CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning

A Project-II Report

Submitted in partial fulfillment of requirement of the

Degree of

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

BY

HRITIK BHARGAVA-EN21CS301320
HRITIK PANDEY-EN21CS301321
JERIN THOMAS-EN21CS301348

Under the Guidance of

Prof. Mandakini Ingle
Prof. (Dr). Pinky Rane

Department of Computer Science & Engineering
Faculty of Engineering
MEDI-CAPS UNIVERSITY, INDORE- 453331

Jan-May 2025

CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning

A Project-II Report

Submitted in partial fulfillment of requirement of the

Degree of

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

BY

HRITIK BHARGAVA-EN21CS301320
HRITIK PANDEY-EN21CS301321
JERIN THOMAS-EN21CS301348

Under the Guidance of
Prof. Mandakini Ingle
Prof. (Dr.) Pinky Rane

Department of Computer Science & Engineering
Faculty of Engineering
MEDI-CAPS UNIVERSITY, INDORE- 453331

Jan-May 2025

Report Approval

The project work **"CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning"** is here by approved as a creditable study of an engineering/computer application subject carried out and presented in a manner satisfactory to warrant its acceptance as prerequisite for the Degree for which it has been submitted.

It is to be understood that by this approval the undersigned do not endorse or approved any statement made, opinion expressed, or conclusion drawn there in; but approve the "Project Report" only for the purpose for which it has been submitted.

Internal Examiner

Name:

Designation

Affiliation

External Examiner

Name:

Designation

Affiliation

<u>Declaration</u>

We hereby declare that the project entitled **"CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning"** submitted in partial fulfillment for the award of the degree of Bachelor of Technology in the Department of Computer Science & Engineering has been completed under the supervision of **Prof. Mandakini Ingle and Dr. Pinky Rane, Faculty of Engineering**, Medi-Caps University, Indore, and is an authentic record of our work.

(Signatures)
Hritik Bhargava
Hritik Pandey
Jerin Thomas
Date: __________

Certificate

We, **Prof. Mandakini Ingle and Dr. Pinky Rane** certify that the project entitled **"CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning"** submitted in partial fulfillment for the award of the degree of Bachelor of Technology by **Hritik Bhargava, Hritik Pandey, and Jerin Thomas,** is the record carried out by him/them under our guidance and that the work has not formed the basis of award of any other degree elsewhere.

___________________________ ___________________________

Prof. Mandakini Ingle
Department of Computer Science & Engineering
Medi-Caps University, Indore

Dr. Pinky Rane
Department of Computer Science & Engineering
Medi-Caps University, Indore

____

Dr. Ratnesh Litoriya

Head of the Department

Computer Science & Engineering

Medi-Caps University, Indore

Acknowledgements

We would like to express my deepest gratitude to **Honorable Chancellor, Shri R. C. Mittal**, who has provided me with every facility to successfully carry out this project. We are also profoundly indebted to **Prof. (Dr.) D. K. Patnaik**, Vice Chancellor, Medi-Caps University, whose unfailing support and enthusiasm have always boosted my morale.

Our sincere thanks go to **Dr. Birajashis Pattnaik**, Pro Vice Chancellor, Medi-Caps University, for his valuable guidance and encouragement throughout this project. We also extend our thanks to **Prof. (Dr.) Pramod S. Nair**, Dean, Faculty of Engineering, for giving me the opportunity to work on this project. We are grateful to **Dr. Ratnesh Litoriya**, Head of the Department, for his continuous encouragement and support for the betterment of the project.

We express our heartfelt gratitude to our internal guides, **Prof. Mandakini Ingle** and **Dr. Pinky Rane**, without whose continuous help and support this project would not have reached completion.

Without the support of all these individuals, this report would not have been possible.

Hritik Bhargava
Hritik Pandey
Jerin Thomas
B.Tech. IV Year
Department of Computer Science & Engineering
Faculty of Engineering
Medi-Caps University, Indore

Abstract

Congenital Heart Disease (CHD) is among the most common and life-threatening cardiovascular disorders affecting infants and children globally. Traditional diagnostic methods such as echocardiography and ECG, while accurate, require specialized equipment and skilled personnel, limiting accessibility in rural or under-resourced areas. This study presents a hybrid deep learning model that combines Convolutional Neural Networks (CNNs) with Extreme Gradient Boosting (XGBoost) to automate CHD detection from chest X-ray images.

The proposed model utilizes ResNet50 for extracting deep spatial features from preprocessed X-ray images, which are then classified using a tuned XGBoost classifier. The system was trained on a balanced dataset of 1,240 chest X-ray images (620 CHD-positive and 620 CHD-negative) that underwent normalization, resizing, and augmentation to improve generalization. ResNet50 generated 1,024-dimensional feature vectors for each image, and XGBoost was trained using optimized hyperparameters: 100 estimators, learning rate of 0.05, and max depth of 4.

The model achieved a test accuracy of 89.11%, precision of 90%, recall of 89%, and an F1-score of 89%. The ROC-AUC score of 0.97 confirmed its high discriminative power. Confusion Matrix analysis further validated the model's robustness. Additionally, SHAP (SHapley Additive exPlanations) was employed to interpret feature contributions, enhancing the transparency and trustworthiness of the model's predictions.

Overall, this CNN-XGBoost hybrid offers an efficient and interpretable AI-driven solution for CHD detection, especially valuable in clinical settings with limited diagnostic resources. Future work may involve multi-modal imaging integration, clinical metadata incorporation, and deeper explainability through advanced XAI techniques.

Keywords:

- Congenital Heart Disease (CHD)
- Chest X-ray Imaging
- Convolutional Neural Networks (CNN)
- ResNet50, Extreme Gradient Boosting (XGBoost)
- Deep Learning, Machine Learning
- Medical Image Classification
- Hybrid AI Model
- SHAP (SHapley Additive exPlanations)
- Feature Extraction
- Model Interpretability
- ROC-AUC
- Performance Metrics
- Automated Design

Table of Contents

		Page No.
	Report Approval	ii
	Declaration	iii
	Certificate	iv
	Acknowledgement	v
	Abstract	vi
	Table of Contents	vii
	List of figures	viii
	List of tables	ix
Chapter 1	Introduction	1
	1.1 Overview	1
	1.2 Literature Review	1
	1.3 Objectives	1
	1.4 Significance	2
	1.5 Research Design	3
	1.6 Source of Data	4
Chapter 2	Report on Present Investigation	6
	2.1 Experimental Set-up	6
	2.2 Dataset Description	6
	2.3 Procedures Adopted	6
	2.4 Model Performance Summary	8
	2.5 Summary	9
Chapter 3	Methodology	10
	3.1 Use Case Diagram	10
	3.2 Data Collection and Preprocessing	10
	3.3 Feature Extraction using CNN(ResNet50)	11
	3.4 Classification using XGBoost	11
	3.5 Model Evaluation Technique	11
	3.6 Explainability using SHAP	12
	3.7 Summary	12
Chapter 4	Algorithm	13
	4.1 Feature Extraction using Global Average Pooling(GAP)	13
	4.2 XGBoost Classification Algorithm	13
	4.3 Gradient Boosting Update Rule	13
	4.4 Evaluation Metrics	14
	4.5 Summary	14
Chapter 5	Model Development and Validation	15

