Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 06, June - 2025

Heart Disease Detection using Machine Learning Classification Techniques in E- Healthcare Systems

Jayakumar M,
Department of Electronics Engineerning,
Madras Institute of Technology,Anna University,
Chennai,TamilNadu,India.

Dr Sridevi C, Associate Professor,
Department of Electronics Engineerning, Madras
Institute of Technology,Anna University,
Chennai,TamilNadu,India.

Dhanushkumar K,
Department of Electronics Engineerning,
Madras Institute of Technology,Anna University,
Chennai,TamilNadu,India.

Surendrakumar S,
Department of Electronics Engineerning, Madras
Institute of Technology,Anna University,
Chennai,TamilNadu,India.

Abstract—**Heart disease continues to be a major cause of mortality worldwide, making early detection vital for successful treatment and prevention. As digital healthcare solutions become more prevalent, there is an increasing demand for intelligent systems to aid in medical diagnosis, particularly in regions with limited access to expert healthcare professionals. This project introduces a machine learning-based method for detecting heart disease through medical images like MRI and CT scans. Rather than depending on conventional health metrics such as blood pressure or cholesterol levels, this system emphasizes the analysis of visual patterns in heart images. By employing deep learning techniques, specifically Convolutional Neural Networks (CNNs), the model learns to differentiate between healthy and diseased heart conditions. Pre-trained architectures like ResNet or EfficientNet are fine-tuned on a labeled dataset to enhance performance and ensure high classification accuracy. The main objective of this project is to support e-healthcare platforms by offering a rapid, dependable, and automated method for heart disease detection. By incorporating this image-based classification model into remote healthcare services, we aim to assist medical professionals in making faster decisions, ultimately improving patient outcomes. To further refine diagnostic accuracy, this system also classifies affected heart conditions into specific disease types such as Myocardial Infarction, Coronary Artery Disease, Valve Disorder, Arrhythmia, and Cardiomyopathy. This level of classification allows for targeted clinical intervention, providing valuable insights that extend beyond simple detection. Additionally, by utilizing transfer learning, the model achieves high accuracy even with limited training data, making it a practical and scalable solution for real-world medical applications.**

Keywords—**Heart Disease Prediction, Machine Learning in Medical Imaging, DICOM Image Classification, Support Vector Machine for Diagnosis, Pseudo-Labeling Techniques, Feature Selection in Healthcare AI, e-Healthcare Decision Support Systems**

I. INTRODUCTION

Heart disease encompasses a wide range of cardiovascular issues that impact the heart's structure and its ability to function properly. This includes conditions affecting the heart muscles, valves, blood vessels, and the electrical system that controls heartbeats. The World Health Organization (WHO) reports that cardiovascular diseases are the top cause of death worldwide, responsible for nearly 18 million deaths each year. Despite progress in medical science, early detection of heart disease remains difficult due to the complexity of symptoms and the variability in how the disease manifests in different individuals. Traditionally, diagnosing heart disease involves clinical tests like electrocardiograms (ECGs), echocardiography, blood tests, and invasive procedures such as angiograms. Recently, medical imaging techniques, particularly Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), have become valuable tools for obtaining detailed structural and functional information about the heart. These imaging methods provide non-invasive ways to evaluate cardiac conditions, but interpreting these images still requires skilled radiologists and cardiologists. Manual analysis is time-consuming and prone to human error, especially when subtle abnormalities are present. With the advent of artificial intelligence and machine learning, there is increasing interest in creating systems that can automatically analyze medical images and aid in diagnosing diseases. This project investigates how machine learning, especially deep learning models, can be trained on heart MRI and CT scan images to accurately detect and classify heart diseases. The aim is to offer a supportive tool for e-healthcare platforms, facilitating faster diagnosis in remote or resource-limited areas and ultimately easing the burden on healthcare systems.

II. LITERATURE SURVEY

J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare" This study highlights the critical role of precise heart disease detection for prompt medical treatment. The authors introduce a machine learning-based strategy for predicting heart disease, employing various classifiers such as Decision Trees (DT), Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN). The dataset comprises patient details like age, gender, type of chest pain, and cholesterol levels. The research assesses the performance of these classifiers, comparing their accuracy, sensitivity, and specificity. The findings suggest that machine learning techniques can serve as a dependable tool for early diagnosis, though they stress the importance of optimizing models to enhance prediction accuracy. The paper also emphasizes the significance of e-healthcare in automating diagnostics and enhancing patient care.

G. G. Geweid and M. A. Abdallah, “A New Automatic Identification Method of Heart Failure Using Improved Support Vector Machine Based on Duality Optimization Technique” In this paper, the authors introduce an automatic system for identifying heart failure using an enhanced Support Vector Machine (SVM) model improved by duality optimization. While traditional SVM methods are effective, they can be constrained by their sensitivity to parameters and dataset size. The duality optimization technique is applied to fine-tune the SVM’s hyperparameters and boost its performance in heart failure prediction. The method is tested on multiple heart disease datasets, and the results indicate that it surpasses standard SVM models in classification accuracy and robustness. This study highlights the potential of SVMs in heart disease detection and the necessity of optimization techniques to improve predictive reliability.

S. Mohan, C. Thirumalai, and G. Srivastava, “Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques” This paper investigates the combination of multiple machine learning algorithms to enhance the accuracy of heart disease prediction. The authors propose a hybrid approach that merges the strengths of various classifiers like SVM, k-NN, and Decision Trees to create an ensemble model. The research demonstrates that integrating these models significantly improves prediction performance compared to using individual classifiers. The ensemble method employs a weighted voting scheme to determine the final prediction based on the outputs of the constituent models. The study illustrates the importance of hybrid models in managing complex datasets and achieving superior classification results in heart disease prediction.

