Intelligent Diagnosis and Classification of Gastrointestinal Ailments using Endoscopic Imaging and Deep Learning

Mr. Gopalakrishna C
Associate Professor
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

Kushal Jetty K V
Student
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

Koushik R
Student
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

N P Prasidh
Student
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

Abhinandan C
Student
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

Mr. Sangareddy B. Kurtakoti
Assistant Professor
Computer Science and Engineering
Adichunchanagiri Institute of Technology
Chikmagalur, Karnataka, India

Abstract—Gastrointestinal (GI) disorders pose significant health risks, potentially leading to severe complications such as colorectal cancer if not diagnosed early. Endoscopy is the primary diagnostic tool for examining the GI tract, but manual evaluation is time-consuming and prone to human error, increasing the risk of missed abnormalities. This study presents a deep learning-based system that utilizes a Convolutional Neural Network (CNN) with a pre-trained ResNet101V2 model to automatically detect and classify GI tract anomalies. The model is trained on the KVASIR-V2 dataset, comprising 8,000 annotated endoscopic images, achieving a classification accuracy of 99%. The results highlight the system's potential to enhance diagnostic accuracy and reduce reliance on manual interpretation, making it a valuable tool for computer-assisted detection of GI disorders.

Index Terms—Alimentary Tract, Endoscopy, Deep Learning, Convolutional Neural Networks (CNN), ResNet101V2.

I. INTRODUCTION

Gastrointestinal (GI) ailments affect millions globally, with GI issues causing over 3 million cases and 2 million deaths annually. Conditions like polyps and ulcerative colitis can lead to complications including cancer if not diagnosed early.

Endoscopy is a primary diagnostic tool using a camera-equipped tube to capture digestive tract images. Manual assessment is time-consuming and requires expertise, making AI-assisted analysis promising for automated anomaly detection.

The human GI system includes:

- Anatomical Markers: Normal cecum, pylorus, and z- line
- Pathological Findings: Polyps, esophagitis, and ulcera- tive colitis
- Polyp Removal Cases: Dyed lifted polyps and dyed resection margins

AI applications in medical imaging can enhance diagnostic accuracy and patient outcomes. Convolutional Neural Networks effectively classify medical images by extracting features from endoscopic visuals.

A. Contributions

Our contributions include:

1) A computer-assisted framework for detecting GI tract abnormalities
2) A deep learning approach using fine-tuned ResNet101V2 CNN
3) Performance evaluation using accuracy, recall, precision and F1-score metrics

B. Paper Organization

The remainder of this document is organized as follows: Section II presents a literature review of related studies. Section III elaborates on the proposed methodology, including dataset preprocessing and model architecture. Section IV discusses the experimental findings, performance assessment, and comparison with established approaches. Ultimately, Section V concludes the study and directions for future research.

II. LITERATURE SURVEY

In these years, deep learning has revolutionized computer-assisted diagnosis (CAD) for gastrointestinal (GI) diseases. Numerous researches have investigated different deep neural network designs deep learning techniques, feature extraction methods, combined strategies, multi-modal approaches, or model adaptation techniques.

Gaurish Anand and Dev Gupta [1] proposed a [ML]machine learning based classification model for GI diseases using the KVASIR-V2 dataset. They implemented a Voting Classifier approach and achieved 88.19% accuracy, reducing human error in disease detection. However, the method was limited in generalizability to other datasets.

Varalaxmi [2] introduced a modern CNN approach using ResNet50 and EfficientNetB7 for GI disease detection. Their method improved diagnostic speed and accuracy, achieving 88.05% accuracy. Despite its success, it required a large dataset, making it computationally expensive.

Iyer [3] built a deep learning model for predicting gastrointestinal diseases using GI endoscopic images. Their approach leveraged the KVASIR dataset, employing data augmentation, transfer learning, and parameter tuning to enhance classification performance. The model achieved 96.89% accuracy, demonstrating the efficacy of deep learning in real-time medical diagnostics. However, their method exhibited high computational costs and relied on the quality of pre-trained models for optimal performance.

Sharmila and Geetha [4] applied deep learning techniques for GI anomaly detection using endoscopic images. Their approach leveraged a CNN-based model with ResNet101, significantly enhancing diagnostic accuracy. Using the KVASIR dataset, they demonstrated an accuracy of 98.37%, showcasing the potential of deep learning for automated GI disease classification. However, their method relied on high-quality labeled data and required substantial computational resources.

