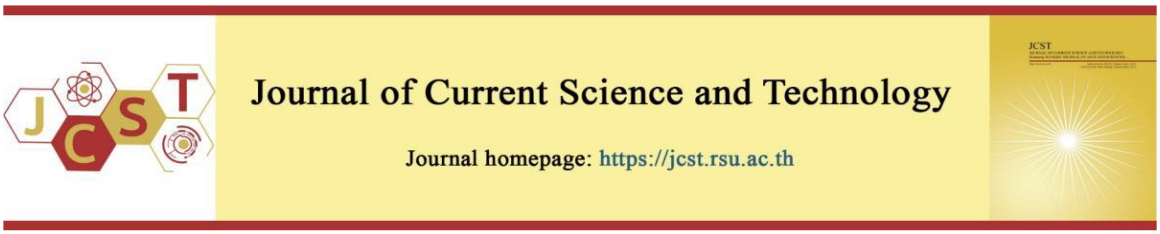
Journal of Current Science and Technology, July - September 2025 Vol. 15 No. 3, Article 126

ISSN 2630-0656 (Online)

Cite this article: Kumar, R., Singh, A., Khamparia, A. (2025). Visual feature refinement with MECNET for gastrointestinal cancer classification. *Journal of Current Science and Technology*, *15*(3), Article 126. https://doi.org/10.59796/jcst.V15N3.2025.126



Visual Feature Refinement with MECNET for Gastrointestinal Cancer Classification

Ravi Kumar[1,*], Amritpal Singh[1], and Aditya Khamparia[2]

[1]Department of Computer Science Engineering, Lovely Professional University, Punjab 144411, India
[2]Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Amethi 226025, India

[*]Corresponding author; E-mail: ravi.42100244@lpu.in

Received 28 March 2025; Revised 20 May 2025; Accepted 25 May 2025; Published online 15 June 2025

Abstract

Early detection and classification of gastrointestinal tract pathologies are crucial for better prognosis and reduced mortality rates, such as in colorectal cancer. In this paper, we introduce MECNET, a new hybrid deep learning framework for efficient classification of endoscopic images. The proposed framework integrates the feature refiner module with state-of-the-art CNN architectures such as VGG19, ResNet50, and EfficientNet for improved performance in image analysis and classification tasks. The feature refiner module successively applies grayscale, Gaussian, and LPQ filters to extract meaningful texture features, which play an important role in differentiating disease categories. Our proposed scheme has been tested on several available datasets, namely WCE, Kvasir, GastroVision, and SCPolyp including 13,000 images from four categories: normal colon, polyps, esophagus, and ulcerative conditions. The MECNET model attained an appreciable performance metric, outperforming state-of-the-art methods at accuracy and F1 scores of 97.4% and 97.34% on the WCE test set and 97.2% and 97.26% on the Kvasir test set, respectively. This proves that MECNET does not only excel in classification but also generalizes well across diverse datasets. The novelty of this work lies inincorporating a feature refiner module with established CNN architectures and utilizinga hybrid ensemble approach. This approach will provide a boost to the model's performance. The proposed framework addresses key challenges in medical image classification: improving feature extraction by making full use of advanced transfer learning techniques.

Keywords: *gastrointestinal diseases; colorectal cancer; CNN; multiclass classification; deep learning*

1. Introduction

The integration of artificial intelligence (AI) and related technologies such as machine learning, deep learning, etc., into healthcare has brought about a transformative shift in how medical data is interpreted by professionals (Bordbar et al., 2023). The accumulation of the data has enabled AI systems to identify patterns and insights that may not be obvious to human practitioners and, thereby improving diagnostic accuracy. Medical imaging is one such area that has advanced tremendously by the incorporation of AI, thereby improving disease diagnosis accuracy (Kumar et al., 2023). Colorectal cancer (CRC) is one of the most common gastrointestinal cancers worldwide (Sabah, & Hassan, 2024). It is the second most frequent cancer in womenand the third most common in males, with a very high mortality rate (Sung et al., 2021). CRC cases are expected to rise by 80% by 2035; hence, effective prevention, early detection, and treatment strategies are critical. Significant developments in medical image analysis for gastrointestinal pathology have been driven by ensemble deep-learning techniques. The AI-assisted system aims to ease the burden on endoscopists and clinicians, supporting them in making rapid and accurate diagnostic decisions. Deep learning, particularly convolutional neural networks (CNNs), has enhanced image-based