A. U. Haq, J. P. Li, M. H. Memon, A. Malik, T. Ahmad, A. Ali, S. Nazir, I. Ahad, M. Shahid, et al., “Feature Selection Based on L1-Norm Support Vector Machine and Effective Recognition System for Parkinson’s Disease Using Voice Recordings” This paper centers on the use of machine learning for diagnosing Parkinson’s disease through voice recordings. It presents a novel feature selection method based on L1-norm SVM, which aids in identifying the most pertinent features from the voice data for classification.The research utilizes a recognition system that integrates the L1-norm SVM with additional classifiers to enhance the accuracy of Parkinson’s disease detection. The authors emphasize the difficulty of feature selection in medical datasets and suggest a robust framework to boost diagnostic precision. Although the study centers on Parkinson's disease, the methods discussed are adaptable for heart disease classification, especially in terms of feature selection and model optimization.

J. Li and H. Liu, “Challenges of Feature Selection for Big Data Analytics,” IEEE Intelligent Systems, vol. 32, no. 2, pp. 9–15, 2017 Li and Liu explore the challenges and strategies of feature selection in big data analytics. The paper underscores the necessity of choosing the most pertinent features from extensive datasets to enhance the efficiency and accuracy of machine learning models. In the realm of heart disease prediction, the paper's insights into feature selection can be applied to pinpoint the most critical medical attributes (e.g., cholesterol levels, heart rate, etc.) that influence prediction accuracy. The authors also examine various feature selection algorithms and their effectiveness in big data contexts, providing valuable guidance for enhancing healthcare prediction models' performance.

A. U. Haq, J. Li, M. H. Memon, J. Khan, S. U. Din, I. Ahad, R. Sun, and Z. Lai, “Comparative Analysis of the Classification Performance of Machine Learning Classifiers and Deep Neural Network Classifier for Prediction of Parkinson’s Disease,” 2018 In this study, the authors evaluate the classification performance of machine learning models and deep neural networks (DNN) for predicting Parkinson’s disease. The research includes various classifiers such as SVM, Naive Bayes, and Decision Trees, alongside DNNs. The findings indicate that while traditional machine learning models perform well, DNNs significantly surpass them in classification accuracy, particularly with large datasets. This paper offers insights into the effectiveness of deep learning models like DNNs, which can be applied to heart disease prediction tasks, where the complexity of medical images or patient data requires advanced deep learning techniques for improved accuracy.

A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, “A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms,” Mobile

Information Systems, vol. 2018, 2018 This paper introduces a hybrid intelligent system framework for heart disease prediction, which combines various machine learning algorithms, including SVM, Random Forest, and Decision Trees. The hybrid approach seeks to enhance predictive accuracy by utilizing the strengths of multiple classifiers. The study also incorporates a feature selection process to improve model efficiency and reduce computational complexity. The authors demonstrate that their hybrid system outperforms traditional single-algorithm classifiers in terms of accuracy and robustness. This hybrid approach is particularly beneficial in the context of heart disease prediction, where diverse datasets and multiple input features can gain from an ensemble method to improve accuracy.

III. DATASETS USED

The dataset used in this project is a large and diverse collection of anonymized DICOM images taken from CT and MRI scans of the human heart. It includes images showing various cardiac conditions and anatomical views, making it highly valuable for medical research and machine learning applications. Most images range in resolution from 256×256 to 512×512 pixels and come with useful metadata like scan type and orientation. In some cases, the dataset also includes labels for specific heart diseases, such as cardiomyopathy or myocardial infarction, and segmentation maps to support image analysis tasks. All personal information has been carefully removed to protect patient privacy, though general demographic details like age range and gender may still be available. Overall, this dataset provides a strong foundation for developing and testing deep learning models focused on cardiac image classification, segmentation, and abnormality detection.

IV. METHODOLOGY

A. Existing Methodology

The proposed research presents an intelligent system designed for precise heart disease (HD) detection through a comprehensive approach that includes data preprocessing, optimal feature selection, and machine learning classification. This system is built upon the Cleveland Heart Disease dataset. To assess the model's reliability, a stringent validation method—Leave-One-Subject-Out Cross-Validation (LOSO CV)—is utilized. The system's performance is evaluated using various metrics such as accuracy, sensitivity, specificity, Matthews Correlation Coefficient (MCC), and processing time.The primary goal is to improve clinical decision-making by offering a lightweight, accurate, and efficient automated diagnostic tool. By minimizing data dimensionality and optimizing model inputs, the system ensures precise classification of patients with or without heart disease. The binary classification output is denoted as '1' for the presence of heart disease and '0' for its absence

