**DESIGN AND IMPLEMENTATION OF A LUNG CANCER DETECTION SYSTEM**

**BY**

|  |  |
| --- | --- |
| **OLUWUYI, OLUMIDE ISMAIL** | **21/0319** |
| **NJOKU, MCKENZIE AKACHUKWU** | **21/0680** |
| **ACHUMNU, AKACHUKWU GIFT** | **21/3022** |

**Acronyms**

CT- Computed Tomography

MRI- Magnetic resonance imaging

AI- Artificial Intelligence

ML- Machine Learning

DL- Deep Learning

CNN- Convolutional Neural Network

LCDS- Lung Cancer Detection System

**1.1 Introduction**

In modern medical settings, early detection and diagnosis of lung cancer are indispensable yet challenging. X-rays, and CT scans, are the traditional diagnostic methods widely used for identifying abnormalities in the human lungs [1]. Those practices, however often involve manual assessments and can be labor intensive which may introduce challenges related to accuracy and reliability. As healthcare environments evolve, there is an increasing demand for more efficient and technologically advanced diagnostic methods.

This study embarks on designing and implementing a lung cancer detection system intended for streamlining diagnosis in medical institutions. This system presents a unique and innovative method to make diagnosis faster. Manual and archaic methods are far from accurate, timely, or efficient. Using recent advancements in AI, this project aims to address these challenges and represents a significant step forward in applying artificial intelligence within healthcare.

**1.2 Background of the Study**

According to [1], Lung cancer is a dangerous and critical condition that occurs within the human lungs, it involves some vital and important recognition and diagnosis with the help of medical imaging. However, it is very astonishing that such a tedious task of medical image analysis has often been an ignored area. A myriad of diverse imaging modalities, such as MRI, CT, and radiography ultrasound, occasionally help describe abnormality in human organs where lung cancer exists. The standard method of reviewing and analyzing medical images, nonetheless largely depends on manual interpretation from trained experts. The human lung is designed as a rather vital organ for the exchange of oxygen and carbon dioxide, substances so important to support metabolic pathways and overall well-being further. Considering such an important role, an innovative approach to the detection of lung cancer begins with the innovative approach provided for traditional medical imaging methods. This system would analyze patterns based on features to identify abnormalities in medical images. It utilizes some high-level algorithms and neural networks for more precise detection, capturing, and analysis of these features with increased effectiveness, hence providing an even better way of early detection of cancers. Although limited in this area, the traditional methods formed a basis for the lung cancer detection system before more automated and technologically advanced systems such as the lung cancer detection system were established [2].

Artificial Intelligence has long been a cornerstone in medical imaging. In the mid-1980s, the Kurt Rossmann Laboratories at the Department of Radiology, University of Chicago, laid the bedrock for the use of computerized analysis to enhance the accuracy and efficiency of interpreting medical images [3]. This early work found significant enhancement in the work of Kunio Doi, who played an important role in the development of image-based techniques for computer-aided diagnosis in the detection of lung cancers [4]. Since then, several other works have started building on these foundational techniques to further the field of detecting lung cancer to its current state. During the last years, especially the breakthroughs of AI in DL, significant breakthroughs took place when algorithms resorted to static methods to develop problem-solving solutions. It was these breakthroughs that have succeeded in implementing ML in commercial applications and healthcare where suitable solutions can be opted for specific data types and challenges [5].

Early diagnosis of lung cancer is particularly challenging due to the absence or non-specific nature of symptoms. Lung tumors are categorized either as benign (non-cancerous) or malignant (cancerous). Benign tumors grow slowly within their defined boundaries and generally pose minimal harm, whereas malignant tumors can spread aggressively through the circulatory system. Moreover, size, shape, texture, and density add to the variability, complicating health professionals' interpretation. In light of such issues and others, automated systems are needed to detect lung cancer. Automated systems can help enhance diagnostic accuracy significantly by decreasing human error and more consistency in interpreting complex imaging data, leading to quicker and more reliable tumor detection.

