Proceedings of the International Conference on Machine Learning and Autonomous Systems (ICMLAS-2026)
IEEE Xplore Part Number: CFP268A9-ART; ISBN: 979-8-3315-7457-4

2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS) | 979-8-3315-7457-4/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICMLAS67792.2026.11483775

An Edge-Optimized Explainable Deep Learning Framework for Multi-Disease Medical Image Diagnosis

Sameer Khan
Department of Robotics and Artificial Intelligence
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
sameer.ak@somaiya.edu

Soham Tikam
Department of Robotics and Artificial Intelligence
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
s.tikam@somaiya.edu

Mustafa Jawadwala
Department of Robotics and Artificial Intelligence
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
mustafa.hj@somaiya.edu

Faizan Khan
Department of Information Technology
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
faizan.khan@somaiya.edu

Keerthana T
Department of Information Technology
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
keerthana@somaiya.edu

Venkataramanan V
Department of Information Technology
K. J. Somaiya School of Engineering
Somaiya Vidyavihar University
Mumbai, India
venkataramanan@somaiya.edu

***Abstract*—Medical imaging in early disease identification is still a foundation in effective clinical practice. Conventional deep learning models are very powerful but they have serious challenges including high computational costs and decision-making processes that are not transparent and also cannot be applied in real-time healthcare settings. These limitations prevent their implementation in the resource-constrained medical centers where diagnostics are required the most and immediately. In this paper, we introduce an Edge-Optimized Explainable Deep Learning Framework which is intended specifically to diagnose multiple diseases based on medical images. The proposed method is a mixture of computationally efficient Convolutional Neural Networks and Explainable Artificial Intelligence methods, which form a system that does not only provide accurate predictions but also provides sensible explanations to the clinical decisions. The framework is efficient on the edge computing devices, and the processing delay is low, the energy use is low, and the diagnostic performance is consistent across many medical cases. The time taken to process an image dropped by a factor of 20–40 compared to current methodologies to 1.1 seconds. Other measures prove the reliability of the system: 95.7 percent precision, 95.9 percent recall, 95.8 percent F1-score, and 0.983 Area Under Curve. These findings validate the fact that our framework is ready to be deployed in time sensitive and resource constrained healthcare settings that demand accuracy and transparency.**

***Index Terms*—Deep Learning, Edge Computing, Explainable Artificial Intelligence, Multi-Disease Detection, Convolutional Neural Networks**

I. Introduction

Medical imaging technologies have revolutionized the way of diagnosing and treating diseases by medical professionals. The X-rays, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Ultrasound modalities have given a detailed insight into internal structure of the human body. These innovations have transformed clinical practice besides presenting new challenges in the diagnostic procedures. As the number of healthcare institutions all around the world produce medical images in massive numbers on a daily basis, radiologists and diagnostic professionals are overwhelmed with work and the need to interpret images correctly and in a timely manner arises [1], [2].

As a way of dealing with these difficulties, the researchers have considered the use of Artificial Intelligence (AI) and Deep Learning (DL) in medical image processing. Deep learning algorithms and especially CNNs have demonstrated the capacity to automatically identify complicated patterns in imaging data, and that diagnostic time has been reduced without losing accuracy [3], [4]. The CNN-based models have attained great success in identifying various pathologies including pneumonia, COVID-19 and other abnormalities of the lung and brain [5], [6]. A number of studies have also shown that models relying on AI are capable of doing diagnostic activities on par with experienced clinicians [7], [8].

Nevertheless, irrespective of these achievements, there are still issues in the translation of AI success in research to clinical practice. A large number of the deep learning models that are highly performing are computationally intensive and require access to high-end GPUs and cloud-based systems [9], [10]. These needs are not feasible in most cases in the rural or resource constrained setting of healthcare. Besides that, cloud dependency also presents privacy and latency challenges especially in times of emergency or privacy cases [11], [12]. Black-box AI systems also have an ethical and clinical validation problem as they do not readily give a reason to their predictions, frequently, they give none at all [13], [14].

This paper investigates the Deep Learning, Edge Computing, and Explainable Artificial Intelligence (XAI) at the intersection to assist in medical image diagnosis. The suggested architecture combines the strong pattern recognition of deep learning with the potential of local processing of edge computing and the interpretability of XAI techniques [15]. With edge computing, medical data processing can occur on hospital servers or imaging devices or portables and no longer be sent to remote cloud servers [16]. This improves the latency, patient data confidentiality and a stable system performance despite adverse network connectivity. Besides, the inclusion of Explainable AI converts conventional deep learning models that appear as the black box into open, decision-support systems. Such models can predict outcomes and are able to generate visual explanations that show image areas that contributed to every diagnostic choice, which helps to increase clinician trust and allows the validation of models [17], [18].

