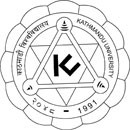
INTEROPERABLE SYSTEM FOR PNEUMONIA DIAGNOSIS USING MACHINE LEARNING, FUZZY LOGIC, DICOM AND FHIR



by

**Sulav Baral (032109-22)**

Masters in Health Informatics

A Thesis

Submitted to

Department of Health Informatics

Kathmandu University

Under the supervision of

Rabindra Bista, Ph.D.

Associate Professor

In Partial Fulfillment of the Requirements

for the Degree of

Masters in Health Informatics

November 25th, 2024

# THESIS EVALUATION

This thesis, submitted by **Sulav Baral** in partial fulfillment of the requirements for the Master of Health Informatics Degree from Kathmandu University, has been read by the Faculty Evaluation Committee under whom the work has been done and is now approved.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Rabindra Bista, Ph.D.

Associate Professor

Department of Computer Science and Engineering

Kathmandu University

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Er. Milan Chikanbanjar

Senior Lecturer

Khwopa Engineering College

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Santosh Khanal

Acting Head of Department

Department of Health Informatics

Kathmandu University

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Engineering at Kathmandu University and is now approved.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Prof. Dr. Manish Pokharel

Dean

School of Engineering

Kathmandu University

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Date:

# PERMISSION

Title: Interoperable System for Pneumonia Diagnosis using Machine Learning, Fuzzy Logic, DICOM and FHIR

Department: Health Informatics

Degree: Master in Health Informatics

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from Kathmandu University-School of Engineering, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the supervisors who supervised my thesis work or, in his (or her) absence, by the head of the department or the dean of the School of Engineering. It is understood that any copying publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and Kathmandu University for any scholarly use that may be made of any material in my thesis.

Sulav Baral

Date:

# ACKNOWLEDGMENTS

First and foremost, I would like to thank everyone who has helped me in one way or another to complete this thesis, “Interoperable System for Pneumonia Diagnosis using Machine Learning, Fuzzy Logic, DICOM and FHIR”. I owe special thanks to my academic advisor, Prof. Dr. Rabindra Bista sir, who has been able to give his opinion about this project constructively.

I would also like to thank the faculty members of the Department of Health Informatics, especially Mr. Sanjog Sigdel sir, whose experience in Artificial Intelligence and Medical Image Processing has greatly assisted me throughout the research. Their initial comments helped shape the intended research's focus and design ahead of its implementation.

Finally, I appreciate the academic resources provided by the Kathmandu University Library. They were instrumental in retrieving important reference materials and will continue to aid the research as it is further developed.

This report is the first step in a process that I hope will lead to fruitful cooperation in the healthcare industry.

# ABSTRACT

Pneumonia is recognized as the foremost infectious disease which is responsible for high mortality rate across all age demographics. As per WHO (World Health Organization), the age group prone to Pneumonia are children under the age of 5 and adults above the age of 65. Early diagnosis of Pneumonia is essential in order to reduce its global spread. For this, CXR (Chest X-Ray) is seen as the most reliable radiology image as its mostly useful in emergency situations. This work provided an innovative AI (Artificial Intelligence) based solution to Pneumonia diagnosis by combining Convolution Neural Network (CNN) with Capsule Network (CapsNet) and Fuzzy logic, generating a system that is accurate and interpretable. The combination of Explainable AI (XAI) via Grad-CAM for CXR, and an expert system-based Symptom Diagnosis with Fuzzy rules allowed transparency in the AI’s decision-making process. Furthermore, the data-sharing gaps in healthcare was solved by populating AI-based results with additional diagnosis metadata into a Digital Imaging and Communication in Medicine (DICOM) file, which can also be converted into Fast Healthcare Interoperability Resources (FHIR) compliant format as JavaScript Object Notation (JSON) to enable seamless interoperability across clinical systems with different message sharing systems. This study improved the accuracy of pneumonia diagnosis and enabled smooth interoperability among different clinical systems, enhancing access to healthcare data. By using TensorFlow’s DenseNet201 model and CapsNet, this system accurately classified CXR images as Normal or Pneumonia with 92.31% accuracy with an Area Under Curve (AUC) score of 0.97 and an overall time complexity of O(n2). By dynamically adjusting Fuzzy membership values, as well as providing a feature to input custom Pneumonia symptoms to incorporate with ever-changing Pneumonia symptoms, the system achieved a Mean Absolute Error (MAE) of 0.135 and 97% accuracy and Root Mean Square Error (RMSE) of 0.175 and 96.1% accuracy, indicating that the symptom evaluations remained responsive and reliable.

Keywords: *Pneumonia diagnosis, Chest X-ray, Convolution Neural Network, Capsule Network, Fuzzy logic, Explainable AI, DICOM, FHIR*

Table of Contents

[THESIS EVALUATION ii](#_Toc183848426)

[PERMISSION iii](#_Toc183848427)

[ACKNOWLEDGMENTS iv](#_Toc183848428)

[ABSTRACT v](#_Toc183848429)

[LIST OF FIGURES viii](#_Toc183848430)

[LIST OF TABLES ix](#_Toc183848431)

[LIST OF ABBREVIATIONS x](#_Toc183848432)

[CHAPTER 1: INTRODUCTION 1](#_Toc183848433)

[1.1 Introduction to AI-based Pneumonia diagnosis 1](#_Toc183848434)

[1.2 Problem Statement 2](#_Toc183848435)

[1.3 Motivation 2](#_Toc183848436)

[1.4 Objectives 2](#_Toc183848437)

[CHAPTER 2: LITERATURE REVIEW 3](#_Toc183848438)

[2.1 Related Work 3](#_Toc183848439)

[2.1.1 Convolution Neural Network for CXR Diagnosis 3](#_Toc183848440)

[2.1.2 Explainable AI for Transparency 5](#_Toc183848441)

[2.1.3 Fuzzy Logic Integration for Symptom-based Diagnosis 6](#_Toc183848442)

[2.1.4 DICOM for Medical Data Storage and Interoperability 8](#_Toc183848443)

[2.1.5 HL7 for Medical Data Interoperability 9](#_Toc183848444)

[2.1.6 Integration of HL7- FHIR with DICOM-SR 10](#_Toc183848445)

[2.2 Gap Analysis 11](#_Toc183848446)

[CHAPTER 3: METHODOLOGY 12](#_Toc183848447)

[3.1 Research Methodology 12](#_Toc183848448)

[3.2 System Design 13](#_Toc183848449)

[3.3 Clinical System Flow 15](#_Toc183848450)

[3.4 Data Pre-Processing 16](#_Toc183848451)

[3.4.1 Data Acquisition 16](#_Toc183848452)

[3.4.2 Contrast Limited Adaptive Histogram Equalization (CLAHE) 17](#_Toc183848453)

[3.4.3 Random Contrast Adjustment 18](#_Toc183848454)

[3.4.4 Normalization and Resizing 18](#_Toc183848455)

[3.4.5 Data Augmentation 18](#_Toc183848456)

[3.4.6 Summary of Data Pre-Processing 19](#_Toc183848457)

[3.5 Machine Learning 19](#_Toc183848458)

[3.5.1 Capsule Network 19](#_Toc183848459)

[3.5.2 TensorFlow Architecture to Implement Transfer Learning 21](#_Toc183848460)

[3.6 Machine Learning Evaluation and Prediction 23](#_Toc183848461)

[3.6.1 Performance Evaluation 23](#_Toc183848462)

[3.6.2 XAI Integrated Classification with Fuzzy Sets 25](#_Toc183848463)

[3.6.3 Fuzzy Rules with Dynamic Membership Adjustment 26](#_Toc183848464)

[3.7 Creating DICOM file and Conversion to FHIR 29](#_Toc183848465)

[CHAPTER 4: RESULTS 34](#_Toc183848466)

[4.1 Simulation Environment 34](#_Toc183848467)

[4.1.1 Libraries Used 34](#_Toc183848468)

[4.2 Component Testing 35](#_Toc183848469)

[4.2.1 Component Testing of Image Transfer 35](#_Toc183848470)

[4.2.2 Component Testing of FHIR Structure 36](#_Toc183848471)

[4.3 Pneumonia Diagnosis with Dense201 and CapsNet 39](#_Toc183848472)

[4.4 Dynamic Adjustment of Fuzzy Based Symptoms 39](#_Toc183848473)

[4.5 System Efficiency Analysis 40](#_Toc183848474)

[4.5.1 Time and Space Complexity of Trainable Machine File 40](#_Toc183848475)

[4.5.2 Time & Space Complexity of Dynamic Membership Adjustment Algorithm 42](#_Toc183848476)

[4.6 Machine Learning Accuracy Validation 43](#_Toc183848477)

[CHAPTER 5: CONCLUSION AND FUTURE WORK 45](#_Toc183848478)

[REFERENCES 47](#_Toc183848479)

# LIST OF FIGURES

[Figure 3. 1 Design Science Research Framework 12](#_Toc183367503)

[Figure 3. 2 System Design of Purposed Pneumonia Diagnosis Application 14](#_Toc183367504)

[Figure 3. 3 Pneumonia Diagnosis and Follow-up Workflow 15](#_Toc183367505)

[Figure 3. 4 Snapshot of Chest-Xray images collected for Classification 17](#_Toc183367506)

[Figure 3. 5 Individual Capsules Consisting of Image Features from CXR Image 21](#_Toc183367507)

[Figure 3. 6 Hybrid DenseNet201 and Capsule Network implemented in TensorFlow 22](#_Toc183367508)

[Figure 3. 7 Confusion Matrix for CXR Classification 23](#_Toc183367509)

[Figure 3. 8 AUC Curve from Pneumonia Diagnosis 24](#_Toc183367510)

[Figure 3. 9 Implemented GRAD-CAM in Pneumonia Diagnosis 25](#_Toc183367511)

[Figure 3. 10 Snapshot of Pneumonia Symptoms Data Frame with Membership Values 28](#_Toc183367512)

[Figure 3. 11 Flow Diagram of DICOM Creation 30](#_Toc183367513)

[Figure 3. 12 Snapshot of Orthanc Web Server Which Handles DICOM Files 31](#_Toc183367514)

[Figure 4. 1 Comparison Between Original and DICOM Images 36](#_Toc182514614)

[Figure 4. 2 Extracted DICOM Metadata and Mapped FHIR Values 37](#_Toc182514615)

[Figure 4. 3 Successful Validation of JSON Structure in Simplifier.net 38](#_Toc182514616)

[Figure 4. 4 MAE and RMSE of each Fuzzy Based Pneumonia Symptom 39](#_Toc182514617)

# LIST OF TABLES

[Table 2. 1 Previous Studies utilizing Dense Architecture for Disease Diagnosis 4](#_Toc182474423)

[Table 2. 2 Comparison of Fuzzy based Symptom Adjustment with Previous Works 7](#_Toc182474424)

[Table 3. 1 Data Augmentation Techniques with Specified Value and its Meaning 18](#_Toc183075239)

[Table 3. 2 Capsule Network Parameter Description with Values 20](#_Toc183075240)

[Table 3. 3 Accuracy Comparison of Hybrid Classification Approach 21](#_Toc183075241)

[Table 3. 4 Fuzzy Membership Range for Pneumonia Symptoms 26](#_Toc183075242)

[Table 3. 5 Severity Impact of Fuzzy Rules on Pneumonia Symptoms 27](#_Toc183075243)

[Table 3. 6 DICOM Metadata and its Respective Tags 31](#_Toc183075244)

[Table 3. 7 Mapping of DICOM Metadata Fields to FHIR Resources 32](#_Toc183075245)

[Table 4. 1 Hardware and Software Specifications 34](#_Toc184029830)

[Table 4. 2 Comparison of Membership Adjustment Algorithm with Previous Works 40](#_Toc184029831)

[Table 4. 3 Time Complexity Analysis using Big O Notation 41](#_Toc184029832)

[Table 4. 4 Space Complexity Analysis using Big O Notation 42](#_Toc184029833)

[Table 4. 5 Time Complexity of Dynamic Membership Adjustment Algorithm 43](#_Toc184029834)

[Table 4. 6 Space Complexity of Dynamic Membership Adjustment Algorithm 43](#_Toc184029835)

[Table 4. 7 Comparison of Validation Accuracy with Different DenseNet Variants 43](#_Toc184029836)

[Table 4. 8 Accuracy Comparison of DenseNet201 with Different Train to Test Ratio 44](#_Toc184029837)

# LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
| Short Form | Meaning |
| AI | Artificial Intelligence |
| AUC | Area Under the Curve |
| CAD | Computer Aided Design |
| CapsNET | Capsule Neural Network |
| CDA | Clinical Document Architecture |
| CLAHE | Contrast-Limited Adaptive Histogram Equalization |
| CNN | Convolution Neural Network |
| CSS | Cascading Style Sheets |
| CXR | Chest X-Ray |
| CXR | Chest X-ray |
| DICOM | Digital Imaging and Communications in Medicine |
| DSR | Design Science Research |
| EHR | Electronic Health Records |
| FHIR | Fast Healthcare Interoperability Resources |
| Grad-CAM | Gradient-Weighted Class Activation Mapping |
| HIS | Hospital Information System |
| HL7 | Health Level 7 |
| HTML | Hypertext Markup Language |
| JPEG | Joint Photographic Experts Group |
| JSON | JavaScript Object Notation |
| LLM | Large Language Model |
| ML | Machine Learning |
| NEMA | National Electrical Manufacturers Association |
| NLP | Natural Language Processing |
| QIDO-RS | Query based on ID for DICOM Objects |
| RGB | Red, Green and Blue |
| SNOMED | Systematized Nomenclature of Medicine–Clinical Terminology |
| SVG | Scalable Vector Graphics |
| WADORS | Web Access to DICOM Objects |
| XAI | Explainable Artificial Intelligence |

# CHAPTER 1: INTRODUCTION

This section covers the overview of AI (Artificial Intelligence) based diagnosis, and the of Pneumonia and its application in Medical Imaging Technology such as DICOM (Digital Imaging and Communications in Medicine), and also the importance of interoperability standards such as FHIR (Fast Healthcare Interoperability Resources) in data exchange, following the issue related to such methods in modern settings and the objective that concerns the research question as well as the motivation behind this proposed research in the vicinity of Pneumonia Diagnosis.

## 1.1 Introduction to AI-based Pneumonia diagnosis

As per WHO, Pneumonia continues to pose as a global threat, particularly between the age group of children under the age of 5 and adults above 65 (McAllister et al., 2019). Early diagnosis of pneumonia is critical to ensure proper treatment and increase survival rates, particularly in vulnerable populations such as Children and the elderly (Meedeniya et al., 2022). Chest X-ray imaging is the most frequently used method for diagnosing pneumonia (Gambato et al., 2023); however, the examination of chest X-rays is a challenging task and is prone to subjective variability (Kiran et al., 2024) . Despite significant development of AI-based diagnosis, Pneumonia continues to pose. For example, in 2019, pneumonia was responsible for over 700,000 deaths in children under five, particularly in vulnerable regions, giving importance to preventative measures and timely medical intervention (Cilloniz et al., 2023; Sharrow et al., 2023). There is a need for an effective AI-based diagnosis that provides transparency in decision-making through Explainable Artificial Intelligence (XAI) and can track major pneumonia symptoms that are responsible for the high mortality rate (Chumbita et al., 2020; Sheu et al., 2023). Moreover, AI-based diagnostics must not only be limited to transparent decision-making but also have support for seamless data interoperability to exchange records between different Healthcare institutes through known standards such as FHIR (Labkoff et al., 2024), especially when medical images are stored in DICOM format due to the reason that FHIR is flexible enough to share and integrate data from any AI-based diagnosis with other clinical systems (Ait Abdelouahid et al., 2023).

## 1.2 Problem Statement

The healthcare industry currently lacks adaptation to the ever-growing symptoms of Pneumonia that differs from person to person based on geographical locations, and also lacks transparency in diagnosis. Furthermore, existing AI based systems are limited in their capability to give clear feedback on how a certain decision was made, especially when based on a complex AI model. This gap between Doctors and AI-based Diagnosis becomes a challenge when attempting to understand the reason behind diagnosis (Ambaliya et al., 2025). Additionally, medical institutions often lack unified interoperability and still rely on legacy systems that limits reliable communication resulting in poor patient care, even though Image storage technology such as DICOM is on the rise (Shivshankar et al., 2024).

## 1.3 Motivation

Following the transparency problem of AI-based diagnosis of Pneumonia and the seamless exchange of medical images, the research question that arises is “How can an AI-based, interoperable system utilizing Explainable AI and Fuzzy logic improve the accuracy, transparency, and integration of chest X-ray diagnoses for pneumonia in real-time clinical settings?”. This research is driven by the aim of not only improving Pneumonia Diagnosis transparency, but also to ensure compatibility between FHIR and DICOM for improved patient outcomes in practical clinical applications.