Alruban and Alabdulkreem [5] utilized a nature-inspired optimization algorithm, integrating bilateral filtering, Enhanced ShuffleNet, and Spotted Hyena Optimizer with a Stacked LSTM network for GI disease classification. Their model achieved 98% accuracy, making it suitable for real-time applications. However, it remained computationally expensive and dataset-dependent.

Shahriar Hossain [6] proposed a complex deep learning model combining Swin Transformer and Xception with trans- fer learning. Their study reported 87.23% accuracy, highlight- ing the potential of Vision Transformers (ViTs) in medical imaging. However, their model required extensive computational resources, limiting its practical deployment.

Although these studies yield promising outcomes in detecting GI diseases, most methods encounter computational limitations, restricted dataset generalizability, and dependency on pre-trained models. The primary objective of this work is to enhance classification accuracy while reducing computational complexity through an optimized ResNet101V2 architecture.

A summary comparing current methods and their performance metrics is presented in Table I.

TABLE I
COMPARISON OF GI DISEASE CLASSIFICATION METHODS

Method	Authors	Model Used	Accuracy (%)
Voting Classifier (KVASIR-V2)	Gupta et al.	Machine Learning	88.19
CNN with ResNet50, EfficientNetB7	Varalaxmi et al.	Deep Learning	88.05
Transfer Learning with Expanded Dataset	Iyer et al.	Transfer Learning	96.89
GI Anomaly Detection using CNN	Sharmila, Geetha.	CNN-Based Model	95.75
Nature-Inspired Algorithm (ShuffleNet, LSTM)	Alruban et al.	Hybrid Model	98.00
Hybrid Model (Swin Transformer + Xception)	Hossain et al.	ViT-Based Model	87.23
Proposed Method	**Our Work**	**ResNet101V2**	**99**

III. PROPOSED METHOD

A. Dataset

The proposed model is trained on the KVASIR-V2 dataset, a publicly available gastrointestinal endoscopic image dataset. It contains a total of 8,000 annotated images spanning different GI tract abnormalities. The images were collected from real-world endoscopic procedures and manually labeled by expert gastroenterologists.

To enhance model generalization, we applied data augmentation techniques, including random rotation (0°–30°), horizontal and vertical flipping. These augmentations help mitigate overfitting and improve the model's robustness to variations in real-world endoscopic imaging.

- Anatomical Landmarks: Normal Z-line, Normal Pylorus, Normal Cecum.
- Pathological Findings: Esophagitis, Polyps, Ulcerative Colitis.
- Polyp Removal Cases: Dyed Lifted Polyps, Dyed Resection Margin.

B. Image Preprocessing

To ensure consistent input dimensions, images were resized to 224×224 pixels. Preprocessing steps included:

- Normalization: Pixel values were scaled between 0 and 1.
- Augmentation: Rotation, flipping, contrast adjustment to enhance model robustness.
- Data Validation: Removing noisy or corrupted images.

C. Proposed Model

A Convolutional Neural Network (CNN) utilizing the ResNet101V2 architecture served as the foundation for feature extraction and classification. The model pipeline includes:

- Feature Extraction: CNN layers extract hierarchical features from endoscopic images.
- Classification: Fully connected layers categorize images into eight distinct classifications.
- Evaluation: Performance was assessed based on accuracy, precision, recall, and F1-score.

The data flow diagram for classification is shown in Fig- ure 2.

D. CNN Architecture

The CNN model consists of:

- Input Layer: Receives resized 224×224 RGB images.
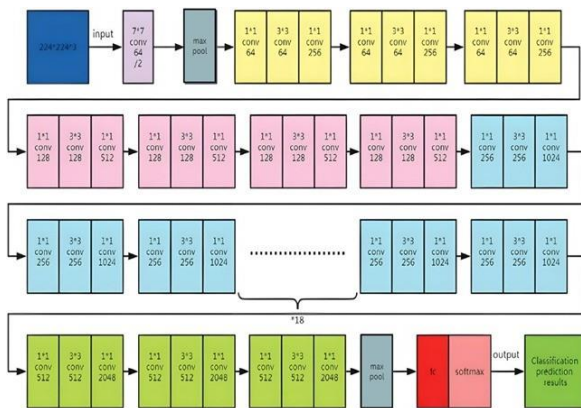- Convolutional Layers: Four Conv2D layers extract feature maps.