A. Block Diagram

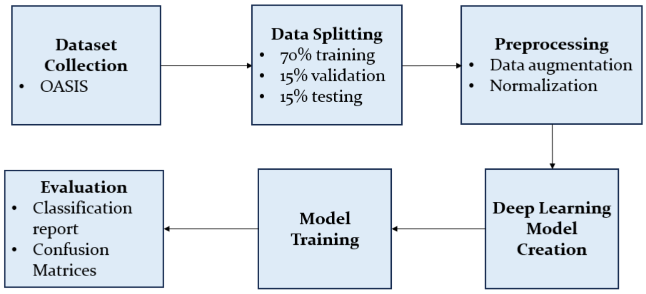


Fig. 1 Block Diagram

The block diagram depicts the sequential process for classifying heart disease from DICOM images using deep learning methods. The procedure starts with Image Acquisition, where a set of DICOM images is collected from pertinent datasets. Along with the images, essential metadata such as patient age and pseudo-labels (for instance, derived from pixel intensity or clinical indicators) are extracted. Next is Data Preprocessing, a vital phase where the raw DICOM pixel data is standardized and resized to a consistent format, which is typically necessary for deep learning models. Additional procedures like image augmentation (including rotation, flipping, and shifting) are employed to artificially enlarge the dataset and minimize overfitting. Next, the dataset is split into training and validation sets, typically using an 80:20 ratio. It's crucial to maintain class balance to avoid bias and ensure that data from the same patient doesn't appear in both sets to prevent data leakage. In the Machine Learning Model phase, a pre-trained convolutional neural network (CNN) architecture, such as EfficientNetB0, ResNet50, or DenseNet121, is selected and configured. These models, originally trained on extensive datasets like ImageNet, are fine-tuned through transfer learning to identify features in medical images. During the Model Training phase, the network is trained on preprocessed data using data generators to efficiently manage memory and batch sizes. Throughout this process, metrics like accuracy and loss are monitored across epochs to ensure effective learning and prevent overfitting. Finally, in the Evaluation and Visualization step, the model's performance is evaluated using tools like the confusion matrix and statistical measures such as accuracy and F1-score. Visualization techniques are also employed to display predictions on actual images, aiding in interpreting how well the model differentiates between normal and affected cases.

B. Data Preprocessing

The original Cleveland dataset contains 303 records with 75 attributes. After comprehensive data cleaning and normalization, the dataset is refined to 297 instances and 13 highly relevant features. The preprocessing phase involves addressing missing values, eliminating redundant or irrelevant attributes, and applying normalization techniques like Standard Scalar and Min-Max Scalar. These methods ensure uniform scaling of all features, enhancing the learning performance of machine learning models.Effective preprocessing enables the classifier to concentrate on significant data patterns while avoiding bias from unbalanced feature scaling. It also substantially reduces the risk of overfitting, resulting in more generalizable outcomes across different subjects.

To boost model efficiency and lower computational costs, the system employs five feature selection techniques. Four standard methods—Relief, Minimum Redundancy Maximum Relevance (mRMR), Least Absolute Shrinkage and Selection Operator (LASSO), and Low-Level Best First Search (LLBFS)—are utilized. Additionally, a novel technique called Fast Conditional Mutual Information Maximization (FCMIM) is introduced.FCMIM functions by selecting features with the highest mutual information in a non-redundant manner. This method captures the most relevant features that significantly contribute to heart disease detection, even in high-dimensional datasets. It greatly enhances the precision and speed of machine learning models while maintaining clinical reliability.

C. Models

Six classifiers are implemented and tested for performance: Support Vector Machine (SVM) with linear and RBF kernels, Logistic Regression, k-Nearest Neighbors (k-NN), Artificial Neural Networks (ANN), Naïve Bayes, and Decision Trees. Each classifier is trained on different subsets of selected features from the five feature selection methods, ensuring optimal input for each model. These classifiers encompass a broad range of machine learning paradigms probabilistic models, neural networks, instance-based learning, and linear separators—making the evaluation comprehensive and robust. This diversity allows for identifying the best-performing model across various metrics.

LOSO CV is employed to validate model performance. In this approach, each subject is used as a test instance once, while the remaining subjects are used for training. This results in a robust estimation of generalization ability, particularly useful in medical datasets with limited instances.

D. Training Process

Data Loading And Preprocessing:

The initial phase of the pipeline begins with loading chest DICOM images from the designated dataset directory. This is done using the pydicom library, where each image file is parsed to extract its pixel data via the pixel array attribute. Given that DICOM images often exhibit varied intensity distributions, a min-max normalization is applied to scale the pixel values uniformly within the 0–255 range. This normalization ensures consistency across all images, facilitating more stable training. Following this, each image is resized to 224×224 pixels using OpenCV's cv2.resize() function to match the required input dimensions for the selected deep learning architecture .Since the original DICOM images are typically grayscale (i.e., single-channel), and the model expects them to be reshaped and color-converted, the images are stored in a NumPy array and further normalized to the [0, 1] range (using float32 type).This final normalization step enhances training performance and model convergencehree- channel RGB input, grayscale images are converted into 3-channel format using cv2.cvtColor()Once the images are reshaped and their colors converted, they are stored in a NumPy array and normalized to a range of [0, 1] using the float32 type. This final normalization step improves training performance and aids in model convergence.

PSEUDO-LABEL GENERATION:

In the absence of ground truth clinical annotations, such as confirmed diagnoses, a pseudo-labeling method is used to simulate supervision. For each image, the mean pixel intensity is calculated after RGB conversion and normalization. This mean intensity acts as a rough indicator of potential abnormalities. To create a binary classification system, the global median of these intensity values across the entire dataset is determined. Images with intensity values above the median are given a pseudo-label of 1 (Heart Disease Affected), while those at or below the median are labeled 0 (Normal). Although not medically definitive, this technique provides a weak but useful signal that allows the model to start learning discriminative patterns.

TRAIN-TEST SPLITTING:

With pseudo-labels assigned, the dataset is divided into training and validation subsets using scikit-learn's train_test_split function. An 80:20 ratio is used, ensuring that most of the data is used for training while a portion is reserved for model evaluation. Stratification is applied based on the pseudo-labels to maintain class balance in both subsets, helping to prevent the model from becoming biased toward a dominant class. Along with the image arrays and labels, associated metadata such as pseudo ages and DICOM filenames are also split accordingly, enabling downstream performance tracking and visualization.

DATA AUGMENTATION:

To enhance model generalization and reduce overfitting, data augmentation is applied to the training images using Keras' ImageDataGenerator. This augmentation simulates real-world variations through random image transformations, including:

1. Rotations of up to ±20° (to simulate varied patient orientations)
2. Zooming in/out by 20% (to mimic differences in scan magnification)
3. Shifts in width and height (10%) to account for misaligned imaging
4. Horizontal flipping, which introduces variation in anatomical symmetry.