## 1.4 Objectives

* To implement Pneumonia Diagnosis using suitable ML algorithms which yields higher accuracy
* To investigate method for tackling biasness of AI based models in determining Pneumonia severity level
* To implement FHIR-based interoperability for seamless integration of DICOM metadata

# CHAPTER 2: LITERATURE REVIEW

The growing trend of AI-based disease diagnosis demands an effective but reliable system to provide accurate and transparent results, especially for Pneumonia, one of the major causes of high mortality rates among adults and children. The literature here will discuss the data collection and previous research in Diagnosis.

## 2.1 Related Work

The manual diagnosis of Pneumonia from chest radiography has been tedious due to overlapping symptoms with other respiratory conditions and the inter-observer variability among radiologists (Elemraid et al., 2014; Hopstaken et al., 2004). AI can improve diagnostic accuracy and speed by leveraging computer-aided diagnosis (CAD) and deep learning algorithms to support radiologists in identifying important indicators on X-rays more effectively (Kwon et al., 2021), and can enable early intervention through more accessible, accurate diagnostics​. Furthermore, the use of XAI is especially important in medical imaging where understanding the specific image features that contribute to a pneumonia diagnosis can help clinicians to understand the reason behind the diagnosis, thus gaining wider acceptance in clinical settings (Ghuse & Monga, 2024). By mapping DICOM metadata to FHIR standards, healthcare providers can achieve seamless communication and integration of imaging data (Kamel & Nagy, 2018).

### 2.1.1 Convolution Neural Network for CXR Diagnosis

Different radiologists may interpret X-ray images differently, leading to inconsistent diagnoses and affecting reliability. Studies reveal that diagnosing pneumonia in outpatient settings is challenging due to the lack of Chest X-rays. Signs like crackling sounds, fever, chest pain, shortness of breath, fast heart and breathing rates, and a runny nose are helpful, yet Doctors are more likely to order a Chest X-ray if a patient has these symptoms, despite 35% of pneumonia cases having normal X-rays (Evertsen et al., 2010; Wootton & Feldman, 2014). Due to such variation in Diagnosis, Unsupervised learning methods in AI such as Convolution Neural Network (CNN) have been widely preferred for Chest X-ray (CXR) to distinguish between Healthy and Pneumonia.

A notable work in Pneumonia detection is the use of CNN’s DenseNet-169 model for feature extraction and SVG as a classifier, achieving an AUC of 0.8002 (Varshni et al., 2019). Similarly, the work of (Rahman et al., 2020) utilized four pre-trained CNN models (AlexNet, ResNet18, DenseNet201, SqueezeNet) to classify 5247 chest X-ray images of normal, bacterial pneumonia, and viral pneumonia proved DenseNet201 being the most effective CNN model for CXR for achieving 98% accuracy. In the use of pre-trained CNNs such as Xception, InceptionResNet, and DenseNet for Pneumonia classification, DenseNet also provides equivalent accuracy (Ravi et al., 2023). Over time, DenseNet seems reliable for AI-based Diagnosis as seen in the work of (Sanghvi et al., 2023) which achieved 99.1% accuracy. Furthermore, COVID-19 and Pneumonia detection also utilized Dense201 getting an accuracy of 95.34%, as in the works of (Chutia et al., 2024). Table 2.1 below discusses related works in Radiology diagnosis that utilize DenseNet model:

Table 2. 1 Previous studies Utilizing Dense Architecture for Disease Diagnosis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Preferred Model | Target Disease | Accuracy | Remarks |
| (Kundu et al., 2021) | DenseNet-121 | Pneumonia | 98.81% | Achieved high accuracy; used data augmentation to improve model generalization. |
| (Rochmawanti & Utaminingrum, 2021) | DenseNet-121 | Tuberculosis, Pneumonia, Cardiomegaly, COVID-19 | Tuberculosis: 89.2%, Pneumonia: 90.4%, Cardiomegaly: 89.8%, COVID-19: 98.6% | Best accuracy achieved with 224x224 image resolution; Global Average Pooling and dropout used to mitigate overfitting; batch normalization accelerates training |
| (Naronglerdrit et al., 2021) | DenseNet-201 | COVID-19 | Multiclass classification: 96.76%,  Binary classification:  99.9% | DenseNet-201 performed well for the identification of COVID-19 cases compared to other models. |
| (Anakha et al., 2021) | DenseNet201 | COVID-19 | 96.54% | DenseNet201 outperformed other models with the best accuracy for COVID-19 detection from X-rays |
| (Sanghvi et al., 2023) | DenseNet201 | COVID-19 and Pneumonia | 99.1% | High accuracy (99.1%), sensitivity (98.5%), and specificity (98.95%) was achieved via DenseNet201 |
| (Chutia et al., 2024) | DenseNet201 | Lung diseases: Pneumothorax, and Atelectasis | 95.34% | DenseNet201 proved efficient compared to other models |

However, prior research on deep CNNs for chest X-ray diagnosis has mainly focused on binary classification tasks, detecting specific diseases, and overlooking the need for multitask classification. Small, biased datasets have hindered model generalizability and accuracy, emphasizing the necessity for more extensive, diverse data. The area under the curve metrics is crucial for a more comprehensive understanding of deep learning models' capabilities in chest X-ray diagnosis, as presented by (Mann et al., 2023). For a disease like pneumonia that is significantly growing, the pre-trained CNN models have shown remarkable metrics but are lacking transparency in terms of ML-based decision-making when performing diagnosis, leaving the Healthcare personnel in a dilemma (Shrimali, 2024; J. Tang et al., 2024).

### 2.1.2 Explainable AI for Transparency

Normally, Pneumonia diagnoses begin with chest X-ray exams by specialists, followed by reports to clinicians for conclusions based on analysis and symptoms. The process can be complex, leading to disagreements. Symptoms vary based on causes and patients, with rapid changes in disease conditions adding to detection challenges as stated by (Ruuskanen et al., 2011; X. Wang et al., 2017). A study by (Ren et al., 2021) had used over 35,000 pneumonia cases from the CheXpert image dataset, and used CNNs DenseNet121 with Bayesian Network.

Other than Pneumonia symptoms, it’s necessary for healthcare personnel to understand the transparency behind the Chest X-ray classification to ensure accurate and reliable patient care, as notable in the works of (Ahmed et al., 2022; Ihongbe et al., 2024; S. Roy et al., 2023; Zou et al., 2023). Also in the work of (Ullas & S, 2024), the ResNet50 model achieved an accuracy of 81.41% on Chest X-rays whose result is further clarified by the use of GRAD-CAM heatmaps which produced distinct visualization between a Normal and Pneumonia category, where the researchers conclude such an approach saves time for clinicians, enhancing patient care in healthcare settings.

This sort of practice will obliviously enhance decision-making. Still, when it comes to becoming more dependent on X-rays, the clinical setting will lack diversity in decision-making as Grad-CAM indeed offers visual insights, but doesn’t fully explain how clinical symptoms influence each prediction, which could limit clinicians’ trust and usability in critical cases (Höhl et al., 2024). On the other hand, the dependence on pre-defined symptoms means that rapid changes in disease conditions add more complexity when the ML model fails to adapt to new symptoms.

### 2.1.3 Fuzzy Logic Integration for Symptom-based Diagnosis

A fuzzy set is a class of objects with a chain of grades between zero and one that is based on mathematical operations like inclusion, union, intersection, complement, relation and convexity (Zadeh, 1965). Medical applications have utilized Fuzzy sets or rules to govern the output of Pneumonia levels such as low, mild, or severe. Rather than binary values, medical symptoms are expressed as degrees of membership rather than strict binary values, so that the diagnostic process can be performed iteratively. “The system can suggest further investigations based on preliminary diagnoses thus making it useful in overlapping symptoms”, as explained in the work (Adlassnig, 1980).

Fuzzy Logic follows the concept of an Expert System, as mentioned in the work of (Shi et al., 1999), and uses human-based knowledge to make diagnosis decisions (Konar & Jain, 2001). The biggest advantage of combining Fuzzy with ML based diagnosis is that while the latter identifies patterns in the medical data, the former handles imprecise information, as mentioned in the research of (Akerkar & Sajja, 2010; Gorzalczany & Piasta, 1999; Senol & Yildirim, 2010). Going forward, (Hasan et al., 2010) developed a web-based fuzzy expert system that also diagnosed Pneumonia. The users can input symptoms, and the system uses fuzzy rules to calculate disease probabilities, considering major symptoms and patient history as catalyst factors, and is similar to the works of (Biswas, 2010; Chinniah & Muttan, 2010; Santra et al., 2019). While Fuzzy logic provides good interpretability of symptoms via the utilization of human-based decision-making, it isn’t standalone effective in the clinical setting, regarding transparency of ever-growing Pneumonia symptoms.

High symptom variability in studies can complicate defining membership functions and rules, leading to inconsistencies and inaccuracies in diagnosis. So, it’s necessary to combine Fuzzy with ML (Machine Learning) methods for pattern analysis. Such an approach, as in the works of (Salem et al., 2022), uses TFKNN (Tuned Fuzzy K-Nearest Neighbors) for the Diagnosis of diabetes to estimate the likelihood of each instance belonging to a specific class, allowing the model to handle uncertain cases more effectively, and the result proved that TFKNN outperformed other ML models with an F1 score of 93.18%, with an overall accuracy of 90.63%, also demonstrating a higher specificity (85.00%) and an average AUC of 94.13%. Also mentioned by (Zheng et al., 2022) whose work was concerned with a survey on deep learning and fuzzy systems, has highlighted the need for high-quality data to prevent overfitting and the need for precise membership functions, such as Fuzzy rules, in complex situations.

However, as radiology images are prone to noise, the accuracy of past models doesn’t guarantee valid results. Even deep learning models like CNN-based architectures lack spatial attention mechanisms to focus on critical regions in chest X-rays, and upon considering this drawback, (A. Roy et al., 2024) proposed a Fuzzy Attention-aided Deep Neural Network (FA-Net), utilized attention modules to highlight spatial regions critical for classification of Pneumonia severity with remarkable accuracy and AUC of 94.28% and 97.89% for binary classification (Normal or Pneumonia) and AUC of 90.94% for multi-class Pneumonia classification, outperforming previous transfer learning models.

Regardless of advanced ML algorithms and significant accuracy, the Fuzzy alone can’t adapt to the ever-growing symptoms as the dataset can vary from region to region several studies on the integration of fuzzy logic with deep learning in medical image analysis reveal both its potential and limitations. A review by (Lu et al., 2024) mentions that while Fuzzy logic improves interpretability in medical AI, it faces challenges in adaptability, data dependency, and rule complexity, making it less effective in dynamic clinical settings. To tackle this issue, this work has utilized a dynamic membership adjustment with contextual weighting, which adjusts the Pneumonia symptom as per the trend and the recorded membership values. The table below discusses the previous research on such algorithms.

Table 2. 2 Comparison of the Implemented Dynamic Adjustment Algorithm with Previous Works

|  |  |  |
| --- | --- | --- |
| Study | Algorithm used | Limitations |
| (Chandra & Bhardwaj, 2024) | Altered Type II fuzzy membership function for the enhancement of medical images | Primarily centered on improving image quality; it may not specifically tackle real-time modifications to membership in diagnostic algorithms. |
| (Li et al., 2023) | Accelerated Fuzzy C-Means with affinity filtering and membership scaling | The concern was on clustering task, which may not be directly relevant for adjusting memberships dynamically in control systems |
| (Lagunes et al., 2020) | Firefly Algorithm with dynamic parameter adjustment using fuzzy logic | Risk of early convergence; necessitates adjustment of fuzzy rules |
| (Cerrada et al., 2002) | Dynamically adaptive fuzzy modeling utilizing parameter tuning based on gradient descent | May lead to local minima; necessitates careful choice of learning rates |

The proposed algorithm in this work focuses on the need for real time membership adjustments in Pneumonia medical diagnosis. The use of exponential smoothing along with contextual weighting avoids problem such as early convergence and parameter sensitivity. Furthermore, adapting percentile-based threshold values ensures the algorithm can handle noisy datasets. Therefore, this work offers a thorough solution for Pneumonia diagnostics where real time adaptability of symptoms become essential.

### 2.1.4 DICOM for Medical Data Storage and Interoperability

DICOM is a universal protocol for medical records that standardizes formats, protocols, and data structures that include services for image management over networks, enabling communication between systems from different manufacturers and promoting equipment selection based on features over compatibility, making it highly preferred for storage of radiology images. In 1982, NEMA and ACR collaborated on a medical image exchange standard, released in 1985, revised in 1988, and renamed "DICOM" in the latest version for digital imaging communication in medicine (Elion, 1995).

Going back, this protocol has proven reliable even to US Veterans Affairs (VA) for working with radiology images, making image acquisition, storage, and display seamless for clinicians, and enabling interoperability to connect devices from different manufacturers (Bidgood et al., 1997; Kuzmak & Dayhoff, 2000). The DICOM-SR,  introduced to improve the documentation of diagnostic images and waveforms, further extended the capabilities of the original DICOM with a focus on radiology reports to be integrated directly with patient imaging records, enhancing interoperability among healthcare organizations (Bidgood, 1998).

Concerning the need for remote availability of patient data, clinicians can upload DICOM files and instantly retrieve or populate patient demographics online, as one such web based DICOM application has been discussed in (Lee et al., 2005), with the only limitation referring to low internet bandwidth at that time which resulted in low quality image transfer or on-purpose loss in image quality. Similar web based DICOM application are mentioned in (Kulkarni et al., 2012; Monteiro et al., 2013) with a concern for data sharing flexibility between different operating systems where web technology such as HTML5 and JavaScript played a vital role in development phase.

DICOM has been preferred for CXR diagnosis due to good compatibility for Neural Network based diagnosis. One such work by (O’Quinn et al., 2019) utilized CXR datasets stored as DICOM file, and the accuracy achieved were remarkable, which concluded X-ray being fit enough for DICOM file. Similar other works that utilized DICOM standard for Neural Network based diagnosis (Joshi et al., 2021; Niehues et al., 2021), also concluded CXR being fit for storage and retrieval of medical data due to remarkable accuracy metrics.

The concern isn’t with the DICOM’s capability itself but regarding exchange of diagnostics metadata with other organizations who don’t have DICOM readers or those EHR (Electronic Health Record) with limited support. This limits radiologists’ ability to provide timely insights that can hinder decision making in a clinical setting. Also, DICOM-SR is way too static and isn’t as good as FHIR for real time interoperability in the scenario where clinicians may need rapid updates. For example, delays in radiology updates will lead to delays in important interventional procedures such as the case of trauma. In follow up imaging, real time information exchange is necessary to make decision based on latest information, such as infections or pulmonary diseases.

To bridge this gap, creating standardized tools to translate DICOM metadata into a standard interoperability framework enhances data exchange for better decision-making and ensures smooth workflow in emergencies.

### 2.1.5 HL7 for Medical Data Interoperability

HL7® (Health Level 7) sets global standards for transferring health data to enhance patient outcomes. Not all healthcare institutions utilize DICOM applications and the data exchange becomes complicated in real time scenarios, demanding a framework for integrating Radiology Information Systems (RIS) to enhance Healthcare Information Systems (HIS) by combining HL7 CDA (Clinical Document Architecture) and DICOM standards, as in the works of (Koncar, 2007), that effectively translated DICOM images and reported into CDA-compliant XML documents accessible through WADO, thus enabling modular and cost-effective HIS.

With concern to this work that utilizes AI-based diagnostic and reporting, HL7 lacked the necessary structure to handle complex imaging metadata due to lack of a linear message model, which doesn’t support DICOM-SR’s nesting of imaging metadata too well, losing reliability in emergency diagnosis, as mentioned in (Mildenberger et al., 2002; K. C. Wang et al., 2014).

Due to HL7’s incompatibility with the DICOM-SR and AI-based diagnosis application, the HL7- FHIR® (Fast Healthcare Interoperability Resources) was introduced as the most effective interoperability standard till date with support for imaging study modality.

### 2.1.6 Integration of HL7- FHIR with DICOM-SR

The HL7® FHIR® standard defines how healthcare information can be exchanged between different computer systems regardless of how it is stored in those systems (*Index - FHIR v5.0.0*, 2024), and is known for its flexibility and vast modalities in radiology. The FHIR ImagingStudy resource such as such as study.uid, series.uid, and instance.uid keeps the track of each imaging exam (CXR or MRI) with unique IDs for future accessibility (*ImagingStudy - FHIR v5.0.0*, 2024). The FHIR endpoint serves as a central access point that is connected to a DICOMWeb server for image retrieval through WADO-RS URLs (Web Access to DICOM Objects – RESTful).