Manual analysis of imaging scans is often slow and error-prone. An automated system provides a more modern and effective solution, increasing the speed and accuracy of lung cancer diagnosis while reducing the risk of human error in interpretation.

**1.3 Statement of the Problem**

In the field of lung cancer diagnostics, the primary goal is to ascertain early diagnosis without conducting invasive tests. However, this remains a major obstacle. Though CT, MRI, and X-ray have promoted diagnostics, they are solely dependent on human interpretation which is laborious and prone to human error.

Modern solutions, involving the use of machine learning, have been proposed to enhance to improve the diagnostic process, however, these systems face several issues, such as limited generalization capabilities, dependence on high-quality imaging, and data imbalance, where there is a higher number of negative cases than positive. This wrongly influences the timely detection of lung cancer and leaves high false positive/negative rates. Moreover, the requirement for high-end hardware restricts access to many of these systems in numerous healthcare settings.

In light of the issues above, there is a clear need for a lung cancer detection system that would provide more precision while supporting real-time implementation working efficiently on low-end computers so that it can be availed to a greater number of hospitals.

**1.4 Methodology**

The methodology for the lung cancer detection system involves several essential steps. The Agile software development model provides for flexibility as well as continuous improvement and thus we will divide our project into several sprints making use of this model. Patient information such as CT scan images among other related clinical information will be assembled and utilized by the system in question. Three categories of these CT images are normal cases, benign cases and malignant cases. Therefore, data will be separated into training, validation and testing sets for the sake of making a detection model. In order to make it more accurate and efficient an appropriate algorithm that could detect lung cancer will be selected then trained on datasets that were prepared earlier. A user-friendly interface will be designed to allow medical personnel to upload images, submit patients’ details and receive diagnostic results quickly without spending too much time.

**1.5 Aim and Objectives**

This project aims to develop a modern, accurate, and efficient method for detecting lung cancer. The system will identify and classify lung cancer, there aiding in early detection and improving patient management.

The specific objectives are:

i. To design a lung cancer detection system.

ii. To implement and test the design.

iii. To evaluate the performance of the system using accuracy, and speed.

**1.6 Justification of the Study**

Transitioning to an advanced system for the detection of lung cancer is necessary based on a number of urgent problems evident in the previous approaches. At the core of traditional and modern techniques are such pitfalls as data imbalance, high false positives, and negatives, and the high cost of the involved hardware. A new system will thus be justified by identifying the following advantages:

1. The multiclass classification system based on the test for correct classification of the lung condition into three classes: benign or non-cancerous, malignant or cancerous, or normal.  
2. Early detection: The system's clearer operations will quicken the diagnostic process for timely and accurate detection which will help doctors to determine the appropriate therapy.  
3. Better Accuracy: The 'oversampling' technology gets better after being tuned by different class weights which is in turn provides an accuracy of about 90%. The latter technology is so decisive that it distinguishes normal from malignant tumors, thus the system gives a much more accurate diagnosis of lung cancer.

4. Low Cost: The proposed system will provide economic efficiency on low-end hardware so as to enable quality lung cancer detection even in resource-poor institutions.  
5. Low Error Rates: Consequently, there will be minimal errors in diagnosis, translating to better patient outcomes.  
6. User-Friendly Interface: Comprehensive and understandable user interface of the results and probabilities predicted to be understood by the health care provider for better decision-making.  
7. Clinical Efficiency: the system will be designed in a manner to ensure that it can easily be integrated into different healthcare environments that may have limited access to advanced technology with the objective of allowing the healthcare professionals' further contributions to overall technological advancement in the health sector.

