IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Explainable Ai-Based Classification Of Capsule Endoscopy Images

[1]K. V. V. Siva Prasad, [2]Y. Priyanka,[3]G. Asrith, [4]A. Pavan Kalyan, [5]S. Mahaboobvali
[1]Assistant Professor, [2]Final Year B.Tech Student, [3]Final Year B.Tech Student, [4]Final Year B.Tech Student, [5]Final Year B.Tech Student,
Department of CSE - AIML,
Aditya College of Engineering & Technology (A), Surampalem, Andhra Pradesh, India

Abstract: The Endoscopy AI System is a deep learning-based application designed to assist in the detection of gastrointestinal diseases from endoscopy images. The system utilizes a Convolutional Neural Network (CNN) based on the MobileNetV2 architecture to classify images into four categories: Normal, Bleeding, Ulcer, and Polyp. The model is trained on labeled medical image datasets and optimized to provide accurate and fast predictions.

To enhance transparency and interpretability, the system integrates Explainable AI (XAI) techniques such as Grad-CAM, Layer-CAM, and LIME. These methods generate visual explanations by highlighting the regions of the image that influence the model's decision, enabling better understanding and trust in the prediction process.

A web-based interface is developed using the Django framework, allowing users to upload endoscopy images and receive real-time predictions along with confidence scores and visual explanations. The system also maintains prediction history and includes an admin dashboard for monitoring and analysis.

Experimental results demonstrate that the proposed system achieves high accuracy and efficient performance, making it suitable for assisting medical professionals in early diagnosis and reducing manual effort. The integration of deep learning with explainable AI highlights the potential of intelligent systems in real-world healthcare applications.

Index Terms - Endoscopy, Deep Learning, MobileNetV2, Explainable AI, Image Classification, Grad-CAM, LIME, Django, Medical Imaging, Computer-Aided Diagnosis

I. INTRODUCTION

Gastrointestinal diseases are among the most common health issues worldwide, and early detection plays a crucial role in preventing severe complications such as ulcers, bleeding, and cancer. Endoscopy is a widely used medical procedure that allows doctors to visually examine the digestive tract using specialized imaging devices. However, manual analysis of endoscopy images is time-consuming, requires expert knowledge, and may lead to human errors due to fatigue and subjective interpretation.

With the rapid advancement of Artificial Intelligence (AI) and Deep Learning (DL), automated systems have emerged as effective tools for medical image analysis. Convolutional Neural Networks (CNNs) have shown significant success in extracting complex features from medical images and performing accurate classification. These systems can assist medical professionals by providing faster and more consistent diagnostic support.

This project proposes an Endoscopy AI System that utilizes a deep learning model based on the MobileNetV2 architecture to classify endoscopy images into four categories: Normal, Bleeding, Ulcer, and Polyp. The system is designed to process images efficiently and provide real-time predictions through a web-based interface.

To address the challenge of interpretability in deep learning models, the system integrates Explainable AI (XAI) techniques such as Grad-CAM, Layer-CAM, and LIME. These techniques generate visual explanations by highlighting important regions in the image that influence the model's decision, improving transparency and trust in the system.

The proposed system is implemented using the Django framework, enabling seamless interaction between users and the prediction model. By combining deep learning with explainable AI and a user-friendly interface, the system aims to enhance diagnostic efficiency, reduce manual effort, and support healthcare professionals in decision-making.

II. SYSTEM ARCHITECTURE

The proposed Endoscopy AI System follows a modular architecture designed for scalability, reliability, and efficient medical image processing. The system is divided into four major layers: Frontend Layer, Backend Layer, Machine Learning Layer, and Data Storage Layer.

A. Frontend Layer

The frontend layer provides a user-friendly web interface for users and administrators. It allows users to upload endoscopy images, view prediction results, and access prediction history.

The interface is developed using HTML, CSS, and Bootstrap to ensure responsiveness and ease of use. It displays classification results along with confidence scores and visual explanations generated by Explainable AI techniques.

B.Backend Layer

The backend layer is implemented using the Django web framework. It handles application logic, routing, authentication, and communication between the frontend and the machine learning model.

The backend processes user requests, manages image uploads, and returns prediction results. Secure login and authentication mechanisms ensure that only authorized users can access system functionalities.

C. Machine Learning Layer

The machine learning layer performs image classification using a Convolutional Neural Network (CNN) based on the MobileNetV2 architecture.

The model processes preprocessed endoscopy images and classifies them into four categories:

- Normal
- Bleeding
- Ulcer
- Polyp

Additionally, Explainable AI (XAI) techniques such as Grad-CAM, Layer-CAM, and LIME are applied to generate visual explanations, highlighting important regions of the image that influence the prediction

D. Data Storage Layer

The data storage layer uses SQLite database (default Django database) to store user details, prediction history, and system logs.