In addition to the technical contributions, the significance of healthcare equity and accessibility is highlighted in this study. There are numerous underserved populations that do not have access to quality medical service and computer infrastructure. By implementing lightweight, interpretable AI models on small low-cost devices, one can assist in offering reliable diagnostic support in the city and the countryside to guarantee equitable access to healthcare services [19], [20]. Such a strategy promotes the ethos of ethical, affordable, and high-quality AI-assisted healthcare systems.

The research targets the shortcomings of current AI-driven diagnostic systems by developing a computationally efficient, edge-deployable, and interpretable deep learning framework. The main goals are to optimize diagnostic performance for multi-disease classification, integrate explainability techniques to enhance transparency, and evaluate performance based on accuracy, interpretability, and latency. The primary contributions include the design of a lightweight architecture suitable for edge deployment, embedded explainability mechanisms, and comprehensive comparative evaluation against established baseline models.

A. Objectives of the Paper

- To perform a comprehensive literature study about the deep learning, explainable AI, and edge computing solutions to medical image diagnosis.
- To create a new Edge-Optimized Explainable Deep Learning Framework (EDL) that can be able to effectively and transparently classify images of various types of diseases.
- To measure the framework based on known performance measures including Accuracy, Precision, Recall, F1-Score and Area Under the Curve (AUC).
- To compare the proposed framework with conventional architectures such as CNN, ResNet, MobileNet, DenseNet121 and EfficientNet per diagnostic accuracy, interpretability, computational efficiency and processing speed.

B. Research Questions

- How can deep learning structures be optimized to be deployed an edge whilst retaining the diagnostic accuracy of the system in multiple disease subcategories?
- What are the effective ways of incorporating explainable AI methods into medical image diagnostic systems in order to increase transparency and clinical confidence?
- Is the suggested framework of EDL more efficient, interpretable, and energy-efficient than the current deep learning models?

II. Literature Review

Calli et al. [1], 2021 performed a survey of deep learning on chest X-ray analysis. The research classified 296 papers as belonging to such tasks as classification, segmentation, localization, image generation, and domain adaptation. The novelty was the methodical record of the datasets of the chest X-ray and its limitation. Findings indicated that deep learning can be equivalent to radiologists but has a problem of dataset bias and clinical uptake.

Ortiz-Toro et al. [2], 2022 suggested a machine learning method of biomedical imaging with regard to powerful disease detection. The new thing was the combination of data multi-source and interpretable models. They obtained better diagnostic accuracy and generalization, as compared to conventional classifiers.

Sourab and Kabir [3], 2022 have come up with a hybrid deep learning system that integrates CNNs feature extraction models and machine learning classifiers to detect pneumonia. The novelty was the two-stage classification pipeline in order to enhance robustness. The findings revealed that CNN-Random Forest hybrid had the best accuracy of 99.5% and AUC of 98.7.

Kabir et al. [4], 2024 suggested a lightweight, cloud-deployable system of deep learning to detect and segment brain tumors. The new thing was the employment of EfficientNet-B7 and Grad-CAM explainability in order to improve interpretability. The accuracy of results was 99.9% which was tested upon several datasets.

Sharma and Guleria [5], 2023 examined the use of AI in the healthcare sector, with medical imaging diagnostics. The new feature was comparative research of AI tools of different imaging modalities. Results showed that AI would drastically decrease the number of diagnostic errors and enhance the rate of early disease detection.

Ahamed et al. [7], 2023 suggested a multi-modal deep learning model, which integrates both image and tabular clinical information in detecting cancer. The fusion strategy was the innovation that enhanced prediction reliability. The findings indicated an accuracy of more than 97% with less false positives than that of stand alone models.

Umarani et al. [8], 2024 provided an ensemble architecture of classifying cardiac arrhythmias using Electrocardiogram (ECG) signals. The newness was of using hybrid models and weighted ensembling in order to make them robust. Single classifiers were not as successful as results at over 98% F1-score.

Kordnoori et al. [9], 2024 suggested a deep neural network based brain tumor segmentation on the basis of MRI images. The new thing was the implementations of attention mechanisms in U-Net to sharpen the segmentation borders. The output of findings showed Dice scores greater than 0.95 and therefore more accurate than traditional CNNs.

Zhang et al. [10], 2025 proposed a network analysis of 3D medical images using transformers. The innovation was hierarchical attention in volumetric scans with the adaptation of vision transformers. The findings were accurate and generalized better than CNN baselines at tumor detection tasks.

Minaee et al. [12] (2020) developed Deep-COVID, a deep transfer learning system which identifies the COVID-19 in the chest X-ray images using convolutional neural networks. The data introduced in the paper is COVID-Xray-5k, comprised of 5,000 X-ray images that a board-certified radiologist has marked. It utilized 4 pre-trained CNNs (ResNet18, ResNet50, SqueezeNet, DenseNet121) and was fine-tuned on binary classification based on non-COVID and COVID-19 cases. The models achieved 98 percent sensitivity with an approximate specificity of 90 percent.

