International Journal of Intelligent Communication and Computer Science
Vol. 1, No. 2, 2023, Pages 19-36.

Detection of Lungs Disease Using Deep Learning Approach: A Systematic Review

Sandhya Devi [1] and Pallavi Khatri[1]

Review Paper

Department of Computer Science and Engineering, ITM University, Gwalior, INDIA
Email: sandhyadeviphd@gmail.com
Received: 2 Oct 2023, Revised: 11 Nov 2023, Accepted: 14 Nov 2023

Abstract:

This comprehensive review paper investigates the evolving landscape of deep learning methodologies applied to the detection of four critical medical conditions: pneumonia, COVID-19, cancer, and joint diseases, utilizing various imaging modalities such as X-rays and CT scans. We systematically categorize and analyze recent advancements in deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models, tailored to address the unique challenges posed by each disease category. The survey encompasses the development of robust models capable of distinguishing between normal and pathological conditions. For pneumonia, we explore the nuances of deep learning applications in pneumonia detection, addressing challenges such as subtle patterns and class imbalance. In the context of COVID-19, we review methodologies specifically designed for the timely and accurate identification of the virus from medical imaging, contributing to the global efforts in pandemic management. Additionally, the review delves into the application of deep learning in cancer detection, covering a spectrum of malignancies including lung cancer. We analyze the potential of deep neural networks (DNNs) in aiding early diagnosis, prognosis, and treatment planning for cancer patients. Furthermore, we investigate the role of DNNs in joint disease detection, addressing conditions such as arthritis. The review encompasses the challenges unique to joint disease imaging, and the adaptation of deep learning models to effectively discern abnormalities in musculoskeletal structures. By synthesizing the current state of knowledge across these diverse medical domains, this review aims to provide a valuable resource for researchers, clinicians, and policymakers.

Keywords: Lung disease, COVID-19, CNN, DNN, Cancer, Pneumonia

1. Introduction

Lung diseases represent a significant global health challenge, encompassing a diverse range of disorders with multifaceted origins. This paper delves into the epidemiology and classification of these conditions, emphasizing their far-reaching impact on public health and healthcare systems. Traditional diagnostic methods, including spirometry and imaging, face limitations in accurately capturing the complexities of lung diseases [1]. Consequently, the emergence of innovative technologies, such as medical imaging advancements, molecular diagnostics, and wearable technologies, has paved the way for more sophisticated diagnostic approaches [2].

Deep learning plays a pivotal role in advancing the prediction of lung diseases through a comprehensive and multifaceted approach that integrates medical imaging and clinical data. The foundation of this process lies in the meticulous collection of a diverse dataset encompassing various lung conditions and normal cases [3]. This dataset serves as the bedrock for training a robust deep learning model. The initial step involves thorough preprocessing of the data, a crucial aspect that contributes to the model's accuracy. This includes tasks such as image resizing to standardize dimensions, normalization to ensure consistent data ranges, and

clinical data cleansing to remove inconsistencies. The goal is to create a high-quality dataset that is suitable for the deep learning model's training. The selection of appropriate model architecture is a critical decision in this process. CNNs are often employed for image analysis, enabling the model to extract intricate patterns and features from medical images [4]. For processing clinical data, Recurrent Neural Networks (RNNs) or Transformer-based models are commonly integrated, allowing the model to capture temporal dependencies and complex relationships within the data [4]. Hybrid models, which combine both image and clinical data, have gained popularity for their potential to enhance predictive accuracy by leveraging the strengths of both types of information. The training phase involves using a portion of the dataset to teach the model to recognize patterns and make accurate predictions. To ensure the model's generalizability, it undergoes evaluation using validation sets. Fine-tuning is then applied to prevent overfitting, a process crucial for achieving optimal performance in real-world scenarios.
Performance metrics such as accuracy, precision, and recall are employed to assess the proficiency of the model. Interpretability techniques are also crucial for shedding light on the factors influencing the model's predictions, providing transparency and aiding healthcare professionals in understanding and trusting the model's outputs. Once rigorously evaluated, the model can be deployed in clinical settings, offering valuable support to healthcare professionals in the early diagnosis and prognosis of lung diseases. However, ethical, privacy, and regulatory considerations must remain paramount throughout the entire process to ensure patient confidentiality and data security. Continuous improvement is achieved through periodic updates, incorporating new data to enhance the model's relevance and efficacy in healthcare applications. This iterative approach ensures that the model stays current with advancements in medical knowledge and technology.

2. Lungs Diseases

Lung Cancer: Lung Cancer, a formidable malignancy originating in the lung tissue cells, is intricately linked to prolonged exposure to cigarette smoke and other carcinogens. The two primary classifications are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), each posing distinct challenges in diagnosis and treatment [5]. The disease can manifest in any part of the lung, potentially spreading to neighboring tissues and organs, underscoring the importance of early detection. Recognizable symptoms include persistent coughing, chest pain, shortness of breath, and hemoptysis. Timely intervention and treatment are pivotal in improving patient outcomes, highlighting the critical role of healthcare professionals in identifying and managing lung cancer cases.

Pneumonia: Pneumonia, a prevalent respiratory ailment, manifests as an inflammatory lung condition primarily triggered by various infections, including bacteria, viruses, fungi, or parasites [6]. This widespread affliction disrupts the normal physiological functioning of the lungs, as it involves the infiltration of the air sacs with pus or other fluids. This intrusion impedes the essential process of oxygenation, compromising the efficient exchange of oxygen and carbon dioxide in the bloodstream. The symptoms of pneumonia exhibit considerable variability, ranging from mild to severe, and can include fever, cough, chest pain, and difficulty breathing. The severity of these symptoms is contingent upon factors such as the specific causative agent and the individual's overall health. In severe cases, pneumonia can lead to significant respiratory distress and complications, underscoring the critical importance of accurate diagnosis and timely intervention. The diverse etiology of pneumonia underscores the need for tailored treatment approaches, as different causative agents may respond differently to specific therapeutic interventions. Given the variability in symptom presentation and the array of potential pathogens, accurate and prompt diagnosis becomes paramount for effective management. This necessitates a comprehensive understanding of the patient's medical history, risk factors, and potential exposure to infectious agents.

COVID-19: COVID-19 Pneumonia represents a distinctive manifestation, caused by the novel coronavirus SARS-CoV-2. This virus, with a predilection for the respiratory system, can lead to severe pneumonia in certain cases [6]. Symptomatically akin to other types of pneumonia, COVID-19 pneumonia has unique features such as the loss of taste and smell, and it can escalate to severe respiratory distress. Mitigating the spread of COVID-19 and its associated pneumonia has been a global priority, with vaccination campaigns and stringent hygiene practices proving essential. The proactive approach to prevention has played a pivotal role in reducing both the incidence of COVID-19 and the severity of associated pneumonia, underscoring the significance of public health measures in mitigating the impact of infectious diseases.

3. State of the Art Methods Details:

The current methodology relies on the utilization of lung CT scans as the initial source images. These CT scans serve as the foundation upon which a series of image preprocessing techniques are applied (Figure 1). During this preprocessing phase, various operations and transformations are performed to enhance the quality and suitability of the images for subsequent analysis. This crucial step aids in the removal of noise, the enhancement of relevant features, and the standardization of image characteristics.
Following the image preprocessing phase, the next critical step is the extraction of essential features from the now-enhanced lung images. This feature extraction process involves identifying and isolating key patterns, structures, and characteristics within the images that are indicative of the underlying health conditions. These features may include texture patterns, lesion shapes, density variations, and other visual cues that can help distinguish between different medical conditions.