The most notable work incorporating DICOMweb and FHIR (S.-T. Tang et al., 2023), explains that an effective medical imaging workflow starts with the HIS (Hospital Information System) creating a ServiceRequest on the FHIR server for an X-ray or MRI, following the RIS (Radiology Information System) which then schedules the appointment and informs the DICOMweb Server after creating Encounter resource on the FHIR server.

To reclassify, this work utilized AI based diagnosis of CXR for Pneumonia classification and the diagnosis metadata is stored as a DICOM file, which is later converted to FHIR structure for seamless interoperability. Previously, (Blezek et al., 2021) proposed a system called ROCKET that displayed AI-generated results within the clinical workflow where FHIR observation resource were used for quantitative measurements such as volume or area, and each observation was coded with relevant LOINC (Logical Observation Identifiers Names and Codes or SNOMED CT codes to indicate the meaning of the measurement, where fields such as valueQuantity and referenceRange documented the measurement values and provided context. But comparatively, the work lacked a broader approach as it was limited to Body composition analysis and followed a static approach instead of real-time adaptability, which is critical for tracking symptom progression.

Even though FHIR and DICOMWeb are optimized for storing and retrieving image data, they lack the necessary infrastructure to handle the system demands of AI-driven diagnosis. Advanced AI-based systems such as predictive modeling, image segmentation, and anomaly detection require special metadata that isn’t natively supported in the current FHIR-DICOM framework, which results in fragmented workflow as radiologists must use external tools for AI analysis, complicating data handling and the risk of delays in an emergency (Tabari et al., 2024).

## 2.2 Gap Analysis

While the advancements in AI-based Radiology diagnosis have shown remarkable results, the integration of such diagnosis with DICOM and FHIR remains a challenge in terms of providing real-time diagnosis and seamless interoperability in the field of radiology, especially for CXR. Indeed, the recent AI applications for Pneumonia diagnosis are capable of producing effective outputs but these models can't be called clinically reliable since they are mostly based on small and biassed datasets making the application on a diverse population quite difficult to implement. While Fuzzy rules have improved the diagnosis since they follow the concept of the Expert system making human-like decisions, and XAI trending for Transparency, the latter is still limited to image highlighting which doesn’t necessarily validate clinical outcomes. On the other hand, the DICOM, though it supports nested data structure for reliable information storage and retrieval, is not flexible for real-time Symptom analysis, and there is a need for an effective HL7-FHIR-based implementation that traces the metadata containing AI-based diagnosis, especially for the organizations with different interoperability standard.

This work addresses the critical gaps in AI-based Chest X-Ray (CXR) diagnosis by combining Dense201 and Capsule Network (CapsNet) with Fuzzy logic for a symptom-based evaluation, bypassing demerits of biased classifications, as it allows a human-like decision-making, and providing transparency to the clinicians, while incorporating DICOM for AI based medical data storage, and enabling seamless interoperability through FHIR.

# CHAPTER 3: METHODOLOGY

This part will discuss the approach and ideas behind the developed system, both the front end and the backend structure which illustrates how certain decisions were made.

## 3.1 Research Methodology

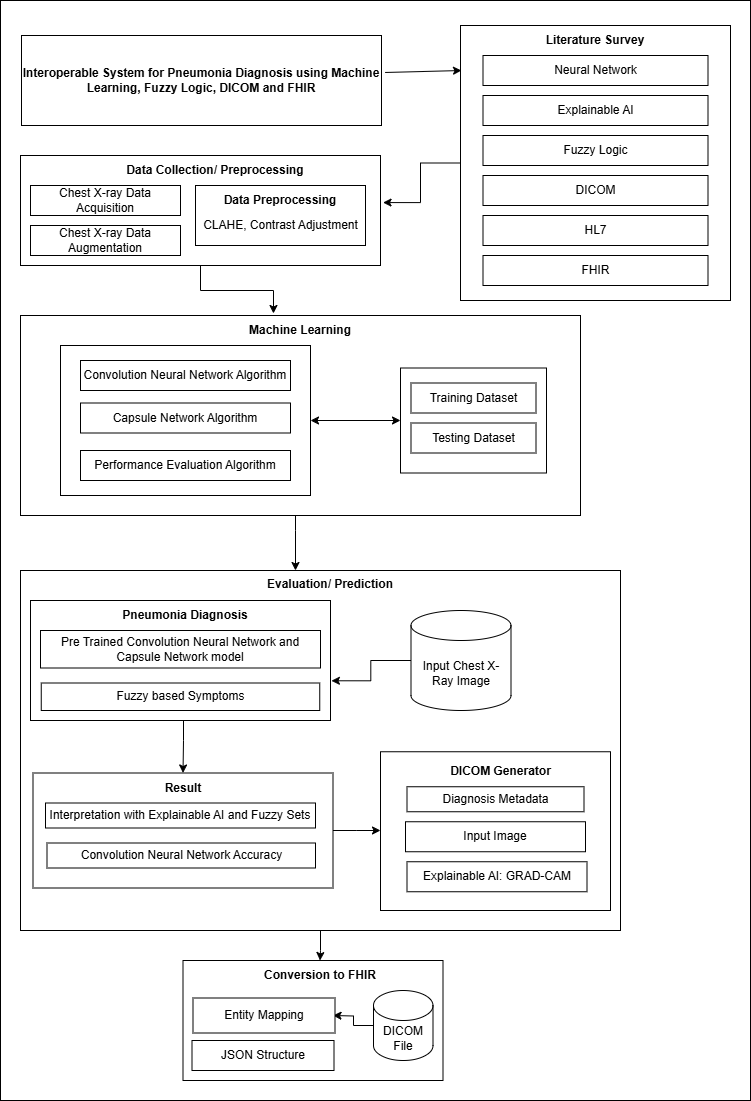


Figure 3. 1 Design Science Research Framework

The Design Science Research (DSR) framework, shown in Figure 3.1, highlights the methodology used in this work. Firstly, the literature review on Neural Networks, Explainable AI, Fuzzy Logic, DICOM, HL7, and FHIR was analyzed, mostly in the realm of AI-based CXR diagnosis for Pneumonia, after which the research problem was identified.

The second phase involved Data collection. This data consists of numerous CXR images categorized into Pneumonia and Normal. This dataset was split into training and testing sets and Image preprocessing technique such as Contrast Limited Adaptive Histogram Equalization (CLAHE), and Contrast Adjustment were applied following the Image Augmentation.

In third phase, the Machine Learning algorithm Convolution Neural Network (CNN) performed feature extraction on the training and testing dataset and the image features were transferred to Capsule Network for preserving the image’s spatial features for better accuracy. After the training was over, a model file consisting of learnable parameters was loaded for the evaluation/prediction part.

This file combined with the Fuzzy sets of input diagnosed an input CXR image as either Normal or Pneumonia category. An Explainable AI was implemented for both CNN and Fuzzy to visualize the diagnosis so as to address the gap of limited transparency. The accuracy scores, XAI Base64 image data and Diagnosis information were populated into a DICOM file, which can be converted to FHIR for achieving seamless interoperability.

## 3.2 System Design

This work utilized a combined approach for Pneumonia diagnosis, incorporating a hybrid CNN and CapsNet for CXR classification in either Normal or Pneumonia, whereas the user-based Pneumonia symptoms were further computed by Fuzzy rules for determining severity of Pneumonia as either low, mild or severe.

The web-based application for Pneumonia Diagnosis consisted of Initial Diagnosis service followed by DICOM converter and Post diagnosis service for the follow up. It also incorporated a DICOM to FHIR converter for achieving a JSON structure which is compliant with other Healthcare systems for Interoperability. Figure 3.2 below gives a basic overview of the purposed system architecture for this work.

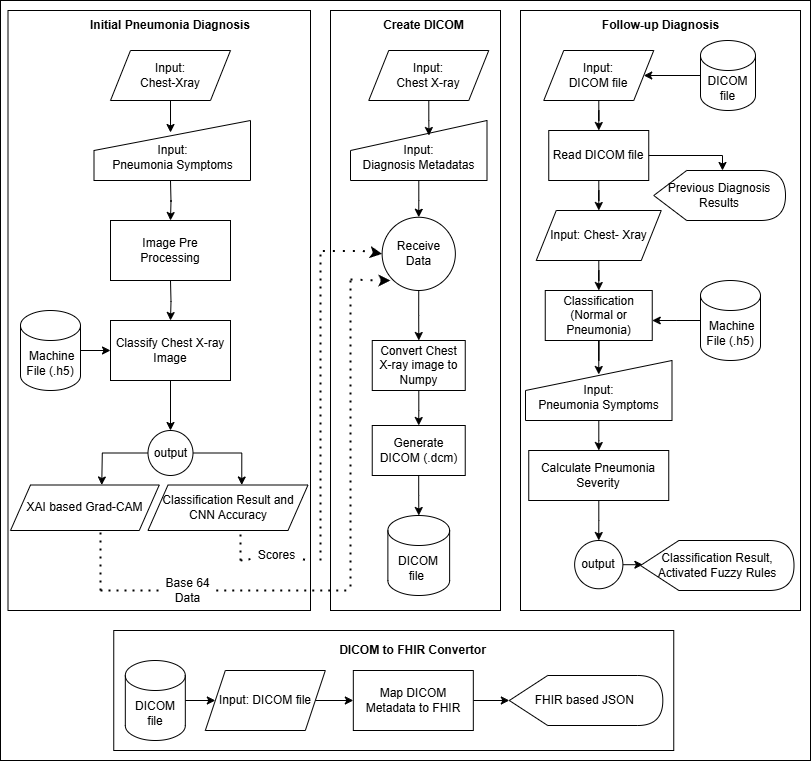


Figure 3. 2 System Design of Purposed Pneumonia Diagnosis Application

The initial Pneumonia Diagnosis provided a feature to add a chest X-Ray image and manual Fuzzy based symptoms based on a Patient’s condition. The Chest X-ray image went through Image Preprocessing operations which classified the image as either Normal or Pneumonia through a pre-trained Machine file containing all the necessary image features. The CNN-CapsNet based classifier and Fuzzy membership values together output the Patient’s condition as low mild or normal level of Pneumonia, but if the classification is Normal, then the output is as it is. Also, a GRAD-CAM of chest X-ray along with CNN accuracy was generated.

The CNN-CapsNet accuracy metrics and the Fuzzy membership values were passed as read only file to the DICOM generator module. The DICOM metadata corresponding to a Patient’s Pneumonia diagnosis information was input manually along with the original Chest X-Ray. Finally, a valid DICOM file was generated.

This generated DICOM was also retrieved for follow up phase, and the implemented mechanism, which is same as the initial diagnosis, computed the latest Diagnosis result based on new Symptom values, and the DICOM file was updated. This phase also incorporated transparency in Pneumonia diagnosis by generating the ML’s decision behind the output, which seemed clinically significant in healthcare scenario.

Additionally, the uploaded DICOM file was converted to FHIR for increasing flexibility with other healthcare institution. A generated JSON format contained all recorded values from the Pneumonia Diagnosis through the concept of entity mapping.

## 3.3 Clinical System Flow

This work is designed to incorporate with a typical Hospital or Clinical scenario. The flow begins when a Patient registers his condition after visiting the Radiology Department, eventually receiving a CXR for Initial Diagnosis. The Diagnostic System implemented in this work performs the AI based diagnosis, starting with Image Pre-Processing following the Convolution Neural Network based classification that also generated an XAI based GRAD-CAM that provides a computer-vision based visualization behind the AI’s decision. This system includes initial Fuzzy based symptoms that is input manually by the practitioner. All the study and Patient details are stored as a DICOM file, that can be retrieved later for follow up purpose. The patients who follow up with new condition provide information about their recovery or difficulty over a certain time period.

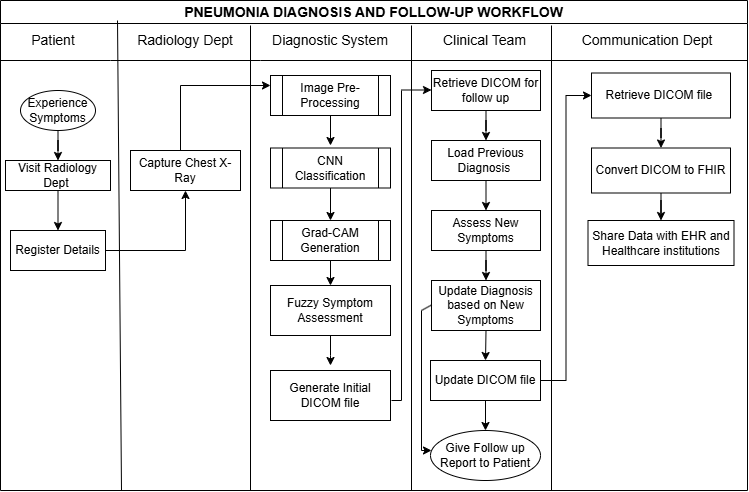


Figure 3. 3 Pneumonia Diagnosis and Follow-up Workflow

Upon completing a patient’s symptom, a new severity is determined by the Diagnostic System and the DICOM file is updated. This DICOM file can be retrieved by the Hospital’s communication team and can be converted to FHIR structure for achieving Interoperability.

## 3.4 Data Pre-Processing

The procedure of CXR data Acquisition for Pneumonia diagnosis, as well as image pre-processing and image data augmentation are as follows:

### 3.4.1 Data Acquisition

This research utilized the Mendely’s CXR image data from University of California San Diego by (Kermany et al., 2018). The dataset consisted of 5912 total images, out of which 5,216 images was split for training purpose which were divided into 1,341 Normal and 3,875 Pneumonia category.

The dataset was adjusted as Training to Testing to Validation in a ratio of 88:11:1. This split is a standard ML based practice to focus on training data and to ensure the model gets sufficient patterns and features to understand image patterns (Sivakumar et al., 2024).

The large proportion of training data (88%) provided diverse patterns which resulted in increased accuracy. The proportion of small testing data (11%) performed an unbiased evaluation between the two binary CXR image classes, avoiding overfitting in spite of large variable number of images.

Around 624 images were split for testing purpose, further divided into 234 Normal and 390 Pneumonia images. The remaining 72 images were left for validation purpose. These chest X-rays were part of routine clinical care provided to the pediatric patients between one and five years old, and were labeled according to diagnoses verified by two expert radiologists to guarantee high diagnostic accuracy by omitting the low-quality images or those with vast quantity of image noise generated during the follow up.

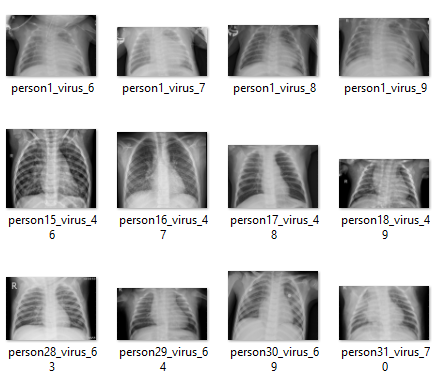


Figure 3. 4 Snapshot of Chest-X-ray images collected for Classification

Furthermore, the demographic details of these CXR images are given below:

* **Image Type**: anterior-posterior (AP) chest X-ray images
* **Number of Images**: 5192
* **Categories**: Normal and Pneumonia
* **Image Format**: JPEG
* **Image Resolution**: 72 dpi to 96 dpi
* **License**: Creative Commons Attribution 4.0 International (CC BY 4.0) license

### 3.4.2 Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE was applied to improve CXR image quality, especially in low-contrast areas of the CXR, utilizing Clip Limit of 2.0 and the image split into an 8 by 8 tile, to prevent over-amplification in the image region and the tile grid size to divide the image for applying CLAHE operation. The clip limit was calculated as:

After calculating Clip limit via Eq. 3.1, the mapping function calculates new intensity for each pixel in the tile, enhancing local contrast as:

Once each tile has been processed by Eq. 3.2, the tiles are blended together in order to form the final equalized image.

### 3.4.3 Random Contrast Adjustment

A random contrast adjustment was performed in the CXR image to improve generalization for Neural Network based diagnosis. The contrast factor was between the ranges 0.8-1.2. This was calculated as:

The adjusted image is returned by Eq.3.3

### 3.4.4 Normalization and Resizing

Normalization adjusts the pixel values in images to a similar range, typically [0,1]. For an 8-bit image whose maximum pixel value is 255, the Normalization (Eq 3.4) is performed as:

Normalization keeps pixel data’s small and consistent for a faster Neural Network operation. Also, both the training and testing images are rescaled to a resolution of 224 by 224 pixels to reduce training time and computational resource.

