Multi-Domain Medical Image Classification Using EfficientNet-B3

Dr. Alugonda Rajani, Dasam Raj Kumar

Assistant Professor, Department Of Electronics & Communivcation Engineering, University College Of Engineering Kakinada, JNTUK, AndhraPradesh, India, rajani.alugonda@jntucek.ac.in

Department Of Electronics & Communication Engineering, University College Of Engineering Kakinada, JNTUK, AndhraPradesh, India, rajkumardasam21@gmail.com

Abstract:- Medical image analysis has seen considerable progress with the integration of deep learning approaches, especially convolutional neural networks (CNNs), which are known to be successful in extracting relevant spatial information from complex biomedical data. However, most existing models are generally developed for a particular application and are not flexible enough to be used on multiple disparate datasets in a single system. In real-world medical applications, ECE images for cardiac analysis, EEG-based graphical representations for seizure analysis, and facial images for emotion analysis are quite disparate in terms of structural information and feature distributions, which make it quite challenging for existing CNN models to generalize across domains without resorting to multiple models or repeating training. These application-specific models are quite computationally intensive and less efficient in terms of deployment. To overcome these hurdles, this study proposes a unified classification platform based on the EfficienyNet-B3 model, which is capable of handling multiple biomedical imaging applications using a single backbone model. The proposed platform incorporates a dynamic key-driven selection module that enables runtime switching between applications without requiring the loading of multiple independent models, thus overcoming redundancy while preserving excellent predictive accuracy and scalability for real-world applications.

Keywords: *EfficientNet-B3, Biomedical Image Classification, Heart Disease Detection, Epileptic Seizure Detection, Facial Emotion Recognition.*

1. Introduction

The demand for intelligent healthcare monitoring systems has increased the importance of automated biomedical signal and medical image interpretation. Deep learning methods, especially Convolutional Neural Networks (CNNs), have demonstrated their ability to learn complex spatial patterns from biomedical data, which enhances clinical applications through better performance.

Several studies have focused on specific prediction tasks within individual domains, especially in epileptic seizure detection, where hybrid architectures combing convolutional and recurrent neural layers have been explored to capture both spatial and temporal information from EEG signals.

Rye and Joe [1] created a Densenet-LSTM framework which achieved competitive results through its sequential feature modelling approach, while Usman et al. [2] proved that CNN- based feature extraction methods deliver better performance than traditional machine learning techniques. Deep learning strategies have been used in both stress assessment and physiological monitoring, which Arza et al. [3] demonstrated through their work on physiological biomarkers that create strong feature representation. Batbaatar et al. [4] established a neural framework for semantic-emotion analysis through deep feature learning while Srinivasu et al. [5] developed interpretable artificial intelligence methods for stroke prediction which support cardiovascular risk evaluation. The existing deep models demonstrate effectiveness, yet their current design remains focused on specific tasks and requires separate architectural solutions for each biomedical modality, which leads to higher computational needs and restricts system expansion in diverse healthcare settings.

2. Related Work

The systems which include standard CNNs, VGGNet, ResNet and NASNet and their deep learning architectures have achieved widespread use in medical image classification because their feature learning abilities produce strong results. The models require extensive parameter counts and high computational demands which result in reduced performance for real-world usage.

EfficientNet applies its compound scaling approach which integrates network depth and width and input resolution to improve accuracy while keeping computational costs down. EffievcientNet-B3 provides an optimal performance solution which utilizes reduces system resources to achieve high classification efficiency through its lower memory needs and fewer parameters. Existing research does not establish a unified system that enables model selection for different biomedical datasets while allowing users to choose tasks during operation.