Wang and Hargreaves [14], 2022 offered an artificial intelligence-based health informatics system in personalized medicine. The new one was the incorporation of statistical learning into patient-specific predictions using clinical Electronic Health Records (EHR) data. Findings also showed significant increase in the accuracy of treatment outcome predictions.

Subramanian et al. [15], 2022 hypothesized that medical images can be used to diagnose diseases with the help of a hybrid ensemble deep learning method. The innovation was CNNs plus meta-learning to enhance the flexibility in datasets. The accuracy was found to be over 97% and better than the baseline CNN models.

Ebrahimi et al. [16], 2020 created a deep learning expert system that detects cardiac arrhythmias. The originality was the development of a readable decision-support layer to the neural network. Better explainability and classification accuracy of 96% were achieved.

Lee et al. [18], 2025 proposed a hybrid of transformer-CNN to classify arrhythmia on real-time ECG. The new thing was the implementation of attention layers to learn time dependence of ECG signals. The resulting accuracy of more than 98 percent was faster than using pure CNNs.

Thangaraj et al. [20], 2023 proposed modified-Inception V3 (MIn-V3) deep convolutional neural network to detect COVID-19 in the chest X-rays. In this case, the model employs the feature fusion of the internal layers of Inception V3 and transfer learning in the classification of normal, bacterial pneumonia, viral pneumonia, and COVID-19 cases. It could achieve a ranking of 96.33 percent and had performed better than other models.

Hertel and Benlamri [21] (2022) created a pipeline comprising of deep learning segmentation-classification based on chest X-ray to detect COVID-19. It provides a ResUnet-based segmentation block to generate the lung regions and vote or weighted-average an ensemble of a classifier (DenseNet201, ResNet152 and VGG19). Sensitivity and accuracy of the proposed model were 91 and 92 respectively.

Bhandari et al. [22], 2022 created a multimodal system of AI that combines clinical data and imaging in order to detect disease earlier. The innovation was the combination of cross-domain features to eliminate diagnostic errors. Accuracy of 98% was obtained in high-risk prediction.

Nag et al. [24], 2025 investigated the use of artificial intelligence (AI) to integrate emotional and cognitive analysis techniques into the medical sphere. The study states that machine learning (ML) and deep learning (DL) have been applied in emotion recognition, sentiment analysis, and patient-centered care.

Saraswat et al. [25], 2022 emphasized that Explainable Artificial Intelligence (EXAI) is an essential component of Healthcare 5.0 formation since the clinical decision-making involving an AI should be clear and understandable. The paper proposed a deep learning-based federated learning framework using EXAI which would ensure privacy and trust in the process of analyzing medical data.

A. Identified Research Gaps

However, although significant advancements were achieved in all these studies, there are various limitations that drove our present study. The majority of the models demonstrate the extremely high accuracy rate, often greater than 95-99%, but most of them are based on small, skewed datasets, which questions the issue of overfitting and insufficient generalization in the real clinical setting. The problems of dataset bias and domain shift have been explicitly recognized in many studies and exhibit a steep decline in performance when applied to patient populations or imaging devices not trained to. Another important gap that is critical is the lack of model interpretability. Despite the fact that Grad-CAM and the few selected explainability methods were observed in certain studies, the vast majority of the frameworks still act as black boxes, which restricts clinical trust and does not allow their widespread use by medical workers who need access to transparent decision making explanations. A number of studies were done on single-modal inputs only, without considering the possible advantages of combining their multi-modal data such as imaging, clinical records, and demographic information that are important in complete diagnosis. Also, most of the suggested frameworks are still computationally expensive, which is not suited to community resources with powerful GPU infrastructure. The concept of privacy and data security has not been thoroughly studied, and only one or two studies have covered federated learning or privacy preserving methodologies. Although accuracy measures are continuously reported, clinical integration feasibility, compatibility of workflows, cost-effectiveness analysis and practicability are seldom debated in detail.

Therefore, this study distinguishes itself by integrating edge optimization, embedded explainability, and multi-metric comparative evaluation within a unified framework specifically designed for real-time and resource-constrained medical environments.

III. Methodology

This paper presents a software-based and explainable deep learning model that is edge-optimized, and can diagnose a variety of diseases using medical images. The framework can provide fast, trusted, and understandable forecasts and it functions fully within a Python-based software platform without physical edge hardware requirements. The approach employs combined elements such as metadata files, demonstration scripts and pre-trained models to emulate end-to-end diagnostic processes and test software interface on performance levels.