### 3.4.5 Data Augmentation

Data augmentation enhances performance of deep learning models by creating duplicate trainable images or mimicking real imaging conditions. This approach improves the diversity of synthetic images, especially CXR images which requires special attention, increasing accuracy as a result (Dong et al., 2022; Shen et al., 2023).

Table 3. 1 Data Augmentation Techniques with Specified Value and its Meaning

|  |  |  |
| --- | --- | --- |
| Parameter | Value/Range | Meaning |
| rotation\_range | 20 | Random image rotation by 20 degrees |
| width\_shift\_range | 0.2 | Horizontal image shift by 20% of its width |
| height\_shift\_range | 0.2 | Horizontal image shift by 20% of its height |
| shear\_range | 0.1 | Image slanting by 10% |
| zoom\_range | 0.2 | Zooms into or out of image by 20% |
| horizontal\_flip | True | Horizontal image flip |
| vertical\_flip | True | Vertical image flip |
| brightness\_range | [0.9, 1.1] | Adjust brightness range creating different lightning condition |
| fill\_mode | nearest | Fills the empty space due to transformations with nearest pixels |

The image data augmentation was applied using the TensorFlow’s pre-defined ImageDataGenerator() library function. These techniques (sub modules) were applied to majority of training CXR images.

### 3.4.6 Summary of Data Pre-Processing

The specific steps described above were essential in the Pneumonia diagnosis in order to produce quality input into the implemented neural network model. The CXR images used in this research consist of 5192 images in two classes: Normal and Pneumonia which was assessed by two radiologists in training, testing and validation datasets. To improve the image quality, particularly in terms of low-contrast areas, Contrast Limited Adaptive Histogram Equalization (CLAHE) was performed by applying clip limit of 2.0 and an 8 by 8 tile to avoid over-amplification of image pixels. Also, random adjustments to the contrast level were also made to generalize the image with a contrast factor (0.8-1.2). After Normalization operation to reduce image intensity and resizing to 224 by 224 pixels for efficient training time, the TensorFlow’s ImageDataGenerator performed the data augmentation by creating equivalent images through rotating, shifting, shearing, zooming and changing brightness of the Image dataset. These preprocessing steps in combination enhances training by simulating various real-world conditions of the CXR image, resulting an effective trained model.

## 3.5 Machine Learning

Following the image data preprocessing, the ML model, trained using TensorFlow was used to classify CXR image into either Pneumonia or Normal via DenseNet201, as well as the transfer of features into a Capsule Network. The architecture developed to implement the Hybrid learning has been discussed below.

### 3.5.1 Capsule Network

A capsule is a group of neurons with an activity vector representing the instantiation parameters of an entity, with active capsules making predictions for higher-level capsules through transformation matrices (Sabour et al., 2017).

Each individual capsules hold significant image features that may have been ignored during the training, as radiology images are prone to pin point accurate diagnosis.

This work utilized two separate Python modules for Capsule Network mechanism. The first one being “Primary Capsule” class that takes feature maps from the previous convolution layers and reshaped them into small capsule vectors, as in Figure 3.5, that represent different features in images. Secondly, the “Capsule Layer” class that performed the dynamic routing allowed the network to capture image relationships between parts and whole of CXR, making the network more flexible to spatial transformations in the input.

The Table 3.2 below discusses the parameters used in the Capsule Network implementation.

Table 3. 2 Capsule Network Parameter Description with Values

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Class | Value | Meaning |
| Number of Capsules | Primary Capsule | 576 | Number of Capsules each with distinct features |
| Dimension of Capsule | Primary Capsule | 16 | Dimension of Capsule vector that defines the feature space |
| Input Shape | Primary Capsule | (batch, height, width, channels) | Shape of the input to passed in this layer from the previous Convolution layer. |
| Convolution Filters | Primary Capsule | Number of capsules \* dimension of capsules | Total filters that generate feature capsules. |
| Kernel Size | Primary Capsule | 3 | Size of convolutional kernel or image patch |
| Strides | Primary Capsule | 1 | Length by which the Kernel moves |
| Padding | Primary Capsule | same | Maintaing consistent output shape compared to input shape |
| Number of Capsules | Capsule Layer | 10 | Number of high-level features or classes. |
| Dimension of Capsules | Capsule Layer | 32 | Dimension of each capsule vector for capture of more complex details. |
| Routings | Capsule Layer | 3 | Number of loops assigned for dynamic routing |

The hybrid CNN (DenseNet-201 model), whose weights and features are transferred to a Capsule Network (Bodapati & Rohith, 2022; Mittal et al., 2020; Zhang et al., 2024), resulted in the model accuracy of 92%.

Initially, a CNN sequential model consisting of 4 Convolution layers was implemented alongside CapsNet , and Fuzzy-c means (FCM) was used for final evaluation via random cluster points (Rohmayani & Rahayu, 2022). It resulted in an accuracy of 88.27%, eventually losing its significance in this research.

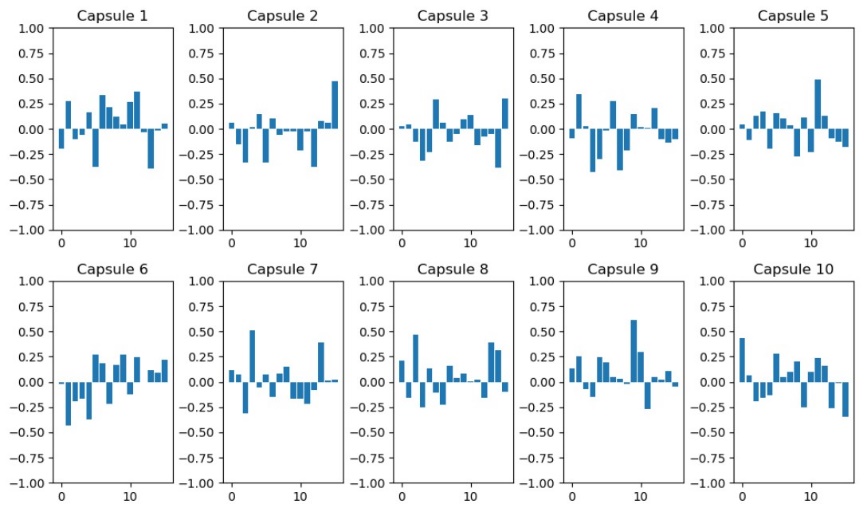


Figure 3. 5 Individual Capsules Consisting of Image Features from CXR Image

The Table 3.1 below highlights the experimental result of different approaches utilized in this work.

Table 3. 3 Accuracy Comparison of Hybrid Classification Approach

|  |  |  |  |
| --- | --- | --- | --- |
| Hybrid Approach | Optimizer | Activation Function | Accuracy |
| DenseNet-201 | Adam | Sigmoid | 89.13% |
| DenseNet-201 + CapsNet | Adam | Sigmoid | **92.31%** |
| CNN (4 x conv2d) + CapsNet + Fuzzy-c means | Adam | Sigmoid | 88.27% |

It’s significant from Table 3.1 that Dense201 architecture combined with Capsule Network became the best fit for the Pneumonia CXR diagnosis.

### 3.5.2 TensorFlow Architecture to Implement Transfer Learning

This work utilized the TensorFlow’s DenseNet-201 architecture as the core feature extractor, which consisted of 201 layers of neural networks and 20,242,984 parameters.

The input RGB image, sized 224 by 224 was passed into the DenseNet201 layer, but not before preprocessing and augmentation operations. The layer started feature extraction with an initial convolution operation that applies 64 filters that reduces the spatial dimensions of the input image. The MaxPooling layer downscales the image size to increase computational efficiency. Similarly, a CXR image moves through sequence of dense blocks and transition layers, where the gradual decrease in dimension is noticed as in Figure 3.6.

Each DenseNet201 blocks consist of 6, 13, 48 and 32 layers of 1 by1 and 3 by 3 convolutions, which is followed by a transition layer that compresses the output to 128, 256 and 512 filters. The DenseNet201 feature extraction concludes after Dense Block 4 with a resulting image size of 7 by 7 with 1920 filters. After the DenseNet201 transfer, a custom Conv2D layer with 256 filters and a MaxPooling2D layer further refine the features, and these are passed to a Primary Capsule layer that organizes the trainable weights into 576 capsules, each with 16 dimensions.

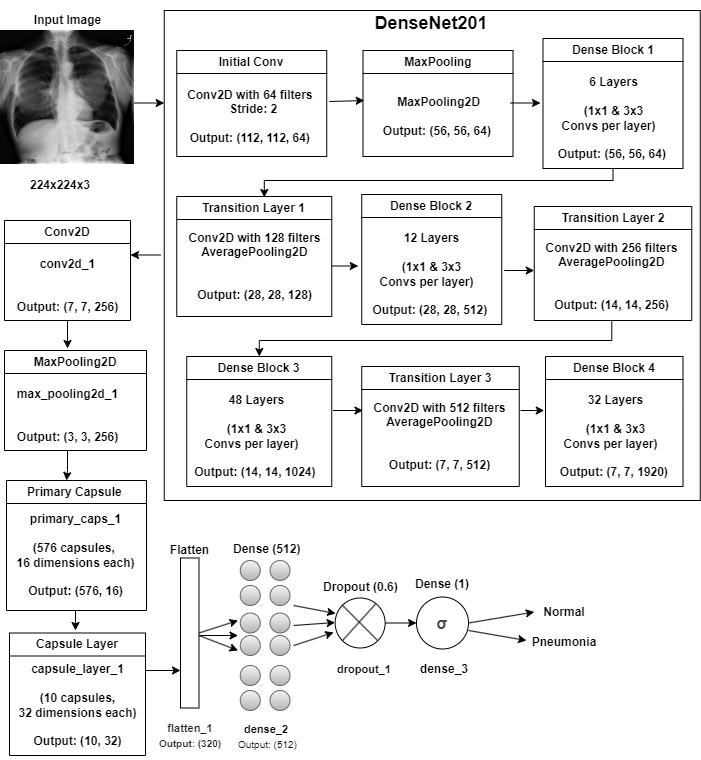


Figure 3. 6 Hybrid DenseNet201 and Capsule Network implemented in TensorFlow

The output from Primary Capsule gets passed to a Capsule Layer that reduces it to 10 capsules with 32 dimensions, whose result is a fully connected neurons with a refined image features for classification. This output is flattened into a 1D vector before performing a full connection in the dense layer with 512 units.

An implemented dropout layer with a rate of 0.6 prevents overfitting by randomly deactivating nodes during training. Finally, the last dense layer consisting of a single unit (due to binary classification) produces a probability score of an image belonging to either Normal or Pneumonia via Sigmoid activation function (σ).

## 3.6 Machine Learning Evaluation and Prediction

With the completion of Neural Network Training, a machine file (.h5) was generated containing the features and learnable parameters captured during that phase. The accuracy evaluation metrics as well as the prediction outputs have been discussed below.

### 3.6.1 Performance Evaluation

The performance was computed based on a 4 by 4 Confusion Matrix. Out of 624 Test images, 217 (True Negative) CXR were correctly classified into Normal whereas 359 (True Positive) were correctly classified as Pneumonia category. However, 17 (False Positive) Normal cases were falsely classified as Pneumonia and 31 (False Negative) Pneumonia cases were classified as Normal. The accuracy metrics used to evaluate the model performance were Recall, Precision and Accuracy scores.

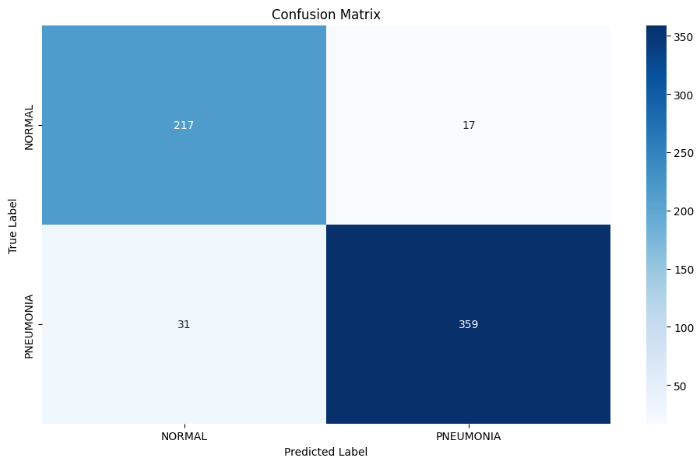


Figure 3. 7 Confusion Matrix for CXR Classification

**Recall:** Out of all positive classes how much was predicted correctly?

**Precision:** Ratio of correctly predicted cases (true positives) to all predicted positive cases (true positives + false positives)

**Accuracy:** Ratio of correctly predicted cases (both true positives and true negatives) to the total number of cases.

[Eq 3.7]

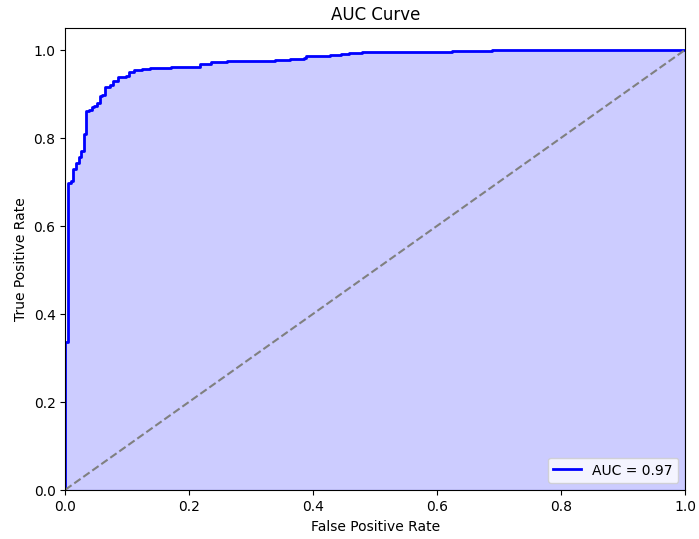


Figure 3. 8 AUC Curve from Pneumonia Diagnosis

The Recall, Precision and Accuracy calculations, as used in Eq 3.5, Eq 3.6 and Eq 3.7 were referred from (Powers, 2011).

Considering the Confusion matrix form Figure 3.7, the Recall, Precision and Accuracy scores calculated were 0.92, 0.95 and 0.92 respectively. Also, the AUC (Area Under Curve) score, as shown in Figure 3.8 was 0.97, that indicated a better performing model.

### 3.6.2 XAI Integrated Classification with Fuzzy Sets

The CNN+CapsNet machine file (.h5) obtained from the TensorFlow training procedure was utilized for Initial Diagnosis where the input is a clear high quality CXR image, which can be in either .jpeg or .png format. It was responsible for distinguishing between Pneumonia or Normal alongside the Pneumonia Symptoms defined as Fuzzy sets which are Breathlessness, Sputum Production, Fever Duration, Fever Value and Hemoptysis.

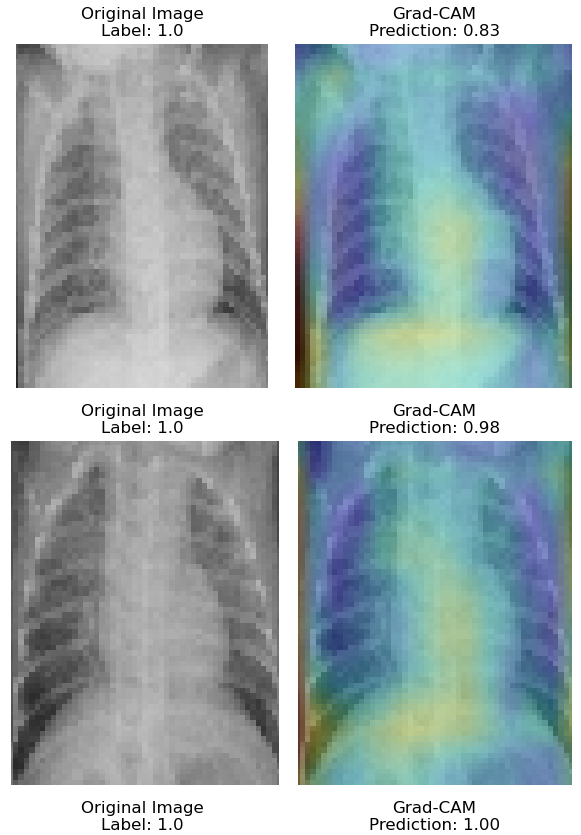


Figure 3. 9 Implemented GRAD-CAM in Pneumonia Diagnosis

Just like the backend, the input image was preprocessed using CLAHE, normalized and resized to 224 by 224 before the classification. The machine file firstly classified the CXR as either Normal or Pneumonia. If the initial classification was Normal, the result is returned as it is, otherwise for the case of Pneumonia, the Fuzzy rules then further calculated the final membership score. If membership score < 0.33; the result is Low Pneumonia, else if score <0.66; the result is Mild Pneumonia. Else the result is Severe Pneumonia. Along with the result, the XAI’s GRAD-CAM was generated to clarify the AI’s decision regarding the generated output in order to increase clinical reliability. The CNN scores, the Fuzzy Symptom values and the GRAD-CAM Base64 encoded data was passed to the DICOM generating module as read-only values.