3. Proposed Methodology

3.1 System Overview

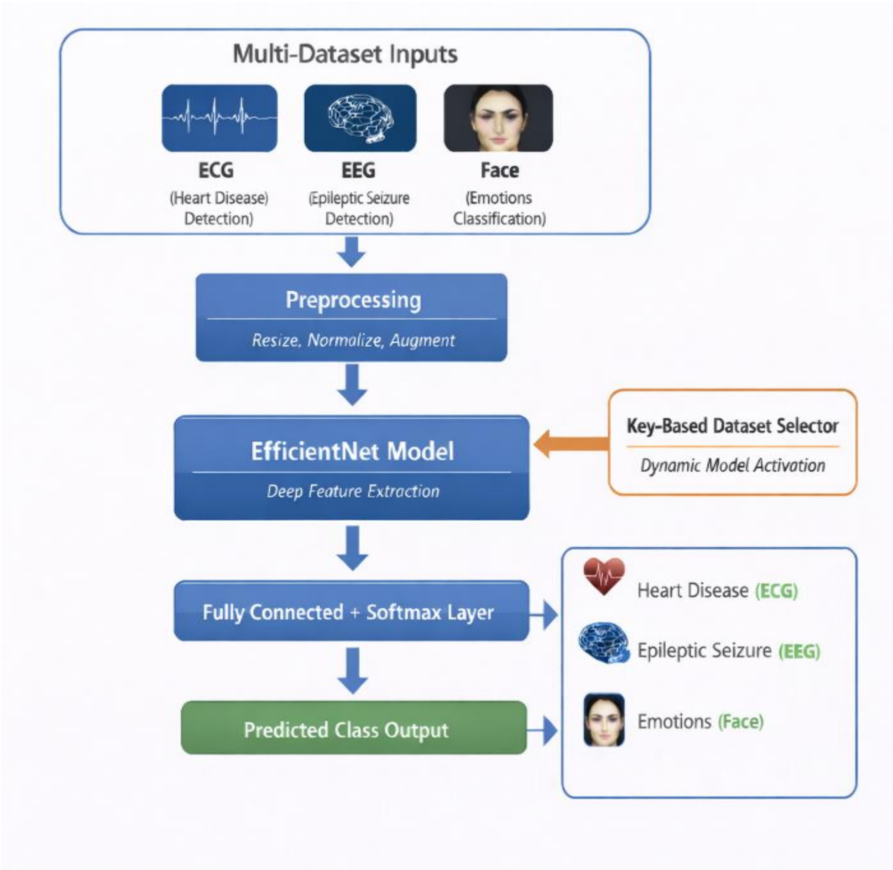

Fig-1: Block Diagram of Proposed Methodology

The created system will function through its three main elements which include preprocessing input images for their initial state and extracting their essential features through the EfficientNet-B3 model while using a dynamic Key-based system to choose which datasets to analyse. The user begins the process by entering a numeric key through the terminal interface which allows the system to access the appropriate dataset for evaluation purposes. The system uses key 1 to access ECG data which detects heart disease, while key 2 uses EEG data to identify seizures and key 3 accesses facial image data for emotion recognition purposes. The framework enables multi-domain classification through its single trained architecture which combines multiple tasks while achieving better design efficiency and operational performance in real-world situations.

3.2 Preprocessing

All datasets were converted into image format and standardized to match the input resolution required by EfficientNet-B3. Preprocessing steps include:

3.2.1. Image Resizing

The EfficientNet-B3 architecture requires image resizing because it needs specific input dimensions. The different acquisition environments which produce ECG and EEG and facial emotion images lead to their distinct resolution and aspect ratio differences which disrupt model training consistency. The preprocessing step creates a standardized input format which enables stable feature extraction and enhances classification accuracy for different biomedical datasets.

3.2.2. Intensity Normalization

The process of intensity normalization reduces existing variations in pixel value distributions by applying different values to ECG, EEG and facial image datasets which originate from distinct acquisition settings and lighting conditions and contrast levels. The process of standardizing pixel values to a common numerical range enables the training process to achieve higher stability in gradient updates.