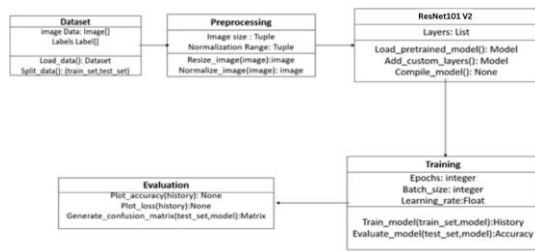Fig. 1. Detailed ResNet101V2 Model Architecture with Custom Classification Head.

Fig. 2. Simplified CNN Data Flow

Fig. 3. Dataset Images

- Pooling Layers: Max-Pooling (2×2) down-samples the features.
- Fully Connected Layer: A dense layer with 256 neurons for classification.
- Softmax Layer: Outputs the final probabilities across the eight classes.

E. Transfer Learning with ResNet101V2

To improve accuracy, ResNet101V2 was fine-tuned by:

- Replacing the final classification layer to suit eight-class classification.
- Freezing initial layers and training only the fully connected layers.
- Optimizing with Adam (learning rate: 0.001).

F. Hyperparameters and Training Setup

The model was trained with:

- Optimizer: Adam (learning rate = 0.001)
- Loss Function: Categorical Cross-Entropy
- Batch Size: 64
- Epochs: 100

Fig. 4. Loss of Training and Validation Over Epochs
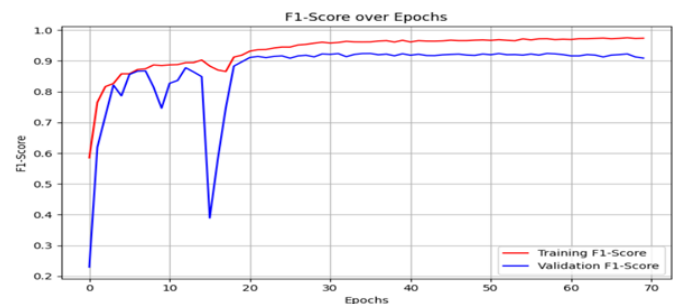
Fig. 5. F1-Score Trends for Different Classes

G. Classification Results

After fine-tuning, the model achieved a validation accuracy of 99%, significantly improving upon previous approaches. Figures 3 display the confusion matrix and sample predictions.
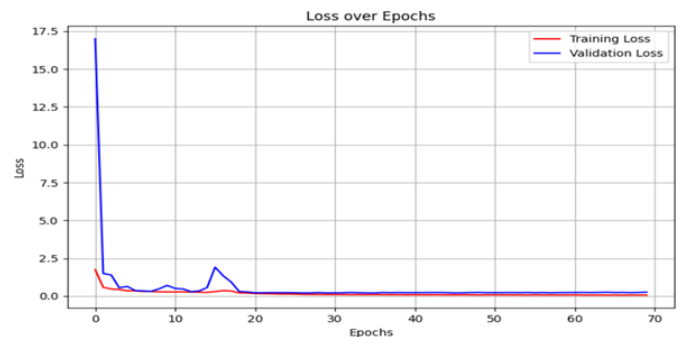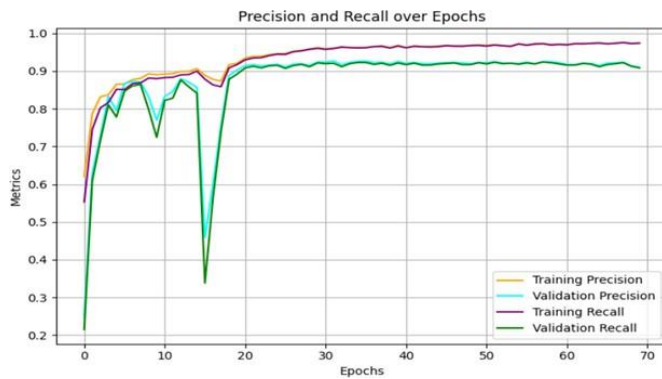
Fig. 6. Precision and Recall Trends for Different Classes

Method	Authors	Model Used
Voting Classifier (KVASIR-V2)	Gupta et al.	Machine Learnin
CNN with ResNet50, EfficientNetB7	Varalaxmi et al.	Deep Learning
Transfer Learning with Expanded Dataset	Iyer et al.	Transfer Learnin
GI Anomaly Detection using CNN	Sharmila, Geetha	CNN-Based Mod
Nature-Inspired Algorithm (ShuffleNet,LSTM)	Alruhan et al.	Hybrid Model
Hybrid Model (Swing Transformer + Xception)	Hossain et al.	ViT-Based Mode
Proposed Method	Our Work	ResNet101V2

