AI-Based Medical Image Classification Using Fine-Tuned EfficientNetB0

Introduction:

Medical imaging is a cornerstone of modern diagnostics, but manually sorting and classifying large image sets is slow and error-prone. This project presents an AI-driven approach for **automatically classifying images as medical or non-medical**, processing both **web URLs** and **PDF documents** through a fine-tuned EfficientNetB0 model integrated into an interactive Streamlit interface.

Objectives:

- **Automated Image Classification** – Accurately identify medical vs. non-medical images.
- **Dual Input Support** – Classify images from URLs or PDF files.
- **Performance Tracking** – Provide inference time, throughput, and model size.
- **User-Friendly Interface** – Deliver results through a clean Streamlit dashboard.
- **Scalability** – Handle multiple images efficiently.

Methodology:

Data Input

- **URL Mode**: Scrapes images from the given web page.
- **PDF Mode**: Extracts images embedded in uploaded PDF files.

Model Architecture

- **Base Model**: EfficientNetB0 pre-trained on ImageNet.
- **Fine-Tuning Steps**:
    - Replaced the final classification layer with a custom fully connected layer for binary classification (medical / non-medical).
    - Used **transfer learning** to leverage pre-trained weights while adapting to the medical image domain.
    - Unfroze later layers of the network to improve feature learning on the target dataset.
    - Applied **data augmentation** (rotation, flipping, brightness adjustment) to improve generalization.
    - Used **Adam optimizer** with a low learning rate for stable convergence.
    - Employed **cross-entropy loss** as the loss function.
    - Validated on a separate dataset to prevent overfitting.

Processing Pipeline

1. Extract images from input source.
2. Preprocess (resize, normalize).
3. Classify with the trained model.
4. Display predictions and confidence scores.
5. Show performance metrics in the dashboard.

Features of the Application:

- Supports URL and PDF image classification.
- Real-time performance monitoring (time, throughput, model size).
- Displays predictions with confidence values.
- Stores classification history.
- Handles errors gracefully**.**

Evaluation:

The model was tested on a dataset of 465 images (267 medical, 198 non-medical).

Classification Report:

Class	Precision	Recall	F1-Score	Support
Medical	1.00	1.00	1.00	267
Non-Medical	1.00	0.99	1.00	198
Accuracy			1.00	465
Macro Avg	1.00	1.00	1.00	465
Weighted Avg	1.00	1.00	1.00	465

Performance Metrics

URL-based Classification

- Found: 36 images in 20.41 seconds (including extraction)
- Total Inference Time: 7.96 sec
- Avg Time/Image: 221.15 ms

- Throughput: 4.52 images/sec
- Model Size: 15.58 MB

PDF-based Classification

- Total Inference Time: 0.26 sec
- Avg Time/Image: 132.28 ms
- Throughput: 7.56 images/sec
- Model Size: 15.58 MB

These results show near-perfect accuracy and fast processing speeds, making the system suitable for real-time applications.