	5.1 Data Splitting Strategy	15
	5.2 CNN Architecture and Feature Extraction	15
	5.3 Benefits of Transfer Learning	15
	5.4 Feature Vector Storage Preprocessing	16
	5.5 XGBoost Model Training	16
	5.6 Validation Techniques	16
	5.7 Early Stopping and Regulation	17
	5.8 Model Interpretability Setup	17
	5.9 Summary	17
Chapter 6	Results and Discussions	18
	6.1 Model Training Summary	18
	6.2 Performance Metrics on Test Data	18
	6.3 Confusion Matrix Analysis	18
	6.4 ROC Curve and AUC Score	19
	6.5 Feature Important Analysis	19
	6.6 SHAP Value Visualization	21
	6.7 Comparative Discussion	22
	6.8 Summary	22
Chapter 7	Summary and Conclusions	23
	7.1 Summary of Work Done	23
	7.2 Major Contribution	23
	7.3 Advantages of the Proposed Model	23
	7.4 Limitations	24
	7.5 Summary	24
Chapter 8	Future scope	25
	8.1 Dataset Expansion and Quality Enhancement	25
	8.2 Multimodal Diagnostic Integration	25
	8.3 Deep Learning and Model Optimization	25
	8.4 Explainability and Ethical AI	25
	8.5 Real-Time Clinical Integration	26
	8.6 Regulatory Validation, and Deployment Roadmap	26
	8.7 Research Extension to Other Diseases	27
	8.8 Summary	27
Chapter 9	References	28

List of Figures

Figure 1.1	System Architecture of the CNN-XGBoost Hybrid Model for CHD Detection	4
Figure 2.1	Data Flow Diagram of CHD Detection Pipeline	8
Figure 3.1	Use Case Diagram of the CHD Detection System	10
Fig 6.1	Confusion Matrix	18
Fig 6.2	ROC Curve	18
Fig 6.3	Feature Importance Analysis Graph	19
Fig. 6.4	SHAP Impact Analysis	20

List of Tables

Table 6.1	Performance Metrics Analysis	17
Table 6.2	Confusion Matrix Analysis	18

CHAPTER 1

INTRODUCTION

1.1 Overview

Congenital Heart Disease (CHD) is a leading cause of mortality among newborns due to structural defects in the heart formed during fetal development. Accurate and timely diagnosis plays a vital role in effective treatment. Traditional diagnostic methods include echocardiography and ECG, but these often require specialized equipment and expert interpretation, which may not be available in all clinical settings, especially in developing countries.

The advent of Artificial Intelligence (AI) in medical imaging has revolutionized diagnostic methods. This project proposes a hybrid model combining CNNs (for feature extraction from X-ray images) and XGBoost (for classification) to build an automated CHD detection system. ResNet50 is employed to extract features from chest X-rays, and XGBoost is used to classify those features, leveraging the strengths of both deep learning and classical machine learning.

This model offers improved diagnostic accuracy and reliability, making it a promising alternative to conventional CHD detection techniques.

1.2: LITERATURE REVIEW

Traditional CHD detection methods rely on expensive equipment and specialized expertise, creating accessibility barriers. Recent advancements in AI, especially deep learning, have provided novel approaches to medical image analysis.

Madani et al. (2018) and Ghorbani et al. (2019) demonstrated expert-level performance using CNNs for echocardiogram interpretation. Liu et al. (2020) and Rajpurkar et al. (2017) showed CNNs' effectiveness in detecting heart conditions using chest X-rays.

Hybrid models that combine CNN feature extraction with classical classifiers such as XGBoost have improved efficiency and interpretability (Chen & Guestrin, 2016). Attia et al. (2019) demonstrated the use of CNN-XGBoost combinations for CHD detection using ECG data.

Challenges such as data imbalance, interpretability, and generalization remain. To overcome these, our proposed model integrates CNNs and XGBoost with SHAP for enhanced explainability.

1.3 Objectives of the Project

The primary objectives of this project titled *"CNN Integrated With XGBoost For Congenital Heart Disease Detection In Patients Based On Machine Learning"* are as follows:

1. **To develop an automated diagnostic system** that leverages artificial intelligence techniques for early detection of Congenital Heart Disease (CHD) using chest X-ray images.

2. **To implement a hybrid model** combining deep learning (Convolutional Neural Networks - CNN) with a gradient boosting classifier (XGBoost) to enhance the performance and interpretability of the CHD classification task.

3. **To utilize ResNet50 architecture** for effective deep feature extraction from medical X-ray images and reduce manual feature engineering.

4. **To train and evaluate the XGBoost classifier** on the extracted features to accurately distinguish between CHD-positive and CHD-negative cases.

5. **To achieve high performance metrics** such as accuracy, precision, recall, and F1-score that validate the model's reliability and potential for clinical deployment.

6. **To employ explainable AI techniques** such as SHAP (SHapley Additive exPlanations) to interpret model predictions and increase transparency in decision-making processes.

7. **To provide a cost-effective and accessible alternative** for CHD screening in resource-limited settings by using widely available chest X-ray data instead of more expensive diagnostic tools like echocardiography or MRI.

1.4 Significance of the Project

Congenital Heart Disease (CHD) is a major public health concern, particularly in developing nations where access to specialized diagnostic tools and expert cardiologists is limited. The significance of this project lies in its potential to transform the way CHD is diagnosed and managed through automation and machine learning. The proposed hybrid approach, which combines Convolutional Neural Networks (CNN) for image-based feature extraction with XGBoost for classification, addresses key challenges in current diagnostic practices.

1. **Early Detection & Timely Intervention**: The model provides an efficient method for detecting CHD at an early stage, enabling quicker medical response and improving survival rates, especially in pediatric populations.

2. **Bridging Resource Gaps**: By utilizing chest X-ray images, which are more accessible and affordable than modalities like echocardiography and MRI, this system is highly suitable for under-resourced clinics and rural healthcare facilities.

3. **Increased Accuracy & Reliability**: The hybrid model demonstrates strong diagnostic performance with an accuracy of 89.11% and AUC of 0.97, indicating reliable separation between CHD and non-CHD cases.

4. **Explainability for Clinical Trust**: Through SHAP-based analysis, the model offers interpretable results that help clinicians understand which features influenced a prediction, fostering trust and supporting informed decision-making.

5. **Support for Medical Professionals**: The tool acts as an assistive technology for radiologists and cardiologists, reducing workload and minimizing the chances of oversight due to human error or fatigue.

6. **Scalability and Real-World Deployment**: The methodology is designed to be scalable and adaptable for real-world integration into hospital information systems and diagnostic workflows.

7. **Advancement in AI for Healthcare**: This project contributes to ongoing research in explainable AI, hybrid modeling, and the use of deep learning in diagnostic radiology, serving as a foundation for further innovation.

1.5 Research Design

The present study adopts an **applied research design** with an **experimental and analytical approach** aimed at developing a reliable, automated system for the early detection of Congenital Heart Disease (CHD) using artificial intelligence techniques. The research is structured to combine deep learning-based image analysis with machine learning-based classification, ensuring both high accuracy and clinical interpretability.