TABLE II
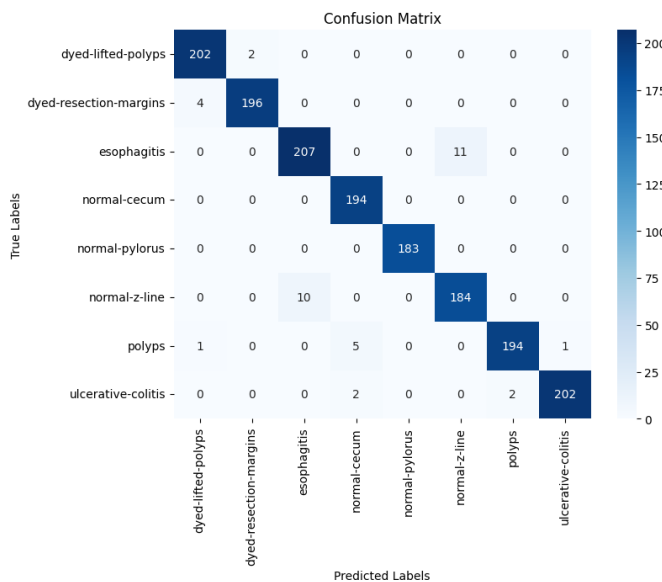COMPARISON BETWEEN THE PROPOSED METHOD AND
OTHER METHODS

Fig. 7. Confusion Matrix

IV. RESULTS AND DISCUSSION

A. Dataset

The KVASIR-V2 dataset is used in this study for training and evaluating the proposed gastrointestinal (GI) disease classification model. The dataset comprises 8000 endoscopic images, categorized into eight classes under three primary groups:

- Anatomical Landmarks: Normal z-line, normal pylorus, and normal cecum.
- Pathological Findings: Esophagitis, polyps, and ulcerative colitis.
- Polyp Removal Cases: Dyed Lifted polyps and Dyed Resection margin.

Each category features 1,000 images, with resolutions ranging from 720×576 to 1920×1072 pixels. Prior to training, the dataset was divided into 60% for training and 40% for validation, ensuring reliable model evaluation.

B. Experimental Setup

The proposed deep learning model was implemented using Python and Keras with TensorFlow as the backend. The following hardware and software specifications were used:

- Processor: Intel Core i5 (10th Gen) @ 3.0 GHz
- RAM: 8 GB
- GPU for Preprocessing: NVIDIA GeForce GPU
- Training Environment: Google Colab with NVIDIA TI-TAN RTX GPU (25GB RAM)

C. Performance Metrics

The model was assessed using accuracy, precision, recall, and F1-score on the validation dataset.

D. Comparison with Previous Works

The achieved accuracy of 99% rivals earlier studies in GI disease classification:

Table II juxtaposes the efficacy of the suggested model with established methods.

E. Evaluation Metrics

To assess the efficacy of the proposed model, we employed the following metrics:

Classification Accuracy: To assess the efficacy of the proposed model, we employed the following metrics:

$$Accuracy = \frac{C_{correct}}{C_{total}} \quad (1)$$

Positive Predictive Value (Precision): Indicates the fraction of relevant positive predictions out of all predicted positives.

$$Precision = \frac{C_{true\text{-}positive}}{C_{true\text{-}positive} + C_{false\text{-}positive}} \quad (2)$$

Recall: Represents the fraction of actual positive cases that were correctly classified.

$$Recall = \frac{C_{true\text{-}positive}}{C_{true\text{-}positive} + C_{false\text{-}negative}} \quad (3)$$

F1-Measure: A balanced score that considers both precision and recall, providing an overall measure of model perfor- mance.

$$F1\text{-}Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (4)$$

F. Training Performance Visualization

To analyze efficacy of the suggested model, the following graphs show the training and validation trends across multiple epochs.

G. Results

The data was partitioned with 60% allocated for training, 20% designated for validation and 20% evaluation. The pro- posed CNN model proficiently distinguished between actual and predicted images, As illustrated in Figure 4, the model demonstrated high accuracy in detecting polyps and normal cecum, while Z-lines and esophagitis were also accurately classified. Nonetheless, errors were evident in differentiating Dyed Resection margins and Dyed Lifted polyps, probably be- cause of their comparable characteristics. Achieving an overall accuracy rate of 99%, the findings from our suggested method exhibit better performance than earlier studies, as illustrated in Table 1.

