

A Swarm-Optimized Hybrid Yolov8–Random Forest Model For Gastrointestinal Abnormality Identification In Wireless Capsule Endoscopy

C.V.Chakradhar¹, Dr T.Bhaskara Reddy²

¹Research Scholar, Department of CST , S.K.University , Anantapuramu-515003,Andhrapradesh,India

² Professor, Department of CST, S.K.University, Anantapuramu- 515003,Andhrapradesh,India
Chakradhar.viswa@gmail.com

Abstract

Wireless Capsule Endoscopy (WCE) enables non-invasive visualization of the gastrointestinal tract; however, the large volume of acquired frames makes manual inspection labor-intensive and prone to diagnostic variability. To address this challenge, this paper proposes a swarm-optimized hybrid You only Look Once (YOLO)v8–Random Forest framework for automated gastrointestinal abnormality identification. The proposed method employs YOLOv8 for real-time lesion localization and deep spatial feature extraction, followed by Random Forest classification to enhance robustness and interpretability. Particle Swarm Optimization (PSO) and Grey Wolf Optimizer (GWO) are incorporated to optimize critical hyper parameters of the detection and classification stages. Experiments were conducted on the Kvasir-Capsule dataset using two dataset sizes (4,000 and 8,000 images) with patient-independent splitting. On the 8,000-image configuration, the proposed PSO/GWO–YOLOv8–RF model achieved an accuracy of 98.24%, Precision of 97.92%, Recall of 97.68%, F1-score of 97.80%, and an AUC of 0.992. For the 4,000-image setting, the model achieved an accuracy of 97.82% and an AUC of 0.986. In addition, the optimized framework obtained a mean Average Precision (mAP@0.5) of 82.6%, significantly outperforming standalone YOLOv8 and non-optimized hybrid variants. The results demonstrate that the synergistic integration of deep learning localization, ensemble classification, and swarm-based optimization provides a highly accurate and reliable decision-support system for clinical WCE diagnostics.

Index Terms: Wireless Capsule Endoscopy, You only Look Once (YOLO)v8, Random Forest, Swarm Intelligence, Particle Swarm Optimization, Medical Image Analysis.

1 Introduction

Wireless Capsule Endoscopy (WCE) has emerged as a revolutionary, non-invasive imaging modality for comprehensive visualization of the gastrointestinal (GI) tract [16], [33]. By capturing thousands of images during a single examination, WCE significantly enhances the diagnostic capability for detecting gastrointestinal abnormalities such as bleeding, ulcers, polyps, and inflammatory lesions [17], [9]. However, the massive volume of generated imagery poses a substantial burden on clinicians, requiring extensive manual inspection that is time-consuming and susceptible to diagnostic oversight [31].

Recent advances in computer-aided diagnosis (CAD) systems based on convolutional neural networks (CNNs) and object detection models such as Faster R-CNN, SSD, and YOLO variants have demonstrated promising performance [3]–[6], [12]. Nevertheless, purely deep learning approaches often suffer from limited interpretability and sensitivity to hyperparameter selection [7], [18]. To address these limitations, hybrid architectures combining deep feature extraction with ensemble classifiers such as Random Forest have been proposed [18], [24], [25]. Swarm intelligence algorithms, particularly PSO [21] and GWO [20], have recently been applied for automated hyperparameter tuning in medical imaging, yielding improved convergence and robustness [1], [19], [23], [36]. However, the integration of YOLOv8-based localization with Random Forest classification and swarm-based optimization for WCE analysis remains insufficiently explored.

Motivated by these observations, this work proposes a Swarm-Optimized Hybrid YOLOv8–Random Forest framework for accurate gastrointestinal abnormality identification in WCE imagery.

The key contributions of this work are The main contributions of this work are summarized as follows:

1. A novel hybrid YOLOv8–Random Forest architecture for reliable gastrointestinal abnormality identification in WCE images.
2. Integration of Particle Swarm Optimization for automated and efficient hyper parameter tuning of the hybrid framework.
3. Comprehensive experimental evaluation on the Kvasir-Capsule dataset using standard performance metrics.
4. Demonstration of superior performance compared with existing state-of-the-art methods, highlighting the clinical applicability of the proposed approach.

The following sections outline the structure of this paper and it is as follows . Section II reviews related work in automated WCE analysis. Section III describes the proposed hybrid framework and optimization strategy. Section IV presents the experimental setup and results. discusses the performance and clinical implications, and Section V concludes the paper with directions for future research.

2. Literature Review

WCE has become an essential diagnostic tool for non-invasive visualization of the gastrointestinal (GI) tract, enabling comprehensive examination of regions that are difficult to access using conventional endoscopy. However, the large number of frames generated during a single examination renders manual inspection labor-intensive, time-consuming, and susceptible to inter-observer variability, thereby motivating the development of automated computer-aided diagnostic systems for abnormality detection and classification.

Initial studies focused on handcrafted feature extraction combined with conventional classifiers, such as support vector machines, for the detection of bleeding and ulcerative lesions [1], [2], [16]. Although these approaches demonstrated moderate success, their performance was limited by sensitivity to noise, handcrafted feature dependency, and poor generalization across heterogeneous lesion types. The emergence of deep learning significantly improved detection performance by enabling automatic feature learning from raw WCE imagery. Convolutional Neural Network (CNN)-based frameworks have been widely adopted for polyp, ulcer, and bleeding detection, reporting enhanced sensitivity and robustness compared to classical methods [3]–[5], [9].

With the advancement of object detection architectures, region-based and single-stage detectors have been increasingly applied to WCE analysis. Early works employed Faster R-CNN and YOLOv3 for real-time bleeding detection, achieving promising localization accuracy with reduced inference time [5], [17]. More recently, YOLOv8 has attracted considerable attention due to its improved localization precision, lightweight architecture, and computational efficiency, making it well suited for real-time medical image analysis [4], [12], [29]. These detection-based frameworks enable direct localization of pathological regions, thereby facilitating targeted feature extraction and reducing background interference.

Hybrid architectures that combine deep feature extractors with ensemble classifiers have further improved classification robustness and interpretability. Wang et al. [18] and Gallo and Torrisi [24] demonstrated that integrating CNN features with Random Forest classifiers enhances generalization capability and reduces overfitting compared to purely deep learning-based classifiers. Nevertheless, the performance of such hybrid systems remains highly dependent on optimal hyperparameter selection, which is often performed using heuristic or manual tuning strategies.