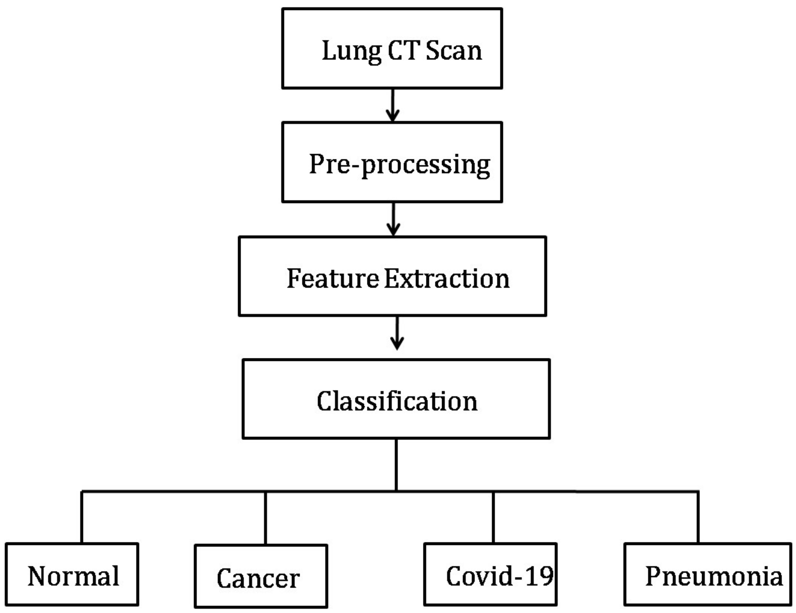

Figure 1: Existing Lungs disease prediction methods

Once the relevant features have been successfully extracted, the classification phase comes into play. In this stage, a variety of machine learning and artificial intelligence techniques are employed to classify the lung images into distinct categories. The primary categories of interest typically encompass normal lung conditions, lung cancer, COVID-19 infection, and pneumonia.
Classification methods may encompass a diverse range of algorithms such as deep neural networks (DNN), support vector machines (SVM), decision trees (DT), or ensemble methods. These algorithms are trained on a labeled dataset that includes examples of each condition, enabling them to learn the patterns and relationships between the extracted features and the corresponding medical diagnoses. Through this training process, the classification models become proficient at accurately categorizing new, unseen lung images based on the extracted features.

4. Distinguishing between Lung Cancer, Pneumonia, and COVID-19:

Distinguishing between lung cancer, pneumonia, and COVID-19 through image processing and DNNs involves tailoring the analytical approach to the unique characteristics associated with each medical condition (Figure 2). In the case of lung cancer, the focus lies on extracting features related to tumor size, shape, and texture, necessitating the utilization of DNN architectures capable of capturing intricate patterns indicative of malignancies, such as U-Net or ResNet. For pneumonia, the emphasis shifts towards identifying infiltrates, consolidations, and opacities in the affected lung areas. Appropriate CNN architectures with attention mechanisms may be employed to recognize characteristic pneumonia patterns. COVID-19, being a viral pneumonia, requires specialized consideration for features like ground-glass opacities, bilateral involvement, and peripheral distribution patterns. Model design should encompass the evolving nature of COVID-19 cases, possibly incorporating temporal information to track the progression of the disease. Integrating clinical data, such as patient history and symptomatology, enhances diagnostic accuracy across all conditions. Continuous validation on diverse datasets, encompassing various manifestations and stages, is pivotal for developing robust and accurate DNN models capable of effectively differentiating between lung cancer, pneumonia, and COVID-19 from medical images.

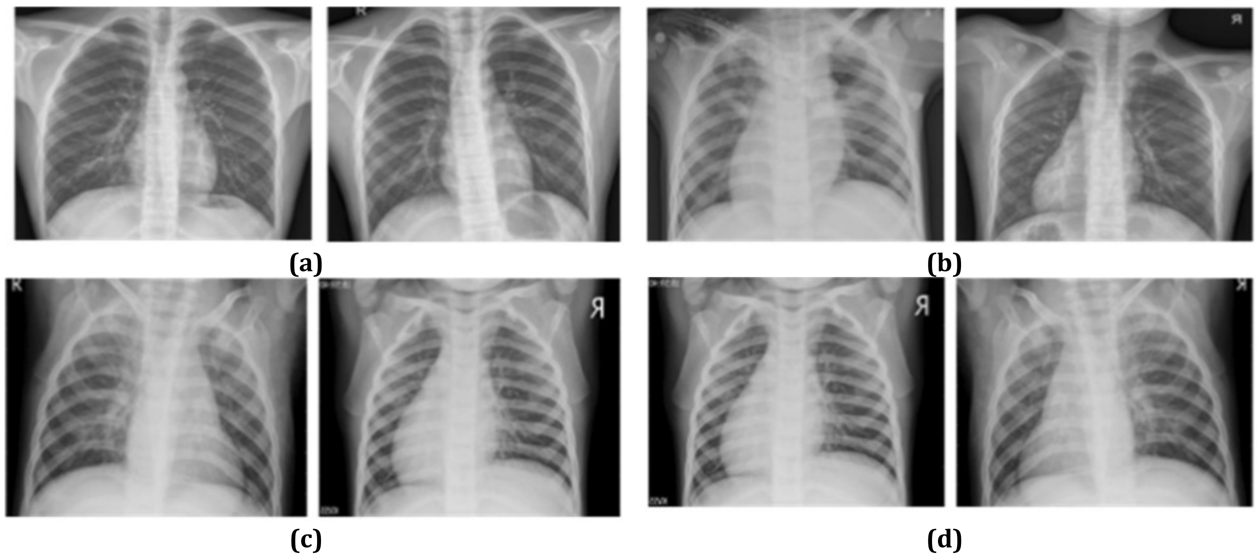

Figure 2: (a) Normal (b) COVID -19 (c) Pneumonia (d) Cancer

3. Notable Research Works

In this dedicated segment, our exploration ventures into the extensive body of existing literature focused on early diagnosis and detection studies related to the selected lung diseases within the scope of our research. This comprehensive review aims to synthesize and critically analyze the wealth of knowledge available, shedding light on the various approaches undertaken for the timely identification of pneumonia, COVID-19, and lung cancer.

The subsequent sections of our review are meticulously designed to unravel distinct machine learning and deep learning methodologies employed in the context of each specific lung disease. In Table 1, we delve into the nuances of methodologies for pneumonia detection, systematically presenting and analyzing advancements in the field. This section emphasizes the diverse approaches undertaken to address challenges such as subtle patterns and class imbalance in pneumonia diagnosis.

Moving forward, Table 2 is dedicated to exploring methodologies tailored for the detection of COVID-19 in medical imaging. This section delves into the unique challenges posed by the timely and accurate identification of the virus, contributing crucial insights to the global efforts in managing the ongoing pandemic. The methodologies discussed encompass a range of deep learning and machine learning techniques optimized for COVID-19 detection from chest X-rays and other relevant imaging modalities.

Table 3 further extends our exploration into the realm of lung cancer detection, covering a spectrum of malignancies. This section analyzes the potential of deep neural networks (DNNs) and other advanced models in aiding early diagnosis, prognosis, and treatment planning for patients with various forms of lung cancer. The methodologies discussed here highlight the evolving landscape of lung cancer detection through medical imaging.

Recognizing the complexity that arises when multiple diseases are considered simultaneously, the final sub-section (Table 4) is exclusively dedicated to investigations encompassing more than one of the mentioned diseases. This inclusive analysis aims to provide a holistic perspective on integrated diagnostic approaches, shedding light on the challenges and opportunities associated with concurrently addressing pneumonia, COVID-19, and lung cancer.

3.1 Pneumonia

Hashmi et al. [7] introduced an effective model designed for the detection of pneumonia, utilizing chest X-ray images. This innovative model incorporated a sophisticated weighted classifier that seamlessly amalgamated predictions from a diverse array of models, such as ResNet18, Xception, and DenseNet121. Leveraging the power of transfer learning, the model underwent meticulous fine-tuning, culminating in an outstanding test accuracy of 98.43% when applied to previously unseen data. This approach not only underscores the significance of ensemble techniques in pneumonia detection but also highlights the efficacy of transfer learning in enhancing the model's performance on novel datasets, showcasing its potential for robust and accurate diagnostic outcomes.