This centralized database supports efficient data retrieval, user management, and monitoring of prediction results.

III. IMPLEMENTATION

A. Image Upload and Prediction Module

The image upload module allows users to submit endoscopy images through the web interface. Once an image is uploaded, it is validated and processed before being passed to the trained deep learning model.

The system preprocesses the image by resizing it to 224 × 224 pixels and normalizing pixel values. The MobileNetV2 model then predicts the class label and generates a confidence score.

B. Explainable AI Module

The system integrates Explainable AI techniques to improve transparency:

- Grad-CAM generates heatmaps showing important regions
- Layer-CAM provides refined visual explanations
- LIME explains predictions at a local level

These techniques help users understand how the model arrived at its prediction.

C. User Dashboard and History Module

The system provides a dashboard where users can:

- View prediction results
- Access previous predictions
- Analyze confidence scores

The history module stores all predictions for future reference.

D. Admin Module

The system includes an admin dashboard that allows administrators to monitor user activity, view prediction history, and analyze system performance.

E. Security and Authentication

The system uses Django's built-in authentication for secure login and user management. Role-based access control is implemented, and user data is securely handled.

IV. RESULTS

The Endoscopy AI System was evaluated based on prediction accuracy, response time, and model interpretability.

A. Performance Metrics

- The system achieves an accuracy of approximately **92%–95%**
- Prediction response time is less than **2 seconds**
- The system provides confidence scores for each prediction

B. Prediction Accuracy

The model successfully classifies endoscopy images into four categories with high accuracy.

- Normal images are classified with high confidence
- Bleeding and Ulcer cases are detected effectively
- Slight variation may occur for visually similar classes

C. Explainable AI Results

The system generates visual explanations using XAI techniques:

- Grad-CAM highlights important image regions
- LIME explains prediction behaviour
- Improves interpretability and trust

D. System Efficiency

- Faster than manual diagnosis
- Reduces workload on medical professionals
- Provides real-time predictions

Table 1 Performance evaluation summary

Metric	Proposed System	Existing/Manual System
Prediction Response Time	<2 sec	Slow (Manual analysis)
Model Accuracy	92%-95%	Variable (depends on expertise)
Confidence Output	Available	Not Available
Explainability	Available (XAI techniques)	Not Available