studies by providing tools for classification and detection tasks across various medical domains. Convolutional neural networks (CNNs) have emerged as strong image analysis tools, particularly for classification and computer vision tasks (Pogorelov et al., 2017; Kumari, & Wasim, 2023). Researchers have employed a variety of CNN architectures, including AlexNet, VGG16, DenseNet, ResNet, and EfficientNet, to tackle diverse image classification challenges, demonstrating their effectiveness across domains (Kumar et al., 2023). Wu et al., (2023) applied an ensemble of Inception-v3, ResNet101, and DenseNet201 on the CE-MRI benchmark dataset for multiclass brain tumor classification, incorporating genetic algorithms to optimize accuracy with reduced parameters. Sima, & Cincar (2021) achieved an accuracy of 71.33% using ShuffleNet and ResNet-50 on the Kvasir v2 dataset, while Islam et al., (2022) reported a notable accuracy of 93.22% for polyp classification using pre-trained CNNs with SVM. He et al., (2023) made significant advancements in early gastric cancer detection with the E-YOLO architecture, achieving an accuracy of 94.16%. Similarly, Srivastava et al., (2022) introduced FocalConvNet, achieving an F1 score of 67.34% in pathology detection, while Rani et al., (2022) used Inception ResNetV2 for bleeding detection with an accuracy of 95.62% on the WCE dataset. Yue et al., (2023) implemented a CNN on the Hyper-Kvasir model and achieved accuracy of 90.75% and F1 score of 65.41%. Bordbar et al., (2023) set a remarkable 97% accuracy benchmark on the WCE dataset using DenseNet. Park et al., (2023) implemented InceptionNet-V3 with Star-GAN (accuracy: 94.9%) further advancing the field. Mushtaq et al., (2023) and Jha et al., (2023) introduced innovative models and datasets, respectively, contributing to the growing body of knowledge in gastrointestinal image classification. Chae, & Cho (2023) proposed transformer based model for polyp classification. They and implemented vision transformer for this task. Their model achieved F1 Score of 87 for early intestine cancer. The trend continued, with Navale et al., (2024), achieving an accuracy of 94.93% using a diverse set of models. Kumar et al., (2024) proposed a deep-learning model for CRC classification with a genetic algorithm for weight optimization. The main limitation of the model was its increased complexity. Kaur, & Kumar, (2024) achieved 94% accuracy with VGG16 for image classification.Despite the extraordinary achievements of these deep learning methods, there remains room for improvement.

The research gaps in existing literature include a scarcity of public datasets with good-quality images, challenges in image processing, and lack of uniform standards for image collection, all of which have limited the efficiency of deep learning techniques.

2. Objectives

The specific objectives of this study are:

1. To enhance image preprocessing by implementing text removal and applying a Feature Refiner Module (FRM) using grayscale, Gaussian, and LPQ filters to improve the quality of diagnostic features in endoscopic images.

2. To design and train a Modified Ensemble Convolutional Neural Network (MECNET) that integrates pre-trained CNN architectures (EfficientNet and ResNet50) with the FRM, enabling high-precision multiclass classification of gastrointestinal images.

3. To evaluate the performance of MECNET on multiple publicly available datasets (WCE, Kvasir, GastroVision, SCPolyp) and assess its generalizability across different GI disease categories.

4. To compare MECNET's performance with existing state-of-the-art deep learning models, demonstrating improvements in accuracy, precision, recall, and F1 score.

3. Material and Methods

This study presents a comprehensive methodology for the classification of gastrointestinal (GI) tract conditions using a deep learning framework, MECNET. The methodology encompasses five key components: dataset acquisition, image pre-processing, feature refinement, model development, and performance evaluation.

We utilized multiple publicly available endoscopic image datasets representing a diverse range of GI conditions, including normal colon, polyps, esophageal disorders, and ulcerative colitis. Image preprocessing stepssuch as text removal, noise reduction, resizing, and normalization were applied to enhance image quality and ensure consistency across datasets. A Feature Refiner Module (FRM) was introduced to improve feature extraction by applying grayscale transformation, Gaussian filtering, and Local Phase Quantization (LPQ), targeting the enhancement of texture and spatial details critical for accurate classification.

For model development, we implemented and fine-tuned an ensemble of pre-trained convolutional

neural networks (CNNs), specifically ResNet50 and EfficientNet, integrated with the FRM to form the proposed MECNET architecture. Transfer learning and hyperparameter optimization were employed to improve generalization and performance. The models were trained and evaluated using standard metrics, accuracy, precision, recall, and F1 score, on both seen and unseen datasets to ensure robustness and reliability. The following subsections detail each methodological component in depth.

3.1 Dataset

Accurate detection and classification of colorectal cancer (CRC) using deep learning models heavily depend on the availability of high-quality, diverse medical imaging datasets. However, acquiring and maintaining such datasets is inherently challenging due to privacy concerns, annotation requirements, and variability in imaging equipment. To address these limitations and ensure robust model training, this study employed four established and diverse endoscopic image datasets: WCE, Kvasir, GastroVision, and SCPolyp, supplemented by additional anonymized clinical images obtained from collaborating medical institutions.

These datasets encompass a wide range of gastrointestinal tract regions, including the colon, sigmoid, rectum, and esophagus. Prior to analysis, all images underwent preprocessing to remove noise and improve quality using standardized enhancement techniques. The images were then organized into four diagnostic categories to support multiclass classification:

- Normal Cases: Images of healthy colonic mucosa, showing smooth surfaces and intact vascular patterns (Figure 1a). These serve as baseline controls within the dataset.
- Polyp Cases: Images illustrating abnormal polypoid growths in the colon or rectum, which carry malignant potential if left untreated (Figure 1b).
- Esophagus Cases: Images of the esophagus displaying signs of conditions such as esophagitis, characterized by mucosal inflammation and structural disruption (Figure 1c).
- Ulcerative Cases: Images depicting ulcerative colitis, an inflammatory bowel disease affecting the colon and rectum, typically presenting with mucosal bleeding and granularity (Figure 1d).