Table 1: Summary of Pneumonia detection methods

Study	Methodology	Model(s) Used	Dataset	Results
Hashmi et al. [7]	Efficient model for pneumonia detection using chest X-ray images. Weighted classifier combining ResNet18, Xception, DenseNet121, etc. Transfer learning for fine-tuning.	ResNet18, Xception, DenseNet121, etc.	Chest X-ray images	Accuracy = 98.43%
Stephen et al. [8]	Pneumonia detection in chest X-ray images using a CNN. Extensive parameter and hyperparameter tuning.	CNN	Chest X-ray images	Accuracy = 93.73%.
ElShennawy et al. [9]	CNN model for pneumonia detection using chest X-ray images. ResNet152V2 and MobileNetV2 as extraction models.	ResNet152V2, MobileNetV2	Chest X-ray images	Accuracy = 99.22%.
Szepesi et al. [10]	Transfer learning models (InceptionV3, ResNet50) for pneumonia detection in children's chest X-ray images.	InceptionV3, ResNet50	Children's chest X-ray images	ResNet50: 89.06% accuracy, InceptionV3: 90.94% accuracy.
Qaimkhani et al. [11]	Deep learning technique for early-stage lung disease identification using CNN.	ANN, CNN, VGG19	Chest X-ray images	Accuracy = 97%.
Abubeker et al. [12]	Dense CNN-160, ResNet-121, and VGG-16 ensemble models.	Dense CNN-160, ResNet-121, VGG-16	Chest X-ray images	Accuracy of 97.69%, 100% recall, and 0.9977 .

Stephen et al. [8] crafted a pneumonia detection model tailored for chest X-ray images, employing the prowess of a CNN. In a dedicated effort to optimize the model's performance, the researchers conducted a thorough tuning of parameters and hyperparameters. The meticulous tuning process resulted in compelling outcomes, showcasing a a validation accuracy of 93.74%. This attention to parameter fine-tuning underscores the commitment to achieving high accuracy rates in both training and validation phases, affirming the model's effectiveness in discerning pneumonia patterns from chest X-ray imagery.

ElShennawy et al. [9]] contributed to the field of medical imaging by developing a CNNmodel specifically tailored for the detection of pneumonia through the analysis of chest X-ray images. In their innovative approach, they incorporated two powerful extraction models, namely ResNet152V2 and MobileNetV2, to enhance the feature extraction capabilities of their CNN architecture. ResNet152V2 and MobileNetV2 are renowned architectures in the realm of deep learning, known for their ability to capture intricate patterns and hierarchical representations in images. By leveraging these sophisticated models, ElShennawy et al. aimed to enhance the discerning power of their CNN in identifying subtle nuances indicative of pneumonia within chest X-ray images. The CNN model devised by ElShennawy and team exhibited outstanding performance, achieving an impressive accuracy rate of 99.22%. This high accuracy underscores the robustness and efficacy of their model in accurately distinguishing between normal and pneumonia-affected chest X-ray images. The utilization of state-of-the-art extraction models, coupled with meticulous design and training, contributed to the exceptional success of their CNN in the challenging task of pneumonia detection.

Szepesi et al. [10] delved into the realm of pediatric radiology with a specific focus on detecting pneumonia in chest X-ray images of children. In their study, the researchers opted for a transfer learning approach, leveraging well-established models such as InceptionV3 and ResNet50 to enhance the performance of their pneumonia detection system. Transfer learning involves utilizing pre-trained models on large datasets and fine-tuning them for a specific task, which is particularly valuable when dealing with limited data, as is often the case in medical imaging. In this context, InceptionV3 and ResNet50 served as the foundation for feature extraction in Szepesi et al.'s approach. The outcomes of their study revealed noteworthy results. ResNet50, a popular convolutional neural network architecture, demonstrated a commendable accuracy of 89.06% in pneumonia detection. However, the study's standout performer was InceptionV3, which outperformed ResNet50 by achieving a higher accuracy rate of 90.94%.

Qaimkhani et al. [11] made significant strides in the realm of early-stage identification of lung diseases by introducing a deep learning approach, particularly focused on the analysis of chest X-rays. The objective of their study was to leverage the power of deep neural networks to detect and classify lung diseases in their nascent stages, offering the potential for timely and accurate diagnosis. In their innovative approach, Qaimkhani et al. employed pre-trained CNN models as feature extractors. The use of pre-trained models implies that these networks had been previously trained on large datasets, enhancing their ability to capture relevant features in chest X-ray images. Furthermore, the study incorporated various classifiers, including ANN, CNN, and the VGG19 model. VGG19, a variant of the VGG (Visual Geometry Group) architecture, has proven to be effective in image classification tasks. In this study, VGG19 emerged as the most successful classifier, achieving the highest accuracy rate of 97%.

Abubeker et al. [12] presented a notable contribution for pneumonia detection in chest X-ray images. Their innovative framework integrated an ensemble of dense CNN with 160 layers (dense CNN-160), ResNet-121, and the well-established VGG-16 model. This ensemble approach was engineered to leverage the strengths of multiple architectures in order to enhance the system's diagnostic capabilities. The performance metrics reported by Abubeker et al. speak to the robustness of their framework. The system achieved an outstanding accuracy of 97.69%, indicating its high. Notably, the perfect recall rate of 100% signifies the framework's ability to correctly identify all instances of pneumonia, minimizing false negatives and showcasing its reliability. The AUC was reported as 0.9977. This high AUC score indicates a strong ability of the model to discriminate between different classes, further confirming the effectiveness of the proposed framework.

3.2 COVID-19

Ramadhan et al. [13] proposed a VGG16-CNN model with the primary objective of classifying chest X-ray images from three distinct public datasets containing COVID-19 patient data. In their study, binary classification was performed on each dataset, and the results were particularly noteworthy. The VGG16-CNN model achieved impressive accuracies of 97%, 98.73%, and 99.76% for the first, second, and third datasets, respectively. This suggests a high level of proficiency in distinguishing COVID-19 cases from non-COVID-19 cases across different datasets.

Jain et al. [14] conducted a comprehensive study comparing the performance of various CNN models—Xception, InceptionV3, and ResNeXt—for the specific task of detecting COVID-19 in chest X-ray images. Among the models assessed, the Xception model stood out, demonstrating the highest recognition accuracy at 97.97%. This underscores the effectiveness of the Xception architecture in accurately identifying COVID-19 cases from X-ray images.

Hussain et al. [15] contributed to the field by developing the CoroDet CNN model, designed to classify COVID-19 cases at different levels, including binary, three-class, and four-class classification. Leveraging a dataset comprised of chest X-rays from COVID-19 patients, the CoroDet model achieved remarkable accuracies, attaining 99.1% in binary classification, 92.4% in three-class classification, and 91.2% in four-class classification. This versatility in classification suggests its potential utility across various diagnostic scenarios.

Nayak et al. [16] explored the application of deep learning-based CNN models, including GoogleNet, SqueezeNet, and VGG-16, for the detection of COVID-19 in X-ray images. Their experiments yielded notable results, with GoogleNet achieving 98.62% accuracy, VGG-16 reaching 96.15% accuracy, and AlexNet surpassing others with an accuracy of 99.05%. This comparative analysis provides valuable insights into the strengths of different CNN architectures in the context of COVID-19 detection.

Ahmed et al. [19] introduced a FRBS designed to facilitate the prompt identification of COVID-19 through the analysis of clinical data. Their innovative approach yielded a commendable prediction accuracy of 88.78%, highlighting the efficacy of employing fuzzy logic-based systems in the realm of healthcare. This impressive accuracy underscores the system's capability to harness intricate clinical information, offering a promising avenue for the timely and precise detection of individuals afflicted with COVID-19. The utilization of fuzzy logic in this context represents a significant advancement in medical technology, as it enables the incorporation of imprecise and uncertain information inherent in clinical datasets. By leveraging fuzzy logic, the FRBS developed by Ahmed et al. demonstrates a nuanced understanding of the complex relationships within clinical data, contributing to its high predictive accuracy. This robust performance holds considerable implications for public health, offering a potential solution for the early identification of COVID-19 cases. The success of the FRBS in achieving an accuracy of 88.78% suggests its viability as a valuable tool in the broader landscape of infectious disease detection.