3.2.3. Data Augmentation

Data argumentation helps to train model better because it enables to create new data samples which helps them analyse biomedical data that has insufficient samples. Controlled transformations are introduced to simulate realistic variations in ECG, EEG, and facial images, including minor rotations, scaling, brightness, adjustments, and horizontal flipping for facial data. The operations created new data variations while maintaining all important clinical Structural elements.

3.2.4. Noise Handling

The system uses noise handling methods to reduce the impact of artifacts which commonly exist in ECG, EEG, and facial image datasets. The data exists as image yet background distortions plus recording noise plus contrast irregularities create problems that reduce the quality of feature extraction. The solution uses light smoothing and filtering techniques which enhance visual clarity without creating excessive changes to the data.

3.3 EfficientNet-B3 Architecture

The EfficientNet-B3 model uses a compound scaling method to achieve efficient classification results which requires less computational power than traditional methods. The system employs Mobile Inverted Bottleneck (MBconv) blocks which use their depth-wise separable and pointwise convolution methods to achieve efficient computations while delivering strong feature representation capabilities. The block use Batch Normalization to stabilize training and the Swish activation function to enable efficient nonlinear learning processes while the Squeeze-and-Excitation mechanism dynamically increase important channel feature visibility. The use of residual skip connections in the design enables better gradient transmission throughout the system while reducing the performance issues that affect deep network models.

A global Average Pooling layer creates a compact representation of spatial information from hierarchical features which it sends to a fully connected layer that uses SoftMax activation to perform final classification. The model starts with ImageNet pretrained weights which it uses to improve convergence before training on biomedical image datasets for domain-specific pattern recognition.