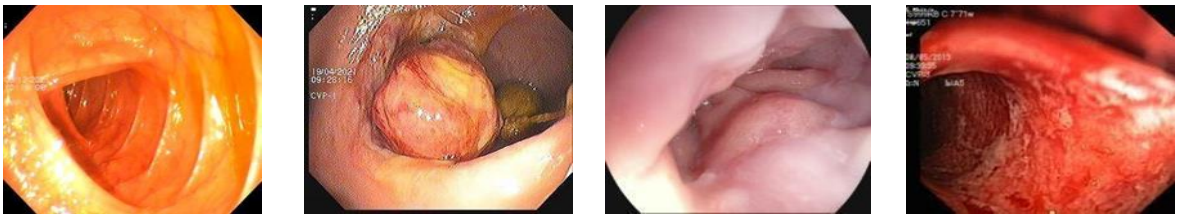
This curated dataset, representing four clinically relevant GI categories, serves as a foundational resource for training and evaluating the proposed deep learning model, MECNET. Its diversity and structure enable a comprehensive assessment of the model's classification performance across varying pathological conditions.

The dataset utilized in this study comprises a comprehensive collection of endoscopic images categorized into four distinct classes. The WCE dataset contributes 1,500 images per category. The Kvasir dataset provides 1,000 images per category. The SCPolyp dataset includes 500 images per category. The GastroVision dataset offers 150 images per category. In total, the dataset contains 13,000 images, representing a diverse range of gastrointestinal conditions. For model evaluation, we implemented and tested the frameworks both with and without the novel Feature Refiner module. This rigorous evaluation process enables a thorough assessment of model performance and the impact of the Feature Refiner on classification accuracy and predictive capabilities.

3.2 Image Preprocessing

Text Removal from Images:

Medical images frequently contain text annotations or labels that can obscure areas of interest, potentially impairing the accuracy of model training and prediction. To address this issue, it is essential to remove any text present in the images. We employed a standard watermark removal technique implemented in Python to achieve this. After text removal, the images were resized and normalized to ensure consistency and optimal quality. The dataset was then partitioned into training, validation, and test sets to facilitate rigorous model evaluation and performance assessment.



(a) (b) (c) (d)

Figure 1 Representative images from the dataset used for classification
(a) Normal colon image with smooth mucosa and visible vasculature, (b) Colon image showing a polyp,
(c) Esophageal image depicting erosive esophagitis,
(d) Ulcerative colitis image with bleeding, granular mucosa and ulcerations

Feature Refinement

The classification of gastrointestinal diseases from endoscopic images follows a systematically structured pipeline consisting of five critical stages, each designed to optimize model performance and ensure robust, accurate classification. The datasetfor this study was sourced from publicly available repositories, covering four diagnostic categories: normal colon, polyps, esophagus, and ulcerative conditions. The images were carefully curatedto eliminate low-quality samples, followed by preprocessing steps, including resizing, intensity normalization, and data augmentation, to standardize the dataset. For multi-scale feature extraction, a custom Feature Refiner Module (FRM) was used, involving Local Phase Quantization (LPQ) and Gaussian filtering.

Model Training

CNN architectures (VGG19, ResNet50, EfficientNet), with transfer learning were fine-tuned for gastrointestinal image classification, and hyper-parameter tuning was employed to optimize model performance. Finally, the final model was evaluated using accuracy, precision, and F1-score, and Grad-CAM visualizations were employed to enhance interpretability for clinical use.

3.3 Implementation Details

A feature refiner module was integrated with modified ResNet50 and EfficientNet models, both pre-trained on ImageNet dataset, to build the enhanced ensemble convolutional neural network model MECNET. By adapting the top layers of the pre-trained models, we developed this model by adding additional convolutional layers, dense layers, a softmax layer, and output layers. Several hyperparameters were introduced to improve model robustness, including Gaussian noise, dropout rates, and L1 regularization. To optimize performance, these hyper-parameters were fine-tuned. Ensemble models were constructed using the most effective selected models. Additionally, the ensemble framework was further refined to enhance accuracy. Two types of images were input into the ensemble model: images processed through the feature refiner module and original images from the dataset. This approach significantly improved the model's overall performance. In the results section, detailed performance metrics and results are provided. This module is designed to extract features in both spatial and frequency domains using a specialized filter. Before being input into the neural network, the images were processed through this module to enhance feature representation. The module employs a combination of grayscale, Gaussian, and Local Pattern Quantization (LPQ) filters to extract salient features from the input images. The mathematical formulation of the grayscale filter is provided in Equation 1.

$$\mathbf{G(i, j \mid d, \theta)} = \sum_{m=1}^{M} * \sum_{n=1}^{N} * \{1, \text{ if } I(m,n)=i \text{ and } I(m+\Delta x, n+\Delta y)= j, \quad (1)$$
$$0, \text{ otherwise}$$

In this formulation, III denotes the input image, while MMM and NNN represent its height and width, respectively. The terms Δx\Delta xΔx and Δy\Delta yΔy correspond to displacements along the x-axis and y-axis, parameterized by a specified distance and orientation.