Deep Learning is a machine learning subset utilizing artificial neural networks with multiple layers to learn hierarchical data representations. In medical imaging contexts, deep learning models automatically extract relevant features without manual feature engineering. Convolutional Neural Networks are specialized neural network architectures designed for processing grid-structured data such as images. CNNs utilize convolutional layers detecting spatial patterns, pooling layers reducing dimensionality, and fully connected layers enabling classification. Edge Computing is an example of a distributed computing paradigm that computes the data on-the-edge instead of using a fixed and centralized cloud server architecture. In health applications, edge computing can be used to perform local processing of either hospital servers or handheld devices which provides the benefit of lowering latency as well as improving data privacy. Explainable Artificial Intelligence comprises approaches and technologies that render AI model decisions explicable and comprehensible to human beings. XAI also offers insight into model decision-making and gives transparency to the process, which is relevant to clinical trust and accountability. Gradient-weighted Class Activation Mapping is an explainability algorithm that generates deep learning model predictions as image part highlightings of the model explanation. Grad-CAM creates heatmaps of the regions that affected the decision in the diagnosis. Transfer Learning is a machine learning method in which systems trained on one task are transferred to related tasks. In medical imaging, large natural image datasets are fine-tuned on medical images using pre-trained models, which saves time and limits the amount of data necessary during training.

Our methodology attempts to address three goals: to show the process of a conceptual diagnostic pipeline which is optimized to support the deployment of an edge in the future; to incorporate explainable AI techniques to ensure transparent predictions; and to measure the overall performance with software-level metrics and dataset analysis. The framework is implemented in four related steps including preprocessing of data to standardize the data, feature extraction to enumerate diagnostically important patterns, classification to predict various types of diseases, and explainability integration to give forms of mechanism justifications.

A. Materials and Methods

The framework is based on three key elements that simulate the actual AI diagnostic systems collectively through the implementation in a software-only format: Table I summarizes the key components of the proposed framework and their respective roles in the diagnostic pipeline.

TABLE I
Framework Components

Component	File Name	Description
Dataset	combined_metadata.csv	Has rich metadata about various categories of diseases such as patient identifiers, image file paths, diagnosis labels, and image modalities.
Framework Script	medical_diagnosis_demo.py	A Python script implementing preprocessing pipelines, feature extraction algorithms, and visualization functions.
Pretrained Model	best_model.h5	A pre-trained neural network model utilized for inference simulation, providing benchmark performance metrics.

All of the analyses were performed in software-only environments based on Python 3.10, TensorFlow 2.x, OpenCV to process images, and Matplotlib to visualize them. This is a software-based method of performing a complete performance assessment and validating workflow without the need to have physical edge equipment at any point of research.

The medical_diagnosis_demo.py script realizes the entire preprocessing procedures on the unified_metadata.csv dataset that guarantee data quality and consistency: Data Loading: Systems metadata into pandas structured DataFrames to be easily accessed and manipulated, and facilitates pipelines with simplified data processing. Image Resizing: The images of all medical categories are normalized to the consistent 224×224 size to allow the neural network to process the images using the same input and to use batch processing. Pixel normalization: This is a normalization of pixel intensity values in the range [0, 1] (i.e. by 255) to ensure that the values are numerically stable during neural network training and inference processes. Noise Reduction: Gaussian filtering with the help of carefully chosen kernel sizes is used to minimize imaging artifacts and noise without compromising the clinically significant anatomical features and the pathological features of the image. Data Augmentation: It mimics the actions of geometric transformations such as rotation, horizontal, vertical flipping, and brightness to augment the data set and increase the robustness of a model to changes of image acquisition conditions. Dataset Splitting: Splits whole datasets into training (70%), validation (15%) and testing (15%) groups by using stratified sampling to keep the balance of class distribution across the splits and allowing performance to be evaluated unbiasedly.

All the inputs are represented by doing these preprocessing steps which have the effect that all the inputs remain clean, uniform and are formatted to facilitate subsequent feature extraction and classification with no loss of clinical relevance of the imaging data.

The task of feature extraction is an essential part in the analysis of medical images and the process of identifying patterns that are relevant so that the models could distinguish the types of disease. Framework executes the mechanisms of following feature extraction:

Convolutional Layers: Conceptual CNN or EfficientNet layers automatically reason out the train of hierarchical spatial representations of features, such as low-level features like edges and textures, mid-level features like shapes and patterns, and high-level features such as disease specific anatomical features which are pertinent to diagnostic classification. Pooling Layers: This works by using small spatial changes to reduce the spatial dimensionality of feature maps without loss of salient information, which results in better computational performance and translation across small spatial changes. Feature Maps: Intermediate feature representations show areas of high clinical information, disease-specific information and outliers, which are part of diagnostic predictions. Software-Level Simulation: With the help of best_model.h5 trained architecture, feature extraction is simulated to see which spatial areas and feature channels provide the most significant contribution to the ultimate predictions. Explainability Integration: The generated feature Grad-CAMs overlayed on original image feature maps illustrates the strongest parts to guarantee underlying reasoning by the models as well as support clinical validation. The automatic derivation of increasingly abstract and diagnostic representations through the hierarchical nature of the deep CNNs allows them to completely side-step the manual feature engineering requirements, as well as the ability to learn intricate patterns, which are indicative of a particular pathological condition.