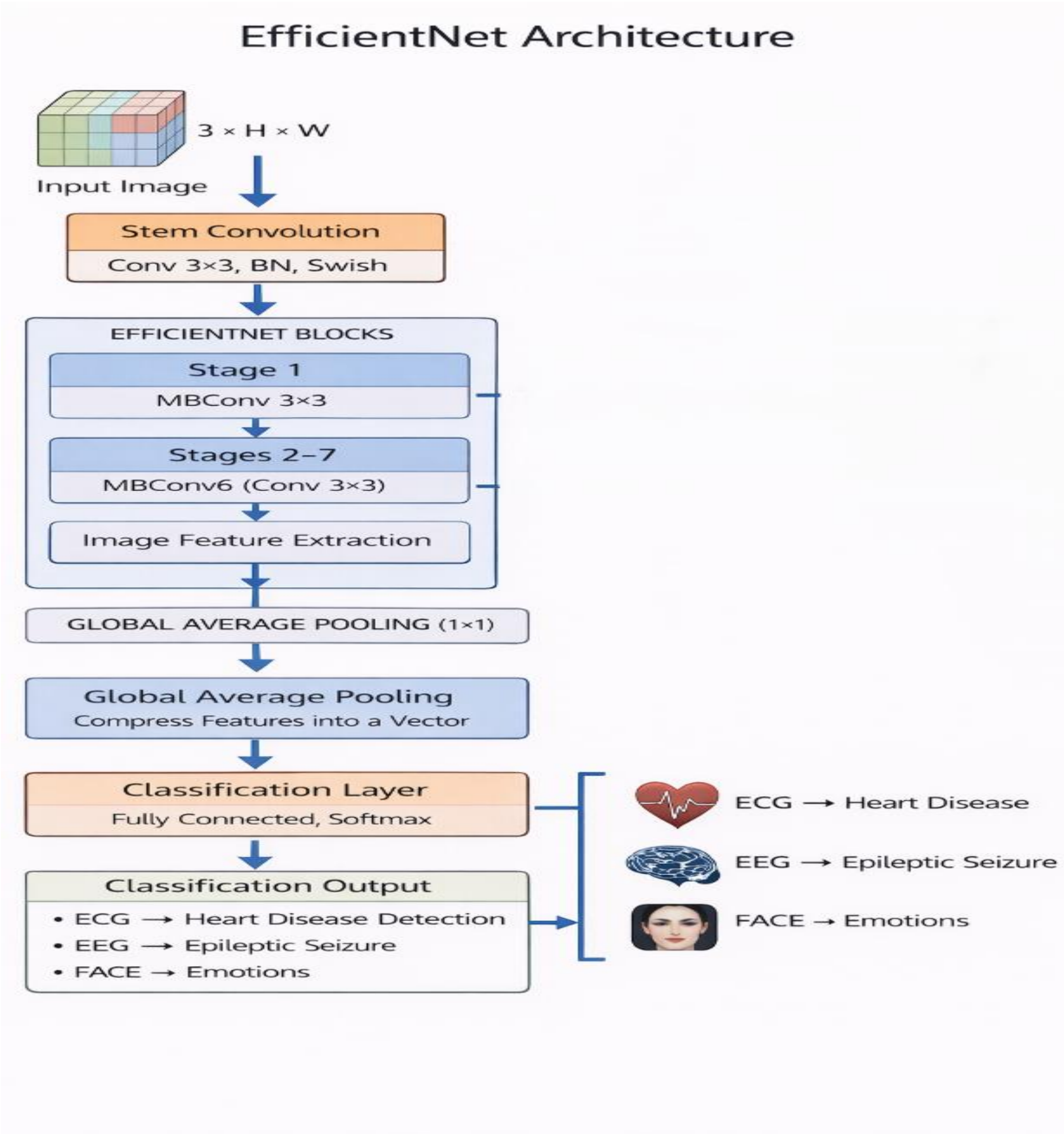

Fig-2: EfficientNet Architecture

Let a biomedical input image be denoted as

$$X \in R^{H*W*C}$$

Where H, W, and C represent the height, width, and number of channels, respectively. A convolutional neural network (CNN) classifier can be expressed as

$$\hat{y} = F(x;\theta)$$

where θ denotes the trainable parameters of the network and $\hat{y}$ is the predicted class label. For multi-class classification with k categories, the final fully connected layer produces logic z_i, which are transformed into probabilities using the SoftMax function:

$$P(y = i \mid z) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \; i = 1, 2, \ldots, K$$

Here, z_i represents the raw output for class i, and P (y = i | z) denotes the predicted probability of that class.

Rather than constructing separate models for ECG-based heart disease detection, EEG-based seizure classification, and facial emotion recognition, this study adopts a unified EfficientNet-B3 backbone capable of processing heterogeneous biomedical datasets within a single framework. EfficientNet employs compound scaling to jointly adjust network depth (d), width (w), and input resolution (r) using a scaling coefficient ϕ:

$$d = \alpha^{\emptyset}, \; w = \beta^{\emptyset}, \; r = \gamma^{\emptyset}$$

subject to the constraint

$$\alpha * \beta^2 * \gamma^2 \approx 2$$

This scaling strategy maintains a balance between representational capacity and computational efficiency.

To enable flexible deployment, a dynamic key-based selector is incorporated for runtime dataset selection. Let the chosen dataset be defined as

$$D_k \in \{D_{ECG}, D_{EEG}, D_{FACE}\}$$

based on user input key k. The overall classification mapping can therefore be written as

$$\hat{y} = F(x \mid D_{k;} \theta)$$

Where F denotes the shared EfficientNet-B3 model across all domains.

The proposed framework achieves an overall classification accuracy of 96.30%, indicating strong generalization across ECG, EEG, and facial images datasets. Performance is evaluated using the standard accuracy metric:

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

The primary contributions of this work include:

- A unified EfficientNet-B3 architecture for multi-domain biomedical image classification
- Integration of a runtime key-based dataset selection mechanism
- Simultaneously handling of ECG, EEG, and facial emotion datasets within a single trained model
- Achievement of 96.30% overall accuracy across heterogeneous domains

3.4 Dynamic Terminal-Based Key Selector

The study presents its primary innovation through its runtime implementation of dynamic dataset selection which allows multiple tasks to be executed within one operational system. The system uses a terminal interface to receive user-defined numeric keys which determine biomedical classification instead of requiring separate trained models for each biomedical application. The system permits users to pick their needed datasets together with their respective tasks while maintaining a common architectural framework. The system decreases deployment difficulties while increasing efficiency to enable hospitals to perform biomedical research across multiple domains in a cost-effective manner.