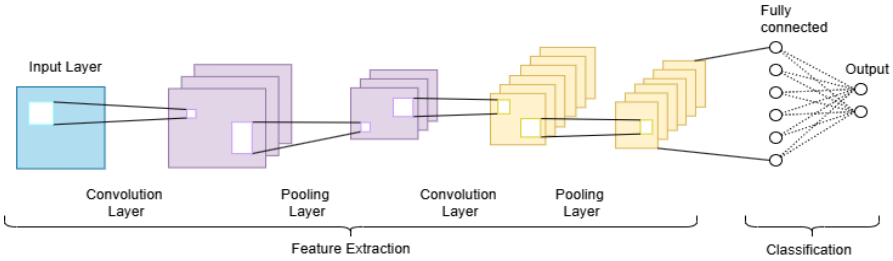


Figure 2 CNN architecture for gastrointestinal image classification. The architecture includes convolutional layers with ReLU activation, followed by pooling layers, flattening, and fully connected layers, culminating in a softmax output layer for multiclass prediction

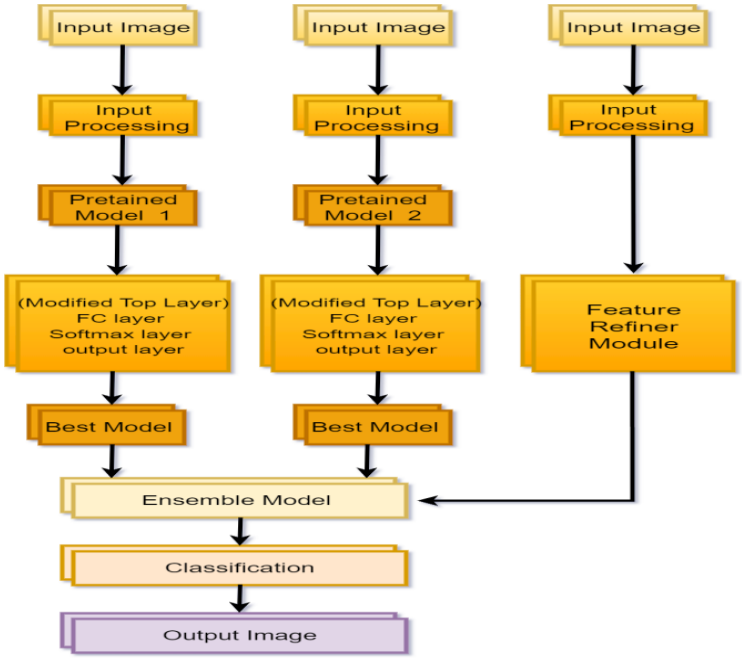


Figure 3 Methodological workflow (MECNET):Two best pre-trained deep learning models with modified top layers, combined with Feature refined module (FR) for optimal feature extraction and classification

The Local Phase Quantization (LPQ) algorithm functions as a texture descriptor by capturing the local phase information of an image. LPQ is calculated within a pixel's immediate neighborhood as given in equation 2, and the resulting patterns are quantized into binary format. The Local Binary Pattern (LBP) operator is then applied to these patterns to generate the final LPQ code. Furthermore, Gabor filters are employed in image processing to analyze textures across multiple scales and orientations, as defined in Equation 3. They are especially adept at capturing textural features, like edges, lines, and patterns within an image. The Feature Refiner module was applied separately to each of the three-color channels of the image to extract texture features. These processed images were then input into the ensemble model. All models were implemented using Python on Google Colab. The models selected for this study include VGG19 (Simonyan, & Zisserman, 2014), ResNet50 (He et al., 2016), EfficientNet (Tan, & Le, 2019), an ensemble model (EM), and Modified Ensemble Convolutional Neural Network (MECNET). Algorithms 1 and 2 present the implementation process of proposed method.

$$LPQ(x,y) = \mathrm{SN}\,(I(x+1,y) - I(x,y))\ \Theta\ \mathrm{SN}\,(I(x,y-1) - I(x,y)) \quad (2)$$

I(x, y) specify the pixel intensity at coordinates (x,y) in the grayscale image. SN is the sign function that yields -1 for negative values, 0 for zero, and 1 for positive values, while **Θ** signifies the bitwise XOR operation.

$$G(x, y \mid f,\theta)=\exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\emptyset\right) \quad (3)$$

Where: x'= x cos(θ) + y sin(θ), y'= -x sin(θ) + y cos (θ), γ= aspect ratio, σ= Standard deviation of Gaussian envelope,λ= wavelength of sinusoidal factor,∅= phase offset**.**

Algorithm 1 (Image pre-processing)
Input Colon and Rectal Images:
Normal Raw Medical Images
Output:
Dataset with Preprocessed Image Details

Begin
1. Load colon and rectal images:
$I = \{I_0, I_1, I_2, ..., I_n\}$ // List of image instances in the initial dataset
2. Apply Text Removal Method to remove any text:
For each image in the folder:
Remove text or watermark using image processing techniques (Thresholding & inpainting.)
3. Resize all images to a standard size:
Resize all images to either 224x224 or 256x256
4. Add all pre-processed images to the final dataset:
Add all processed images (normal and upgraded) to the final dataset for further model training or analysisreturn dataset
End Algorithm

Algorithm 2 (Image, Classification)
Input:
Colon and Rectal Images
Output:
Predicted Class: Cancerous, Non-Cancerous, Polyps, Normal, Ulcerative, Esophagus