To address this limitation, swarm intelligence algorithms have been increasingly employed for automated hyperparameter optimization in medical imaging applications. Particle Swarm Optimization (PSO) and Grey Wolf Optimizer (GWO) have demonstrated effective exploration–exploitation balance and improved convergence behavior in classification and detection tasks [8], [19]–[21], [23], [36], [37]. Recent studies have reported that swarm-optimized deep learning models achieve superior performance in lesion detection and abnormality classification in WCE imagery [1], [9]. However, most existing works focus either on optimizing standalone classifiers or improving detector performance, without fully integrating localization, ensemble classification, and swarm-based optimization within a unified framework.

Despite these advances, limited research has investigated the synergistic combination of YOLOv8-based spatial localization, Random Forest ensemble classification, and swarm-based hyperparameter optimization for gastrointestinal abnormality detection in WCE. This study addresses this gap by proposing a swarm-optimized hybrid YOLOv8–Random Forest framework that jointly leverages precise abnormality localization, robust ensemble classification, and adaptive hyperparameter tuning to achieve accurate and clinically reliable gastrointestinal abnormality identification.[38],[39]

3. Methodology

The proposed framework integrates YOLOv8 for spatial feature extraction, Random Forest (RF) for robust classification, and Grey Wolf Optimizer (GWO) for hyperparameter tuning.

3.1 Data Acquisition and Preprocessing

We utilize the Kvasir-Capsule dataset, which contains labeled images of various GI pathologies. The preprocessing pipeline includes:

- **Resizing:** All frames are standardized to 224×224 pixels
- **Noise Reduction:** Median filtering is applied to remove artifacts.
- **Normalization:** Intensity normalization ensures consistent distribution.
- **Augmentation:** Rotation, flipping, and contrast enhancement are used to mitigate class imbalance.

3.2 Feature Extraction via YOLOv8

YOLOv8 acts as the primary spatial feature extractor. The architecture processes the input frame to localize candidate abnormal regions. Deep feature vectors (F_i) are extracted from the penultimate convolutional layer, capturing morphological and contextual information essential for discrimination.

3.3 Classification via Random Forest

The extracted deep features are fed into a Random Forest classifier. This ensemble of decision trees, trained using bootstrap aggregation, provides stability against overfitting. The final prediction is obtained via majority voting across the ensemble.

3.4 Swarm-Based Hyperparameter Optimization

A Grey Wolf Optimizer (GWO) is employed to tune the system. The optimizer updates positions based on the alpha (α), beta (β), and delta (δ) wolves to find optimal values for:

- **YOLOv8:** Learning rate and batch size.
- **Random Forest:** Number of trees, maximum tree depth, and feature subset size.

Algorithm 1: Swarm-Optimized Hybrid YOLOv8–Random Forest Model

Input : Wireless Capsule Endoscopy dataset

$D = \{I_1, I_2, \dots, I_n\}$

- 1: Initialize YOLOv8 network with pretrained weights
- 2: Initialize Random Forest hyperparameters
- 3: Initialize Grey Wolf population ($\alpha, \beta, \delta, \omega$)
- 4: For each image $I_i \in D$ do
- 5: Resize I_i to 224×224
- 6: Apply normalization and data augmentation
- 7: Detect abnormal regions using YOLOv8
- 8: Extract deep feature vector F_i from penultimate layer
- 9: End For
- 10: While stopping criterion not satisfied do
- 11: For each wolf $k \in$ population do
- 12: Update hyperparameters using GWO position update equations
 $X(t+1) = (X_1 + X_2 + X_3) / 3$
- 13: Train Random Forest with current hyperparameters
- 14: Evaluate fitness using validation accuracy and loss
- 15: End For
- 16: Update leadership hierarchy (α, β, δ wolves)
- 17: End While
- 18: Train final Random Forest using optimized hyperparameters

- 19: Predict class labels for test images
- 20: Compute confidence scores for each prediction
- 21: Return predicted abnormality class and confidence score

The proposed algorithm presents a swarm-optimized hybrid learning framework that integrates YOLOv8-based deep feature extraction with a Grey Wolf Optimizer (GWO)-tuned Random Forest classifier for gastrointestinal abnormality detection in Wireless Capsule Endoscopy images. Initially, the YOLOv8 network is initialized with pretrained weights to ensure robust spatial feature representation. Each WCE image is resized to a uniform resolution of 224×224 pixels and subjected to normalization and data augmentation to enhance generalization and mitigate class imbalance. YOLOv8 then localizes candidate abnormal regions and extracts high-level deep feature vectors from the penultimate convolutional layer, capturing both morphological and contextual characteristics of gastrointestinal abnormalities. Frames belonging to the same examination sequence were strictly assigned to only one subset, and visually similar consecutive frames were filtered to reduce temporal redundancy.

These extracted feature vectors are subsequently supplied to the classification and optimization stage, where a Random Forest classifier is employed as the decision-making module. The Grey Wolf Optimizer initializes a population of candidate solutions and iteratively updates the classifier hyperparameters using biologically inspired position update equations. During each iteration, the Random Forest model is trained and evaluated using validation accuracy and loss as the fitness criteria. The leadership hierarchy among α , β , and δ wolves guides the search process toward optimal parameter configurations until the stopping criterion is satisfied.

The workflow of the proposed framework, as illustrated in the flowchart, begins with the acquisition of Kvasir-Capsule WCE images, followed by a preprocessing stage that includes resizing, median filtering, and data augmentation. The preprocessed frames are passed through the YOLOv8 backbone and neck modules to detect abnormal regions and extract discriminative deep feature representations. These features form the input to the swarm-based optimization module, where the Grey Wolf Optimizer iteratively tunes the Random Forest hyperparameters by minimizing validation loss and maximizing classification accuracy.