4. Results

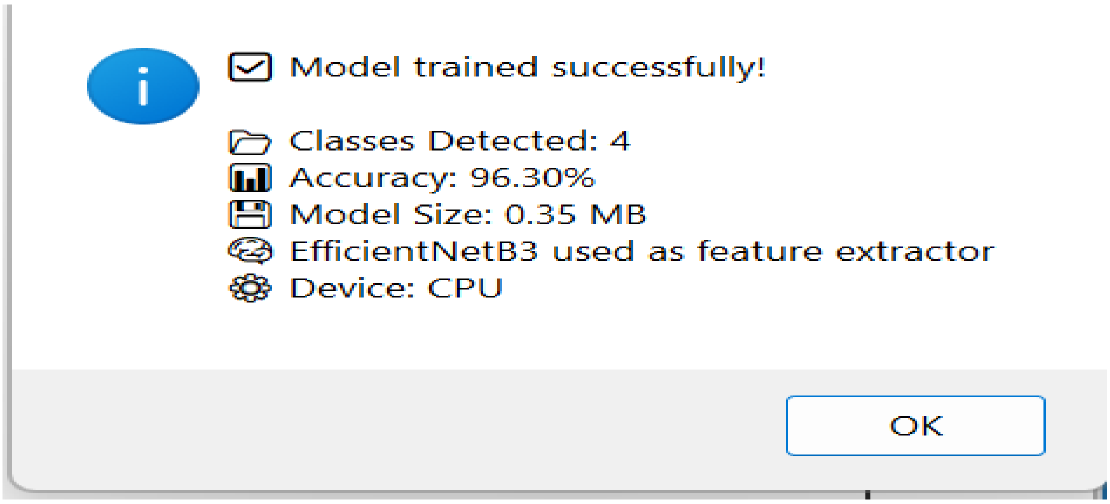

Fig-3: Parameters Of the Trained Model

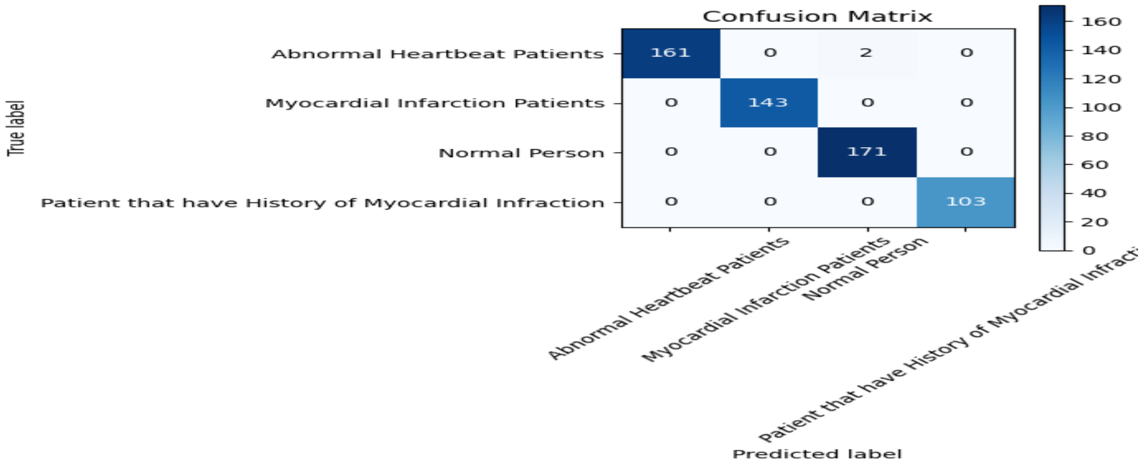

Fig-4: Confusion Matrix

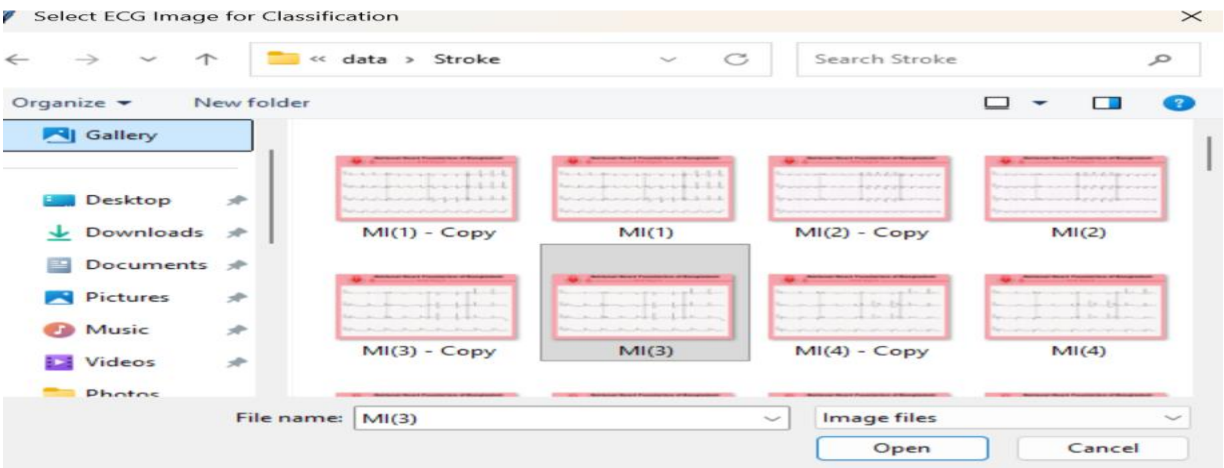

Fig-5: Input From Dataset

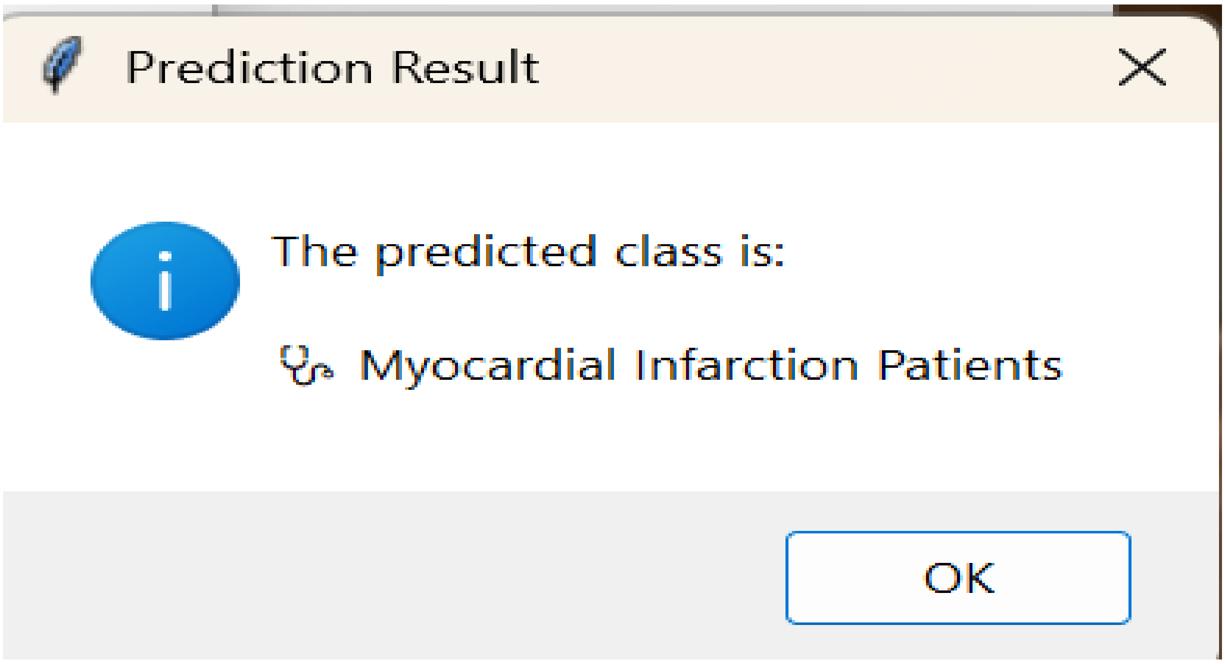

Fig-6: Classified Output

5. Conclusion

The research studies biomedical image classification through its application of EfficientNet-B3 architecture as the standard backbone for the study. The system introduces a dynamic terminal-based key selection mechanism to enable multi-domain operation through a single trained model which eliminates the requirement for different architectures that handle multiple modalities. The unified design approach achieves two advantages by decreasing the need for duplicate computing resources and making it easier to deploy systems in real-world healthcare settings. The experimental results demonstrates that the framework achieves 96.30% accuracy across ECG, EEG, facial image datasets while using compound scaling and transfer learning for efficient