Begin
1. Extract features using the Feature refiner module.
2. Apply and evaluate deep learning models to Train dataset:
For each model in Model List (e.g., ResNet50, EfficientNetB0 etc.):
Train model using Train dataset
Evaluate model on Validation dataset to measure performance
3. Select the best-performing model:
Choose the model with the highest performance on the Validation dataset
4. The features extracted from Deep Learning models and a Feature refiner module are combined and passed to the proposed model (MECNET).
5. MECNET was evaluated on validation set.
6. Test the optimal (MECNET) model:
Test the selected model on Test dataset to evaluate final accuracy and performance metrics
7. Output predicted class for each test image:
For each test image in TestSet:
Apply the optimal model
Return predicted classes (Normal, Polyps, Ulcerative, Esophagus)
End Algorithm

Transfer learning was employed across all models using weights pre-trained on ImageNet. The ensemble model was constructed by integrating the two best-performing models: EfficientNet and ResNet50. The best results were obtained with hyper-parameters setting such as batch size of 32, 100 epochs, a dropout rate of 0.2, and L2 regularization, and the Adam optimizer as detailed in Table 1. Models were fine-tuned by adjusting hyperparameters to achieve optimal performance.

Table 1 Selected hyperparameters for training MECNET and baseline models

Sr No.	Parameter	Value
1	Batch Size	**32**/64/128
2	Epochs	30 -100
3	dropout	0.1-0.5
4	Regularization	L1/**L2**
5	Optimizer	**Adam**/SGD
6	learning rate	0.001 -0.01
7	Loss function	Categorical sparse Entropy

3.4 Model Evaluation

Binary and multiclass classifications were performed and the model evaluation was done based on accuracy, precision, recall, and F1 score. The calculations for accuracy, precision, recall, and F1 score were performed according to the formulas provided in Equations 4, 5, 6, and 7 respectively.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (4)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (5)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (6)$$

$$\text{F1 score} = \frac{(2* \text{Precision}* \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (7)$$

4. Results and Discussion

The selected models were implemented with the same configuration on chosen datasets and the results were tabulated and discussed in this section.

4.1 Model Performance on Datasets

The performance of the models, based on the chosen parameters across multiple datasets, is displayed in both tabular and graphical formats. Table 2 outlines the results for the training and validation datasets.

Our proposed model achieved an accuracy of 99.6% (95% CI: 0.9947-0.9973) on the training set and 99.37% (95% CI: 0.9912-0.9962) on the validation set, outperforming all other examined models. The results for the WCE and Kvasir test sets are presented in Table 3. The observed performance aligns with the training and validation datasets, demonstrating the model's effective generalization to unseen data. The MECNET model outperforms all other models in terms of accuracy, precision, recall, and F1 score.

The models were also tested on data from various medical sources, and results were evaluated. Model Evaluation on the GastroVision dataset is presented in Table 4.

Our proposed model, MECNET, outperformed the other models, achieving an accuracy of 97.3% (95% CI: 0.9659–0.9801) and an F1 score of 97.2% (95% CI: 0.9648–0.9792) on SCPolyp dataset. Detailed results are presented in Table 5.

Table 2 Performance of individual and ensemble models on training and validation datasets

Models	Accuracy		Precision		Recall		F1	
	Train	Val	Train	Val	Train	Val	Train	Val
Efficient Net	0.989	0.983	0.985	0.983	0.985	0.981	0.985	0.982
ResNet50	0.994	0.982	0.992	0.982	0.991	0.981	0.992	0.982
VGG19	0.961	0.947	0.958	0.946	0.948	0.946	0.953	0.946
OEM	0.995	0.973	0.994	0.972	0.993	0.972	0.973	0.972
MECNET	***0.996***	***0.993***	***0.995***	***0.992***	***0.994***	***0.992***	***0.994***	***0.992***

Table 3 Test set performance of models on WCE and Kvasir datasets

Models	Accuracy		Precision		Recall		F1	
	WCE	Kvasir	WCE	Kvasir	WCE	Kvasir	WCE	Kvasir
EfficientNet	0.9681	0.967	0.654	0.965	0.962	0.960	0.963	0.962
ResNet50	0.951	0.95	0.958	0.958	0.957	0.951	0.958	0.954
VGG19	0.945	0.932	0.941	0.930	0.940	0.942	0.940	0.936
OEM	0.96	0.9603	0.962	0.962	0.961	0.959	0.961	0.960
MECNET	***0.974***	***0.972***	***0.974***	***0.972***	***0.973***	***0.972***	***0.973***	***0.972***

Table 4 Classification results of models on the GastroVision dataset

Model	Accuracy	Precision	Recall	F1
EfficientNet	96.3	96.1	96	96.1
ResNet50	94.9	95.6	95.3	95.5
VGG19	93.4	93.1	93.9	93.7
OEM	95.8	96	96.1	95.9
MECNET	**97.5**	**97.3**	**97.4**	**97.3**

Table 5 Performance of models on the unseen SCPolyp dataset

Model	Accuracy	Precision	Recall	F1
Efficient Net	0.963	0.962	0.963	0.963
ResNet50	0.961	0.962	0.961	0.961
VGG19	0.931	0.931	0.932	0.931
OEM	0.965	0.965	0.964	0.964
MECNET	***0.973***	***0.972***	***0.973***	***0.972***