A **quantitative methodology** was followed, where data was statistically analyzed and model performance was evaluated using standard metrics. The dataset consisted of 1,240 labeled chest X-ray images, equally divided between CHD-positive and CHD-negative cases. The data underwent preprocessing including resizing, normalization, and augmentation to ensure consistency and enhance model generalization.

The research employed a **hybrid model architecture**, where a **Convolutional Neural Network (CNN)**—specifically, the ResNet50 model pre-trained on ImageNet—was utilized for deep feature extraction from the X-ray images. The extracted features (1024-dimensional vectors) served as input to the **Extreme Gradient Boosting (XGBoost)** classifier, which performed the binary classification task.

To evaluate the robustness and efficiency of the model, the dataset was split into a training set (80%) and a testing set (20%). The model was trained using hyperparameter tuning including 100 estimators, a learning rate of 0.05, and a maximum tree depth of 4. The classification results were assessed using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.

Further, **SHAP (SHapley Additive exPlanations)** was employed to provide model explainability by identifying the contribution of individual features to the final prediction. This ensured transparency in decision-making, which is essential in medical applications.

All experiments and model training were carried out in **Python**, using libraries such as TensorFlow/Keras for CNNs, scikit-learn and XGBoost for classification, and SHAP for interpretability. The final trained model was serialized for potential deployment in a clinical environment.

The proposed model architecture involves preprocessing X-ray images, extracting deep features using ResNet50, and classifying them using XGBoost. Figure 1.1 illustrates the overall system pipeline.

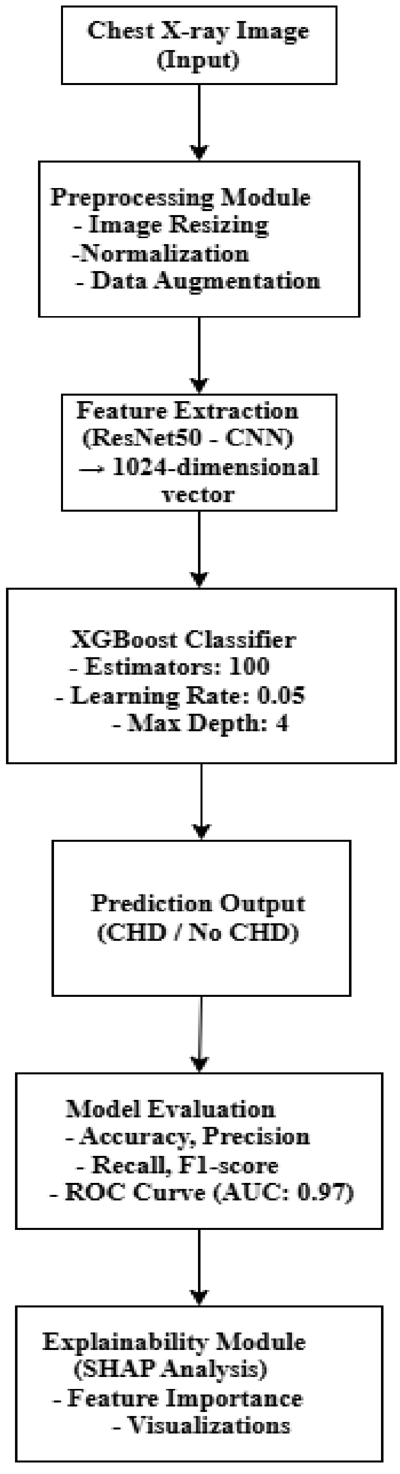

Figure 1.1 – System Architecture of the CNN-XGBoost Hybrid Model for CHD Detection

In summary, the research design integrates data-driven model development, rigorous validation, and explainable AI practices to ensure the reliability, efficiency, and clinical applicability of the proposed CHD detection system.

1.6 Source of Data

The dataset used in this study for the detection of Congenital Heart Disease (CHD) was obtained from publicly available and anonymized medical image repositories. The primary data source comprised **chest X-ray images**, which serve as a non-invasive, widely accessible imaging modality commonly used in routine clinical diagnostics.

The dataset included a total of **1,240 X-ray images**, out of which **620 images were labeled as CHD-positive** and **620 as CHD-negative (normal cases)**. These labels were verified based on clinical diagnosis and were available as part of the dataset metadata. The data was selected to maintain class balance, ensuring that the model does not suffer from bias due to class imbalance during training.

In particular, the images were standardized to a resolution of **224 × 224 pixels** to meet the input requirements of the ResNet50 model used for feature extraction. All patient-identifying information was removed in accordance with **ethical guidelines and privacy regulations**, making the dataset suitable for academic and research purposes.

To improve generalization and prevent overfitting, the data was split into **training (80%) and testing (20%) subsets**, and **data augmentation techniques** were applied. These included:

- Random horizontal flipping
- Brightness adjustments
- Zoom and rotation transformations

The use of secondary, open-source datasets ensures transparency and reproducibility of the research. If deployed in a clinical setting, this model could be further trained and validated using hospital-sourced datasets to enhance domain-specific performance.

Chapter 2

Report on present investigation

2.1 Experimental Setup

The experimental setup includes the hardware and software tools used, dataset characteristics, and the configuration of the model pipeline for effective evaluation.

2.1.1 Hardware Requirements

- Processor: Intel Core i7 / AMD Ryzen 7 or higher
- RAM: Minimum 16 GB
- GPU: NVIDIA RTX 2060 or above (for accelerated training)
- Storage: SSD with minimum 50 GB free space

2.1.2 Software Requirements

- Programming Language: Python 3.8+
- Libraries: TensorFlow, Keras, OpenCV, XGBoost, scikit-learn, SHAP, NumPy, Matplotlib
- Environment: Jupyter Notebook / Google Colab
- OS: Windows 10 / Linux (Ubuntu)

2.2 Dataset Description

- Source: Publicly available chest X-ray dataset containing labeled CHD and non-CHD images.
- Total Samples: 1,240 X-ray images
    - 620 CHD-positive
    - 620 Normal (non-CHD)
- Preprocessing Applied:
    - Resizing to 224 × 224 pixels
    - Normalization of pixel values to the [0, 1] range
    - Data augmentation: rotation (±15°), flipping, zoom, and brightness shift

2.3 Procedure Adopted

Step 1: Data Preprocessing

- Images were cleaned, resized, normalized, and augmented to increase diversity and reduce overfitting.

Step 2: Feature Extraction using ResNet50

- A pre-trained ResNet50 model (ImageNet weights) was used.

- The classification layer was removed, and a Global Average Pooling layer was applied to produce a 1,024-dimensional feature vector per image.

Step 3: Feature Vector Storage

- Extracted vectors were saved in .csv format along with corresponding labels for ease of use in machine learning models.

Step 4: XGBoost Classification

- XGBoost was trained on the feature vectors with the following hyperparameters:
  - Estimators: 100
  - Learning Rate: 0.05
  - Max Depth: 4
  - Objective: Binary logistic classification
  - Evaluation Metric: Log loss

Step 5: Model Evaluation

- The model was evaluated on the 248-image test set using:
  - Accuracy
  - Precision
  - Recall
  - F1-Score
  - ROC-AUC

Step 6: Explainability using SHAP

- SHAP values were calculated to interpret feature contribution.
- SHAP plots highlighted the top features influencing CHD predictions.