The augmented images provide more diverse training inputs, encouraging the model to learn robust, invariant features. Notably, the validation data is not augmented; it is only rescaled to maintain consistency for fair performance evaluation.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 06, June - 2025

GLOBAL AVERAGE POOLING 2D:

Global Average Pooling 2D condenses each feature map into a single value by averaging its spatial elements. Instead of flattening and employing dense layers, it produces a vector where each element is

DROPOUT (REGULARIZATION TECHNIQUE):

Dropout is a regularization technique that randomly "drops" (sets to zero) a portion of the neurons during training to prevent overfitting. This compels the network not to depend on specific neurons, enhancing generalization. Where p is the keep probability (e.g., 0.5), and ri and xi is a binary mask. During inference, dropout is disabled, and outputs are scaled accordingly

$$\tilde{x}i = xi \cdot ri \qquad ri \sim Bernoulli(p)$$

RELU ACTIVATION FUNCTION (RECTIFIED LINEAR UNIT):

ReLU is a non-linear activation function that replaces all negative values with zero while leaving positive values unchanged. ReLU aids the network in learning complex patterns while being computationally efficient. It is widely used due to its simplicity and effectiveness in avoiding the vanishing gradient problem

$$f(x) = max(0, x)$$

MODEL DESIGN COMPARISON

ARCHITECTURE STYLE:

MODEL	DESIGN	STRUCTURE
RESNET 50	Residual learning through skip connections	Deep network with 50 layers; uses residual blocks and identity shortcuts
DENSENET 121	Dense connections between all previous and current layers	Encourages feature reuse in a compact architecture
EFFICIENTNETB0	Compound scaling of depth, width and resolution	Uses Squeeze and Excitation and MBConv block for efficiency

PARAMETERS AND SIZE:

MODEL	DEPTH	PARAMETERS	SPEED
RESNET50	50 layers	~25 million	Moderate
DENSENET-121	121 layers	~8 million	Slow(Due to dense connections)
EFFICIENTNETB0	237 layers (lightweight)	~5.3 million	Fastest(Highly optimized)

PERFORMANCE AND USES:

MODEL	ACCURACY	OVERFITTING	BEST FOR
RESNET50	High	Less(residuals help generalization)	General purpose in image processing
DENSENET- 121	Higher than ResNet50	Less(Due to feature reuse)	Memory efficient tasks
EFFICIENT-NETB0	Very High	Less(Due to regularization and scaling)	Mobile/Embedded systems, highly efficient tasks

MODEL COMPILATION:

The model is compiled with the Adam optimizer, chosen for its adaptive learning rate capabilities that adjust during training. The binary cross-entropy loss function is employed, suitable for binary classification tasks. Performance is tracked using accuracy as the primary metric. To enhance training dynamics, a ReduceLROnPlateau callback is incorporated. This scheduler monitors the validation loss and automatically reduces the learning rate if improvements plateau, aiding the model in converging to better optima.

ADAM OPTIMIZER:

Adam is an optimization algorithm that merges the benefits of Momentum and RMSProp. It maintains running averages of both gradients (first moment) and their squares (second moment) to adapt learning rates and is widely used in deep learning due to its fast convergence and minimal tuning.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) y^2$$

$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

MODEL TRAINING:

Training is conducted over 50 epochs using the augmented training data and validated against the unaugmented validation set. The model trains in mini-batches, updating its internal weights with each pass to minimize the binary cross-entropy loss. Throughout training, both accuracy and loss are monitored for the training and validation sets. The learning rate is dynamically reduced when necessary, using the ReduceLROnPlateau strategy, which activates if the validation loss does not improve for three consecutive epochs. This adaptive approach stabilizes learning and reduces the risk of overfitting or early stagnation.

E. Performance Evaluation

The evaluation process of an image classification model involves assessing how well the model performs on a given task of classifying images into predefined categories. It is essential to determine how accurate and reliable the model is on both the training data (used for learning) and unseen test data (used for evaluation). Here's an overview of the key steps and metrics in the evaluation process:

1) Precision: Precision quantifies the accuracy of the model in predicting a specific emotion class. It is the ratio of true positive predictions to the total predicted positives for that class. High precision indicates a low rate of false positives.

$$Precision = \frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Positives(FP)}$$

2) ***Recall:*** Recall measures the model's ability to correctly identify all instances of a particular class. It is the ratio of true positive predictions to the total actual positives for that class. High recall indicates a low rate of false negatives.

$$Recall = \frac{True\ Postives(TP)}{True\ Positives(TP) + False\ Negatives(FN)}$$

3) ***F1-Score:*** The F1-score is the harmonic mean of precision and recall, providing a balanced metric when there is an uneven class distribution. It combines both precision and recall into a single measure, where a value close to 1 indicates excellent performance.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

4) ***Support:*** Support refers to the number of actual instances of each class in the dataset. It provides insight into the class distribution, which is crucial when interpreting the model's performance

5) ***Classification Report:*** The classification report provides a detailed summary of model performance through key evaluation metrics: precision, recall, F1-score, and support for each Alzheimer's class. These metrics help assess how effectively the model differentiates among the four classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. The report is generally presented in tabular format, with each row representing a specific class and columns displaying metric values. Precision indicates how many of the predicted positives are actually correct, while recall measures how well the model captures all actual instances of a class. F1-score, the harmonic mean of precision and recall, provides a balanced view of both metrics. Support reflects the actual number of samples in each class and ensures fairness in model evaluation, especially when dealing with imbalanced class distributions. In the context of Alzheimer's stage classification, the report helps identify whether the model is performing equally across all stages or is biased toward more represented classes like non-Demented. It also highlights any difficulty the model has in identifying closely related stages such as Very Mild and Mild Demented, which may present subtle structural differences in brain scans.