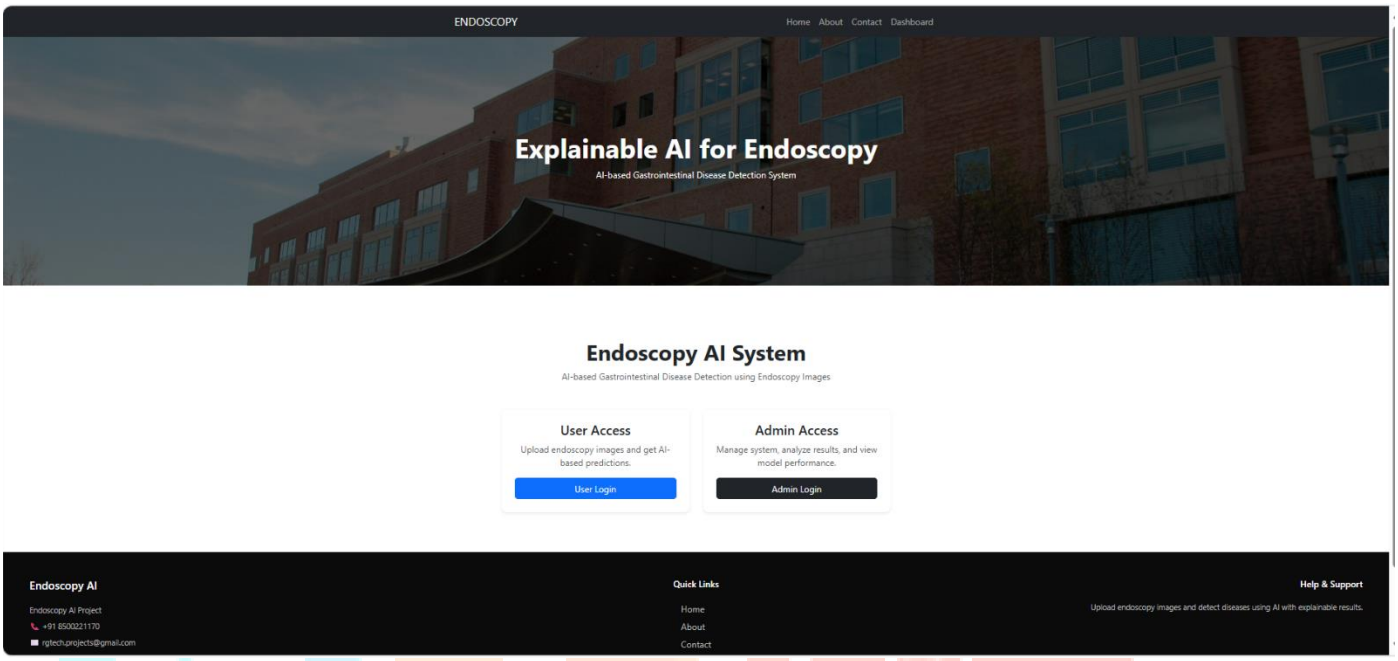

Figure 1.0: Explainable Ai-Based Classification Of Capsule Endoscopy Images Home Page