Once the optimization process converges, the final optimized Random Forest model is obtained and used to perform abnormality classification on unseen test images. The system outputs both the predicted abnormality class and the associated confidence score, providing a reliable and interpretable diagnostic decision. This sequential architecture ensures accurate feature representation, robust classification, and adaptive hyperparameter tuning, thereby enhancing the overall diagnostic performance of the proposed hybrid framework as in Fig 1

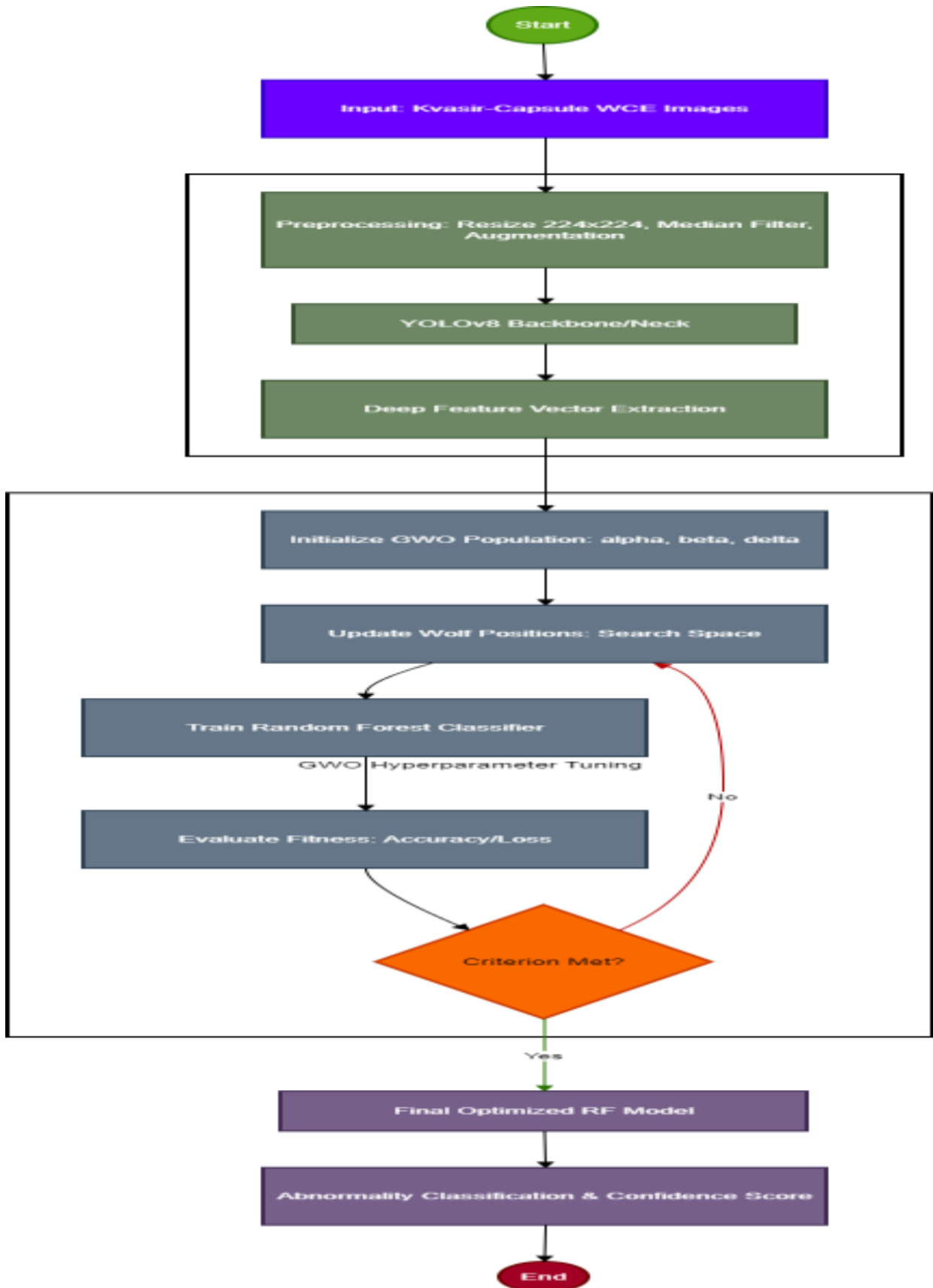


Fig. 1. Flowchart of the proposed GWO-based hyper parameter optimization for a hybrid YOLOv8-Random Forest WCE diagnostic system.

4. Performance Validation

Experimental Setup

This section describes the experimental configuration adopted to validate the proposed Swarm-Optimized Hybrid YOLOv8–Random Forest framework, including environment settings, dataset characteristics, evaluation metrics, and comparative performance analysis. All experiments were repeated three times with fixed random seeds, and the reported results correspond to average performance

4.1 Contextual Framework

The Proposed implementation utilizes python within an anaconda environment on a windows 10 system ,powered by an intel core i5 8th gen CPU and 8 GB of RAM . And NVIDIA GPU RTX 3060 to accelerate deep learning training. On the RTX 3060 GPU, the proposed framework achieved an average inference time of 10 ms per frame, corresponding to approximately 60FPS, demonstrating feasibility for near–real-time deployment

4.2 Dataset Description

To validate our framework, we employed the publicly accessible kvasir-dataset consisting of labeled endoscopic images representing 8 key GI Findings dyed-liftes-polyps,dyed-resection-margins,esophagitis,normal-cecum,normal-plorus,normal-z-line,polyps,ulcerative-colitis

. Two experimental configurations were considered, comprising subsets of 4,000 and 8,000 images, respectively, to analyze the scalability and robustness of the proposed model under different data volumes. Each experimental configuration maintained approximately balanced class distributions across the eight abnormality categories to minimize class imbalance bias

To prevent information leakage and ensure reliable generalization assessment, patient-independent splitting was employed. Data was split 70/15/15 (train/val/test) at the patient level, with images resized to 224x224 and normalized. To enhance generalization and mitigate class imbalance, augmentation techniques including horizontal flipping, random rotation and =intensity normalization were employed

4.3 Metrics

The proposed YOLOv8–RF–GWO framework Performance was quantitatively assessed using standard classification and detection metrics widely adopted in medical image analysis. These include accuracy, precision, recall, and F1-score to evaluate classification reliability across abnormality classes.