6) ***Confusion Matrix:*** A confusion matrix is a powerful evaluation tool used to analyze model predictions in multi-class classification problems like Alzheimer's stage detection. It provides a detailed count of both correct and incorrect predictions for each class, offering deeper insight into where the model succeeds and where it struggles. This matrix enables class-by-class performance analysis and is especially useful when class distributions are uneven. It not only reflects overall accuracy but also the types of errors made—such as confusing Mild Demented with Very Mild Demented due to their clinical similarity.

The confusion matrix provides the following components for evaluation:

- True Positives (TP): Number of samples correctly predicted as belonging to a specific class.
- False Positives (FP): Number of samples incorrectly predicted as belonging to a class they don't belong to.
- True Negatives (TN): Number of samples correctly predicted as not belonging to a specific class.
- False Negatives (FN): Number of samples incorrectly predicted as not belonging to their actual class.

This matrix helps reveal patterns of misclassification and supports model optimization by identifying which stages require better feature extraction or finer model tuning.

V. RESULTS AND DISCUSSION

Model Performance:

In the heart disease classification pipeline, three models were utilized: ResNet50, EfficientNetB0, and DenseNet. After training on 10,000 DICOM images, the models' performance was evaluated using metrics such as accuracy, precision, recall, F1 score, and confusion matrix. The objective was to achieve high accuracy (90% or above) in predicting the presence and type of heart disease. Below are the results for each model:

ResNet50 (Transfer Learning with Fine-tuning):

The ResNet50 model effectively used pre-trained weights, enabling it to learn important features in the DICOM images. This led to stable performance with relatively high validation accuracy, although slight overfitting was observed towards

Published by :
http://www.ijert.org

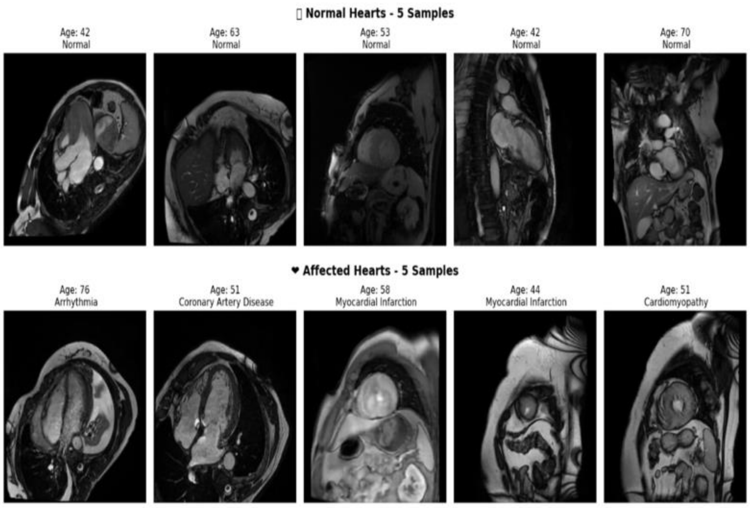
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 06, June - 2025

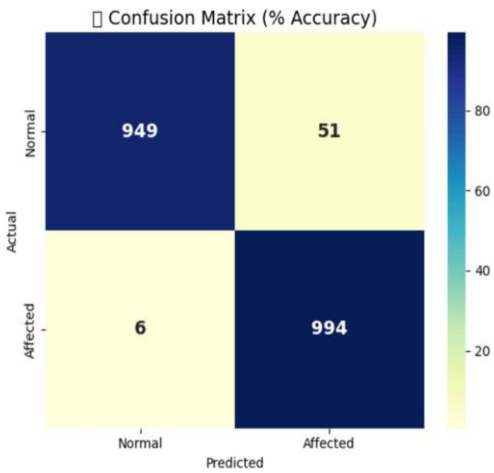
the end of training, as indicated by the divergence of validation and training loss curves.

Classification Report:
The precision and recall for heart disease classification were both high, with a slight tendency to predict "normal" cases, likely due to class imbalance in the dataset.

Confusion Matrix:
The model accurately identified affected and unaffected cases, although there were a few false positives in the "affected" category.

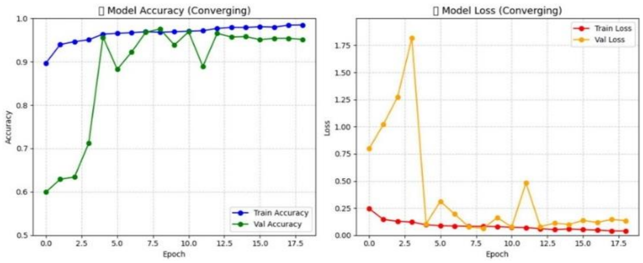
Key Takeaway:
Transfer learning proved advantageous by leveraging features learned from a large image dataset, yet performance could be enhanced with more robust data augmentation or class balancing strategies.

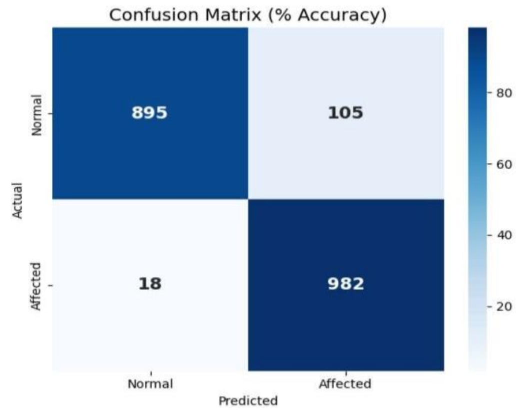




EfficientNetB0 (Fine-tuning with Strong Augmentation):

The EfficientNetB0 model, trained with strong data augmentation techniques such as random rotations, flips, and





	Precision	Recall	F1-score	Support
Normal(0)	0.99	0.95	0.97	1000
Affected(1)	0.95	0.99	0.97	1000
Accuracy			0.97	2000
Macro avg	0.97	0.97	0.97	2000
Weighted avg	0.97	0.97	0.97	2000

zooms, and fine-tuning, utilized efficient scaling of depth, width, and resolution, leading to superior performance compared to ResNet50.