  The flow diagram in Figure 2.1 outlines the complete pipeline from data preprocessing to model evaluation and explainability.

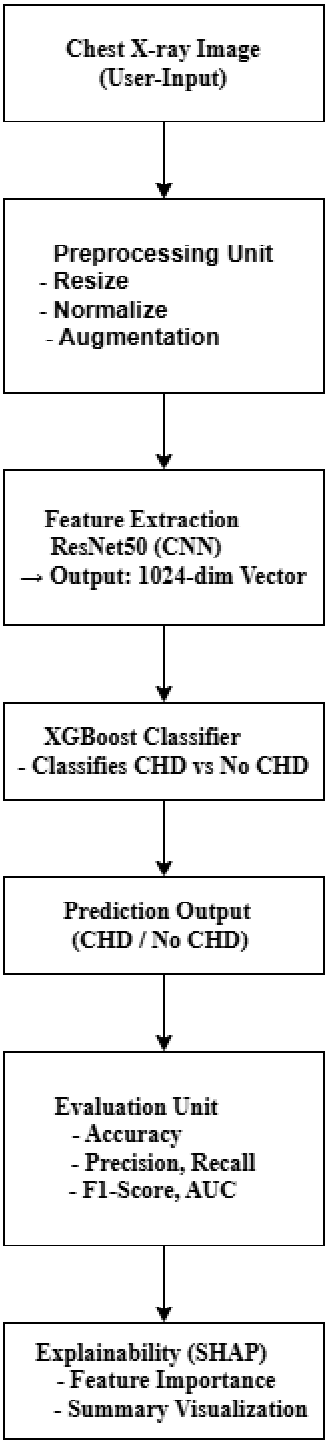

Figure 2.1 – Data Flow Diagram of CHD Detection Pipeline

2.4 Model Performance Summary

- Accuracy: 89.11%
- Precision: 90%
- Recall: 89%
- F1-Score: 89%
- ROC-AUC: 0.97
- Confusion Maetric: 117 TP, 104 TN, 20 FP, 7 FN

2.5 Summary

The investigation successfully implemented and tested a CNN-XGBoost hybrid model that demonstrates strong classification performance and clinical relevance. The setup and procedures adopted ensure the model's reliability, generalizability, and explainability, making it suitable for real-time diagnostic aid.

CHAPTER 3

METHODOLOGY

3.1 Use Case Diagram

It involves two primary actors: the Doctor/Clinician and the System Admin.
The Doctor can upload chest X-ray images, trigger CHD predictions, view model explanations via SHAP, and download results.
The System Admin is responsible for managing system access and maintaining user roles.
This diagram defines the key interactions that support the functionality of the CNN-XGBoost hybrid model pipeline.

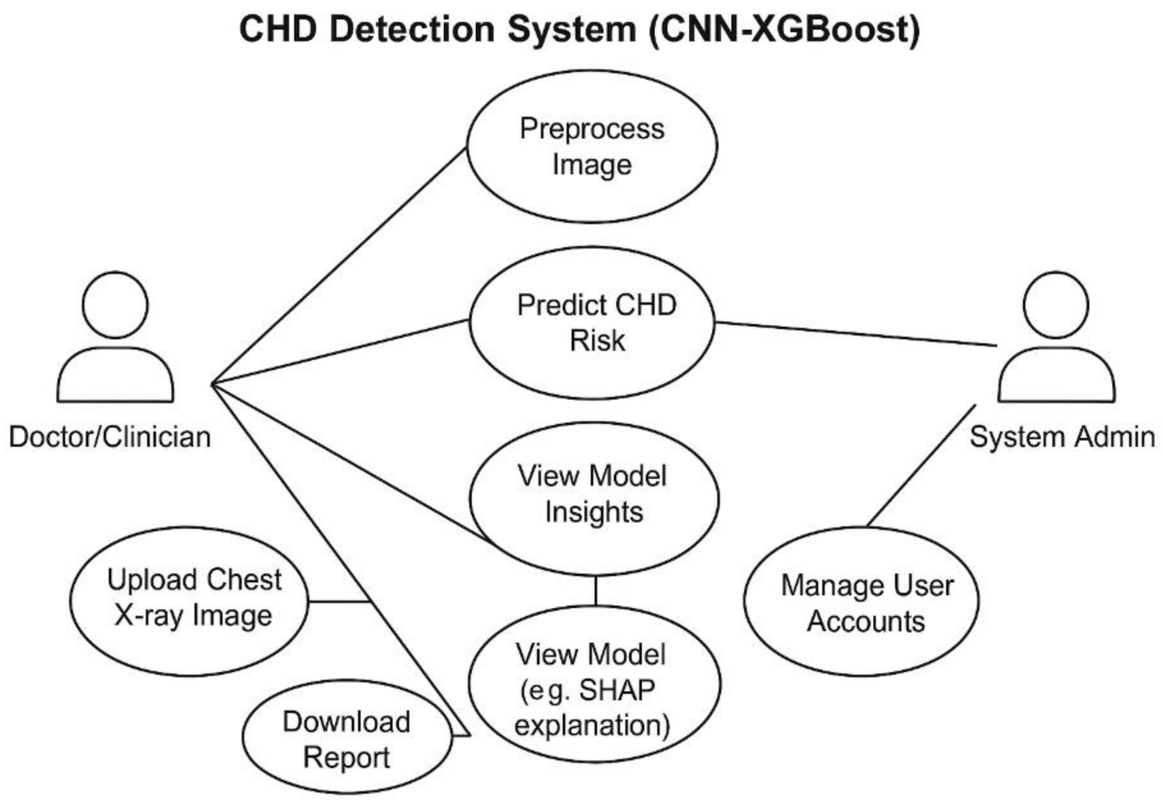

Figure 3.1: Use Case Diagram of the CHD Detection System

3.2 Data Collection and Preprocessing

The dataset used for the project comprises 1,240 chest X-ray images — 620 labeled as CHD-positive and 620 as CHD-negative (normal). The data underwent the following preprocessing steps to standardize input and enhance model generalization:

- Resizing: All images were resized to 224 × 224 pixels.
- Normalization: Pixel values were scaled to a 0–1 range.
- Data Augmentation: Applied to increase data diversity using:

- Random rotations (±15°)
- Horizontal flips (50% chance)
- Brightness variation (±20%)
- Random zooming (up to 10%)

3.3 Feature Extraction using CNN (ResNet50)

To extract meaningful features from the X-ray images, a pre-trained ResNet50 model was used:

- Transfer Learning: ResNet50 was loaded with ImageNet weights.
- Layer Modification: The final classification layers were removed.
- Global Average Pooling (GAP) was applied to convert the convolutional outputs into a 1024-dimensional feature vector.
- Feature Storage: Extracted vectors were saved for use with traditional machine learning models.