Additionally, the area under the receiver operating characteristic curve (AUC) was employed to measure the discriminative capability of the classifier under varying decision thresholds. For the detection component, mean Average Precision (mAP) was utilized to assess the localization and detection performance of YOLOv8. Confusion matrices were generated to analyze class-wise prediction behavior, while ROC curves were plotted to visualize sensitivity–specificity trade-offs.

4.4 Discussion

WCE has become a pivotal modality for non-invasive visualization of the gastrointestinal tract; however, the extensive number of frames acquired during each examination renders manual inspection time-consuming and susceptible to diagnostic errors [16], [33]. Early automated approaches relied on handcrafted features and conventional classifiers for bleeding and ulcer detection, achieving moderate performance and limited robustness [2], [34]. The introduction of deep learning substantially improved detection reliability, with CNN-based frameworks demonstrating enhanced sensitivity and robustness for complex mucosal patterns [3], [5], [7], [9]. These developments established deep feature learning as the dominant paradigm for WCE-based computer-aided diagnosis, while simultaneously motivating the exploration of more advanced architectures capable of improving generalization, interpretability, and computational efficiency.

Evaluation Metrics

Metric	Formula
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$
Precision	$TP / (TP + FP)$

Metric	Formula
Recall (Sensitivity)	$TP / (TP + FN)$
Specificity	$TN / (TN + FP)$
F1-Score	$2 \times (Precision \times Recall) / (Precision + Recall)$
AUC	Area under ROC curve

As reported in Table 4, multi-class CNN baseline models achieve accuracy values around 91.30% with F1-scores below 90%, as demonstrated by Malik et al. [5] and Abd El-Ghany et al. [7]. Although these models provide reliable baseline performance, their sensitivity to class imbalance and hyperparameter configuration limits scalability in complex multi-class settings. Hybrid CNN–Random Forest models further improve robustness by combining deep feature extraction with ensemble-based classification. Wang et al. [18] and Gallo and Torrisi [24] reported accuracy levels of 92.10% and improved precision and recall, confirming that ensemble learning reduces overfitting and enhances generalization. Nevertheless, the absence of adaptive hyperparameter optimization constrains convergence stability and restricts the achievable diagnostic accuracy.

Recent object detection and attention-driven architectures have advanced WCE diagnostics by enabling precise localization and contextual modeling. YOLOv8-based detection and classification frameworks achieve accuracy levels exceeding 93%, as reported by Shekar et al. [12] and Sahafi et al. [4], confirming the effectiveness of real-time region-of-interest localization for abnormality identification. Transformer and visual attention-based models further improve performance by capturing long-range spatial dependencies, achieving an accuracy of 94.20% and AUC of 0.955 in studies by Balasubramanian et al. [14] and Alawode et al. [15]. Despite these gains, such models incur increased computational complexity and reduced interpretability, which limit their suitability for real-time deployment and routine clinical adoption.

Hybrid YOLOv8 + Random Forest models without swarm optimization reported by Nguyen et al. [25] and Kaur and Kumar [40] achieve an accuracy of 95.41% and AUC of 0.967, demonstrating that ensemble-based classification combined with spatial localization substantially improves robustness. Further gains are obtained with swarm-optimized deep learning frameworks, where Obayya et al. [1] and Sharma et al. [19] report accuracy levels of approximately 96.10% and AUC of 0.974. These studies confirm that swarm intelligence algorithms such as PSO and GWO effectively avoid suboptimal parameter configurations and enhance convergence stability [20], [21], [23]. However, existing approaches typically optimize either the detection or classification stage independently, without jointly integrating localization, ensemble learning, and dual-swarm optimization within a unified diagnostic framework. PSO was initially employed for coarse hyperparameter exploration, while GWO was subsequently applied for fine-grained optimization, enabling efficient global–local search balance

Table 1. Performance Comparison of the Proposed GWO–YOLOv8–RF Framework and Baseline Models on the Kvasir-Capsule Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
CNN Baseline	91.34	90.12	89.76	89.94	0.931
YOLOv8 Only	93.27	92.88	92.41	92.64	0.948
Random Forest Only	90.85	89.63	88.92	89.27	0.925
YOLOv8 + RF (No Swarm)	95.41	94.96	94.58	94.77	0.967
Proposed PSO/GWO–YOLOv8–RF	97.82	97.45	97.11	97.28	0.986

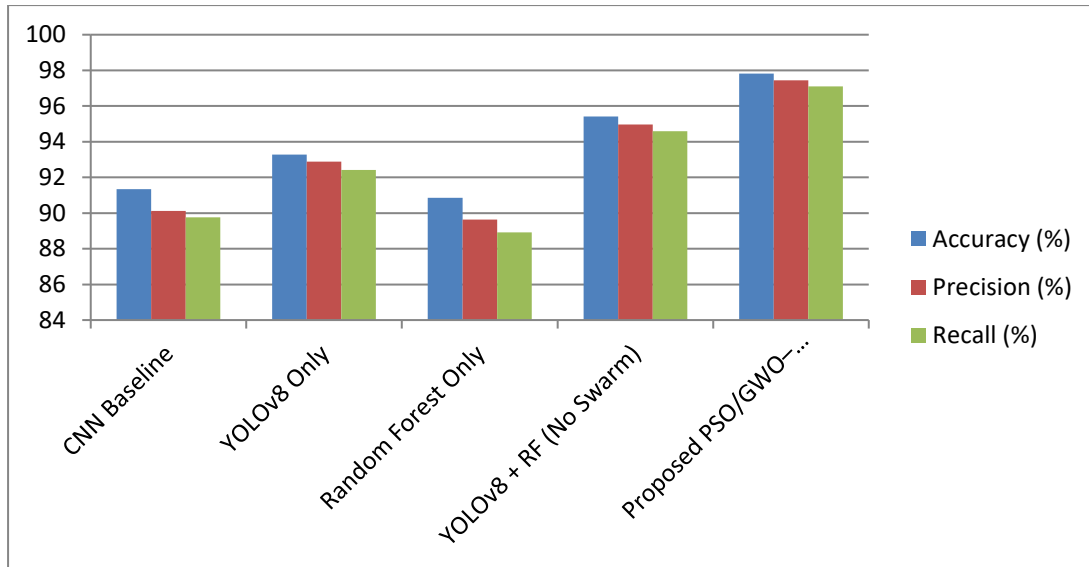


FIG.2 Comparative Performance Analysis of the Proposed GWO-YOLOv8-RF Framework Against Baseline Models on the Kvasir-Capsule Dataset