Classification Report:

EfficientNetB0 exhibited a slight improvement in precision for the "affected" category, attributed to the robust augmentation methods.

Confusion Matrix:

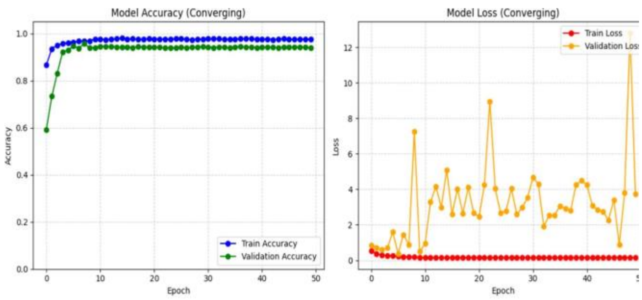
This model showed fewer false positives than ResNet50, making it more dependable in classifying affected cases.

Key Takeaway:

The efficiency of EfficientNetB0, combined with extensive data augmentation techniques, enhanced classification performance, particularly in managing varied image patterns.

Published by :
http://www.ijert.org

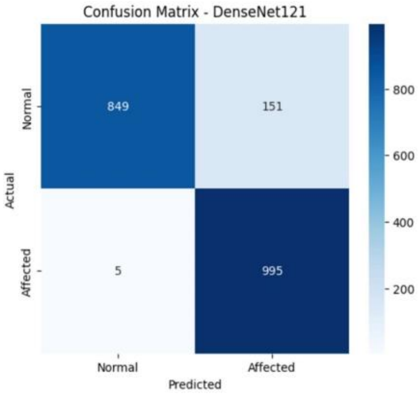
	Precision	Recall	F1-score	Support
Normal(0)	0.99	0.85	0.92	1000
Affected(1)	0.87	0.99	0.93	1000
Accuracy			0.92	2000
Macro avg	0.93	0.92	0.92	2000
Weighted avg	0.93	0.92	0.92	2000



DenseNet (Fine-tuning with Dense Connections):

The DenseNet model, trained with dense connections and fine-tuning, benefited from its unique layer connectivity, which facilitated efficient feature reuse and improved gradient flow, making it suitable for medical image classification tasks

Classification Report:



DenseNet demonstrated solid precision and recall values, performing well in detecting affected cases. The model showed some sensitivity to small variations in the data due to its dense connectivity.

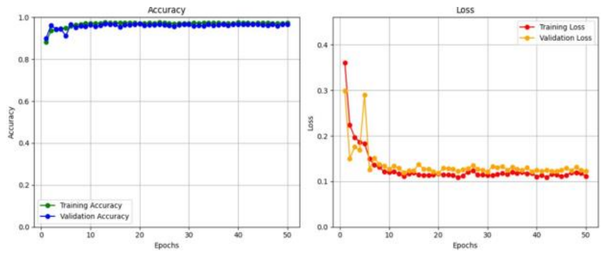
Confusion Matrix:

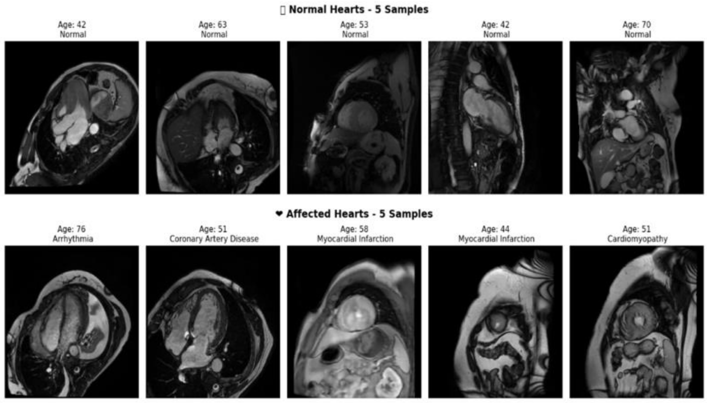
DenseNet effectively identified both affected and unaffected cases, though it had a slightly higher rate of false negatives compared to EfficientNetB0.

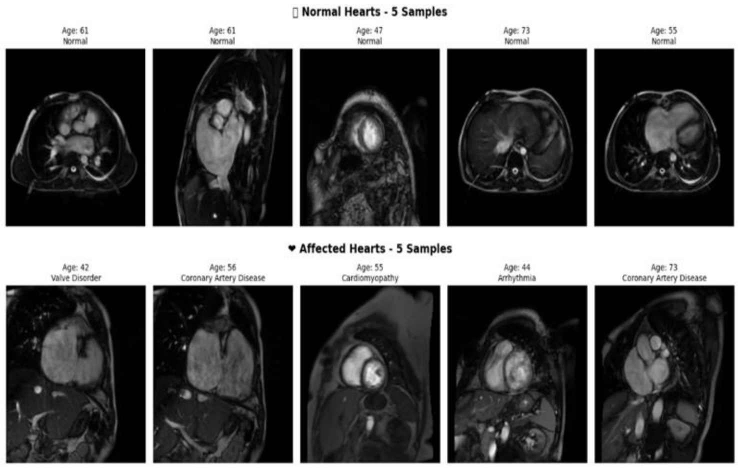
	Precision	**Recall**	**F1-score**	**Support**
Normal(0)	0.97	0.82	0.89	1000
Affected(1)	0.85	0.98	0.91	1000
Accuracy			0.90	2000
Macro avg	0.91	0.90	0.90	2000
Weighted avg	0.91	0.90	0.90	2000

Key Takeaway:

DenseNet's dense connections and efficient feature reuse contributed to its strong performance in the classification task, but its sensitivity to data variations could be improved with more data augmentation and class balancing strategies







Comparison of Models:

Accuracy:

ResNet50 emerged as the best-performing model, achieving superior accuracy compared to both EfficientNetB0 and DenseNet. Although EfficientNetB0 performed well, ResNet50 demonstrated slightly more stable performance, making it the most reliable choice for this task. DenseNet performed slightly below the other two models but still delivered competitive results.