Implementation of the advanced lung cancer detection system tries to solve with significant advantages to the health institutes and an improvement in patient care and disease diagnosis

**1.7 Significance of the Study**

The significance of this study lies in its potential to revolutionize lung cancer detection in medical settings. By providing an efficient, effective, easy-to-use, and accessible system, much of the time and resources used now by manuals in diagnostic processes will be saved; in addition, the implementation of techniques for dealing with data imbalances underlines the commitment to practicality and broader adoption. This system could serve as a model for similar applications in various healthcare institutions, and open the gateway to broader acceptance of artificial intelligence applications in medical diagnosis.

**1.8 Operational Terms**

1. Lung Cancer: A type of cancer that originates in the lungs, formed through the ungoverned division of cells, leading to the generation of tumors, which can metastasize to other parts of the body.  
2. Lung Cancer Detection System: A software program or platform meant for the detection and classification of lung cancer from imaging data for early diagnosis and better management of the patients.  
3. Prognosis: The forecasted or probable level of recovery that can be expected in a health condition together with the time expected to reach the particular level of recovery.  
4. Imbalanced Data: A situation where, within a dataset, cases are distributed unequally; some classes, such as non-cancerous ones, are more frequent than others, like early-stage cancer, therefore leading to biased model performance.  
5. Machine Learning: A subfield of artificial intelligence that incorporates algorithms which are trained on datasets for the generation of models that may perform tasks that range from the categorization of images to analysis of data or predicting outcomes.  
6. Model: A mathematical representation developed, typically using machine learning algorithms, trained on patterns or predictions that the input data should make.  
7. Classification: A category of supervised machine learning in which, for given input data, the model predicts the correct label. The model is first trained with labeled data and tested on test data to make predictions on new unseen data.  
8. Prediction: The output of a machine learning algorithm that has been trained on historic data to make an estimate or forecast for the future based on new data.

**1.9 Organization of Chapters**

**Chapter 1**: Introduction: This chapter establishes the framework for the thesis by providing background information on the subject, presenting the research questions or hypotheses, and emphasizing the significance of the study.

**Chapter 2**: Literature Review: Provides a broad overview analysis of current works in the paradigm, reviews relevant literature on lung cancer detection methods, discusses the limitations of existing proposed methods, and outlines the benefits of advanced techniques.

**Chapter 3**: Methodology: Explores the systematic methodologies employed in developing the lung cancer detection system. Discusses the chosen approach, system analysis intricacies, and design principles, serving as a blueprint for implementation.

**Chapter 4**: System Implementation and Testing: When putting the suggested system into practice, this is the phase where software development, testing, and deployment constraints are given top attention.

**Chapter 5**: Summary, Conclusion, and Recommendations: This chapter brings together many of the points that have been used in previous chapters for a generalized conclusion. It summarizes the findings of the study, discusses the implications for healthcare, and offers recommendations for future research and system improvements.

**REFERENCES**

1. Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *Journal of Infection and Public Health*, *13*(9), 1274–1289.
2. Alsheikhy, A. A., Said, Y., Shawly, T., Alzahrani, A. K., & Lahza, H. (2023). A CAD System for Lung Cancer Detection Using Hybrid Deep Learning Techniques. *Diagnostics (Basel, Switzerland)*, *13*(6), 1174.
3. Giger, M. L., Doi, K., MacMahon, H., Metz, C. E., & Yin, F. F. (1990). Pulmonary nodules: computer-aided detection in digital chest images. *RadioGraphics*, *10*(1), 41–51.
4. Armato, S. G., Giger, M. L., Moran, C. J., Blackburn, J. T., Doi, K., & MacMahon, H. (1999). Computerized Detection of Pulmonary Nodules on CT scans. *RadioGraphics*, *19*(5), 1303–1311.
5. Nitha, V. R., & Chandra. (2023). ExtRanFS: An Automated Lung Cancer Malignancy Detection System Using Extremely Randomized Feature Selector. *Diagnostics*, *13*(13), 2206–2206.
6. Reddy, N. S., & Khanaa, V. (2023). Intelligent deep learning algorithm for lung cancer detection and classification. *Bulletin of Electrical Engineering and Informatics*, *12*(3), 1747–1754.