



Pneumonia Detection Using Cnns And Random Forest Algorithm

¹Najmusher H, ²Yogitha H T, ³Sarika A P, ⁴Vaishnavi Mugale

¹Assistant Professor, ²Student, ⁴Student, ⁴Student

¹Department of Computer Science and Engineering,

¹HKBK College of Engineering, Bangalore, India

Abstract: This study has been undertaken to investigate the determinants of stock returns in Karachi Stock Exchange (KSE) using two assets pricing models the classical Capital Asset Pricing Model and Arbitrage Pricing Theory model. To test the CAPM market return is used and macroeconomic variables are used to test the APT. The macroeconomic variables include inflation, oil prices, interest rate and exchange rate. For the very purpose monthly time series data has been arranged from Jan 2010 to Dec 2014. The analytical framework contains.

Index Terms - Pneumonia Detection, Convolutional Neural Network (CNN), Chest X-ray Analysis, Healthcare AI, Random Forest Algorithm, Machine Learning.

I. INTRODUCTION

Pneumonia is an acute inflammatory condition of the lungs, typically caused by bacterial, viral, or fungal infections. Previously discovery of pneumonia is captious for effective treatment and reducing mortality rates. working advancements in machine knowledge and deep knowledge, have automated styles for discovery of pneumonia have shown great pledge in amending individual delicacy, speed, and vacuity. We'll use combination of two algorithm in this design for pneumonia discovery Convolutional Neural Networks (CNN) Convolutional Neural Networks (CNNs) are a class of deep education algorithms. This is particularly well- suited for image division tasks. In pneumonia disinterring, CNNs are generally used to break down medical imaging data, corresponding as casketX- shafts. CNNs root hierarchical features from images, enabling them to honor complex patterns and also this CNN will distinguish between normal and infected lung handkerchief. The use of CNNs minimizes the need for homemade point birth, making the process largely effective and secure. Random Forest Algorithm Random Forest combines beaucoup decision trees to fashion healthy prophecies while CNNs transcend at image- rested point birth, Random Forest can be used to classify the features corkscrewed by CNNs. The combination of CNNs and Random Forest can enhance performance CNNs for point-rich image processing and Random Forest algorithm for type. Combined Approach Integration of both CNNs and Random Forest algorithm takes advantage of their reciprocal strengths point birth CNNs pull deep, high- stratum features from casketX-shaft images. Bracket These features are also fed into a Random Forest classifier to make the final predicting. This methodology increases individual delicacy and ensures that the system is important to variability in image quality and other factors. The integration of both the algorithm will give Improved Accuracy Combining CNN and Random Forest can achieve progressive delicacy than using either strategy alone. automation Reduces the burden on healthcare professionals by automating the unearthing process. In summary, the integration of CNNs and Random Forest for pneumonia exhuming represents a significant step forward in medical diagnostics, offering a steady, fruitful, and scalable result to a global health challenge.

II. RELATED WORK

Recent progress in computer vision and deep learning has made automated medical image analysis increasingly reliable.[1] Specifically, pneumonia detection from chest X-ray images has drawn considerable attention due to its clinical importance and the potential for reducing diagnostic time. One of the widely used approaches involves Convolutional Neural Networks (CNNs), which have been trained to classify chest radiographs by learning complex patterns in lung textures. Researchers have found that even basic CNN architectures can detect signs of pneumonia with high accuracy, especially when large annotated datasets are available. [5] In addition to classification, some researchers have applied segmentation techniques to identify affected regions in the lungs. Models based on U-Net and similar architectures have shown promise in highlighting infection areas, offering visual cues that can support clinical decision-making. Other works have explored hybrid models and ensemble methods to enhance robustness. Data augmentation and preprocessing techniques, such as histogram equalization and lung field cropping, are often used to improve model input quality. Despite these advancements, ensuring model reliability, interpretability, and clinical integration remains a challenge. This motivates ongoing research into lightweight, accurate, and interpretable models for pneumonia detection using deep learning. To enhance accuracy and efficiency, researchers have increasingly utilized transfer learning techniques. In this approach, models like Resnet, VGG, and Inception—originally trained on large image datasets—are adapted to medical imaging tasks by retraining them on chest X-ray images. This method reduces the need for extensive training from scratch and improves the model's ability to generalize, particularly when working with smaller or specialized datasets