Naqvi et al. [20] conducted a comprehensive study that assessed the widespread impact of COVID-19 on a global scale. Their investigation covered diverse aspects such as healthcare, public safety, economics, industry, and travel restrictions, providing a holistic understanding of the multifaceted consequences of the pandemic.

Nasiri and Hasani [21] proposed a deep learning approach combining the DenseNet169 deep neural network (DNN) with the extreme gradient boosting (XGBoost) algorithm for precise classification of X-ray images as COVID-19 positive. The binary classification outperformed ternary classification, achieving a notable 98.23% accuracy, 99.78% specificity, and 92.08% sensitivity in binary classification. In ternary classification, the model achieved 89.70% accuracy, 100% specificity, and 95.20% sensitivity. This demonstrates the efficacy of their hybrid approach in accurately identifying COVID-19 cases from X-ray images.

Khan et al. [22] introduced a novel approach called STM-RENet, a channel-boosted CNN for COVID-19 detection from chest X-ray images. The scheme incorporated split–transform–merge (STM) with region and edge-based (RE) operations, achieving a maximum accuracy of 96.53% and an F-score of 95% across three datasets. This innovative CNN architecture, combining spatial and spectral operations, showcases promising results in accurately identifying COVID-19 cases in chest X-ray images.

Table 2: Summary of COVID-19 detection methods

Study	Methodology	Model(s) Used	Dataset	Results
Ramadhan et al. [13]	VGG16-CNN for COVID-19 classification in chest X-ray images. Binary classification on three datasets.	VGG16	Chest X-ray images of COVID-19 patients	Dataset 1: 97% accuracy, Dataset 2: 98.73% accuracy, Dataset 3: 99.76% accuracy.
Jain et al. [14]	Comparison of CNN models (Xception, InceptionV3, ResNeXt) for COVID-19 detection in chest X-ray images.	Xception, InceptionV3, ResNeXt	Chest X-ray images	Xception: 97.97% accuracy.
Hussain et al. [15]	CNN model "CoroDet" for COVID-19 classification with 2, 3, and 4 class levels.	CoroDet	Chest X-rays of COVID-19 patients	Binary: 99.1% accuracy, Three-class: 92.4% accuracy, Four-class: 91.2% accuracy.
Nayak et al. [16]	CNN models (GoogleNet, SqueezeNet, VGG-16) for COVID-19 detection in X-ray images.	GoogleNet, SqueezeNet, VGG-16	X-ray images of COVID-19 patients	GoogleNet: 98.62% accuracy, VGG-16: 96.15% accuracy, AlexNet: 99.05% accuracy.
Zagrouba et al. [17]	SVM for COVID-19 outbreak modelling.	SVM	COVID-19 outbreak data	Training: 98.88% accuracy, Testing: 96.79% accuracy.
Rahman et al. [18]	Supervised machine learning-based predictive model for COVID-19 outbreak.	SVM	COVID-19 outbreak data	Validation: 98.4% accuracy.
Ahmed et al. [19]	Fuzzy rule-based system (FRBS) for early identification of COVID-19 using clinical data.	FRBS	Clinical data	Accuracy: 88.78%, Precision: 72.22%, Sensitivity: 68.42%, Specificity: 93.67%, F1-Score: 69.28%.
Naqvi et al. [20]	Comprehensive study on the impact of COVID-19 globally.	N/A	N/A	N/A
Nasiri and Hasani [21]	Deep learning (DenseNet169 DNN) with XGBoost for COVID-19 classification in X-ray images.	DenseNet169 DNN + XGBoost	X-ray images	Binary: 98.23% accuracy, 99.78% specificity, 92.08% sensitivity; Ternary: 89.70% accuracy, 100% specificity, 95.20% sensitivity.
Khan et al. [22]	Channel boosted CNN for COVID-19 detection from chest X-ray images.	STM-RENet	Chest X-ray images	Accuracy: 96.53%, F-score: 95%.

3.3 Cancer

Alsinglawi et al. [23] conducted a study focused on predicting lengths of stay for lung cancer patients using machine learning algorithms. Their analysis, carried out on the MIMIC-III dataset, involved Logistic Regression, Random Forest, and XGBoost. Notably, Random Forest, when combined with Synthetic Minority Over-sampling Technique (SMOTE), demonstrated exceptional performance with an AUC score of 98%, signifying a high level of accuracy in forecasting the lengths of stay for individuals with lung cancer.

Venkatesh et al. [24] explored the prediction accuracy of various algorithms in forecasting lung cancer cases, utilizing the SEER dataset. The algorithms employed included K-Nearest Neighbors, Decision Trees, and Adaboost. Their findings highlighted Adaboost as the most effective, achieving a remarkable prediction accuracy of 98.2%, showcasing its robust performance in anticipating instances of lung cancer within the SEER dataset.

Table 3: Summary of Cancer detection methods

Study	Methods/Classifiers Used	Dataset	Key Findings
Alsinglawi et al. [23]	Logistic Regression (LR), RF, XGBoost	MIMIC-III	Random Forest with SMOTE achieved an outstanding AUC score of 98%, demonstrating high accuracy in predicting lengths of stay for lung cancer patients.
Venkatesh et al. [24]	K-Nearest Neighbors, DT, Adaboost	SEER dataset	Adaboost displayed the highest prediction accuracy of 98.2% in forecasting lung cancer cases in the SEER dataset.
Vikas et al. [25]	SVM, RF	Not specified	Support Vector Machine demonstrated superior computational efficiency and prediction performance with impressive metrics, including 98% accuracy.
Puneet et al. [26]	XGBoost, GridSearchCV, Logistic Regression, SVM, and more	Lanzhou University dataset	XGBoost outperformed other classifiers, achieving a prediction accuracy rate of 92.16%.
Sim et al. [27]	DT, Logistic Regression, Bagging, RF, AdaBoost	Clinical data from lung cancer surgeries	AdaBoost and Random Forest emerged as top models, with AdaBoost achieving an accuracy rate of 94.8% and an AUC value of 94.9%.
Patra [28]	Radial Basis Function Network, SVM, ANN, RF	UCI repository	RBF exhibited the highest accuracy rate of 81.25% and excelled in precision, recall, and F1-score.
P.R. et al. [29]	Naive Bayes, SVM, DT, LR	Dataset from Data World	Support Vector Machine demonstrated the highest accuracy rate of 99.2%, showcasing its precision in early lung cancer diagnosis.
Wu et al. [30]	RF	Lanzhou University dataset	Random Forest achieved a notable accuracy rate of 95.7%, showcasing its efficiency in distinguishing and identifying lung cancer cases.
Faisal et al. [31]	Naive Bayes, Neural Network, SVM, Gradient Boosted Tree, RF	UCI repository	Gradient Boosted Tree excelled among ensemble methods, achieving a remarkable accuracy rate of 90% and high precision, recall, and F1-score.
Safiyari et al. [32]	Bagging, Dagging, AdaBoost, MultiBoosting,	SEER dataset	AdaBoost stood out as the top performer, achieving an AUC of 94.9% and an accuracy rate of 88.98% in predicting lung cancer survival rates.

Vikas et al. [25] delved into the realm of computational efficiency and prediction performance in the context of lung cancer. Although the specific dataset was not specified, their study involved the application of SVM and RF. Impressively, Support Vector Machine stood out, demonstrating superior computational efficiency and prediction accuracy, achieving a noteworthy rate of 98%.