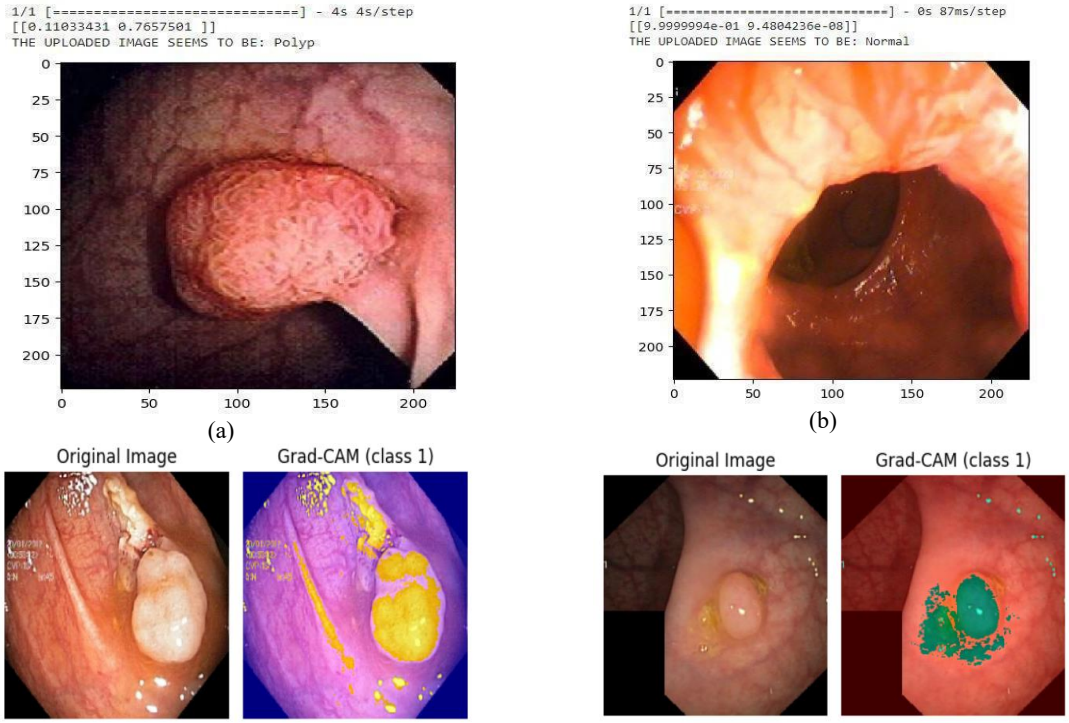


Figure 4 Prediction and visualization results from the MECNET model. (a) Predicted polyp with 96% confidence, (b) Predicted normal colon with 99% confidence, (c) Grad-CAM heatmap using AUTUMN colormap, (d) Grad-CAM heatmap using JET colormap to localize pathology

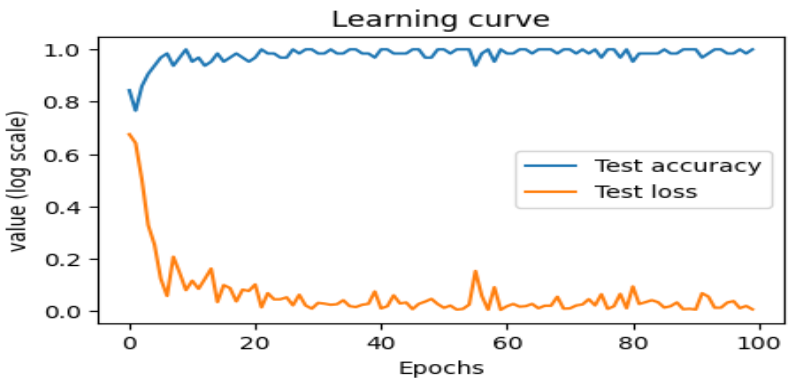


Figure 5 Learning curve of MECNET on the test dataset

After testing, several sample images were individually passed to the model, which correctly predicted most of them. The predicted output of the proposed model is shown in Figure 4.

4.2 Discussion

The proposed MECNET model demonstrates high efficacy in the classification of gastrointestinal (GI) tract conditions, particularly for early colorectal cancer detection. This performance is driven by the integration of a Feature Refiner Module (FRM), which enhances image texture features through grayscale, Gaussian, and Local Phase Quantization (LPQ) filtering. These preprocessing steps significantly improve the discriminative power of extracted features, enabling better differentiation between subtle tissue variations in normal, polyp, ulcerative, and esophageal images.

A key advantage of MECNET lies in its ensemble architecture, which fuses the strengths of EfficientNet and ResNet50. By combining pre-trained convolutional layers and customizing the upper layers with fine-tuned hyperparameters and dropout, the model achieves excellent generalization on both seen and unseen datasets. Notably, MECNET achieved an F1 score of 97.34% on WCE, 97.26% on Kvasir, and 97.2% on SCPolyp, outperforming each of its individual backbone models by 2–4%. This confirms that the ensemble not only stabilizes predictions but also compensates for the limitations of each individual architecture.