Sl. No	Author(s)/Year	Methodology	Key Benefits	Limitations	Performance
1	M. Ali, 2024	EfficientNetV2L	Automated diagnosis with strong accuracy	Needs high-quality X-rays and powerful hardware	High accuracy and efficiency
2	Yaseliani, 2022	Decision Tree + Random Forest + KNN + Logistic Regression	Detects unsafe water, useful in isolated areas	Varies with input quality across water sources	Reliable with ~81% accuracy
3	H. Ren, 2021	NodeMCU + GPS + Bluetooth + Vital Sign Sensors	Real-time health tracking and pet location monitoring	Depends on sensor accuracy and integration	Efficient alerts, precise data flow
4	Jianpeng Zhang, 2021	Chest X-rays + CAAD model	Operates with fewer labeled images, helpful early on	Limited to chest Xrays; needs more clinical proof	Sensitivity: 93%
5	N. Shilpa, 2024	X-rays + CLAHE + 5 CNNs	Delivers fast and precise predictions	Requires validation across age brackets	High accuracy achieved
6	Yeongbong Jin, 2023	CycleGANs + Chest X-rays	Can assist in algorithmic development for pneumonia stages	Limited scope physiologically; variability among cases	Recall: 97%, F1score: 87.8% (real), 82.3% (synthetic)
7	Adnan Hussain et al., 2023	Deep Ensemble Strategy with Chest X-ray Analysis	Combines multiple deep learning models for more reliable diagnosis	computationally expensive and requires large memory for ensemble processing	98.43%

8	T. S. Arulananth, 2024	Chest X-rays + DenseNet + Normalization	Works well in resource-limited settings	Needs interpretability and broader testing	Accuracy: 97.03%, AUC: 96.34%
9	Kang, 2019	CT-based CAP detection	Compact, integrationfriendly	Challenges in distinguishing CAP from COVID-19	Accuracy: 95.5%, Sensitivity: 96.6%
10	A. Tripathi, 2024	MobileNetV2L2 + CLAHE + CNN + ELU	Boosts performance and diagnosis precision	Requires structured inputs and limited to classification	Training: 99.53%, Validation: 100%, Test: 95.51%
11	Chowdhury et al., 2020	Deep CNNs + Data Augmentation	High sensitivity; useful in quick screening	Drops in multi-class cases, especially with COVID-19	Deep CNNs show generalization on augmented data
12	Yaoming Lai & Guangming.L,2021	NCIP-Net + Deep CNN + DCNN	Cloud-ready analytics, effective on limited data	Low training data limits performance	AUC (Test): 0.919, External AUC: 0.908
13	Dimpy Varshni et al., 2019	DenseNet-169 + ChestX-ray14	Enhances pneumonia classification	Excludes patient history/clinical info	AUC: 0.8002
14	Kosasih et al., 2015	Wavelet Transform	Low-cost, suitable for low-resource zones	Environmentally sensitive results, limited X-ray use	Sensitivity: 94%, Specificity: 63%
15	W. Khan, 2021	ChexNet (DenseNet121)	Handles large datasets; highly sensitive	Risk of bias due to class imbalance	Accuracy up to 98.82%,AUC:0.99726

Table 1: Comparison of various methodologies for Pneumonia disease detection

III. SUMMARY OF OUTCOMES

[2] With the increase in colorful medical discovery algorithms which are being used in the medical field there would be colorful anomalies set up the opinion systems. Then we use the combined form of the CNN and the Random Forest algorithm to prove the effectiveness in using both the algorithms while detecting pneumonia through the x-rays reports and also the symptoms that are being observed for a prolonged time in a case. This helps us to get into the medical opinion more directly also only pertaining with the reports of the x-ray which might not give accurate results as anticipated.

3.1 Enhanced Diagnostic Accuracy

[11] The integration of the CNN and Random Forest algorithms have redounded in bettered delicacy and robust bracket compared to using either model singly. Then CNN helps to prize the colorful features from the casketx-ray images and Random Forest helps in the bracket grounded on the features being handed.