### 3.6.3 Fuzzy Rules with Dynamic Membership Adjustment

A patient who performs a follow-up needs the metadata updated, considering that the patient’s condition may improve or worsen. The fuzzy logic method implemented in this severity assessment model begun by defining further symptoms regarding their intensity levels. For example, breathlessness ranged from 0 (absent) to 1 (severe), and fever values span from 35 degrees Centigrade to 43 degrees Centigrade.

Each symptom variable was further divided into fuzzy membership functions like “poor”, “average” and “good”, which indicate the category within the value ranges. Table 3.4 discusses the range of the Pneumonia symptoms

Table 3. 4 Fuzzy Membership Range for Pneumonia Symptoms

|  |  |
| --- | --- |
| **Pneumonia Symptom** | **Fuzzy Range** |
| Breathlessness | 0-1 |
| Sputum production | 0-1 |
| Fever duration (days) | 0-30 |
| Fever value (centigrade) | 35-42 |
| Hemoptysis | 0-1 |
| Cough Severity | 0-1 |
| Chest Pain | 0-1 |
| Fatigue | 0-1 |
| Appetite Loss | 0-1 |
| Confusion | 0-1 |

The output variable was “Pneumonia Severity” that represented the severity of pneumonia as either Low, Mild or Severe.

The Fuzzy range for this output severity ranged between 0 (no severity) and 1 (maximum severity) with an incremental factor of 0.1. The final pneumonia severity was calculated as per Eq. 3.13.

This division reflected the significant nature of medical symptoms in a clinical context. The Table 3.4 below highlights the Fuzzy-based pneumonia severity, following the diagnosis guideline as per the medical information from (Dean & Florin, 2018; Modi & Kovacs, 2020)

Table 3. 5 Severity Impact of Fuzzy Rules on Pneumonia Symptoms

|  |  |  |  |
| --- | --- | --- | --- |
| Rule Number | Condition | Severity Outcome | Explanation |
| 1 | Breathlessness is "good" or fever value is "good" or oxygen level is "poor" | Severe | High severity is observed if breathlessness and fever are very high, especially with low oxygen levels |
| 2 | Breathlessness is "average" and sputum production is "average" and fever duration is "average" | Mild | These moderate symptoms suggest a mild severity |
| 3 | Breathlessness is "poor" and sputum production is "poor" fever duration is "poor" and oxygen level is "good" | Negligible | Low severity due to minimal symptoms due to “good” i.e. high oxygen levels |
| 4 | Fatigue is "good" or cough severity is "good" or chest pain is "good" | Mild | Mild severity is indicated since fatigue, cough, or chest pain in isolation is a moderate stage |
| 5 | Confusion is "good" and appetite loss is "good" and hemoptysis is "good" | Severe | High level confusion and hemoptysis are indicators of Severe stage Pneumonia. |
| 6 | Fever value is "poor" and oxygen level is "good" | Negligible | Low severity due to low fever with good oxygen, indicating that the patient is in stable condition. |
| 7 | Fever value is "average" and fatigue is "average" and chest pain is "average" | Mild | These average symptoms indicate Mild stage of Pneumonia. |
| 8 | Fever value is "good" and breathlessness is "good" and sputum production is "good" | Severe | High severity due to Fever, breathlessness and Sputum production having high values. |
| 9 | Oxygen level is "poor" and confusion is "good" | Severe | |  | | --- | |  |  |  | | --- | | Low oxygen combined with high confusion is a severe state since both symptoms are critical indicators of hypoxia. | |

However, this work advances the concept of expert system by integrating Explainable AI (XAI) as well as incorporating new symptoms into the fuzzy system, providing transparency on how each rule impacts the final severity calculation. Furthermore, this work has integrated Dynamic Fuzzy membership adjustment to cover the problem of biased Diagnosis form the previous static models. A Data Frame stored all the symptom values and the trend of the growing symptom trend adjusted the membership values.

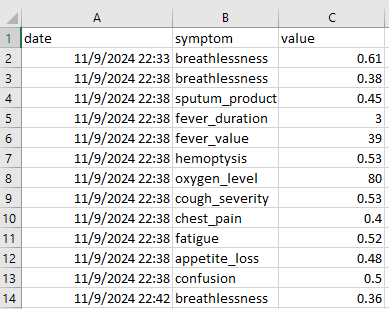


Figure 3. 10 Snapshot of Pneumonia Symptoms Data Frame with Membership Values

To address the gap of Fuzzy based diagnosis, this work has implemented an Algorithm for dynamic membership adjustment with contextual weighting as:

**Step 1:** Check the Data Frame for Trending Symptoms

**Step 2:** Perform Data cleaning by Filtering out values which deviates by more than three standard deviations from the mean using Z-Score (Newbold et al., 2013).

**Step 3:** Calculate recent variance σ2 and Mean μ (Walpole & Myers, 1978), for the last 10 data points (n=10).

* If data points < 10; use variance and mean of all available data

**Step 4:** Calculate the smoothing factor S (Hyndman & Athanasopoulos, 2018) with initial base smoothing B defined as 0.3

**Step 5:** Apply the exponential smoothing (Gardner, 2006)

* Initialize first value of the weighted symptom data
* For each symptom data point x*i,* calculate smoothed value Sxi

**Step 6:** Set percentile thresholds values with respect to recent mean value μ

* If μ < 0.5; set lower percentile = 10% and upper percentile = 90%
* If μ > 0.5; set lower percentile = 15% and upper percentile = 85%

**Step 7:** Calculate adjusted minimum and maximum range as per the percentile thresholds (Bornmann et al., 2013)

* Adjusted minimum Amin is the 10th or 15th percentile of x*i*
* Adjusted maximum Amax is the 90th or 85th percentile of x*i*

**Step 8:** Generate Adjusted range Ar of values from the adjusted minimum to the adjusted maximum with an increment of 0.1

**Step 9:** Limit Adjusted range Ar within the specified minimum and maximum bounds and return the clipped range

The output was the severity level from the final fuzzy scores ranging from 0 to 1 and the categories are “negligible”, “mild” or “severe”. The model applied fuzzy rules to capture the complex patterns among symptoms, which were calculated on how different combinations influence severity. Each symptom was first fuzzified into degrees of fuzzy sets which were then applied to the rules (refer to Table 3.5).

After all the rules were applied, the fuzzy outputs were combined to produce a single numeric severity score. The calculation was done by taking the sum of each rule's fuzzy severity value multiplied by its weight, divided by the sum of the weights, as shown in Eq 3.13.

## 3.7 Creating DICOM file and Conversion to FHIR

In the Figure 3.11 below, the flow is pretty simple to understand. The user uploads a chest X-ray image and enters Metadata like Patient information and study-related details which is mandatory for a valid DICOM file. The CNN accuracy, Fuzzy based Symptoms and GRAD-CAM Base64 compressed Data from the diagnosis is automatically fed into the DICOM as read-only data.

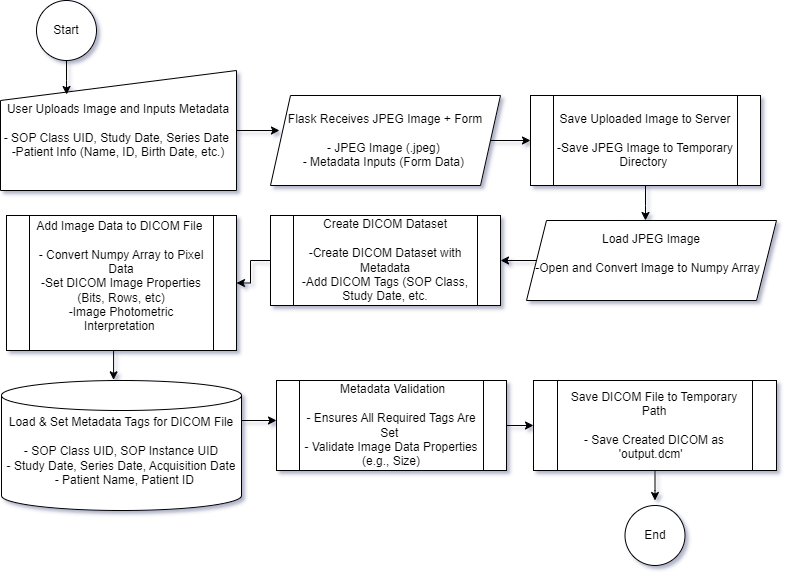


Figure 3. 11 Flow Diagram of DICOM Creation

Once the user was done submitting both the image and the metadata, the Flask server received both of them and the image is temporarily saved on the server. The System then fetched this JPEG image and converted it into a NumPy array for processing. Using this conversion, the system further created a DICOM dataset where metadata like Study Date, Patient Name, etc.. are added with their respective DICOM tags, as shown in Table 3.6. The system also checked if the image data resolution is valid. Only after proper validation, the DICOM file was saved to a temporary path under a designated file name, typically “output.dcm”. This generated DICOM be used for future Diagnosis. As shown in Figure 3.12, this is the snapshot of the generated DICOM (output.dcm) that was uploaded to Orthanc server (A genuine DICOM handler Application):



Figure 3. 12 Snapshot of Orthanc Web Server Which Handles DICOM Files

The Table 3.2 below contains all the implemented DICOM fields along with its Equivalent DICOM Tags, as defined in (National Electrical Manufacturers Association (NEMA), 2023), that was implemented in this research work, which is as follows:

Table 3. 6 DICOM Metadata and its Respective Tags

|  |  |  |
| --- | --- | --- |
| **Field Name** | **Equivalent DICOM Tag** | **Description** |
| Patient Name | (0010,0010) | Patient's full name |
| Patient ID | (0010,0020) | Patient’s Unique Identifier |
| Patient Birth Date | (0010,0030) | Patient’s Date of Birth |
| Patient Sex | (0010,0040) | Patient’s Gender (M, F or O) |
| Study Date | (0008,0020) | The date the study was performed |
| Series Date | (0008,0021) | The date the series was performed |
| Acquisition Date | (0008,0022) | The date the acquisition was performed |
| SOP Class UID | (0008,0016) | Identifies the type of SOP (Service Object Pair) |
| SOP Instance UID | (0008,0018) | Unique identifier for the DICOM instance |
| Modality | (0008,0060) | Specifies the modality type (e.g., CT, MR) |
| Institution Name | (0008,0080) | Institution where the study was performed |
| Referring Physician's Name | (0008,0090) | Name of the referring physician |
| Study Description | (0008,1030) | A description of the study performed |
| Series Description | (0008,103E) | A description of the series |
| Body Part Examined | (0018,0015) | The body part being imaged (e.g., chest) |
| Image Rows | (0028,0010) | Number of rows in the image |
| Image Columns | (0028,0011) | Number of columns in the image |
| Pixel Spacing | (0028,0030) | The physical distance between the centers of each pixel |
| Bits Allocated | (0028,0100) | Number of bits allocated for each pixel sample |
| Photometric Interpretation | (0028,0004) | Specifies the intended interpretation of the pixel data (e.g., MONOCHROME2) |
| Manufacturer | (0008,0070) | The manufacturer of the equipment used for the study |
| Frame of Reference UID | (0020,0052) | Unique identifier for the frame of reference |
| Study Instance UID | (0020,000D) | Unique identifier for the study |
| Series Instance UID | (0020,000E) | Unique identifier for the series. |
| Software Versions | (0018,1020) | Specifies the Version of the Software |
| CNN Accuracy | (9999,0010) | Custom tag for storing the CNN accuracy as a percentage |

Continuing further, the generated DICOM file was converted to an equivalent FHIR output. A valid JSON structure, defined in the vicinity of Pneumonia Diagnosis, was found compatible with the HL7-FHIR implementation guidelines. The system was designed to read a DICOM file and extract the metadata. Table 3.7 describes the implemented FHIR resources:

Table 3. 7 Mapping of DICOM Metadata Fields to FHIR Resources

|  |  |
| --- | --- |
| **DICOM Field** | **Equivalent FHIR Field** |
| TransferSyntaxUID | CUSTOM TAG |
| SOPClassUID | ImagingStudy.series.instance.sopClass |
| SOPInstanceUID | ImagingStudy.series.instance.uid |
| StudyDate | ImagingStudy.started |
| SeriesDate | ImagingStudy.series.started |
| AcquisitionDate | ImagingStudy.series.started |
| ContentDate | CUSTOM TAG |
| StudyTime | CUSTOM TAG |
| SeriesTime | CUSTOM TAG |
| AcquisitionTime | CUSTOM TAG |
| ContentTime | CUSTOM TAG |
| Modality | ImagingStudy.modality |
| Manufacturer | CUSTOM TAG |
| PatientName | Patient.name |
| PatientID | Patient.id |
| PatientBirthDate | Patient.birthDate |
| PatientSex | Patient.gender |
| StudyInstanceUID | ImagingStudy.uid |
| SeriesInstanceUID | ImagingStudy.series.uid |
| PatientAge | Patient.age |
| ReferringPhysicianName | ImagingStudy.referrer |
| InstitutionName | CUSTOM TAG |
| StudyDescription | ImagingStudy.description |
| SeriesDescription | ImagingStudy.series.description |
| ManufacturerModelName | CUSTOM TAG |
| SoftwareVersions | CUSTOM TAG |
| CNN Accuracy | ImagingStudy.series.extension |

The FHIR resources library, implemented in this work, was imported from (Islam, 2019/2024). While most DICOM fields have a specific FHIR tag, the “CUSTOM TAG” fields don’t have direct co-ordination in the respective FHIR specification. These particulars can be defined as FHIR extensions that allows custom data or information to be included without breaking interoperability.

# CHAPTER 4: RESULTS

The purpose of this work is to highlight the simulated environment and component testing of individual modules. Furthermore, a comparative study will indicate how efficient this work is compared to previous research.

## 4.1 Simulation Environment

The hardware, software, and programming modules to develop this system are discussed here.

Table 4. 1 Hardware and Software Specifications

|  |  |
| --- | --- |
| Hardware Specification | Software Specification |
| CPU: Intel Core i7 10700k | OS: Windows 10, Home Edition |
| RAM: 16 GB | Framework: Visual Studio Code, Sublime Text Editor, Jupyter Notebook |
| Graphics Card: NVIDIA RTX 3070, 8 GB | Programming Language: Python, JavaScript, HTML (Web scripting) |

### 4.1.1 Libraries Used

The libraries imported for CNN + CapsNet based classification, and the integration of Fuzzy rules, as well as the incorporation of XAI have been discussed below.

* **TensorFlow**

Tensorflow was used to define and load deep learning models, such as Capsule Network and CNN models. This library is the core of Data augmentation and image preprocessing methods used for model training, and were trained efficiently to classify the CXR images (Bharati et al., 2020; Salam et al., 2021) . It was also used to build custom layers like PrimaryCaps and CapsuleLayer for the Capsule Network.

* **NumPy**

NumPy (Harris et al., 2020) was used for array based operations like flattening model features, Image Preprocessing operation (CLAHE), image normalization and calculating fuzzy input factors for fuzzy logic classification.

* **Matplotlib**

Matplotlib was used for plotting training accuracy plots, Validation loss plots, AUC and ROC plots (Caswell et al., 2021), and specifically for plotting Grad-CAM heatmaps.

* **Scikit-learn (sklearn)**

This library was primarily used for Fuzzy logic to provide support for feature scaling and reshaping data (Varoquaux et al., 2015). Additionally, sklearn provided model evaluation metrics like accuracy, F1-score, and recall.

* **Pydicom**

This library was used for handling the DICOM (Digital Imaging and Communications in Medicine) files, which is the actual standard format for medical images. This library allows users to read, modify, and create DICOM files (Gupta et al., 2021). Specifically, it has been used to extract metadata from existing DICOM files during the conversion of DICOM to FHIR (JSON-based structure).

## 4.2 Component Testing

Component testing ensures that each specific part of an application works well as required, verifying whether the inputs produce the expected outputs. In this research, the focus is on operations involving image conversion and metadata extraction.

### 4.2.1 Component Testing of Image Transfer

The original image quality must remain unhindered when stored as a DICOM file with other Metadata. This work has implemented Python’s Pydicom library for conversion. This ensures lossless transfer of an X-ray image which is stored as a NumPy array.