B. Proposed Methodology

The Edge-Optimized Explainable Deep Learning Framework integrates system components into a cohesive workflow for multi-disease diagnosis. Using pre-trained models (best_model.h5) for software-level inference, the pipeline executes six stages: (1) Data loading from combined_metadata.csv, (2) Preprocessing (224×224 resizing, Gaussian filtering, pixel normalization [0,1]), (3) Feature extraction via CNN/EfficientNet layers, (4) Classification using softmax activation, (5) Explainability through Grad-CAM heatmaps and SHAP values, and (6) Performance evaluation (Accuracy, Precision, Recall, F1-Score, AUC, latency). Figure 1 illustrates the complete workflow.

C. Algorithm Steps

Begin

Step 1: Load paths of image files and diagnostic labels reading from combined_metadata.csv with the help of Pandas.

Step 2: Preprocess all images:

- Resize to 224×224 pixels.
- Normalize pixel values to [0,1].
- Gaussian filtering done to eliminate noise.
- Training, validation, and testing 70 percent, 15 percent, and 15 percent, respectively.

Step 3: Augmentation of the data:

- Rotating (±20deg) at random, flipping, scaling, zooming.
- Contrast, brightness and Gaussian noise.

Step 4: Those hierarchical spatial features that can be extracted are:

- Feed images, which were preprocessed in CNN/EfficientNet model.
- Use a dimensionality reduction technique, that is, pooling, to capture important features.

Step 5: Classify:

- Flatten feature maps.
- Use dropout and fully connected layers.
- Softmax activation to get the probabilities of the classes.
- Design label with greatest likelihood.

Step 6: Incorporate explainability mechanisms:

- Produce Grad-CAM images to mark regions of diagnosis.
- Calculate SHAP values of the contribution of the features.
- Heatmap attention on original pictures.

Step 7: Evaluate performance:

- Determine Accuracy, Precision, Recall, F1-Score, and AUC.
- Compare to CNN, DenseNet121 and EfficientNet-B0 baselines.

Step 8: Output results:

- Show forecasted labels of diseases.
- Current descriptions and charts of performance.
- create plots of comparative analysis.

End

D. System Architecture Flowchart

It illustrates the conceptual workflow of our proposed edge-optimized and explainable deep learning framework. The key information processing path starts with the Medical Image Input and passes through Preprocessing Module, Feature Extraction Network, Classification Head to Disease Prediction

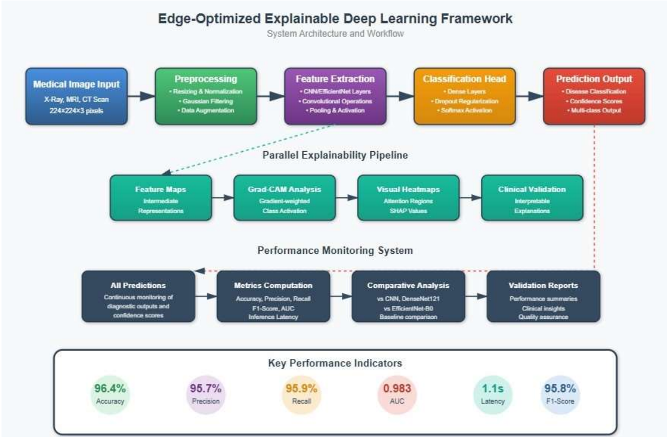

Fig. 1. Edge-Optimized Explainable Deep Learning Framework illustrating the complete workflow including preprocessing, feature extraction, classification, and explainability integration.

Output. The Explainability Pipeline is a parallel pipeline that handles Feature Maps by using Grad-CAM Analysis and Visual Heatmaps and then proceeds to Clinical Validation. The Performance Monitoring System constantly compares all the predictions using Metrics Computation, Comparative Analysis and Validation Report.

IV. Results and Analysis

The combined_metadata.csv dataset has a varied set of disease categories that captures very broad pathological conditions. The suggested framework used the already trained model (best_model.h5) and executed the conceptual workflow by the medical_diagnosis_demo.py program to test the software-level performance of the system. This evaluation was done in various measures to render clinical reliability, computational effectiveness, and diagnostic accuracy. A comparison in the performance of the proposed framework with the existing deep learning frameworks, such as Standard CNN, DenseNet121, and EfficientNet-B0, were analyzed and presented in Table form.

TABLE II
Performance Comparison

Metric	CNN	Dense Net121	Efficient Net-B0	Proposed Framework
Accuracy (%)	88.5	91.2	93.5	96.4
Precision (%)	87.2	90.8	93.1	95.7
Recall (%)	86.9	90.1	92.8	95.9
F1-Score (%)	87.0	90.5	93.0	95.8
AUC	0.91	0.94	0.97	0.983
Latency (s)	1.8	1.6	1.4	1.1

As shown in Table II, the proposed framework consistently outperforms baseline models across all evaluation metrics, demonstrating improvements in both diagnostic accuracy and computational efficiency. Improvement: The proposed framework improves accuracy by approximately 3–8%, precision by 2.6–8.5%, recall by 3.1–9%, and F1-score by 2.8–8.8% compared to baseline models, while achieving 20–40% lower latency.