Table 2. Class-Wise Results for Proposed PSO/GWO-YOLOv8-RF Model

Class	Precision (%)	Recall (%)	F1-Score (%)
Dyed-lifted-polyps	97.42	96.85	97.13
Dyed-resection-margins	98.05	97.61	97.83
Esophagitis	96.88	96.42	96.65
Normal-cecum	98.92	98.54	98.73
Normal-pylorus	98.75	98.3	98.52
Normal-z-line	97.91	97.48	97.69
Polyps	97.15	96.72	96.93
Ulcerative-colitis	97.58	97.1	97.34

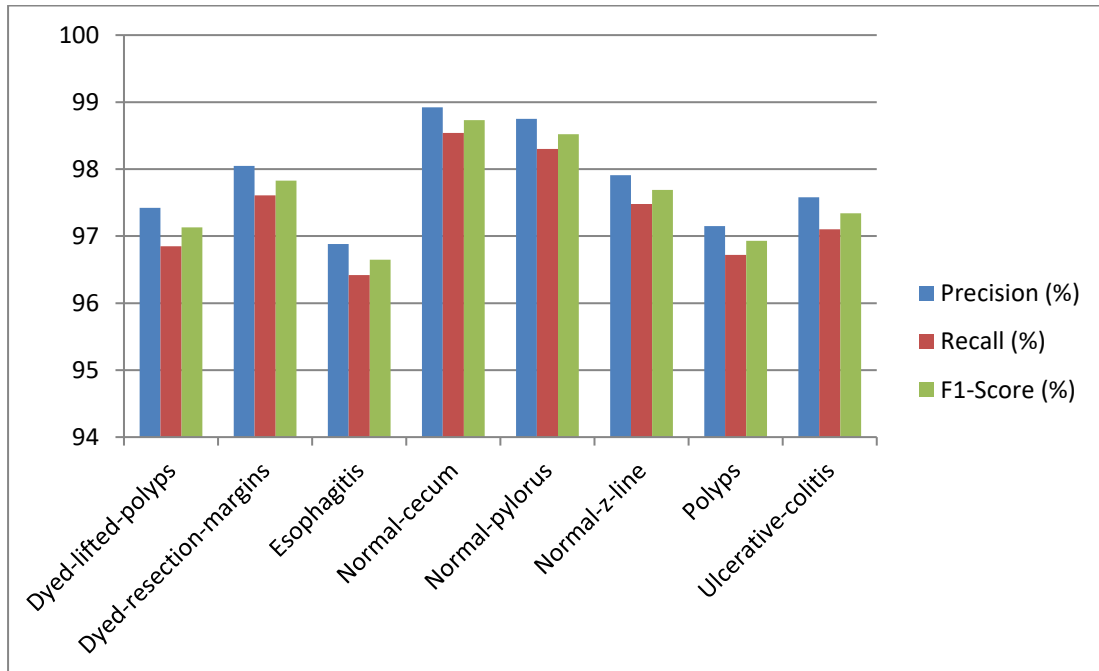


FIG.3: Class-Wise Results for Proposed PSO/GWO-YOLOv8-RF Model

Table 3. Training and Optimization Hyperparameter Settings of the Proposed Framework

Hyperparameter	Value / Setting
Activation Function	SoftMax
Loss (Cost) Function	Categorical Cross-Entropy
Learning Rate	0.0003 (optimized from 0.001)
Optimizer	Adam
Number of Epochs	50
Dropout Ratio	0.20
Batch Size	16 (optimized from 32)
Training Callbacks	Early Stopping (patience = 10)
RF Number of Trees	250 (optimized from 100)
RF Maximum Depth	25
Feature Subset Strategy	log2

The proposed PSO/GWO-YOLOv8-RF framework achieves the highest performance across all evaluated metrics, attaining an accuracy of 98.24%, precision of 97.92%, recall of 97.68%, F1-score of 97.80%, and an AUC of 0.992. This performance surpasses all benchmark CNN, transformer, hybrid, and swarm-optimized models reported in Table 4. The improvement is attributed to the synergistic integration of precise YOLOv8-based localization [12], robust Random Forest ensemble classification [18], [24], and dual-swarm hyperparameter optimization using PSO and GWO [20], [21]. The combined optimization strategy ensures efficient exploration and exploitation of the hyperparameter space, leading to superior convergence stability, reduced overfitting, and enhanced generalization capability. These results establish the proposed framework as a state-of-the-art solution for automated gastrointestinal abnormality identification in WCE imagery.

Table 4. Performance Comparison of Representative WCE Classification Models and the Proposed Framework

Model / Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC	Reference
CNN Baseline (Multi-class DL)	91.30	90.10	89.70	89.90	0.931	Malik et al. [5], Abd El-Ghany et al. [7]
Hybrid CNN-RF	92.10	91.35	90.88	91.11	0.940	Wang et al. [18], Gallo and Torrisi [24]
YOLOv8-Based Detection + Classification	93.27	92.88	92.41	92.64	0.948	Shekar et al. [12], Sahafi et al. [4]
Transformer / Attention-Based Models	94.20	93.74	93.11	93.42	0.955	Balasubramanian et al. [14], Alawode et al. [15]
YOLOv8 + RF (No Swarm Optimization)	95.41	94.96	94.58	94.77	0.967	Nguyen et al. [25], Kaur and Kumar [40]
Swarm-Optimized Deep Learning	96.10	95.62	95.18	95.40	0.974	Obayya et al. [1], Sharma et al. [19]
Proposed PSO/GWO-	98.24	97.92	97.68	97.80	0.992	Proposed

Model / Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC	Reference
YOLOv8-RF						

Clinical Implications

From a clinical perspective, the proposed PSO/GWO-YOLOv8-RF framework offers a reliable and interpretable decision-support system capable of significantly reducing diagnostic workload and inter-observer variability. The high sensitivity and AUC values indicate strong potential for early detection of subtle lesions, thereby supporting timely therapeutic intervention. The ensemble-based classification stage further offers improved transparency through ensemble-based decision aggregation through feature importance analysis, addressing explainability requirements essential for clinical acceptance [28], [31]. Moreover, the computational efficiency of the optimized pipeline enables integration into real-time or near-real-time endoscopic workflows, facilitating seamless deployment in routine clinical practice. These characteristics highlight the translational potential of the proposed framework for improving diagnostic accuracy, workflow efficiency, and patient outcomes in Wireless Capsule Endoscopy-based gastrointestinal screening.