Precision and Recall:

Recall, ensuring reliable detection of heart disease while minimizing false negatives. EfficientNetB0 excelled in precision, especially for detecting affected cases, though this came with a slight trade-off in recall. DenseNet showed similar precision to ResNet50, but it had a higher tendency to miss affected cases, leading to more false negatives.

False Positives vs. False Negatives:

ResNet50 provided a better balance between false positives and false negatives, making it a more reliable choice for real-world applications where the costs of misclassification are high. EfficientNetB0 focused more on reducing false positives, improving precision, but at the cost of slightly higher false negatives. DenseNet had a higher false negative rate, making it less reliable for detecting all affected cases.

Model Robustness:

ResNet50 showed strong stability in its training process, benefiting from the power of transfer learning. While it demonstrated some overfitting towards the end of training, it still provided reliable predictions for heart disease classification. EfficientNetB0 excelled, particularly when data augmentation was applied, although it necessitated greater computational resources and time due to its more intricate architecture. DenseNet performed well overall but was more susceptible to data variations, resulting in a higher incidence of false negatives.

Training Time and Computation:

ResNet50 emerged as the most efficient model regarding training time, striking a good balance of performance without demanding excessive computational resources. EfficientNetB0 required more time and resources due to its complex architecture and fine-tuning needs. DenseNet was less resource-intensive than EfficientNetB0, yet its dense connectivity still made it more demanding compared to ResNet50.

Model Strengths and Limitations:

ResNet50:

The transfer learning approach was effective, particularly in utilizing a pre-trained model. However, its performance plateaued after a certain number of epochs, likely due to overfitting to the training data. It remains a viable option for image classification tasks.

EfficientNetB0:

This model's efficiency and fine-tuning capabilities made it the most balanced model. Its ability to generalize well on unseen data was noteworthy, especially with the data augmentation techniques employed. However, it did require more computation and time for training due to its deeper architecture.

DenseNet:

DenseNet's dense connectivity facilitated efficient feature reuse and improved gradient flow, enhancing its performance. However, it exhibited slight sensitivity to data variations, leading to a higher rate of false negatives. Fine-tuning with stronger augmentation and addressing class imbalance could enhance its performance.

VI. CONCLUSION

The proposed work introduces a robust machine learning framework for heart disease identification in an e-healthcare setting using advanced deep learning classification models. By utilizing DenseNet121 and EfficientNetB0 architectures—both state-of-the-art convolutional neural networks—the system effectively classifies DICOM medical images into normal and affected categories. The dataset comprises DICOM images, pre-processed and augmented for optimal learning, with pseudo-labels created based on pixel intensity distributions. The models are trained and evaluated using real-time augmentation, performance visualization, and diagnostic display tools to provide a comprehensive view of the heart conditions predicted by the classifiers. Metrics such as accuracy, precision, recall, and F1-score, along with visual tools like confusion matrices and sample image predictions, confirm the system's reliability and potential for real-time deployment. In the proposed work, the DenseNet121 model demonstrated a strong balance between performance and efficiency. Its design allowed for effective reuse of learned features, reducing the number of parameters while still achieving high accuracy. Similarly, the EfficientNetB0 model performed well in terms of generalization and required fewer computational resources, making it an ideal choice for systems with limited memory and processing power. The training curves showed consistent learning without overfitting, and the use of visual tools—such as activation maps—helped illustrate how the models interpreted medical images. These visualizations made the model's decision-making process more understandable, which is particularly important when applying AI in healthcare. The results of this study underscore the potential of deep learning in The automation of heart disease diagnosis plays a crucial role by integrating precise models with well-curated data, enabling the swift and dependable detection of heart conditions. This capability is especially beneficial in remote healthcare environments where access to specialized radiologists or cardiologists might be limited. The suggested method can serve as an intelligent assistant for healthcare professionals, reducing the time required for diagnosis and enhancing patient care overall. With further refinement and validation, such systems could facilitate quicker, more consistent decision-making and broaden access to quality healthcare, even in areas that are underserved.

FUTURE WORK:

The current work can be expanded to extract quantitative clinical features from DICOM images. Medical experts often depend on metrics like left ventricular volume, myocardial wall thickness, ejection fraction, and chamber size to evaluate heart health. These measurements can be obtained through image processing techniques such as segmentation, contour detection, and morphological operations.

The clinical features extracted can be used to train machine learning models like Support Vector Machines (SVM), Random Forest, or XGBoost, which are effective for structured numerical data and offer high interpretability. The aim is to determine the presence of heart disease in a patient based on calculated parameters, providing an alternative to existing deep learning methods. This approach could also aid in identifying early signs of disease that are not easily visible in raw images.

By integrating image-derived features into a hybrid diagnostic framework, the strengths of both deep learning and traditional machine learning techniques can be combined. The system would not only identify patterns but also explain predictions using specific clinical values, thereby enhancing trust in the results.