Puneet et al. [26] focused on predictive modeling for lung cancer using a variety of algorithms such as XGBoost, GridSearchCV, Logistic Regression, and SVM. Their study utilized the Lanzhou University dataset, and among the classifiers, XGBoost emerged as the top performer with a notable prediction accuracy rate of 92.16%.

Sim et al. [27] undertook a comprehensive analysis utilizing clinical data derived from lung cancer surgeries, aiming to discern predictive models for surgical outcomes. Their approach involved the application of various machine learning algorithms, including DT, LR, Bagging, RF, and AdaBoost, to scrutinize the dataset and identify robust models for outcome prediction. Among the diverse algorithms employed, the results highlighted AdaBoost and Random Forest as the standout performers in terms of predictive accuracy. AdaBoost, in particular, demonstrated an impressive accuracy rate of 94.8%. Accuracy is a crucial metric that reflects the model's ability to make correct predictions, and a high accuracy rate suggests the effectiveness of the model in capturing patterns within the clinical data. Additionally, the reported AUC value of 94.9% for AdaBoost is indicative of its strong discriminatory power. AUC is a widely used metric in binary classification tasks, reflecting the model's ability to distinguish between positive and negative instances. In this context, the high AUC value reinforces the robustness of AdaBoost in predicting outcomes related to lung cancer surgeries. The findings suggest that ensemble learning methods, such as AdaBoost and Random Forest, prove to be particularly effective in handling the complexity of clinical data associated with lung cancer surgeries. The ability of these models to integrate information from multiple sources and provide accurate predictions underscores their potential utility in supporting clinical decision-making.

Patra [28] focused on predicting lung cancer using the Radial Basis Function Network (RBF) and various other algorithms, leveraging data from the UCI repository. Their investigation aimed to assess the performance of different algorithms in predicting lung cancer outcomes, with a specific emphasis on accuracy and other key metrics. Among the algorithms explored, the RBF stood out as the top performer, achieving the highest accuracy rate of 81.25%. Accuracy is a fundamental metric that indicates the model's overall correctness in predicting outcomes. The statement suggests that RBF excelled in these metrics, further emphasizing its effectiveness in predicting lung cancer outcomes with a high degree of precision and recall. This research contributes valuable insights into the potential of RBF and other algorithms for accurate and reliable predictions in the challenging domain of lung cancer outcomes.

P.R. et al. [29] delved into early lung cancer diagnosis, utilizing algorithms such as Naive Bayes, SVM, DT, and LR on a dataset from Data World. Their findings highlighted Support Vector Machine as the most precise, achieving the highest accuracy rate of 99.2% in the context of early lung cancer detection.

Wu et al. [30] focused on distinguishing and identifying lung cancer cases using Random Forest and the Lanzhou University dataset. Their study demonstrated the efficiency of Random Forest, achieving a notable accuracy rate of 95.7%, underscoring its effectiveness in the classification task related to lung cancer.

Faisal et al. [31] explored lung cancer prediction through ensemble methods, utilizing algorithms like Naive Bayes, Neural Network, SVM, Gradient Boosted Tree, and Random Forest on the UCI repository. Their study revealed that Gradient Boosted Tree excelled among the ensemble methods, achieving a remarkable accuracy rate of 90%, accompanied by high precision, recall, and F1-score.

Safiyari et al. [32] investigated the prediction of lung cancer survival rates using ensemble methods on the SEER dataset. Among the various methods explored, AdaBoost emerged as the top performer, achieving an AUC of 94.9% and an accuracy rate of 88.98%, demonstrating its efficacy in predicting lung cancer survival outcomes.

3.4 Joint Lungs disease detection

Bhandari et al. [33] collaboratively conducted a study with the overarching goal of classifying chest X-ray images into potential cases of COVID-19, pneumonia, and tuberculosis (TB). Their innovative approach integrated deep learning techniques coupled with an Explainable Artificial Intelligence (XAI) framework, enhancing interpretability. The dataset employed for this study comprised a substantial 7132 chest X-ray images. Utilizing a robust 10-fold cross-validation methodology, the research yielded noteworthy results, showcasing an average test accuracy of 94.31% and a validation accuracy of 94.54% These findings underscore the efficacy of their approach in accurately categorizing diverse respiratory conditions.

Venkataramana et al. [34] proposed a sophisticated multi-level classification system that employed two distinct models. The first model was dedicated to binary classification for tuberculosis (TB) and pneumonia, while the second model focused on categorizing different types of pneumonia. The extensive dataset included 14,693 images, and the study demonstrated an accuracy of 95.7% before balancing, which notably increased to 96.6% post-application of the SMOTE technique. This approach showcases the significance of addressing dataset imbalances to enhance the accuracy of classification systems.

Hasan et al. [35] presented a dedicated model designed for the precise detection of pneumonia in the specific context of COVID-19 patients. Their approach incorporated Convolutional Neural Network (CNN) architecture, utilizing VGG16 both as a feature extractor and a classifier during the model training process. This strategic use of deep learning techniques aimed to capture intricate patterns in chest X-ray images associated with pneumonia in the context of COVID-19. During the implementation of their model, Hasan et al. leveraged various machine learning tools to enhance its performance. One notable tool employed was LabelBinarizer, a technique used for one-hot encoding labeled X-ray images. This process aids in transforming categorical labels into a binary matrix format, facilitating the training of the CNN model. The reported average accuracy of 91.69% attests to the model's proficiency in distinguishing between different classes of X-ray images, specifically those indicative of pneumonia in COVID-19 cases. Accuracy serves as a key metric in evaluating the model's overall correctness in its predictions. Furthermore, the model demonstrated a high sensitivity of 95.92% in predicting pneumonia. Sensitivity, also known as recall, signifies the model's ability to correctly identify positive instances, in this case, accurately detecting pneumonia cases within the COVID-19 context. The notable sensitivity emphasizes the effectiveness of the model in minimizing false negatives, a crucial aspect when dealing with medical diagnoses.

Ibrokhimov and Kang [36] made significant strides in the realm of medical imaging by pioneering the development of a deep learning-based diagnosis system specifically designed for detecting pneumonia using X-ray images. Their approach involved the application of transfer learning techniques, leveraging pre-trained models such as VGG19 and ResNet50. The dataset used in their study was extensive, encompassing 11,956 samples from COVID-19 cases, 11,263 samples from viral or bacterial pneumonia cases, and 10,701 samples from normal cases. This diverse and substantial dataset allowed for a comprehensive evaluation of the model's performance across different classes, reflecting real-world scenarios. In their comparative analysis, VGG19 emerged as the superior performer when compared to ResNet50. VGG19 exhibited an impressive average accuracy of 96.6% across all classes, showcasing its robustness in accurately classifying X-ray images into the categories of COVID-19, viral or bacterial pneumonia, and normal cases. This remarkable accuracy underscores the effectiveness of VGG19 in leveraging knowledge gained from pre-training on large datasets to enhance diagnostic capabilities in the challenging task of pneumonia detection.

Bashar et al. [37] introduced a cutting-edge deep learning approach aimed at diagnosing COVID-19 and pneumonia through the analysis of chest X-ray images. To achieve this, they harnessed the power of a substantial public dataset comprising 21,165 chest X-ray images, implementing a sophisticated methodology to enhance the diagnostic accuracy of their model. In their comprehensive approach, the researchers employed various image enhancement techniques and data augmentation methods to preprocess the dataset. Image enhancements are crucial for improving the visibility of features in medical images, and data augmentation aids in creating a more robust model by generating additional training samples through transformations.

