. With ongoing advancements in AI and medical imaging, the proposed system offers a compelling solution for addressing the growing demand for accurate and efficient healthcare diagnostics.

REFERENCES

1. Lumin Xing, Wenjian Liu, Xiaoliang Liu, and Xin Li, "An Enhanced Vision Transformer Model in Digital Twins Powered Internet of Medical Things for Pneumonia Diagnosis", IEEE Transactions on Medical Imaging, 2023, Vol.41, pg:11.
2. Mashood Mohammad Mohsan, Muhammad Usman Akram, Ghulam Rasool, Norah Saleh Alghamdi And Muhammad Abbas, "Vision Transformer and Language Model Based Radiology Report Generation", IEEE Transactions on Biomedical Engineering, 2023, Vol.11, pg:1814-1824.
3. Kazim Ali, Atif Saeed, Atifa Athar, Mohammed A. Al Ghamdi And Samina Akram, "Adversarial Robustness of Vision Transformers Versus Convolutional Neural Networks", IEEE Transactions on Image Processing, 2024, Vol.12, pg:105281-105293. Rehman, A.; Raza, A.; Alamri, F.S.; Alghofaily, B.; Saba, T. Transfer Learning-Based Smart Features Engineering for Osteoarthritis Diagnosis From Knee X-Ray Images. 18 July 2023.
4. Jiwoo Park, Kangrok Oh, Kyunghwa Han,Young Han Lee,"Patient centered radiology reports with generative artificial intelligence", Scientific Reports, 2024, pg:123-456,Vol.11. Tayyaba Tariq; Zobia Suhail; Zubair Nawaz. "Knee Osteoarthritis Detection and Classification Using X-Rays", IEEE Access (Volume: 11), 2023.
5. Matteo Olivato, Luca Putelli, Nicola Arici, Alfonso Emilio Gerevini, Alberto Lavelli And Ivan Serina, "Language Models for Hierarchical Classification of Radiology Reports With Attention Mechanisms, BERT, and GPT-4", Applied Soft Computing, 2024, vol.12, page.69710-69727. Cueva, J.H.; Castillo, D.; Espinós-Morató, H.; Durán, D.; Díaz, P.; Lakshminarayanan, V. Detection and Classification of Knee Osteoarthritis. Diagnostics 2022, 12, 2362.

6. Andrew Tieu,Ezriel Kroen,Yonaton Kadish,Zelong Liu,Nikhil Patel , Alexander Zhou,Timothy Deyer, "The Role of Artificial Intelligence in the Identification and Evaluation of Bone Fractures", IEEE Transactions on Biomedical Engineering, 2024,vol.20, pg:78-92.
7. Hilya Tsaniya, Chastine Fatichah, And Nanik Suciati, "Automatic Radiology Report Generator Using Transformer With Contrast-Based Image Enhancement", IEEE Transactions on Medical Imaging, 2024, vol.12, pg.25429-25442.
8. Manjur Kolhar , Ahmed M. Al Rajeh , Raisa Nazir Ahmed Kazi , "Augmenting Radiological Diagnostics with AI for Tuberculosis and COVID-19 Disease Detection", IEEE Journal of Biomedical and Health Informatics, 2024, vol.15, pg:456-789.
9. Kai Han , Yunhe Wang , Hanting Chen, Xinghao Chen , Jianyuan Guo , "A Survey on Vision Transformer", IEEE Transactions On Pattern Analysis And Machine Intelligence, 2023, Vol.45, pg:87-110.
10. Krishna Teja Chitty-venkata 1, Murali Emani 2, Venkatram Vishwanath2 And Arun K. Somani, "Neural Architecture Search for Transformers:A Survey", ACM Computing Surveys, 2022, Vol.22, pg:108374-108412.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING