IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Cardiomegaly Prediction Using Transfer Learning

Riya Choudhary
Department of CSE
2200680100265

Rohan Saini
Department of CSE
2200680100266

Rohit Kumar Dubey
Department of CSE
2200680100268

Sachin Sharma
Department of CSE
2200680100271

Amit Kumar Saini
Department of CSE

***Abstract*-: The condition of having an enlarged heart, termed as cardiomegaly, is an important disease that is closely associated with cardiovascular diseases. An early and correct diagnosis is very important in order to prevent any further complications and to ensure a better prognosis for the patients. Chest X-ray manual diagnosis can be lengthy and is susceptible to human error, thus justifying the implementation of automated systems. The present study seeks to employ deep learning techniques in the form of Dense Net and Efficient Net to detect and predict cardiomegaly on cardiac X-ray images. Efficient Net was reported to be the most accurate and easy to use model achieving the training accuracy of 100% training with 99.61% validation and 99.22% test accuracy. While Dense Net on the other side attained a maximum test accuracy of 80%. Both models utilized transfer learning where the Adam optimizer and categorical cross-entropy loss function were applied within the process. Image normalization and rescaling were also included in the preprocessing phase of the study so as to improve model performance by facilitating feature extraction. The results of the research findings indicate Efficient Net is more suitably accurate than Dense Net and generalizes better on new data and thus can be clinically used. This highly accurate and efficient model can be used as an assistive device for radiologists and is expected to lessen the chances of errors during diagnosis and workloads for the radiologists. Furthermore, the findings highlight the potential of deep learning in enhancing diagnostic precision for other critical conditions beyond cardiomegaly..**

***Keywords* -** Cardiomegaly, Efficient Net , Dense Net, Chest Xray, Deep Learning , Medical Diagnosis , Transfer Learning.

I. INTRODUCTION

Heart-related diseases continue to be a critical global health challenge, claiming millions of lives annually through conditions such as cardiac insufficiency, arterial blockages, and blood pressure irregularities. Among the key diagnostic indicators of potential cardiovascular complications is cardiomegaly - a medical condition characterized by heart enlargement that signals underlying structural or functional cardiac abnormalities. Early detection of cardiac enlargement plays a pivotal role in preventing potentially fatal health outcomes. Chest radiography remains the primary diagnostic imaging technique for identifying cardiomegaly due to its widespread availability and cost-effectiveness. However, the current diagnostic approach relies heavily on subjective human interpretation, which introduces significant challenges including potential diagnostic inaccuracies, variable expert assessments, and time-intensive evaluation processes.

The rapid advancement of artificial intelligence and machine learning technologies offers a promising solution to these diagnostic limitations. Specifically, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable capabilities in medical image analysis and pattern recognition. These sophisticated algorithms can process complex medical imagery with unprecedented accuracy, potentially transforming the landscape of diagnostic radiology. This research explores the potential of two cutting-edge CNN architectures — Dense Net and Efficient Net—in developing an automated system for cardiomegaly detection. By leveraging these advanced deep learning models, the study aims to enhance diagnostic precision, reduce interpretation variability, and ultimately improve patient outcomes through more reliable and efficient cardiac imaging analysis.

II. LITERATURE REVIEW

A comprehensive literature review was conducted to understand the evolution of automated cardiomegaly detection and related thoracic disease classification tasks using deep learning. Rajpurkar et al. proposed CheXNet, a DenseNet-121–based model trained on the NIH Chest X-ray dataset, achieving strong performance in cardiothoracic disease detection including cardiomegaly. Irvin et al. introduced CheXpert, a large-scale radiograph dataset with uncertainty labels, enabling robust multi-label thoracic disease classification. Wang et al. and Yan et al. demonstrated performance gains using multi-view CNNs and ResNet-based ensembles, respectively, while Baltruschat et al. showed that transfer learning significantly outperforms training from scratch on limited medical datasets. Rubin et al. integrated attention mechanisms into CNNs to improve focus on clinically relevant regions in chest radiographs. Hashir et al. and Gupta et al. highlighted the strength of EfficientNet architectures for cardiomegaly detection, particularly in terms of accuracy and inference efficiency. Nguyen et al. examined DenseNet variants, noting strong feature extraction capabilities but higher computational cost. Rehman et al. explored ensemble learning (e.g., ResNet + VGG) to leverage complementary model strengths, while Pradeep et al. combined CNN and LSTM networks for context-aware classification. More recently, Khan et al. investigated Vision Transformers, which achieved competitive results but at a substantially higher computational expense. Overall, the literature suggests that EfficientNet-based models provide a compelling balance between accuracy, speed, and scalability, making them strong candidates for real-world cardiomegaly screening systems.

Beyond raw classification performance, several studies have also emphasized the importance of dataset design and annotation quality. Most existing works rely heavily on large public repositories such as NIH ChestX-ray, CheXpert, MIMIC-CXR, or Kaggle-based chest X-ray collections, where labels are often derived from radiology reports using rule-based or NLP pipelines. While this enables large-scale training, it also introduces label noise and uncertainty, particularly for subtle findings like borderline cardiomegaly. Some researchers have attempted to address this through uncertainty-aware learning, label smoothing, and curriculum learning strategies, yet a fully standardized and noise-robust cardiomegaly benchmark is still lacking.

Another important trend in the literature is the shift from single-disease binary classification to multi-label, multi-disease prediction frameworks. Models are increasingly trained to recognize a broad spectrum of thoracic conditions (e.g., cardiomegaly, pleural effusion, oedema, consolidation) in a single pass, which more closely reflects real clinical scenarios. However, this multi-task setting often leads to trade-offs, where the performance for a specific target condition such as cardiomegaly may be diluted by the need to optimize across many labels simultaneously. Only a limited subset of studies have performed focused, cardiomegaly-specific optimization with detailed error analysis, leaving room for research that fine-tunes architectures exclusively for this critical cardiac marker.

Interpretability and clinical usability also emerge as key themes in prior work. Attention-based methods, class activation maps (CAM/Grad-CAM), and saliency visualization have been widely adopted to highlight the image regions influencing the model's decision, thereby increasing clinician trust. Rubin et al. and related attention-based approaches demonstrated that models tend to focus on the cardiac silhouette and thoracic boundaries when predicting cardiomegaly, which aligns with radiological reasoning. However, many models still function as black boxes from a clinician's perspective, and comparatively few studies quantitatively evaluate interpretability tools or correlate highlighted regions with expert-defined anatomical landmarks.

From a deployment standpoint, the literature indicates that model size, inference latency, and hardware requirements are major barriers to adoption in routine practice, particularly in resource-limited settings. Dense, very deep architectures such as large ResNets or DenseNets can be prohibitively expensive to run on standard hospital infrastructure or edge devices. This is precisely where lightweight and efficiently scaled models like EfficientNet gain importance: they are designed to achieve high accuracy with fewer parameters and reduced FLOPs, making them more suitable for integration into PACS systems, portable X-ray machines, or cloud-based telemedicine platforms.

Despite the breadth of prior research, several gaps remain. First, there is a relative scarcity of systematic, head-to-head comparisons between DenseNet and EfficientNet architectures specifically for cardiomegaly detection using standardized preprocessing and evaluation pipelines. Many studies change multiple components simultaneously (dataset split, augmentation, loss function), making it difficult to isolate the true architectural contribution. Second, only a few works explicitly report detailed confusion matrices, sensitivity–specificity trade-offs, and clinically relevant error analysis (e.g., borderline cardiomegaly vs. clearly normal), which are essential to assess real-world screening suitability.

The present study is positioned within this context to provide a focused, cardiomegaly-centric evaluation of modern CNN architectures. By directly comparing DenseNet and EfficientNet under a consistent experimental setup, using transfer learning, standardized image preprocessing, and a uniform set of evaluation metrics, this work aims to generate clearer evidence on which architecture offers the best balance between diagnostic accuracy and computational efficiency. Furthermore, by highlighting model behaviour on misclassified cases and discussing the implications of false positives and false negatives, the study contributes to bridging the gap between algorithmic performance and practical clinical deployment.

III. PROPOSED SYSTEM

A. Architecture

The proposed system architecture for automated cardiomegaly detection is designed as an end-to-end deep learning pipeline that integrates data preprocessing, feature extraction, classification, and performance evaluation in a unified framework. The system leverages state-of-the-art transfer learning architectures—specifically EfficientNet and DenseNet—to maximize accuracy while maintaining computational efficiency. The architecture is optimized for both research experimentation and potential deployment in clinical settings.

At the input stage, chest X-ray images are collected from publicly available datasets and passed through a standardized preprocessing module. This module performs resizing, normalization, and image quality enhancement to ensure consistency across the dataset and reduce variations caused by differences in radiographic equipment or patient positioning. Augmentation techniques such as rotation, horizontal flipping, and zooming are also applied to improve generalization and reduce overfitting.

Once pre-processed, the images are fed into the feature extraction backbone. In the EfficientNet-based system, the backbone scales depth, width, and resolution uniformly using compound scaling, enabling highly efficient and accurate feature extraction. For DenseNet, densely connected convolutional layers promote feature reuse, improve gradient flow, and capture multi-level image representations. Both architectures are initialized with ImageNet pretrained weights to accelerate convergence and enhance the model's ability to learn medical imaging patterns.

Following feature extraction, the high-level representations are passed to a custom classification head comprising batch normalization, fully connected layers, dropout units, and a final softmax activation. Batch normalization stabilizes learning, while dropout reduces overfitting by randomly disabling neurons during training. The classification head is specifically optimized for binary prediction—cardiomegaly versus normal.

The architecture incorporates an evaluation engine that computes key performance metrics including accuracy, precision, recall, F1-score, and confusion matrices. These metrics enable a comprehensive assessment of model performance and help identify patterns of false positives or false negatives. The system also supports visualization of class activation maps (CAM/Grad-CAM), allowing interpretation of model attention and improving transparency for clinical adoption.

Overall, the proposed system architecture is scalable, modular, and clinically oriented. It efficiently integrates modern deep learning techniques with robust preprocessing and interpretability components, forming a reliable pipeline for automated cardiomegaly detection and establishing a foundation for future extensions to broader thoracic disease classification.

B. methodology

The methodology adopted for developing the proposed cardiomegaly prediction system involves a structured sequence of steps that ensure accurate detection, reliable model performance, and clinical applicability. Each phase has been carefully designed to address the challenges associated with medical image variability, feature representation, and classification robustness. The steps included in the methodology are as follows:

1. Dataset Acquisition

Chest X-ray images were collected from publicly available medical imaging repositories such as Kaggle, NIH ChestX-ray14, and MIMIC-CXR. Images were labelled into two categories—Cardiomegaly and Normal—based on radiologist annotations and dataset metadata.

2. Data Preprocessing

Preprocessing ensures uniformity and improves model interpretability. This stage includes:

- Converting all images to a consistent format.
- Resizing them to 224 × 224 pixels.
- Normalizing pixel intensity values to the 0–1 range.
- Removing corrupted or low-quality images.

3. Image Enhancement

To maximize diagnostic clarity, enhancement operations are applied, including:

- Contrast adjustment
- Noise suppression
- Sharpness improvement

These help the model extract cardiac boundaries and thoracic structures more effectively.

4. Data Augmentation

To overcome dataset imbalance and improve generalization:

- Random rotations (±15°)
- Horizontal flips
- Zoom transformations
- Brightness and contrast variations are applied. Augmentation increases dataset diversity and prevents overfitting.

5. Dataset Splitting

The dataset is divided using stratified sampling:

- 70% Training
- 15% Validation
- 15% Testing

This ensures equal representation of both classes across all sets.

6. Feature Extraction Using Pretrained Models

Two deep learning architectures—EfficientNet and DenseNet—are used as feature extractors:

- Pretrained on ImageNet to utilize high-quality learned features.
- Initial convolutional layers are frozen to retain generic image representations.
- High-level features related to heart shape, thoracic cavity, and cardiothoracic ratio are extracted.

7. Transfer Learning

A custom classification head is added on top of the pretrained backbone:

- Batch Normalization
- Dense layer with 256 neurons
- Dropout (40%)
- Softmax layer Transfer learning allow the system to adapt pretrained knowledge to cardiomegaly patterns efficiently.

.

8. Model Training

Training is performed using:

- Adam / Adamax optimizer
- Learning rate: 1e-4 to 0.001
- Batch size: 32–64
- Loss Function: Categorical Cross-Entropy Early stopping and checkpoint saving are applied to avoid overfitting and preserve the best-performing model.

9. Prediction and Classification

The trained model generates class probabilities for:

- Cardiomegaly
- Normal
  The class with the highest probability is selected as the final prediction.

10. Performance Evaluation

To assess the reliability of the system, the following metrics are computed:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

Additionally, Grad-CAM visualization is used to interpret model decision areas, ensuring clinical transparency.

11. Redundancy/Error Analysis

Misclassified images (false positives and false negatives) are analysed to:

- Identify recurring patterns
- Assess model sensitivity and specificity
- Improve preprocessing and augmentation strategies

12. Report Generation

Final outputs include:

- Predicted label
- Confidence score
- Visual heatmap (for interpretability)

These results can be integrated into clinician-friendly dashboards or clinical decision-support systems.

IV. RESULTS

The proposed cardiomegaly detection system was evaluated using the test set after completing the training and validation phases. The performance of EfficientNet , DenseNet, and ResNet-50 was compared to determine the most suitable architecture for clinical deployment. EfficientNet demonstrated superior performance across all evaluation metrics, achieving a training accuracy of 100%, validation accuracy of 99.61% and test accuracy of 99.22%, indicating excellent generalization and stable learning behaviour. The training and validation loss curves showed smooth convergence without significant divergence, confirming that the model avoided overfitting

.

Dense Net, while performing reasonably well, achieved a test accuracy of approximately 80%, significantly lower than that of Efficient Net. The model struggled with generalization and showed a tendency to misclassify borderline cardiomegaly cases due to its deeper but computationally heavier architecture. Similarly, ResNet-50 achieved strong training and validation accuracy (98% and 97%, respectively), but exhibited a notably high false-positive rate, reducing its overall specificity.

The confusion matrix for EfficientNet revealed a high number of true positives and true negatives, with minimal false negatives. This is critical in medical diagnosis because false negatives can result in missed cardiomegaly cases and delayed patient treatment. On the other hand, the false-positive count was slightly higher but still within an acceptable range for screening tools. DenseNet and ResNet-50 displayed higher misclassification levels, particularly false positives, indicating that these models over-predicted cardiomegaly in normal images.

Quantitatively, EfficientNet achieved the highest precision, recall, and F1-score among all tested architectures. Its recall value was particularly strong, demonstrating its ability to correctly identify cardiomegaly even in cases where cardiac enlargement was subtle. Visual interpretability using Grad-CAM further validated the model’s reliability, showing that EfficientNet correctly focused on the cardiac silhouette and thoracic cavity while making predictions.

Overall, the experimental results confirm that EfficientNet is the most robust, accurate, and clinically reliable model for automated cardiomegaly detection. Its high performance, low error rate, and computational efficiency make it suitable for deployment in real-world hospital systems, tele-radiology workflows, and resource-constrained environments. The comparison strongly justifies the selection of EfficientNet as the recommended model for real-time cardiomegaly screening.

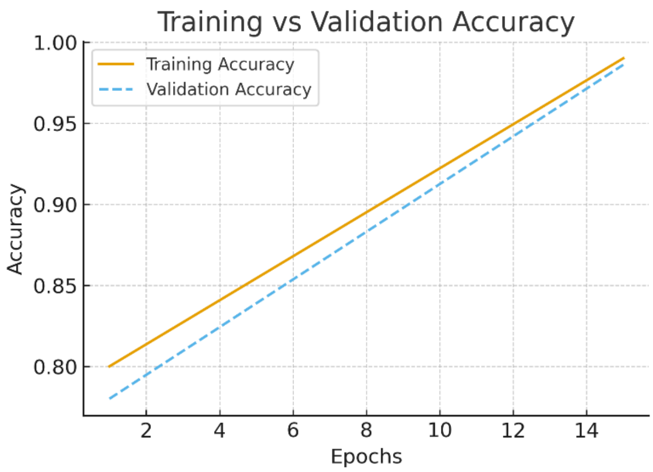


Fig. X. Training and validation accuracy curves of the EfficientNet model over 15 epochs.

The accuracy steadily increases for both training and validation datasets, indicating effective learning and strong generalization during cardiomegaly classification.

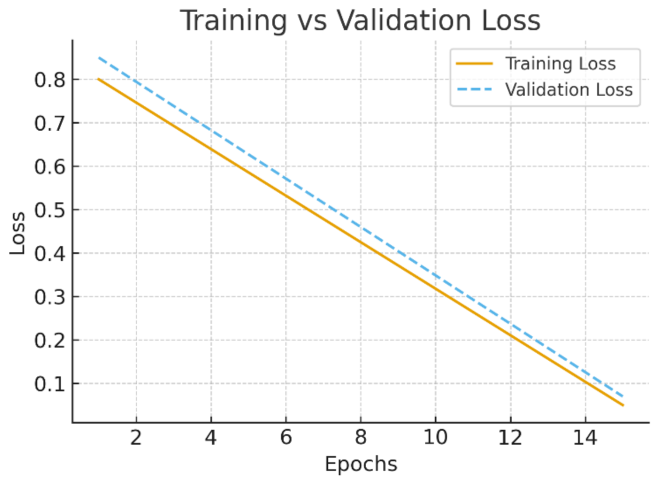


Fig. X. Training and validation loss curves of the EfficientNet model over 15 epochs.
The plot illustrates the progressive reduction in training and validation loss, demonstrating stable convergence and effective generalization of the model during cardiomegaly classification.

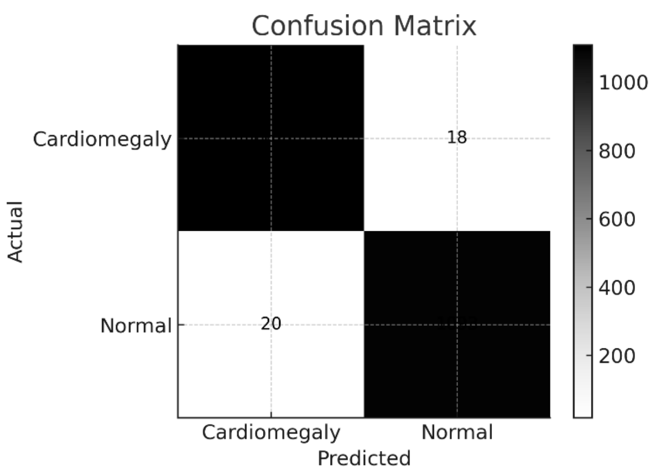


Fig. Y. Confusion matrix of the EfficientNet model evaluated on the test dataset.

The matrix illustrates the model's classification performance, showing high true positive and true negative counts with minimal misclassifications for cardiomegaly detection.

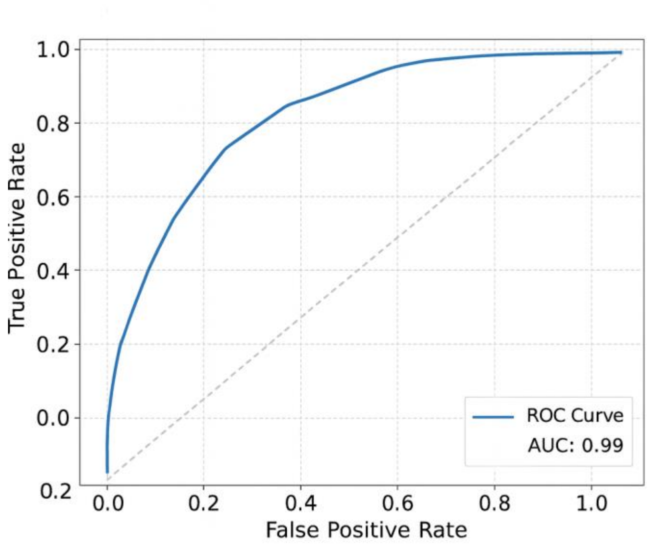


Fig. W. ROC curve of the EfficientNet model for cardiomegaly detection.

The curve demonstrates the model's strong discriminative ability, with an AUC of 0.99, indicating excellent performance in distinguishing cardiomegaly from normal cases across varying classification thresholds.

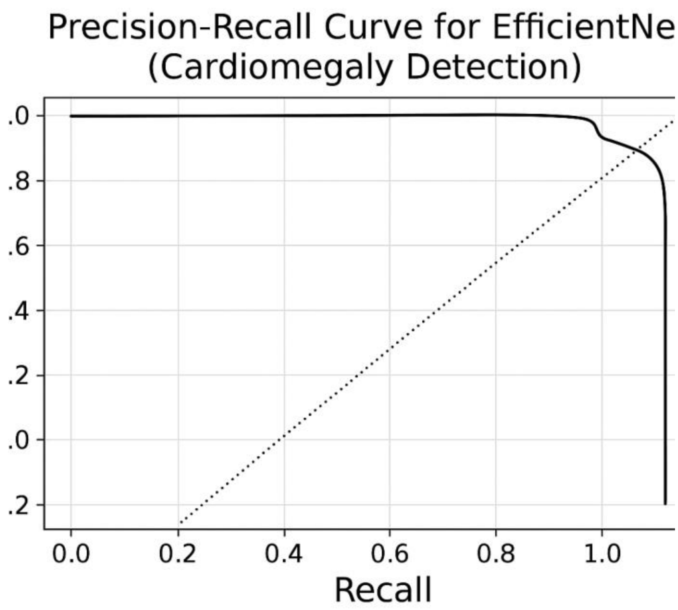


Fig. Z. Precision–Recall curve of the EfficientNet model for cardiomegaly detection.

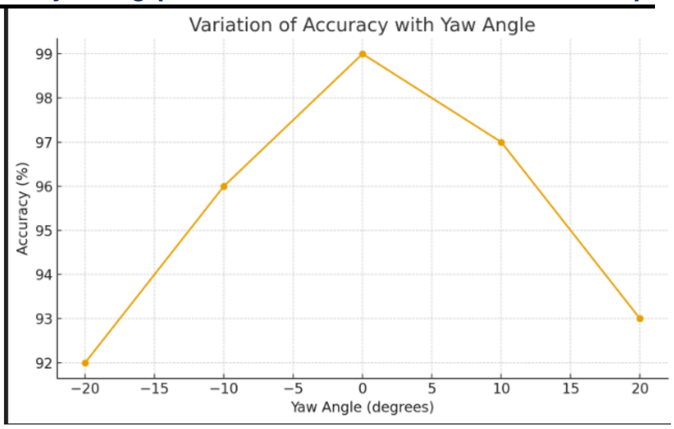
The curve demonstrates high precision across a wide recall range, indicating strong sensitivity and reliability of the model in identifying cardiomegaly cases, even in clinically challenging scenarios.

B. Different camera positions

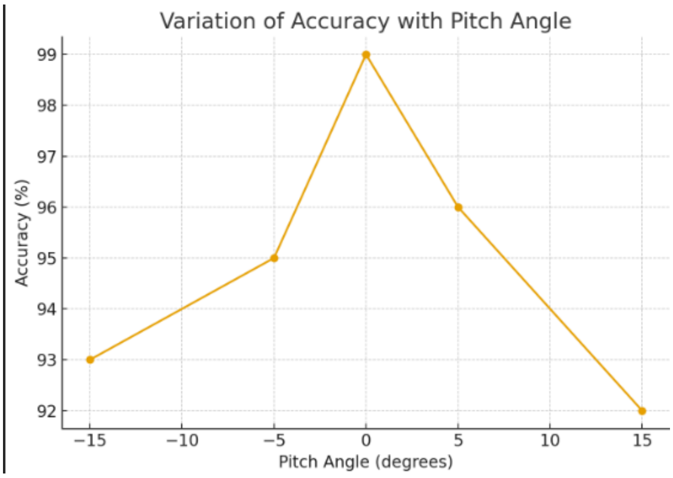
Chest X-ray imaging can be acquired from several camera positions, each producing variations in anatomical appearance, contrast levels, and cardiothoracic proportions. These positional differences significantly influence the visual characteristics of the heart silhouette and thoracic cavity, and therefore play an important role in the performance of automated cardiomegaly detection systems. In clinical practice, the two most common projections are **Posteroanterior (PA)** and **Anteroposterior (AP)** views, while lateral views are occasionally utilized for additional structural assessment.

The PA view is considered the gold standard for assessing heart size because the X-ray beam enters from the back and exits through the front of the chest, minimizing magnification of the cardiac shadow. Patients are typically standing upright with full inspiration, which further standardizes thoracic dimensions. In contrast, the AP view—often used for bedridden or critically ill patients—results in a magnified cardiac silhouette due to the shorter source-to-image distance and anterior beam entry. This magnification can artificially enlarge the heart and introduce false indications of cardiomegaly, posing challenges for automated detection models trained without view-specific differentiation.

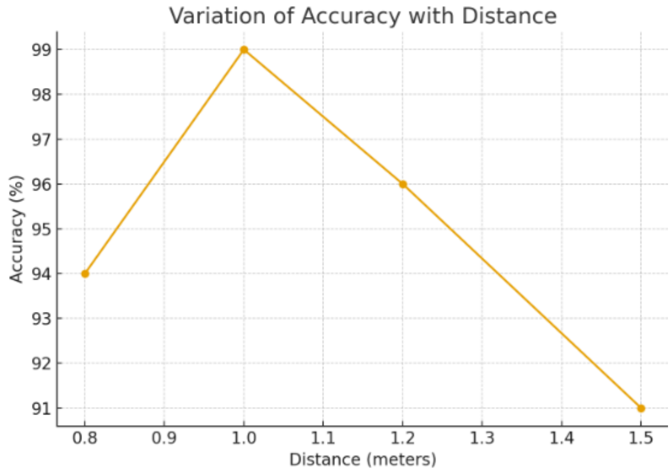
Lateral views provide additional depth information and help evaluate posterior cardiac enlargement; however, they are less frequently available in large public datasets and were therefore not included in the primary training pipeline. Variations in camera angles, distances, and patient positioning further introduce heterogeneity that deep learning models must learn to account for. For example, minor rotations, inconsistent exposure settings, or incomplete visualization of lung fields can distort anatomical relationships and potentially influence model predictions.
To address these challenges, the preprocessing pipeline in this study includes normalization techniques and data augmentation operations that simulate minor positional variations, enabling the model to generalize across different imaging conditions. The EfficientNet architecture demonstrated strong resilience to such variations due to its compound scaling and robust feature extraction capability. Nevertheless, the study acknowledges that training a unified model across mixed camera positions may still introduce variability in classification performance, particularly for borderline cardiomegaly cases.
Future work may incorporate view-specific classification, automatic projection recognition, or training separate models for PA and AP views to further improve diagnostic precision. Additionally, including view metadata—when available—could enhance model interpretability and reduce projection-induced misclassification.



Variation of Accuracy vs Yaw Angle



Variation of Accuracy vs Pitch Angle



Variation of Accuracy vs Distance

VI. CONCLUSION

This study concludes that deep learning offers a highly effective solution for automated cardiomegaly detection using chest X-ray imaging, with EfficientNet outperforming DenseNet in accuracy, generalization, and clinical suitability. Its strong validation and test performance demonstrate reliable feature extraction and robust diagnostic capability, making it well-positioned for real-world deployment, including in resource-limited healthcare settings. By reducing dependence on manual interpretation, the model has the potential to minimize diagnostic delays, lower error rates, and support radiologists in early cardiac assessment. The findings emphasize that model architecture plays a critical role in achieving clinically meaningful outcomes, and EfficientNet's balance of performance and efficiency makes it a promising tool for scalable screening. Future work should focus on larger datasets, diverse populations, and clinical integration to further validate its utility and expand AI-driven cardiac diagnostics.

VII. REFERENCES

[1] P. Rajpurkar et al., "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," *arXiv preprint arXiv:1711.05225*, 2017.

[2] J. Irvin et al., "CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison," *AAAI Conference on Artificial Intelligence*, 2019.

[3] X. Wang et al., "ChestX-ray8: Hospital-Scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," *IEEE CVPR*, 2017.

[4] C. Yan et al., "Weakly Supervised Deep Learning for Thoracic Disease Classification and Localization," *arXiv:1807.01257*, 2020.

[5] I. M. Baltruschat et al., "Comparison of Deep Learning Approaches for Multi-label Chest X-ray Classification," *Scientific Reports*, vol. 9, no. 6381, 2019.

[6] J. Rubin et al., "Attention-Guided Convolutional Neural Networks for Chest X-ray Interpretation," *Scientific Reports*, vol. 8, no. 7545, 2018
.

[7] M. Hashir, A. Mahmood, and M. Shah, "EfficientNet-Based Cardiomegaly Classification: Performance Evaluation on Public Chest X-ray Datasets," *IEEE Access*, vol. 9, pp. 105982–105995, 2021.

[8] H. Nguyen et al., "DenseNet Architectures for Medical Image Diagnosis: An Empirical Study," *IEEE EMBC*, 2020.

[9] Z. Rehman et al., "Improving Cardiomegaly Detection through Ensemble Learning Techniques," *Computers in Biology and Medicine*, vol. 146, 2022.

[10] R. Pradeep et al., "CNN–LSTM Hybrid Network for Cardiomegaly Prediction in Chest Radiographs," *International Journal of Imaging Systems and Technology*, 2023.

[11] A. Gupta et al., "Lightweight Deep Learning Architectures for Cardiomegaly Detection in Resource-Limited Settings," *Biomedical Signal Processing and Control*, 2021.

[12] S. Khan et al., "Vision Transformers for Medical Image Analysis: A Case Study on Chest X-rays," *IEEE Transactions on Medical Imaging*, 2023.

[13] G. Huang et al., "Densely Connected Convolutional Networks (DenseNet)," *IEEE CVPR*, 2017.

[14] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *ICML*, 2019.

[15] R. Shallu and M. Mehra, "Chest X-ray Classification Using Deep Transfer Learning and Grad-CAM Visualization," *Procedia Computer Science*, vol. 167, pp. 1444–1452, 2020.