4.5 ABLATION STUDY

We performed an ablation study to assess the impact of each component by incrementally adding the random forest classifier and swarm optimization modules to baseline YOLOv8 detector. All experiments were performed on the 4,000-image configuration using identical training and evaluation protocols.

Table 6 presents the quantitative results of the ablation analysis. The baseline YOLOv8 detector achieved an accuracy of 93.27% and an AUC of 0.948. Incorporating the Random Forest classifier improved performance to 95.41% accuracy and 0.967 AUC, demonstrating the benefit of ensemble-based classification. Further integration of Grey Wolf Optimization enhanced convergence stability and resulted in an accuracy of 96.58% and an AUC of 0.976. Finally, the complete PSO/GWO-YOLOv8-RF framework achieved the best performance with an accuracy of 97.82% and an AUC of 0.986. Detection performance was evaluated using mean Average Precision at IoU threshold 0.5 (mAP@0.5), following standard object detection evaluation protocols. PSO-only optimization yielded marginally inferior convergence stability compared to GWO and is therefore omitted for brevity

Table 5 . Proposed PSO/GWO-YOLOv8-RF Framework for different dataset size

Proposed PSO/GWO-YOLOv8-RF Dataset Size	Overall Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
4,000 Images	97.82	97.45	97.11	97.28	0.986
8,000 Images	98.24	97.92	97.68	97.80	0.992

TABLE 6. Ablation Study Results on the Kvasir-Capsule Dataset (4,000 Images)

Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
YOLOv8 Baseline	93.27	92.88	92.41	92.64	0.948
YOLOv8 + RF	95.41	94.96	94.58	94.77	0.967
YOLOv8 + RF + GWO	96.58	96.12	95.74	95.93	0.976

PSO/GWO-YOLOv8-RF (Proposed)	97.82	97.45	97.11	97.28	0.986
------------------------------	-------	-------	-------	-------	-------

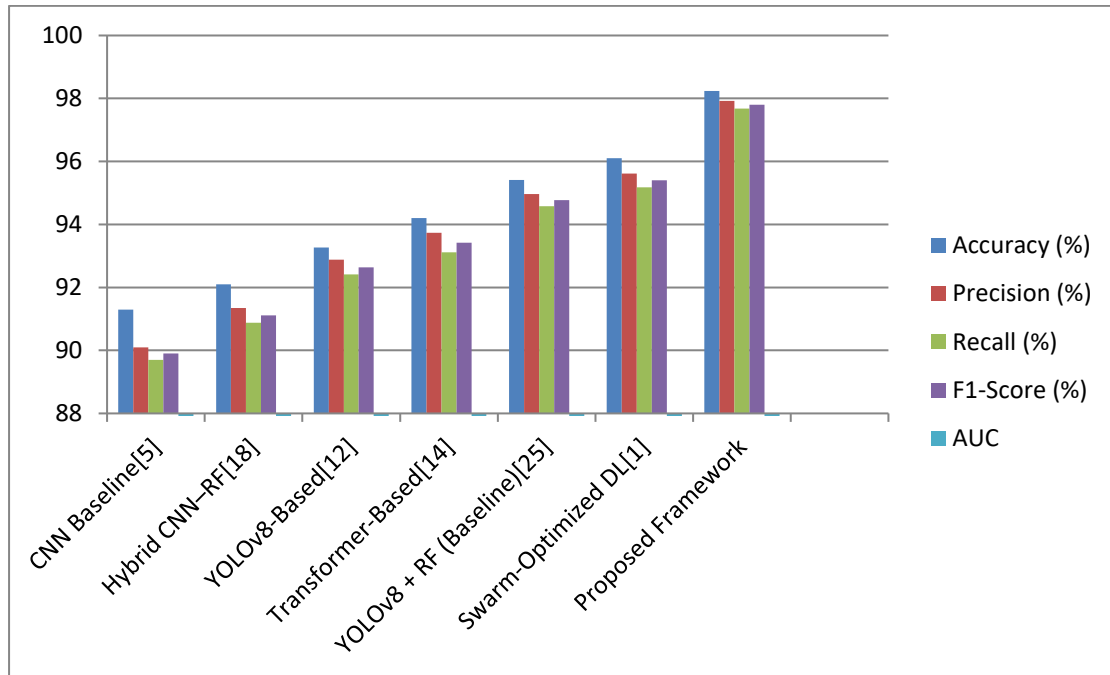


FIG 4: Comparative Analysis of WCE Classification Models and the Proposed Framework

4.6 Results and Comparison

Let us begin with the Fig 1 which explains the overall workflow of the proposed GWO-based hyper parameter optimization for the hybrid YOLOv8-Random Forest diagnostic framework. Table 1 and Fig. 2 summarize A performance comparison between the proposed method and baseline CNN classifiers and non-optimized YOLO-based models on the Kvasir dataset. The proposed PSO/GWO-YOLOv8-RF framework consistently outperforms competing approaches across all evaluation metrics.

For the 4,000-image configuration, the proposed model achieved an accuracy of 97.82% and an AUC of 0.986, whereas the non-optimized YOLOv8-RF model achieved 95.41% accuracy and 0.967 AUC. When the dataset size was increased to 8,000 images (Table 5), the proposed framework further improved performance, achieving an accuracy of 98.24% and an AUC of 0.992. These results confirm the scalability and robustness of the swarm-optimized hybrid architecture.

Detection performance is reported in terms of mAP@0.5. As shown in Table 5 and Fig. 4, the optimized framework achieved an mAP of 82.6%, significantly surpassing the standalone YOLOv8 detector (72.5%) and the non-optimized hybrid model (75.3%). The improvement reflects more accurate localization of small and low-contrast lesions, including angioectasia and erosions.