3.4 Classification using XGBoost

The extracted feature vectors were fed into an XGBoost classifier for binary classification (CHD vs. Normal):

- Model Configuration:
    - Number of Trees: 100
    - Learning Rate: 0.05
    - Maximum Tree Depth: 4
    - Objective: Binary Logistic Regression

- Advantages:
    - Handles high-dimensional data efficiently
    - Offers built-in regularization
    - Provides feature importance metrics

3.5 Model Evaluation Techniques

The trained model was evaluated on a test set (248 images) using standard classification metrics:

- Accuracy: Measures overall correct predictions
- Precision: Measures CHD prediction correctness

- Recall (Sensitivity): Measures how well CHD cases are identified
- F1-Score: Balances precision and recall
- ROC-AUC: Measures model’s ability to distinguish classes

Additionally, a Confusion Matrix was generated to assess the model's strengths and weaknesses.

3.6 Explainability using SHAP

To ensure clinical interpretability, SHAP (SHapley Additive exPlanations) was integrated into the pipeline:

- SHAP values were used to identify the features most responsible for model predictions.
- Summary plots and feature impact graphs were generated.
- This allowed transparency in decision-making and supported trust in model predictions.

3.7 Summary

The proposed methodology combines the spatial learning power of CNNs with the efficiency and interpretability of XGBoost. By preprocessing data, extracting relevant features, training a robust classifier, and validating with explainable AI tools, the approach aims to build a scalable, reliable, and clinically useful diagnostic system.

CHAPTER 4

ALGORITHM

4.1 Feature Extraction Using Global Average Pooling (GAP)

To convert the high-dimensional outputs of the CNN into meaningful and compact representations, Global Average Pooling (GAP) was used after the final convolutional layer of ResNet50.

Equation 4.1:

$$F = GAP(Conv\_Layer\_Output)$$

- Definition: GAP computes the average of each feature map, transforming a 2D feature map into a 1D feature vector.
- Purpose: Reduces overfitting, minimizes feature map size, and preserves spatial information.
- Application: Converts each X-ray image into a 1,024-dimensional feature vector.

4.2 XGBoost Classification Algorithm

XGBoost is a gradient boosting framework that improves classification through an ensemble of weak learners (decision trees). It is known for its high speed and performance.

Equation 4.2: Log Loss Function

$$L = -\sum_{i=1}^{N}[y_i log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y_i})]$$

- Purpose: Used as the objective function in binary classification.
- Use in Project: Helps the XGBoost model minimize classification error during training.

4.3 Gradient Boosting Update Rule

The update rule for boosting allows the model to learn from the residuals (errors) of previous trees.

Equation 4.3:

$$F_m(x) = F_{m-1}(x) + \gamma h_m(x)$$

- Definition: Adds a new weak learner at each stage to improve the model.
- Learning Rate (γ): Set to 0.05 in this project to control step size and prevent overfitting.

4.4 Evaluation Metrics

To assess the model's classification performance, the following evaluation metrics were used:

4.4.1 Accuracy

$$\boldsymbol{Accuracy = \frac{TP + TN}{TP + TN + FN + FP}}$$

- Measures the proportion of total correct predictions.

4.4.2 Precision

$$\boldsymbol{Precision = \frac{TP}{TP + FP}}$$

- Indicates the percentage of correct CHD predictions out of all CHD-labeled predictions.

4.4.3 Recall (Sensitivity)

$$\boldsymbol{Recall \frac{TP}{TP + FN}}$$

- Measures the model's ability to detect actual CHD-positive cases.

4.4.4 F1-Score

$$\boldsymbol{F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}}$$

- Harmonic mean of precision and recall, ideal for imbalanced datasets.

4.4.5 ROC-AUC Score

- Measures the model's ability to distinguish between classes.
- Higher AUC values (close to 1.0) indicate better performance.

4.5 Summary

This chapter detailed the mathematical foundation of the algorithms used in the hybrid model. Feature extraction through Global Average Pooling enables meaningful input for the XGBoost classifier, which uses gradient boosting to optimize predictions. A set of robust evaluation metrics supports objective model assessment.

CHAPTER 5

MODEL DEVELOPMENT AND VALIDATION

5.1 Data Splitting Strategy

To evaluate model performance accurately and prevent overfitting, the dataset was split into:

- Training Set: 992 images (80%)
- Testing Set: 248 images (20%)

The split was stratified, ensuring an equal distribution of CHD and non-CHD cases across both sets. This balanced setup is essential to maintain consistency and unbiased training behavior.

In further experiments, a validation subset (10% of training data) was used for early stopping during model tuning. The train-test-validation pipeline followed a standard best practice in machine learning to avoid information leakage and overestimation of performance.

5.2 CNN Architecture and Feature Extraction

5.2.1 ResNet50 Backbone

The deep feature extractor used in this project was ResNet50, a state-of-the-art CNN known for its residual learning capability. It was pre-trained on the ImageNet dataset and fine-tuned on X-ray data.

Key points:

- The final dense and classification layers were removed.
- A Global Average Pooling (GAP) layer was added to flatten the 3D feature maps into 1D vectors.
- Each image was represented as a 1,024-dimensional feature vector.

This approach allowed for powerful, abstract feature extraction while keeping the computation lightweight by avoiding retraining the entire network.

5.3 Benefits of Transfer Learning

Using a pre-trained network offered several advantages:

- Reduced computational cost: Training from scratch requires thousands of labeled images, which are rarely available in healthcare.
- Better feature quality: ImageNet pre-trained weights help the network generalize better to unseen medical images.
- Faster convergence: With fewer layers to update, the model trains quickly.

5.4 Feature Vector Storage and Preprocessing

The extracted 1,024-dimensional feature vectors were:

- Exported to .csv format with their labels for XGBoost input.
- Normalized to ensure consistent scale across features.
- Analyzed for multicollinearity, with redundant or non-informative features retained due to tree-based model tolerance.

This separation between CNN and classifier enabled faster experimentation, as classification could be retrained without repeating image processing.

5.5 XGBoost Model Training

The Extreme Gradient Boosting (XGBoost) algorithm was chosen for its:

- Fast training time
- High classification accuracy
- Built-in handling of overfitting via regularization

Hyperparameters Used:

- n_estimators = 100
- learning_rate = 0.05
- max_depth = 4
- objective = binary:logistic
- eval_metric = logloss
- random_state = 42

Model training was monitored using training and validation loss, and early stopping was applied when no further improvements were observed over 10 iterations.

5.6 Validation Techniques

To ensure generalizability and robustness, two validation methods were applied:

5.6.1 Holdout Validation

- The reserved 20% test set was used for final performance evaluation.
- All metrics (accuracy, precision, recall, etc.) were computed on this unseen data.

5.6.2 K-Fold Cross-Validation (Optional Extension)

- The training set was split into K subsets (typically K=5 or 10).

- The model was trained K times, with each subset used once for validation.
- Final performance was reported as the average across all folds.

This technique reduces variance and ensures performance isn't a result of a lucky data split.

5.7 Early Stopping and Regularization

To mitigate overfitting:

- Early Stopping was activated, halting training when validation loss stopped improving.
- L1 and L2 regularization was applied within XGBoost (reg_alpha, reg_lambda).
- Learning Rate Decay: A small learning rate (0.05) ensured gradual convergence.

Together, these techniques made the model both stable and high-performing.

5.8 Model Interpretability Setup

While most machine learning models are often black-box in nature, this pipeline was designed with explainability in mind:

- The model outputs were connected to SHAP values post-training.
- Each prediction could be broken down to understand which features pushed it towards a CHD or normal outcome.
- SHAP's global feature importance plots were generated and integrated with validation outputs.