The inclusion of text removal and intensity normalization in the image preprocessing pipeline also contributes to increased accuracy by reducing background noise and standardizing input quality. These steps ensured that critical diagnostic features were preserved while minimizing irrelevant artifacts that might confuse the model.

Compared to prior works, MECNET consistently shows stronger performance. For instance, Park et al., (2023) reported 94.9% accuracy using InceptionNet-V3 on Kvasir, while Bordbar et al., (2023) achieved 97% using DenseNet on WCE. Our model surpasses both benchmarks, confirming that the hybrid design of MECNET offers a tangible improvement in clinical decision support. While these studies used either standalone architectures or simple ensembles, MECNET's deep integration of refined feature inputs and architectural diversity plays a decisive role in its superior accuracy.

Despite these promising outcomes, MECNET is not without limitations. The combined use of ResNet50 and EfficientNet increases model complexity, resulting in higher computational costs during training. For example, MECNET took approximately 490 seconds per epoch on the Kvasir dataset, compared to 359 and 455 seconds for EfficientNet and ResNet50, respectively. Moreover, the trainable parameters in MECNET significantly exceed those in its base models. This complexity could pose barriers to deployment in resource-constrained environments.

To mitigate these limitations, future work should explore lightweight alternatives, such as knowledge distillation or pruning techniques, which could reduce model size while maintaining high performance. Additionally, although MECNET has been tested on diverse public datasets, it has yet to be evaluated in clinical settings with real-world hospital data. Such validation will be critical to assessing the model's robustness in variable imaging conditions and its readiness for integration into diagnostic workflows.

Table 6 Comparative accuracy of MECNET and existing models on GI image classification datasets.

Sr. No.	Author	Dataset	Accuracy
1	UÇan et al., (2022)	Kvasir	93.5%
2	Yue et al., (2023)	Kvasir	90.75%
3	Park et al., (2023)	Kvasir	94.9%
4	Mary et al., (2023)	Kvasir	94.21%
5	**MECNET (Proposed)**	**Kvasir**	**97.2%**
6	Rani et al., (2022)	WCE	95.62
7	Bordbar et al., (2023)	WCE	97%
8	**MECNET (Proposed)**	**WCE**	**97.4%**

Compared to recent studies, MECNET achieves higher classification performance across key GI datasets. As shown in Table 6, previous models such as Park et al., (2023) (94.9% on Kvasir) and Rani et al., (2022) (95.62% on WCE) achieved strong performance; however, MECNET surpasses these benchmarks, achieving 97.2% and 97.4%, respectively. This suggests that the integration of a Feature Refiner Module with ensemble deep learning contributes significantly to overall accuracy and robustness across diverse datasets.

In summary, MECNET introduces a robust, generalizable, and high-performing framework for gastrointestinal disease classification. Its innovative use of refined image preprocessing, ensemble modeling, and transfer learning marks a significant advancement over current state-of-the-art methods. By continuing to reduce model complexity and expand validation datasets, MECNET holds great potential for future clinical deployment

5. Conclusion

This study presents MECNET, a hybrid deep learning framework that demonstrates high efficacy in the automatic classification of gastrointestinal (GI) tract diseases, with particular focus on colorectal cancer detection. By integrating a Feature Refiner Module with advanced CNN architectures such as VGG19, ResNet50, and EfficientNet, MECNET achieves superior performance across multiple benchmark datasets in terms of accuracy, precision, recall, and F1 score. The adoption of robust preprocessing techniques and multi-scale feature extraction significantly enhances the model's capability to interpret complex medical images, while its strong generalization on unseen data underscores its potential for real-world clinical deployment.

The results establish MECNET as a competitive and scalable solution in the field of medical image analysis, contributing to the advancement of AI-assisted diagnostic systems. However, the study also acknowledges limitations, including increased model complexity, the reliance on publicly available datasets, and the lack of clinical validation. Future work should prioritize model optimization for real-time applications, integration with diverse clinical datasets, and prospective evaluation in healthcare settings.

Overall, this research contributes to the growing field of AI in gastroenterology and offers a promising direction for enhancing early detection and diagnosis of GI diseases. MECNET's strong performance and modular design pave the way for further innovations in intelligent diagnostic tools aimed at improving patient care and reducing diagnostic burdens in clinical practice.

6. Conflict of Interest

The authors declare that they have no conflict of interest.

7. Funding

No funding was used for this study.

8. References

Bordbar, M., Helfroush, M. S., Danyali, H., &Ejtehadi, F. (2023). Wireless capsule endoscopy multiclass classification using three-dimensional deep convolutional neural network model. *BioMedical Engineering OnLine*,*22*(1), Article 124. https://doi.org/10.1186/s12938-023-01186-9

Chae, J. W., & Cho, H. C. (2023). Enhanced classification of gastric lesions and early gastric cancer diagnosis in gastroscopy using multi-filter autoaugment. *IEEE Access*, *11*, 29391-29399. https://doi.org/10.1109/ACCESS.2023.3260983

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA. https://doi.org/10.1109/CVPR.2016.90

He, R., Li, C., Yang, X., Yang, J., Jiang, T., Grzegorzek, M., & Sun, H. (2023). *Application of graph-curvature features in auxiliary diagnosis for histopathological image identification of gastric cancer* [Preprint]. SSRN. https://doi.org/10.2139/ssrn.4525425