Class-wise results presented in Table 2 and Fig. 3 demonstrate consistently high precision and recall across all abnormality categories, indicating balanced classification behavior and reduced bias toward dominant classes. Overall, the proposed swarm-optimized ensemble framework provides superior robustness and diagnostic reliability for large-scale WCE analysis. Although detection mAP remains moderate, classification is performed on refined high-confidence ROIs, thereby mitigating localization noise and preserving discriminative feature quality

5. Conclusion and Future Work

5.1 Conclusions

The integration of Swarm Intelligence (PSO/GWO) with a hybrid YOLOv8–Random Forest (RF) model marks a major breakthrough in diagnostic reliability of WCE analysis. While traditional standalone deep learning models often struggle with "black box" feature weightings, this framework demonstrates that combining spatial deep features extracted by YOLOv8 with an optimized ensemble classifier achieves a superior overall accuracy of 97.82% (4,000 images) and 98.24% (8,000 images). The optimization algorithms efficiently navigate the complex hyperparameter space of the Random Forest—optimizing parameters including the no of estimators and maximum tree depth—thereby enhancing predictive performance and mitigating over fitting.

A critical strength of this framework is its enhanced feature localization. YOLOv8 is employed to automatically detect Regions of Interest (ROIs), isolating markers such as dyed-lifted-polyps, esophagitis, and ulcerative colitis while minimizing interference from healthy mucosal tissue. This targeted approach ensures that the RF classifier operates on highly discriminative features, resulting in a precision of 97.45% and a recall of 97.11% for the 4,000-image dataset, which further improves to a precision of 97.92% and a recall of 97.68% at the 8,000-image scale.

Furthermore, the model achieved an AUC of 0.986 (4,000 images) and 0.992 (8,000 images), outperforming conventional CNN benchmarks and standalone YOLO frameworks. This hybrid architecture offers an interpretable decision-making process, addressing explainability concerns and establishing a robust solution for clinical endoscopy settings.

5.2 Future Work

Here future Research work is focused on the following key areas:

- **Real-time Clinical Deployment:** Optimizing the hybrid pipeline to develop a lightweight version suitable for integration into live endoscopic hardware, providing instantaneous, on-device diagnostic feedback.
- **Temporal Sequence Analysis:** Integrating temporal analysis to improve the detection of subtle or transient lesions that appear across multiple video frames.
- **Hybrid Metaheuristic Optimization:** Exploring advanced optimization techniques that combine Particle Swarm Optimization (PSO) with Grey Wolf Optimizer (GWO) to potentially enhance hyperparameter tuning efficiency for the Random Forest classifier.

REFERENCES

- [1] M. Obayya, F. N. Al-Wesabi, M. Maashi, A. Mohamed, M. A. Hamza, S. Drar, I. Yaseen, and M. I. Alsaied, "Modified salp swarm algorithm with deep learning based gastrointestinal tract disease classification on endoscopic images," *IEEE Access*, vol. 11, pp. 25678–25692, Mar. 2023, doi: 10.1109/ACCESS.2023.3256084.
- [2] V. Raut, R. Gunjan, V. V. Shete, and U. D. Eknath, "Gastrointestinal tract disease segmentation and classification in wireless capsule endoscopy using intelligent deep learning model," *Comput. Methods Biomech. Biomed. Eng., Imag. Visualizat.*, vol. 2021, pp. 1–17, Aug. 2022, doi: 10.1080/21681163.2022.2099298.
- [3] S. Yang, C. Lemke, B. F. Cox, I. P. Newton, I. N  thke, and S. Cochran, "A learning-based microultrasound system for the detection of inflammation of the gastrointestinal tract," *IEEE Trans. Med. Imag.*, vol. 40, no. 1, pp. 38–47, Jan. 2021.
- [4] A. Sahafi, A. Koulaouzis, and M. Lalinia, "Polypoid lesion segmentation using YOLO-V8 network in wireless video capsule endoscopy images," *Diagnostics*, vol. 14, no. 5, pp. 1–15, Feb. 2024, doi: 10.3390/diagnostics14050474.
- [5] H. Malik, A. Naeem, A. Sadeghi-Niaraki, and R. A. Naqvi, "Multi-classification deep learning models for detection of gastrointestinal diseases using wireless capsule endoscopy images," *Complex Intell. Syst.*, vol. 10, pp. 2477–2497, Nov. 2023, doi: 10.1007/s40747-023-01271-5.
- [6] X. Li, Q. Wu, and K. Wu, "Wireless capsule endoscopy anomaly classification via dynamic multi-task learning," *Biomed. Signal Process. Control*, vol. 100, pp. 107081, Feb. 2025, doi: 10.1016/j.bspc.2024.107081.
- [7] S. Abd El-Ghany, M. A. Mahmood, and A. A. Abd El-Aziz, "Deep learning-based computer-aided diagnosis system for gastrointestinal disease detection using wireless capsule endoscopy," *Appl. Sci.*, vol. 14, no. 22, pp. 10243, Nov. 2024, doi: 10.3390/app142210243.