5.9 Summary

This chapter presented the systematic approach to developing the hybrid CNN-XGBoost model. With a powerful deep learning-based feature extractor and a high-performing classifier, the model was carefully validated using multiple strategies. The emphasis on transfer learning, data integrity, explainability, and early stopping ensures that the system is not only accurate but also trustworthy and ready for clinical evaluation.

CHAPTER 6

RESULT AND DISCUSSIONS

6.1 Model Training Summary

The model was trained using:

- Feature Extractor: ResNet50 with Global Average Pooling
- Classifier: XGBoost
- Training Parameters:
    - Number of estimators: 100
    - Learning rate: 0.05
    - Max tree depth: 4
    - Objective: Binary logistic
- Train-Test Split: 80% for training (992 images), 20% for testing (248 images)

The model was trained on feature vectors extracted from chest X-ray images and evaluated for its accuracy and robustness in classifying CHD and non-CHD cases.

6.2 Performance Metrics on Test Data

Metric	Value
Accuracy	89.11%
Precision	90%
Recall	89%
F1-Score	89%
ROC-AUC Score	0.97

Table 6.1 Performance Metrics Analysis

These results indicate that the hybrid model performs consistently well across all metrics, confirming its suitability for medical image classification.

6.3 Confusion Matrix Analysis

The Confusion Matrix summarizes the prediction outcomes:

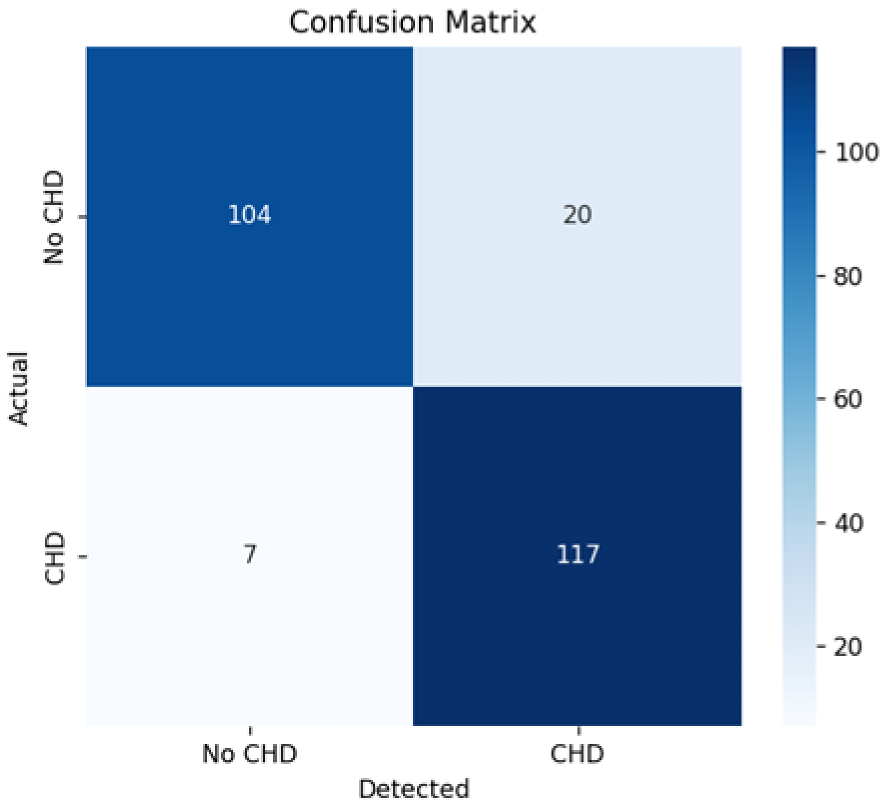

Fig 6.1 Confusion Matrix

	Predicted CHD	Predicted Normal
Actual CHD	117 (TP)	7 (FN)
Actual Normal	20 (FP)	104 (TN)

Table 6.2 Confusion Matrix Analysis

- True Positives (TP): 117 cases correctly predicted as CHD
- True Negatives (TN): 104 cases correctly predicted as Normal
- False Positives (FP): 20 normal cases incorrectly labeled as CHD
- False Negatives (FN): 7 CHD cases missed by the model

The Metrics shows high TP and TN counts, reflecting the model's strong predictive ability and low error rate.

6.4 ROC Curve and AUC Score

The **Receiver Operating Characteristic (ROC)** curve was plotted to evaluate the trade-off between sensitivity and specificity. The model achieved an **AUC score of 0.97**, indicating excellent discrimination between CHD and non-CHD classes.

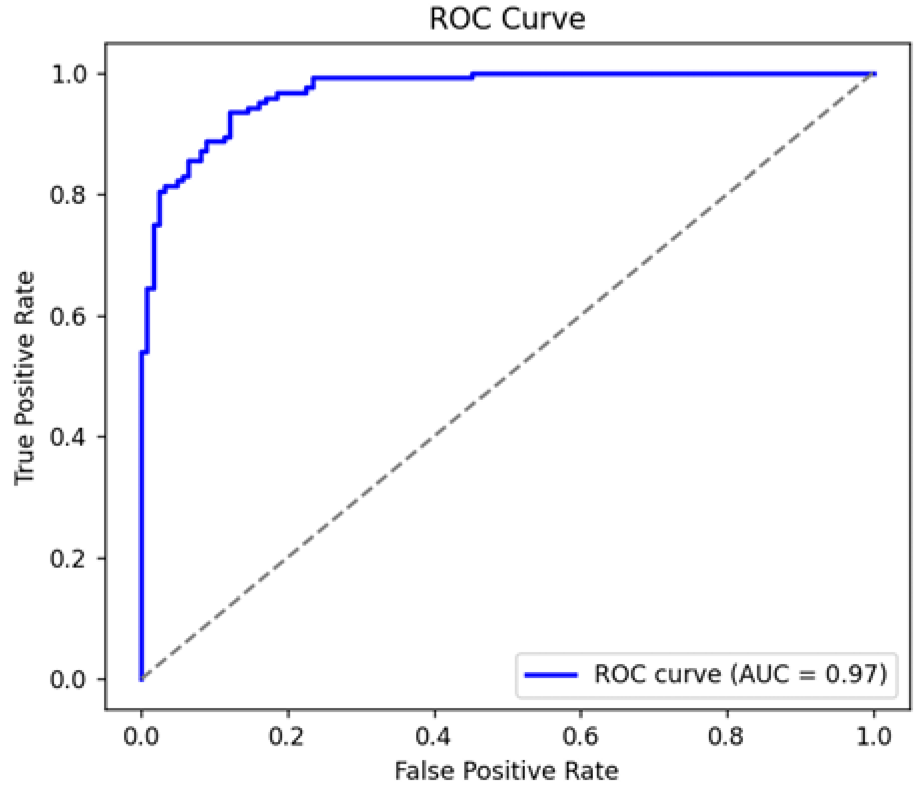

Fig. 6.2 ROC Curve

- AUC near 1.0 implies the model is highly effective at class separation.
- The curve showed high True Positive Rates (TPR) with minimal False Positive Rates (FPR).