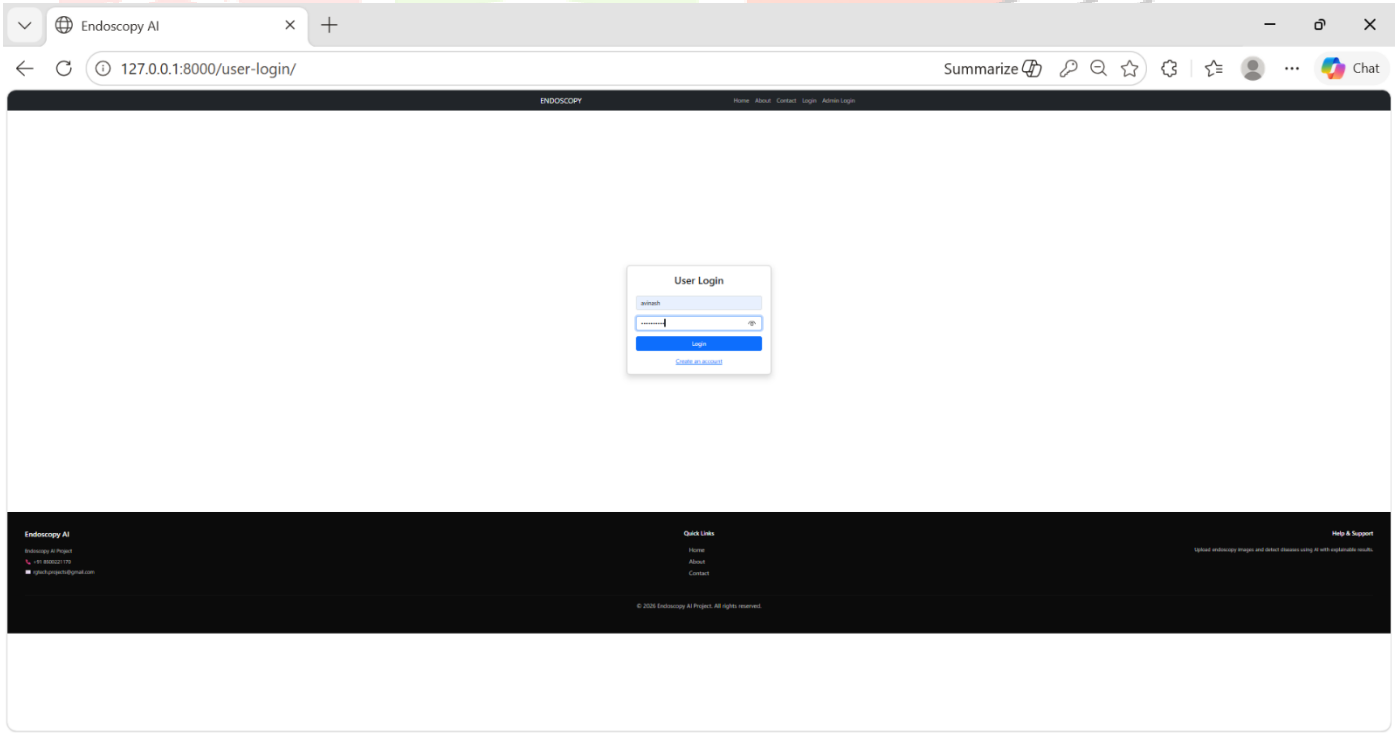

Figure 1.1: User Login Page

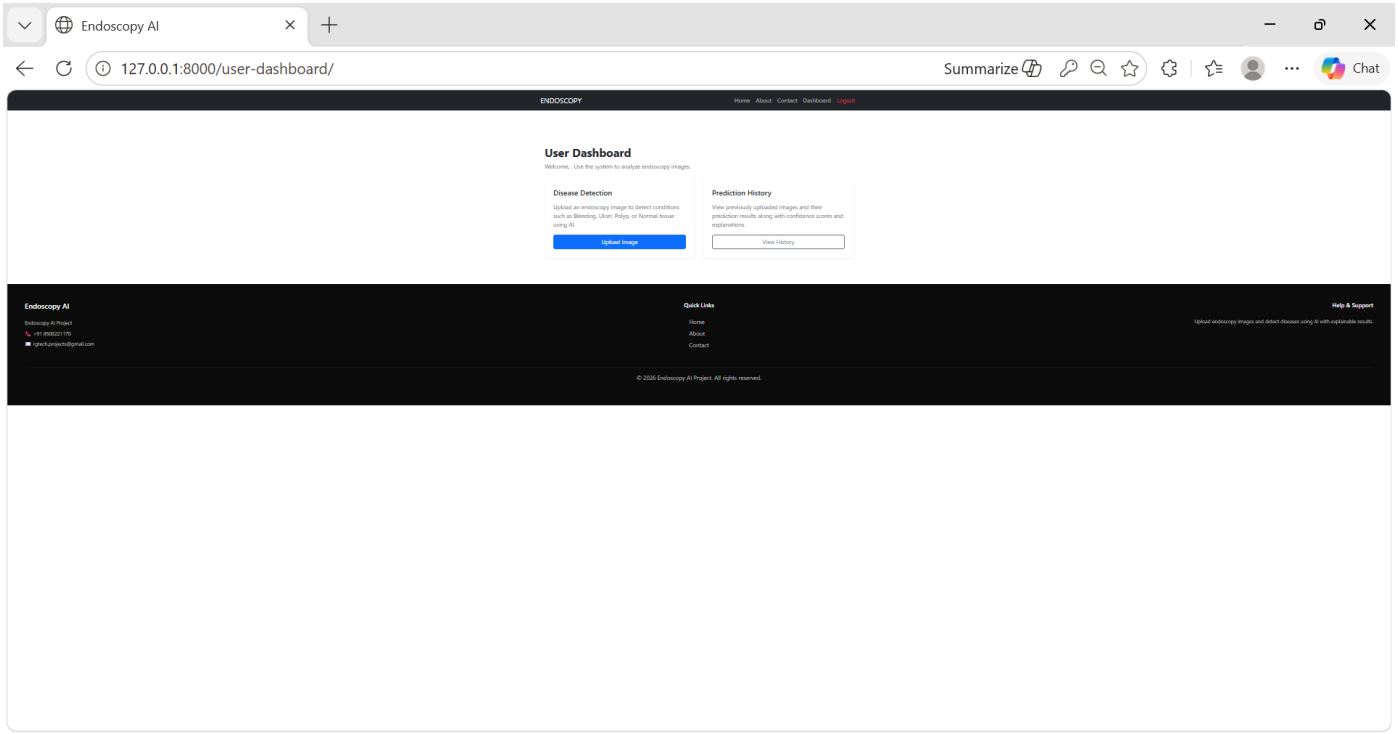

Figure 1.2: User Dashboard

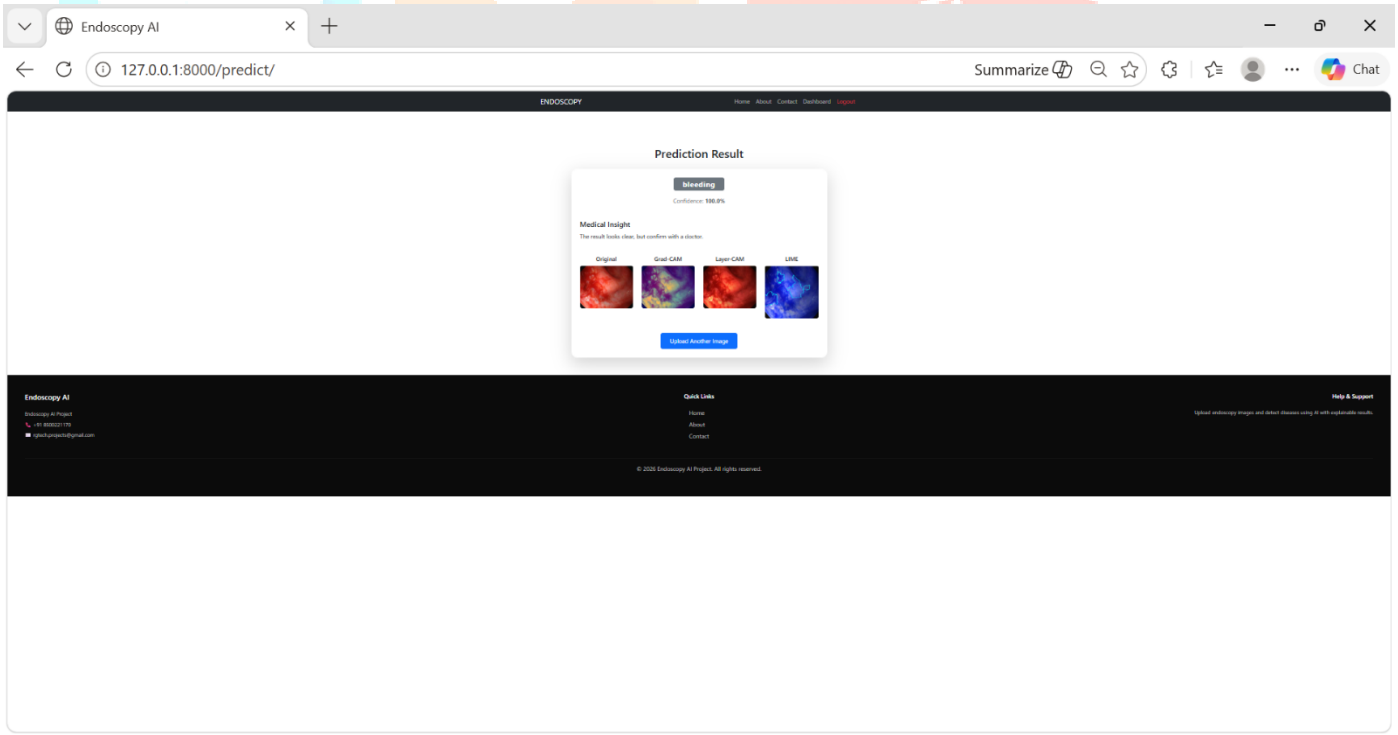

Figure 1.3: User Prediction Page

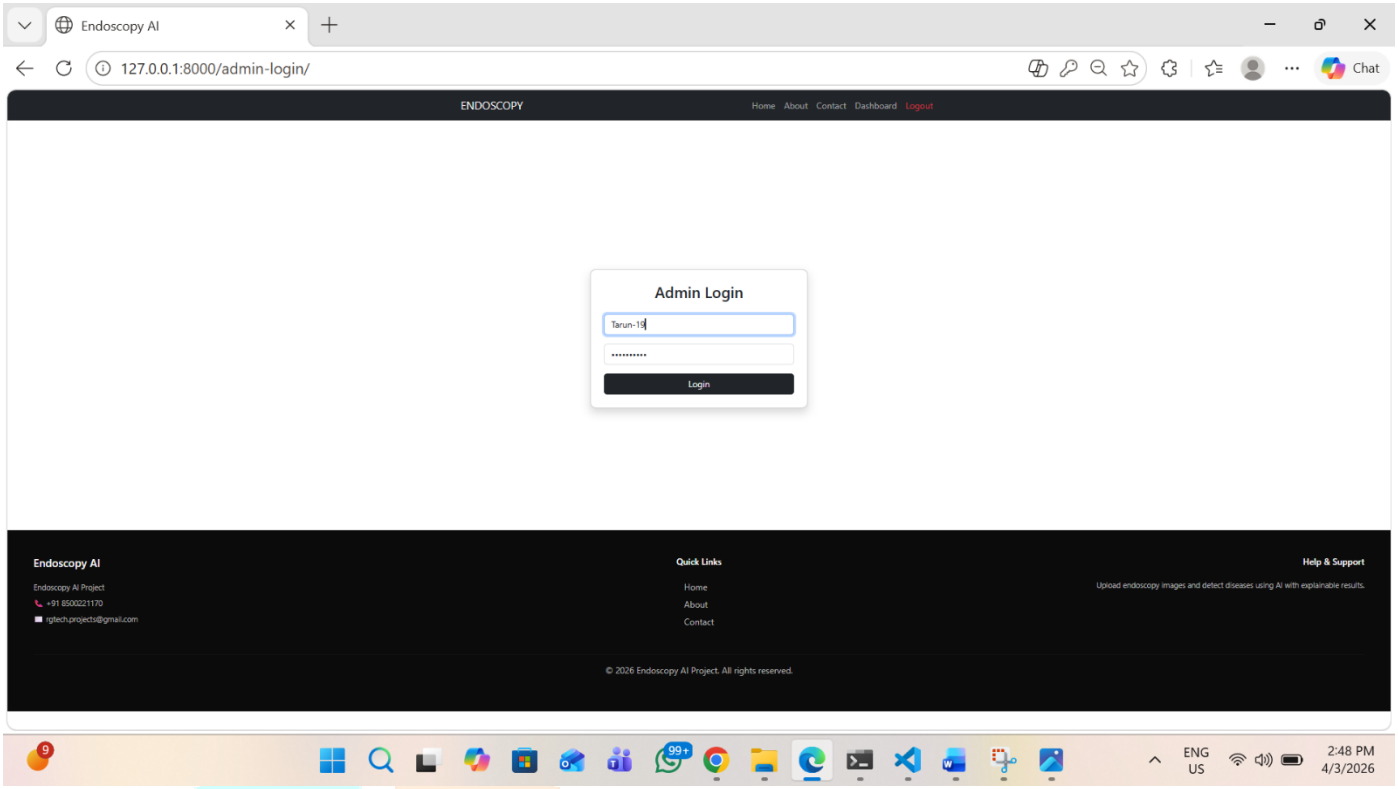

Figure 1.4: Admin Login Page

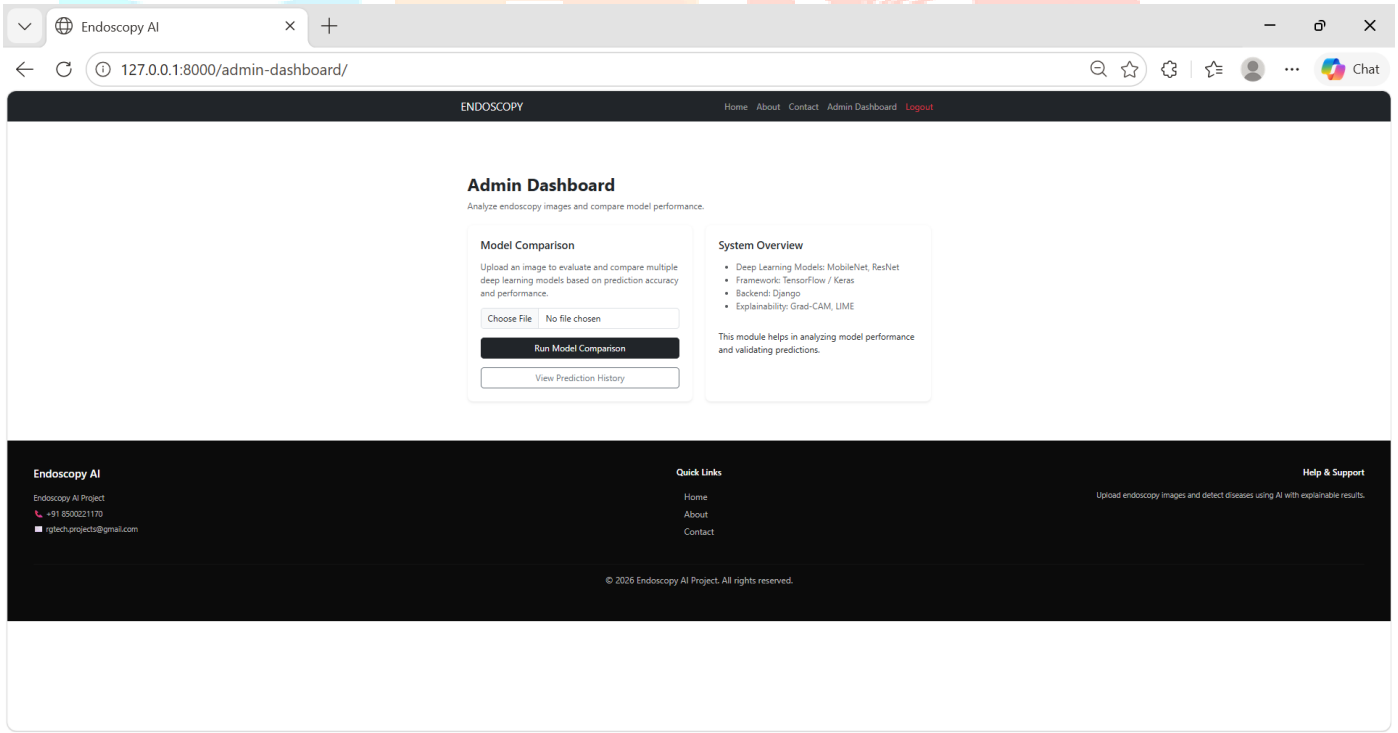

Figure 1.5: Admin Dashboard

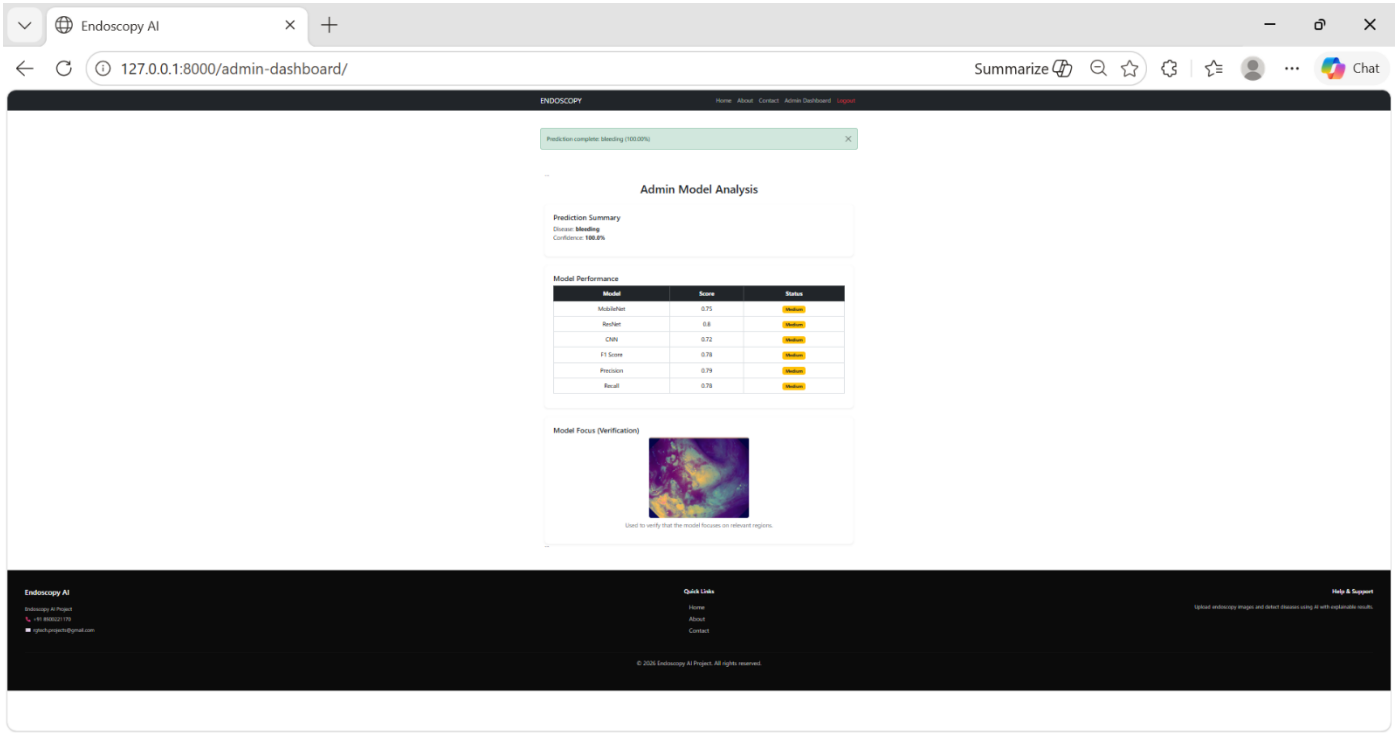

Figure 1.6: Admin Prediction Result Page

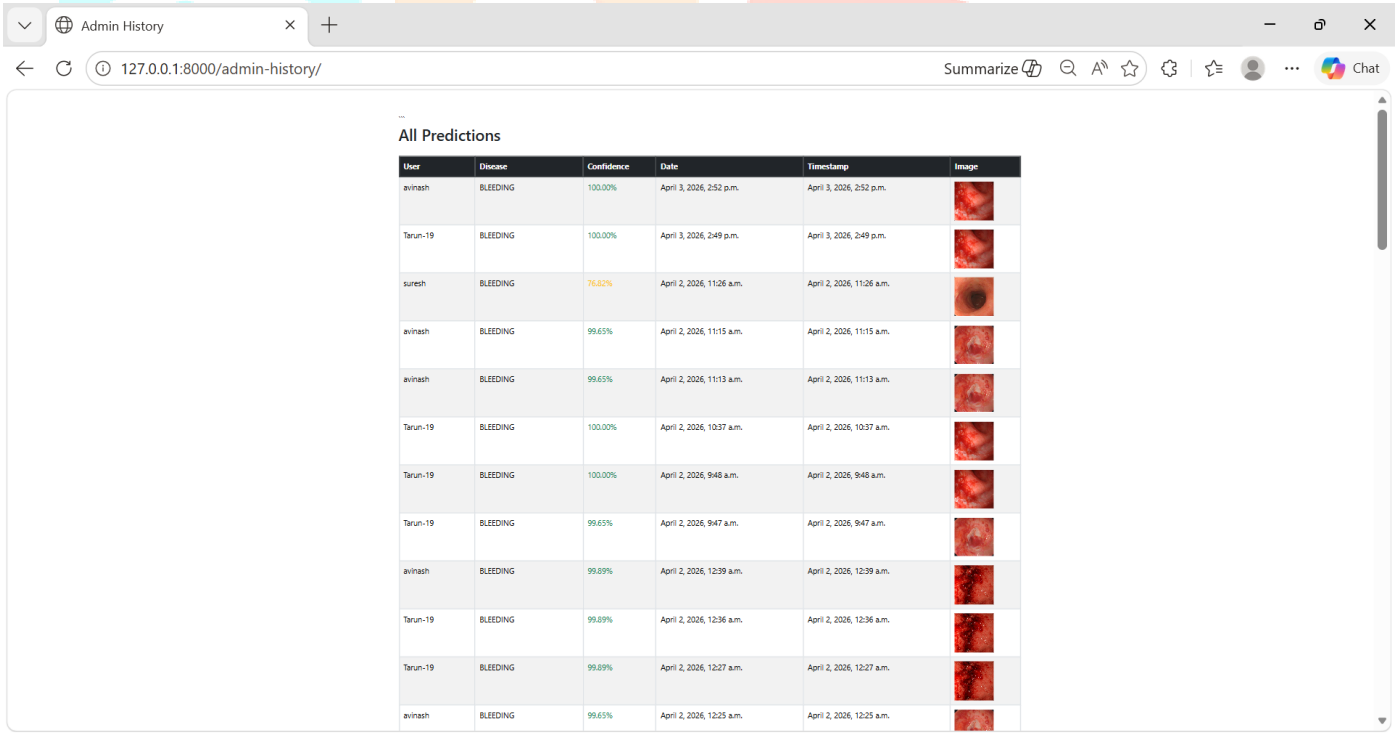

Figure 1.7: Prediction History Page

V. CONCLUSION

This project presents an Endoscopy AI System designed to assist in the detection and classification of gastrointestinal diseases using deep learning techniques. The system utilizes a Convolutional Neural Network (CNN) based on the MobileNetV2 architecture to classify endoscopy images into four categories: Normal, Bleeding, Ulcer, and Polyp.