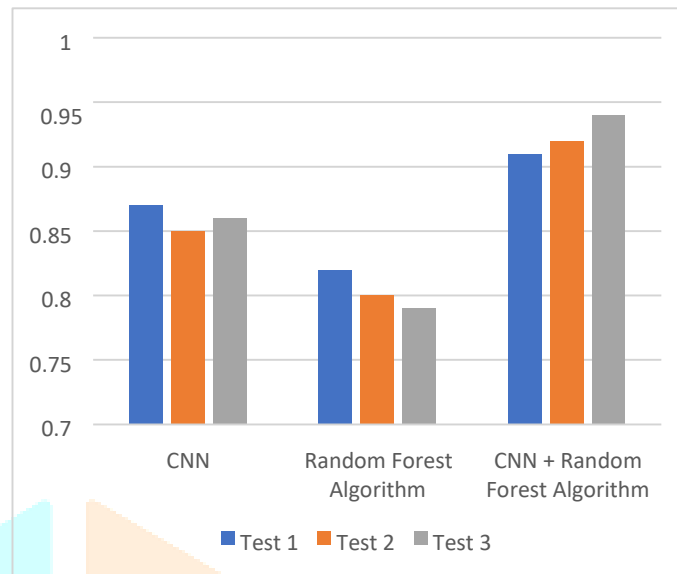


Fig. 3 Comparison of miniature interpretation benchmarks

3.2 Operative Robotization

[1] Since the model is significantly automated for the discovery of pneumonia it helps in minimizing the time for homemade interpretation. It also helps to reduce the workload from healthcare professionals, especially in high- demand settings where the medical time would not be suitable to break down the undivided reports to find the factual medical affiliated opinion.

3.3 Robustness to Image Variability

[14] The system is suitable to demonstrate adaptability against the inconsistencies in image quality which is veritably common in real- world clinical data. Then the robustness contributes to the trustability of the model in different healthcare surroundings. Images for the casketx rays are n't of the RGB images theX-rays do n't give the proper view of the casket the algorithm which is being used helps in the specialized image analyses and give proper affair for the medical professional to understand the factual opinion

3.4 Scalability and Deployment Readiness

[9] The methodology which is being designed then can be scalable to both clinical and remote healthcare setting as we're suitable to prognosticate the complaint with the help of symptoms as well which is substantially useful incase of remote medical camps where the casketx-rays might not be available.

3.5 Implicit for Broder Applications

[15] While the frame developed substantially focuses on the pneumonia discovery there's a way for applying it in analogous ways similar as to descry the other respiratory conditions one of the good illustration could be covid- 19. The opens to the unborn advancements using this mongrel model for the medical image analysis.

3.6 Reduction in Human Intervention

[13] Since the mongrel model as a web interface helps in getting the cases data including the symptoms,x-rays reports, Doctor movables , medical conventions etc it helps maintain the data directly and also give proper medical advises and helps the cases to follow up with the croaker and use the specified drug with the time period as mentioned in his lines. All these colorful features handed by this design also helps both the cases and the medical professionals to easily understand the opinion and continue their treatment consequently.

3.7 Flowchart

[7] The flowchart outlines a technical workflow for an AI-driven pneumonia detection system using chest X-rays. The process initiates with user authentication through registration or login, followed by access to or creation of a patient profile. Once the user uploads a chest X-ray image, the system integrates this image with patient symptom data using either a rule-based approach or a machine learning model for pre-processing. A Convolutional Neural Network (CNN) specifically trained for pneumonia detection then analyzes the X-ray to identify potential signs of infection. Based on the model's inference, the system categorizes the diagnostic outcome as either "pneumonia likely," "pneumonia unlikely," or suggests further testing if the result is inconclusive. A comprehensive, printable medical report is then generated, incorporating the diagnosis, associated symptoms, X-ray image, and clinical recommendations. Depending on the diagnosis, the system provides targeted follow-up instructions, which may include home care guidelines, urgent hospital visit advisories, or medication recommendations. Finally, users have the option to schedule a medical appointment by selecting an available time slot with a healthcare provider through an integrated booking feature.