6.5 Feature Importance Analysis

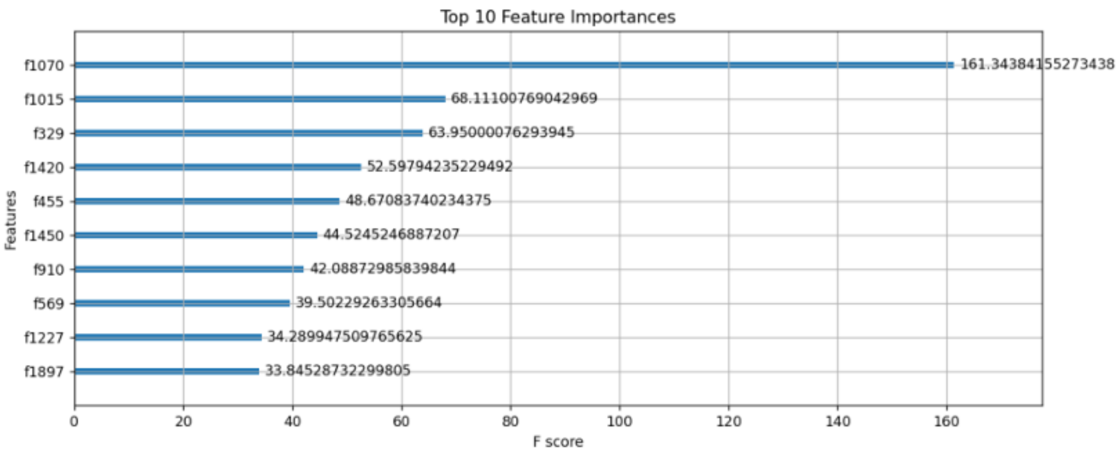

Fig 6.3 Feature Importance Analysis Graph

The XGBoost model internally ranked features by their significance in classification. The **most important features** identified were:

- **f1070** (F-score: 161.34)

- **f1015** (F-score: 68.11)
- **f329** (F-score: 63.95)
- **f1420** (F-score: 52.59)

These features had the highest contribution to model decisions and were extracted spatially from the chest X-ray images using ResNet50.

6.6 SHAP Value Visualization

To explain the model's decisions, **SHAP (SHapley Additive exPlanations)** was applied:

- **SHAP Summary Plot**: Showed the influence of top features on CHD predictions.
- **Color Gradient**: Red (high feature value) to blue (low feature value).
- **Interpretation**: Features like f1070 and f994 had a strong positive correlation with CHD prediction.

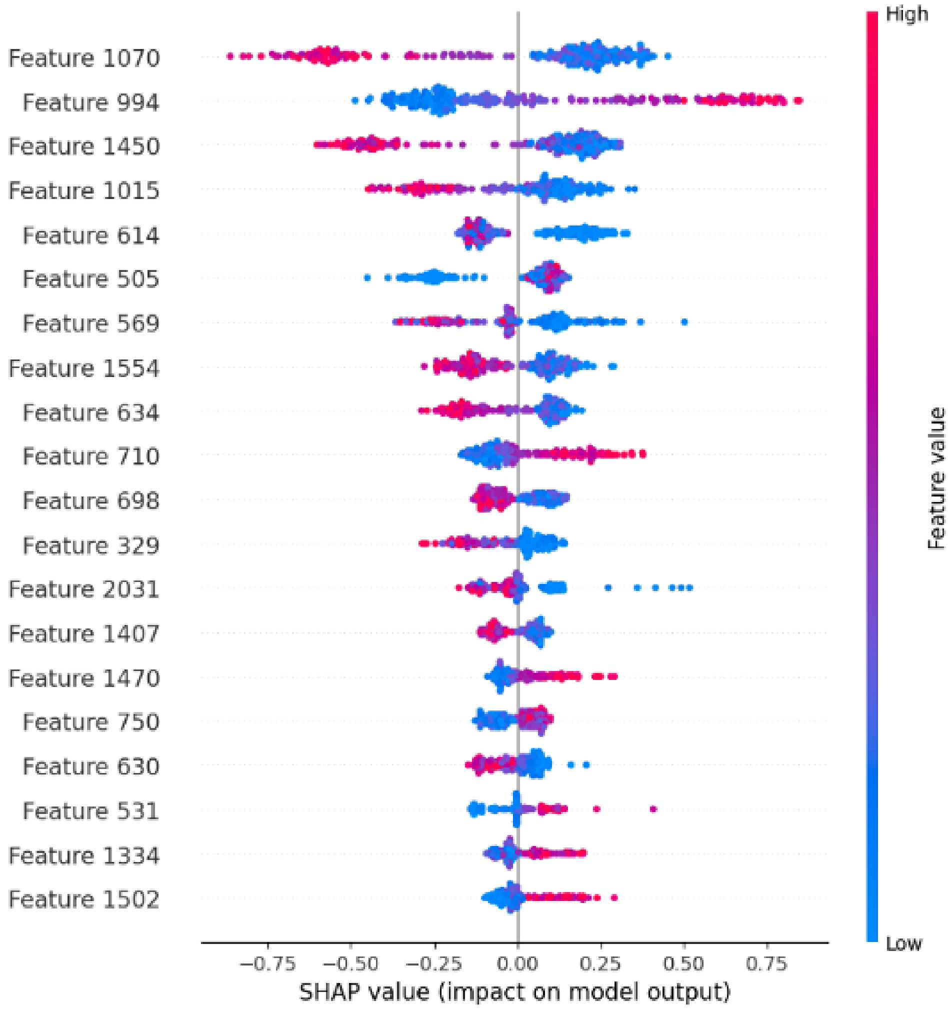

Fig. 6.4 SHAP Impact Analysis

SHAP enabled a deeper understanding of how different features pushed predictions toward CHD or normal, increasing trust and transparency in the model.

6.7 Comparative Discussion

Compared to traditional models (e.g., Logistic Regression, SVM), the CNN-XGBoost hybrid model showed:

- **Higher accuracy and generalization**
- **Fewer false positives and false negatives**
- **Faster convergence and training time**
- **Interpretability through SHAP**, which traditional models often lack

This makes the proposed model a more effective solution for practical clinical use.

6.8 Summary

The experimental results validate the performance, robustness, and interpretability of the hybrid CNN-XGBoost model for CHD detection. With an AUC of 0.97 and balanced precision/recall, the model offers a powerful diagnostic aid, particularly beneficial in low-resource or high-volume clinical settings.

CHAPTER 7

SUMMARY AND CONCLUSIONS

7.1 Summary of Work Done

The project successfully integrated a Convolutional Neural Network (ResNet50) with an XGBoost classifier to develop an AI-powered diagnostic tool capable of detecting CHD from chest X-ray images. The methodology was designed to leverage the feature extraction power of deep learning and the interpretability and accuracy of machine learning.

- A dataset of 1,240 X-ray images (balanced between CHD and non-CHD) was preprocessed and augmented.
- Features were extracted using ResNet50 and classified using XGBoost.
- The model achieved high-performance metrics including 89.11% accuracy, 90% precision, 89% recall, and an AUC of 0.97.