Islam, J., Bhuiyan, S., Hossain, A., Surja, A. S., & Iqbal, M. S. (2021). Classification of gastric precancerous diseases using hybrid CNN-SVM. In *2021 3rd International Conference on Electrical & Electronic Engineering (ICEEE)*. IEEE, Rajshahi, Bangladesh. https://doi.org/10.1109/ICEEE54059.2021.9718790

Jha, D., Sharma, V., Dasu, N., Tomar, N. K., Hicks, S., Bhuyan, M. K., ... & de Lange, T. (2023). GastroVision: A multi-class endoscopy image dataset for computer-aided gastrointestinal disease detection. In *Workshop on Machine*

Learning for Multimodal Healthcare Data (pp. 125–140). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47679-2_10

Kaur, P., & Kumar, R. (2024). Performance analysis of convolutional neural network architectures over wireless capsule endoscopy dataset. *Bulletin of Electrical Engineering and Informatics*,*13*(1), 312-319. https://doi.org/10.11591/eei.v13i1.5858

Kumar, R., Singh, A., &Khamparia, A. (2023). Multiclass classification of gastrointestinal colorectal cancer using deep learning. In *International Conference on Innovative Computing and Communication*. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-4071-4_48

Kumar, R., Singh, A., &Khamparia, A. (2024). MEGANET: Improved framework with nature inspired approach for colorectal cancer polyp classification. *Intelligent Decision Technologies*, *18*(2), 825-836. https://doi.org/10.3233/IDT-240187

Kumari, R., & Wasim, J. (2023). A deep learning approach for human facial expression recognition using Residual Network–101. *Journal of Current Science and Technology*, *13*(3), 517–532. https://doi.org/10.59796/jcst.V13N3.2023.2152

Mary, X. A., Raj, A., Evangeline, C. S., Neebha, T. M., Kumaravelu, V. B., & Manimegalai, P. (2023). Multi-class Classification of Gastrointestinal Diseases using Deep Learning Techniques. *The Open Biomedical Engineering Journal*, *17*(1). Article e187412072301300.

Mushtaq, D., Madni, T. M., Janjua, U. I., Anwar, F., &Kakakhail, A. (2023). An automatic gastric polyp detection technique using deep learning. *International Journal of Imaging Systems and Technology*, *33*(3), 866-880. https://doi.org/10.1002/ima.22850

Navale, S., Anjanikar, A., Mohammed, A., SV, A., & Gite, S. (2024). Automated classification of gastrointestinal abnormalities using convolutional neural networks. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(12s), 348-360. https://ijisae.org/index.php/IJISAE/article/view/4521

Park, H. C., Hong, I. P., Poudel, S., & Choi, C. (2023). Data augmentation based on generative adversarial networks for endoscopic image classification. *IEEE Access*, *11*, 49216-49225. https://doi.org/10.1109/ACCESS.2023.3275173

Pogorelov, K., Randel, K. R., Griwodz, C., Eskeland, S. L., de Lange, T., Johansen, D., ... & Halvorsen, P. (2017). Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, Taipei, Taiwan. https://doi.org/10.1145/3083187.3083212

Rani, K., Devi, G., Kumar, S., Figueiredo, I. N., & Figueiredo, P. N. (2022). Classification of wireless capsule endoscopy images for bleeding using deep features fusion. In *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. IEEE, Maldives. https://doi.org/10.1109/ICECCME55909.2022.9987916

Sabah, Y. H., & Hassan, F. (2024). Comparison of serum vitamins C, D, and E in Iraqi colorectal cancer patients with and without non-alcoholic fatty liver disease to healthy individuals. *Journal of Current Science and Technology, 14*(3), Article 56. https://doi.org/10.59796/jcst.V14N3.2024.56

Sima, I., &Cincar, K. (2021). Transfer learning-based classification of gastrointestinal polyps. In *2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE)*, IEEE, Kragujevac, Serbia. https://doi.org/10.1109/BIBE52308.2021.9635497

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556

Srivastava, A., Tomar, N. K., Bagci, U., & Jha, D. (2022). Video capsule endoscopy classification using focal modulation guided convolutional neural network. In *2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, Shenzen, China. https://doi.org/10.1109/CBMS55023.2022.00064

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and

mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209-249. https://doi.org/10.3322/caac.21660

Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*, PMLR, Long Beach, California.

UÇan, M., Kaya, B., & Kaya, M. (2022). Multi-class gastrointestinal images classification using EfficientNet-B0 CNN model. In *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, IEEE, Sakhir, Bahrain. https://doi.org/10.1109/ICDABI56818.2022.10041447

Wu, M., Liu, Q., Yan, C., & Sen, G. (2023). Multi-Classification of brain tumors on magnetic resonance images using an ensemble of pre-trained convolutional neural networks. *Current Medical Imaging Reviews*, *19*(1), 65-76. https://doi.org/10.2174/1573405618666220415122843

Yue, G., Wei, P., Liu, Y., Luo, Y., Du, J., & Wang, T. (2023). Automated endoscopic image classification via deep neural network with class imbalance loss. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1-11. https://doi.org/10.1109/TIM.2023.3264047