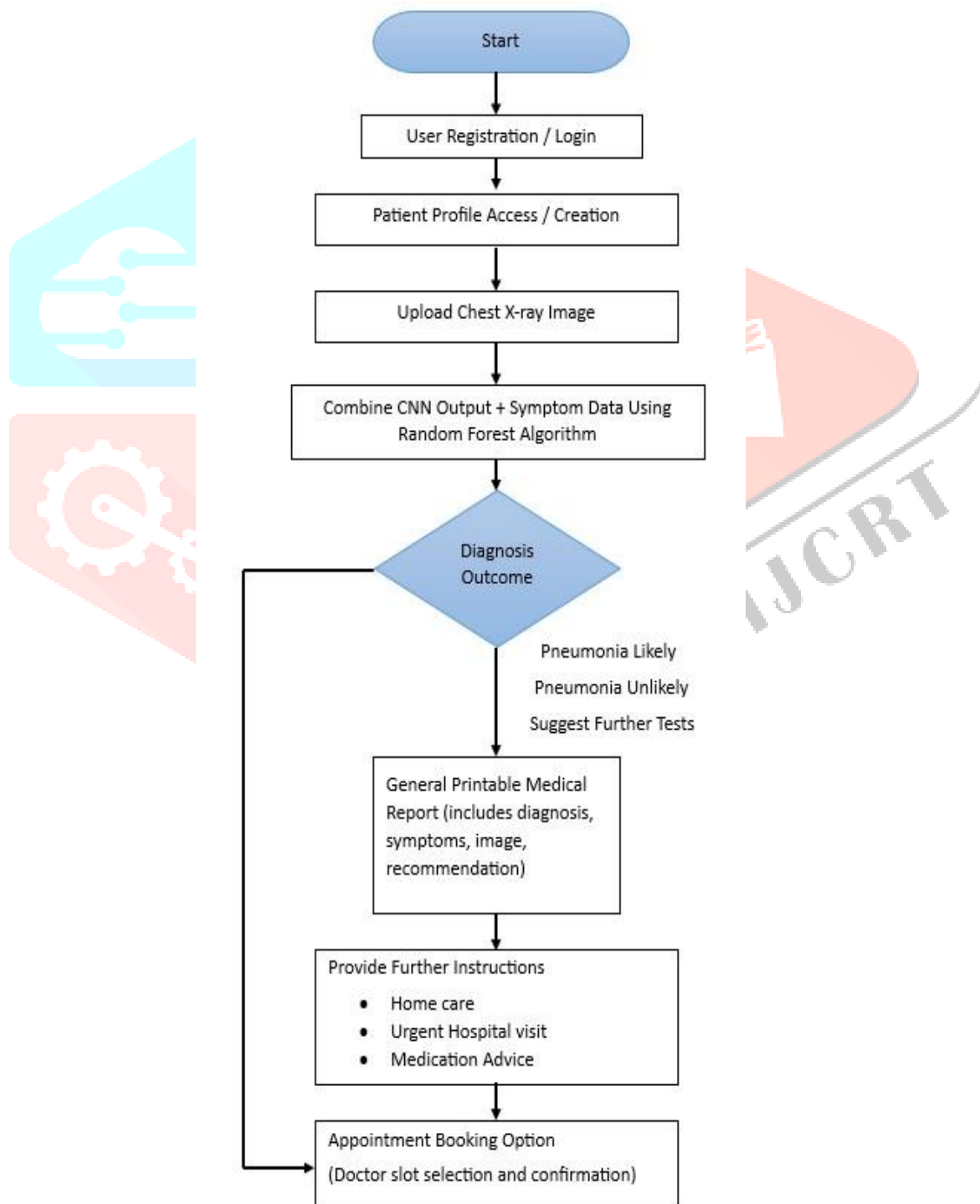


Fig. 4. Pneumonia Detection & Diagnosis Workflow

IV. EXISTING RESEARCH GAPS

Identified Research Gaps in Pneumonia Detection

4.1 Limited Generalization Across Datasets

[15] Most deep learning models for pneumonia detection are trained and validated on a single dataset (e.g., ChestX-ray14 or Covid). This limits their generalization capability across different populations, imaging devices, or clinical settings. There is a need for multi-source or cross-domain models that can adapt to diverse datasets.

4.2 Lack of Explainability in Deep Models

[9] While deep learning models show promising accuracy, many work as “black boxes,” offering little insight into their decision-making process. Explainable AI (XAI) approaches are needed to ensure clinicians understand and trust the model predictions, especially for critical diagnoses like pneumonia.

4.3 Underutilization of Clinical Metadata

[14] Most studies rely solely on imaging data (X-ray or CT) without integrating relevant patient metadata such as age, gender, smoking history, or symptoms. Fusing clinical and image data may enhance diagnostic performance and provide personalized insights.

4.4 Neglect of Class Imbalance and Rare Case Representation

[2] Datasets are often skewed, with more normal cases than pneumonia-positive ones. Many models overlook the imbalance issue, which can result in biased predictions. Techniques like focal loss, data augmentation, and generative models can be further explored to address this.