The integration of Explainable AI (XAI) techniques such as Grad-CAM, Layer-CAM, and LIME enhances the transparency of the model by providing visual explanations for predictions. This helps users understand the decision-making process and improves trust in the system.

The system is implemented as a web-based application using the Django framework, enabling real-time image upload, prediction, and result visualization. The inclusion of prediction history and an admin dashboard further improves usability and system monitoring.

The results demonstrate that the system achieves high accuracy and fast response time, making it suitable for assisting medical professionals in early diagnosis and reducing manual effort.

Future work may include expanding the dataset, adding more disease categories, improving model accuracy, and integrating real-time video analysis for advanced medical applications.

VI. LITERATURE REVIEW

Recent research has shown significant advancements in the use of deep learning for medical image analysis, particularly in gastrointestinal disease detection using endoscopy images.

Several studies have utilized Convolutional Neural Networks (CNNs) for automated classification of endoscopy images. These models have demonstrated high accuracy in identifying abnormalities such as bleeding, ulcers, and polyps. Transfer learning models like ResNet, MobileNet, and Vision Transformers have been widely adopted to improve performance.

Researchers have also focused on improving the interpretability of deep learning models through Explainable AI (XAI) techniques. Methods such as Grad-CAM, Grad-CAM++, Layer-CAM, and LIME provide visual explanations by highlighting important regions in medical images, helping clinicians understand model predictions.