In Figure 4.1, it is proved that no image data was lost during the conversion. As per the details on the top, the mean pixel intensity difference between the original and DICOM transferred image is 0.0, which means that the pixel values are identical. The middle section presents the visual comparison between the original image and the image converted into DICOM format. Both images are seen having the same resolution and appear identical. The bottom part is the histogram of the pixel intensity values for both images. These diagrams illustrate the frequency distribution of pixel values from 0 to 255. As per the images, both histograms have similar shapes and distributions, confirming that the pixel intensities have been preserved accurately during the conversion.

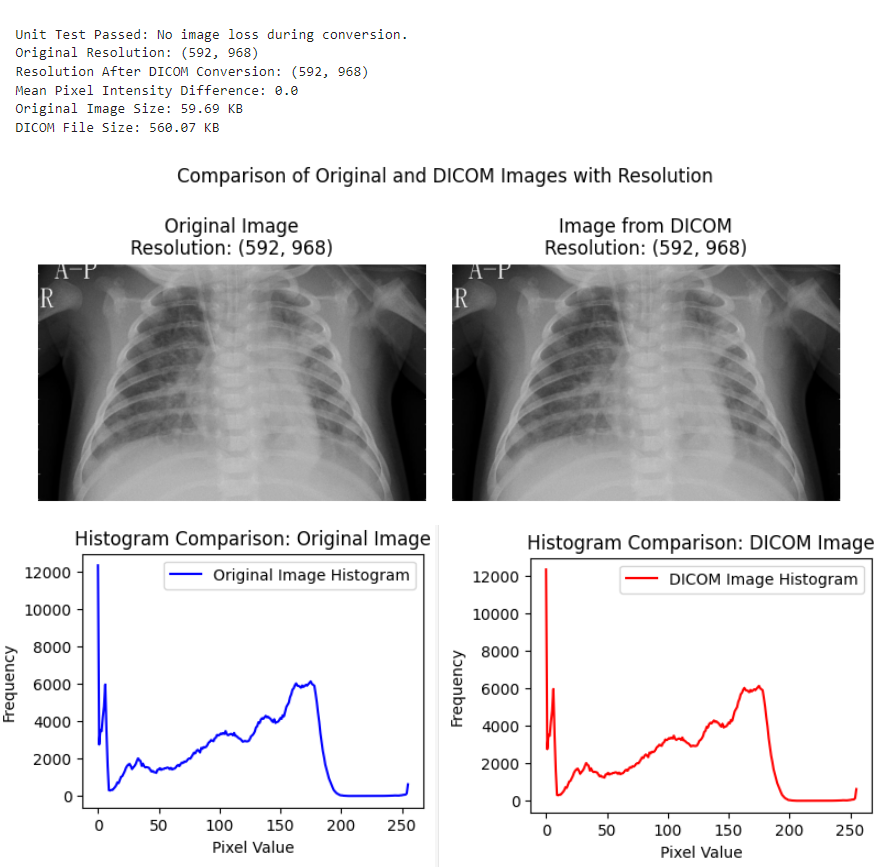


Figure 4. 1 Comparison Between Original and DICOM Images

This concludes that there is indeed no loss in image quality of Chest X-ray when converting to NumPy array and storing them in DICOM file. The Pydicom library is well optimized to handle such operations.

### 4.2.2 Component Testing of FHIR Structure

The research implements a JSON structure that maps each specified Fields of DICOM file as seen in Table 3.2 earlier. With the help of Pydicom, it’s easier to trace the relevant metadata to the appropriate part of FHIR’s JSON structure. Figure 4.2 below is an illustration regarding unit testing of metadata mapping in Python before generating the appropriate format:

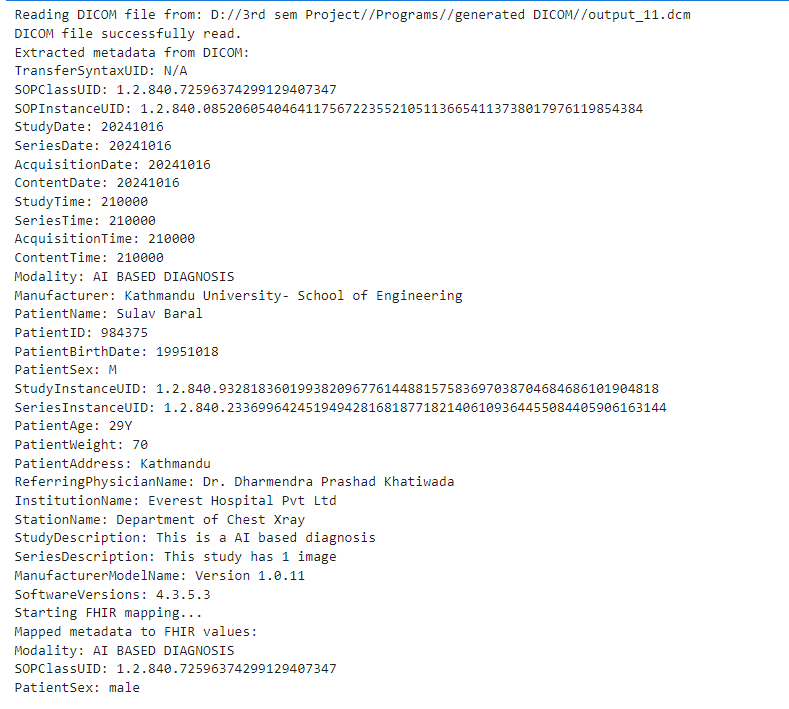


Figure 4. 2 Extracted DICOM Metadata and Mapped FHIR Values

This metadata includes detailed information such as the patient’s personal details, imaging study descriptions, and the manufacturer’s software version in order to clarify the diagnosis involved for proper health data interoperability between systems.

Next, the generated DICOM has been verified in an online copy and paste validator to check if the generated structure is valid or not. These validators check if the generated FHIR structure is compliant with HL7 standards or not as well as validation of appropriate data in appropriate fields. As for this work, some Metadata of DICOM (such as Accuracy Scores) don’t have a relevant format. So, FHIR allows the use of custom extensions to fill up such details, making information more flexible for interoperability.

One such website to check validation is: [*https://simplifier.net*](https://simplifier.net)*.* Figure 4.3 demonstrates the result of FHIR structure validation.



Figure 4. 3 Successful Validation of JSON Structure in Simplifier.net

This structure was the result of original JSON being compatible with the HL7-FHIR guidelines. Thus, this work achieved an interoperable structure for active communication among healthcare organization.

## 4.3 Pneumonia Diagnosis with Dense201 and CapsNet

Aforementioned in Methodology section, the Dense201 and CapsNet based training yielded an accuracy of 92.31%, and an AUC score of 0.97.

It’s evident that the pre trained machine file produced significant result in classifying between two classes, increasing model reliability in clinical settings. This is due to high AUC score, as shown in Figure 3.8, which indicates almost perfect agreement for the CXR dataset in use.

## 4.4 Dynamic Adjustment of Fuzzy Based Symptoms

The implemented dynamic membership adjustment with contextual weighting algorithm for a total of 121 symptom values score produced an average Mean Absolute Error (MAE) of 0.135 and Root Mean Square Error (RMSE) of 0.175, which indicates that the algorithm’s prediction or value adjustments align reasonably well with actual vales since the average errors are relatively low.

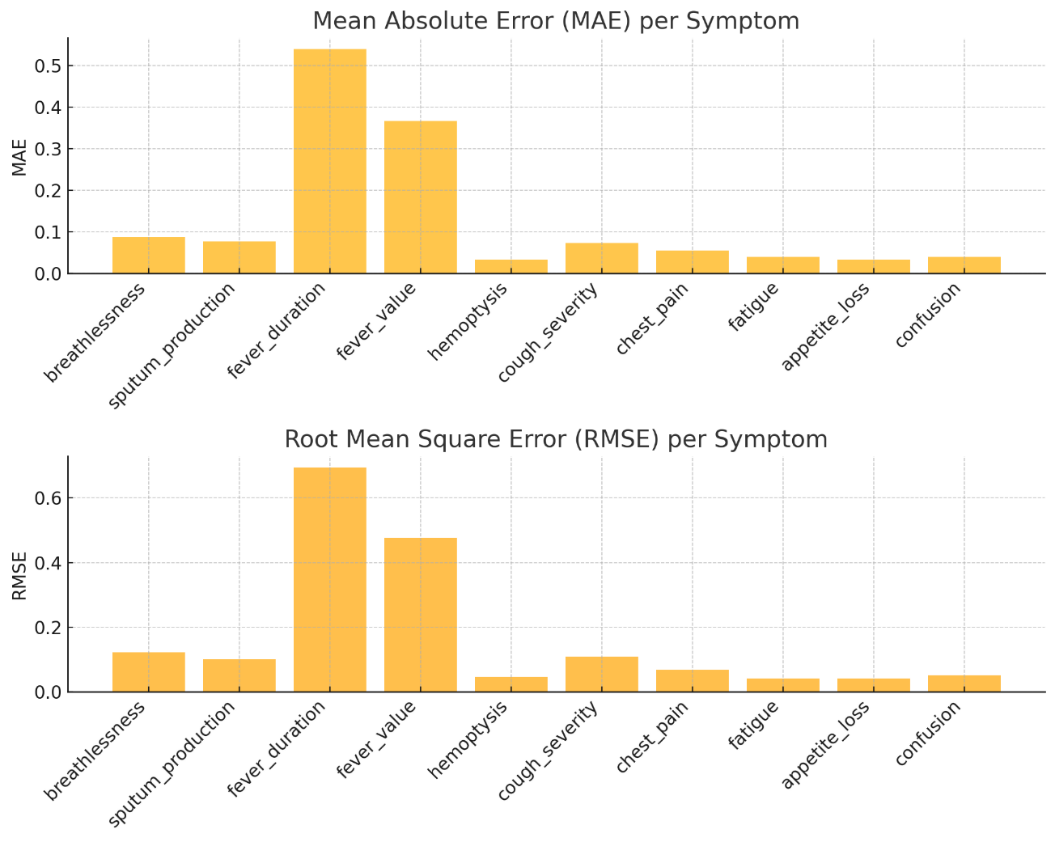


Figure 4. 4 MAE and RMSE of each Fuzzy Based Pneumonia Symptom

For calculating the accuracy, the average range of symptoms was defined. Different symptoms have different range, such as fever value (35 to 42) and fever duration (0 to 30 degree centigrade), and the rest have 1. For a total of N number of symptoms, the Average range is:

Upon current observation of RMSE and MAE, the accuracy from both methods is:

From Eq [4.2] and Eq [4.3], the Accuracy achieved from MAE and RMSE were 97% and 96.1% respectively.

Table 4. 2 Comparison of Membership Adjustment Algorithm with Previous Works

|  |  |  |  |
| --- | --- | --- | --- |
| Study | Algorithm used | Accuracy | Limitations |
| This work (2024) | Dynamic membership adjustment with contextual weighting | 96%-97% | Other clinical factors for Pneumonia diagnosis require separate processing, increasing computational complexity |
| (Burnashev et al., 2023) | Decision Tree with Predefined Fuzzy Sets | Not Defined | No Dynamic Rule Adjustment for overlapping symptoms |
| (Ieracitano et al., 2022) | Fuzzy-Enhanced Deep Learning (CovNNet) | 81% | Limited to CXR image classification, lacks symptom-based adjustment |
| (Arani et al., 2019) | Fuzzy Logic with Static Rules | 93% | Lacks sensitivity to Pneumonia symptom adjustment |

## 4.5 System Efficiency Analysis

When implementing any application, it’s important to analyze its processing time and the space consumed. The analysis of Pneumonia Diagnosis can be described through the Big O Notation. The Big O is a method to clarify how the required time or space grows as the size of the input increases, giving an idea about how efficient is the implemented algorithm for large inputs.

### 4.5.1 Time and Space Complexity of Trainable Machine File

In Table 4.1 below, the calculations of each component have been described. The formula for the time complexity is derived based on its input size, and filter size. Number of channels, and the number of operations involved. The operations whose time complexity is O (1) have been neglected (such as Dropout and Fuzzy Inputs).

Table 4. 3 Time Complexity Analysis using Big O Notation

|  |  |  |  |
| --- | --- | --- | --- |
| Component | The formula for Time Complexity | Implemented Values | Calculated Time Complexity |
| Resizing Image | O (h⋅w) | Width: 224, Height:224, Color Channel:3 | O (224 x 224 x 3)  ≈ 150,000 ops |
| CLAHE Preprocessing | O(m×n) | Clip Limit: 2.0, Tile Grid Size: (8, 8), Image Size: 224x224 | O (224 x 224 x 3) ≈ 150,000 ops |
| DenseNet201 | O(h×w×c×f) across all layers | Input: 224 x 224 x 3, Initial Conv (64 filters, stride 2), Dense Block 1 (6 layers), Dense Block 2 (12 layers), Dense Block 3 (48 layers), Dense Block 4 (32 layers), transition layers with Conv2D filters (128, 256, 512), final dense layer | O (11.1M+0.2M+1.1M+0.1M+0.97M+0. 05M+2.3M+0.02M+0.7M+512) ≈ 16.55 million ops |
| Conv2D Layer | O (h⋅w⋅c⋅k2.f) per layer | Input: (7, 7, 1920), Filters: 256, Kernel: (3, 3) | O(7×7×1920×9×256) ≈ 44.2 million ops |
| MaxPooling Layers | O (h⋅w⋅c) per layer | Output: (3, 3, 256) | O (3×3×256) = 2,304 ops |
| Primary Capsule Layer | O (h.w.c.f) | Output: (576, 16), Parameters: 2,360,320 | 2.36 million ops |
| Capsule Layer (Routing) | O(n×m) | Output: (10, 32), Parameters: 5,120 | 5,120 ops |
| Flattening Layer | O(n) | Output: (320) | 320 ops |
| Dense Layer | O(n×m) | Input: 320, Output: 512, Parameters: 164,352 | 164,352 ops |
| Grad-CAM Generation | O (m⋅n⋅c) | Last Conv2D layer output (7x7x256) | O (7⋅7⋅512) ≈ 12,544 ops |
| **Total** | **O(h⋅w⋅c⋅k2) +O(r⋅d2) +O(n)** |  | **~63.61 million ops** |

For Space complexity, the consideration factors are memory required to store functions, weights, and intermediate results which are stored as floating-point values at each layer. Each element occupies 4 bytes regarding the 32-bit floating point precision in Windows 10.

Table 4. 4 Space Complexity Analysis using Big O Notation

|  |  |  |  |
| --- | --- | --- | --- |
| Component | Formula for Space Complexity | Implemented Values | Calculated Space Complexity |
| Resizing Image | O(h×w×c) | Width: 224, Height: 224, Color Channel: 3 | O(224 x 224 x 3) ≈ 150,528 elements |
| CLAHE Preprocessing | O(m×n×c) | Image Size: 224 x 224 x 3 | O(224×224×3) ≈ 150,528 elements |
| DenseNet201 | Sum of Parameters in all layers | Initial Conv with 64 filters, Individual Dense blocks, Transition layers and Final Dense layer | ≈ 18,321,984 elements |
| Conv2D Layer | O(h×w×f) | Output Shape: (7, 7, 256) | O(7×7×256) ≈ 12,544 elements |
| MaxPooling Layers | O(h×w×c) | Output: (3, 3, 256) | O(3×3×256) ≈ 2,304 elements |
| Primary Capsule Layer | O(h×w×f) | Output: (576, 16) | O(576×16=) ≈ 9,216 elements |
| Capsule Layer (Routing) | O(n×m) | Output: (10, 32) | O(10×32) ≈ 320 elements |
| Flattening Layer | O(n) | Output: (320) | 320 elements |
| Dense Layer | O(n×m) | Input: 320, Output: 512 | O(320 x 512) ≈ 164,352 elements |
| Grad-CAM Generation | O(m×n×c) | Last Conv2D Layer Output: (7, 7, 512) | O(7×7×512) ≈ 12,544 elements |
|  | **Total** |  | **~18,824,640 elements** |

### 4.5.2 Time & Space Complexity of Dynamic Membership Adjustment Algorithm

Earlier from the Methodology part, the implemented algorithm of dynamic membership adjustment takes two factors: n as the number of data points for the symptom and m as the size of generated range interval between the adjusted minimum and the adjusted maximum value. The time complexity of component with constant operation type as O(1) was negated. The calculated Time complexity, as shown in Table 4.3 was as follows:

Table 4. 5 Time Complexity of Dynamic Membership Adjustment Algorithm

|  |  |  |
| --- | --- | --- |
| Component | Formula for Time Complexity | Implemented Values |
| Removing Empty values and filtering outliers | O(n) | Removing Empty values and Z-score filtering |
| Exponential smoothing | O(n) | Iteration through Symptom data |
| Calculating percentiles | O(nlogn) | Sorting required for percentiles |
| Generating adjusted range | O(m) | Range based on adjusted minimum and adjusted maximum |
| Clipping adjusted range | O(m) | Clipping the values within the minimum value and maximum value |
| **Total Time Complexity** | **O(nlogn+m)** | |

Similarly, the space complexity was calculated as shown in Table 4.4 below:

Table 4. 6 Space Complexity of Dynamic Membership Adjustment Algorithm

|  |  |  |
| --- | --- | --- |
| Component | Formula for Space Complexity | Implemented Values |
| Array of symptom data | O(n) | Stored filtered symptom data |
| Array of weighted data | O(n) | Stored smoothed data |
| Generated range | O(m) | Range of array between adjusted minimum and adjusted maximum |
| **Total Space Complexity** | **O(n+m)** | |

## 4.6 Machine Learning Accuracy Validation

The image preprocessing and data augmentation techniques remained unchanged. Table 4.7 below highlights the validation accuracy achieved.