4.5 Limited Real-Time and Edge Deployability

Despite high accuracy in controlled environments, many models are too computationally heavy for deployment on edge devices or in real-time clinical applications. Research is needed on lightweight architectures, model compression, and deployment frameworks.

4.6 Scarcity of Longitudinal and Progression Tracking Models

[12] Current models primarily focus on binary classification (pneumonia or not). Few studies track the progression or regression of pneumonia over time using sequential imaging. Developing models that monitor disease evolution can aid treatment planning.

4.7 Insufficient External Validation and Clinical Trials

A significant gap exists in external validation using unseen hospital data and real-world clinical trials. Most research remains confined to academic settings without robust clinical evaluation, limiting real-life applicability.

4.8 Inadequate Handling of Noisy or Low-Quality Images

[8] Real-world X-ray images often contain noise, artifacts, or suboptimal quality, yet many models assume ideal conditions. Robust pre-processing and denoising techniques tailored to clinical settings are underexplored.

4.9 Ethical and Bias Considerations

[7] There is limited discussion on ethical concerns such as demographic bias, data privacy, and informed consent. More work is needed on fairness-aware AI models that ensure equitable healthcare outcomes.

V. FUTURE SCOPE

[11] The future scope of a pneumonia detection project using a combination of Convolutional Neural Networks (CNNs) and Random Forest algorithms is vast, given the increasing focus on AI-driven healthcare solutions. Here's an overview of the potential directions and advancements:

5.1 Development of Comprehensive Diagnostic Tools

[14] Integration with some Other Imaging Methods: Expanding the project to include MRI, CT scans, or even ultrasound imaging for more comprehensive diagnostic compass. [15] Multi-Disease Diagnosis: Enhancing the model to detect other respiratory conditions, such as lung cancer, tuberculosis, or COVID-19, alongside pneumonia.

5.2 Deployment in Real-World Applications

[9] Telemedicine: Including this model into telemedicine platforms to provide remote diagnostic capabilities, particularly in underprivileged areas..

5.3 Enhancement of Model Capabilities

[5] Explainability: Incorporating explainable AI (XAI) techniques to make the decision-making process of the model transparent and understandable for clinicians. Performance Optimization: Reduce false positives/negatives, examine different ways to enhance model accuracy, and optimize computational efficiency.

5.4 Personalization and Precision Medicine

Patient-Specific Models: Incorporating patient-specific factors, such as medical history, age, or any genetic problems to improve diagnostic accuracy. Treatment Monitoring: Using the model to monitor the effectiveness of treatments over time through imaging data.

5.5 Scalability and Accessibility

Global Reach: Adapting the model for use in various healthcare systems worldwide, considering differences in population demographics and imaging equipment. Cost-Effectiveness: Developing low-cost solutions that are affordable for healthcare providers in resource-limited settings. By pursuing these directions, the project can evolve into a transformative tool for pneumonia diagnosis, contributing significantly to the global fight against respiratory diseases.

VI. CONCLUSION

Here we detect the pneumonia detection using the CNN and Random Forest algorithm particularly increases the role of artificial intelligence particularly machine learning and deep learning, in enhancing diagnostic accuracy and efficiency in diagnostic accuracy and efficiency. Convolutional Neural Networks (CNNs) have demonstrated strong capabilities in extracting hierarchical image features from chest X-rays, leading to high performance in classification tasks. On the other hand, Random Forest algorithms, known for their robustness and interpretability, have been effectively used either independently or in hybrid models to improve prediction accuracy by leveraging structured features. With the Overall integration of CNNs and Random Forests offers a promising approach to developing reliable, scalable, and interpretable systems for early pneumonia detection. However, further research is needed to address challenges related to dataset diversity, model generalization, and clinical deployment.