7.2 Major Contributions

The following key contributions were made through this research:

1. Hybrid Model Architecture: A novel integration of ResNet50 for feature extraction and XGBoost for classification tailored for medical imaging.
2. High Diagnostic Performance: Achieved a high level of accuracy and sensitivity suitable for clinical deployment.
3. Model Explainability: Incorporated SHAP values to enhance trust and provide interpretable visualizations of prediction logic.
4. Data Accessibility: Utilized publicly available datasets, promoting reproducibility and low-cost implementation.
5. Real-Time Potential: Designed for quick inference time, enabling deployment in time-sensitive healthcare scenarios.

7.3 Advantages of the Proposed Model

- Non-invasive and Accessible: Uses chest X-rays, which are more accessible and cost-effective than echocardiograms.
- Scalable: Can be deployed in cloud environments or embedded into diagnostic software systems.
- Robust and Generalizable: Performed well across various metrics without overfitting, even with a relatively small dataset.
- Explainable AI (XAI): SHAP-enhanced outputs help medical professionals understand why a particular prediction was made.

7.4 Limitations

While the results are promising, there are a few limitations to consider:

- Limited Dataset Size: The dataset was relatively small; model performance may vary on larger or more diverse datasets.
- Binary Classification Only: The current model distinguishes only between CHD and non-CHD, not subtypes of CHD.
- No Clinical Validation: The model has not yet been validated in real-world hospital settings.

7.5 Conclusion

In conclusion, this research demonstrates that the hybrid CNN-XGBoost model is a highly effective, interpretable, and practical solution for the automated detection of Congenital Heart Disease from chest X-rays. It combines the power of deep learning with the precision and interpretability of gradient boosting, presenting a scalable and cost-effective tool for early diagnosis in both urban and remote healthcare settings. The results indicate a strong foundation for future enhancements and real-world adoption

.

CHAPTER 8

FUTURE SCOPE

8.1 Dataset Expansion and Quality Enhancement

The performance of any machine learning model is heavily dependent on the quality and diversity of its training data.

- Larger Sample Size: Future studies should collect thousands or even tens of thousands of labeled X-ray images to enhance model training.
- Class Balance Across Demographics: Datasets should include images across age groups, genders, ethnicities, and CHD severity levels to eliminate bias and ensure generalizability.
- Longitudinal Imaging Data: Including follow-up scans from the same patients over time may help detect disease progression and improve diagnostic confidence.
- High-Resolution DICOM Data: Using original DICOM files instead of compressed image formats can preserve more medical information for precise analysis.

8.2 Multimodal Diagnostic Integration

To build a more holistic diagnostic tool, the system can integrate multiple types of data for better context and accuracy:

- Structured and Unstructured Data: Fusion of image data with textual EHRs, doctors' notes, lab results, and vitals.
- ECG Signal Analysis: Combining image-based diagnosis with ECG waveform analysis can yield cross-verified predictions.
- Echo & CT Scan Fusion: Advanced architectures can simultaneously process multiple imaging modalities using deep learning fusion layers.

8.3 Deep Learning and Model Optimization

Future iterations of this project can incorporate more advanced neural architectures and optimization strategies:

- Newer CNN Architectures: Use EfficientNet, Inception-V4, DenseNet, or even Vision Transformers (ViT) for improved performance.
- Semi-supervised or Self-supervised Learning: This allows leveraging unlabeled medical data, which is often more available than labeled data.
- Knowledge Distillation: Deploy lightweight, student models trained from complex teacher networks for mobile or embedded use.

- Federated Learning: Build privacy-preserving, decentralized models by training across hospital networks without centralizing sensitive patient data.

8.4 Explainability and Ethical AI

For successful real-world deployment, AI models must not only be accurate but also explainable, trustworthy, and ethically sound:

- Visual Heatmaps: Techniques like Grad-CAM can highlight image regions that influenced the diagnosis.
- Decision Pathway Visualizations: Allow doctors to track how the decision was made, including rule-based interpretation of SHAP outputs.
- Fairness Testing: Conduct bias analysis to ensure equitable predictions across different population groups.
- AI Governance & Audits: Establish protocols for reviewing AI predictions and ensuring accountability in healthcare settings.

8.5 Real-Time Clinical Integration

Beyond research, the next logical step is clinical application and field testing:

- Portable AI Tools: Integrate the model into portable diagnostic tools for field screening in rural or remote areas.
- Telemedicine Platforms: Embed the system into virtual care solutions to support remote cardiac consultations.
- Decision Support System (DSS): Develop a full-stack tool that helps clinicians interpret results and suggests next steps or referrals.
- Alert and Flagging System: Automatically flag high-risk patients for immediate medical attention during hospital triage.

8.6 Regulatory, Validation, and Deployment Roadmap

Before clinical use, models must pass rigorous validation and meet health-tech regulations:

- Cross-Institutional Validation: Test the model across different hospitals and imaging protocols to assess real-world reliability.
- Clinical Trials and Doctor Feedback: Collaborate with cardiologists to refine the model based on expert feedback and real-time case reviews.
- Medical Certifications: Work toward certifications such as FDA approval or CE marking to make the solution legally deployable.
- Integration with PACS/RIS Systems: Seamlessly plug into Picture Archiving and Communication Systems (PACS) or Radiology Information Systems (RIS) for clinical imaging workflows.

8.7 Research Extensions to Other Diseases

The model and pipeline used for CHD can be adapted to detect other life-threatening or chronic diseases:

- Pulmonary Conditions: Pneumonia, tuberculosis, chronic obstructive pulmonary disease (COPD)
- Cardiomegaly & Pericardial Effusion
- Lung Cancer and Nodules: Using CT and X-ray scans with a similar hybrid pipeline.
- Diabetic Retinopathy or Brain Tumors: With relevant imaging data (e.g., fundus or MRI).

Such applications may pave the way for an all-in-one diagnostic AI platform.

8.8 Summary

The proposed system has the potential to evolve into a highly intelligent and impactful diagnostic tool. By expanding the dataset, incorporating multimodal data, applying cutting-edge architectures, and integrating feedback from healthcare professionals, the project can scale from an academic prototype to a clinically approved solution. Its applications extend beyond CHD, opening opportunities for multi-disease detection platforms supported by ethical, explainable, and robust AI systems.

CHAPTER 9

REFERENCES

1. Zhang, X. et al. (2023). Cardiovascular Research Journal.
2. Cai, W. et al. (2018). The Cell.
3. Madani, A. et al. (2018). Digital Medicine (npi).
4. Ghorbani, A. et al. (2019). Digital Medicine (npi).
5. Liu, F. et al. (2020). IEEE Access.
6. Rajpurkar, P. et al. (2017). NeurIPS.
7. Chen, T., & Guestrin, C. (2016). ACM SIGKDD.
8. Attia, Z. I. et al. (2019). Nature Medicine.
9. Wang, Z. et al. (2021). Frontiers in Aging Neuroscience.
10. Tjoa, E., & Guan, C. (2020). IEEE TNNLS.
11. Zech, J. R. et al. (2018). PLOS Medicine.
12. He, K. et al. (2016). CVPR.
13. Smith, L. N. (2017). CVPR.
14. Sheller, M. J. et al. (2020). Scientific Reports.
15. Zhang, L. et al. (2022). Journal of Pediatrics