computation. The results demonstrate strong generalization abilities across various biomedical input structures. Future work may explore extending the frame work to additional medical imaging modalities, integrating multimodal data fusion strategies, and optimizing the architecture for real-time embedded or edge-based healthcare systems.

Refrences

[1] S. Ryu and I. Joe, "A hybrid DenseNet-LSTM model for epileptic seizure prediction," Applied Sciences, vol.11, no.16, p.7661, Aug.2021.

[2] S. M. Usman, S. Khalid, and M. H. Aslam, "Epileptic seizures prediction using deep learning techniques," IEEE Access, vol.8, pp.39998-40007, 2020.

[3] A. Arza, J. M. Garzon-Rey, J. Lazaro, E. Gil, R. lopez-Anton, C. de la camara, P. Laguna, R. Bailon, and J. Aguilo, "Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress," Medical & Biological Engineering & Computing, vol.57, no.1, pp.271-287, Jan.2019.

[4] E. Batbaatar, M. Li, and K. H. Ryu, "Semantic-emotion neural network for emotion recognition from text," IEEE Access, vol.7, pp.111866-118878, 2019.

[5] P. N. Srinivasu, U. Sirisha, K. Sandeep, S. P. Praveen, L. P. Maguluri, and T. Bikku, "An interpretable approach with explainable AI for heart stroke prediction," Diagnostics, vol.14, no.2, p.128, Jan.2024.

[6] Alugonda Rajani, “Automated detection of ECG noise signal and denoising using CEEMD algorithm” vol.8,pp 0-8 ,2020/12
[7] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, “Towards in sighting cybersecurity for healthcare domains: A comprehensive review of recent practices and trends,” Cyber Security Applications, vol.1, Art.no.100016, Dec.2023.
[8] Alugonda Rajani , “ANNET:A Light Weight Neural Network For ECG Error Detection In Edge Sensors” pp 156-162,2023/9
[9] S. Ponnada, T. K. Tak, P. R. Kshirsagar, P. S. Rao, and A. Dayal, “Expanding applications of TinyML in versatile assistive devices: From XGBoost transfer learning,” IEEE Access, 2024, doi:10.1109/Access.2024.3496791
[10] Alugonda Rajani , “A Novel assessment for SVM classification method for unsupervised ECG diagnosis Environments” vol .12 pp 0-3,2020/12