Table 4. 7 Comparison of Validation Accuracy with Different DenseNet Variants

|  |  |
| --- | --- |
| DenseNet-121 | 94.37% |
| DenseNet-169 | 96.21% |
| DenseNet-201 | **98.56%** |

Furthermore, the Dataset with 17229 images were split into different Training to Testing ratios for DenseNet-201. The results are shown in Table 4.8 below:

Table 4. 8 Accuracy Comparison of DenseNet201 with Different Train to Test Ratio

|  |  |  |
| --- | --- | --- |
| Train to Test Ratio | Validation Accuracy | Validation Loss |
| 90:10 | 98.13% | 6.12% |
| 60:40 | 97.30% | 7.65% |
| 70:30 | 97.62% | 6.25% |
| 85:15 | **98.57%** | **4.27%** |

From the Table 4.8, the Train to Test ratio of 85:15 yielded the best validation accuracy of 98.57%. But most importantly, the validation loss of 4.27% indicates the model achieved a near perfection metric for the CXR classification.

# CHAPTER 5: CONCLUSION AND FUTURE WORK

The work purposed an advanced pneumonia diagnosis tool by combining DenseNet201, CapsNet and Fuzzy Logic. The CXR image classifier achieved an accuracy of 92.31% and an AUC score of 0.97, making it efficient for classification between Normal and Pneumonia cases from Chest X-Rays. The system addressed the gap of Fuzzy rule-based symptom diagnosis following the implementation of a dynamically adjusting Fuzzy membership algorithm values based on the ever-changing Pneumonia symptoms by ensuring that evaluation remained responsive and reliable, as the algorithm achieved a MAE of 0.135 and 97% accuracy as well as a RMSE of 0.175 and 96.1 accuracy%. Through the integration of XAI, the system used Grad-CAM to highlight the key area in CXR which can help clinicians to understand the reason behind the AI’s decision, and also incorporated an expert system based Fuzzy rules for increased transparency behind the systems decision. Additionally, this work also addresses the data accessibility issue of CXR images by populating the AI based accuracy information with other diagnosis data in a DICOM file that also can be converted into a FHIR-compliant JSON, thus enabling seamless interoperability between different healthcare systems.

To summarize, the purposed work addresses these gaps of Pneumonia diagnosis:

* Clinical Reliability: The combination of DenseNet201, CapsNet and Fuzzy logic for improved accuracy increased reliability of AI-based pneumonia diagnosis.
* Transparency: The integration of XAI clarified the decision behind an AI-based Pneumonia diagnosis, increasing reliability among clinicians.
* Data Accessibility and Interoperability: Storing the AI-based results and CXR diagnosis information in a DICOM file guaranteed flawless data accessibility, and conversion of DICOM to FHIR-compliant JSON enabled seamless interoperability across healthcare systems.

Despite covering the diagnosis gaps, the study had some limitations. Primarily, the accuracy and error metrics were achieved on a secondary CXR image dataset, meaning that the machine learning model was unable to capture the variability of Pneumonia patterns in real world cases.

Also, the issue of static Pneumonia diagnosis was partially addressed despite the integration dynamic membership adjustment algorithm. This is due to lack of diverse patient data from different geographical location, decreasing the clinical significance of the purposed system as the work was unable to integrate more complex symptom pattern. Furthermore, the integrated XAI may still not improve clinical reliability as the work failed to address a human like response to explain the decision behind the AI based diagnosis via Natural Language Processing (NLP).

**Future Work**

With concern to a few limitations, these are the suggestions for further improvement of this work:

* Enhancing XAI through the integration of LLM (Large Language Model) for generating a complete Pneumonia diagnosis report.
* Use Primary CXR image data and collect diverse Pneumonia symptom data to increase the clinical significance of this diagnosis application.
* Experiment with various activation function for result comparison.
* Investigate advanced neural network architecture for getting different result.

# REFERENCES

Adlassnig, K. P. (1980). A fuzzy logical model of computer-assisted medical diagnosis. *Methods of Information in Medicine*, *19*(3), 141–148.

Ahmed, S. B., Solis-Oba, R., & Ilie, L. (2022). Explainable-AI in Automated Medical Report Generation Using Chest X-ray Images. *Applied Sciences*, *12*(22), Article 22. https://doi.org/10.3390/app122211750

Ait Abdelouahid, R., Debauche, O., Mahmoudi, S., & Marzak, A. (2023). Literature Review: Clinical Data Interoperability Models. *Information*, *14*(7), Article 7. https://doi.org/10.3390/info14070364

Akerkar, R., & Sajja, P. (2010). *A Neuro-Fuzzy Decision Support System for Selection of Small Scale Business*. *81*, 306–315. https://doi.org/10.1007/978-3-642-14058-7\_31

Ambaliya, M., Chauhan, S., Paliwal, M., Shastri, A., & Sabale, K. (2025). Enhancing Pneumonia Detection Transparency: Exploring Explainable AI Model. In V. Shrivastava, J. C. Bansal, & B. K. Panigrahi (Eds.), *Power Engineering and Intelligent Systems* (pp. 465–478). Springer Nature. https://doi.org/10.1007/978-981-97-6710-6\_35

Anakha, B. M., Shaji, G., & Geetha, S. (2021). Detecting COVID-19 from Chest X-Ray Images using Deep Learning. *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, 1–4. https://doi.org/10.1109/ISCON52037.2021.9702491

Arani, L. A., Sadoughi, F., & Langarizadeh, M. (2019). An Expert System to Diagnose Pneumonia Using Fuzzy Logic. *Acta Informatica Medica*, *27*(2), 103–107. https://doi.org/10.5455/aim.2019.27.103-107

Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*, *20*, 100391. https://doi.org/10.1016/j.imu.2020.100391

Bidgood, W. D. (1998). Clinical importance of the DICOM structured reporting standard. *The International Journal of Cardiac Imaging*, *14*(5), 307–315. https://doi.org/10.1023/A:1006073709957

Bidgood, W. D., Jr, Horii, S. C., Prior, F. W., & Van Syckle, D. E. (1997). Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging. *Journal of the American Medical Informatics Association*, *4*(3), 199–212. https://doi.org/10.1136/jamia.1997.0040199

Biswas, S. S. (2010). *A Soft Computing Model for Physicians’ Decision Process* (No. arXiv:1005.4025). arXiv. http://arxiv.org/abs/1005.4025

Blezek, D. J., Olson-Williams, L., Missert, A., & Korfiatis, P. (2021). AI Integration in the Clinical Workflow. *Journal of Digital Imaging*, *34*(6), 1435–1446. https://doi.org/10.1007/s10278-021-00525-3

Bodapati, J. D., & Rohith, V. N. (2022). ChxCapsNet: Deep capsule network with transfer learning for evaluating pneumonia in paediatric chest radiographs. *Measurement*, *188*, 110491. https://doi.org/10.1016/j.measurement.2021.110491

Bornmann, L., Leydesdorff, L., & Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics*, *7*(1), 158–165. https://doi.org/10.1016/j.joi.2012.10.001

Burnashev, R., Enikeeva, A., Amer, I. F., Akhmedova, A., Bolsunovskaya, M., & Enikeev, A. (2023). Building a Fuzzy Expert System for Assessing the Severity of Pneumonia. In K. Arai (Ed.), *Intelligent Systems and Applications* (pp. 380–396). Springer International Publishing. https://doi.org/10.1007/978-3-031-16075-2\_27

Caswell, T. A., Droettboom, M., Lee, A., Sales De Andrade, E., Hunter, J., Hoffmann, T., Firing, E., Klymak, J., Stansby, D., Varoquaux, N., Hedegaard Nielsen, J., Root, B., May, R., Elson, P., Seppänen, J. K., Dale, D., Lee, J.-J., McDougall, D., Straw, A., … Ivanov, P. (2021). matplotlib/matplotlib: REL: v3.4.0. *Zenodo*. https://doi.org/10.5281/zenodo.4638398

Cerrada, M., Aguilar, J., Colina, E., & Titli, A. (2002). An approach for dynamical adaptive fuzzy modeling. *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE’02. Proceedings (Cat. No.02CH37291)*, *1*, 156–161 vol.1. https://doi.org/10.1109/FUZZ.2002.1004978

Chandra, N., & Bhardwaj, A. (2024). Medical image enhancement using modified type II fuzzy membership function generated by Hamacher T-conorm. *Soft Computing*, *28*(9), 6753–6774. https://doi.org/10.1007/s00500-023-09535-5

Chinniah, P., & Muttan, D. S. (2010). *ICD 10 Based Medical Expert System Using Fuzzy Temporal Logic* (No. arXiv:1001.1979). arXiv. http://arxiv.org/abs/1001.1979

Chumbita, M., Cillóniz, C., Puerta-Alcalde, P., Moreno-García, E., Sanjuan, G., Garcia-Pouton, N., Soriano, A., Torres, A., & Garcia-Vidal, C. (2020). Can Artificial Intelligence Improve the Management of Pneumonia. *Journal of Clinical Medicine*, *9*(1), Article 1. https://doi.org/10.3390/jcm9010248

Chutia, U., Tewari, A. S., Singh, J. P., & Raj, V. K. (2024). Classification of Lung Diseases Using an Attention-Based Modified DenseNet Model. *Journal of Imaging Informatics in Medicine*, *37*(4), 1625–1641. https://doi.org/10.1007/s10278-024-01005-0

Dean, P., & Florin, T. A. (2018). Factors Associated With Pneumonia Severity in Children: A Systematic Review. *Journal of the Pediatric Infectious Diseases Society*, *7*(4), 323–334. https://doi.org/10.1093/jpids/piy046

Dong, H., Zhu, B., Zhang, X., & Kong, X. (2022). Use data augmentation for a deep learning classification model with chest X-ray clinical imaging featuring coal workers’ pneumoconiosis. *BMC Pulmonary Medicine*, *22*(1), 271. https://doi.org/10.1186/s12890-022-02068-x

Elemraid, M. A., Muller, M., Spencer, D. A., Rushton, S. P., Gorton, R., Thomas, M. F., Eastham, K. M., Hampton, F., Gennery, A. R., Clark, J. E., & Group, on behalf of the N. E. of E. P. R. I. S. (2014). Accuracy of the Interpretation of Chest Radiographs for the Diagnosis of Paediatric Pneumonia. *PLOS ONE*, *9*(8), e106051. https://doi.org/10.1371/journal.pone.0106051

Elion, J. L. (1995). DICOM media interchange standards for cardiology: Initial interoperability demonstration. *Proceedings. Symposium on Computer Applications in Medical Care*, 591–595.

Evertsen, J., Baumgardner, D. J., Regnery, A., & Banerjee, I. (2010). Diagnosis and management of pneumonia and bronchitis in outpatient primary care practices. *Primary Care Respiratory Journal: Journal of the General Practice Airways Group*, *19*(3), 237–241. https://doi.org/10.4104/pcrj.2010.00024

Gambato, M., Scotti, N., Borsari, G., Bertoja, J. Z., Gabrieli, J.-D., Cassai, A. D., Cester, G., Navalesi, P., Quaia, E., & Causin, F. (2023). Chest X-ray Interpretation: Detecting Devices and Device-Related Complications. *Diagnostics*, *13*(4), 599. https://doi.org/10.3390/diagnostics13040599

Gardner, E. S. (2006). Exponential smoothing: The state of the art—Part II. *International Journal of Forecasting*, *22*(4), 637–666. https://doi.org/10.1016/j.ijforecast.2006.03.005

Ghuse, N. D., & Monga, S. (2024). A Review on Pneumonia Detection Using Artificial Intelligence Techniques. *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, 1–5. https://doi.org/10.1109/AIMLA59606.2024.10531467

Gorzalczany, M. B., & Piasta, Z. (1999). Neuro-fuzzy approach versus rough-set inspired methodology for intelligent decision support. *Information Sciences*, *120*(1), 45–68. https://doi.org/10.1016/S0020-0255(99)00070-5

Gupta, Y., Costa, C., Pinho, E., Silva, L. A. B., Mandal, S., & Neugebauer, U. (2021). Dicomization of LSM fluorescence composite microscopic image with its bioimaging information. *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 533–538. https://doi.org/10.1109/CBMS52027.2021.00086

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hasan, M. A., Sher-E-Alam, K. M., & Chowdhury, A. (2010). Human Disease Diagnosis Using a Fuzzy Expert System. *ArXiv*. https://www.semanticscholar.org/paper/Human-Disease-Diagnosis-Using-a-Fuzzy-Expert-System-Hasan-Sher-E-Alam/02159107e1697e2cea5541f751481e65c8de6828

Höhl, A., Obadic, I., Fernández-Torres, M.-Á., Oliveira, D., & Zhu, X. X. (2024). *Recent Trends Challenges and Limitations of Explainable AI in Remote Sensing*. 8199–8205. https://openaccess.thecvf.com/content/CVPR2024W/XAI4CV/html/Hohl\_Recent\_Trends\_Challenges\_and\_Limitations\_of\_Explainable\_AI\_in\_Remote\_CVPRW\_2024\_paper.html

Hopstaken, R. M., Witbraad, T., Engelshoven, J. M. A. van, & Dinant, G. J. (2004). Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory tract infections. *Clinical Radiology*, *59*(8), 743–752. https://doi.org/10.1016/j.crad.2004.01.011

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.