Datasets such as Kvasir and Kvasir-Capsule have played a crucial role in training and evaluating these models. These datasets contain labeled endoscopy images representing various gastrointestinal conditions.

Despite the progress, challenges remain in terms of dataset availability, model reliability, and interpretability. The integration of deep learning with Explainable AI is considered a promising approach to improve trust and usability in real-world healthcare applications.

The proposed system builds upon these advancements by combining MobileNetV2-based classification with Explainable AI techniques and integrating them into a user-friendly web application.

ACKNOWLEDGMENT

We express our sincere gratitude to our project guide and faculty members for their continuous support, valuable guidance, and encouragement throughout the development of this project.

We would also like to thank our institution for providing the necessary resources and infrastructure to successfully complete this work. Special thanks to our friends and peers for their support and constructive feedback.

REFERENCES

[1] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *International Conference on Learning Representations (ICLR)*, 2015.

[2] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *MICCAI*, 2015.

[3] D. T. V. Nguyen et al., “Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection,” *ACM MMSys*, 2017.

[4] R. R. Selvaraju et al., “Grad-CAM: Visual Explanations from Deep Networks,” *ICCV*, 2017.

[5] M. T. Ribeiro et al., “Why Should I Trust You? Explaining the Predictions of Any Classifier,” *KDD*, 2016.

[6] A. Howard et al., “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” *CVPR*, 2018.