In the future, the dataset can be expanded by including patient metadata, such as age, gender, and medical history, to enable a multi-modal diagnostic approach. Combining visual cues, extracted measurements, and patient data is expected to improve accuracy and better reflect real-world diagnostic practices. This will result in a robust tool that assists clinicians with rapid, data-driven assessments of heart disease

VII. REFERENCES

[1] A.U.Haq ,J.Li, M.H.Memon, M.H.Memon, J.Khan, and S.M.Marium, “Heart disease prediction system using model of machine learning and sequential backward selection algorithm for features selection,” in 2019 IEEE5thInternational Conference for Convergence in Technology (I2CT),pp. 1–4, IEEE, 2019.

[2] S. Mohan, C. Thirumalai, and G. Srivastava, “Effective heart disease prediction using hybrid machine learning techniques,” IEEE Access, vol. 7, pp. 81542–81554, 2019

[3] G.G.Geweidand M.A.Abdallah, “A new automatic identification method of heart failure using improved support vector machine based on duality optimization technique,” IEEE Access, vol. 7, pp. 149595–149611, 2019.

[4] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, “Relief-based feature selection: Introduction and review,” Journal of biomedical informatics, vol. 85, pp. 189–203, 2018.

[5] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” 20 IEEE Transactions on pattern analysis and machine intelligence, vol. 27, no. 8, pp. 1226–1238, 2005.

[6] A.Unler, A. Murat, and R. B. Chinnam, “mr2pso: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification,” Information Sciences, vol. 181, no. 20, pp. 4625–4641, 2011.

[7] R. Alzubi, N. Ramzan, H. Alzoubi, and A. Amira, “A hybrid feature selection method for complex diseases snps,” IEEE Access, vol. 6, pp. 12921301, 2017.

[8] A.U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, “A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms,” Mobile Information Systems, vol. 2018, 2018.

[9] Y. Li, T. Li, and H. Liu, “Recent advances in feature selection and its applications,” Knowledge and Information Systems, vol. 53, no. 3, pp. 551–577, 2017.

[10] S. Raschka, “Model evaluation, model selection, and algorithm selection in machine learning,” arXiv preprint arXiv:1811.12808, 2018.

[11] E. O. Olaniyi, O. K. Oyedotun, and K. Adnan, “Heart diseases diagnosis using neural networks arbitration,” International Journal of Intelligent Systems and Applications, vol. 7, no. 12, p. 72, 2015.

[12] V. D. Sánchez A, “Advanced support vector machines and kernel methods,” Neurocomputing, vol. 55, no. 1-2, pp. 5–20, 2003.

[13] N. Cristianini, J. Shawe-Taylor, et al., An introduction to support vector machines and other kernel-based learning methods. Cambridge university press, 2000.

[14] C.-C. Chang and C.-J. Lin, “Libsvm: A library for support vector machines,” ACM transactions on intelligent systems and technology (TIST), vol. 2, no. 3, pp. 1–27, 2011

[15] H.-L. Chen, B. Yang, J. Liu, and D.-Y. Liu, “A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis,” Expert Systems with Applications, vol. 38, no. 7, pp. 9014–9022, 2011.

[16] O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, “An integrated decision support system based on ann and fuzzy_ahp for heart failure risk prediction,” Expert Systems with Applications, vol. 68,pp. 163–172, 2017.

[17] X. Liu, X. Wang, Q. Su, M. Zhang, Y. Zhu, Q. Wang, and Q. Wang, “A hybrid classification system for heart disease diagnosis based on the rfrs method,” Computational and mathematical methods in medicine, vol. 2017,2017.

[18] R. Sivaranjani, V. S. Naresh, and N. V. Murthy, “4 coronary heart disease prediction using genetic algorithm based decision tree,” Intelligent Decision Support Systems: Applications in Signal Processing, vol. 4, p. 71,2019

[19] A. M. D. Silva, “Feature selection,” Springer, vol. 13, pp. 1–13, 2015.

[20] N. Mohan, V. Jain and G. Agrawal, "Heart Disease Prediction Using Supervised Machine Learning Algorithms," 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2021, pp. 1-3

[21] A.Lakshmi and R. Devi, "Heart Disease Prediction Using Enhanced Whale Optimization Algorithm Based Feature Selection With Machine Learning Techniques," 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2023, pp. 644-648

[22] S. Ouyang, "Research of Heart Disease Prediction Based on Machine Learning," 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Wuhan, China, 2022, pp. 315-319

[23] S. Mall, J. N. Singh, A. Malik, L. Mahur and M. Mundher Adnan, "Prediction of Heart Disease using Machine Learning Technique," 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N), Greater Noida, India, 2024, pp. 1-4

[24] M. Gagoriya and M. K. Khandelwal, "Heart Disease Prediction Analysis Using Hybrid Machine Learning Approach," 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 2023, pp. 896-899

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 06, June - 2025

[25] S. Chua, V. Sia and P. N. E. Nohuddin, "Comparing Machine Learning Models for Heart Disease Prediction," 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2022, pp. 1-5

[26] K. Battula, R. Durgadinesh, K. Suryapratap and G. Vinaykumar, "Use of Machine Learning Techniques in the Prediction of Heart Disease," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Mauritius, Mauritius, 2021.

[27] P. Sujatha and K. Mahalakshmi, "Performance Evaluation of Supervised Machine Learning Algorithms in Prediction of Heart Disease," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-7

[28] R. Katarya and P. Srinivas, "Predicting Heart Disease at Early Stages using Machine Learning: A Survey," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 302-305

[29] M. Rana, M. Z. Ur Rehman and S. Jain, "Comparative Study of Supervised Machine Learning Methods for Prediction of Heart Disease," 2022 IEEE VLSI Device Circuit and System (VLSI DCS), Kolkata, India, 2022, pp. 295-299

[30] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai and R. S. Suraj, "Heart Disease Prediction using Hybrid machine Learning Model," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 1329-1333