H. Deployment Challenges

Despite high classification accuracy, real-world deployment poses challenges. Model interpretability remains a key con- cern, as deep learning models often function as black boxes. Additionally, integrating the model into pathology workflows requires usability enhancements, including intuitive UI design and explainability methods. Computational efficiency must also be considered to ensure smooth deployment in resource-constrained environments.

V. CONCLUSION

The proposed deep learning model effectively classified gastrointestinal (GI) diseases by utilizing endoscopic images, achieving exceptional classification accuracy. The fine-tuned ResNet101V2 architecture, combined with image preprocess- ing and augmentation methods, facilitated efficient feature extraction and enhanced classification performance. The model realized a validation accuracy of 99%, surpassing numerous existing approaches in GI disease detection. The utilization of the AdamW optimizer advanced model convergence, assuring improved generalization and stability.

Key findings of this research include:

- Strong classification of GI anomalies into eight cate- gories.
- Enhanced precision and recall for detecting patholog- ical conditions like esophagitis, polyps, and ulcerative colitis.
- Effective feature learning utilizing deep CNN archi- tectures, augmenting medical image-based diagnostics.

This study illustrates that deep learning can serve as a dependable tool for automated GI disease classification, po- tentially assisting medical professionals in expediting and enhancing the accuracy of Page 1 of 2 diagnoses. Future endeavors will concentrate on the real-time implementation of the model in clinical settings, as well as further enhancements utilizing larger datasets.

REFERENCES

[1] D. Gupta, G. Anand, P. Kirar, and P. Meel, "Classification of endoscopic images and identification of gastrointestinal diseases," in Proceedings of the 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON). IEEE, 2022, pp. 231–235.

[2] G. Varalaxmi, S. R. Baddam, E. S. Yalamarthi, K. Swaraja, K. R. Madhavi, and C. Sujatha, "Diagnosis of gastrointestinal diseases using modern cnn techniques," in 2023 IEEE 8th International Conference for Convergence in Technology (I2CT). IEEE, 2023, pp. 1–6.

[3] S. Iyer, D. Narmadha, G. N. Sundar, S. J. Priya, and K. M. Sagayam, "Deep learning model for disease prediction using gastrointestinal-endoscopic images," in Proceedings of the 2023 4th International Conference on Signal Processing and Communication (ICSPC). IEEE, 2023, pp. 133–137.

[4] V. Sharmila and S. Geetha, "Detection and classification of gi-tract anomalies from endoscopic images using deep learning," in Proceed- ings of the 2022 IEEE 19th India Council International Conference (INDICON). IEEE, 2022, pp. 1–6.

[5] A. Alruban, E. Alabdulkreem, M. M. Eltahir, A. R. Alharbi, I. Issaoui, and A. Sayed, "Endoscopic image analysis for gastrointestinal tract disease diagnosis using nature inspired algorithm with deep learning approach," in Proceedings of the IEEE Conference on Access Technolo- gies. IEEE, 2023, pp. 130 022–130 030.

[6] S. Hossain, M. Fahim-Ul-Islam, R. Rahman, and A. Chakrabarty, "Gastrointestinal insights redefined: An integrated hybrid model fusing vision transformer and transfer learning," in 2024 6th International Con- ference on Electrical Engineering and Information & Communication Technology (ICEEICT). IEEE, 2024, pp. 19–24.

[7] A. C. Mu¨ller and S. Guido, Introduction to Machine Learning with Python: A Guide for Data Scientists, 1st ed. O'Reilly Media, 2017.

[8] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.

[9] A. Ge´ron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd ed. O'Reilly Media, 2019.

[10] P. Documentation, torch.nn.Module, PyTorch, 2021. [Online]. Available: https://pytorch.org/docs/stable/generated/torch.nn.Module.html

[11] ——, torch.nn.ReLU, PyTorch. [Online]. Available: https://pytorch.org/docs/stable/generated/torch.nn.ReLU.html

[12] S. learn Documentation, confusion matrix, Scikit-learn. [Online]. Avail- able: https://scikit-learn.org/stable/modules/generated/sklearn.metrics. confusion matrix.html

[13] M. Documentation, matplotlib.pyplot, Matplotlib. [Online]. Available: https://matplotlib.org/stable/api/pyplot summary.html

[14] S. Documentation, seaborn.heatmap, Seaborn. [Online]. Available: https://seaborn.pydata.org/generated/seaborn.heatmap.html