REFERENCES

- [1] Mudasir Ali, Mobeen Shahroz, Urooj Akram, Muhammad Faheem Mushtaq, Stefanía Carvajal Altamiranda, Silvia Aparicio Obregon, Isabel De La Torre Díez, And Imran Ashraf. Pneumonia Detection Using Chest Radiographs with Novel EfficientNetV2L Model. 2024 (10.1109/ACCESS.2024.3372588)
- [2] Adnan Hussain, Sareer Ul Amin, Hunjoo Lee, Asma Khan, Noreen Fayyaz Khan, And Sanghyun Seo. An Automated Chest X-Ray Image Analysis for Covid (10.1109/ICECCT.2019.8869364) 19 and Pneumonia Diagnosis Using Deep Ensemble Strategy. 2023 (10.1109/ACCESS.2023.3312533)
- [3] Amrita Tripathi, Tripty Singh, Rekha R. Nair, And Prakash Duraisamy. Improving Early Detection and Classification of Lung Diseases With Innovative MobileNetV2 Framework. 2024 (10.1109/ACCESS.2024.344057)
- [4] Mohammad Yaseliani, Ali Zeinal Hamadani, Abtin Ijadi Maghsoodi, And Amir Mosavi. Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers. 2022 (10.1109/ACCESS.2022.3182498)
- [5] Hao Ren, Aslan B. Wong, Wanmin Lian, Weibin Cheng, Ying Zhang, Jianwei He, Qingfeng Liu, Jiasheng Yang, Chen Jason Zhang, Kaishun Wu, And Haodi Zhang. Interpretable Pneumonia Detection by Combining Deep Learning and Explainable Models with Multisource Data. 2021 (10.1109/ACCESS.2021.3090215)
- [6] Jianpeng Zhang, Yutong Xie, Graduate Student Member, Guansong Pang, Zhibin Liao, Johan Verjans, Wenxing Li, Zongji Sun, Jian He, Yi Li, Chunhua Shen, Member, and Yong Xia. Viral Pneumonia Screening on Chest X-Rays Using Confidence-Aware Anomaly Detection. 2021 (10.1109/TMI.2020.3040950)

- [7] N. SHILPA, W. AYEESHA BANU, AND PRAKASH B. METRE. Revolutionizing Pneumonia Diagnosis: AI-Driven Deep Learning Framework for Automated Detection From Chest X-Rays. 2024 (10.1109/ACCESS.2024.3498944)
- [8] Yeongbong Jin, Woojin Chang, And Bonggyun Ko. Generating Chest X-Ray Progression of Pneumonia Using Conditional Cycle Generative Adversarial Networks. 2023 (10.1109/ACCESS.2023.3305994)
- [9] Muhammad E. H. Chowdhury, Tawsifur Rahman, Amith Khandakar, Rashid Mazhar, Muhammad Abdul Kadir, Zaid Bin Mahbub, Khandakar Reajul Islam, Muhammad Salman Khan, Atif Iqbal, Nasser Al Emadi, Mamun Bin Ibne Reaz, And Mohammad Tariqul Isla. Can AI Help in Screening Viral and COVID-19 Pneumonia? 2020 (10.1109/ACCESS.2020.3010287)
- [10] T. S. Arulananth, S. Wilson Prakash, Ramesh Kumar Ayyasamy, V. P. Kavitha, P. G. Kuppusamy, And P. Chinnasamy. Classification of Paediatric Pneumonia Using Modified DenseNet-121 Deep-Learning Model. 2024 (10.1109/ACCESS.2024.3371151)
- [11] Dimpny Varshni, Rahul Nijhawan, Kartik Thakral, Ankush Mittal, Lucky Agarwal. Pneumonia Detection Using CNN based Feature Extraction. (10.1109/ICECCT.2019.8869364)
- [12] Keegan Kosasih, Udantha R. Abeyratne , Vinayak Swarnkar, and Rina Triasih. Wavelet Augmented Cough Analysis for Rapid Childhood Pneumonia Diagnosis. 2015 (10.1109/TBME.2014.2381214)
- [13] Wasif Khan, Nazar Zaki, And Luqman Ali. Intelligent Pneumonia Identification From Chest XRays: A Systematic Literature (10.1109/ACCESS.2021.3069937)
- [14] Review. 2021 Yaoming Lai, Guangming Li, Dongmei Wu, Wanmin Lian, Cheng Li, Junzhang Tian, Xiaofen Ma, Hui Chen, Wen Xu, Jun Wei, Yaqin Zhang, And Guihua Jiang. 2019 Novel Coronavirus-Infected Pneumonia on CT: A Feasibility Study of Few-Shot Learning for Computerized Diagnosis of Emergency Diseases. 2020 (10.1109/ACCESS.2020.3033069)
- [15] Hengyuan Kang, Liming Xia, Fuhua Yan, Zhibin Wan, Feng Shi , Huan Yuan, Huiting Jiang, Dijia Wu, He Sui, Changqing Zhang , and Dinggang Shen. Diagnosis of Coronavirus Disease 2019 (COVID-19) With Structured Latent Multi-View Representation Learning. 2020 (10.1109/TMI.2020.2992546)