Table 4: Summary of joint disease detection methods

Study	Methodology	Models/Techniques Used	Dataset	Results
Bhandari et al. [33]	Explainable Artificial Intelligence (XAI).	Deep learning, XAI framework	7132 chest X-ray images	Average test accuracy: 94.31% ± 1.01%, Validation accuracy: 94.54% ± 1.33% (10-fold cross-validation).
Venkataramana et al. [34]	Multi-level classification system: binary classification model for TB and pneumonia, second model for pneumonia types. Synthetic minority oversampling technique (SMOTE) used.	Binary classification, SMOTE	14,693 images	Before balancing: 95.7% accuracy, After balancing: 96.6% accuracy.
Hasan et al. [35]	Model for pneumonia detection in COVID-19 patients using CNN, VGG16 for training. Multiple machine learning tools used.	CNN, VGG16	Labeled X-ray images	Average accuracy: 91.69%, Sensitivity: 95.92%.
Ibrokhimov and Kang [36]	Transfer learning with VGG19 and ResNet50.	VGG19, ResNet50	11,956 COVID-19 samples, 11,263 viral/bacterial pneumonia, 10,701 normal samples	VGG19: Average accuracy of 96.6% across all classes.
Bashar et al. [37]	Deep learning approach for COVID-19 and pneumonia diagnosis using X-ray images.	VGG16, VGG19, GoogleNet, etc.	21,165 chest X-ray images	Highest accuracy: 95.63% using VGG16 on augmented and enhanced dataset.
Baltazar et al. [38]	Optimized five deep learning architectures (InceptionV3, InceptionResNetV2, Xception, VGG, MobileNet).	InceptionV3, InceptionResNetV2, Xception, VGG, MobileNet	Own generated data	InceptionV3: 86% sensitivity, 99% specificity, 91% PPV.
Nasir et al. [39]	Fusion of deep features and Light Gradient Boosting Machine (LightGBM) for COVID-19 detection from chest X-ray images. Dataset of 1125 images used.	Fusion of deep features, LightGBM	1125 chest X-ray images	Two-class accuracy: 98.54%, Three-class accuracy: 91.11%.
Liu et al. [40]	Deep learning model for joint diagnosis of COVID-19 and pneumonia using chest X-ray images. Investigation of various image processing approaches.	Various image processing approaches	Chest X-ray images	Binary classification: 91.5% accuracy, Multivariate classification: 91.11% accuracy.

To further bolster their model's capabilities, the team leveraged multiple transfer learning algorithms. Transfer learning involves utilizing pre-trained models on large datasets and fine-tuning them for a specific task, allowing the model to leverage knowledge learned from diverse datasets. The pinnacle of their achievements came with the VGG16 algorithm, which, when applied to the augmented and enhanced dataset, achieved an outstanding accuracy of 95.63%. The choice of VGG16 and the impressive accuracy obtained highlight the efficacy of this particular architecture in accurately diagnosing COVID-19 and pneumonia from chest X-ray images.

Baltazar et al. [38] focused their efforts on the development of models for detecting COVID-19 and pneumonia, utilizing a dataset generated in-house. In their study, they explored and optimized five diverse deep learning architectures, to determine the most effective model for their specific diagnostic task. The results of their investigation revealed that InceptionV3 emerged as the top-performing architecture among the selected models. InceptionV3 demonstrated impressive performance metrics, with 86% sensitivity, 99% specificity, and a positive predictive value (PPV) of 91%. These metrics are critical in evaluating the effectiveness of a diagnostic model, particularly in the context of medical image analysis. The reported sensitivity of 86% indicates the model's ability to correctly identify true positive cases, specifically in this study, detecting COVID-19 and pneumonia. The high specificity of 99% signifies the model's accuracy in correctly identifying true negative cases, minimizing false positives. The positive predictive value (PPV) of 91% highlights the reliability of the model in correctly identifying positive cases among those predicted.

Nasir et al. [39] introduced a novel approach for detecting COVID-19 from chest X-ray images, combining deep features and the Light Gradient Boosting Machine (LightGBM). In their innovative scheme, the researchers worked with a dataset consisting of 1125 images, demonstrating the effectiveness of their proposed methodology in accurately classifying the images into different categories. In the two-class scenario, distinguishing between COVID-19 and Healthy cases, the proposed scheme achieved an impressive accuracy of 98.54%. This high accuracy reflects the model's ability to make correct predictions in differentiating between normal and COVID-19-affected X-ray images. In the more complex three-class scenario, which involved classifying images into COVID-19, Healthy, and Pneumonia categories, the proposed scheme exhibited a commendable accuracy of 91.11%. This suggests the robustness of the model in handling the nuanced task of categorizing X-ray images into multiple classes, reflecting different respiratory conditions.

Liu et al. [40] contributed to the field of medical image analysis with a sophisticated DNN model designed for the joint diagnosis of COVID-19 and pneumonia using chest X-ray images. Their approach involved a comprehensive exploration of various image processing techniques, showcasing the importance of incorporating diverse methodologies to enhance the diagnostic capabilities of the model. These techniques play a crucial role in capturing intricate details and patterns within the chest X-ray images, enabling the model to make more nuanced and accurate diagnoses. The reported accuracy of the model for both binary and multivariate classification is noteworthy. For binary classification, distinguishing between COVID-19 and pneumonia cases, the model achieved an accuracy of 91.5%. In the more complex multivariate classification scenario, involving the identification of COVID-19, pneumonia, and potentially other conditions, the model attained an accuracy of 91.11%. These high accuracy rates underscore the effectiveness of the proposed deep learning model in accurately identifying and differentiating between respiratory conditions.
The significance of incorporating diverse image processing techniques is evident in the model's success. By leveraging various methodologies, the model can capture a wide range of features and patterns in the X-ray images, enhancing its ability to discern between different medical conditions. This comprehensive and multi-faceted approach reflects the evolving nature of deep learning in medical image analysis and its potential to advance diagnostic capabilities in the identification of COVID-19 and pneumonia cases.

4. Challenges in Lung disease detection

4.1 Image Quality Variability: One of the primary challenges arises from the variability in image quality. Factors such as the type of X-ray or CT machine used, the settings chosen for the scan (e.g., exposure level and resolution), and the presence of artifacts or noise can all affect the quality of the images. High-quality images are characterized by sharp contrast, well-defined structures, and minimal noise, while low-quality images may exhibit blurriness or artifacts. Ensuring consistent model performance across this wide quality spectrum is a non-trivial task.

The exposure level during imaging can significantly impact the appearance of lung structures. Overexposed images may result in areas of pure white, making it challenging to discern details, while underexposed images may be too dark, obscuring important features. Models must be robust enough to handle these variations and still accurately identify abnormalities, such as nodules or infiltrates.

Patients' positioning during imaging can vary, leading to differences in how lung structures are captured. Variations in posture and patient anatomy can affect the way X-rays or CT scans are interpreted. For instance, a patient's arm or ribcage may partially overlap with the lung area, potentially mimicking abnormalities. Models need to be trained to recognize and account for these positional variations.

Addressing these challenges requires careful preprocessing of the data to standardize image quality and ensure that images are correctly aligned. Additionally, data augmentation techniques can be applied to artificially introduce variations in exposure levels and positioning during training, helping the model become more robust.

4.2 Anatomical Variability: The human body is incredibly diverse, and its anatomy can vary significantly from one individual to another. This inherent variability extends to the structure and appearance of the lungs, which are a key focus in medical imaging for the diagnosis of lung diseases. This anatomical diversity presents a challenge when developing deep learning models for accurate and robust disease detection, as these models must contend with the nuanced differences in lung structure and presentation across individuals.

1. **Lung Shape and Size:** The size and shape of the lungs can differ from person to person. Some individuals may have larger or smaller lungs, and the relative position of the organs within the chest cavity can vary. This variation can influence the appearance of lung diseases, as abnormalities may manifest differently in individuals with distinct lung sizes and shapes.
2. **Anatomical Features:** Lung anatomy is not uniform; it includes variations in bronchial branching, blood vessel patterns, and the presence of fissures that separate different lobes of the lung. These anatomical features can create unique patterns in medical images. For instance, the presence or absence of fissures can affect how lung nodules or infiltrates appear in X-rays or CT scans.
3. **Patient History and Health Status:** Factors such as a patient's medical history, underlying health conditions, and lifestyle can influence the appearance of lung diseases. For example, individuals with preexisting lung conditions or a history of smoking may present with different patterns of lung disease, even when affected by the same condition.
4. **Demographic Variability:** Age, gender, and genetic factors can contribute to variability in lung anatomy and, consequently, the presentation of lung diseases. For instance, certain lung diseases may be more prevalent in specific demographic groups, leading to differences in disease characteristics.
5. **Pathological Variations:** Lung diseases themselves can exhibit variations in their appearance. For example, lung cancer can manifest as different types (e.g., adenocarcinoma, squamous cell carcinoma), each with its own distinctive radiological features. These variations challenge models to generalize across the diverse manifestations of the same disease.