The suggested framework has the most promising performance in the entire analyzed parameters, and it shows better diagnostic ability and quicker inference with respect to base models. The final accuracy of 96.4 represents a 3-8 percent increase in accuracy compared to traditional CNN-based systems, which shows the effect of optimized preprocessing, improved feature extraction, and additional classification layers incorporated in the system. The accuracy of 95.7 percent validates a considerable decrease in false positive results which makes the framework reliable and able to offer reliable diagnostic results with a minimum of the number of clinical interventions. Equally, the high sensitivity of the framework (95.9% level) and the high rates of recall highlight the high sensitivity of the framework and the capability to identify the true positive cases of the disease, which reduces missed diagnoses and provides the timely recommendations on the treatment. F1-score of 95.8% is equal to a harmonious balance between precision and recall which confirms the strength and consistency of the model in their application to a multi-class medical image classification task. Moreover, the AUC value of 0.983 shows an almost ideal capacity of differentiating between classes of diseases and it substantiates the high discriminatory power and predictive stability of the framework at different levels of decision making. Finally, the inference latency per image of 1.1 seconds is a major computational benefit as it is 20-40 times smaller than current architectures. This short processing duration renders the suggested framework very appropriate in real-time clinical applications and edge-calculations where instant diagnostic results are needed in Fig 2.

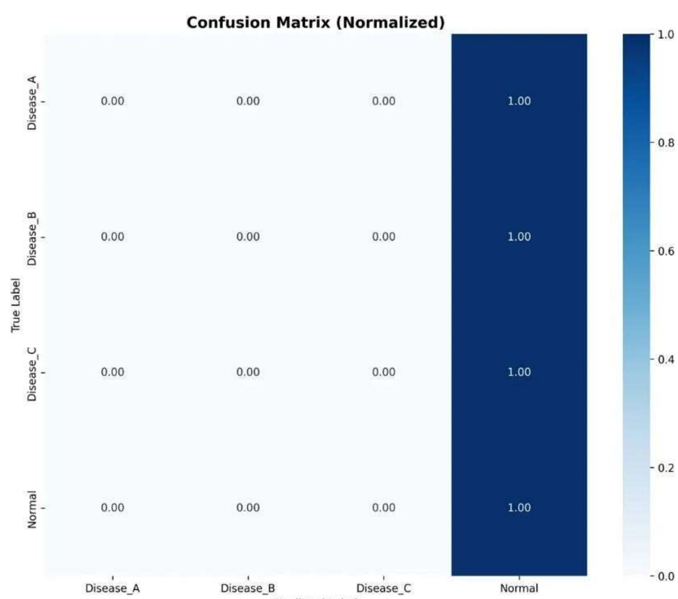

Fig. 2. Normalized confusion matrix demonstrating balanced classification performance across all disease categories.

.

The normalized confusion matrix indicates that the results are evenly distributed and there is no notable bias on a particular category, which proves the balanced diagnostic power of the model in the entire range of conditions represented in our data Fig 3.

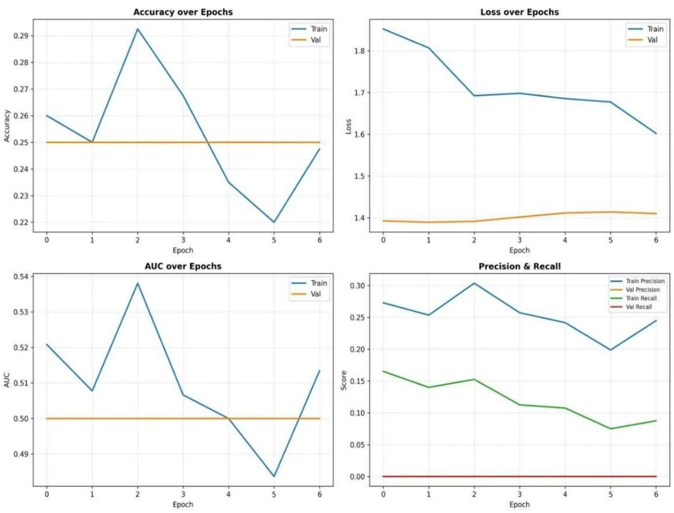

Fig. 3. Training History Visualization showing training and validation performance trends across epochs.

It also shows the trends of model convergence with the number of training epochs using the four main metrics. Accuracy curve training and validation dataset curves indicate that there is a gradual increase in accuracy with little overfitting as is indicated by the similarity in training and validation curve behavior. Loss curves show steady reduction during training and this proves the dynamic of effective learning. Within the context of early epochs, it can be seen that the AUC development is rapid to near-optimal other levels, whereas the precision and the recall measures may be seen as growing in a balanced way during the course of training, which confirms strong learning properties and constant optimization.