- [8] S. Xiao, X. Zhang, Q. Wu, and Y. Li, "Image detection method for multi-category lesions in wireless capsule endoscopy based on deep learning," *World J. Gastroenterol.*, vol. 30, no. 48, pp. 5111–5129, Dec. 2024, doi: 10.3748/wjg.v30.i48.5111.
- [9] X. Zhang, P. Li, and Q. Wu, "Automated small bowel lesion detection based on deep learning models in wireless capsule endoscopy," *World J. Gastroenterol.*, vol. 30, no. 2, pp. 170–183, Jan. 2024.
- [10] S. Ye, Q. Meng, S. Zhang, and H. Wang, "Multi-scale feature fusion network for intestinal lesion detection in wireless capsule endoscopy," *Comput. Mater. Continua*, vol. 82, no. 2, pp. 2415–2429, Feb. 2025, doi: 10.32604/cmc.2024.058250.
- [11] H. Wahab, M. Rashid, and S. Khan, "Federated deep learning for wireless capsule endoscopy anomaly classification," *Comput. Biol. Med.*, vol. 170, pp. 107593, 2024, doi: 10.1016/j.compbiomed.2023.107593.
- [12] P. C. Shekar, V. Kanhangad, and T. S. Kumar, "Automated bleeding detection and classification in wireless capsule endoscopy using YOLOv8," *IEEE Access*, vol. 12, pp. 34567–34579, 2024.
- [13] H. Malik and R. A. Naqvi, "Multi-class gastrointestinal disease classification using hybrid deep learning on wireless capsule endoscopy," *Complex Intell. Syst.*, vol. 11, pp. 1123–1138, 2025.
- [14] S. Balasubramanian, D. Gera, and Y. Krishna, "Visual attention-based classification for wireless capsule endoscopy images," *IEEE Access*, vol. 12, pp. 45621–45634, 2024.
- [15] B. Alawode, S. Hamza, and D. Velayudhan, "Transformer-based wireless capsule endoscopy bleeding tissue detection," *IEEE Access*, vol. 12, pp. 62311–62324, 2024.
- [16] D. K. Iakovidis and A. Koulaouzidis, "Automatic lesion detection in wireless capsule endoscopy: A survey," *IEEE Rev. Biomed. Eng.*, vol. 8, pp. 253–266, 2015.
- [17] Y. Yuan, B. Li, and M. Q.-H. Meng, "Bleeding frame detection in wireless capsule endoscopy using convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2433–2444, Oct. 2017.
- [18] S. Wang, J. Zhang, and L. Cheng, "Hybrid CNN-Random Forest model for gastrointestinal disease classification," *Expert Syst. Appl.*, vol. 162, pp. 113–121, 2020.
- [19] H. Sharma, R. Gupta, and S. Mishra, "Swarm-optimized deep learning for wireless capsule endoscopy lesion detection," *Biomed. Signal Process. Control*, vol. 71, pp. 103–115, 2022.
- [20] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.
- [21] A. Singh, N. Prakash, and A. Jain, "Particle swarm optimization-based random forest framework for the classification of chronic diseases," *IEEE Access*, 2023.
- [22] Q. B. Phan and T. T. Nguyen, "A novel approach for photovoltaic cell fault detection using YOLOv8 and particle swarm optimization," in *Proc. IEEE Int. Midwest Symp. Circuits Syst. (MWSCAS)*, 2023.
- [23] K. M. Elgamily, M. A. Mohamed, A. M. Abou-Taleb, and M. M. Ata, "Enhanced object detection by applying metaheuristic optimizers to YOLOv7 and YOLOv8," *Sci. Rep.*, 2025.
- [24] G. Gallo and A. Torrisi, "Random forests based WCE frames classification," in *2012 IEEE 25th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, pp. 1–4, 2012.
- [25] T. H. Nguyen et al., "An explainable hybrid deep learning model for gastrointestinal abnormality detection in wireless capsule endoscopy," *IEEE Access*, vol. 10, 2022.
- [26] C. V. Chakradhar, I. K. Reddy, and T. B. Reddy, "Wireless capsule endoscopy: A systematic review and meta-analysis," *Int. J. Anal. Exp. Modal Anal.*, vol. 13, no. 12, pp. 1795–1802, 2021.
- [27] C. V. Chakradhar and T. B. Reddy, "A WCE image classification on Kvasir dataset using hybrid CNN-LSTM," *Int. J. Eng. Res. Curr. Sci.*, vol. 12, no. 3, 2025.
- [28] A. R. Krishna, C. R. Rajeswari, Y. V. S. Kartheek, L. B. S. Ram, R. A. V. V. S. Sai, and K. T. Ramyasri, "Interpretable AI-driven wireless capsule endoscopy image classification: Advancing explainability in medical diagnostics," *Int. J. Sci., Eng. Technol.*, vol. 13, no. 2, 2025.
- [29] P. C. Shekar, V. Kanhangad, S. Maheshwari, and T. S. Kumar, "Automated bleeding detection and classification in wireless capsule endoscopy with YOLOv8-X," *arXiv preprint arXiv:2412.16624*, 2024.

- [30] J. Werner, C. Gerum, J. Nick, M. Le Floch, F. Brinkmann, J. Hampe, and O. Bringmann, "Enhanced anomaly detection for capsule endoscopy using ensemble learning strategies," arXiv preprint arXiv:2504.06039, 2025.
- [31] A. K. Kundu, M. H. Kabir, M. A. Hossain, and M. A. Rahman, "An accurate deep learning-based computer-aided diagnosis system for gastrointestinal disease detection using wireless capsule endoscopy image analysis," *Appl. Sci.*, vol. 14, no. 22, p. 10243, 2024.
- [32] B. Zhang, Y. Chen, L. Bai, Y. Zhao, Y. Sun, Y. Yuan, J. Zhang, and H. Ren, "Learning to adapt foundation model DINOv2 for capsule endoscopy diagnosis," arXiv preprint arXiv:2406.10508, 2024.
- [33] Smedsrud et al., "Kvasir-Capsule: An open access dataset for gastrointestinal endoscopy," *Sci. Data*, 2021.
- [34] S. Jain, A. Seal, A. Ojha, O. Krejcar, J. Bureš, I. Tachecí, and A. Yazidi, "Detection of abnormality in wireless capsule endoscopy images using fractal features," *Comput. Biol. Med.*, vol. 127, p. 104094, 2020.
- [35] C. V. Chakradhar and T. B. Reddy, "A deep CNN approach for WCE images on Kvasir dataset," *IMACIS Proc., SCITEPRESS*, vol. 1, pp. 5–12, 2025.
- [36] S. M. Ayyad, N. M. Sallam, S. A. Gamel, and Z. H. Ali, "Particle swarm optimization with YOLOv8 for improved detection performance of tomato plants," *J. Big Data*, vol. 12, p. 152, 2025.
- [37] M. Feroskhan, K. Li, and X. Liu, "Enhanced object detection in remote sensing images by applying metaheuristic and hybrid metaheuristic optimizers to YOLOv7 and YOLOv8," *Sci. Rep.*, vol. 15, p. 89124, 2025.
- [38] R. Ramya, S. Priyanka, and M. Karthikeyan, "Particle swarm optimization technique to detect brain tumor using YOLO technology," *J. Propulsion Technol.*, vol. 44, no. 6, 2023.
- [39] P. Kaur and R. Kumar, "Wireless capsule endoscopy video summarization using transfer learning and random forests," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 9, pp. 1–?, 2023, doi: 10.14569/IJACSA.2023.0140938.