4.3 Small Lesion Detection: Detecting lung diseases, especially those characterized by small lesions or nodules, presents a formidable challenge in the field of medical imaging. These tiny abnormalities may be the early signs of serious conditions like lung cancer or infections, and they require prompt and accurate identification for effective treatment. Expert radiologists, despite their extensive training and experience, can sometimes struggle to detect these subtle cues. This is where deep learning models, with their ability to analyze vast amounts of medical imaging data, play a crucial role.

1. **Size and Scale:** Lesions and nodules can vary in size, with some being just a few millimeters in diameter. These minuscule dimensions make them difficult to distinguish from surrounding healthy tissue, even for human experts who must meticulously analyze numerous image slices.

2. **Low Contrast:** Subtle lesions often exhibit low contrast compared to their surroundings, making them nearly imperceptible in standard medical images. This reduced contrast further compounds the difficulty in detecting them.
3. **Noise and Artifacts:** Medical images can contain noise and artifacts due to various factors, including equipment imperfections or patient motion during imaging. These artifacts can obscure the already faint signals from tiny lesions, creating a challenging environment for both radiologists and machine learning models.
4. **Overlapping Structures:** Lesions may overlap with other anatomical structures, such as blood vessels or adjacent organs, further complicating their identification. Differentiating between the lesion and overlapping structures requires a high level of sensitivity.
5. **Progression Monitoring:** For conditions like lung cancer, it's crucial to detect lesions at an early stage when they are small. Deep learning models need to not only detect small lesions but also monitor their growth over time, which can be crucial for prognosis and treatment planning.

4.4 Noise and Artifacts: The quality and reliability of medical images are paramount in healthcare, as they form the basis for accurate diagnoses and treatment decisions. However, the acquisition and processing of medical images can introduce various sources of noise and artifacts, which, in turn, can pose significant challenges for deep learning models and impact their performance.

1. **Patient Motion:** Patients may not always remain perfectly still during the image acquisition process, especially in procedures like MRI or CT scans that require longer imaging times. Even minor movements, such as breathing or involuntary muscle twitches, can lead to motion artifacts in the images. These artifacts can distort anatomical structures and create false abnormalities, confounding deep learning models.
2. **Image Compression:** To manage the large file sizes associated with medical images, healthcare institutions often employ image compression techniques. While compression reduces storage requirements, it can also introduce compression artifacts, such as blocky patterns or loss of fine details. These artifacts can make it difficult for models to accurately interpret the images.

Addressing these challenges involves a combination of techniques to enhance the quality of medical images and improve deep learning model robustness:

- **Motion Correction:** Post-processing techniques can correct motion artifacts by aligning multiple image frames or slices, reducing the impact of patient motion.
- **Artifact Reduction Algorithms:** Specialized algorithms can be applied to suppress or remove specific types of artifacts, such as those caused by metal objects in MRI.
- **High-Quality Imaging Protocols:** Implementing optimized imaging protocols and quality assurance measures can minimize equipment-related artifacts and inconsistencies.
- **Noise Reduction:** Noise reduction techniques, including denoising algorithms, can be employed to enhance image clarity while preserving diagnostic information.
- **Standardized Image Formats:** Using standardized image formats and avoiding excessive compression can help maintain image fidelity.

To overcome these challenges, deep learning models must be trained on large and diverse datasets that encompass the full spectrum of anatomical variations and disease presentations. Data augmentation techniques can artificially introduce variations in lung anatomy, helping models become more adaptable to real-world cases. Additionally, transfer learning, where models are pre-trained on a broad range of data, can improve their ability to recognize patterns in diverse lung images.

5. Conclusion

In conclusion, this comprehensive review has systematically explored the application of deep learning methodologies in the detection of critical medical conditions, namely pneumonia, COVID-19, cancer, and joint diseases, utilizing various imaging modalities such as X-rays and CT scans. The in-depth analysis covered recent advancements in deep learning architectures, including CNNs, RNNs, and hybrid models, tailored to address the distinct challenges posed by each disease category. The exploration of nuances in pneumonia detection, methodologies for timely identification of COVID-19, and the potential of DNNs in cancer diagnosis underscore the significance of leveraging advanced technologies for improved healthcare outcomes. By synthesizing the current state of knowledge, this review serves as a valuable resource for researchers, clinicians, and policymakers. It critically evaluates the strengths and limitations of existing methodologies, providing insights into research gaps and proposing potential directions for future studies. Through these efforts, the paper contributes to the ongoing advancement of medical imaging, offering a pathway to enhance early detection and management strategies for pneumonia, COVID-19, cancer, and joint diseases, ultimately improving patient care and outcomes in diverse medical domains.

References

1. Gupta, Naman, Deepak Gupta, Ashish Khanna, Pedro P. Rebouças Filho, and Victor Hugo C. de Albuquerque. "Evolutionary algorithms for automatic lung disease detection." *Measurement* 140 (2019): 590-608.
2. Sahu, Tikendra, and Aakanksha S. Choubey. "A Review on Lungs Disease Detection using Image Processing." *i-Manager's Journal on Information Technology* 11, no. 1 (2021): 48.
3. Das, Nilakash, Marko Topalovic, and Wim Janssens. "Artificial intelligence in diagnosis of obstructive lung disease: current status and future potential." *Current opinion in pulmonary medicine* 24, no. 2 (2018): 117-123.
4. Kido, Shoji, Yasusi Hirano, and Noriaki Hashimoto. "Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN)." In *2018 International workshop on advanced image technology (IWAIT)*, pp. 1-4. IEEE, 2018.
5. Hong, Min, Beanbonyka Rim, Hongchang Lee, Hyeonung Jang, Joonho Oh, and Seongjun Choi. "Multi-class classification of lung diseases using CNN models." *Applied Sciences* 11, no. 19 (2021): 9289.
6. Goyal, Shimpy, and Rajiv Singh. "Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques." *Journal of Ambient Intelligence and Humanized Computing* (2021): 1-21.
7. Hashmi, Mohammad Farukh, Satyarth Katiyar, Avinash G. Keskar, Neeraj Dhanraj Bokde, and Zong Woo Geem. "Efficient pneumonia detection in chest xray images using deep transfer learning." *Diagnostics* 10, no. 6 (2020): 417.
8. Stephen, Okeke, Mangal Sain, Uchenna Joseph Maduh, and Do-Un Jeong. "An efficient deep learning approach to pneumonia classification in healthcare." *Journal of healthcare engineering* 2019 (2019).
9. Elshennawy, Nada M., and Dina M. Ibrahim. "Deep-pneumonia framework using deep learning models based on chest X-ray images." *Diagnostics* 10, no. 9 (2020): 649.
10. Szepesi, Patrik, and László Szilágyi. "Detection of pneumonia using convolutional neural networks and deep learning." *Biocybernetics and Biomedical Engineering* 42, no. 3 (2022): 1012-1022.
11. Ahmed, Mohammed Salih, Atta Rahman, Faris AlGhamdi, Saleh AlDakheel, Hammam Hakami, Ali AlJumah, Zuhair AlIbrahim, Mustafa Youldash, Mohammad Aftab Alam Khan, and Mohammed Imran Basheer Ahmed. "Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach." *Diagnostics* 13, no. 15 (2023): 2562.
12. Abubeker, K. M., and S. Baskar. "B2-Net: an artificial intelligence powered machine learning framework for the classification of pneumonia in chest x-ray images." *Machine Learning: Science and Technology* 4, no. 1 (2023): 015036.
13. Ramadhan, Awf A., and Muhammet Baykara. "A Novel Approach to Detect COVID-19: Enhanced Deep Learning Models with Convolutional Neural Networks." *Applied Sciences* 12, no. 18 (2022): 9325.
14. Jain, Rachna, Meenu Gupta, Soham Taneja, and D. Jude Hemanth. "Deep learning based detection and analysis of COVID-19 on chest X-ray images." *Applied Intelligence* 51 (2021): 1690-1700.
15. Hussain, Emtiaz, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, and Mohammad Zavid Parvez. "CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images." *Chaos, Solitons & Fractals* 142 (2021): 110495.
16. Nayak, Soumya Ranjan, Deepak Ranjan Nayak, Utkarsh Sinha, Vaibhav Arora, and Ram Bilas Pachori. "Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study." *Biomedical Signal Processing and Control* 64 (2021): 102365.
17. Zagrouba, Rachid, Muhammad Adnan Khan, Muhammad Aamer Saleem, Muhammad Faheem Mushtaq, Abdur Rehman, and Muhammad Farhan Khan. "Modelling and Simulation of COVID-19