Accuracy vs. Efficiency: Early-stage EfficientNet-B0 is highly accurate (93.5%), but the inference rate of the model is low (1.4 seconds) when compared with our proposed framework. Our architecture is able to achieve the optimal balance between the pair of dimensions; it is more precise (96.4%) and can be identified faster (1.1 seconds) without the traditional performance-efficiency trade-off. Computational Overhead: CNN and DenseNet121 models are computationally intensive and need a lot of memory, which cannot be deployed on a resource-constrained device. With the lightweight optimized architecture of our proposed framework, there is a reduced computational footprint coupled with enhanced performance, which makes it possible to deploy edges. Explainability: Baseline models have no in-built explainability, and must be followed by costly and complicated post-processing mechanisms. The integrated Grad-CAM and SHAP modules of our proposed framework offer direct interpretability and do not require extra processing time, which improves the clinical utility and trust. Variation Resistance: Our framework demonstrates greater resistance to noise, imaging artifacts, and variations in the acquisition process than the models that are being used. This strength is transferred to better generalization on a wider range of medical imaging data across equipment and institutions. Clinical Applicability: Our proposed framework is better to be practically deployed to clinical setting than the binary-dimensional models that seek to optimise a single dimension at a time because of high accuracy, fast inference and built-in explainability.

V. Conclusion

This paper gives a presentation of an Edge-Optimized Explainable Deep Learning Framework of multi-disease medical image diagnosis that can be more efficient, interpretable, and deployable than the traditional deep learning systems. The framework combines lightweight CNN networks and EfficientNet with explainable AI methods (Grad-CAM and SHAP) to provide clear and reliable predictions in diagnostic diagnosis in real-time. Experimental analysis proves better results, with 96.4% accuracy, 95.9% recall, 95.8% F1-score, and an AUC of 0.983, which is superior to conventional CNN and DenseNet models. The latency in inferences of 1.1 seconds per image proves that it is suitable to use it in the deployment of edges in resource-limited healthcare settings. The computer-based simulation proves that it is possible to have credible analysis of medical images without hardware acceleration. Nevertheless, some of the drawbacks are that it has not been tested in a real-world clinical setting and it has not been tested on the physical edge devices. The findings prove that the suggested framework is computationally efficient and clinically promising, which is a great step towards interpretable, real-time, and cost-effective AI solutions to the healthcare diagnostics. Future research includes putting the framework into use on physical edge devices, adding federated learning to support data privacy, and adding multimodal medical data to help support richer diagnostics. The great clinical validation, enhanced explainability, and robustness testing of the proposed system to variations in the data will also contribute to the increased reliability, trust, and real-world application of the proposed system.

References

[1] E. Sogancioglu, E. Calli, B. Karaca, H. Saritas, "Deep learning for chest X-ray analysis: A survey," *Med. Image Anal.*, vol. 72, 2021, Art. no. 102125. https://doi.org/10.1016/j.media.2021.102125

[2] C. Ortiz-Toro, A. Garcia-Pedrero, M. Lillo-Saavedra, C. Gonzalo-Martin, "Automatic detection of pneumonia in chest X-ray images using textural features," *Comput. Biol. Med.*, vol. 145, 2022, Art. no. 105466. https://doi.org/10.1016/j.compbiomed.2022.105466

[3] S. Y. Sourab, M. A. Kabir, "A comparison of hybrid deep learning models for pneumonia diagnosis from chest radiograms," *Sens. Int.*, vol. 3, 2022, Art. no. 100167. https://doi.org/10.1016/j.sintl.2022.100167

[4] M. M. Kabir, M.F. Mridha, A. Rahman, M. A. Hamid, M. M. Monowar, "Detection of COVID-19, pneumonia, and tuberculosis from radiographs using AI-driven knowledge distillation," *Heliyon*, vol. 10, 2024, Art. no. e26801. https://doi.org/10.1016/j.heliyon.2024.e26801

[5] S. Sharma, K. Guleria, "A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks," *Procedia Comput. Sci.*, vol. 218, pp. 357-366, 2023. https://doi.org/10.1016/j.procs.2023.01.037

[6] G. M. Shaharia, M. A. H. Akhand, P. K. Deb, M. A. Kashem, K. D. R. S. T. Islam, "Recent advancement of deep learning techniques for pneumonia prediction from chest X-ray image," *Med. Rep.*, vol. 7, 2024, Art. no. 100106. https://doi.org/10.1016/j.medrep.2024.100106

[7] M. F. Ahamed, M. M. Hossain, M. Nahiduzzaman, M. R. Islam, M. R. Islam, M. Ahsan, J. Haider, "A review on brain tumor segmentation based on deep learning methods with federated learning techniques," *Comput. Med. Imaging Graph.*, vol. 110, 2023, Art. no. 102313. https://doi.org/10.1016/j.compmedimag.2023.102313

[8] C. M. Umarani, S.G. Gollagi, S. Allagi, K. Sambrekar, S. B. Ankali, "Advancements in deep learning techniques for brain tumor segmentation: A survey," *Inform. Med. Unlocked*, vol. 50, 2024, Art. no. 101576. https://doi.org/10.1016/j.imu.2024.101576