Ieracitano, C., Mammone, N., Versaci, M., Varone, G., Ali, A.-R., Armentano, A., Calabrese, G., Ferrarelli, A., Turano, L., Tebala, C., Hussain, Z., Sheikh, Z., Sheikh, A., Sceni, G., Hussain, A., & Morabito, F. C. (2022). A fuzzy-enhanced deep learning approach for early detection of Covid-19 pneumonia from portable chest X-ray images. *Neurocomputing*, *481*, 202. https://doi.org/10.1016/j.neucom.2022.01.055

Ihongbe, I. E., Fouad, S., Mahmoud, T. F., Rajasekaran, A., & Bhatia, B. (2024). Evaluating Explainable Artificial Intelligence (XAI) techniques in chest radiology imaging through a human-centered Lens. *PLOS ONE*, *19*(10), e0308758. https://doi.org/10.1371/journal.pone.0308758

*ImagingStudy—FHIR v5.0.0*. (2024). HL7 FHIR. https://www.hl7.org/fhir/imagingstudy.html

*Index—FHIR v5.0.0*. (2024). HL7 FHIR. https://www.hl7.org/fhir/

Islam, M. N. (2024). *Nazrulworld/fhir.resources* [Python]. https://github.com/nazrulworld/fhir.resources (Original work published 2019)

Joshi, R. C., Yadav, S., Pathak, V. K., Malhotra, H. S., Khokhar, H. V. S., Parihar, A., Kohli, N., Himanshu, D., Garg, R. K., Bhatt, M. L. B., Kumar, R., Singh, N. P., Sardana, V., Burget, R., Alippi, C., Travieso-Gonzalez, C. M., & Dutta, M. K. (2021). A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images. *Biocybernetics and Biomedical Engineering*, *41*(1), 239–254. https://doi.org/10.1016/j.bbe.2021.01.002

Kamel, P. I., & Nagy, P. G. (2018). Patient-Centered Radiology with FHIR: An Introduction to the Use of FHIR to Offer Radiology a Clinically Integrated Platform. *Journal of Digital Imaging*, *31*(3), 327–333. https://doi.org/10.1007/s10278-018-0087-6

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, *172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kiran, Kumar D S, S., C H, C., Datar, S. G., C P, V., & Kumar H S, R. (2024). Developing Adaptive Solutions for Accurate Pneumonia Diagnosis via Chest X-Ray Analysis. *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 1–5. https://doi.org/10.1109/ICDCECE60827.2024.10549708

Konar, A., & Jain, L. C. (2001). An Introduction to Computational Intelligence Paradigms. In L. Jain & P. De Wilde (Eds.), *Practical Applications of Computational Intelligence Techniques* (pp. 1–64). Springer Netherlands. https://doi.org/10.1007/978-94-010-0678-1\_1

Koncar, M. (2007). HL7 and DICOM based integration of radiology departments with healthcare enterprise information systems. *International Journal of Medical Informatics*. https://www.academia.edu/17513392/HL7\_and\_DICOM\_based\_integration\_of\_radiology\_departments\_with\_healthcare\_enterprise\_information\_systems

Kulkarni, D., Doijade, M., Devrukhkar, C., Zilpe, G., & Surana, R. (2012). NetraRIS -a Web based DICOM Viewer. *International Journal of Computer Applications*, *48*. https://doi.org/10.5120/7532-0635

Kundu, R., Das, R., Geem, Z. W., Han, G.-T., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PLOS ONE*, *16*(9), e0256630. https://doi.org/10.1371/journal.pone.0256630

Kuzmak, P. M., & Dayhoff, R. E. (2000). The use of Digital Imaging and Communications in Medicine (DICOM) in the integration of imaging into the electronic patient record at the Department of Veterans Affairs. *Journal of Digital Imaging*, *13*(Suppl 1), 133. https://doi.org/10.1007/BF03167644

Kwon, T., Lee, S. P., Kim, D., Jang, J., Lee, M., Kang, S. U., Kim, H., Oh, K., On, J., Kim, Y. J., Yun, S. J., Jin, K. N., Kim, E. Y., & Kim, K. G. (2021). Diagnostic performance of artificial intelligence model for pneumonia from chest radiography. *PLOS ONE*, *16*(4), e0249399. https://doi.org/10.1371/journal.pone.0249399

Labkoff, S., Oladimeji, B., Kannry, J., Solomonides, A., Leftwich, R., Koski, E., Joseph, A. L., Lopez-Gonzalez, M., Fleisher, L. A., Nolen, K., Dutta, S., Levy, D. R., Price, A., Barr, P. J., Hron, J. D., Lin, B., Srivastava, G., Pastor, N., Luque, U. S., … Quintana, Y. (2024). Toward a responsible future: Recommendations for AI-enabled clinical decision support. *Journal of the American Medical Informatics Association*, *31*(11), 2730–2739. https://doi.org/10.1093/jamia/ocae209

Lagunes, M. L., Castillo, O., Valdez, F., & Soria, J. (2020). Comparison of Fuzzy Controller Optimization with Dynamic Parameter Adjustment Based on of Type-1 and Type-2 Fuzzy Logic. In O. Castillo & P. Melin (Eds.), *Hybrid Intelligent Systems in Control, Pattern Recognition and Medicine* (pp. 47–56). Springer International Publishing. https://doi.org/10.1007/978-3-030-34135-0\_4

Lee, W.-J., Yang, C.-Y., Liu, K.-L., Liu, H.-M., Ching, Y.-T., & Chen, S.-J. (2005). Establishing a Web-Based DICOM Teaching File Authoring Tool Using Open-Source Public Software. *Journal of Digital Imaging*, *18*(3), 169. https://doi.org/10.1007/s10278-005-5171-z

Li, D., Zhou, S., & Pedrycz, W. (2023). Accelerated Fuzzy C-Means Clustering Based on New Affinity Filtering and Membership Scaling. *IEEE Transactions on Knowledge and Data Engineering*, *35*(12), 12337–12349. IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2023.3273274

Lu, J., Ma, G., & Zhang, G. (2024). Fuzzy Machine Learning: A Comprehensive Framework and Systematic Review. *IEEE Transactions on Fuzzy Systems*, *32*(7), 3861–3878. IEEE Transactions on Fuzzy Systems. https://doi.org/10.1109/TFUZZ.2024.3387429

Mann, M., Badoni, R. P., Soni, H., Al-Shehri, M., Kaushik, A. C., & Wei, D.-Q. (2023). Utilization of Deep Convolutional Neural Networks for Accurate Chest X-Ray Diagnosis and Disease Detection. *Interdisciplinary Sciences: Computational Life Sciences*, *15*(3), 374–392. https://doi.org/10.1007/s12539-023-00562-2

McAllister, D. A., Liu, L., Shi, T., Chu, Y., Reed, C., Burrows, J., Adeloye, D., Rudan, I., Black, R. E., Campbell, H., & Nair, H. (2019). Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: A systematic analysis. *The Lancet Global Health*, *7*(1), e47–e57. https://doi.org/10.1016/S2214-109X(18)30408-X

Meedeniya, D., Kumarasinghe, H., Kolonne, S., Fernando, C., Díez, I. D. la T., & Marques, G. (2022). Chest X-ray analysis empowered with deep learning: A systematic review. *Applied Soft Computing*, *126*, 109319. https://doi.org/10.1016/j.asoc.2022.109319

Mildenberger, P., Eichelberg, M., & Martin, E. (2002). Introduction to the DICOM standard. *European Radiology*, *12*(4), 920–927. https://doi.org/10.1007/s003300101100

Mittal, A., Kumar, D., Mittal, M., Saba, T., Abunadi, I., Rehman, A., & Roy, S. (2020). Detecting Pneumonia Using Convolutions and Dynamic Capsule Routing for Chest X-ray Images. *Sensors*, *20*(4), Article 4. https://doi.org/10.3390/s20041068

Modi, A. R., & Kovacs, C. S. (2020). Community-acquired pneumonia: Strategies for triage and treatment. *Cleveland Clinic Journal of Medicine*, *87*(3), 145. https://doi.org/10.3949/ccjm.87a.19067

Monteiro, E., Costa, C., & Oliveira, J. (2013). *A DICOM viewer based on web technology*. 167–171. https://doi.org/10.1109/HealthCom.2013.6720660

Naronglerdrit, P., Mporas, I., & Sheikh-Akbari, A. (2021). COVID-19 detection from chest X-rays using transfer learning with deep convolutional neural networks. *Data Science for COVID-19*, 255. https://doi.org/10.1016/B978-0-12-824536-1.00031-9

National Electrical Manufacturers Association (NEMA). (2023). *DICOM Standard Part 3: Information Object Definitions* (No. DICOM 2023a). National Electrical Manufacturers Association (NEMA). https://www.dicomstandard.org/current/

Newbold, P., Carlson, W. L., & Thorne, B. M. (2013). *Statistics for business and economics*. Pearson. https://thuvienso.hoasen.edu.vn/handle/123456789/9470

Niehues, S. M., Adams, L. C., Gaudin, R. A., Erxleben, C., Keller, S., Makowski, M. R., Vahldiek, J. L., & Bressem, K. K. (2021). Deep-Learning-Based Diagnosis of Bedside Chest X-ray in Intensive Care and Emergency Medicine. *Investigative Radiology*, *56*(8), 525. https://doi.org/10.1097/RLI.0000000000000771

O’Quinn, W., Haddad, R. J., & Moore, D. L. (2019). Pneumonia Radiograph Diagnosis Utilizing Deep Learning Network. *2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT)*, 763–767. https://doi.org/10.1109/ICEICT.2019.8846438

Powers, D. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv*. https://www.semanticscholar.org/paper/Evaluation%3A-from-precision%2C-recall-and-F-measure-to-Powers/6d03a38c5ddb7c7cd1ceb59b28907dc918c5d83a

Rahman, T., Chowdhury, M. E. H., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., & Kashem, S. (2020). Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray. *Applied Sciences*, *10*(9), Article 9. https://doi.org/10.3390/app10093233

Ravi, V., Acharya, V., & Alazab, M. (2023). A multichannel EfficientNet deep learning-based stacking ensemble approach for lung disease detection using chest X-ray images. *Cluster Computing*, *26*(2), 1181–1203. https://doi.org/10.1007/s10586-022-03664-6

Ren, H., Wong, A. B., Lian, W., Cheng, W., Zhang, Y., He, J., Liu, Q., Yang, J., Zhang, C. J., Wu, K., & Zhang, H. (2021). Interpretable Pneumonia Detection by Combining Deep Learning and Explainable Models With Multisource Data. *IEEE Access*, *9*, 95872–95883. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3090215

Rochmawanti, O., & Utaminingrum, F. (2021). Chest X-Ray Image to Classify Lung diseases in Different Resolution Size using DenseNet-121 Architectures. *Proceedings of the 6th International Conference on Sustainable Information Engineering and Technology*, 327–331. https://doi.org/10.1145/3479645.3479667

Rohmayani, D., & Rahayu, A. (2022). Classification of X-Ray Images of Normal, Pneumonia, and Covid-19 Lungs Using the Fuzzy C-Means (FCM) Algorithm. *Journal of Applied Intelligent System*, *7*, 16–25. https://doi.org/10.33633/jais.v7i1.5512

Roy, A., Bhattacharjee, A., Oliva, D., Ramos-Soto, O., Alvarez-Padilla, F. J., & Sarkar, R. (2024). FA-Net: A Fuzzy Attention-aided Deep Neural Network for Pneumonia Detection in Chest X-Rays. *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, 338–343. https://doi.org/10.1109/CBMS61543.2024.00063

Roy, S., Pal, D., & Meena, T. (2023). Explainable artificial intelligence to increase transparency for revolutionizing healthcare ecosystem and the road ahead. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *13*(1), 4. https://doi.org/10.1007/s13721-023-00437-y

Ruuskanen, O., Lahti, E., Jennings, L. C., & Murdoch, D. R. (2011). Viral pneumonia. *The Lancet*, *377*(9773), 1264–1275. https://doi.org/10.1016/S0140-6736(10)61459-6

Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic Routing Between Capsules. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper\_files/paper/2017/hash/2cad8fa47bbef282badbb8de5374b894-Abstract.html

Salam, M. A., Taha, S., & Ramadan, M. (2021). COVID-19 detection using federated machine learning. *PLoS ONE*, *16*(6), e0252573. https://doi.org/10.1371/journal.pone.0252573

Salem, H., Shams, M. Y., Elzeki, O. M., Abd Elfattah, M., F. Al-Amri, J., & Elnazer, S. (2022). Fine-Tuning Fuzzy KNN Classifier Based on Uncertainty Membership for the Medical Diagnosis of Diabetes. *Applied Sciences*, *12*(3), Article 3. https://doi.org/10.3390/app12030950

Sanghvi, H. A., Patel, R. H., Agarwal, A., Gupta, S., Sawhney, V., & Pandya, A. S. (2023). A deep learning approach for classification of COVID and pneumonia using DenseNet-201. *International Journal of Imaging Systems and Technology*, *33*(1), 18–38. https://doi.org/10.1002/ima.22812

Santra, D., Basu, S., Mandal, J., & Goswami, S. (2019). *Lattice-Based Fuzzy Medical Expert System for Low Back Pain Management*. https://doi.org/10.48550/arXiv.1909.03983

Senol, C., & Yildirim, T. (2010). Fuzzy-neural networks for medical diagnosis. *Int. J. of Reasoning-Based Intelligent Systems*, *2*, 265–271. https://doi.org/10.1504/IJRIS.2010.036873

Shen, Z., Ouyang, X., Xiao, B., Cheng, J.-Z., Shen, D., & Wang, Q. (2023). Image synthesis with disentangled attributes for chest X-ray nodule augmentation and detection. *Medical Image Analysis*, *84*, 102708. https://doi.org/10.1016/j.media.2022.102708

Sheu, R.-K., Pardeshi, M. S., Pai, K.-C., Chen, L.-C., Wu, C.-L., & Chen, W.-C. (2023). Interpretable Classification of Pneumonia Infection Using eXplainable AI (XAI-ICP). *IEEE Access*, *11*, 28896–28919. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3255403

Shi, Y., Eberhart, R., & Chen, Y. (1999). Implementation of evolutionary fuzzy systems. *Trans. Fuz Sys.*, *7*(2), 109–119. https://doi.org/10.1109/91.755393

Shivshankar, S., Makhija, N., & Mathusudhanan, P. (2024). Digital Imaging and Communication in Medicine (DICOM): Biomedical and Health Informatics: Imaging and Interoperability Using HL7 and DICOM. In M. A. Chaurasia, P. Balaji, & A. C. Frery (Eds.), *Smart Healthcare and Machine Learning* (pp. 299–317). Springer Nature. https://doi.org/10.1007/978-981-97-3312-5\_20

Shrimali, S. (2024). Optimizing Chest X-ray Analysis for Pneumonia Detection Through Comparative Evaluation of Transfer Learning CNN Architectures. *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, 1–5. https://doi.org/10.1109/I2CT61223.2024.10543957

Sivakumar, M., Parthasarathy, S., & Padmapriya, T. (2024). Trade-off between training and testing ratio in machine learning for medical image processing. *PeerJ Computer Science*, *10*, e2245. https://doi.org/10.7717/peerj-cs.2245

Tabari, P., Costagliola, G., Rosa, M. D., & Boeker, M. (2024). State-of-the-Art Fast Healthcare Interoperability Resources (FHIR)–Based Data Model and Structure Implementations: Systematic Scoping Review. *JMIR Medical Informatics*, *12*(1), e58445. https://doi.org/10.2196/58445

Tang, J., Zhang, B., Liu, J., Dong, Z., Zhou, Y., Meng, X., & Toe, T. T. (2024). Pneumonia Image Classification: Deep Learning and Machine Learning Fusion. *2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 440–447. https://doi.org/10.1109/ICAIBD62003.2024.10604438

Tang, S.-T., Tjia, V., Noga, T., Febri, J., Lien, C.-Y., Chu, W.-C., Chen, C.-Y., & Hsiao, C.-H. (2023). Creating a Medical Imaging Workflow Based on FHIR, DICOMweb, and SVG. *Journal of Digital Imaging*, *36*(3), 794–803. https://doi.org/10.1007/s10278-021-00522-6

Ullas, G. S., & S, K. (2024). XAI aided Comparative Analysis of X-ray and CT Scans. *Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing*, 596–600. https://doi.org/10.1145/3675888.3676118

Varoquaux, G., Buitinck, L., Louppe, G., Grisel, O., Pedregosa, F., & Mueller, A. (2015). Scikit-learn: Machine Learning Without Learning the Machinery. *GetMobile: Mobile Comp. and Comm.*, *19*(1), 29–33. https://doi.org/10.1145/2786984.2786995

Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019). Pneumonia Detection Using CNN based Feature Extraction. *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1–7. https://doi.org/10.1109/ICECCT.2019.8869364

Walpole, R. E., & Myers, R. H. (1978). *Probability and statistics for engineers and scientists* (2d ed). Macmillan. https://cir.nii.ac.jp/crid/1130853398484951313

Wang, K. C., Kohli, M., & Carrino, J. A. (2014). Technology Standards in Imaging: A Practical Overview. *Journal of the American College of Radiology*, *11*(12), 1251–1259. https://doi.org/10.1016/j.jacr.2014.09.014

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. *arXiv:1705.02315*. https://doi.org/10.48550/arXiv.1705.02315

Wootton, D., & Feldman, C. (2014). The diagnosis of pneumonia requires a chest radiograph (x-ray)—Yes, no or sometimes? *Pneumonia*, *5*(1), Article 1. https://doi.org/10.15172/pneu.2014.5/464

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, H., Lv, Z., Liu, S., Sang, Z., & Zhang, Z. (2024). Cn2a-capsnet: A capsule network and CNN-attention based method for COVID-19 chest X-ray image diagnosis. *Discover Applied Sciences*, *6*(4), 190. https://doi.org/10.1007/s42452-024-05796-3

Zheng, Y., Xu, Z., & Wang, X. (2022). The Fusion of Deep Learning and Fuzzy Systems: A State-of-the-Art Survey. *IEEE Transactions on Fuzzy Systems*, *30*(8), 2783–2799. IEEE Transactions on Fuzzy Systems. https://doi.org/10.1109/TFUZZ.2021.3062899

Zou, L., Goh, H. L., Liew, C. J. Y., Quah, J. L., Gu, G. T., Chew, J. J., Kumar, M. P., Ang, C. G. L., & Ta, A. W. A. (2023). Ensemble Image Explainable AI (XAI) Algorithm for Severe Community-Acquired Pneumonia and COVID-19 Respiratory Infections. *IEEE Transactions on Artificial Intelligence*, *4*(2), 242–254. IEEE Transactions on Artificial Intelligence. https://doi.org/10.1109/TAI.2022.3153754