Outbreak Prediction Using Supervised Machine Learning." *Computers, Materials & Continua* 66, no. 3 (2021).
18. Atta-ur-Rahman, A., K. Sultan, I. Naseer, R. Majeed, D. Musleh, M. Abdul Salam Gollapalli, S. Chabani, N. Ibrahim, S. Yamin Siddiqui, and M. Adnan Khan. "Supervised machine learning-based prediction of COVID-19." *Computers, Materials & Continua* 69, no. 1 (2021): 21-34.
19. Ahmed, Mohammed Imran Basheer, Atta-ur Rahman, Mehwash Farooqui, Fatimah Alamoudi, Raghad Baageel, and Amjad Alqarni. "Early Identification of COVID-19 Using Dynamic Fuzzy Rule Based System." *Mathematical Modelling of Engineering Problems* 8, no. 5 (2021).
20. Naqvi, Rizwan Ali, Muhammad Faheem Mushtaq, Natash Ali Mian, Muhammad Adnan Khan, Muhammad Ali Yousaf, Muhammad Umair, and Rizwan Majeed. "Coronavirus: A mild virus turned deadly infection." *Computers, Materials and Continua* 67, no. 2 (2021): 2631-2646.
21. Nasiri, Hamid, and Sharif Hasani. "Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost." *Radiography* 28, no. 3 (2022): 732-738.
22. Khan, Saddam Hussain, Anabia Sohail, Asifullah Khan, and Yeon-Soo Lee. "COVID-19 detection in chest X-ray images using a new channel boosted CNN." *Diagnostics* 12, no. 2 (2022): 267.
23. Alsinglawi, Belal, Osama Alshari, Mohammed Alorjani, Omar Mubin, Fady Alnajjar, Mauricio Novoa, and Omar Darwish. "An explainable machine learning framework for lung cancer hospital length of stay prediction." *Scientific reports* 12, no. 1 (2022): 607.
24. Venkatesh, Surrajkumar Prabhu, and Lilly Raamesh. "Predicting lung cancer survivability: A machine learning ensemble method on seer data." (2022).
25. Vikas, Prabhpreet Kaur. "Lung cancer detection using chi-square feature selection and support vector machine algorithm." *International Journal of Advanced Trends in Computer Science and Engineering* (2021).
26. Chauhan, Anamika. "Detection of lung cancer using machine learning techniques based on routine blood indices." In *2020 IEEE international conference for innovation in technology (INOCON)*, pp. 1-6. IEEE, 2020.
27. Sim, Jin-ah, Young Ae Kim, Ju Han Kim, Jong Mog Lee, Moon Soo Kim, Young Mog Shim, Jae Ill Zo, and Young Ho Yun. "The major effects of health-related quality of life on 5-year survival prediction among lung cancer survivors: applications of machine learning." *Scientific reports* 10, no. 1 (2020): 10693.
28. Patra, Radhanath. "Prediction of lung cancer using machine learning classifier." In *Computing Science, Communication and Security: First International Conference, COMS2 2020, Gujarat, India, March 26–27, 2020, Revised Selected Papers 1*, pp. 132-142. Springer Singapore, 2020.
29. Radhika, P. R., Rakhi AS Nair, and G. Veena. "A comparative study of lung cancer detection using machine learning algorithms." In *2019 IEEE international conference on electrical, computer and communication technologies (ICECCT)*, pp. 1-4. IEEE, 2019.
30. Wu, Jiangpeng, Xiangyi Zan, Liping Gao, Jianhong Zhao, Jing Fan, Hengxue Shi, Yixin Wan, E. Yu, Shuyan Li, and Xiaodong Xie. "A machine learning method for identifying lung cancer based on routine blood indices: qualitative feasibility study." *JMIR medical informatics* 7, no. 3 (2019): e13476.
31. Faisal, Muhammad Imran, Saba Bashir, Zain Sikandar Khan, and Farhan Hassan Khan. "An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer." In *2018 3rd international conference on emerging trends in engineering, sciences and technology (ICEEST)*, pp. 1-4. IEEE, 2018.
32. Safiyari, Ali, and Reza Javidan. "Predicting lung cancer survivability using ensemble learning methods." In *2017 intelligent systems conference (IntelliSys)*, pp. 684-688. IEEE, 2017.
33. Bhandari, Mohan, Tej Bahadur Shahi, Birat Siku, and Arjun Neupane. "Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI." *Computers in Biology and Medicine* 150 (2022): 106156.
34. Venkataramana, Lokeswari, D. Venkata Vara Prasad, S. Saraswathi, C. M. Mithumary, R. Karthikeyan, and N. Monika. "Classification of COVID-19 from tuberculosis and pneumonia using deep learning techniques." *Medical & Biological Engineering & Computing* 60, no. 9 (2022): 2681-2691.
35. Hasan, MD Kamrul, Sakil Ahmed, ZM Ekram Abdullah, Mohammad Monirujjaman Khan, Divya Anand, Aman Singh, Mohammad AlZain, and Mehedi Masud. "Deep learning approaches for detecting pneumonia in COVID-19 patients by analyzing chest X-ray images." *Mathematical Problems in Engineering* 2021 (2021): 1-8.

36. Ibrokhimov, Bunyodbek, and Justin-Youngwook Kang. "Deep learning model for COVID-19-infected pneumonia diagnosis using chest radiography images." *BioMedInformatics* 2, no. 4 (2022): 654-670.
37. Bashar, Abul, Ghazanfar Latif, Ghassen Ben Brahim, Nazeeruddin Mohammad, and Jaafar Alghazo. "COVID-19 pneumonia detection using optimized deep learning techniques." *Diagnostics* 11, no. 11 (2021): 1972.
38. Baltazar, Lei Rigi, Mojhune Gabriel Manzanillo, Joverlyn Gaudillo, Ethel Dominique Viray, Mario Domingo, Beatrice Tiangco, and Jason Albia. "Artificial intelligence on COVID-19 pneumonia detection using chest xray images." *Plos one* 16, no. 10 (2021): e0257884.
39. Nasiri, Hamid, Ghazal Kheyroddin, Morteza Dorrigiv, Mona Esmaeili, Amir Raeisi Nafchi, Mohsen Haji Ghorbani, and Payman Zarkesh-Ha. "Classification of COVID-19 in chest X-ray images using fusion of deep features and LightGBM." In *2022 IEEE world AI IoT congress (AIIoT)*, pp. 201-206. IEEE, 2022.
40. Liu, Xiangbin, Wenqian Wu, Jerry Chun-Wei Lin, and Shuai Liu. "A Deep Learning Model for Diagnosing COVID-19 and Pneumonia through X-ray." *Current Medical Imaging* 19, no. 4 (2023): 333-346.