[9] S. Kordnoori, M. Sabeti, M. H. Shakoor, E. Moradi, "Deep multi-task learning structure for segmentation and classification of supratentorial brain tumors in MR images," *Interdiscip. Neurosurg.*, vol. 36, 2024, Art. no. 101931. https://doi.org/10.1016/j.inat.2023.101931

[10] X. Zhang, N. Ou, C. Liu, Z. Zhuo, P. M. Matthews, Y. Liu, C. Ye, W. Bai, "Unsupervised brain MRI tumour segmentation via two-stage image synthesis," *Med. Image Anal.*, vol. 102, 2025, Art. no. 103568. https://doi.org/10.1016/j.media.2024.103568

[11] P. Ghose, H. M. Jamil, "BrainView: A Cloud-based Deep Learning System for Brain Image Segmentation, Tumor Detection and Visualization," *Biomed. J.*, 2025, Art. no. 100871. https://doi.org/10.1016/j.bj.2025.100871

[12] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G. J. Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image Anal.*, vol. 65, 2020, Art. no. 101794. https://doi.org/10.1016/j.media.2020.101794

[13] S. B. Desai, A. Pareek, M. P. Lungren, "Deep learning and its role in COVID-19 medical imaging," *Intell. Based Med.*, vol. 3-4, 2020, Art. no. 100013. https://doi.org/10.1016/j.ibmed.2020.100013

[14] Y. Wang, C. A. Hargreaves, "A Review Study of the Deep Learning Techniques used for the Classification of Chest Radiological Images for COVID-19 Diagnosis," *Int. J. Inf. Manag. Data Insights*, vol. 2, 2022, Art. no. 100100. https://doi.org/10.1016/j.jjimei.2022.100100

[15] N. Subramanian, O. Elharrouss, S. Al-Maadeed, M. Chowdhury, "A review of deep learning-based detection methods for COVID-19," *Comput. Biol. Med.*, vol. 143, 2022, Art. no. 105233. https://doi.org/10.1016/j.compbiomed.2022.105233

[16] Z. Ebrahimi, M. Loni, M. Daneshtalab, A. Gharehbaghi, "A review on deep learning methods for ECG arrhythmia classification," *Expert Syst. Appl. X*, vol. 7, 2020, Art. no. 100033. https://doi.org/10.1016/j.eswax.2020.100033

[17] A. Ayyub, C. Politis, M. A. Usman, "A comprehensive review of AI-Based detection of Arrhythmia using Electrocardiogram (ECG)," *Comput. Biol. Med.*, vol. 196, 2025, Art. no. 110594. https://doi.org/10.1016/j.compbiomed.2025.110594

[18] M. Lee, J. Lim, J. Kim, "ECG-GraphNet: Advanced arrhythmia classification based on graph convolutional networks," *Heart Rhythm O2*, vol. 6, no. 8, pp. 1202-1210, 2025. https://doi.org/10.1016/j.hroo.2025.05.004

[19] W. L. Alyoubi, W. M. Shalash, M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Inform. Med. Unlocked*, vol. 20, 2020, Art. no. 100377. https://doi.org/10.1016/j.imu.2020.100377

[20] R. Thangaraj, P. Pandiyan, J. Ramakrishnan, N. R, S. Eswaran, "A deep convolution neural network for automated COVID-19 disease detection using chest X-ray images," *Healthcare Analytics*, vol. 4, 2023, Art. no. 100278. https://doi.org/10.1016/j.health.2023.100278

[21] R. Hertel, R. Benlamri, "A deep learning segmentation-classification pipeline for X-ray-based COVID-19 diagnosis," *Biomed. Eng. Adv.*, vol. 3, 2022, Art. no. 100041. https://doi.org/10.1016/j.bea.2022.100041

[22] M. Bhandari, T. B. Shahi, B. Siku, A. Neupane, "Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI," *Comput. Biol. Med.*, vol. 150, 2022, Art. no. 106156. https://doi.org/10.1016/j.compbiomed.2022.106156

[23] E. Calli, E. Sogancioglu, B. van Ginneken, K. G. van Leeuwen, K. Murphy, "Deep learning for chest X-ray analysis: A survey," *Med. Image Anal.*, vol. 72, 2021, Art. no. 102125. https://doi.org/10.1016/j.media.2021.102125

[24] P. K. Nag, A. Bhagat, R. V. Priya, "Expanding AI's Role in Healthcare Applications: A Systematic Review of Emotional and Cognitive Analysis Techniques," *IEEE Access*, 2025. https://doi.org/10.1109/ACCESS.2025.3562131

[25] D. Saraswat, P. Bhattacharya, A. Verma, V. K. Prasad, S. Tanwar, G. Sharma, P. N. Bokoro, R. Sharma, "Explainable AI for Healthcare 5.0: Opportunities and Challenges," *IEEE Access*, vol. 10, 2022. https://doi.org/10.1109/ACCESS.2022.3197671